

heartBEATS: A Hybrid Energy Approach for Real-Time B-spline Explicit Active Tracking of Surfaces

Daniel Barbosa^a, João Pedrosa^{a,*}, Brecht Heyde^a, Thomas Dietenbeck^b, Denis Friboulet^b, Olivier Bernard^b, Jan D'hooge^a

^aLaboratory on Cardiovascular Imaging and Dynamics, KU Leuven, Belgium

^bUniversité de Lyon 1, CREATIS, CNRS UMR5220, INSERM U630; INSA-Lyon, France

Abstract

In this manuscript a novel method is presented for left ventricle (LV) tracking in three-dimensional ultrasound data using a hybrid approach combining segmentation and tracking-based clues. This is accomplished by coupling an affine motion model to an existing LV segmentation framework and introducing an energy term that penalizes the deviation to the affine motion estimated using a global Lucas-Kanade algorithm. The hybrid nature of the proposed solution can be seen as using the estimated affine motion to enhance the temporal coherence of the segmented surfaces, by enforcing the tracking of consistent patterns, while the underlying segmentation algorithm allows to locally refine the estimated global motion. The proposed method was tested on a dataset composed of 24 4D ultrasound sequences from both healthy volunteers and diseased patients. The proposed hybrid tracking platform offers a competitive solution for fast assessment of relevant LV volumetric indices, by combining the robustness of affine motion tracking with the low computational burden of the underlying segmentation algorithm.

Keywords: real-time image segmentation, B-spline, left ventricle segmentation and tracking

*Corresponding Author: João Pedrosa, Medical Imaging Research Center, Herestraat 49, bus 7003, B-3000 Leuven, Belgium; Email: joao.pedrosa@kuleuven.be

Introduction

Despite the existence of several functional indices, ejection fraction remains the most widely used parameter to assess global cardiac function (Thomas and Popovic, 2006). While Cardiac Magnetic Resonance Imaging (CMRI) remains the gold standard to assess cardiac morphology and function, it is still impractical for large patient throughput. Thanks to the technological maturation of 2D matrix array transducers, 3D ultrasonic imaging systems have been gradually introduced into clinical practice over the past decade. When compared with the conventional 2D echocardiographic (2DE) examination, real-time 3D echocardiography (RT3DE) offers some important advantages. Indeed, it overcomes several known limitations of conventional 2D echocardiographic imaging, since it avoids foreshortening, out-of-plane motion and the need of geometric assumptions for volume estimation. As a result, it has already been shown that RT3DE offers superior performance in the assessment of global morphology and function than 2DE, when compared against CMRI measurements (Muraru et al., 2010). Thus, the ability of combining accurate 3D volumetric assessment of cardiac morphology and function with the intrinsic benefits of ultrasound imaging makes RT3DE a very useful and promising tool for clinical daily routine.

It should be noticed that the added dimensionality of RT3DE data also poses some challenges in the data analysis pipeline, when compared with conventional 2D echocardiography. The manual analysis of RT3DE data remains cumbersome and time consuming, which has triggered the development of several software suites in order to reduce the burden on the operating physician while extracting relevant cardiac diagnostic information. However, even state-of-the-art commercial solutions still require some degree of user interaction both at the initialization step and for correction of the segmentation/tracking results (Muraru et al., 2010).

Nonetheless, several relevant methods have been proposed in order to increase the automation degree and decrease the total analysis time of RT3DE data. For instance, Dikici and Orderud (2012) have recently extended the Kalman-based, computationally efficient tracking framework initially introduced in Orderud et al. (2007), in order to account for advanced edge detectors to track the position of the left ventricular endocardial border throughout the cardiac cycle. Leung et al. (2010) proposed a fully automatic method for 3D echocardiographic data segmentation based on 3D Active Appearance Models (AAM). The same authors have equally proposed a tracking framework able to take into account statistical cardiac motion models to improve the performance of an optical flow based tracking algorithm (Leung et al., 2011). Yang et al. (2011) have proposed a computational framework targeted to robust and fast 3D tracking of deformable objects without any user interaction, by combining the input of multiple collaborative trackers. Different commercial solutions also exist in the field, typically dependent on user interaction for single frame segmentation which is then propagated throughout the heart cycle using different tracking strategies as for example in Tomtec’s and Toshiba’s commercial tools (Pedrosa et al., 2016). On the other hand, other approaches such as the newest available tools from Siemens and Philips, use large statistical models to perform the segmentation in 4D.

B-Spline Explicit Active Surfaces (BEAS) is a segmentation framework able to segment 3D data in real-time, introduced originally in Barbosa et al. (2012b). Despite the promising results of this approach on the segmentation of RT3DE data in both end-diastolic and end-systolic frames shown in Barbosa et al. (2013a), the direct application of this algorithm to track the left ventricle throughout the cardiac cycle has some intrinsic flaws. Indeed, segmentation-oriented energy functionals are designed to recover objects from a background in static images. Although they can be used in multi-static scenarios to perform tracking on subsequent images, there

is no guarantee that the temporal coherence of the patterns being tracked will be captured.

With this in mind, a novel hybrid framework which combines both segmentation as well as motion-oriented clues is proposed. This is accomplished by coupling an affine motion model to the segmented LV surface and introducing an energy term that penalizes the deviation to the affine motion estimated using a global Lucas-Kanade algorithm. The hybrid nature of the proposed solution can be seen as using the estimated affine motion to enhance the temporal coherence of the segmented surfaces, by enforcing the tracking of consistent patterns, while the underlying segmentation algorithm allows to locally refine the estimated global motion. The proposed solution offers a competitive approach for fast assessment of relevant LV volumetric indices, by combining the robustness of affine motion tracking with the low computational burden of BEAS.

The present manuscript is structured as follows. First, a revision of the fundamental notions of image segmentation using BEAS is presented, while also introducing the key novelty of our work, a hybrid tracking platform relying on both segmentation-based energies and tracking-oriented clues. It is shown that this coupling can be done via affine transformation of the coordinate system associated with the segmented LV surface. The key parameter values chosen in the implementation of the proposed algorithm are then addressed. In the Results section, an evaluation of the performance of the method using a dataset composed of 24 4D ultrasound exams is performed. In the Discussion, the main findings of the experiments are discussed and the performance of the proposed algorithm is compared against the most relevant prior work in literature. Finally, the Conclusions section gives the main conclusions and perspectives of this work.

