Robust Segmentation of Intima–Media Borders With Different Morphologies and Dynamics During the Cardiac Cycle

Shen Zhao, Zhifan Gao, Heye Zhang, Yaoqin Xie, Jianwen Luo, Dhanjoo Ghista, Zhanghong Wei, Xiaojun Bi, Huahua Xiong, Chenchu Xu, and Shuo Li

***Abstract*—Segmentation of carotid intima-media (IM) borders from ultrasound sequences is challenging because of unknown image noise and varying IM border morphologies and/or dynamics. In this paper, we have developed a state-space framework to sequentially segment the carotid IM borders in each image throughout the cardiac cycle. In this framework, an H_∞ filter is used to solve the state-space equations, and a grayscale-derivative constraint snake is used to provide accurate measurements for the H_∞ filter. We have evaluated the performance of our approach by comparing our segmentation results to the manually traced contours of ultrasound image sequences of three synthetic models and 156 real subjects from four medical centers. The results show that our method has a small segmentation error (lumen intima, LI: 53 $\pm$ 67 μm; media-adventitia, MA: 57 $\pm$ 63 μm) for synthetic and real sequences of different image characteristics, and also agrees well with the manual segmentation (LI: bias = 1.44 μm; MA: bias = −3.38 μm). Our approach can robustly segment the carotid ultrasound sequences with various IM border morphologies, dynamics, and unknown image noise. These results indicate the potential of our framework to segment IM borders for clinical diagnosis.**

***Index Terms*—Carotid artery, H_∞ filter, IM borders, snake algorithm, ultrasound image sequences.**

Manuscript received August 9, 2017; revised October 13, 2017; accepted November 16, 2017. Date of publication November 21, 2017; date of current version August 31, 2018. This work was supported in part by the National Key Research and Develop Program of China under Grant 2016YFC1301700 and Grant 2016YFC0105102, in part by the Science Technology and Innovation Committee of Shenzhen for Research Projects under Grant JCYJ20151030151431727 and Grant JCYJ20170306090501763, in part by the Shenzhen Key Technical Research Project under Grant JSGG20160229203812944, and in part by the Beijing Center for Mathematics and Information Interdisciplinary Sciences. *(Corresponding authors: Heye Zhang; Yaoqin Xie.)*

S. Zhao and Y. Xie are with the Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: z-s-06@163.com; yq.xie@siat.ac.cn).

Z. Gao and H. Zhang are with the Institute of Advanced Computing and Digital Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: zf.gao@siat.ac.cn; heye.zhang@gmail.com).

J. Luo is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China (e-mail: luo_jianwen@tsinghua.edu.cn).

D. Ghista is with the University 2020 Foundation, MA, USA (e-mail: d.ghista@gmail.com).

Z. Wei is with the Shenzhen People's Hospital, Shenzhen, China (e-mail: weizhhsz@126.com).

X. Bi is with the Tongji Hospital, Wuhan 430030, China (e-mail: bixiaojun185@aliyun.com).

H. Xiong is with the Second Hospital of Shenzhen, Shenzhen 518055, China (e-mail: dennis8710@163.com).

C. Xu is with the Anhui University, Hefei, China (e-mail: 443779952@qq.com).

S. Li is with the University of Western Ontario, London, ON N6A 3K7, Canada (e-mail: slishuo@gmail.com).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JBHI.2017.2776246

I. Introduction

THE development of atherosclerosis can lead to fatal cardiovascular diseases such as heart attack or failure, which is still the major cause of mortality throughout the world [1]. A number of clinical studies have shown that the carotid intima-media thickness (IMT) could be used as a risk indicator to monitor the status of atherosclerosis [2]–[4]. As a non-invasive and low-cost technique, ultrasound imaging has been frequently applied to examine carotid IMT [5]. Furthermore, different segmentation algorithms have been developed to detect the IMT from ultrasound images [6]–[8]. Recently, a state-of-the-art multi-ethnic study reported that the changes of the IMT during the cardiac cycle are related to pulse pressure during systole, which might be better associated with cardiovascular risk [9]. As a result, the reliability and reproducibility of IMT segmentation during the cardiac cycle has attracted great attention [3]. However, it remains challenging to acquire reliable and reproducible segmentation of the intima-media border throughout the complete cardiac cycle because of the uncertainties in ultrasound data [10], [11]. The uncertainties in ultrasound data introduced by different ultrasound systems, scanners, and physicians impose great difficulties in robustly tracking the change of the IMT during the cardiac cycle. Therefore, there is great demand to develop a robust segmentation framework that can quantify the change of the IMT during the cardiac cycle from different ultrasound image sequences.

