

# Robust Shape Tracking With Multiple Models in Ultrasound Images

Jacinto C. Nascimento, *Member, IEEE*, and Jorge S. Marques

**Abstract**—This paper addresses object tracking in ultrasound images using a robust multiple model tracker. The proposed tracker has the following features: 1) it uses multiple dynamic models to track the evolution of the object boundary, and 2) it models invalid observations (outliers), reducing their influence on the shape estimates. The problem considered in this paper is the tracking of the left ventricle which is known to be a challenging problem. The heart motion presents two phases (diastole and systole) with different dynamics, the multiple models used in this tracker try to solve this difficulty. In addition, ultrasound images are corrupted by strong multiplicative noise which prevents the use of standard deformable models. Robust estimation techniques are used to address this difficulty. The *multiple model data association* (MMDA) tracker proposed in this paper is based on a bank of nonlinear filters, organized in a tree structure. The algorithm determines which model is active at each instant of time and updates its state by propagating the probability distribution, using robust estimation techniques.

**Index Terms**—Image analysis, low level features, multiple model data association (MMDA), segmentation, tracking, ultrasound images.

## I. INTRODUCTION

OBJECT tracking in a video sequence is a solved problem when the object shape and motion parameters change slowly and the object boundary is easily detected in the image domain. However, these assumptions are not valid in most cases. Occlusions, multiple motions, and missing boundaries are among the most common difficulties found in object tracking.

This paper addresses a difficult problem in the field of medical image analysis: automatic tracking of the left ventricle (LV) in  $2 - D + T$  (space + time) ultrasound B-Mode images. Three main difficulties are considered in this paper. First, we assume that the object motion is not always smooth and may exhibit abrupt changes. Second, the images are corrupted by strong multiplicative noise, leading to many misdetections and false alarms: many detected features do not belong to the object. These features should be considered as outliers and neglected by the tracker since they jeopardize the performance of the tracker and often lead to misleading results. Finally, we assume that the object shape changes during the tracking process. This issue is

Manuscript received July 5, 2007; revised November 15, 2007. This work was supported in part by FCT under project Heart 3D and FCT ISR/IST plurianual funding, POSC, FEDER. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zoltan Kato.

The authors are with the Instituto Superior Tecnico, Instituto de Sistemas e Robotica (ISR/IST), 1049-001 Lisboa, Portugal.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2007.915552

addressed using a shape-space representation of the object contour to reduce the degrees of freedom.

These difficulties are treated in an unified framework as follows: i) *prediction with multiple dynamic models* is incorporated in the tracker. The tracker deals with abrupt motion changes by using a bank of switched dynamic models, each of them tailored to a specific shape/motion regime. This allows us to tackle complex motions where the dynamics significantly change; ii) *filtering with probabilistic data association*. Since the success of the tracker also depends on the observations extracted from the image and how they are handled, data association is crucial to deal with measurement uncertainty, providing a robust mechanism to attenuate the influence of outliers. Data association has its roots in Control theory and it is often used for tracking point targets in cluttered environments [1]. A validation region is used to select the set of admissible observations. Then an association probability (confidence degree) is assigned to each observation using a statistical model of the valid observations and outliers; iii) finally, a *shape-space model* is used to restrict the variation of the object contour. We will assume that the object shape belongs to a subspace with a low dimension which can be learned from the data. This representation should allow several types of shape deformations. In most of the cases, this representation is appropriate, since the number of modes capturing the main shape deformations is usually much smaller than the original number of feature components used to describe the shape.

The approach proposed herein is able to separate the uncertainty associated to the system dynamics and the uncertainty produced by shape deformation. This is a key issue since these two features are conceptually different. Multiple dynamic models can describe shape changes in systole and diastole phases, whereas the statistical shape model constrains the range of shape deformations and global motion (e.g., translation, rotation). Thus, we can determine which changes are due to global translation/rotation or to the contraction/expansion of the cardiac muscle.

Recent advances in this area include the work [2], which decouples the uncertainties associated to the system dynamics from the shape model, introducing a unified framework for fusing shape information (subspace shape model), system dynamics (dynamic model), and the measurements model (heteroscedastic noise).

The main contribution of the current paper is a new framework for efficient tracking of objects based on multiple dynamic models. The proposed method is able to cope with missing boundaries, occlusions and motion nonlinearities using multiple models. Furthermore, it overcomes the difficulties associated with clutter and multiplicative noise.

The paper is organized as follows. Section II describes related work. Section V describes the proposed tracker based on switched dynamic models. Section VI presents experimental results and a comparison with a state of the art tracker. Section VII concludes the paper.

## II. RELATED WORK

The evolution of the LV during the cardiac cycle provides valuable information which allows the assessment of the heart function (e.g., the ejection fraction i.e., the amount of blood pumped by the heart in each cycle) and the detection of illnesses [3]. Several attempts have been made to estimate the walls of the LV during the cardiac cycle in an automated way, in order to allow a quantitative evaluation of the cardiac function. This is, however, a difficult problem since the tracker has to deal with significant shape deformations, rapid motion during the systole (contraction) phase and low image quality due to the presence of non-Gaussian multiplicative noise produced by scattering. Furthermore, the ultrasound images of the heart often show an edge drop-out effect, i.e., some regions of the ventricle boundary may not be seen.

Several attempts have been made to solve this problem and track the heart cavities in ultrasound sequences. Most of them are based on deformable contours [4]–[10]. One of the first attempts [9] uses the snake algorithm and optical flow information extracted from the sequence of ultrasound images. This initial work was extended in many ways and alternative shape models have been proposed. The following directions were considered.

*Boundary representation:* The boundary of the left ventricle has been approximated by a spline curve (e.g., B-splines [7], Laplacian splines [11]) or by a Fourier series expansion whose coefficients are updated at every frame [10]. The Fourier series is more adapted to the elongated shapes of the ventricle. The next step was the use of learning methods to improve the shape model. Since the ventricle has typical deformation modes which are valid for most patients, the shape deformation can be learned from training data using PCA. This approach is followed in active shape models, e.g., [8], [12], [13], and shape subspace approaches [14]. Other alternatives include the use of deformable templates [4] and multiple contours to represent the epicardium and endocardium boundaries [15].

*Image features:* Different kinds of visual cues have been used for tracking purposes. Some works use optical flow [9], [16], [17] but it is usually difficult to obtain reliable optical flow estimates from noisy ultrasound images. Other alternatives include region based methods which characterize the statistical properties of the ultrasound image inside and outside the ventricle [18]. Image gradient [5], [19] and edges [20] have also been used. In the latter case, it is assumed that there is an intensity transition associated to the boundary of the ventricle. This approach also has some difficulties. The multiplicative noise present in the endocardium produces false edge points which are not associated to anatomical details. Furthermore, some images exhibit the edge drop-out effect which means that the boundary of the endocardium is not detected in a given region. The first difficulty can be alleviated by

preprocessing techniques (e.g., median [9], mean [17], integrated backscatter (IBS) [19], filtering or denoising methods [21]). This operation reduces speckle noise, but the key difficulties remain to be solved by contour estimation methods.

*Dynamic models:* Time dependence is a valuable source of information and it is used in most of the recent approaches [5], [16], [22]. This dependence can be modeled by assuming that the parameters of the elastic curve (e.g., spline control points or Fourier coefficients) are produced by a dynamic linear system. In most cases, the estimation is performed by Kalman filtering [23], assuming a linear model for the motion and for the observations. When the shape parameters are described by a nonlinear equation, particle filtering can be used instead [24]. The model can also be improved by noticing that the observations (e.g., edge points) have space-varying statistics i.e., their accuracy varies along the contour and during the cardiac cycle.

*Heteroscedastic noise:* Sophisticated strategies accounting for heteroscedastic, i.e., inhomogeneous (space varying) and anisotropic (directional) noise [25] have been recently proposed to improve the quality of the tracker [2], [22]. The proposed tracker deals with heteroscedastic noise and decouples motion information from statistical shape constraints, providing a fusion framework for these three cues.

We follow a different approach in this paper to deal with abrupt motion changes and image noise during the tracking operation. The proposed tracker is based on multiple dynamic models: we assume that parameter evolution is described by two dynamic models each of them tailored to a specific motion regime (systole/diastole). Shape estimation is performed by robust filtering methods which account for invalid (outlier) observations.

Multiple dynamic models has been studied in Control theory as a way to represent nonstationary dynamic behaviors [26]–[28]. This type of models is also known as switched dynamic models and they have been applied to deal with abrupt changes in dynamic systems [29]–[31] and tracking of manoeuvring targets with radar [1], [32].

The probability distribution of the state vector given the sequence of observations is no longer Gaussian in this case. It is a mixture of Gaussians with an exponentially growing number of modes. Therefore, exact inference is infeasible and approximate methods must be used to evaluate the state. This difficulty can be overcome by using mode merging and elimination algorithms i.e., by discarding modes with negligible probabilities and merging modes with similar parameters [1], [31]. Another alternative is the use of particle filtering techniques [24].

The use of switched dynamical models in active contours is scarce. Multiple dynamic models have been used to improve shape tracking in presence of abrupt shape and motion changes [24], [33]. They have also been used for heart segmentation in static images in [34], assuming that the contour to be estimated is the trajectory of a random point which moves along the object boundary. Multiple dynamic models are used to predict different shapes of the cavities boundary.

The tracker proposed in this paper combines multiple dynamic models and robust shape tracking using data association. Multiple models are used to represent the heart motion in systole

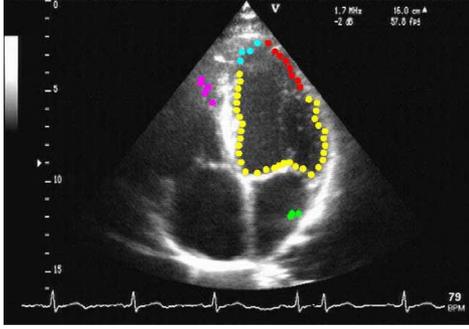


Fig. 1. Features linked to form five strokes (each color is a different stroke).

and diastole phases and robust estimation allows the tracker to cope with outlier features detected in the image. We first compute the edge points in the image and link them in strokes. Then we explicitly assume that many of them are outliers and should be neglected by the tracker. This is done by associating a binary label to each stroke. The binary labels are unknown and have to be estimated. In addition, we assume that the evolution of the shape parameters is complex and cannot be accurately described by a single dynamic model leading to a robust multi-model tracker which deals with multiple dynamics and outlier observations.

### III. PROBLEM STATEMENT

We wish to track the boundary of the left ventricle in a sequence of ultrasound images. The main difficulty lies in the presence of invalid features (edge points) detected in the image which do not belong to the ventricle boundary, as well as missing data. The first difficulty would disappear if every detected feature was classified correctly as valid or invalid. Since we do not have this information, all possible combinations (interpretations) of valid and invalid low level features should be considered. Unfortunately, the number of edge points,  $N$ , is very large and the number of label sequences ( $2^N$ ) is too large to be considered. To simplify the problem, we associate groups of features in strokes, drastically decreasing the number of interpretations.

Features are obtained by considering search lines orthogonal to the estimated contour of the left ventricle. All the abrupt transitions of intensities are then detected using a binary mask. This is similar to the method proposed in [35]. Edge segments (strokes) are then obtained by matching feature points using some spatial restrictions (e.g., distance from the predicted contour). Fig. 1 illustrates the detected features organized in five different strokes. Details are given in Section V.