Methodology

B-Spline Explicit Active Surfaces (BEAS)

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.* $x_1 = \psi(x_2, \dots, x_n) = \psi(\mathbf{x}^*)$. Following the mathematical formalism introduced in the work of Barbosa et al. (2012b), ψ was defined as a linear combination of B-spline basis functions:

$$\psi(\mathbf{x}^*) = \sum_{\mathbf{k} \in \mathbb{Z}^{n-1}} c[\mathbf{k}] \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k} \right), \quad (1)$$

where \mathbf{x}^* is the point of coordinates $\{x_2, \dots, x_n\}$ and $\beta^d(\cdot)$ the uniform $(n-1)$ -dimensional B-spline of degree d . The knots of the B-splines are located on a rectangular grid defined on the chosen coordinate system, with a regular spacing given by h . The coefficients of the B-spline representation are gathered in $c[\mathbf{k}]$.

In the present work, a modified version of the localized means separation energy is used, which takes advantage of the darker appearance of blood with respect to the myocardial tissue, introduced in Barbosa et al. (2013a), which is expressed as:

$$E_L = \int_{\Omega} \delta_{\phi}(\mathbf{x}) \int_{\Omega} B(\mathbf{x}, \mathbf{y}) (u_x - v_x) \, d\mathbf{y} \, d\mathbf{x}, \quad (2)$$

where $B(\mathbf{x}, \mathbf{y})$ is a mask function where the local means inside and outside the interface Γ , u_x and v_x respectively, are estimated. $\delta_{\phi}(\mathbf{x})$ is the Dirac operator applied to the level-set like function $\phi(\mathbf{x}) = \psi(\mathbf{x}^*) - x_1$, which is defined over the image domain Ω . This segmentation functional can be directly minimized wrt. the B-

spline coefficients $c[\mathbf{k}]$ according to:

$$\frac{\partial E_L}{\partial c[\mathbf{k}_i]} = \int_{\Gamma} \left(\frac{(\bar{I}(\mathbf{x}^*) - u_x)}{A_u} + \frac{(\bar{I}(\mathbf{x}^*) - v_x)}{A_v} \right) \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k}_i \right) d\mathbf{x}^*, \quad (3)$$

where A_u and A_v are the areas inside and outside the interface Γ used to estimate the local means u_x and v_x , respectively. For clarity sake, $\bar{I}(\mathbf{x}^*)$ corresponds to the image value at the position $\mathbf{x} = \{\psi(\mathbf{x}^*), x_2, \dots, x_n\}$. For the complete derivation of equation (3) the reader is referred to the original paper by Barbosa et al. (2013a). For the present work, ψ is defined in the spherical space, *i.e.* $\rho = \psi(\theta, \varphi)$. Further details regarding the fundamental formalism of BEAS and the derivation of the segmentation energy minimization strategy can be found in Barbosa et al. (2012b) and Barbosa et al. (2013a) respectively.

Fast left ventricular affine motion estimation

Three-dimensional cardiac motion assessment is a very active research field and different algorithms have been proposed. While block-matching is a very popular approach in current commercial software suites (Jasaityte et al., 2013), the recent trends in the research community show a tendency towards solutions based on elastic registration and optical flow algorithms, as highlighted in the recent comparative study by De Craene et al. (2013). In the present manuscript, a global 3D extension of the method proposed by Sühling et al. (2004) for the estimation of the local affine motion is used, as introduced in Barbosa et al. (2013b).

As noted in the seminal work of Lucas and Kanade (1981), the least squares solution of the optical flow equation is equivalent to the first order Taylor expansion of the minimization of the sum of squared differences between two subsequent frames. Therefore, optical flow motion estimation algorithms build upon the assumption that

the intensity of a particular point in a moving pattern does not change over time and thus any difference in the local appearance of a region over a sequence is uniquely defined by the underlying motion. Let $I(x_1, x_2, x_3, t)$ denote the pixel intensity at location $\mathbf{x} = [x_1, x_2, x_3]$ and time t in a 4D image dataset. This assumption can then be formulated as (Horn and Schunck, 1981):

$$I_{x_1}(\mathbf{x}, t)u(\mathbf{x}, t) + I_{x_2}(\mathbf{x}, t)v(\mathbf{x}, t) + I_{x_3}(\mathbf{x}, t)w(\mathbf{x}, t) + I_T(\mathbf{x}, t) = 0, \quad (4)$$

where $\nabla I = [I_{x_1}, I_{x_2}, I_{x_3}]$ is the local image spatial gradient and I_T corresponds to the temporal derivative. u , v and w are the x_1 -, x_2 - and x_3 -components of the optical flow that we wish to estimate.

Taking the affine motion model as defined in Sühling et al. (2004), the 3D affine motion on frame t can be estimated by minimizing the following energy term:

$$E_M(t) = \int \mathcal{W}(x_1 - c_1, x_2 - c_2, x_3 - c_3) (I_{x_1}u + I_{x_2}v + I_{x_3}w + I_T)^2 d\mathbf{x}, \quad (5)$$

where \mathcal{W} is a local window function centered in the position $\mathbf{c} = [c_1, c_2, c_3]$ and

$$u(\mathbf{x}, t) = u_0 + u_1(x_1 - c_1) + u_2(x_2 - c_2) + u_3(x_3 - c_3), \quad (6)$$

$$v(\mathbf{x}, t) = v_0 + v_1(x_1 - c_1) + v_2(x_2 - c_2) + v_3(x_3 - c_3), \quad (7)$$

$$w(\mathbf{x}, t) = w_0 + w_1(x_1 - c_1) + w_2(x_2 - c_2) + w_3(x_3 - c_3), \quad (8)$$

encode the local motion field along respectively x_1 , x_2 and x_3 . Parameters u_0 , v_0 and w_0 correspond to the motion at the window center and u_1 , u_2 , u_3 , v_1 , v_2 , v_3 , w_1 , w_2 and w_3 are respectively the first order spatial derivatives of u , v and w . These parameters then define the affine transform for frame t associated with the local

motion field $[u, v, w]$ as expressed by the augmented matrix

$$\mathbf{M}_t = \begin{bmatrix} 1 + u_1 & u_2 & u_3 & u_0 \\ v_1 & 1 + v_2 & v_3 & v_0 \\ w_1 & w_2 & 1 + w_3 & w_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

By differentiating (5) with respect to the affine motion field components, it can be shown that the minimization of the weighted least-squares criterion in (5) can be expressed as the solution of $\mathbf{A}^T \mathcal{W} \mathbf{A} \mathbf{d} = \mathbf{A}^T \mathcal{W} \mathbf{b}$, whose components are expanded at the bottom of the page. Note that $\langle a, b \rangle$ denotes the continuous analogue of the dot product, expressed as $\int a(\mathbf{x})b(\mathbf{x})d\mathbf{x}$ and that \mathbf{A} , \mathbf{d} and $\mathbf{A}^T \mathcal{W} \mathbf{b}$ are functions of (\mathbf{x}, t) though this was omitted for simplicity.