The main task of carotid ultrasound image segmentation is to extract the borders of the intima-media (IM) complex, namely, the lumen intima(LI) and media-adventitia(MA) borders on the far wall of the artery [12]. Comprehensive reviews of IMT segmentation algorithms for ultrasound images can be found in [10], [11]. Most of these works focus on segmenting images/sequences without atheromatous plaques [6], [7], [13]–

[16], while a promising integrated system is specially designed to segment plaques in ultrasound sequences in [17]. From the aspect of methodology, these works can be grouped into edge detectors [7], [13], dynamic programming [14], [15], [18], snakes [6], [16], [19]–[22], and neural networks [23]. Among these methods, the snake algorithm, which evolves the contours under the action of external and internal forces, is a popular method to detect image features (lines and edges) in ultrasound images. The deformable snake model yields a smoothed curve [24], which is beneficial for LI/MA segmentation [16]. Closed snakes are used to segment IM borders with plaques in [17], which is different from the open snakes commonly used to segment IM borders without plaques.

However, the uncertainties in ultrasound image sequences still hinder the snake algorithm from segmenting the IM borders correctly. The image sequences are obtained using different ultrasound systems and scanners, which may yield images with different image quality and characteristics. For example, in some ultrasound sequences, the LI borders can be affected by speckle noise in the lumens, while in others, the MA borders can be influenced by the ambiguous appearance of the adventitia. The noisy lumen can generate erroneous edges, while the blurred adventitia can hide the correct borders. Since the snake algorithm uses image gradients to evolve the contour, the noise and/or ambiguous borders can hinder segmentation of the LI/MA borders by producing erroneous image gradients or weakening the correct image features. Another problem is that the snakes are sensitive to the initial contours. In each frame of the sequence, the snake should be initialized in the vicinity of the border of interest; otherwise, it cannot be attracted to the correct boundary [19]. Thus, for the segmentation of an ultrasound sequence, initialization might be problematic if the temporal consistency is not taken into consideration.

The state-space approach, however, is appropriate to segment ultrasound sequences based on temporal consistency. The consecutive frames in an ultrasound sequence can be initialized using the results of the previous frame [25] in state-space approaches. The Kalman filter is used to solve the state-space equations for the optimal estimation of IM border positions by minimizing the covariance of the system and measurement noise in the state space [26]. The Kalman filter works well when the system and measurement noise can be assumed to be Gaussian distributed with known statistics. However, ultrasound image sequences are often affected by unknown and uncertain speckle noise. The wide variation of ultrasound systems and scanning instruments makes it difficult to obtain the statistical features of this noise, and it is not likely that the errors caused by this noise are Gaussian distributed. The determination of parameters in the state-space equations can be a problem because of these uncertainties. Additionally, if the movement of the IM borders in consecutive frames is large, the initialization of the snake positions in the latter frame by simply using the results of the previous frame could still cause problems. Therefore, a need exists for the development, implementation, and evaluation of an approach that is more robust to different types of noises, IM border morphologies and dynamics for segmentation of ultrasound sequences during cardiac cycles.

In this paper, we have proposed a H_∞ grayscale-derivative constraint snake by integrating the H_∞ filter and the grayscale-derivative constraint snake algorithm into an adaptive framework. The contributions of our work in the context of IM segmentation are: 1) segmentation of the IM border during the cardiac cycle with the help of the H_∞ filter, which can generate the optimal estimation for the IM border positions. The worst-case estimation error is minimized, and the requirement that the noise is Gaussian distributed is eliminated; 2) our H_∞ grayscale-derivative constraint snake algorithm can adapt to the image grayscale, curvature, and movement in consecutive frames. The parameters do not need to be adjusted manually according to the image characteristics of different sequences in different centers. The algorithm is accurate and robust to uncertain noise.

