Fast Tracking of the Left Ventricle Using Global Anatomical Affine Optical Flow and Local Recursive Block Matching

Daniel Barbosa[1], Denis Friboulet[2], Jan D'hooge[3], and Olivier Bernard[2]

[1] Life and Health Sciences Research Institute (ICVS), University of Minho, Portugal
[2] Lab on Cardiovascular Imaging & Dynamics, Department of Cardiovascular Sciences, KU Leuven, Belgium
[3] Université de Lyon, CREATIS; CNRS UMR5220; INSERM U630; Université Lyon 1; INSA-LYON, France

Abstract. We present a novel method for segmentation and tracking of the left ventricle (LV) in 4D ultrasound sequences using a combination of automatic segmentation at the end-diastolic frame and tracking using both a global optical flow-based tracker and local block matching. The core novelty of the proposed algorithm relies on the recursive formulation of the block-matching problem, which introduces temporal consistency on the patterns being tracked. The proposed method offers a competitive solution, with average segmentation errors of 2.29 and 2.26mm in the training (#=15) and testing (#=15) datasets respectively.

1 Introduction

While magnetic resonance imaging remains the gold standard for cardiac morphology and function assessment, several studies have shown that real-time 3D echocardiography (RT3DE) is a competitive modality for this clinical task [1]. Indeed, RT3DE offers superior performance when compared to conventional 2D ultrasound imaging on the visualization of the entire left ventricle (LV), thus avoiding several pitfalls of 2D echocardiography such as foreshortening, out-of-plane motion and the need of geometric assumptions for volume estimation [2]. On the other hand, and due to the intrinsic physical limits of acoustical wave propagation, 3D ultrasound imaging requires advanced beam-forming strategies to sweep the entire scan volume with a number of ultrasound pulses compatible with real-time imaging. As a consequence, image quality may be impaired as compared to conventional 2D echo. Simultaneously, the increased dimensionality of the data poses some challenges on the data analysis pipeline, which has triggered a significant effort from both industrial and academic research teams on the development of automated software packages for LV volumetric assessment [3, 4]. However, even state-of-the-art commercial solutions still require some degree of user interaction both at the initialization step and for a correction step of the segmentation/tracking results [5]. Thus, tools allowing automatic fast 3D LV segmentation are still needed [6].

Our prior approaches to LV tracking focused on the problem from a global perspective, since we have developed an algorithm modeling the LV motion during the cardiac cycle as an affine transformation. Despite its interesting performance for the quantification of global functional indices [7], there is still room for improved tracking performance using local refinement of the globally deformed LV surface. We have initially used a hybrid method relying on the combination of the global tracking-based algorithm with a local refinement based on segmentation-oriented clues. We have shown that this positively contributes to the tracking performance [8], but it remains limited to the assumption that the boundary position matches the optimal value of the associated data attachment term. Since often the physicians do not delineate the LV boundaries at maximum local contrast positions, the development of accurate segmentation energies is challenging. Thus, in the present paper we follow the global tracking plus local refinement strategy but we drive the local tracking using a block matching approach rather than a segmentation-based term.

2 Methods

2.1 Automatic LV Segmentation at End-Diastole

B-Spline Explicit Active Surfaces (BEAS) is a real-time segmentation framework recently introduced [9]. The fundamental concept of the BEAS framework is to regard the boundary of an object as an explicit function, where one of the coordinates of the points within the surface, x_1, is given explicitly as a function of the remaining coordinates, i.e. $x_1 = \psi(x_2, \cdots, x_n) = \psi(\mathbf{x}^*)$. Following [9], ψ was defined as a linear combination of B-spline basis functions, where the segmented surface is explicitly controlled through the B-spline coefficients $c[\mathbf{k}]$, where $\mathbf{k}$ defined the position of the B-spline kernel spanned on a grid.