A global formulation of the algorithm introduced by Sühling *et al.* could be employed by considering $\mathcal{W}(\mathbf{p}) = 1, \forall \mathbf{p}$. Such assumption would yield the global affine transform between the two subsequent images. Nonetheless, increasing the span of the window function \mathcal{W} intrinsically hampers the underlying assumption that the affine motion model is constant within this region. This is particularly problematic in the case of echocardiographic data, since it is known that the motion patterns of the blood and the surrounding tissues (e.g. pericardium and valves) are significantly different from the ones within the myocardium, thus violating the fundamental as-

$$\mathbf{A} = \begin{bmatrix} I_{x_1} \\ I_{x_2} \\ I_{x_3} \\ x_1 I_{x_1} \\ x_2 I_{x_1} \\ x_3 I_{x_1} \\ x_1 I_{x_2} \\ x_2 I_{x_2} \\ x_3 I_{x_2} \\ x_1 I_{x_3} \\ x_2 I_{x_3} \\ x_3 I_{x_3} \end{bmatrix}, \quad [\mathbf{A}^T \mathcal{W} \mathbf{A}]_{ij} = \langle A_i \mathcal{W}, A_j \rangle, \quad \mathbf{d} = \begin{bmatrix} u_0 \\ v_0 \\ w_0 \\ u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \\ v_3 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix}, \quad \mathbf{A}^T \mathcal{W} \mathbf{b} = - \begin{bmatrix} \langle \mathcal{W}, I_{x_1} I_T \rangle \\ \langle \mathcal{W}, I_{x_2} I_T \rangle \\ \langle \mathcal{W}, I_{x_3} I_T \rangle \\ \langle x_1 \mathcal{W}, I_{x_1} I_T \rangle \\ \langle x_2 \mathcal{W}, I_{x_1} I_T \rangle \\ \langle x_3 \mathcal{W}, I_{x_1} I_T \rangle \\ \langle x_1 \mathcal{W}, I_{x_2} I_T \rangle \\ \langle x_2 \mathcal{W}, I_{x_2} I_T \rangle \\ \langle x_3 \mathcal{W}, I_{x_2} I_T \rangle \\ \langle x_1 \mathcal{W}, I_{x_3} I_T \rangle \\ \langle x_2 \mathcal{W}, I_{x_3} I_T \rangle \\ \langle x_3 \mathcal{W}, I_{x_3} I_T \rangle \end{bmatrix}$$

sumption of constant motion within the region of interest (ROI). In order to allow estimating the global affine transformation of the LV between subsequent frames, the existing segmentation framework will be used and the formalism introduced in Barbosa et al. (2013b) to define a ROI only around the segmented surface from the convolution expanded so that:

$$\mathcal{W}(\mathbf{x}) = \delta_\phi(\mathbf{x}) * N(\mathbf{x}), \quad (10)$$

where $\delta_\phi(\mathbf{x})$ implicitly defines the segmented surface and $N(\mathbf{x})$ is simply a neighborhood function defined as a 3D cube centered in \mathbf{x} . An example of the resulting $\mathcal{W}(\mathbf{x})$ is shown in Figure 1.

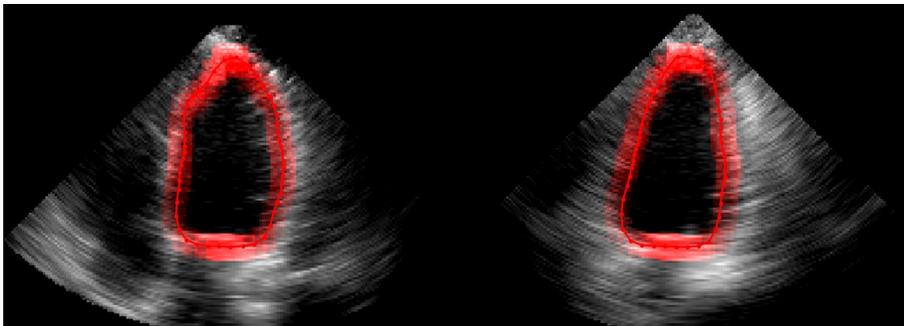


Figure 1: Anatomical ROI for affine motion estimation.

Hybrid framework for fast left ventricle tracking

The integration of the motion information estimated with the aforementioned optical flow algorithm within the existing segmentation framework will enhance its tracking performance. This is achieved since the temporal coherence of the patterns being tracked is the fundamental driver of affine motion estimation, whereas the existing segmentation framework is mostly relying on the identification of salient features in static images. By combining the two, a hybrid framework will allow