Our proposed H_∞ grayscale-derivative constraint snake segmentation method is evaluated using 156 ultrasound sequences from four medical centers. First, our algorithm is trained on 40 ultrasound sequences obtained from three medical centers. Then, the performance of our method is tested on the other 116 sequences from all four medical centers without any manual adjustments of the parameters. In particular, both the training set and the testing set include sequences with variant morphologies and/or dynamics (slanted, noisy, blurred, large movement between consecutive images). Sequences with atheromatous plaques, which are excluded in many classic and state-of-the-art studies [6], [7], [13]–[16], are also included in our evaluation. This means that sequences with and without plaques are segmented without changing the method or the parameters. The extensive validation of our algorithm using sequences of different morphologies and dynamics can also help to ensure that our model is not over-fitted. The segmented results are compared with the manual tracing method as well as 4 other methods that are used in recent works for similar applications: snake, Kalman snake [20], dynamic programming [15], and level set method using Chan-Vese energy functional [27].

II. Methodology

A. Pre-Processing

Before implementing the H_∞ grayscale-derivative constraint snake algorithm, the ultrasound sequences are processed as follows: (1) Normalization. The grayscale of each frame in the ultrasound sequence is normalized into [0, 255]. This eliminates the influence of inhomogeneity of intensity distributions across frames and/or sequences introduced by different gain settings, ultrasound systems, and scanners [19]. (2) Image filtering. A Gaussian low-pass filter is applied in every frame. This is used for noise reduction and attenuation of the high intensity noisy pixels. (3) ROI definition. A user-defined region of interest is selected with size 360 pixels $\times$ 260 pixels, except for a few sequences whose IM borders are too poorly visible.

B. The State-Space Framework

The vertical positions of each column in the image are regarded as state-space variables, and their dynamics can be described by the state-space equations ([26, p. 128]). The

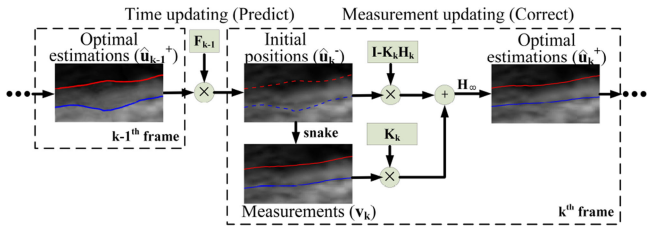

Fig. 1. Flowcharts of our method. In the time updating process, $\hat{\mathbf{u}}_k^-$ is predicted using $\hat{\mathbf{u}}_{k-1}^+$. In the measurement updating process, the measurement $\mathbf{v_k}$ is first obtained using the grayscale-derivative constraint snake; then, the optimal IM border positions $\hat{\mathbf{u}}_k^+$ are determined from $\hat{\mathbf{u}}_k^-$ and $\mathbf{v_k}$ by H_∞ filter.

state-space equations are solved using the H_∞ method, which consists of a time updating step and a measurement updating step, as illustrated in Fig. 1.

1) Time Updating Step: In the time updating step, the initial snake contour (*a priori* state estimate) $\hat{\mathbf{u}}_\mathbf{k}^-$ and its error covariance $\mathbf{P_k}^-$ are calculated using (1).

$$\begin{aligned}\hat{\mathbf{u}}_k^- &= \mathbf{F_{k-1}}\hat{\mathbf{u}}_{k-1}^+ \\ \mathbf{P_k}^- &= \mathbf{F_{k-1}}\mathbf{P_{k-1}}^+\mathbf{F_{k-1}}^T + \mathbf{Q_{k-1}}\end{aligned} \tag{1}$$

where $\mathbf{P_{k-1}}^+$ represents the *a posteriori* error covariance matrix of the previous frame, and $\mathbf{Q_{k-1}}$ is the covariance matrix of the process noise $\mathbf{q_{k-1}}$.

2) Measurement Updating Step: In the measurement updating step, the optimal (or *a posteriori*) state estimate $\hat{\mathbf{u}}_k^+$ is computed using *a priori* state $\hat{\mathbf{u}}_k^-$ and the measurement of the current frame $\mathbf{v_k}$ by minimizing the cost function J:

$$J = \frac{\sum_{k=0}^{N-1}\left|\left|\tilde{\mathbf{u}}_k - \hat{\mathbf{u}}_k^+\right|\right|^2}{\left|\left|\tilde{\mathbf{u}}_0 - \hat{\mathbf{u}}_0^+\right|\right|_{P_0^{-1}}^2 + \sum_{k=0}^{N-1}(||\mathbf{q_k}||_{Q_k^{-1}}^2 + ||\mathbf{r_k}||_{R_k^{-1}}^2)} \tag{2}$$

