Tracking the left ventricle in ultrasound images based on total variation denoising *

Jacinto C. Nascimento João M. Sanches Jorge S. Marques
{jan,jmrs,jsm}@isr.ist.utl.pt
Instituto de Sistemas e Robótica - Instituto Superior Técnico

Abstract. Tracking the Left Ventricle (LV) in ultrasound sequences remains a challenge due to speckle noise, low SNR and lack of contrast. Therefore, it is usually difficult to obtain accurate estimates of the LV cavities since feature detectors produce a large number of outliers. This paper presents an algorithm which combines two main operations: i) a novel denoising algorithm based on the Lyapounov equation and ii) a robust tracker, based on an outlier feature model. Experimental results are provided, showing that the proposed algorithm is computationally efficient and leads to accurate estimates of the LV.

1 Introduction

The left ventricle (LV) boundary estimation plays an important role in clinical diagnosis since it allows to extract relevant measures of the heart dynamic behavior, among which the ejection fraction and local wall motion.

Ultrasound imaging is a popular technique to observe the dynamical behavior of the heart. However, the low signal-to-noise ratio (SNR) and the multiplicative nature of the noise (speckle) corrupting the ultrasound images, make the LV segmentation a difficult task.

The major edge detection algorithms fail due to the presence of multiplicative noise in heart ultrasound imagery. The strongest edges are often not located on the endocardium. In [1] it is proposed the instantaneous coefficient of variation (ICOV) providing good segmentation results, but the so called problem of "edge dropout" still remains (this is typical in the diastole phase). Therefore, noise reduction must be applied before edge detection. Several techniques have been proposed to reduce the speckle noise without distorting the relevant clinical details, e.g., Bayesian methods [2], mixture distribution of the Rician pdf with the inverse Gaussian as a mixture distribution (RiIG) [3], soft thresholding [4], wavelet based methods [5], wavelet soft-shrinking [6], median filtering [7], and anisotropic diffusion [8].

Even though the denoising algorithms significantly reduce the speckle noise, advanced tracking techniques are needed to segment the LV boundary. Prior art in segmentation of echocardiographic sequences of the heart includes active shape models [9], or level set techniques [10].

* This work was supported by Fundacao para a Ciência e a Tecnologia (ISR/IST pluriannual funding) through the POS Conhecimento Program which includes FEDER funds.

In this paper we join a novel edge preserving total variation (TV) based denoising algorithm and a robust tracker [11]. The denoising algorithm must process a large number of ultrasound images in an efficient way. This is obtained by formulating the filtering operation as the solution of a Sylvester/Lyapunov equation for which there are fast and computationally efficient algorithms described in the literature.

The robustness of the tracker is obtained by using feature grouping (line segments), which are labeled as valid or invalid. Since the labels are unknown they are replaced by their probabilities computed using a probabilistic model of the observations. A data association filter is then used to update the contour parameters under the presence of outliers.

The tracking algorithm proposed in the paper was assessed using a set of image sequences, segmented by medical doctors. These images, are used as a ground truth to compute FOM (figures of merit).

The paper is organized as follows: Section 2 describes the overall system. Sections 3, 4 and 5 describe the pre-processing, feature detection and tracking steps respectively. Section 6 describes experimental and section 7 concludes the paper.

2 System Overview

The proposed system aims is to track the boundary of the left ventricle during the cardiac cycle. The system input is a sequence of ultrasound images sampled at 25Hz.

The system performs three main operations: i) *denoising* : to reduce the speckle noise and enhance the contrast; due to the large amount of data to be processed a novel algorithm was developed to perform this task, ii) *feature detection* : detects intensity transitions along orthogonal lines radiating from the contour. Transitions are obtained by applying a matched filter to the intensity profiles and computing the local maxima [12], and iii) *tracking* : based on a robust tracking algorithm which fits a deformable curve (quadratic B-spline) to the points detected in the image. This algorithm must be able to deal with a large number of outliers and to interpolate the boundary when no features are detected due to low contrast of the heart boundary. This is specially important close to the apex and in the presence of sudden motion changes (e.g., in the mitral valve). A recent tracking algorithm is used in this step.

3 Pre-Processing

The performance of the tracker depends on the SNR of the input images which have multiplicative noise.

The goal of the pre-processing step is to reduce the noise without losing relevant information. In this paper a MAP criterion is used to estimate the original images from the noisy ones. This approach is usually slow and computationally demanding, furthermore, there is a large number of images to process.