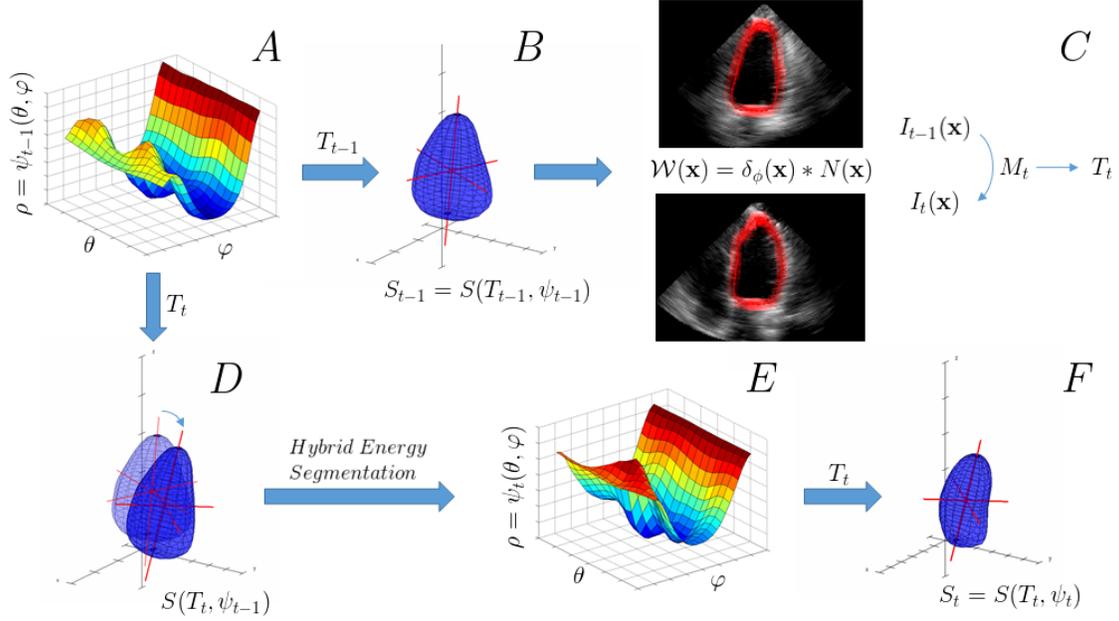


Figure 2: Conceptual description of the proposed hybrid algorithm. The explicit function ψ_{t-1} (A) defining the LV surface \mathbf{S}_{t-1} through \mathbf{T}_{t-1} (B) is used to estimate the affine transformation \mathbf{M}_t between two consecutive frames in the anatomical ROI $\mathcal{W}(\mathbf{x})$ (C). This gives \mathbf{T}_t , the LV global pose for frame t . This affine deformation is then applied to \mathbf{S}_{t-1} (shown in greater transparency) giving $\mathbf{S}(\mathbf{T}_t, \psi_{t-1})$ (D). Hybrid energy segmentation is then applied to evolve ψ_{t-1} according to equation 15 to obtain ψ_t (E), which, through \mathbf{T}_t , defines the final LV surface for frame t , \mathbf{S}_t (F).

a synergistic collaboration between tracking-based and segmentation-based clues, increasing the overall performance and robustness. In the present sub-section the manner through which the hybrid tracking strategy can be implemented via affine deformation of the coordinate system associated with the segmented LV surface is described. A conceptual description of the proposed approach is illustrated in Figure 2.

First, the inherent characteristics of the BEAS segmentation framework should be recalled: BEAS models the LV object (in 3D) through an explicit function ψ_t in the spherical coordinate system, thus a function of azimuthal and elevation angles θ and φ . As such, to translate this explicit function ψ_t into the LV surface \mathbf{S}_t in

the Cartesian coordinate system, not only does the spherical to Cartesian conversion need to be performed, but the orientation and position of the spherical coordinate system (in which the BEAS surface is defined) relative to the Cartesian system have to be taken into account. The LV surface can thus be defined as:

$$\mathbf{S}_t = \mathbf{S}(\mathbf{T}_t, \psi_t) = \mathbf{T}_t \begin{bmatrix} \psi_t(\theta, \varphi) \cos(\theta) \sin(\varphi) \\ \psi_t(\theta, \varphi) \sin(\theta) \sin(\varphi) \\ \psi_t(\theta, \varphi) \cos(\varphi) \end{bmatrix}, \quad (11)$$

where \mathbf{T}_t is the augmented matrix comprising the orientation of the LV long axis and its center position on frame t . This relationship is evident in Figure 2-A and B, where the explicit function ψ_{t-1} is transformed into \mathbf{S}_{t-1} by employing equation 11. In this equation it is clear that, though ψ_t controls the local shape of the LV surface, this surface can also be globally deformed through \mathbf{T}_t .

Given that the estimated motion field can be expressed as an affine transformation, the estimated affine transformation between subsequent frames, \mathbf{M}_t , will be coupled to the underlying spherical to Cartesian transformation, using the recursive formulation:

$$\mathbf{T}_t = \mathbf{M}_t \mathbf{T}_{t-1}, \quad (12)$$

where \mathbf{T}_1 is the augmented matrix of orientation and center position of the LV in the end-diastolic frame. By applying the current estimate of \mathbf{T}_t to the underlying spherical to Cartesian transformation, the entire surface can be intrinsically deformed according to the estimated affine transformation through both translation, rotation and scaling. This is in fact the equivalent of deforming the original coordinate frame of the LV object according to the global affine motion estimated with the anatomical optical flow algorithm detailed previously, as illustrated in Figure 3. In the proposed

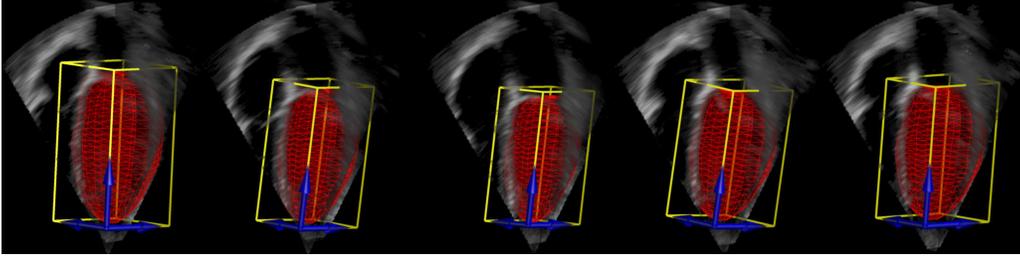


Figure 3: Affine deformation of the left ventricular surface along time. For clarity sake, the coordinate system associated with the LV object was translated to the corner of the bounding box to facilitate its visualization.

hybrid framework, the LV surface \mathbf{S}_{t-1} (Figure 2-B) is used to determine $\mathcal{W}(\mathbf{x})$ which in turn is used to estimate the affine transformation \mathbf{M}_t and obtain the new LV surface pose \mathbf{T}_t (Figure 2-C). This augmented matrix \mathbf{T}_t is then applied to ψ_{t-1} to deform the LV surface according to the affine motion estimated as shown in Figure 2-D.