where $\hat{\mathbf{u}}_k^+$ is the *a posteriori* estimate of the current state $\mathbf{u_k}$, $\hat{\mathbf{u}}_0^+$ is the estimate of the initial state $\mathbf{u_0}$, $\mathbf{P_0}$ is the covariance of the initial estimate, $\mathbf{Q_k}$, and $\mathbf{R_k}$ are the covariance matrices of the noise terms $\mathbf{q_k}$ and $\mathbf{r_k}$, respectively. This minimization is implemented by setting a user-specified upper bound θ for the cost function ($J < \frac{1}{\theta}$). The partial derivative of J with respect to $\hat{\mathbf{u}}_k^+$ is obtained; then, $\hat{\mathbf{u}}_k^+$ is calculated by finding the minimum point of J. The H_∞ filter gain $\mathbf{K_k}$ and *a posteriori* error covariance matrix $\mathbf{P_k}^+$ are then calculated:

$$\begin{aligned}\hat{\mathbf{u}}_k^+ &= \hat{\mathbf{u}}_k^- + \mathbf{K_k}(\mathbf{v_k} - \mathbf{H_k}\hat{\mathbf{u}}_k^-) \\ \mathbf{K_k} &= \mathbf{P_k}^+\mathbf{H_k}^T\mathbf{R_k}^{-1} \\ \mathbf{P_k}^+ &= \mathbf{P_k}^-(\mathbf{I} - \theta\mathbf{L_k}\mathbf{P_k} + \mathbf{H_k}^T\mathbf{R_k}^{-1}\mathbf{H_k}\mathbf{P_k})^{-1}\end{aligned} \tag{3}$$

It can be seen from (3) that the H_∞ estimation is less reliant on modelling (namely, it is robust to modelling error) because the H_∞ filter gives more attention to the measurements as $\mathbf{K_k}$ increases. However, a brute increment of $\mathbf{K_k}$ will also result in poor estimation. The H_∞ filter provides the optimal way to increase $\mathbf{K_k}$ because the estimation error has an upper bound. Thus, the H_∞ filter is more robust to error in the system and noise models in the state-space equations, which is required for IM segmentation.

C. The Acquisition of Measurements

Since the robustness of the H_∞ filter to error in the system and the noise model originates from its larger weight on measurement $\mathbf{v_k}$, precise measurements are very important. Therefore, the snake algorithm is modified using grayscale and derivative constraints to acquire accurate measurements of the IM borders in each frame in the sequence. Before discussing the grayscale-derivative constraint snake, we first provide a brief overview of the classical snake.

1) The Classical Snake Algorithm and Its Limitations: A snake model is a contour geometrically represented by the spatial coordinates $\mathbf{v}(s) = [x(s), y(s)]$ $(x, y) \in \Re^2$ of an image, where $s \in [0, 1]$ is the parametric domain. By minimizing the energy function (4), the snake dynamically adapts itself to fit the shape of a target contour.

$$\begin{aligned}E_{snake} &= \int_0^1 [\mathbf{G_{int}}(\mathbf{v}(s)) + \mathbf{G_{ext}}(\mathbf{v}(s))]\,ds \\ \mathbf{G_{int}}(\mathbf{v}(s)) &= \alpha(s)\left|\mathbf{v}'(s)\right|^2 + \beta(s)\left|\mathbf{v}''(s)\right|^2 \\ \mathbf{G_{ext}}(\mathbf{v}(s)) &= w_l\mathbf{G_l} + w_e\mathbf{G_e} + w_t\mathbf{G_t}\end{aligned} \tag{4}$$

where $\mathbf{G_{ext}}(\mathbf{v}(s))$ and $\mathbf{G_{int}}(\mathbf{v}(s))$, respectively, denote the external and internal energy fields. They are both matrices of the same size as the image $\mathbf{I}$. The two terms in $\mathbf{G_{int}}(\mathbf{v}(s))$ represent the elastic energy and bending energy, and the three terms in $\mathbf{G_{ext}}(\mathbf{v}(s))$ represent the energies of lines, edges, and terminations, respectively (w_l, w_e and w_t represent the weight factors of these energies). A more detailed calculation of these energy terms can be found in [24]. To solve $\mathbf{v}(s)$, (4) is discretized, and the gradient descent method is used to minimize E_{snake}. Let $\frac{\partial u}{\partial t} = \frac{\partial E_{snake}}{\partial u} = 0$, the snake can be moved iteratively towards the local minima of $\mathbf{G_{ext}}$ (the edges of the images) according to the following equation:

