10.1109/ULTSYM.2009.0346

Tracking the Endocardial Border in Artifact-Prone 3D Images

K.Y. Esther Leung*, Mikhail G. Danilouchkine*, Marijn van Stralen*,
Nico de Jong*†, Antonius F.W. van der Steen*†, and Johan G. Bosch*
*Biomedical Engineering, Thoraxcenter, Erasmus MC, Rotterdam, the Netherlands
†ICIN - Interuniversity Cardiology Institute of the Netherlands, Utrecht, the Netherlands

***Abstract*— Echocardiography is a commonly-used, safe, and noninvasive method for assessing cardiac dysfunction and related coronary artery disease. The analysis of echocardiograms, whether visual or automated, has traditionally been hampered by the presence of ultrasound artifacts, which obscure the moving myocardial wall. In this study, a novel method is proposed for tracking the endocardial surface in 3D ultrasound images. Artifacts which obscure the myocardium are detected in order to improve the quality of cardiac boundary segmentation. The expectation-maximization algorithm is applied in a stationary and dynamic, cardiac-motion frame-of-reference, and weights are derived accordingly. The weights are integrated with an optical-flow based contour tracking method, which incorporates prior knowledge via a statistical model of cardiac motion. Evaluation on 35 three-dimensional echocardiographic sequences shows that this weighed tracking method significantly improves the tracking results. In conclusion, the proposed weights are able to reduce the influence of artifacts, resulting in a more accurate quantitative analysis.**

I. Introduction

A. Artifacts in Cardiac Ultrasound

Echocardiography is a commonly-used, fast, and relatively inexpensive imaging modality for assessing left ventricular dysfunction and underlying coronary artery disease. The analysis of echocardiograms, whether visual or automated, has traditionally been hampered by the presence of ultrasound artifacts and anomalies, such as side-lobe artifacts, reverberations, shadowing, and near-field clutter [1]. For accurate assessment of the heart, it is highly desirable to identify such artifacts.

Side-lobes of the main, central ultrasound beam may cause echoes coming from the side-lobe areas to appear as if they come from the main beam. This results in an arched 'edge' in the image, and can be clearly seen if the object in the side-lobe has a strong reflection. Reverberations occur when the sound beam bounces back and forth between reflectors in the imaged area before it is received by the transducer. This artifact appears as one or more echo targets directly behind the reflector. A more troublesome anomaly is shadowing and drop-outs, which causes loss of signal of the imaged structures. Shadowing is caused by objects in the ultrasound path which cause high reflections or attenuation. In transthoracic cardiac images, this is usually caused by the rib-cage, which prevents the imaging of the myocardium. Drop-outs occur when the imaged structure is parallel to the ultrasound beam. Near field clutter arises from reverberations of high-amplitude signals close to the transducer. This hampers the imaging of objects in this area, and is particularly noticeable in echocardiograms acquired from the apical window, where the near-field clutter often obscures the left ventricular apex.

For removing ultrasound artifacts, much effort involving hardware improvements has been described in the literature. For example, different pulsing schemes and frequencies have been proposed, such as second harmonic imaging [2] and coded excitation [3]. To improve the overall visibility of the myocardium in echocardiograms, contrast agents are often employed [4]. Image post-processing methods have also been proposed, e.g. for identifying reverberations [5], [6], near-field noise [7], and stationary artifacts in general [8].

B. Study Goal

In this study, we aim at tracking the endocardial borders in 3D echocardiographic sequences. The cardiac motion, estimated using the image intensity information, is applied to propagate an end-diastolic (ED) contour throughout the sequences. The tracking is enhanced by detecting areas in which the myocardium is obscured by typical artifacts such as shadowing, near-field clutter, and static reverberations. The motion is estimated first using a global tracker, which is guided by a statistical model of cardiac motion. This estimation is locally refined using a purely data-driven optical flow approach. The local refinement is applied, depending on the visibility of the cardiac wall. To determine this, we apply a fast expectation-maximization (EM) approach in a stationary and in a dynamic (i.e. based on the initially estimated cardiac motion) frame-of-reference. The EM algorithm is used to analyze local temporal intensity profiles in the images, thus generating weights corresponding with visible and artifact-obscured myocardium. These weights are then used to give a better combination of the global and local tracker.