In order to balance the contribution between tracking and segmentation-based clues, a hybrid segmentation is then performed to this deformed surface. For that purpose, an energy term that penalizes the deviation between the current surface position, $\mathbf{S}(\mathbf{T}_t, \psi_t)$, and the one initialized by applying the affine transformation to the previous segmentation result, *i.e.* $\mathbf{S}(\mathbf{T}_t, \psi_{t-1})$, is proposed:

$$E_A(t) = \int_{\Gamma} (\mathbf{S}(\mathbf{T}_t, \psi_t(\theta, \varphi)) - \mathbf{S}(\mathbf{M}_t \mathbf{T}_{t-1}, \psi_{t-1}(\theta, \varphi)))^2 d\mathbf{x}^*, \quad (13)$$

which, because the global pose $\mathbf{T}_t = \mathbf{M}_t \mathbf{T}_{t-1}$ is shared between the two LV surfaces, can be simplified to:

$$E_A(t) = \int_{\Gamma} (\psi_t(\theta, \varphi) - \psi_{t-1}(\theta, \varphi))^2 d\mathbf{x}^*, \quad (14)$$

The global energy term for optimization can then be defined as:

$$E(t) = E_L(t) + \lambda E_A(t), \quad (15)$$

where λ is a hyperparameter controlling the balance between the segmentation and the tracking-based energies. This energy criterion can be minimized directly wrt. the B-spline coefficients controlling the shape of ψ_t :

$$\frac{\partial E(t)}{\partial c[\mathbf{k}_i]} = \frac{\partial E_L(t)}{\partial c[\mathbf{k}_i]} + \lambda \frac{\partial E_A(t)}{\partial c[\mathbf{k}_i]}, \quad (16)$$

where

$$\frac{\partial E_A(t)}{\partial c[\mathbf{k}_i]} = 2 \int_{\Gamma} (\psi_t(\theta, \varphi) - \psi_{t-1}(\theta, \varphi)) \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k}_i \right) d\mathbf{x}^*, \quad (17)$$

and $\partial E_L(t)/\partial c[\mathbf{k}_i]$ is defined as in (3).

Note that even though the affine transform parameters are not explicitly stated through $E(t)$, these are present in the term $E_A(t)$ as shown in equation (13). This term penalizes the deviation from the LV surface position after the affine deformation, guaranteeing that the data attachment term from the segmentation energy functional only modifies the segmented LV surface in regions where strong image content is available. Thus, LV regions with reduced image content are preferably tracked with the global affine deformations via optical flow estimation, while regions with rich image content rely mostly on the image data. Therefore, there is an intrinsic trade-off between global, robust tracking and localized, accurate surface positioning.

By performing this hybrid segmentation, the explicit function ψ_t is then obtained (Figure 2-E), which, through \mathbf{T}_t , defines the final LV surface for frame t , \mathbf{S}_t (Figure 2-F) and thus enabling to restart the process for the next frame.

Implementation Details

In order to speed-up the estimation of $\mathcal{W}(\mathbf{x})$, the convolution expressed in (10) defining the anatomical ROI where E_M is evaluated was simplified by assuming that $\delta_\phi(\mathbf{x})$ is different from zero only in the positions of the discretized BEAS surface. N was defined as a 11x11x11 cube centered in the target point. The image gradient ∇I was estimated using a gaussian derivative kernel with $\sigma = 1$, implemented as a separable convolution operation. Since differential optical flow approaches are best suited to estimate small displacements, the optimization of $E_M(t)$ was performed using an iterative displacement refinement scheme to improve the accuracy and robustness of the affine motion \mathbf{M}_t estimation (Bouguet, 1999). Five iterations were used in all experiments.

The hyperparameter λ controlling the balance between the segmentation and tracking-based terms in (15) was empirically set to 0.25. As in previous BEAS implementations for the LV (Barbosa et al., 2013a), the angular discretization of the boundary ϕ was set to 24×16 and the B-spline scale h to 2^1 . The mask function $B(\mathbf{x}, \mathbf{y})$ was restricted to the points along the normal direction of the surface at a distance smaller than 16mm as in Barbosa et al. (2013a). The optimization of the global segmentation energy $E(t)$ was implemented in a modified gradient descent with feedback step adjustment as in previous BEAS implementations (Barbosa et al., 2013a).

Experiments and Results

Twenty-four RT3DE exams were acquired using a Siemens Acuson SC2000 rev. 1.5 (Siemens Ultrasound, Mountain View, CA) using a 4Z1c matrix transducer. Volume sequences were acquired during apical scanning and the sonographer aimed

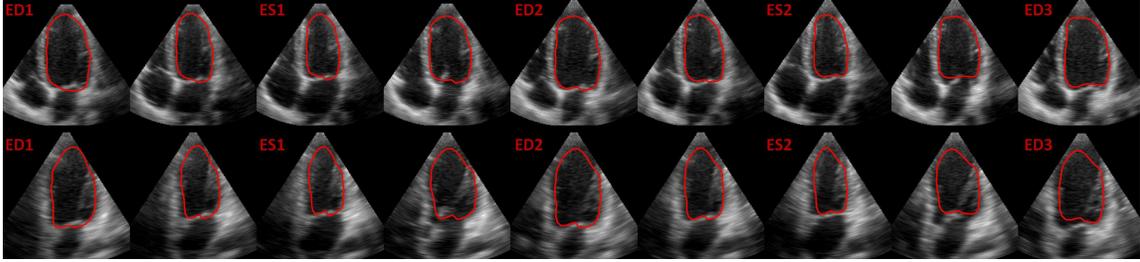


Figure 4: Tracking of the left ventricle in a RT3DE dataset using the proposed hybrid approach (top: apical 4 chambers view, bottom: apical 2 chamber view).

at the inclusion of the entire LV within the pyramidal field of view. Volume rates ranged from 20 to 40 volumes per second. Each sequence was analyzed by three experts using eSie LVA pre-release software (Siemens, Mountain View), who provided manual delineation of the left ventricular chamber at both end-diastolic and end-systolic frames. From these, the corresponding end diastolic (ED) and end systolic (ES) volumes were calculated. The stroke volume (SV) and ejection fraction (EF) were posteriorly computed from the EDV and ESV. The mean value of the three experts was taken as the reference for the aforementioned LV volumetric indices. The described protocol was approved by the institutional review board and patients signed an informed consent.