We assume that each stroke can be classified as either true (valid) or false (invalid). An interpretation of the detected strokes,  $I_i$ , is defined as  $I_i = (I_i^1, \dots, I_i^M)$ , where  $I_i^j = \{0, 1\}$  is the label of the  $j$ th stroke in the interpretation  $i$ .

Let  $\mathbf{y}_t$  be the vector of all image features detected at the time instant  $t$  and let  $\mathbf{y}_t^i$  be a vector with the true features according to the interpretation  $i_t$ . Let  $\mathbf{x}_t$  be a set of parameters defining the

object boundary. It will be assumed that  $\mathbf{y}_t^i$  and  $\mathbf{x}_t$  are related by

$$\mathbf{y}_t^i = \mathbf{C}^i \mathbf{x}_t + \mathbf{v}_t^i \quad (1)$$

where  $\mathbf{C}^i$ , is the observation matrix (specified in the Appendix) and  $\mathbf{y}_t^i$  is the observation vector associated to the  $i$ th interpretation, and  $\mathbf{v}_t^i \sim \mathcal{N}(0, \mathbf{R}^i)$  is a white Gaussian measurement noise. Notice that the matrices  $\mathbf{C}^i$ ,  $\mathbf{C}^j$ , associated with two interpretations  $i, j$ , are different since the observation vectors  $\mathbf{y}_t^i$ ,  $\mathbf{y}_t^j$  often have different dimensions. It is assumed that the evolution of the object parameters,  $\mathbf{x}_t$  can be described by a set of switched dynamic models to be defined in Section IV.

### IV. SWITCHED DYNAMIC MODELS

We assume that the shape parameters  $\mathbf{x}_t$  are described by a set of stochastic difference equations [31]

$$\mathbf{x}_t = \mathbf{A}_{k_t} \mathbf{x}_{t-1} + \mathbf{w}_t \quad (2)$$

where  $k_t \in \{1, \dots, m\}$  is the label of the active model at instant  $t$ ,  $\mathbf{A}_{k_t}$  is the dynamic matrix,  $m$  is the number of steady state models, and  $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}_{k_t})$  is a white Gaussian noise. Each value of  $k_t$  corresponds to a different dynamic model. The state vector  $\mathbf{x}_t$  contains the parameters needed to specify the object contour. It is assumed throughout the paper that the object shape is a transformed version of a reference shape plus a local deformation, the latter being described by a B-spline (see the Appendix). Furthermore, it is assumed that the sequence of active labels  $\{k_t\}$  is a first order Markov process with transition matrix  $\mathbf{T}$  defined by

$$T_{rq} = P(k_t = q \mid k_{t-1} = r) \quad (3)$$

where  $r, q \in \{1, \dots, m\}$  and we assume that  $k_t$  is independent of previous state values  $\mathbf{x}_\tau$ ,  $\tau < t$ .

The hybrid state variable  $z_t = (\mathbf{x}_t, k_t)$  includes the state vector  $\mathbf{x}_t$  and the label of the active model  $k_t$ . The evolution of the hybrid state is characterized by the transition density  $p(\mathbf{x}_t, k_t \mid \mathbf{x}_{t-1}, k_{t-1})$ , which can be split as follows:

$$\begin{aligned} p(\mathbf{x}_t, k_t \mid \mathbf{x}_{t-1}, k_{t-1}) \\ = p(\mathbf{x}_t \mid k_t, \mathbf{x}_{t-1}, k_{t-1}) P(k_t \mid \mathbf{x}_{t-1}, k_{t-1}). \end{aligned} \quad (4)$$

The first factor can be computed from the dynamic (2) while the second term is an entry of the transition matrix  $T_{k_{t-1}, k_t}$ . In the ventricle tracking problem, we shall consider two dynamic models: one for the systolic phase and another for the diastolic phase. The first model contracts the contour while the second expands it.

### V. MULTIPLE MODEL DATA ASSOCIATION TRACKER

The tracking problem can now be formulated as follows: *Given a set of observations  $\mathbf{Y}^t = \{\mathbf{y}_1, \dots, \mathbf{y}_t\}$  which may contain outliers, what are the best estimates of the state vector  $\mathbf{x}_t$  and model label  $k_t$  at time  $t$ ?*

This is a nonlinear filtering problem which is non-Gaussian due to the presence of multiple models and outliers. This

problem can be tackled by propagating the *a posteriori* distribution of the hybrid state  $p(\mathbf{x}_t, k_t | \mathbf{Y}^t)$  for each new frame. This is done in three steps.

- *Prediction*: This step computes the distribution of the hybrid state in the next frame  $p(\mathbf{x}_t, k_t | \mathbf{Y}^{t-1})$ , given past observations.
- *Measurement*: Given a new image, a set of observations (edge strokes) is detected in the vicinity of the predicted contour.
- *Filtering*: The *a posteriori* distribution  $p(\mathbf{x}_t, k_t | \mathbf{Y}^t)$  is updated using the observations.

Let us first assume that we have a single dynamic model and all the observations are valid. In this case, the state vector and the observations are described by a linear dynamic system

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{w}_t \quad (5)$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{v}_t \quad (6)$$

where  $\mathbf{w}_t, \mathbf{v}_t$  are additive white noise processes. This is a well-known case where inference can easily be performed. The *a posteriori* distribution of the state vector is Gaussian  $p(\mathbf{x}_t | \mathbf{Y}^t) = \mathcal{N}(\hat{\mathbf{x}}_t, \mathbf{P}_t)$  with mean vector and covariance matrix

$$\hat{\mathbf{x}}_t = E\{\mathbf{x}_t | \mathbf{Y}^t\} \quad \mathbf{P}_t = E\{(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)^T | \mathbf{Y}^t\} \quad (7)$$

updated by Kalman filtering [36].

We now have to extend these ideas to allow multiple dynamic models switched according to a Markov process and invalid observations (outliers). These two cases will be separately considered.

#### A. Inference With Switched Dynamic Models

Let us consider a switched dynamic model with valid observations

$$\mathbf{x}_t = \mathbf{A}_{k_t}\mathbf{x}_{t-1} + \mathbf{w}_t \quad \mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}_{k_t}) \quad (8)$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{v}_t \quad \mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R}) \quad (9)$$

where  $k_t \in \{1, \dots, m\}$  is the label of the active model at time  $t$ .

Let us assume we have a sequence of observations  $\mathbf{Y}^t = (\mathbf{y}_1, \dots, \mathbf{y}_t)$  produced according to (9). These observations were produced by a sequence of models  $\mathbf{k}^t = (k_1, \dots, k_t)$  which is not known.

If the model sequence  $\mathbf{k}^t$  was known, the *a posteriori* distribution of the state would be Gaussian  $p(\mathbf{x}_t | \mathbf{Y}^t, \mathbf{k}^t) = \mathcal{N}(\hat{\mathbf{x}}_{K^t}, \mathbf{P}_{K^t})$ <sup>1</sup> with mean vector and covariance matrix  $\hat{\mathbf{x}}_{K^t}, \mathbf{P}_{K^t}$  computed by the Kalman filtering.

Since we do not know the sequence of active models  $\mathbf{k}^t$ , all sequences have to be considered. In this case the *a posteriori* distribution is a mixture of Gaussians. Each Gaussian is associated to a different path in the tree

$$p(\mathbf{x}_t, \mathbf{k}^t | \mathbf{Y}^t) = \sum_{\mathbf{k}^{t-1}} c_{K^t} \mathcal{N}(\hat{\mathbf{x}}_{K^t}, \mathbf{P}_{K^t}) \quad (10)$$

<sup>1</sup>In the following the notation,  $u_{K^t}$  means that the variable  $u$  depends on the path  $\mathbf{k}^t$ ;  $\mathbf{k}^t$  being the sequence of model labels up to instant  $t$ .

where

$$c_{K^t} = P\{\mathbf{k}^t | \mathbf{Y}^t\}. \quad (11)$$

The mean vectors and covariance matrices  $\hat{\mathbf{x}}_{K^t}, \mathbf{P}_{K^t}$  are updated by Kalman filtering as before. We can imagine that we perform a Kalman filtering iteration in each branch of the tree. Each filtering step is based on a specific dynamic model. This procedure propagates the mixture modes from the root to the leaves. When we arrive at the leaves, we obtain the mode parameters  $(c_{K^t}, \hat{\mathbf{x}}_{K^t}, \mathbf{P}_{K^t})$ . Only the update of  $c_{K^t}$  is different from usual Kalman filter equations and it is derived in the Appendix. The Kalman filter equations comprise the prediction and filtering steps as follows.

**Prediction:** Computes the modes  $c_{K^t}^- p(\mathbf{x}_t | \mathbf{Y}^{t-1}, \mathbf{k}^t) = c_{K^t}^- \mathcal{N}(\hat{\mathbf{x}}_{K^t}^-, \mathbf{P}_{K^t}^-)$

$$\begin{aligned} \hat{\mathbf{x}}_{K^t}^- &= \mathbf{A}_{k_t} \hat{\mathbf{x}}_{K^{t-1}} \\ \mathbf{P}_{K^t}^- &= \mathbf{A}_{k_t} \mathbf{P}_{K^{t-1}} \mathbf{A}_{k_t}^T + \mathbf{Q}_{k_t} \\ c_{K^t}^- &= \mathbf{T}_{k_{t-1}, k_t} c_{K^{t-1}}. \end{aligned} \quad (12)$$

**Filtering:** Computes the *a posteriori* distribution of the state at time  $t$  for the mode  $\mathbf{k}^t$

$$\begin{aligned} \hat{\mathbf{x}}_{K^t} &= \hat{\mathbf{x}}_{K^t}^- + \mathbf{K}_t(\mathbf{y}_t - \mathbf{C}\hat{\mathbf{x}}_{K^t}^-) \\ \mathbf{P}_{K^t} &= (\mathbf{I} - \mathbf{K}_t\mathbf{C})\mathbf{P}_{K^t}^- \\ c_{K^t} &= \gamma c_{K^t}^- \prod_{n=1}^L \mathcal{E}_t(s_n) \end{aligned} \quad (13)$$

where the superscript “-” denotes the predicted value at time  $t$  given the observation until time  $t - 1$ ,  $\mathbf{K}_t$  is the Kalman gain and

$$\mathcal{E}_t(s_n) = \rho^{-1} \mathcal{N}(\boldsymbol{\nu}_t(s_n); 0, \mathbf{S}_t(s_n)) \quad (14)$$

variables  $\rho, \mathbf{S}, \boldsymbol{\nu}$  are defined in the Appendix.

These are the classic Kalman filter equations assuming  $\mathbf{x}_t$  is produced by a time-varying linear model defined by the label sequence  $\mathbf{k}^t$ .

The computation of the mixture parameters can be done using closed form expressions. The only difficulty which remains is the fact that the number of modes grows exponentially with tracking time  $t$ . Fortunately, there are several ways to prune the tree to keep the number of modes bounded [31]. Two methods which can be used to achieve this goal are *component elimination* and *merging*. Component elimination discards components whose mixing coefficients  $c_{K^t}$  are smaller than a given threshold, normalizing the others. The eliminated components produce negligible contribution to the mixture density. The second method (component merging), merges similar components into a single one. This method avoids the existence of multiple components with close densities. The Kullback–Leibler divergence is used for deciding whether components are similar or not. In this paper, component elimination was implemented, using a threshold  $T = 10^{-3}$ .