II. Methods

A. Global and local motion estimation

Previously, we have proposed an optical-flow-based method for tracking left ventricular endocardial borders throughout the cardiac cycle [9]. The method consists of a *global* tracker which is guided by a statistical cardiac motion model, followed by a *local* refinement using purely data-driven tracking.

1) Local tracker: The local tracker estimates the cardiac motion using differential optical flow [10]. The general optical flow equation describes the velocity $\mathbf{v}$ of an object at position $\mathbf{x}$ as a function of the spatial ($\nabla I \equiv (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial z})$) and temporal image gradients $\partial I/\partial t$ as follows:

$$\nabla I(\mathbf{x}, t) \cdot \mathbf{v}(\mathbf{x}, t) + \partial I(\mathbf{x}, t)/\partial t = 0. \quad (1)$$

In this study, we choose to solve equation (1) by assuming constant velocity in a small region c around $\mathbf{x}$, in which N gradients are sampled, as proposed in [10]:

$$\min_{\mathbf{v}} \sum_{\mathbf{x} \in c} w(\mathbf{x}) \left[\nabla I(\mathbf{x}) \cdot \mathbf{v} + \partial I(\mathbf{x})/\partial t\right]^2, \quad (2)$$

where w represents a local pixel weight. For a translation-only description of $\mathbf{v}$, this can be solved via least-squares, by differentiating equation (2) with respect to each translation component $\mathbf{v} = (v_x, v_y, v_z)$, and equating the result to zero.

2) Global tracker: To estimate the global cardiac motion, we use a statistical model of cardiac motion, embedded in the general optical flow equation [9]. In essence, instead of assuming constant velocity in a small region, the statistical model is used to guide the motion estimation, ensuring global results which are continuous across the whole image C and all time frames T. The motion model is obtained by applying Principal Component Analysis (PCA) on frame-to-frame affine transforms $\mathbf{a}$ throughout the whole cardiac cycle, extracted from a set of training contours. This results in a linear approximation of the affine parameters into an average transform $\bar{\mathbf{a}}$ and eigenvariations $\mathbf{\Phi}$, controlled by the PCA model parameters $\mathbf{p}$:

$$\mathbf{a} \approx \bar{\mathbf{a}} + \mathbf{\Phi}\mathbf{p}. \quad (3)$$

The statistical model is embedded in the optical flow equation by substituting the velocity term with the PCA model, resulting in a modified equation:

$$\min_{\mathbf{p}} \sum_{t \in T} \sum_{\mathbf{x} \in C} w(\mathbf{x}, t) \left[\nabla I(\mathbf{x}, t) \cdot \mathbf{v}(\mathbf{a}(\mathbf{p})) + \partial I(\mathbf{x}, t)/\partial t)\right]^2. \quad (4)$$

This is then solved via least-squares, by differentiation with respect to $\mathbf{p}$.

3) Combining global and local trackers: These trackers are combined by concatenating the estimated spatial transforms of the global T_G and local T_L trackers frame-by-frame. Starting at end-diastole (ED), the contour position $\mathbf{x}^t$ at frame t is:

$$\mathbf{x}^t = T_L^{t-1} \circ T_G^{t-1} \circ \ldots \circ T_L^{ED} \circ T_G^{ED}(\mathbf{x}^{ED}). \quad (5)$$

In this way, a local refinement is applied after the global motion estimation. This method generates an initial, reasonably accurate estimation of the cardiac motion. To improve on this estimation, it makes sense to apply the local refinement only in areas where the cardiac wall is clearly visible; in artifact-obscured areas, a local refinement would degrade the global tracking result. This motivates us to develop a method which separates the visible and artifact-obscured cardiac wall.