The proposed tracking framework was automatically initialized in the end-diastolic frame with the algorithm introduced in Barbosa et al. (2013a). In order to demonstrate the synergistic interaction of the segmentation and tracking-based clues, the proposed solution was compared with the pure segmentation-based approach, by setting λ to zero in (15), and also with a pure global tracking approach, by keeping ψ fixed and adjusting \mathbf{T}_t over time.

The summary of the results for the LV volumetric indices extracted using the proposed hybrid framework against the manual references can be found in Table 1. In the same table, the performances for both the pure affine tracking-based solution

Table 1: LV volumetric indexes extracted using the proposed hybrid approach versus the pure segmentation-based (PS) tracking and the pure global affine optical flow algorithm (LOA: limits of agreement; *, $p < 0.05$, paired t-test against zero).

	Correlation Coefficient (R)			Bland-Altman Analysis LOA (bias $\pm 1.96\sigma$)		
	PS	Affine	Hybrid	PS	Affine	Hybrid
EDV (ml)	0.964	0.966	0.971	-5.76* \pm 25.9	-3.68 \pm 26.2	-2.58 \pm 23.4
ESV (ml)	0.929	0.930	0.950	-8.40* \pm 26.6	2.43 \pm 26.0	-0.60 \pm 23.4
SV (ml)	0.904	0.906	0.934	2.64 \pm 25.6	-6.11 \pm 30.8	-1.99 \pm 24.8
EF (%)	0.734	0.776	0.833	5.35* \pm 16.3	-1.65 \pm 16.1	1.20 \pm 13.3

and the pure segmentation solution are also reported. An example of a RT3DE exam segmented using the proposed hybrid tracking algorithm is given in Figure 4, for 2 consecutive cardiac cycles. The segmentation of the first frame, which included the automatic initialization step, took approximately 1s, while the tracking between subsequent frames was done in 30ms, in a C++ implementation running on an Intel i7 laptop.

The key parameter to be tuned in the proposed algorithm is the hyperparameter λ controlling the balance between the contribution of the segmentation-based and tracking-based terms in equation (16). In order to test the sensitivity of the empirically chosen value, *i.e.* $\lambda = 0.25$, the value of λ was varied between 0 and 0.5, in steps of 0.05 and its corresponding LV tracking results observed. This allowed to assess the influence of removing the tracking-based term from equation (16) on the bottom side of the variation range, while the upper variation range corresponded to doubling the influence of the tracking-based clues. Additionally, the pure global tracking approach, where the influence the tracking-based clues tend to infinity, was also tested. The results of this sensitivity analysis are given in Figure 5.

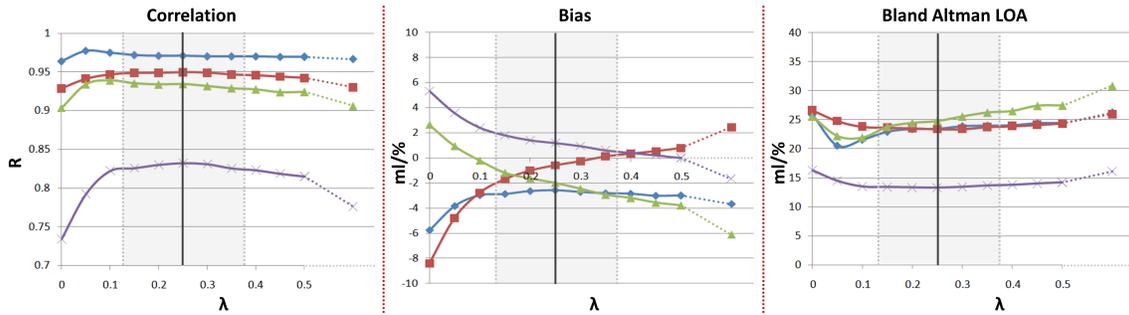


Figure 5: Influence of the balance between the segmentation and tracking-based terms in the overall performance of the proposed hybrid tracking algorithm (left to right: Pearson correlation coefficient, Bland-Altman bias (μ) and limits of agreement (1.96σ); EDV: blue; ESV: red; SV: green; EF: purple). The chosen value of λ is marked as a vertical gray line, while the range corresponding to a $\pm 50\%$ variation of its numeric value is shaded in light gray. Note that the leftmost data points in each plot correspond to a pure segmentation-based approach, while the rightmost ones correspond to a pure tracking approach.

Discussion

The proposed hybrid tracking framework offers competitive performance for the fully automatic quantification of relevant volumetric cardiac indices used in daily practice for assessment of left ventricular morphology and global function. This is supported by the strong correlation for all the estimated volumetric indices. Furthermore, low, non-statistically significant bias and tight limits of agreement were observed by Bland-Altman analysis. Comparing the results from the proposed hybrid approach to the pure segmentation and pure tracking strategies shown on Table 1 it becomes clear that the proposed hybrid approach outperforms both the pure segmentation and the pure tracking approach. There is thus a significant advantage on bringing together the segmentation- and tracking-based clues within the same approach.

Furthermore, the proposed approach compares positively against the pure affine motion estimation algorithm used to estimate the global LV deformation previously introduced in Barbosa et al. (2013b). This clearly indicates the advantages of the re-

finement stage using a hybrid combination of both segmentation and tracking-based clues. Indeed, despite the small inter-frame differences between the pure affine tracking and the proposed hybrid approach, a cumulative effect over the entire cardiac cycle leads to a significant improvement in tracking performance. Such observation is supported by the appreciable reduction in both bias and limits of agreement of stroke volume estimation, to which sums up the strong increase in the ejection fraction correlation against the reference measurements. Therefore, even if visually the added value of the proposed hybrid strategy is not strikingly evident between two frames, the accumulation over the entire cycle improves significantly the tracking performance. Furthermore, the comparison against the previously proposed strategy based on sequential segmentation (Barbosa et al., 2012a), whose results are included in the performance comparison in the following section, clearly highlights the synergistic effect between the tracking-oriented clues and segmentation-based energy terms.