## B. Inference With Outliers

When we have invalid observations (outliers) and multiple model, the problem becomes more difficult. Let us assume for the moment that we know which observations are invalid. In this case, we can discard them since they do not convey useful information. This can be done by throwing away the corresponding lines of the observation matrix  $\mathbf{C}$  defined in the Appendix.

Suppose we have detected  $M_t$  strokes in the  $t$ th frame. Each of these strokes can be valid or invalid. We, therefore, have  $2^{M_t}$  different interpretations of the data. Each of them can be identified by a label  $i_t \in \{1, \dots, 2^{M_t}\}$

The evolution of the state vector and observations can be formulated as

$$\mathbf{x}_t = \mathbf{A}_{k_t} \mathbf{x}_{t-1} + \mathbf{w}_t \quad \mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}_{k_t}) \quad (15)$$

$$\mathbf{y}_t^i = \mathbf{C}_t^i \mathbf{x}_t + \mathbf{v}_t \quad \mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R}_t^i) \quad (16)$$

where  $k_t$  is the label of the active model (unknown) and  $i_t$  the label of the data interpretation (unknown). In this case we have multiple dynamic models and also multiple observation equations and we do not know which pair is active at each time instant.

We could try to apply the same strategy as before to solve the inference problem. We could consider all the admissible sequences of active model/data up to time  $t$  ( $\mathbf{k}^t, \mathbf{i}^t$ ) and express the *a posteriori* distribution as a mixture of Gaussians. However, this is unfeasible. The number of dynamic models is usually small (less than 10) but the number of data interpretations is much larger (hundreds or thousands). A different strategy must be adopted.

The filtering problem with outliers was studied by Bar-Shalom in the context of point target tracking in radar systems [1]. The radar provides echoes corresponding to targets and clutter. The tracker must be able to identify which observations are valid and which are invalid. To address this problem, Bar-Shalom proposed a probability data association (PDA) filter which assigns a probability (confidence degree) to each observation. The PDA filter is not optimal but achieves excellent results in problems where the optimal solution is unfeasible. These ideas can be applied in the context multiple model tracking with outliers.

The main hypothesis is the following. We will assume that the prediction distribution associated to a path  $\mathbf{k}^t$  in the tree is still Gaussian

$$p(\mathbf{x}_t | \mathbf{Y}^{t-1}, \mathbf{k}^{t-1}) = \mathcal{N}(\mathbf{x}_{K^t}^-, \mathbf{P}_{K^t}^-) \quad (17)$$

where the mean vector and covariance matrix  $\mathbf{x}_{K^t}^-$ ,  $\mathbf{P}_{K^t}^-$  depend on the path. Equation (17) is not an exact assumption. The mixture modes in the presence of outliers are no longer Gaussian. However, this hypothesis is accurate enough to allow good results and leads to closed form expressions for the tracker.

Under this hypothesis, inference can be done in a simple way. The prediction step is the same as before [see (12)], the filtering

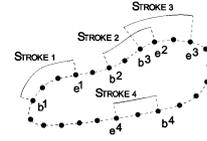


Fig. 2. Stroke characterization. Each stroke has two indices:  $b^j$  indicating the beginning of the  $j$ th stroke and  $e^j$  the end of the  $j$ th stroke.

step is different. We compute the mean vector and covariance matrix for each observation model  $i$  and combine them all.

Filtering:

$$\begin{aligned} \hat{\mathbf{x}}_{K^t} &= \hat{\mathbf{x}}_{K^t}^- + \sum_{i=1}^{m_t} \alpha_{K^t}^i \mathbf{K}_t^i \boldsymbol{\nu}_t^i \\ \mathbf{P}_{K^t} &= \left[ \mathbf{I} - \sum_{i=1}^{m_t} \alpha_{K^t}^i \mathbf{K}_t^i \mathbf{C}_t^i \right] \mathbf{P}_{K^t}^- \\ &\quad + \sum_{i=0}^{m_t} \alpha_{K^t}^i \hat{\mathbf{x}}_{K^t}^i \hat{\mathbf{x}}_{K^t}^{iT} - \hat{\mathbf{x}}_{K^t} \hat{\mathbf{x}}_{K^t}^T \\ c_{K^t} &= \gamma c_{K^t}^- \sum_i \alpha_{K^t}^i \prod_{j=1}^M \prod_{n=b^j}^{e^j} \mathcal{E}_{it}^j(s_n) \end{aligned} \quad (18)$$

where  $\alpha_{K^t}^i = P(i_t | \mathbf{Y}^t, \mathbf{k}^t)$  is the probability of the  $i$ th observation model given the data also known as the association probability,  $\mathcal{E}_{it}^j(s_n)$  is defined in the Appendix. Therefore, all the observation models are used to compute the final estimates of the Gaussian modes but they have different weights. These equations are derived in the Appendix.

The association probabilities can be obtained using the Bayes law provided we have a model for the valid data and for the outliers

$$\alpha_{K^t}^i \triangleq P(i_t | \mathbf{k}^t, \mathbf{Y}^t). \quad (19)$$

## C. Association Probabilities

This section describes how to compute the association probabilities. The association probabilities depend on the data model and outlier model. We have to, therefore, specify how the strokes are generated in the case of valid and outlier strokes, and assign them a probability distribution. It is assumed in this paper that high probabilities are assigned to the interpretations which contain i) long strokes; ii) strokes close to the predicted contour; iii) without stroke overlap, i.e., valid interpretations do not have multiple observations assigned to the same contour sample (see Fig. 2).

We will assume that the data generation model depends on the following variables:  $M$ —number of strokes,  $b^j$ ,  $e^j$ —beginning and ending of the  $j$ th stroke,  $I$ —interpretation and  $\mathbf{y}_t$ —vector of observations. These variables are assumed to be generated as shown in Fig. 3. The first block generates the number of strokes. The second one defines the start and end points of all the strokes. The output of the third block defines the valid/invalid strokes (interpretation). Finally, the fourth block generates the image features.

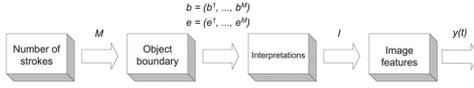


Fig. 3. Data generation model.

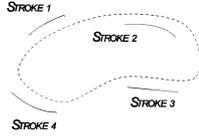


Fig. 4. Initial shape estimate (dashed line) and four detected strokes (solid lines).

The data model is characterized by the joint distribution  $p(\mathbf{y}_t, i_t, \mathbf{b}, \mathbf{e}, M | \mathbf{k}^t, \mathbf{Y}^{t-1})$ . From (19), one can write the association probabilities as follows:

$$\begin{aligned} \alpha_{K^t}^i &= P(i_t | \mathbf{y}_t, \mathbf{b}, \mathbf{e}, M, \mathbf{k}^t, \mathbf{Y}^{t-1}) \\ &= \frac{p(\mathbf{y}_t, i_t, \mathbf{b}, \mathbf{e}, M | \mathbf{k}^t, \mathbf{Y}^{t-1})}{p(\mathbf{y}_t, \mathbf{b}, \mathbf{e}, M | \mathbf{k}^t, \mathbf{Y}^{t-1})} \\ &= c p(\mathbf{y}_t, i_t, \mathbf{b}, \mathbf{e}, M | \mathbf{k}^t, \mathbf{Y}^{t-1}) \end{aligned} \quad (20)$$

where  $c$  is a normalization constant. In this expression, only the variable  $i$  is unknown.

The probability  $\alpha_{K^t}^i$  can be expressed as the product of  $p(\mathbf{y}_t | i_t, \mathbf{b}, \mathbf{e}, M, \mathbf{k}^t, \mathbf{Y}^{t-1})$  and  $P(i_t | \mathbf{b}, \mathbf{e}, M, \mathbf{k}^t, \mathbf{Y}^{t-1})$  which characterizes the data features generated by valid/invalid strokes, as well as the *a priori* probabilities of the data interpretations. In the first case, we assume that all features are independent, that is (see the Appendix)

$$p(\mathbf{y}_t | i_t, \mathbf{b}, \mathbf{e}, M, \mathbf{k}^t, \mathbf{Y}^{t-1}) = \prod_{j=1}^M \prod_{n=b^j}^{e^j} p(\mathbf{y}_t^j(s_n) | i_t) \quad (21)$$

where  $\mathbf{y}_t^j(s_n)$  is the observation of the  $j$ th stroke detected in the vicinity of  $s_n$ . In this paper, it is assumed that  $\mathbf{y}_t^j(s_n)$  has a normal distribution if the stroke is valid and a uniform one, otherwise. This approximation is similar to the one described in [1].

Furthermore, we assume that stroke labels (valid/invalid) are independent and valid strokes are longer than small strokes

$$P(i_t | \mathbf{b}, \mathbf{e}, M) = \prod_{j: I_t^j=1} p^j \prod_{j: I_t^j=0} (1 - p^j), \text{ with } p^j = \beta \frac{l^j}{L} \quad (22)$$

where  $\beta$  is a constant,  $l^j$  is the length of the  $j$ th stroke, and  $L$  is the samples number of the entire contour. This probability becomes larger as the length of the stroke increases.

To understand this model better, let us consider the example of Fig. 4 with four strokes. The dashed line is the initial contour estimate. Assuming that the object undergoes a 2-D rigid motion, a translation should be estimated. Since there is no *a priori* knowledge about the strokes labels, all interpretation must be considered leading to sixteen possible interpretations.

Each interpretation has a binary configuration. For instance, the interpretation  $I_5 = (0, 1, 0, 1)$  means that the strokes  $S_1$ ,

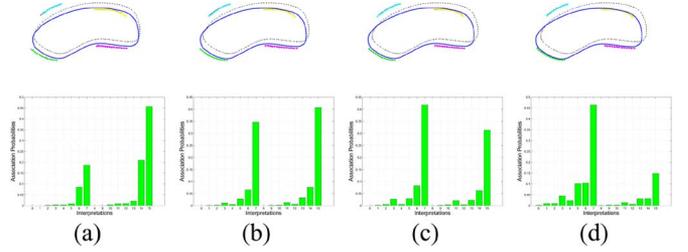


Fig. 5. (First row) Shape estimates and (second row) evolution of the association probabilities.