In the present work, we use a modified version of the localized means separation energy which takes advantage of the darker appearance of the blood with respect to the myocardial tissue [10]. While this approach evolves the contours towards the positions of maximum local contrast, the expert physicians usually prefer to delineate the LV surface closer to the blood-tissue interface. Thus, we have previously introduced a hyper-parameter that allows controlling the balance between the forces exerted by the inner and the outer regions [11]. This allows to globally steer the LV surface position inwards or outwards, in order to better match the manual contouring protocol. Thus, in the present work the B-spline coefficients $c[\mathbf{k}]$ are updated using the following expression:

$$\frac{\partial E_L}{\partial c[\mathbf{k_i}]} = \int_\Gamma \left(\alpha_{in} \left(\bar{I}(\mathbf{x}^*) - u_x \right) + \left(\bar{I}(\mathbf{x}^*) - v_x \right) \right) \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k}_i \right) \, d\mathbf{x}^*, \quad (1)$$

where u_x and v_x are the local means estimated inside and outside a local neighborhood B around each point on the LV surface, respectively. For clarity sake, $\bar{I}(\mathbf{x}^*)$ corresponds to the image value at the position $\mathbf{x} = \{\psi(\mathbf{x}^*), x_2, \cdots, x_n\}$. α_{in} controls the balance between the forces exerted by the inner and outer regions. A value of $\alpha_{in} > 1$ will increase the influence of the inner regions, thus

attracting the segmented LV towards the blood-tissue interface. For the present work, ψ is defined in the spherical space, *i.e.* $\rho = \psi(\theta, \varphi)$. The initialization of the segmentation algorithm was obtained with the method introduced in [10], which provides an ellipsoid approximation to the LV endocardial surface.

2.2 Combination of Global Affine Optical Flow and Recursive Block Matching For Efficient LV Tracking

We propose a two-step tracking strategy. First, we estimate the global deformation between two subsequent frames using the anatomical affine optical flow introduced in [7]. This allows a robust tracking over the cardiac cycle, while also keeping a consistent shape since only affine transformations are allowed. The novelty of the proposed solution lies in the subsequent local refinement based on a block matching approach. Since the larger displacements were already accounted for in the global tracking stage, the block matching can be restricted to a smaller search region.

The traditional block-matching approaches perform an exhaustive search of a given image kernel in the subsequent frame. The position optimizing the chosen similarity criterion (such as sum of absolute differences, normalized cross correlation, among others) is taken as the most likely new position and the inter-frame displacement is computed accordingly. However, errors can cumulate over the entire cardiac cycle and indeed it can be shown that there is an intrinsic optimal relationship between the frame rate and the displacement to be recovered [12]. One of the key drivers of error accumulation is the fact that the tracking problem is posed as a sequential block matching problem, thus not enforcing temporal continuity between multiple frames. Indeed, the kernel being tracked at frame t might be completely distinct from the kernel found in $t - 2$. In the present paper, we introduce the concept of temporally recursive block matching by considering a dynamic version of the block to be searched in the next image. Indeed, instead of considering that the kernel is simply a block from the frame t, we use a dynamic block which is recursively built not only with the image patterns from frame t but also from the previous frames.

Traditional block matching approaches estimate motion by tracking a 3-D kernel between subsequent frames using a sliding window technique and using a given similarity measure to estimate the optimal match. Considering the sum of squared differences, the motion for a given point would be estimated as the minimum over (u, v, w) of:

$$SSD(u,v,w) = \sum_{\mathbf{x} \in \Omega} \left(I(x_1, x_2, x_3, t) - I(x_1 + u, x_2 + v, x_3 + w, t + 1) \right)^2, \quad (2)$$

where $\mathbf{x} \in \Omega$ corresponds to the local region of interest (ROI) being tracked. Typically the search region is limited by constraining the values of the displacement vector (u, v, w) where the SSD is estimated. In this case the optimal displacement (u_t, v_t, w_t) is taken as the minimum of $SSD(u, v, w)$.