Interestingly, the proposed affine coupling allows to deform the spherical discretization grid used to represent the segmented surface according to the estimated LV deformation. This enables capturing valuable information regarding both longitudinal and circumferential global motion of the heart which could not be properly evaluated with the previous segmentation framework. This opens the path towards the extraction of other cardiac global functional indices, such as global longitudinal and circumferential strain. Furthermore, the inclusion of the longitudinal/circumferential deformation via affine transformation of the coordinate system associated with the segmented LV surface provides a seamless integration on the previous formal framework, therefore not requiring any special modification to the underlying mathematical foundations.

The results of the sensitivity analysis demonstrate its robustness towards the

chosen value for the hyperparameter λ controlling the balance between the tracking-based and segmentation-based terms in (16). Indeed, a modification of $\pm 50\%$ of its nominal value does not lead to appreciable modifications in the correlation values for any of the LV volumetric indices considered in the present work. A similar trend is observed for the width of the limits of agreement for the EDV, SV and EF estimates. Nonetheless, the influence of λ is particularly visible in the performance of the ESV estimation. Such observation is explained by the opposite bias of the pure tracking-based and segmentation-based approaches. Indeed, while the pure segmentation-based approach over-estimates the true ESV volume, the pure tracking-based affine optical flow method underestimates it.

Performance comparison

Although the proposed hybrid tracking approach offers promising results, a careful comparison with the values reported in the literature has been done in order to evaluate its competitiveness against currently available solutions. This comparison does not aim to be extensive, but rather informative to the reader on how the proposed fully automatic framework for LV volume analysis compares with existing relevant methods. An overview of the results of the proposed algorithm and its comparison with other methods reported in the literature is shown in Table 2. Nonetheless, it should be noted that a fair and quantitative comparison is not trivial, due to differences in patient population and image quality and due to different acquisition conditions and equipment. We have selected relevant algorithms ranging from pure segmentation-based approaches to more oriented tracking strategies. However, methods with a similar validation approach (*i.e.* where the segmentation results were compared with manual segmentation of RT3DE data) were selected.

Table 2 also shows that the proposed algorithm presents a competitive performance when compared with the most relevant algorithms presented in the literature, both in terms of accuracy and overall computational load. Indeed, performance-wise only the algorithm of Leung et al. (2011) and Yang et al. (2011) provide clearly more accurate results than the ones provided by the developed algorithms. Nonetheless, both these algorithms are not able to run in real-time and rely on statistical shape and motion models. While prior knowledge is a very powerful tool to deal with missing information scenarios, which is often the case in RT3DE data where some of the boundaries are missing, care should be taken to not infer wrong estimations due to excessive influence of the priors. Furthermore, the ability to deal with unseen data is typically related with the amount of different patterns included in the initial learning phase, which implies very tedious and labor-intensive stages towards building such statistical databases. It should also be noted that the algorithm in Leung et al. (2011) is tracking-oriented and, thus, requires LV delineation at ED. While the same authors also propose an automatic ED segmentation algorithm in Leung et al. (2010), the validation of their complete segmentation/tracking framework remains to be done.

In terms of overall running time, only the work of Orderud et al. (2007); Orderud and Rabben (2008); Hansegard et al. (2007) and the framework of Duan et al. (2010) are able to compete with the proposed algorithm, which yields an average computing time of 30 ms per processed frame. On the other hand, level-set based algorithms, such as the ones in Angelini et al. (2005) and Rajpoot et al. (2011), require a significantly larger computational power due to the implicit representation of the evolving interface, with a direct trade-off between shape topology freedom and computational burden. Furthermore, in the current implementation of heart-BEATS there are redundant computations introduced within the anatomical ROI in

(10), since there are overlapping regions. This sums up to the high degree of parallelism in the key algorithmic blocks of the proposed method, which opens the path to further implementation optimizations which would allow to further reduce overall computational time. Therefore, there is still a considerable margin to improve the processing speed of heartBEATS, which will allow the method to be prepared to deal with higher frame-rate 3D acquisitions, currently a hot topic in the ultrasound community (Hasegawa and Kanai, 2011; Tong et al., 2013).

The comparison against the recent work of Zhang et al. (2013) also supports the competitive performance of the proposed algorithm. Despite validating their algorithm in a dataset composed of patients selected for cardiac resynchronization therapy, Zhang et al. (2013) reports relative volume errors of $4.2 \pm 17.4\%$ and $-1.3 \pm 16.8\%$ ($\mu \pm \sigma$) for the segmented LV surfaces considering the input of two different users. In the current study, the proposed algorithm yielded $-4.7 \pm 14.1\%$ considering both EDV and ESV values. Thus, heartBEATS presents a competitive performance when compared to the active shape model-based method of Zhang et al. (2013), while doing so without any kind of user input.

Conclusions

The proposed hybrid segmentation/tracking framework (heartBEATS) combines both segmentation-oriented image information with global tracking clues, for enhanced performance on the tracking of the left ventricular surface throughout the cardiac cycle. Furthermore, it allows assessing the motion components tangential to the LV boundaries, which was a limitation of the existing segmentation algorithm. Lastly, the computational burden is low, pointing towards the feasibility of accurate real-time online tracking.

Table 2: Proposed vs. state-of-the-art algorithms (#: number of exams; ΔT_f : average frame processing time (s); R: correlation coefficient; BA: Bland-Altman analysis; FC: full cycle, NR: not reported).