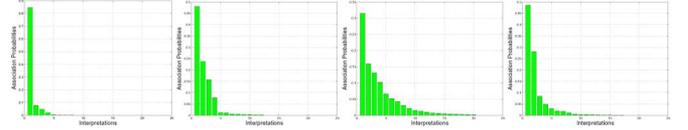
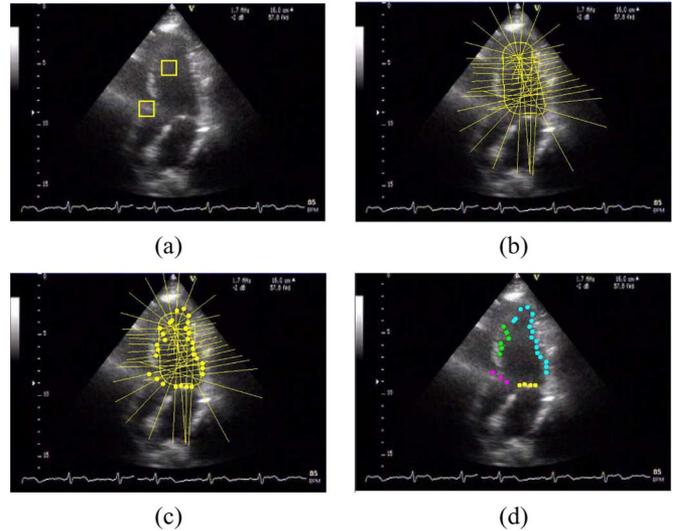


Fig. 6. Association probabilities computed in the frames 15, 32, 46, and 72 in a sequence of ultrasound images (to see the tracking results in the sequence, please refer to Section V).


 Fig. 7. Feature detection process. (a) Two  $15 \times 15$  regions are computed from two points given by the user, in which the mean intensities of the template are collected; (b) orthogonal (and tangential) lines radiating from the predicted contour; (c) orthogonal lines superimposed with detected features (dots); (d) detected features organized in line segments (each color denotes a different line segment).

$S_3$  are considered as outliers and strokes  $S_2$ ,  $S_4$  are considered as reliable (belonging to the boundary shape). In this example, we have sixteen different interpretations ranging from  $I_0 = (0, 0, 0, 0)$  to  $I_{15} = (1, 1, 1, 1)$ .

Since a considerable displacement is required, the translation is recursively updated and the association probabilities are updated in each iteration. Fig. 5 shows the evolution of the shape estimates as well as the association probabilities. Fig. 5(a) and (d) shows the initial and final configurations and Fig. 5(b) and (c) shows intermediate configurations. It can be seen that the tracker solves the ambiguity assigning the highest probability to the interpretation  $I_7 = (0, 1, 1, 1)$  discarding the influence of  $S_1$  and correctly estimating the translation given by the strokes  $S_2$ ,  $S_3$ ,  $S_4$  [Fig. 5(d)]. Although the interpretation  $I_{15} = (1, 1, 1, 1)$  has a significant influence at the beginning

TABLE I  
MULTIPLE MODEL DATA ASSOCIATION TRACKER

### Multiple Model Data Association Tracker

- 1) **Prediction:** (Expand each leaf creating  $m$  new leaves each one associated with a different label  $k_t$ ). For each leaf of the tree, predict the state mean  $\hat{\mathbf{x}}_{K^t}^-$ , covariance  $\mathbf{P}_{K^t}^-$ , and mixture coefficients  $c_{K^t}^-$

$$\hat{\mathbf{x}}_{K^t}^- = \mathbf{A}_{k_t} \hat{\mathbf{x}}_{K^{t-1}},$$

$$\mathbf{P}_{K^t}^- = \mathbf{A}_{k_t} \mathbf{P}_{K^{t-1}} \mathbf{A}_{k_t}^T + \mathbf{Q}_{k_t},$$

$$c_{K^t}^- = \mathbf{T}_{k_{t-1}, k_t} c_{K^{t-1}}.$$

- 2) **Measurement:** Compute the best estimate of the model and state vector in the next frame

$$\hat{k}_t = \arg \max_{k_t} \sum_{\mathbf{k}^t: k_t=j} c_{K^t}^-.$$

$$\hat{\mathbf{x}}_t^- = \gamma \sum_{\mathbf{k}^t: k_t=\hat{k}_t} c_{K^t}^- \hat{\mathbf{x}}_{K^t}^-,$$

$$\hat{\mathbf{y}}_t^- = \mathbf{C} \hat{\mathbf{x}}_t^-.$$

Detect features in the vicinity of  $\hat{\mathbf{y}}_t^-$  and organize them in strokes.

- 3) **Filtering:** For each leaf (component of the state vector) compute
- Association probabilities :  $\alpha_{K^t}^i \triangleq P(i_t | \mathbf{k}^t, \mathbf{Y}^t)$
  - Mean and covariance matrix of the state estimate and mixture coefficients for each tree path

$$\hat{\mathbf{x}}_{K^t} = \hat{\mathbf{x}}_{K^t}^- + \sum_{i=1}^{m_t} \alpha_{K^t}^i \mathbf{K}_t^i \nu_t^i,$$

$$\begin{aligned} \mathbf{P}_{K^t} &= \left[ \mathbf{I} - \sum_{i=1}^{m_t} \alpha_{K^t}^i \mathbf{K}_t^i \mathbf{C}^i \right] \mathbf{P}_{K^t}^- \\ &+ \sum_{i=0}^{m_t} \alpha_{K^t}^i \hat{\mathbf{x}}_{K^t}^i \hat{\mathbf{x}}_{K^t}^{iT} - \hat{\mathbf{x}}_{K^t} \hat{\mathbf{x}}_{K^t}^T, \end{aligned}$$

$$c_{K^t} = \gamma c_{K^t}^- \sum_i \alpha_{K^t}^i \prod_{j=1}^M \prod_{n=b^j}^{e^j} \mathcal{E}_{i_t}^j(s_n).$$

**Mode elimination:** discard the leaves with mixture coefficients below  $10^{-3}$  and normalize the mixture coefficients  $c_{K^t}$ . Merge modes with similar parameters.

it decays afterwards. It is interesting to note that there is a competition between these two interpretations.

Another example is illustrated in Fig. 6. This figure shows the association probabilities using real data (the probabilities are ordered from the highest one to the lowest one) where, once again, the distribution of the association probabilities suggests the most probable interpretation plays an important role but the others cannot be ignored.

### D. Image Measurement

Let us consider feature extraction. Ideally, we would like to detect all the points belonging to the boundary of the endocardium. However, we will also detect many false alarms produced by multiplicative noise and by other structures of the heart. Instead of relying on edge points only, we will link them in order to obtain strokes. Strokes are more reliable and robust

than edge points. For example, small strokes and spurious edges can be easily eliminated.

Furthermore, we are not interested in all the strokes present in the image. Since the ventricle wall changes in a predictable way, stroke detection will be guided by the ventricle estimate. We will first predict the ventricle contour in the next frame and perform stroke detection in the vicinity of the predicted contour.

The detection of the edge points is similar to the one described in [35, Ch. 5, pp. 100–110]. We sample the predicted contour in 40 equally spaced points and detect intensity transitions along directions orthogonal to the predicted contour [see Fig. 7(b)]. This is done by filtering the image profiles along the orthogonal directions using a binary template of length 11

$$\mathcal{R}(d) = \sum_n |p_i(n) - T(n-d)|^2 \quad (23)$$

where

$$T(n) = \begin{cases} I_{\text{avi}}, & -N \leq n < 0 \\ 0, & n = 0 \\ I_{\text{avo}}, & 0 < n \leq N \end{cases} \quad (24)$$

$p_i(n)$  is the intensity profile along the  $i$ th direction,  $I_{\text{avi}}$ ,  $I_{\text{avo}}$  are the average intensities inside and outside the contour evaluated on the first image of the sequence. The average intensity is computed in a  $15 \times 15$  window inside and outside of the LV. This is done for each new sequence [see Fig. 7(a)]. We note that multiple edge points can be detected in some directions. We then link the edge points associated to neighboring directions if their distances to the contour are similar.

Since we have multiple dynamic models, we will first select the most probable model

$$\hat{k}_t = \arg \max_{k_t} P(k_t | \mathbf{Y}^{t-1}) \quad (25)$$

where

$$P(k_t = j | \mathbf{Y}^{t-1}) = \sum_{\mathbf{k}^t: k_t = j} c_{K^t}^- \quad j = 1, \dots, m. \quad (26)$$

After knowing the most probable label, the ventricle shape can be estimated using the observation equation

$$\hat{\mathbf{y}}_t^- = \mathbf{C} \sum_{\mathbf{k}^t: k_t = \hat{k}_t} c_{K^t}^- \mathbf{x}_{K^t}^-. \quad (27)$$

We multiply the  $\mathbf{C}$  matrix by an estimate of the state vector computed using all the leaves of the tree associated to the most probable label.

### E. Multiple Model Data Association Tracker

Let us summarize the equations of the MMDA tracker. The tracker is based on the propagation of the *a posteriori* distribution of the hybrid state  $(\mathbf{x}_t, k_t)$  using a mixture of Gaussians. Three main steps are performed for each frame: prediction, image measurement and filtering which are described in Table I. As described previously, this method accounts for multiple dynamics and invalid observations (outliers). To deal with the first difficulty (multiple dynamics) we increase the number

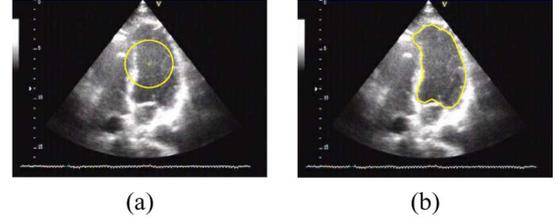


Fig. 8. (a) Initial contour (circle) and (b) contour estimate obtained by the S-PDAF method at the second iteration.

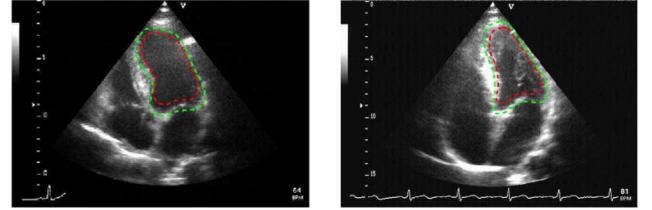


Fig. 9. Contours prediction using model “one” (red dashed line); and model “two” (dashed line) for two different sequences.

of mixture modes in the prediction step, whilst the second difficulty (outliers) is solved by data association methods. The tracker is summarized in Table I.

## VI. EXPERIMENTAL RESULTS

### A. Settings

The MMDA tracker was tested using a set of ten ultrasound sequences of the heart. These sequences correspond to real medical exams performed on ten different people. The sequences were obtained at a frame rate of 15 frames s<sup>-1</sup> using an ultrasound probe operating at 1.7 Mhz. The experimental setup is similar to the one described in [37].

Since ultrasound images are corrupted by multiplicative noise, they were preprocessed using a median filter with a support of  $4 \times 4$  pixels. Other alternative approaches could be used instead (e.g., statistical methods [18]).

The size of the ultrasound sequences used in this study are:

- case #1: 100 frames (3 cycles);
- case #2: 220 frames (8 cycles);
- case #3: 450 frames (25 cycles);
- case #4: 490 frames (26 cycles);
- case #5: 215 frames (9 cycles);
- case #6: 470 frames (19 cycles);
- case #7: 210 frames (8 cycles);
- case #8: 155 frames (7 cycles);
- case #9: 201 frames (10 cycles);
- case #10: 720 frames (42 cycles).

### B. Initialization

In this paper, we have reduced user intervention to a minimum. The user only has to specify a point inside the left ventricle. We then initialize a circular contour centered at this point and apply the S-PDAF method [38] to obtain an initial estimate for the left ventricle boundary. Both steps are shown in Fig. 8.