A Bayesian framework is used with the MAP criterion, and the optimization algebraic problem is formulated as Sylvester/Lyapunov equation for which there are fast em computationally efficient algorithms described in the literature.

The denoising algorithm estimates the original image, X, by minimizing the following energy function $E(Y,X) = -\log[p(Y|X)p(X)]$ where Y is the noisy image, $p(Y|X)$ is the observation model and $p(X)$ is the prior distribution of the unknown image.

Assuming conditional independence of the observations, leads to $p(Y|X) = \prod_{ij}^{N,M} p[y(ij)|x(ij)]$ where $p(y|x) = \frac{y}{x}e^{-y^2/2x}$ is the Rayleigh distribution [13].

An edge preserving prior $p(X)$ was chosen to avoid over-smoothing the transitions. The prior is based on the total variation (TV) function as $p(X) = \frac{1}{Z}e^{-\alpha\sum_{i,j}g(i,j)}$ where $g(i,j) = |\nabla X(i,j)|$ is the gradient magnitude of X at the (i,j) pixel, α is a parameter and Z is a partition function. This gradient magnitude may be approximated by using the first order differences, $g(i,j) = \sqrt{\delta^2_{v_{i,j}} + \delta^2_{h_{i,j}}}$ where $\delta_{v_{i,j}} = x(i,j) - x(i,j-1)$ and $\delta_{h_{i,j}} = x(i,j) - x(i-1,j)$.

The denoised image is obtained by solving the following equation

$$\hat{X} = \arg\min_X E(Y,X) \tag{1}$$

where

$$E(Y,X) = \sum_{i,j}\left[\log\left(\frac{y(i,j)}{x(i,j)}\right) - \frac{y^2(i,j)}{2x(i,j)}\right] + \alpha\sum_{i,j}g(i,j) \tag{2}$$

To find out the minimizer of (2), its stationary points must be computed, i.e., $\nabla E(Y,X) = 0$, which is equivalent to

$$\frac{x(i,j) - x^{ML}(i,j)}{x^2(i,j)} + \frac{\partial}{\partial x(i,j)}\sum_{i,j}g(i,j) = 0, \qquad 0 \le i,j \le N-1, M-1 \tag{3}$$

where $x^{ML}(i,j) = y^2(i,j)/2$ is the *maximum likelihood* (ML) estimate for the Rayleigh distribution. The set of equations (3) is non-linear on X and it is iteratively solved. The fixed point method and the majorize/minimize (MM) algorithm described in [14] leads too the following recursion equation,

$$\frac{x(i,j) - x^{ML}(i,j)}{x^2_{t-1}(i,j)} + \frac{\partial}{\partial x(i,j)}\sum_{i,j}\frac{\delta^2_{v_{i,j}} + \delta^2_{h_{i,j}}}{w_{t-1}(i,j)} = 0, 0 \le i,j \le N-1, M-1 \tag{4}$$

where $x_{t-1}(i,j)$ and $w_{t-1}(i,j) = 1/g_{t-1}(i,j)$ are the image and gradient magnitude reciprocals, respectively, computed in the $(t-1)$-th iteration.

The set of equations (4) can be written in the following matrix notation as shown in [15]

$$X_{t-1} \circledast (X - X^{ML}) + 2\alpha G^{-1}_{t-1} \circledast [\phi_v X + X\phi_h] = 0 \tag{5}$$

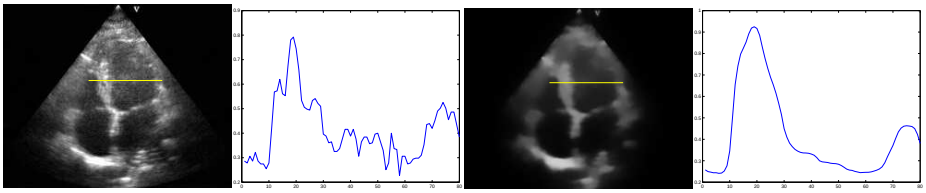
where $\phi_v = \theta_v^T\theta_v$, $\phi_h = \theta_h^T\theta_h$ and $G^{-1}_{t-1}(i,j) = 1/|\nabla X_{t-1}(i,j)|$ is the matrix whose elements are the reciprocals of the gradient magnitudes of X_{t-1}. The operator $\circledast$ stands for Hadamard product, i.e., element wise product. θ_v and