Study	Algorithm/Frames / User Input	#	ΔT_f	R			BA($\mu \pm 2\sigma$)		
				EDV	ESV	EF	EDV	ESV	EF
Prior work									
Angelini et al. (2005)	PS/ED+ES/I	10	NR	0.63	0.62	0.45	16.1±50	6.6±34	0.5±22
Hansegard et al. (2007)	MSS/FC/0	21	0.008	0.91	0.91	0.74	-5.9±21	6.2±19	-7.7±12
Leung et al. (2010)	PS/ED/0	99	NR	0.95	NR	NR	-1.47±40	NR	NR
Leung et al. (2011)	PT/FC/II	35	6	0.982	NR	NR	1.9±14	NR	NR
Yang et al. (2011)	HST/FC/0	67	1.5	NR	NR	NR	1.32±12	1.0±10	NR
Rajpoot et al. (2011)	PS/ED+ES/I	34	NR	NR	NR	NR	-5.0±49	1.2±26	-0.7±14
Rajpoot et al. (2011)	PT/FC/II	34	NR	NR	NR	NR	NR	4.0±40	-3.3±25
Barbosa et al. (2012a)	MSS/FC/0	24	0.05	0.98	0.92	0.78	-3.9±22	-5.0±27	3.4±15
Zhang et al. (2013)	MSS/FC/II	34	10	0.84	NR	NR	NR	NR	NR
Proposed									
heartBEATS	HST/FC/0	24	0.03	0.97	0.95	0.83	-2.6±23	-0.6±23	1.2±13

Note that PS, MSS, PT and HST stand for the algorithm class, namely pure segmentation, multi-static segmentation, pure tracking and hybrid segmentation and tracking. Regarding user input, 0 stands for a fully automatic method, I for minor user input (such as few anatomical landmarks) and II for significant user input, such as manual contouring at the end-diastolic frame.

Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia, through the PhD. grant SFRH/BD/62230/2009 and European Research Council under the European Unions Seventh Framework Program (HeartMAPAS; FP7/2007-2013/ERC Grant Agreement number 281748).

The authors would like to thank Helene Houle (Siemens Healthcare) for providing the volumetric ultrasound data as well as the reference measurements.

Conflict of Interest Statement

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

- Angelini E, Homma S, Pearson G, Holmes J, Laine A. Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function: A clinical study on right and left ventricles. *Ultrasound In Medicine and Biology*, 2005;31:1143–1158.
- Barbosa D, Bernard O, Heyde B, Dietenbeck T, Houle H, Friboulet D, D’hooge J. B-spline explicit active tracking of surfaces (BEATS): Application to real-time 3D segmentation and tracking of the left ventricle in 3D echocardiography. In: *Proceedings of IEEE IUS2012*, 2012a. pp. 224–227.
- 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 in Medicine and Biology*, 2013a;39:89–101.
- 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 Transactions on Image Processing*, 2012b;21:241–251.
- 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: *Proceedings of Functional Imaging and Modeling of the Heart - FIMH2013*, 2013b. pp. 191–199.
- Bouguet J. Pyramidal implementation of the affine lucas kanade feature tracker. Tech. rep., Intel Corporation, 1999.
- De Craene M, Marchesseau S, Heyde B, Gao H, Alessandrini M, Bernard O, Piella G, Porras A, Saloux E, Tautz L, Hennemuth A, Prakosa A, Liebgott H, Somphone O,

- Allain P, Ebeid S, Delingette H, Sermesant M, D'hooge J. 3D strain assessment in ultrasound (Straus): A synthetic comparison of five tracking methodologies. *IEEE Transactions on Medical Imaging*, 2013;32:1632–1642.
- Dikici E, Orderud F. Generalized step criterion edge detectors for kalman filter based left ventricle tracking in 3D+T echocardiography. In: Camara O, Mansi T, Pop M, Rhode K, Sermesant M, Young A (Eds.), *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*. Vol. 7746 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012. pp. 261–269.
- Duan Q, Angelini ED, Laine AF. Real-time segmentation by active geometric functions. *Computer methods and programs in biomedicine*, 2010;98:223–230.
- Hansegard J, Orderud F, Rabben S. Real-time active shape models for segmentation of 3D cardiac ultrasound. In: Kropatsch W, Kampel M, Hanbury A (Eds.), *Computer Analysis of Images and Patterns*. Vol. 4673 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2007. pp. 157–164.
- Hasegawa H, Kanai H. High-frame-rate echocardiography using diverging transmit beams and parallel receive beamforming. *Journal of Medical Ultrasonics*, 2011;38:129–140.
- Horn BK, Schunck BG. Determining optical flow. *Artificial Intelligence*, 1981;17:185 – 203.
- Jasaityte R, Heyde B, D'hooge J. Current state of three-dimensional myocardial strain estimation using echocardiography. *Journal of The American Society Of Echocardiography*, 2013;26:15–28.

- Leung K, 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 In Medicine and Biology*, 2011;37:605–616.
- Leung K, van Stralen M, van Burken G, de Jong N, Bosch J. Automatic active appearance model segmentation of 3D echocardiograms. In: *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI2010)*, 2010. pp. 320–323.
- Lucas B, Kanade T. An iterative image restoration technique with an application to stereo vision. In: *Proc. DARPA Image Understanding Workshop*, 1981. pp. 121–130.
- 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. *European Journal of Echocardiography*, 2010;11:359–368.
- Orderud F, Hansegard J, Rabben S. Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach. In: Ayache N, Ourselin S, Maeder A (Eds.), *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2007*. Vol. 4791 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007. pp. 858–865.
- Orderud F, Rabben S. Real-time 3D segmentation of the left ventricle using deformable subdivision surfaces. In: *Proceedings of IEEE Conference Computer Vision and Pattern Recognition (CVPR2008)*, 2008. pp. 1–8.
- Pedrosa J, Barbosa D, Almeida N, Bernard O, Bosch J, D’hooge J. Cardiac chamber

- volumetric assessment using 3D ultrasound - a review. *Current Pharmaceutical Design*, 2016;22:105–121.
- Rajpoot K, Grau V, Noble JA, Becher H, Szmigielski C. The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking. *Medical Image Analysis*, 2011;15:514 – 528.
- Sühling M, Arigovindan M, Hunziker P, Unser M. Multiresolution moment filters: theory and applications. *IEEE Transactions on Image Processing*, 2004;13:484 – 495.
- Thomas J, Popovic Z. Assessment of left ventricular function by cardiac ultrasound. *Journal of the American College of Cardiology*, 2006;48:2012–2025.
- Tong L, Gao H, D’hooge J. Multi-transmit beam forming for fast cardiac imaging-a simulation study. *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2013;60:1719–1731.
- 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 Transactions on Medical Imaging*, 2011;30:1921 –1932.
- Zhang H, Abiose A, Gupta D, Campbell D, Martins J, Sonka M, Wahle A. Novel indices for left-ventricular dyssynchrony characterization based on highly automated segmentation from real-time 3-D echocardiography. *Ultrasound in Medicine and Biology*, 2013;39:72 – 88.