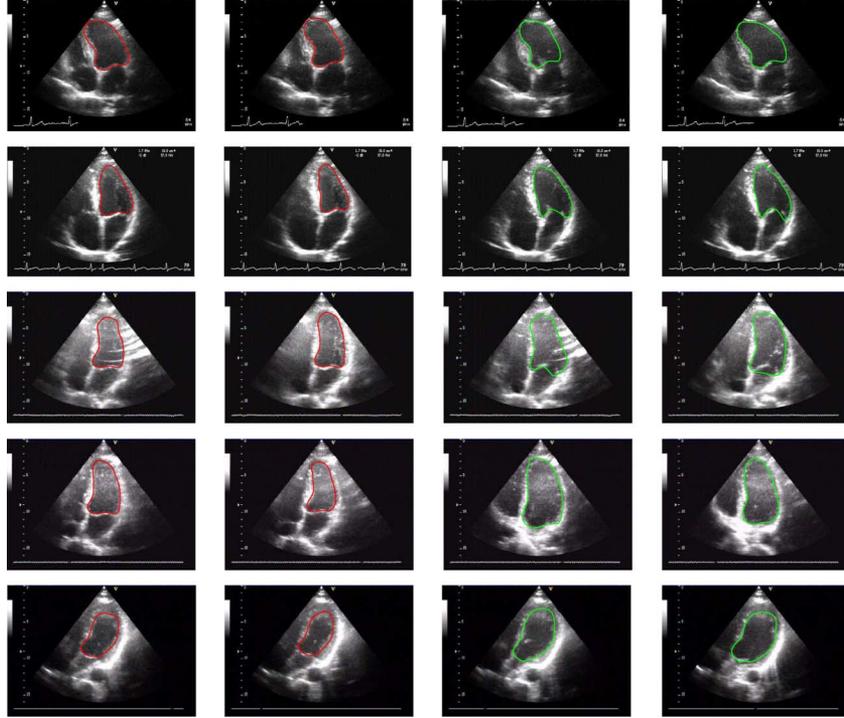


Fig. 10. Output of the MMDA for five patients (cases # 1 to # 5) in systole (red line) and diastole phase (green line). Each row corresponds to a different sequence.

### C. Multiple Model Tracking

We have used two dynamic models (see 8–9) to track the boundary of the left ventricle. The  $\mathbf{A}_k$  matrices used in this experiment are

$$\mathbf{A}_k = \text{diag}(\mathbf{l}_k, \mathbf{1}), \quad k = 1, 2 \quad (28)$$

with size  $(D + 2N_c) \times (D + 2N_c)$ ,  $\mathbf{1}$  is a  $2N_c \times 1$  vector,  $\mathbf{l}_k = [\Delta_k \ 1 \ 1 \ 1]$  is a  $D \times 1$  vector with  $\Delta_1 = 0.9$ ,  $\Delta_2 = 1.1$  which accounts for the contraction and expansion of the reference shape trying to model the motion of the heart in systole and diastole phases. These values were learned from the data using the first sequence of Fig. 10 (see details in [39] which proposes a learning method for switching dynamic models). The values of the matrix  $\mathbf{A}_k$  were obtained using the sequence # 1 for training.

Fig. 9 shows the configuration of the predicted contours using the two dynamic models in two examples. The red line accounts for the contraction, whilst the green line accounts for the expansion. There are two competing models trying to represent the data in each frame.

Figs. 10 and 11 show the results of the MMDA tracker in ten different cases. Each row shows four snapshots of the same sequence, two images extracted from the systole phase and two from the diastole phase. The snapshots were randomly selected in each phase.

These images display the motion of the endocardium and the motion of the mitral valve which exhibits smaller time constants. The color of the contour identifies the most probable model chosen by the tracker: red means model 1 (contraction) and green model 2 (expansion). In the diastole, the ventricle expands and the mitral valve is open. In the systole phase, the ventricle contracts, and the mitral valve closes. The MMDA tracker

automatically chooses the model (contraction or expansion), associated to the phase of the cardiac cycle.

Ventricular systole is the interval between the QRS complex (peak of the ECG signal) and the end of the T wave (the Q-T interval). The systole ends when the aortic valve closes and the mitral valve opens and corresponds to the moment of minimal LV size/volume (known as reduced ejection). In the diastole, the MMDA tracker chooses the expansion model (model “two”—green line). This phase starts after the occurrence of the T-wave (known as isovolumetric relaxation) until a peak of the ECG is observed (known as atrial systole).

The model is not able to accurately track rapid changes of the mitral valve. This can be seen in the first row of the Fig. 10 (diastole phase). The valve suddenly opens and produces speckle noise near its boundary. Therefore, the contour becomes more complex in this region. However, this problem does not affect the choice of the active model.

The remaining rows of the Figs. 10 and 11 show more results of the tracker exhibiting a remarkable robustness in the presence of low signal to noise ratio.

Fig. 12 shows the ideal label sequence (top) and the label estimates chosen by the tracker (down) for the first sequence during three cardiac cycles. Fig. 12(a) shows the estimates obtained without using switching probabilities (all transitions are equiprobable) and Fig. 12(a) shows the estimated obtained using the switching probabilities  $\mathbf{T}$ . The ideal label was built by hand from the ECG signal.

Without transition probabilities the tracker provides a reasonable estimate of the cardiac phase but fails in four frames. This experiment was done assuming that transitions are independent random variables. This is, of course, an unrealistic assumption since there is a correlation between consecutive labels.

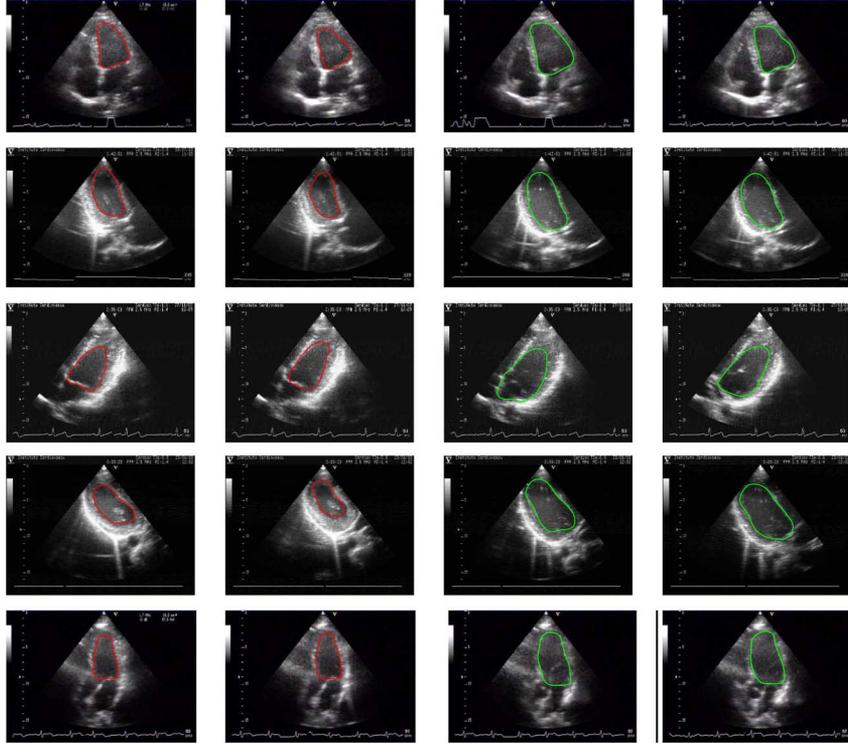


Fig. 11. Output of the MMDA for five patients (cases #6 to #10) in systole (red line) and diastole phase (green line). Each row corresponds to a different sequence.

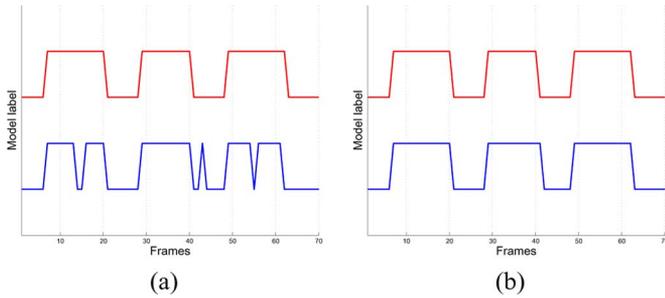


Fig. 12. Ideal label sequence and estimated signal label sequence during three cardiac cycles.

The same experiment was repeated with transition probabilities  $T_{ii} = 0.8$ ,  $i = 1, 2$ . In this case, a much better performance was achieved [see Fig. 12(b)].

#### D. Objective Evaluation

The performance of the tracker was measured by comparing the contour estimates with reference contours provided by a cardiologist of Amadora-Sintra Hospital. For the sake of comparison, we also present the results obtained with a state-of-the-art tracker for the left ventricle recently proposed by Comaniciu *et al.* (see [22]) and applied on the same data.

In these tests, we have selected four images from each cardiac cycle (two images in the systole phase and two images in the diastole phase) and asked a medical doctor to manually define the left ventricle contour for each of these images. This was done for the sequences #5 and #6 [see Figs. 10 and 11].

The cardiologist segmented 80 images: 40 images from each sequence, 20 of them extracted during the systole phase and the

other 20 during the diastole phase. The frames were randomly selected in each of these phases.

Three metrics were used in these tests to compare the output of the tracker with the reference contours: the Hausdorff distance, the average distance, and the Hammoude distance [40] which is used when a single object is present in the image [41]. We will briefly define the metrics.

Let  $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_x}\}$ , and  $\mathcal{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{N_y}\}$ , be two sets of points obtained by sampling the estimated contour and the reference contour. The smallest distance from a point  $\mathbf{x}_i$  to a curve  $\mathcal{Y}$  is defined as

$$d(\mathbf{x}_i, \mathcal{Y}) = \min_j \|\mathbf{y}_j - \mathbf{x}_i\|. \quad (29)$$

The average distance between the two sets  $\mathcal{X}$ ,  $\mathcal{Y}$  is

$$d_{av} = \frac{1}{N_x} \sum_{i=1}^{N_x} d(\mathbf{x}_i, \mathcal{Y}) \quad (30)$$

and the Hausdorff distance between both sets is the largest distance from a point to the other set [42]

$$d_{\max}(\mathcal{X}, \mathcal{Y}) = \max \left( \max_i \{d(\mathbf{x}_i, \mathcal{Y})\}, \max_j \{d(\mathbf{y}_j, \mathcal{X})\} \right). \quad (31)$$

The Hammoude distance is also simple [40]. Let  $\mathcal{X}$ ,  $\mathcal{Y}$  be the image regions inside the two contours. We compute the number of points which belongs to only one of these regions and normalize it by the number of points of the union

$$d_H = \frac{\#((\mathcal{X} \cup \mathcal{Y}) - (\mathcal{X} \cap \mathcal{Y}))}{\#(\mathcal{X} \cup \mathcal{Y})}. \quad (32)$$

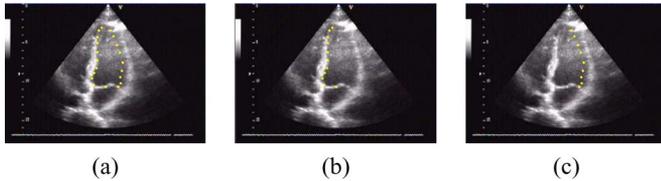


Fig. 13. (a) Ground truth provided by the cardiologist, (b) left side and (c) right side.