In the proposed approach, instead of taking the local region of interest from the previous frame, we recursively combine the kernel of the current frame being tracked with the previously tracked ROI's through:

$$RSSD(u, v, w) = \sum_{\mathbf{x} \in \Omega} \left(\tilde{I}(x_1, x_2, x_3, t) - I(x_1 + u, x_2 + v, x_3 + w, t + 1) \right)^2, \quad (3)$$

where

$$\tilde{I}(x_1, x_2, x_3, t) = 0.5 \cdot \tilde{I}(x_1 - u_{t-1}, x_2 - v_{t-1}, x_3 - w_{t-1}, t-1) + 0.5 \cdot I(x_1, x_2, x_3, t). \quad (4)$$

Using the proposed approach, if the tracking fails at a given frame, it can still recover from the error since the temporal consistency introduced through the recursive estimation of the ROI will steer the block matching in the next frame towards the patterns previously being tracked.

3   Implementation Details

Prior to processing, each 4D ultrasound sequence was re-sampled to guarantee isotropic voxel size, which was set to the smallest voxel dimension of the original dataset. Regarding the underlying segmentation framework, all the parameters were set as originally reported in [9] and [10], with two notable exceptions. First, the size of the sliding plane used to detect the LV base position was set to $rmin$ (*i.e.* 15 mm) rather than $rmax$ (*i.e.* 35 mm), contrarily to what was proposed in [10]. Secondly, and since there was significant differences in the voxel size, the radius of the local neighborhood was set to 16 mm rather than 16 voxels as previously reported in [10]. The parameter controlling the balance between the forces of the inner and outer regions, *i.e.* α_{in}, was empirically set to 1.25. The anatomical optical flow parameters were set as in [7], whereas the sizes of the ROI and search kernels in the block matching stage were set to[7x7x7] and [11x11x11], as proposed in [13]. No sub-pixel accuracy strategies were used for the block-matching stage.

4   Experiments and Results

The proposed pipeline was tested in the database from the CETUS challenge, which is composed by 4D ultrasound sequences from healthy volunteers and patients with myocardial infarction and dilative cardiomyopathy. This data has been acquired with multiple imaging platforms from different vendors, namely a Siemens SC 2000, a Philips i33 and a GE E9. This database is divided into two subsets of 15 exams for training and testing. While for the training dataset the meshes for the reference segmentation are provided, this information is not available for the testing dataset. The performance of the algorithm was assessed using distance metrics to the reference segmentation result, namely mean absolute distance (MAD), Hausdorff distance (HD) and modified Dice coefficient

Table 1. Segmentation performance of the proposed algorithm in the CETUS training dataset, starting from the ground truth surface at ED and evaluating the LV surfaces at end-systole (MAD: mean absolute distance; HD: Hausdorff distance; Dice*: modified Dice coefficient (*i.e.* 1-Dice); R: Pearson correlation coefficient; LOA: Limits of agreement, estimated as $\mu \pm 1.96\sigma$). MAD and HD values given in mm.

Algorithm	MAD	HD	Dice*	SV (ml)		EF (%)	
				R	LOA	R	LOA
Proposed	1.19±0.3	4.38±1.1	0.063±0.03	0.951	3.42±15.0	0.958	0.70±8.6
GAOF+BM	1.20±0.3	4.29±1.0	0.061±0.02	0.944	4.04±16.2	0.953	1.12±9.1
GAOF [7]	1.18±0.3	4.53±1.0	0.063±0.02	0.953	4.48±14.9	0.959	1.53±8.6

Table 2. Segmentation performance of the proposed algorithm in the CETUS database (MAD: mean absolute distance; HD: Hausdorff distance; Dice*: modified Dice coefficient (*i.e.* 1-Dice)). All values given as $\mu \pm \sigma$, MAD and HD values given in mm.