B. Separating visible cardiac wall from artifacts

1) Categorizing temporal image profiles: To separate visible from artifact-obscured cardiac wall, we categorize the temporal image intensity information into 'stable' (stationary to slowly changing), 'wandering' (moderately changing), or 'lost' (fast changing) components, inspired by the work of Jepson et al. [11]. The image intensity I_t at time t is modeled by the mixture probability density function (pdf), which is a sum of the stable (p_s), wandering (p_w), and lost (p_l) components:

$$p(I_t) = m_s p_s + m_w p_w + m_l p_l. \quad (6)$$

The stable component is modeled by a Gaussian pdf p_s:

$$p_s(I_t|\mu_s, \sigma_s^2) = e^{-(I_t-\mu_s)^2/2\sigma_s^2}/(\sigma_s\sqrt{2\pi}), \quad (7)$$

where μ_s and σ_s^2 are the mean and variance of the stable component. The wandering component is drawn from a temporally varying Gaussian pdf p_w, defined as:

$$p_w(I_t|\mu_{w,t}, \sigma_{w,t}^2) = e^{-(I_t-\mu_{w,t})^2/2\sigma_{w,t}^2}/(\sigma_{w,t}\sqrt{2\pi}) \quad (8)$$

The lost component is drawn from a uniform pdf, defined by the maximum and minimum intensity values over all frames:

$$p_l(I_t) = 1/(I_{max} - I_{min}). \quad (9)$$

2) Fast EM approach: The unknown model parameters (i.e. the mixing probabilities (m_s, m_w, m_l) and the mean and variance of the wandering component $(\mu_{w,t}, \sigma_{w,t}^2)$) are estimated with a fast and efficient EM algorithm. This version of the EM algorithm generates linear updates for the parameters per time frame, by considering image intensities in past frames exponentially less important [11]. Two steps are computed for each new image frame: an expectation step, in which the ownership probabilities are estimated using previous estimates of the unknown parameters, and a maximization step, in which the unknown parameters are updated using the ownership probabilities. The ownership probabilities, which can be viewed as the probability that the current observation belongs to a certain component, given the current estimates of the model parameters, are estimated as:

$$o_{i,t}(I_t) = m_{i,t} p_i / p \quad \text{for} \quad i \in \{s, w, l\}. \quad (10)$$

This makes $o_{s,t} + o_{w,t} + o_{l,t} = 1$ at each time step. The mixing probabilities are updated using a constant α:

$$\hat{m}_{i,t} = \alpha o_{i,t}(I_t) + (1-\alpha)\hat{m}_{i,t-1} \quad \text{for} \quad i \in \{s, w, l\}. \quad (11)$$

The wandering component is updated with:

$$\begin{aligned} \mu_{w,t} &= \hat{M}_{1,t}/\hat{M}_{0,t} \quad \text{and} \\ \sigma_{w,t} &= \hat{M}_{2,t}/\hat{M}_{0,t} - \mu_{w,t}^2, \\ \text{where} \quad \hat{M}_{j,t} &= \alpha I_t^j o_{w,t}(I_t) + (1-\alpha)\hat{M}_{j,t-1} \end{aligned} \quad (12)$$

with $\hat{M}_{j,t}$ denoting the jth-order, ownership weighed data moment.

3) Stable component in different frames-of-reference: We apply the categorization of temporal image profiles in a stationary frame-of-reference and a dynamic frame-of-reference. First, the EM ownerships are calculated in a stationary frame-of-reference, which is simply a voxel-wise analysis over time (Fig. 1A). The stable component's ownership o_s will be high for the areas where the myocardium is obscured and low for areas where it is clearly visible. Next, we obtain an initial estimate of the cardiac motion (section II-A) and apply the EM calculation in this dynamic, cardiac-motion-based frame-of-reference (Fig. 1B). Suppose that this cardiac motion estimate is reasonably accurate. o_s will now be low for the obscured areas and high for visible cardiac wall. This motivates us to define a weight which has high values for the visible areas and low values for obscured cardiac wall:

$$w = \frac{1}{2}(o_{s,CF} - o_{s,SF} + 1) \tag{13}$$

where $o_{s,CF}$ is the stable ownership in the dynamic frame-of-reference and $o_{s,SF}$ is the stable ownership in the stationary frame-of-reference. The weight is defined so that it ranges from $[0, 1]$ (since $o_s \in [0, 1]$).