TABLE II  
EVALUATION OF CASE CASE # 5 USING THREE METRICS BETWEEN THE ESTIMATED CONTOUR AND THE REFERENCE CONTOUR PROVIDED BY THE SPECIALIST

case # 5	left side contour		right side contour	
	MMDA	Comaniciu	MMDA	Comaniciu
“Hammoude”	0.20	0.25	0.28	0.15
“average”	2.96	3.53	6.23	4.34
“Hausdorff”	13.78	15.88	26.48	22.49

TABLE III  
EVALUATION OF CASE CASE # 6 USING THREE METRICS BETWEEN THE ESTIMATED CONTOUR AND THE REFERENCE CONTOUR PROVIDED BY THE SPECIALIST

case # 6	left side contour		right side contour	
	MMDA	Comaniciu	MMDA	Comaniciu
“Hammoude”	0.19	0.30	0.28	0.13
“average”	2.7	3.05	6.98	3.69
“Hausdorff”	6.29	11.19	27.7	23.77

The Hausdorff and Hammoude distances are metrics since they verify the axioms of a metric. The average distance is not a metric since it is not symmetric, in general,  $d_{av} = (\mathcal{X}, \mathcal{Y}) \neq d_{av} = (\mathcal{Y}, \mathcal{X})$ .

If we represent the regions  $\mathcal{X}$ ,  $\mathcal{Y}$  by two binary images, the  $d_H$  is equal to the Hamming distance between the binary images, normalized by the area of their sum. The Hammoude distance is a normalized version of the Hamming distance which is widely used to compare binary strings and images in information theory.

These three distances allow a comparison between pairs of contours. To evaluate the algorithm during several frames, we compute the average value of these metrics. Other statistics can be used for this purpose, as well [41], [43].

The quality of ultrasound images is an important issue since it influences the assessment made by a specialist. In this case, the cardiologists considered that the left side of the ventricle usually had a higher quality than the right side and they suggested that a separate evaluation of both regions should be made.

Fig. 13(a) shows the ground truth for one snapshot of the case # 5. Fig. 13(b) and (c) shows the ground truth for the left and right parts of the image.

Tables II and III show the results of both trackers in the estimation of sides of the LV contour in case # 5 and # 6, respectively. It can be seen that the MMDA tracker provides better results for the left side, than Comaniciu’s tracker in the left part of

TABLE IV  
EVALUATION OF CASE CASE # 5 (WHOLE CONTOUR) USING THREE METRICS BETWEEN THE ESTIMATED CONTOUR AND THE REFERENCE CONTOUR PROVIDED BY THE SPECIALIST

case # 5	whole contour	
	MMDA	Comaniciu
“Hammoude”	0.24	0.21
“average”	4.64	3.90
“Hausdorff”	21.41	20.48

TABLE V  
EVALUATION OF CASE CASE # 6 (WHOLE CONTOUR) USING THREE METRICS BETWEEN THE ESTIMATED CONTOUR AND THE REFERENCE CONTOUR PROVIDED BY THE SPECIALIST

case # 6	whole contour	
	MMDA	Comaniciu
“Hammoude”	0.24	0.19
“average”	4.87	3.37
“Hausdorff”	20.26	17.21

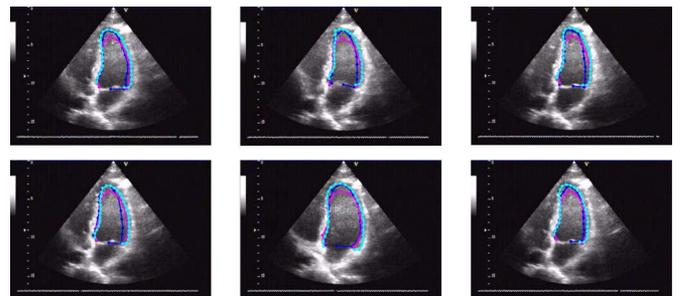


Fig. 14. Ground truth and estimated contours: (blue line) medical ground truth; (cyan line) contour estimates obtained with MMDA; (magenta line) and Comaniciu’s tracker.

the ventricle. However, in right side (i.e., images of poor quality) Comaniciu’s tracker performs better.

Tables IV and V show the performance of both trackers in the whole contour. The results obtained by both trackers are similar.

In case # 5, we obtained Hausdorff distances of 21.41 pixels for the case MMDA, and 20.48 pixels for the case of Comaniciu’s tracker. If we normalize by the average perimeter (249 pixels), we obtain a relative error of 8.59% for the MMDA, while the Comaniciu’s method has error of 8.22%. The average metric leads to an error of 1.86% for the MMDA, and 1.56% for Comaniciu tracker.

In case # 6 (Table V), the LV has an average perimeter of 222 pixels. Concerning the Hausdorff distance, we obtain a relative error of 9.13% for the MMDA and 7.76% for Comaniciu tracker. The average distance is 2.2% for MMDA, and 1.52% for the Comaniciu’s method. We thus conclude that the MMDA has a tracking performance comparable to a state-of-the-art tracker.

Figs. 14 and 15 show the contour estimates of the MMDA tracker, as well as the tracker proposed by Comaniciu. Superimposed to these contours, we also draw (in darker blue) the medical segmentation (ground truth) of the frames.

Fig. 16 shows the three best (top row) and worst contour estimates obtained with MMDA. The Hammoude metric is also shown.

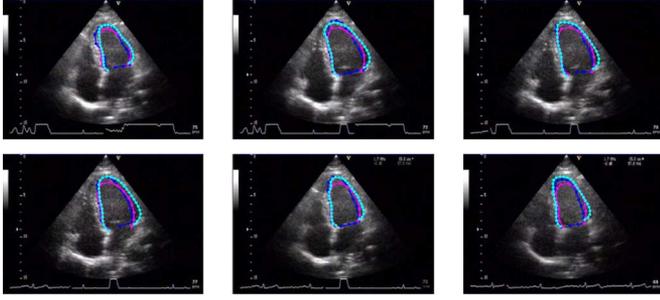


Fig. 15. Ground truth and estimated contours: (blue line) medical ground truth; (cyan line) contour estimates obtained with MMDA; (magenta line) and Co-maniciu's tracker.

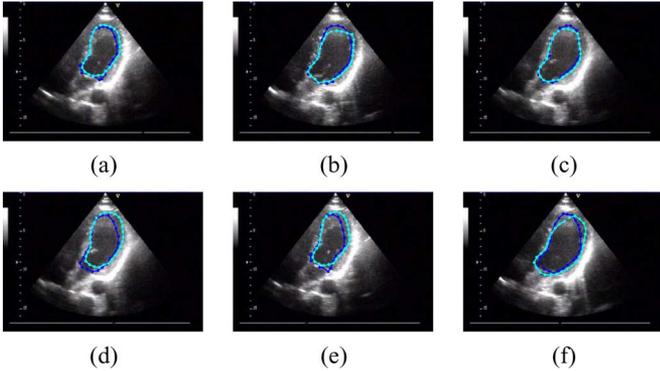


Fig. 16. Three (top row) best and (bottom row) worst contour estimates obtained with MMDA. The Hamoude metric obtained for each case is: (a) 0.16, (b) 0.16, (c) 0.11, (d) 0.21, (e) 0.26, (f) 0.25.

## VII. CONCLUSION

This paper presents a robust algorithm for tracking the boundary of the left ventricle in sequences of ultrasound images. The evolution of the shape and motion parameters is described by a bank of switched dynamic systems which is able to represent complex motion and shape dynamics. The visual features detected in the ultrasound images contain a large number of outliers which do not belong to the boundary of the ventricle. A robust filtering algorithm is proposed which is able to deal with multiple dynamics and invalid observations at the same time. This is accomplished by propagating the *a posteriori* density of the unknown parameters using a tree of probability data association filters.

It was experimentally observed that the proposed method efficiently estimates the ventricle contour in ultrasound images, corrupted by speckle noise and outliers, without losing track. The MMDA tracker provides two levels of information. The output of the first level is an estimate of the heart boundary and the output of the second level is a binary signal which defines the active model used for tracking and discriminates the two phases of the cardiac cycle.

## APPENDIX

*Shape Models:* This Appendix describes the shape model used in this paper based on a shape template (reference shape). The model considers global motion of the template plus a local deformation.

It is assumed in this paper that the object boundary  $v(s) \in \mathbb{R}^2$  is given by<sup>2</sup>

$$v(s) = \mathcal{T}r(s) + d(s) \quad (33)$$

where  $r(s)$  is a reference shape,  $\mathcal{T}$  is a geometric transformation,  $d(s)$  is the curve deformation, and  $s$  is a parameter defining the position of a point in the curve. It is assumed that  $d(s)$  is a B-spline curve [35]

$$d(s) = \sum_{n=1}^{N_c} \theta_n \phi_n(s) \quad (34)$$

where  $N_c$  is the number of control points,  $\phi_n(s)$  are B-spline basis functions, and  $\theta_n \in \mathbb{R}^2$  are the control points.

The global transformation used in this work is an Euclidean similarity. Therefore, the object boundary  $v(s) = (v_x(s), v_y(s))$  is given by

$$\begin{cases} v_x(s) = x_1 r_x(s) - x_3 r_y(s) + x_2 + d_x(s) \\ v_y(s) = x_3 r_x(s) + x_1 r_y(s) + x_4 + d_y(s) \end{cases} \quad (35)$$

where  $\mathbf{x} = [x_1, \dots, x_4]^T$  are the motion parameters.

Suppose we want to compute the position of  $N$  points on the curve and let  $\mathbf{y}_t$  be a vector with their coordinates

$$\mathbf{y}_t = [y_x(s_1), \dots, y_x(s_N), y_y(s_1), \dots, y_y(s_N)]^T. \quad (36)$$

Using (35), we obtain

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t \quad (37)$$

where  $\mathbf{C}$  is a  $2N \times (D + 2N_c)$  matrix,  $\mathbf{x}_t$  is a  $(D + 2N_c) \times 1$  vector given by

$$\mathbf{x}_t = [x_1, \dots, x_D, \theta_{x1}, \dots, \theta_{xN_c}, \theta_{y1}, \dots, \theta_{yN_c}]^T \quad (38)$$

where  $D = 4$  is the number of motion coefficients defining the rotation, scale and translation and  $\theta$  the deformation parameters. The vector  $\mathbf{x}_t$  will be denoted as state vector since it summarizes past information about shape evolution.

The shape matrix  $\mathbf{C}$  is given by

$$\mathbf{C} = \begin{bmatrix} \mathbf{M}_x & -\mathbf{M}_y & \mathbf{B} & \mathbf{O} \\ \mathbf{M}_y & \mathbf{M}_x & \mathbf{O} & \mathbf{B} \end{bmatrix} \quad (39)$$

where

$$\mathbf{M}_x = \begin{bmatrix} r_x(s_1) & 1 \\ \vdots & \vdots \\ r_x(s_N) & 1 \end{bmatrix}, \quad \mathbf{M}_y = \begin{bmatrix} r_y(s_1) & 0 \\ \vdots & \vdots \\ r_y(s_N) & 0 \end{bmatrix} \quad (40)$$

$$\mathbf{B} = \begin{bmatrix} \phi_1(s_1) & \phi_2(s_1) & \dots & \phi_{N_c}(s_1) \\ \phi_1(s_2) & \phi_2(s_2) & \dots & \phi_{N_c}(s_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(s_N) & \phi_2(s_N) & \dots & \phi_{N_c}(s_N) \end{bmatrix} \quad (41)$$

$\mathbf{B}$  is a  $N \times N_c$  interpolation B-spline matrix,  $\mathbf{M}_x$ ,  $\mathbf{M}_y$  are  $N \times 2$  matrices defining the reference shape and  $\mathbf{O}$  is a  $N \times N_c$  null matrix.