Algorithm	Training Dataset			Testing Dataset		
	MAD	HD	Dice*	MAD	HD	Dice*
Proposed	2.29±0.6	8.20±2.4	0.102±0.03	2.26±0.8	8.45±2.5	0.127±0.05
HT [8]	2.47±0.7	8.61±2.4	0.107±0.04	2.36±0.9	8.99±2.8	0.131±0.06
GAOF [7]	2.29±0.6	8.21±2.4	0.102±0.05	2.25±0.7	8.45±2.5	0.127±0.05

(Dice*, estimated as 1-Dice). These metrics were computed directly through the challenge MIDAS platform, thus guaranteeing a common evaluation platform for all the participants in the challenge. The accuracy of the estimation of LV volumetric indices used in clinical routine, namely end-diastolic volumes (EDV), end-systolic volumes (ESV), stroke volumes (SV) and ejection fraction (EF), was equally included. The reported CPU timings refer to a MATLAB-based implementation running on a standard Windows 7 laptop equipped with a dual-core i7-640m processor and 4 GB of RAM.

We have carried an initial experiment using the training dataset in order to evaluate the added value of the proposed dynamic ROI introduced in (3) versus the classical pairwise block matching for local refinement (GAOF+BM), as well as against the global affine anatomical optical flow baseline tracker (GAOF). The results of this initial experiment are presented in Table 1.

The summary of the results for the distance segmentation metrics is presented in Table 2, while the accuracy of the estimation of LV volumetric indices can be found in Tables 3 and 4. Fig. 1 illustrates the local surface errors on a representative sub-set of 3D ultrasound frames. We have equally compared the performance of the proposed method with the baseline global affine optical flow tracker and with the hybrid framework (HT) introduced in [8], where the local refinement is done using segmentation-based clues, as opposed to local tracking-based hints in the proposed algorithm. An example of a 4D RT3DE segmented with the proposed approach is illustrated in Fig. 2

Table 3. Assessment of LV volumetric indices in the training database (R: Pearson correlation coefficient; LOA: Limits of agreement, estimated as $\mu \pm 1.96\sigma$).

Algorithm	EDV (ml)		ESV (ml)		SV (ml)		EF (%)	
	R	LOA	R	LOA	R	LOA	R	LOA
Proposed	0.99	13.8±36.2	0.99	9.4±34.8	0.92	4.3±18.8	0.92	1.1±11.4
HT [8]	0.99	13.8±36.2	0.98	11.7±40.9	0.91	2.0±20.5	0.90	0.3±13.3
GAOF [7]	0.99	13.8±36.2	0.99	9.6±34.5	0.93	4.1±18.2	0.92	1.0±10.9

Table 4. Assessment of LV volumetric indices in the testing database (R: Pearson correlation coefficient; LOA: Limits of agreement, estimated as $\mu \pm 1.96\sigma$).

Algorithm	EDV (ml)		ESV (ml)		SV (ml)		EF (%)	
	R	LOA	R	LOA	R	LOA	R	LOA
Proposed	0.97	-11.2±22.7	0.99	-13.3±13.3	0.50	2.1±20.0	0.93	4.7±7.8
HT [8]	0.97	-11.2±22.7	0.97	-9.6±19.5	0.71	-1.6±18.4	0.84	3.0±11.2
GAOF [7]	0.97	-11.2±22.7	0.99	-13.0±13.0	0.51	1.9±19.8	0.93	4.5±7.7

5   Discussion

The results of the initial experiment put an evidence on the trade-off between local accuracy and global tracking performance. Indeed, while starting from the reference LV surfaces, the global affine optical flow algorithm introduced in [7] presents the lower MAD, the local refinement using block matching improves the local tracking, as supported by the lower HD values. Regarding the added value of the proposed algorithm, the memory effect of the ROI being locally tracked improves the estimation accuracy of the functional LV volumetric indices (SV and EF), reducing the bias and the limits of agreements of the estimated values when compared to the traditional block matching refinement.

For the complete pipeline (ED segmentation + LV tracking over the entire cycle), the method in [8] presents lower tracking performance than the one introduced in the present manuscript, whose local refinement strategy is primarily driven by tracking-based clues. While this observation does not imply that hybrid strategies are not suited to LV tracking problems, it stresses the need to develop more accurate segmentation energies adapted to the complexity of the local appearance of RT3DE images around the endocardial surface, since positions of higher contrast do not always correspond to the true endocardium.