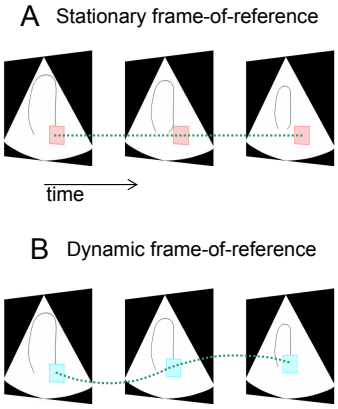

Fig. 1. Pixel-wise application of expectation-maximization algorithm on temporal intensity profiles, in (A) the stationary frame-of-reference and (B) the dynamic frame-of-reference (i.e. based on the initial cardiac motion estimation).

C. Weighed motion estimation

The derived weights w are combined with the motion estimation (section II-A) in two ways. First, the motion estimation is recalculated with w inserted into equations (2) and (4). Second, the local refinement is weighed, before concatenation with the global tracker:

$$T_{L,weighed}(\mathbf{x}) = w(\mathbf{x})T_L(\mathbf{x}). \tag{14}$$

In this way, the motion estimation relies more on the local tracker if the cardiac wall is clearly visible, and more on the global tracker if the cardiac wall is obscured.

D. Experimental details

1) Data description: The proposed tracking method is tested on clinically available 3D echocardiographic sequences, acquired with a Philips Sonos 7500 system (Philips Medical Systems, Best, The Netherlands), equipped with a X4 matrix array transducer. Typical spatial dimensions were 160 $\times$ 144 $\times$ 208 voxels with 1mm $\times$ 1mm $\times$ 0.7mm resolution.

2) Contour delineation: Left ventricular contours were needed for training the statistical motion model, as initialization in ED, and as the ground truth in the validation. Full-cycle endocardial borders were drawn with a previously developed semi-automated method, based on pattern matching and dynamic programming [12]. For the evaluation, the statistical motion model was built in a leave-one-out fashion.

3) Experiments: We compared the tracking accuracy of the weighed motion estimation (section II-C) with the initial motion estimation used to obtain the dynamic frame-of-reference (section II-A). Statistical testing was performed using the paired t-test.

III. Results

An example of tracking can be seen in Fig. 2, showing the merit of the weighed motion estimation in the drop-out areas in the anterior wall. Examples of images with artifacts and the weights are displayed in Fig. 3. The low weights correspond with regions with typical artifacts, which obscure the cardiac wall.

Fig. 4 shows the point-to-surface and volume errors across the cardiac cycle. Much improvement in tracking accuracy can be seen using the weighed method. Bull's eye plots of the local point-to-surface errors are displayed in Fig. 5. The improvement in areas with typical shadowing and near-field artifacts is clearly noticeable. Average errors are summarized in Table I. The proposed weighed estimation outperforms the initial estimation statistically significantly.