θ_h are $n \times n$ vertical and $m \times m$ horizontal difference operators respectively. Therefore, equation (5) can be rewritten as follows

$$\Phi_v X + X\Phi_h + Q_{t-1} = 0 \tag{6}$$

where $\Phi_v = \beta I_N/2 + 2\alpha\phi_v$, $\Phi_h = \beta I_M/2 + 2\alpha\phi_h$, $Q_{t-1} = W_{t-1} \circledast (X_{t-1} - X^{ML}) - \beta X_{t-1}$ and $W_{t-1} = G_{t-1} \circledast / [X_{t-1} \circledast X_{t-1}]$. I_N and I_M are N and M dimensional identity matrices respectively and β is a conditioner parameter to improve the stability of the algorithm (typically $\beta = 1$).

The equation (6) is the so called Sylvester equation for which there are efficient and fast solver algorithms. Fig.1 shows an example of denoised ultrasound



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

Fig. 1. (a) Real image, (b) denoised image using the proposed technique, (c) image profile of (a), image profile of (b).

image using the pre-processing described above.

4 Feature Detection

Feature detection detects line segments belonging to the boundary of the LV. This is done in two steps. First we detect intensity transitions along lines orthogonal to the predicted contour. This is done by template matching. Feature detection along the ith direction is performed by computing the local maxima of the function

$$\mathcal{J}(t_0) = \int_t |p_i(t) - T(t, t_0)|^2 dt \tag{7}$$

where $p_i(t)$ is the image profile taken at the ith direction, t denotes the distance to the object boundary and $T(t, t_0)$ is a template which is obtained off-line. The template T is obtained as follows: $T(t)$ is equal to the typical intensity of the object for $t \leq t_0$ and $T(t)$ is equal to the background intensity for $t > t_0$. In the second step, feature points detected at consecutive lines are grouped, by mutual by mutual favorite pairing, forming image line-segments.

5 Tracking

A deformable curve (B-spline) is used to approximate the LV contour. The parameters of the B-spline at time k, $x_k \in \mathbb{R}$ are estimated from the image features obtained in the previous step using a tracking algorithm.

This is not an easy task since there are many invalid features detected in the ultrasound image and the tracker must able to ignore them and to track valid features only. The Kalman filter fails in this problem since it is not able to separate valid features from invalid ones.

In this paper we have used a data association filter which was recently proposed in [11]. This method considers all the hypothesis of valid/invalid features, H_i, and assigns a probability to each of them (see [11] for the details).

To avoid an exponential growth of hypothesis at different time instants, a simplifying assumption is adopted: it is assumed that the state distribution given past observations is Gaussian, i.e.,

$$p[x_k \mid Y^{k-1}] = \mathcal{N}[x_k; \hat{x}_{k|k-1}, P_{k|k-1}] \tag{8}$$

where $\hat{x}_{k|k-1}$, $P_{k|k-1}$ are the mean and covariance of x_k given past observations Y^{k-1}. This hypothesis was proposed by Bar-Shalom in the context of target tracking [16].

The computation of the state estimate (state mean) given current and past observations is done considering all the hypothesis

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \sum_{i=1}^{m_k} \alpha_{i\,k} K_{i\,k} \nu_{i\,k} \tag{9}$$

This resembles the Kalman filter. In (9) $\hat{x}_{k|k}$ is the estimate of the state vector, $K_{i\,k}$, $\nu_{i\,k}$ are the Kalman gain and the innovation respectively, and $\alpha_{i\,k} \triangleq p(H_{i\,k} \mid Y^k)$ is the *a posteriori* probability of the i-th hypothesis $H_{i\,k}$. The interpretation of equation (9) suggests that we have a bank of Kalman filters each one specialized to each ith data hypothesis.

A recursive equation can also be derived for the covariance matrix (see details in [11])

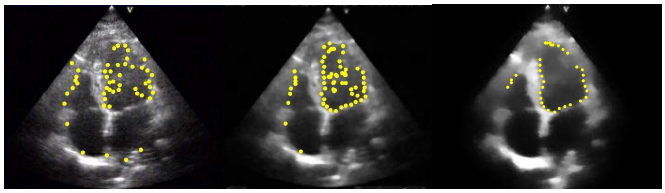
6 Experimental Results

This section shows experimental results obtained with the proposed method. A echocardiographic sequence of the left ventricle is used in this study. The length of the sequence has 490 frames comprising 27 cardiac cycles and each image has 320×240 pixels.