The global tracker [7] remains very competitive when compared with the proposed approach, thus not being trivial to identify the best tracking strategy. Indeed, while the LV motion using the proposed approach appears more realistic, since it presents more complex deformation patterns with local variations, the global tracker remains more insensitive to errors. Since the analyzed dataset only assesses the LV position at ED and ES frames, a more complete 4D analysis would likely be necessary to put in evidence the differences between the global tracking approach and the locally refined LV surface using the proposed method.
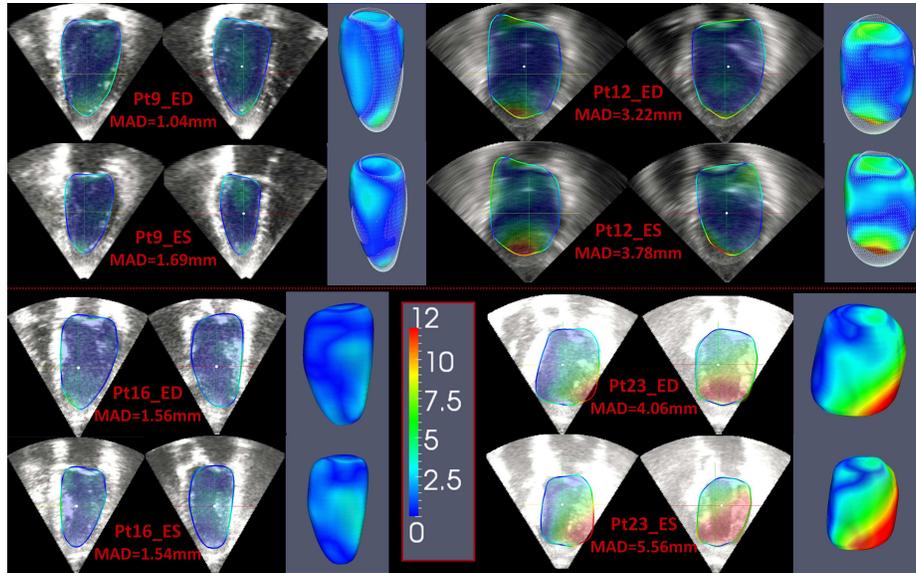
Fig. 1. Best (left) and worse (right) segmentation results for the training (up) and testing (down) datasets. The local segmentation error is color encoded, while the reference LV surface is shown in white for the training dataset cases.

6 Conclusion

The proposed LV segmentation and tracking framework offers competitive performance for the fully automatic detection of LV endocardial borders in 4D ultrasound sequences. This is supported by the low segmentation errors and accurate volume estimates found in a cohort of 30 exams with a wide range of image quality and cardiac functional status. Furthermore, the computational burden of the proposed method remains low, allowing global affine tracking plus local refinement to be done in approximately 2s for two consecutive frames. From our prior experience, this indicates that near-real time performance could be achieved by using a more efficient coding of the algorithm in, for instance, C++.