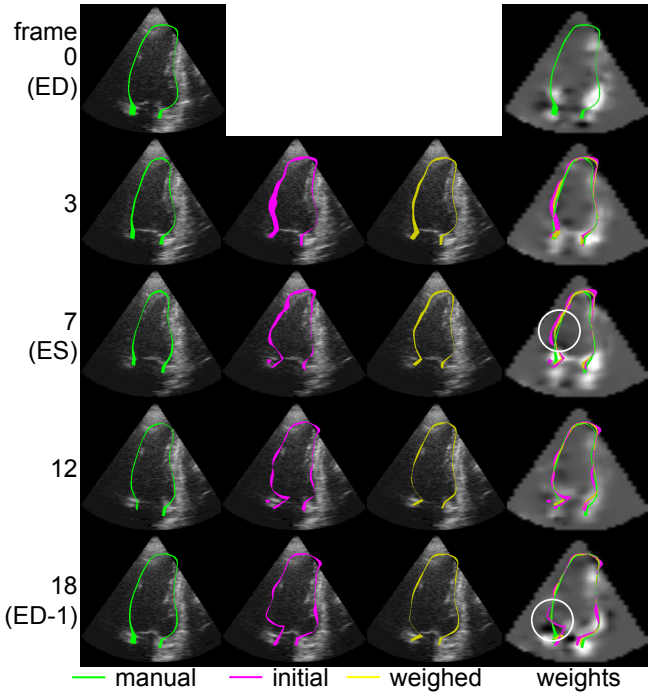

Fig. 2. Example of tracking results in two-chamber view. Notice the improved tracking in the drop-out areas, which have lower weights as seen right column.

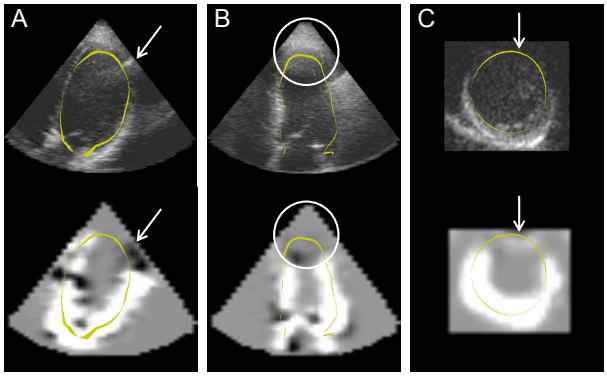

Fig. 3. Examples of images with (A) side lobe artifact, (B) large near-field, and (C) shadowing artifacts and the corresponding weights. Black denotes low weights, white denotes high weights.

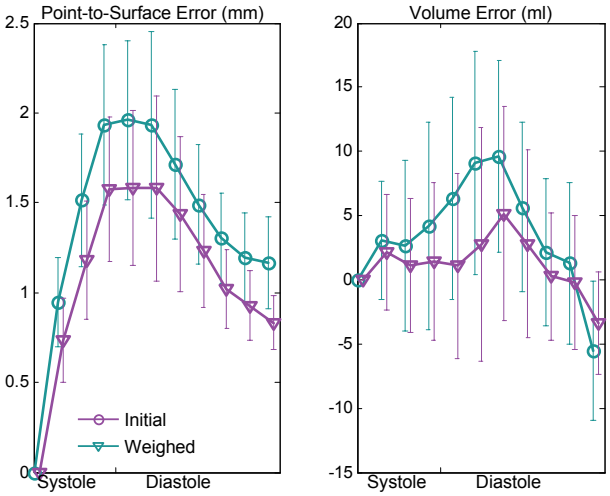

Fig. 4. Results of initial and weighed tracking method, showing improved point-to-surface and volume accuracy.

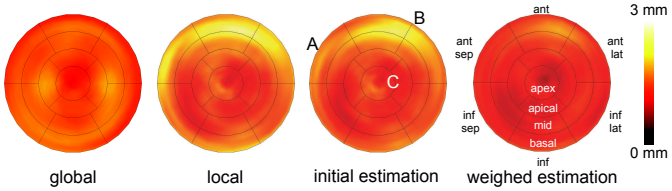

Fig. 5. Bull's eye plots of the point-to-surface errors, showing improved tracking accuracy using the weighed method, especially of (A) the fast moving aortic valve (which is hard to track using the local motion estimation), (B) the anterior region with common shadowing and drop-out, and (C) apical region with near-field artifacts.

TABLE I

AVERAGE ERRORS IN 35 IMAGE SEQUENCES. * DENOTES STATISTICALLY SIGNIFICANTLY BETTER THAN THE INITIAL ESTIMATION ($p < 0.05$).