<sup>2</sup>For the sake of notation simplicity, we omit the time dependence  $t$  on the curves herein described.

In Section VI (experimental results), the following matrices are used:

$$\mathbf{P} = \alpha \mathbf{I}_P, \quad \mathbf{Q} = \beta \mathbf{I}_Q, \quad \mathbf{R} = \gamma \mathbf{I}_R \quad (42)$$

where  $\mathbf{I}_P, \mathbf{I}_Q$  are  $(D+2N_c) \times (D+2N_c)$  identity matrices and  $\mathbf{I}_R$  is a  $2N \times 2N$  identity matrix,  $\alpha, \beta, \gamma$  are constants.

*Mixture Coefficients of MMDA Tracker:* This section addresses the update of  $c_{K^t}$  in the prediction and filtering step.

*Prediction:* The mixture coefficients  $c_{K^t}^-$  can be recursively obtained by

$$\begin{aligned} c_{K^t}^- &= P(\mathbf{k}^t | \mathbf{Y}^{t-1}) = P(k_t, \mathbf{k}^{t-1} | \mathbf{Y}^{t-1}) \\ &= P(k_t | \mathbf{k}^{t-1}, \mathbf{Y}^{t-1}) P(\mathbf{k}^{t-1} | \mathbf{Y}^{t-1}) \end{aligned} \quad (43)$$

since  $P(k_t | \mathbf{k}^{t-1}, \mathbf{Y}^{t-1}) = \mathbf{T}_{k_{t-1}, k_t}$  and  $c_{K^{t-1}} = P(\mathbf{k}^{t-1} | \mathbf{Y}^{t-1})$ , (43) can be rewritten as

$$c_{K^t}^- = \mathbf{T}_{k_{t-1}, k_t} c_{K^{t-1}}. \quad (44)$$

*Filtering:*

$$c_{K^t} \triangleq \frac{p(\mathbf{k}^t, \mathbf{Y}^t)}{p(\mathbf{Y}^t)} = \frac{p(\mathbf{y}_t | \mathbf{k}^t, \mathbf{Y}^{t-1}) p(\mathbf{k}^t, \mathbf{Y}^{t-1})}{p(\mathbf{Y}^t)}. \quad (45)$$

Since the term  $p(\mathbf{k}^t, \mathbf{Y}^{t-1})$  is related with the prediction step, (45) can be written as

$$\begin{aligned} c_{K^t} &= \gamma c_{K^t}^- \int p(\mathbf{y}_t | \mathbf{k}^t, \mathbf{Y}^{t-1}, \mathbf{x}_t) \\ &\quad \times p(\mathbf{x}_t | \mathbf{k}^t, \mathbf{Y}^{t-1}) d\mathbf{x}_t \\ &= \gamma c_{K^t}^- \int \sum_i p(\mathbf{y}_t | i_t, \mathbf{k}^t, \mathbf{Y}^{t-1}, \mathbf{x}_t) \\ &\quad \times P(i_t | \mathbf{k}^t, \mathbf{Y}^{t-1}, \mathbf{x}_t) \\ &\quad \times p(\mathbf{x}_t | \mathbf{k}^t, \mathbf{Y}^{t-1}) d\mathbf{x}_t \\ &= \gamma c_{K^t}^- \sum_i \alpha_{K^t}^i \int p(\mathbf{y}_t | i_t, \mathbf{k}^t, \mathbf{Y}^{t-1}, \mathbf{x}_t) \\ &\quad \times p(\mathbf{x}_t | \mathbf{k}^t, \mathbf{Y}^{t-1}) d\mathbf{x}_t \end{aligned} \quad (46)$$

with  $\gamma = (p(\mathbf{Y}^{t-1})/p(\mathbf{Y}^t))$ .

The observations  $\mathbf{y}_t$  is a set of  $M$  strokes. With the hypothesis made in Section V-B, we can write  $p(\mathbf{y}_t | i_t, \mathbf{k}^t, \mathbf{Y}^{t-1}) = p(\mathbf{y}_t | i_t, \mathbf{k}^t, \mathbf{b}, \mathbf{e}, M, \mathbf{Y}^{t-1})$  where  $\mathbf{b} = \{b^1, \dots, b^M\}$ ,  $\mathbf{e} = \{e^1, \dots, e^M\}$  define the beginning and the end of the strokes. Using the following observation model:

$$p(\mathbf{y}_t | i_t, \mathbf{k}^t, \mathbf{b}, \mathbf{e}, M, \mathbf{Y}^{t-1}) = \prod_{j=1}^M \prod_{n=b^j}^{e^j} p(\mathbf{y}_t^j(s_n) | I_{it}^j, \mathbf{k}^t) \quad (47)$$

where  $\mathbf{y}_t^j(s_n)$  is the feature point belonging to the  $j$ th stroke detected in the vicinity of  $s_n$ . It is assumed that the visual features have uniform distribution in the search area if  $I_{it}^j = 0$  (classified as unreliable) and Gaussian distribution if  $I_{it}^j = 1$  (classified as reliable). Therefore,  $\mathcal{E}_{it}^j(s_n) = p(\mathbf{y}_t^j(s_n) | I_{it}^j, \mathbf{k}^t)$

$$\mathcal{E}_{it}^j(s_n) = \begin{cases} V_t(s_n)^{-1}, & \text{if } I_{it}^j = 0 \\ \rho^{-1} \mathcal{N}(\boldsymbol{\nu}_t^j(s_n); 0, \mathbf{S}_t(s_n)), & \text{otherwise} \end{cases} \quad (48)$$

where  $V_t(s_n)$  is the length of the search area,  $\rho$  is the normalization constant,  $\boldsymbol{\nu}_t^j(s_n) = \mathbf{y}_t^j(s_n) - \mathbf{C}(s_n) \hat{\mathbf{x}}_t^-$  is the innovation associated to the  $j$ th stroke, and  $\mathbf{S}_t(s_n) = \mathbf{C}(s_n) \mathbf{P}_{K^t}^- \mathbf{C}(s_n)^T + \mathbf{R}(s_n)$  is the covariance of the innovation vector where  $\mathbf{C}(s_n)$  and  $\mathbf{R}(s_n)$  are the output matrix and noise covariance associated to the  $n$ th sample of the object contour. Replacing (48) in (47) into (46) leads to

$$c_{K^t} = \gamma c_{K^t}^- \sum_i \alpha_{K^t}^i \prod_{j=1}^M \prod_{n=b^j}^{e^j} \mathcal{E}_{it}^j(s_n). \quad (49)$$

*State Vector and Covariance Matrix of MMDA Tracker:*

Let us compute the expected value of state vector given the current observations, in the case of multiple data interpretations  $i_t$

$$\begin{aligned} \hat{\mathbf{x}}_{K^t} &= E\{\mathbf{x}_t | \mathbf{k}^t, \mathbf{Y}^t\} = \int \mathbf{x}_t p(\mathbf{x}_t | \mathbf{k}^t, \mathbf{Y}^t) d\mathbf{x}_t \\ &= \int \mathbf{x}_t \sum_i p(\mathbf{x}_t, i_t | \mathbf{k}^t, \mathbf{Y}^t) d\mathbf{x}_t \\ &= \sum_i \int \mathbf{x}_t p(\mathbf{x}_t | i_t, \mathbf{k}^t, \mathbf{Y}^t) \\ &\quad \times P(i_t | \mathbf{k}^t, \mathbf{Y}^t) d\mathbf{x}_t. \end{aligned} \quad (50)$$

Let us define the association probability of the  $i$ th interpretation and the state mean  $i_t$  as

$$\alpha_{K^t}^i \triangleq P(i_t | \mathbf{k}^t, \mathbf{Y}^t) \quad (51)$$

$$\hat{\mathbf{x}}_{K^t}^i = E\{\mathbf{x}_t | i_t, \mathbf{k}^t, \mathbf{Y}^t\}. \quad (52)$$

Equation (50) can be rewritten as

$$\hat{\mathbf{x}}_{K^t} = \sum_{i=0}^{m_t} \alpha_{K^t}^i \hat{\mathbf{x}}_{K^t}^i \quad (53)$$

where  $m_t$  is the number of data interpretations at time  $t$ . Therefore, the state estimate  $\hat{\mathbf{x}}_{K^t}$  is a weighted sum of the state estimates  $\hat{\mathbf{x}}_{K^t}^i$  obtained for each interpretation  $i_t$  and updated by Kalman filtering

$$\hat{\mathbf{x}}_{K^t}^i = \hat{\mathbf{x}}_{K^t}^- + \mathbf{K}_t^i \boldsymbol{\nu}_t^i, \quad (54)$$

where  $\hat{\mathbf{x}}_{K^t}^-$  is the component state prediction in the path  $\mathbf{k}^t$ ,  $\mathbf{K}_t^i, \boldsymbol{\nu}_t^i$  are the Kalman gain and innovation associated to the interpretation  $i_t$  on the tree leaf. Replacing (54) in (53) leads to

$$\hat{\mathbf{x}}_{K^t} = \hat{\mathbf{x}}_{K^t}^- + \sum_{i=1}^{m_t} \alpha_{K^t}^i \mathbf{K}_t^i \boldsymbol{\nu}_t^i. \quad (55)$$

The covariance of the state estimate is<sup>3</sup>

$$\begin{aligned} \mathbf{P}_t &= E\{[\mathbf{x}_t - \hat{\mathbf{x}}_t][\mathbf{x}_t - \hat{\mathbf{x}}_t]^T | \mathbf{Y}^t\} \\ &= E\left\{\left(\begin{array}{cccc} \mathbf{x}_t \mathbf{x}_t^T & -\mathbf{x}_t \hat{\mathbf{x}}_t^T & -\hat{\mathbf{x}}_t \mathbf{x}_t^T & +\hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T \\ \mathbf{P}^1 & \mathbf{P}^2 & \mathbf{P}^{2^T} & \mathbf{P}^3 \end{array}\right) | \mathbf{Y}^t\right\} \end{aligned} \quad (56)$$

<sup>3</sup>For the sake of simplicity, we drop the dependence on the path  $K^t$ .

where

$$\mathbf{P}^1 \triangleq E \{ \mathbf{x}_t \mathbf{x}_t^T \mid \mathbf{Y}^t \} = \sum_{i=0}^{m_t} E \{ \mathbf{x}_t \mathbf{x}_t^T \mid i_t, \mathbf{Y}^t \} \alpha_t^i. \quad (57)$$

The first term is

$$\mathbf{P}^1 = \sum_{i=0}^{m_t} \alpha_t^i [\mathbf{P}_t^i + \hat{\mathbf{x}}_t^i \hat{\mathbf{x}}_t^{iT}]. \quad (58)$$