The experiments involve three main steps: *i*) the LV boundary is manually defined by an expert (ground-truth) in several images; *ii*) the sequence is automatically processed by the tracker; *iii*) metrics between automatic and manual boundaries are computed for the sequence. The tests are performed under three options: *i*) without, *ii*) median, *iii*) and Lyapounov pre-processing.

6.1 Ground Truth

To obtain the ground truth, an observer provides a hand-labeled contours for the sequence. Four images in each cardiac cycle are selected for hand labeling: two images in the systole phase and two images in the diastole phase. A total number of 108 contours were manually generated (54 in each phase). The tracker-generated boundaries are compared to the ground truth resulting in an error measurement in each image.



(a) (b) (c)

Fig. 2. Features (dots) detected in three situations: without pre-processing (a), median filtering (b), proposed algorithm (c).

6.2 Error Metrics

Three error metrics are used to compare the tracker-generated boundaries against the boundaries outlined by the observer.

The two curves are represented as sets of points $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_x}\}$, and $\mathcal{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_y}\}$, where $N_y > N_x$. Each $\mathbf{x}_i$ and $\mathbf{y}_i$ is a pair of coordinates of the point in the image plane.

The distance from a point $\mathbf{x}_i$ to the curve $\mathcal{Y}$ is

$$d(\mathbf{x}_i, \mathcal{Y}) = \min_j ||\mathbf{y}_j - \mathbf{x}_i|| \tag{10}$$

The average distance from the contour model $\mathcal{X}$ to the ground truth boundary $\mathcal{Y}$ (ideal contour) is

$$d_{av} = \frac{1}{N_x} \sum_{i=1}^{N_x} d(\mathbf{x}_i, \mathcal{Y}) \tag{11}$$

The Hausdorff distance between the two curves is defined as the maximum distance from a point to the other curve

$$d_{max}(\mathcal{X}, \mathcal{Y}) = \max\Big(\max_i\{d(\mathbf{x}_i, \mathcal{Y})\}, \max_j\{d(\mathbf{y}_j, \mathcal{X})\}\Big) \tag{12}$$

The third metric is the Hammoude [17] measure proposed in the context of ultrasound images and given by

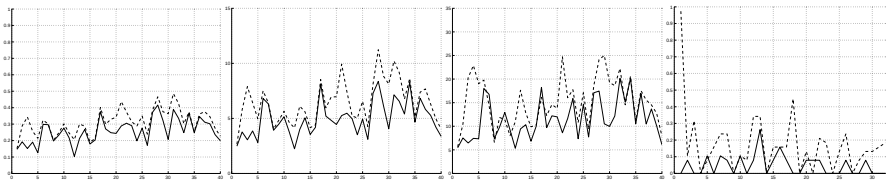
$$d_H = \frac{\sharp\big((X \cup Y) - (X \cap Y)\big)}{\sharp(X \cup Y)} \tag{13}$$

where X, Y are binary images such that all pixels inside the curves have label 1 and remaining pixels have label 0. This metric computes the normalized number of pixels which receive different labels.

Fig. 3 shows the evolution of the metrics for the sequence. The first measure (Fig. 3 (a)) belongs to the interval $[0, 1]$, the remaining ones are expressed in terms of pixels. The dashed line refers to the results obtained by using a median filter in the pre-processing step. The solid line represents the values obtained

by the proposed method, (we do not show the results obtained without pre-processing since they are much worse). Fig. 3 shows the results at specific frames. These frames correspond to the time instants when the cardiac phase switch from systole to the diastole and vice-versa. This figure also shows the situation when the tracker has difficulties to represent the *apex* of the ventricle.

Fig. 3 shows that the denoising technique proposed herein has a much better performance compared with the median filter (the solid line is under the dashed line). The first and second order statistics of the contour metrics are shown in Table 1. Here, it is shown the average computation time associated to the tracker in seconds. We conclude that the proposed pre-processing method is twofold: i) the mean and variance error of the shape estimates is smaller than in the other cases; ii) it allows a faster tracking since less outlier features are detected in the image. See Fig. 2 where it is clearly shown that the number of outliers decrease from (a) (without pre-processing) through (c) (proposed algorithm), the latter preserving the contour.