References

1. Dorosz, J., Lezotte, D., Weitzenkamp, D., Allen, L., Salcedo, E.: Performance of 3-dimensional echocardiography in measuring left ventricular volumes and ejection fraction. J Am Coll Cardiol **59**(20) (2012) 1799–1808
2. Mor-Avi, V., Jenkins, C., Kuhl, H., Nesser, H., Marwick, T., Franke, A., Ebner, C., Freed, B., Steringer-Mascherbauer, R., H.Pollard, Weinert, L., Niel, J., Sugeng, L., Lang, R.: Real-time 3-dimensional echocardiographic quantification of left ventricular volumes: Multicenter study for validation with magnetic resonance imaging and investigation of sources of error. JACC: Cardiovasc Im **1**(4) (2008) 413 – 423
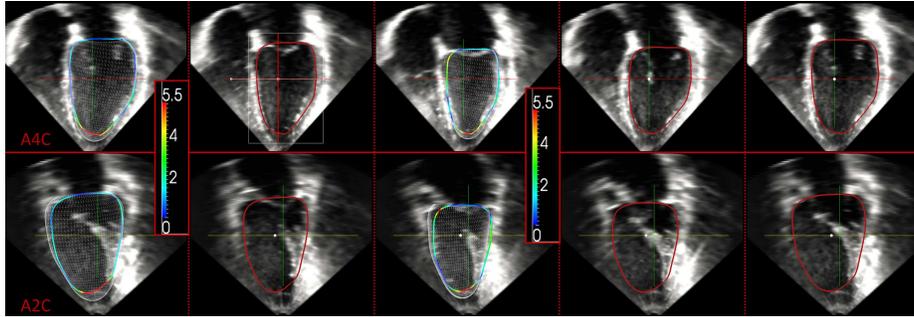
Fig. 2. LV tracking using the proposed approach (red/colormap: segmented/tracked LV surface; white: reference LV surface at end-diastolic and end-systolic instants).

3. Yang, L., Georgescu, B., Zheng, Y., Wang, Y., Meer, P., Comaniciu, D.: Prediction based collaborative trackers (PCT): A robust and accurate approach toward 3D medical object tracking. IEEE T Med Imaging **30**(11) (2011) 1921 –1932

4. Leung, E., Danilouchkine, M., van Stralen, M., de Jong, N., van der Steen, A., Bosch, J.: Left ventricular border tracking using cardiac motion models and optical flow. Ultrasound Med Biol **37**(4) (2011) 605 – 616

5. Muraru, D., Badano, L., Piccoli, G., Gianfagna, P., Del Mestre, L., Ermacora, D., Proclemer, A.: Validation of a novel automated border-detection algorithm for rapid and accurate quantitation of left ventricular volumes based on three-dimensional echocardiography. Eur J Echocardiogr **11**(4) (2010) 359–368

6. Leung, K., Bosch, J.: Automated border detection in three-dimensional echocardiography: principles and promises. Eur J Echocardiogr **11**(2) (2010) 97–108

7. Barbosa, D., Heyde, B., Dietenbeck, T., Friboulet, D., D'hooge, J., Bernard, O.: Fast left ventricle tracking in 3D echocardiographic data using anatomical affine optical flow. In: Proc. of FIMH2013. (2013) 191–199

8. Barbosa, D., Bernard, O., Heyde, B., Dietenbeck, T., Friboulet, D., Dhooge, J.: Hybrid energy approach for real-time b-spline explicit active tracking of surfaces (heartbeats). In: Proc. of IUS2013. (2013) 663 – 666

9. Barbosa, D., Dietenbeck, T., Schaerer, J., D'hooge, J., Friboulet, D., Bernard, O.: B-spline explicit active surfaces: An efficient framework for real-time 3D region-based segmentation. IEEE T Image Process **21**(1) (2012) 241 – 251

10. Barbosa, D., Dietenbeck, T., Heyde, B., Houle, H., Friboulet, D., D'hooge, J., Bernard, O.: Fast and fully automatic 3D echocardiographic segmentation using b-spline explicit active surfaces: Feasibility study and validation in a clinical setting. Ultrasound Med Biol **39**(1) (2013) 89–101

11. Queirós, S., Barbosa, D., Heyde, B., Morais, P., Vilaça, J., Friboulet, D., Bernard, O., D'hooge, J.: Fast automatic myocardial segmentation in 4D cine CMR datasets. Medical Image Analysis (2014) (In Press)

12. Ha, J., Walker, W., Hossack, J.: Determination of an optimal image frame interval for frame-to-frame ultrasound image motion tracking. IEEE T Ultrason Ferr **52**(3) (2005) 386–396

13. Duan, Q., Angelini, E., Herz, S., Ingrassia, C., Costa, K., Holmes, J., Homma, S., Laine, A.: Region-based endocardium tracking on real-time three-dimensional ultrasound. Ultrasound Med Biol **35**(2) (2009) 256–265