The second term in (56) is

$$\begin{aligned} \mathbf{P}^2 &\triangleq - \sum_{i=0}^{m_t} E \{ \mathbf{x}_t \hat{\mathbf{x}}_t^T \mid i_t, \mathbf{Y}^t \} \alpha_t^i \\ &= - \left( \sum_{i=0}^{m_t} E \{ \mathbf{x}_t \mid i_t, \mathbf{Y}^t \} \alpha_t^i \right) \hat{\mathbf{x}}_t^T \\ &= - \left( \sum_{i=0}^{m_t} \hat{\mathbf{x}}_t^i \alpha_t^i \right) \hat{\mathbf{x}}_t^T = -\hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T = \mathbf{P}^{2T}. \end{aligned} \quad (59)$$

The third term is

$$\mathbf{P}^3 \triangleq \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T \sum_{i=0}^{m_t} \alpha_t^i = \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T = -\mathbf{P}^2. \quad (60)$$

Combining (57) and (59) into (56) yields

$$\mathbf{P} = \sum_{i=0}^{m_t} \alpha_t^i [\mathbf{P}_t^i + \hat{\mathbf{x}}_t^i \hat{\mathbf{x}}_t^{iT}] - \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T \quad (61)$$

where

$$\mathbf{P}_t^i = (\mathbf{I} - \mathbf{K}_t^i \mathbf{C}^i) \mathbf{P}^-. \quad (62)$$

Therefore, (61) can be written as

$$\mathbf{P} = \left[ \mathbf{I} - \sum_{i=1}^{m_t} \alpha_t^i \mathbf{K}_t^i \mathbf{C}^i \right] \mathbf{P}_t^- + \sum_{i=0}^{m_t} \alpha_t^i \hat{\mathbf{x}}_t^i \hat{\mathbf{x}}_t^{iT} - \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t^T. \quad (63)$$

#### ACKNOWLEDGMENT

The authors would like to thank Dr. D. Comaniciu for kindly providing tracking results obtained with his method [2] and for helpful comments. They would also like to thank Prof. J. Sanches of the Institute of Systems and Robotics and Prof. F. Pinto from the Faculty of Medicine of the University of Lisbon for the ultrasound images used in this work, as well Dr. A. Freitas from the Hospital Amadora-Sintra for providing the ground truth segmentation of the left ventricle.

#### REFERENCES

[1] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. New York: Academic, 1988.  
 [2] X. Zhou, D. Comaniciu, and A. Gupta, "An information fusion framework for robust shape tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 1, pp. 115–129, Jan. 2005.  
 [3] K. Schmailzl and O. Ormerod, *Ultrasound in Cardiology*. Oxford, U.K.: Blackwell, 1994.

[4] M. Mignotte, J. Meunier, and J.-C. Terdif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable templates and Markov random fields," *Pattern Anal. Appl.*, vol. 4, no. 4, pp. 256–271, 2001.  
 [5] Y. Akgul and C. Kambhamettu, "A coarse-to-fine deformable contour optimization framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 174–186, Feb. 2003.  
 [6] S. Minut and G. Stockman, "Interpolation snakes for border detection in noisy images," in *Proc. Int. Conf. Computer Vision Theory and Applications*, 2006, pp. 25–28.  
 [7] M. A. T. Figueiredo, J. M. N. Leitao, and A. K. Jain, "Unsupervised contour representation and estimation using B-splines and a minimum description length criterion," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1075–1087, Jun. 2000.  
 [8] N. Paragios, M.-P. Jolly, M. Taron, and R. Ramaraj, "Active shape models and segmentation of the left ventricle in echocardiography," presented at the Int. Conf. on Scale Space Theories and PDEs Methods in Computer Vision, 2005.  
 [9] D. Adam, H. Hareuveni, and S. Sideman, "Semiautomated border tracking of cine echocardiogram ventricular images," *IEEE Trans. Med. Imag.*, vol. MI-6, no. 3, pp. 266–271, Sep. 1987.  
 [10] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 11, pp. 1061–1075, Nov. 1992.  
 [11] D. Suter and F. Chen, "Left ventricular motion reconstruction based on elastic vector splines," *IEEE Trans. Med. Imag.*, vol. 19, no. 4, pp. 295–305, Apr. 2000.  
 [12] T. Cootes, C. Taylor, D. Cooper, and J. Haslam, "The use of active shape models for locating structures in medical images," *Image Vis. Comput.*, vol. 12, no. 6, pp. 355–366, Jul. 1994.  
 [13] T. Cootes and C. Taylor, "Statistical models for appearance for medical image analysis and computer vision," in *Proc. SPIE Med. Imag.*, 2001, pp. 236–248.  
 [14] B. Georgescu, X. Zhou, D. Comaniciu, and B. Rao, "Real-time multi-model tracking of myocardium in echocardiography using robust information fusion," presented at the Int. Conf. Medical Image Computing and Computer Assisted Intervention, 2004.  
 [15] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 15, no. 3, pp. 290–298, Jun. 1996.  
 [16] J. McEachen and J. Duncan, "Shape-based tracking of left ventricular wall motion," *IEEE Trans. Med. Imag.*, vol. 16, no. 3, pp. 270–283, Jun. 1997.  
 [17] G. E. Mailloux, F. Langlois, P. Y. Simard, and M. Bertrand, "Restoration of the velocity field of the heart from two-dimensional echocardiograms," *IEEE Trans. Med. Imag.*, vol. 8, no. 2, pp. 143–153, Jun. 1989.  
 [18] J. M. Dias and J. Leitão, "Wall position and thickness estimation from sequences of echocardiograms images," *IEEE Trans. Med. Imag.*, vol. 15, no. 2, pp. 25–38, Feb. 1996.  
 [19] G. Jacob, A. Noble, and A. Blake, "Robust contour tracking in echocardiographic sequence," in *Proc. 6th Int. Conf. Computer Vision*, Bombay, India, 1998, pp. 408–413.  
 [20] N. Friedland and D. Adam, "Automatic ventricular boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans. Med. Imag.*, vol. 8, no. 4, pp. 344–353, Dec. 1989.  
 [21] J. Sanches, J. Nascimento, and J. Marques, "An unified framework for bayesian denoising for several medical and biological imaging modalities," presented at the IEEE Engineering in Medicine and Biology Soc. Annu. Int. Conf., Aug. 2007.  
 [22] D. Comaniciu, X. Zhou, and S. Krishnan, "Robust real-time myocardial border tracking for echocardiography: An information fusion approach," *IEEE Trans. Med. Imag.*, vol. 23, no. 7, pp. 849–860, Jul. 2004.  
 [23] D. Terzopoulos and R. Szeliski, *Tracking With Kalman Snakes*. Cambridge, MA: MIT Press, 1993.  
 [24] M. Isard and A. Blake, "A mixed-state condensation tracker with automatic model-switching," in *Proc. Int. Conf. Computer Vision*, 1998, pp. 107–112.  
 [25] Y. Kanazawa and K. Kanatani, "Do we really have to consider covariance matrices for image features?," in *Proc. 8th Int. Conf. Computer Vision*, 2001, vol. 37, pp. 586–591.  
 [26] K. Narandra and C. Xiang, "Adaptive control of discrete-time systems using multiple models," *IEEE Trans. Autom. Control*, vol. 45, no. 9, pp. 1699–1686, Sep. 2000.  
 [27] K. S. Narandra and J. Balakrishnan, "Adaptive control using multiple models," *IEEE Trans. Autom. Control*, vol. 42, no. 2, pp. 171–187, Feb. 1997.  
 [28] J. M. Lemos, L. M. Rato, and E. Mosca, "Integrating predictive and switching control: Basic concepts and an experimental case study," in *Nonlinear Model Predictive Control*, F. Allgöwer and A. Zheng, Eds. Basel, Switzerland: Birkhäuser Verlag, 2000, pp. 181–190.

- [29] C. Chang and M. Athans, "State estimation for discrete systems with switching parameters," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, no. 3, pp. 418–425, May 1978.
- [30] V. Petridis and A. Kehagias, "A multi-model algorithm for parameter estimation of time varying nonlinear systems," *Automatica*, vol. 34, no. 4, pp. 469–475, 1998.
- [31] J. Tugnait, "Detection and estimation for abruptly changing systems," *Automatica*, vol. 18, no. 5, pp. 607–615, 1982.
- [32] J. Evans and R. Evans, "Image-enhanced multiple model tracking," *Automatica*, vol. 35, no. 11, pp. 1769–1786, 1999.
- [33] J. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background on image processing," in *Proc. IEEE Int. Conf. Image Processing*, Vancouver, BC, Canada, 2002, vol. 3, pp. 82–85.
- [34] P. Abolmaesumi and M. R. Sirouspour, "An interacting multiple model probabilistic data association filter for cavity boundary extraction from ultrasound images," *IEEE Trans. Med. Imag.*, vol. 23, no. 6, pp. 772–784, Jun. 2004.
- [35] A. Blake and M. Isard, *Active Contours*. New York: Springer, 1998.
- [36] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [37] J. Jensen, *Estimation of Blood Velocities Using Ultrasound, a Signal Processing Approach*. Cambridge, MA: Cambridge Univ. Press, 1996.
- [38] J. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background," *IEEE Trans. Multimedia*, vol. 6, no. 8, pp. 852–861, Dec. 2004.
- [39] G. Celeux, J. Nascimento, and J. S. Marques, "Learning switching dynamic models for objects tracking," *Pattern Recognit.*, vol. 37, no. 9, pp. 1841–1853, Sep. 2004.
- [40] A. Hammoude, "Computer-assited endocardial border identification from a sequence of two-dimensional echocardiographic images," Ph.D. dissertation, Univ. Washington, Seattle, 1988.
- [41] V. Chalana and Y. Kim, "A methodology for evaluation of boundary detection algorithms on medical images," *IEEE Trans. Med. Imag.*, vol. 16, no. 10, pp. 642–652, Oct. 1997.
- [42] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.
- [43] C. Alberola-Lopez, M. Martin-Fernandez, and J. Ruiz-Alzola, "Comments on: A methodology for evaluation of boundary detection algorithms on medical images," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 658–660, May 2004.



**Jacinto C. Nascimento** (M'06) received the E.E. degree from the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal, and the M.Sc. and Ph.D. degrees from the Instituto Superior Técnico (IST), Lisbon, in 1995, 1998, and 2003, respectively.

Currently, he is a Postdoctorate Researcher at IST affiliated with the Institute for Systems and Robotics (ISR). His research interests are image processing, shape tracking, robust estimation, medical imaging, and video surveillance.



**Jorge S. Marques** received the E.E., M.Sc., and Ph.D. degrees, and the aggregation title from the Technical University of Lisbon, Lisbon, Portugal, in 1981, 1984, 1990, and 2002, respectively.

Currently, he is an Associate Professor with the Electrical Engineering Department, Instituto Superior Técnico, Lisbon, where he is also a Researcher at the Institute for Systems and Robotics. He has published over 120 papers in international journals and conferences and he is the author of the book *Pattern Recognition: Statistical and Neural Methods*, 2nd ed. (in Portuguese) (IST, 2005). His research interests are in the areas of statistical image processing, shape analysis, and pattern recognition.

Dr. Marques was President of the Portuguese Association for Pattern Recognition (APRP) from 2001–2003 and a member of the IAPR Governing Board.