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

Fig. 3. Metric statistics for the heart sequence, (a) d_H, (b) d_{av}, (c) d_{max}, (d) number of outliers without pre-processing and with the proposed technique. Median filtering (dashed line), proposed method (solid line).

Table 1. Mean and variance values for the metrics shown in the Fig. 3 and the computation time (average per frame) for the three different cases.

	Hammoude metric-d_H			Average Distance-d_{av}			Hausdorff Distance-d_{max}		
	without	Median	Denoising	without	Median	Denoising	without	Median	Denoising
E[.]	0.29	0.28	**0.26**	6.1	5.7	**5.2**	14.2	13.7	**11.7**
var[.]	0.01	0.01	**0.008**	5.5	4.6	**4.0**	38.0	26.5	**22.9**

	without	Median	Denoising
Time_{av}	0.87	0.86	**0.76**

7 Conclusions

This paper proposes a system for tracking the left ventricle using two key operations. The first is a novel denoising algorithm based on the Lyapounov equation. The second is a robust tracker used to estimate the evolution of the LV contour. The robustness is achieved by using data-association within the detected line-segments.

It is concluded from the experimental results that the proposed algorithm manages to accurately track the heart motion in images with a low contrast between the heart cavity and the miocardium. It is also concluded that the denoising algorithm plays an important role and significantly reduces the number of outliers.

References

1. Y. Yu and S. T. Acton, "Edge detection in ultrasound imagery using the instantaneous coefficient of variation," *IEEE Trans. Med. Imag.*, vol. 13, no. 12, pp. 1640–1655, 2004.
2. Z. Zeng and I. Cumming, "Bayesian speckle noise reduction using the discrete wavelet transform," in *Int. geosc. and rem. sens. symp.*, pp. 6–10, 1998.
3. T. Eltoft, "Modeling the amplitude statistics of ultrasonic images," *IEEE Trans. Med. Imag.*, vol. 25, no. 2, pp. 229–240, 2006.
4. S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, "A wavelet based statistical approach for speckle reduction in medical ultrasound images," in *Med. and Biol. Eng. and computing*, vol. 42, pp. 189–192, 2004.
5. I. Duskunovic, A. Pizurica, G. Stippel, W. Philips, and I. Lemahieu, "Wavelet based denoising techniques for ultrasound images," in *Proc. of the IEEE Eng. in Medicine and Biol. Soc. Conference*, vol. 4, pp. 2662–2665, 2000.
6. Y. Yue, M. M. Croitoru, A. Bidani, J. B. Zwischenberger, and J. W. Clark, "Ultrasonic speckle suppression using robust nonlinear wavelet diffusion for LV volume quantification," in *Proc. of the Int. Conf. of the IEEE EMBS*, pp. 1609–1612, 2004.
7. T. Loupas, W. McDicken, and P. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 129–135, 1989.
8. J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache, "Anisotropic filtering for model-based segmentation of 4d cylindrical echocardiographic images," *Pattern Recognition Letters - Special Issue on Ultrasonic Imag. Proc. and Anal.*, vol. 24, pp. 815–828, 2003.
9. X. H. amd N. Paragios and D. Metaxas, "Establishing local correspondences towards compact representation of anatomical structures," *Med. Imag. Comp. and Computer-Assisted Intervention*, 2003.
10. S. Osher and N. Paragios, *Geometric Level Set Methods in Imag., Vision and Graphics*. Springer Verlag, 2003.
11. J. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background," *IEEE Trans. Multimedia*, vol. 6, no. 6, pp. 852–861, 2004.
12. A. Blake and M. Isard, *Active Contours*. Springer, 1998.
13. C. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Trans. on Sonics and Ultrsonics*, vol. SU-25, no. 1, pp. 1–6, 1978.
14. J. M. B. Dias, "Fast GEM wavelet-based image deconvolution algorithm," *IEEE Int. Conf. on Image Proc.*, pp. 961–964, 2003.
15. J. Sanches and J. S. Marques, "Image denoising using the lyapunov equation from non-uniform samples," in *ICIAR*, 2006.
16. Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. Academic Press, 1988.
17. A. Hammoude, "Computer-assited endocardial border identification from a sequence of two-dimensional echocardiographic images," Ph.D. dissertation, Univ. Washington, Seatle, 1988.