	Point-to-surface (mm)	Absolute volume (ml)	Absolute ejection fraction (%)
Initial estimation	1.49±0.52	6.7±5.8	6.0±3.9
Weighed estimation	1.19±0.47*	5.2±4.5*	3.9±2.9*

IV. CONCLUSION

In this study, a method was proposed to distinguish between obscured and clearly visible myocardium, for accurate tracking in echocardiograms. The method categorized temporal intensity profiles in a stationary frame-of-reference and dynamic frame-of-reference. The weights generated by the method were shown to improve cardiac motion estimation, resulting in lower bias and smaller ranges in point and volume errors. The improvement is especially noticeable in areas with common ultrasound artifacts. Although the method was applied in 3D sequences in this study, the method can be used in 2D as well. Based on these promising results, we think that this method shows great potential in improving quantitative analysis in 3D echocardiograms.

ACKNOWLEDGMENTS

We thank Gerard van Burken for the data support provided. Dr. Marco Voormolen, dr. Marcel Geleijnse, dr. Boudewijn Krenning, dr. Attila Nemes, dr. Osama Soliman, and dr. Folkert ten Cate, are gratefully acknowledged for acquiring the images. This research is supported by the Dutch Technology Foundation STW (grant 06666), applied science division of NWO and the Technology Program of the Ministry of Economic Affairs, the Netherlands.

REFERENCES

[1] H. Feigenbaum, W. F. Armstrong, and T. Ryan, *Feigenbaum's Echocardiography*, 6th ed. Philadelphia: Lippincott Williams & Wilkins, 2005.

[2] F. A. Duck, "Nonlinear acoustics in diagnostic ultrasound," *Ultrasound Med. Biol.*, vol. 28, no. 1, pp. 1–18, 2002.

[3] Y. Takeuchi, "An investigation of a spread energy method for medical ultrasound systems. part one: theory and investigation," *Ultrasonics*, vol. 17, pp. 175–182, 1979.

[4] A. Nemes, M. L. Geleijnse, B. J. Krenning, O. I. I. Soliman, A. M. Anwar, W. B. Vletter, and F. J. ten Cate, "Usefulness of ultrasound contrast agent to improve image quality during real-time three-dimensional stress echocardiography," *Am. J. Cardiol.*, vol. 99, no. 2, pp. 275–278, 2007.

[5] M. A. Duarte, J. C. Machado, and W. C. A. Pereira, "A method to identify acoustic reverberation in multilayered homogeneous media," *Ultrasonics*, vol. 41, pp. 683–698, 2003.

[6] N. E. Bylund, M. Andersson, and H. Knutsson, "Interactive 3D filter design for ultrasound artifact reduction," *Proc IEEE Int. Conf. Image Process.*, vol. 3, pp. 728–731, 2005.

[7] T. Hozumi, K. Yoshida, Y. Abe, R. Kanda, T. Akasaka, T. Takagi, T. Yagi, Y. Ogata, and J. Yoshikawa, "Visualization of clear echocardiographic images with near field noise reduction technique: Experimental study and clinical experience," *J. Am. Soc. Echocardiogr.*, vol. 11, pp. 660–667, 1998.

[8] G. Zwirn and S. Akselrod, "Stationary clutter rejection in echocardiography," *Ultrasound Med. Biol.*, vol. 32, no. 1, pp. 43–52, 2006.

[9] K. Y. E. Leung, M. G. Danilouchkine, M. Van Stralen, N. De Jong, A. F. W. Van der Steen, and J. G. Bosch, "Tracking left ventricular borders in 3d echocardiographic sequences using motion-guided optical flow," *SPIE Med. Imaging*, vol. 7259, p. 72590W, 2009.

[10] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. DARPA Image Understanding Workshop*, pp. 121–130, 1981.

[11] A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi, "Robust online appearance models for visual tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1296–1311, 2003.

[12] M. van Stralen, J. G. Bosch, M. M. Voormolen, G. van Burken, B. J. Krenning, R. M. van Geuns, C. T. Lance, N. de Jong, and J. H. Reiber, "Left ventricular volume estimation in cardiac three-dimensional ultrasound: A semiautomatic border detection approach," *Acad. Radiol.*, vol. 12, no. 10, pp. 1241–1249, 2005.