$$\begin{aligned}F_{i,t-1} = {} & \frac{u_{i,t} - u_{i,t-1}}{\delta t} + \frac{\beta}{\delta s^4}u_{i-2,t-1} \\ & - \left(\frac{\alpha}{\delta s^2} + \frac{4\beta}{\delta s^4}\right)u_{i-1,t-1} + \left(\frac{2\alpha}{\delta s^2} + \frac{6\beta}{\delta s^4}\right)u_{i,t-1} \\ & - \left(\frac{2\alpha}{\delta s^2} + \frac{4\beta}{\delta s^4}\right)u_{i+1,t-1} + \frac{\beta}{\delta s^4}u_{i+2,t-1}\end{aligned} \tag{5}$$

where $u_{i,t}$ represents u_i at time t, δt and δs are the time step length and the spatial sampling distance, respectively, and $F_{i,t-1}$ is the value of $\frac{\partial \mathbf{G_{ext}}}{\partial u}$ (which can be regarded as the external force field driving the snake) at the ith point at time $t-1$. By iteratively solving (5), the snake is driven to the desired location, and one frame is segmented. To prevent the point on the snake from gathering together, a limit is added to restrain the displacement of the snake in the radial (y) direction [15].

For an ultrasound sequence with reasonable image quality, the snake algorithm is able to segment the IM borders correctly using a set of two open-contour snakes [17], as shown in Fig. 2(a). However, in images with noisy lumen and/or blurred adventitia, the segmentation may fail because the correct borders vanish and erroneous borders appear (Fig. 2(b)). Moreover, in ultrasound sequences, the IM borders are periodically

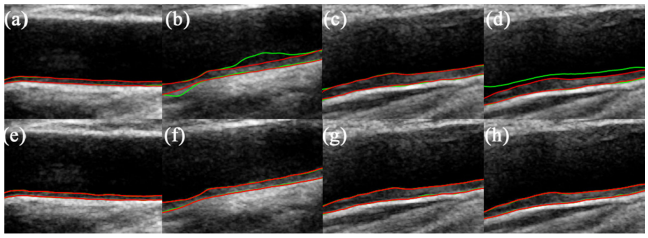

Fig. 2. Segmentation results of the traditional snake algorithm. The green lines are the segmentation results, and the red lines are the ground truth. (a): Correctly segmented borders in images with reasonable quality; (b) Segmentation failure due to noisy lumen and blurred adventitia; (c) and (d): Although one frame is correctly segmented (c), the next frame is not (d) because of large movement between the two frames; (e)–(h): The segmentation using our methods corresponding to (a)–(d). It is no longer an issue that the sequences are affected by noise, blurring, and large movements.

moving up during diastole and moving down during systole. If the IM borders move downwards rapidly (the downward movement is particularly significant at the end of the T peak of the ECG signal in some sequences), the traditional snake method may lose track of the IM borders, and it cannot correctly segment until the borders move back upwards to a relatively close location. For example, the segmentation is correct in Fig. 2(c), but the IM borders move down rapidly in the successive frames in Fig. 2(d). Resultantly, the segmentation fails in Fig. 2(d) as well as in several successive frames until the IM borders move up at diastole.

Thus, although the snake algorithm has existed for two decades, a robust method to address noisy/blurred images and large movement between sequences is needed in the context of current applications. Therefore, grayscale and derivative constraint energies are added to the snake algorithm.

2) Adaptive Grayscale-Derivative Constraint Snake for Consecutive Image Segmentation: Before setting up the force equations for the grayscale-derivative snake, it is essential to establish the concept of the upper and lower neighboring zones of one point in the snake contour. For an arbitrary point $(x,\ y)$ in the contour, the upper neighboring zone is the rectangle region characterized by the lower-middle coordinates, the width and the height, which are $(x,\ y)$, 3, and 5 respectively. Similarly, the lower neighboring zone is a rectangle region whose upper-middle coordinates are $(x,\ y)$. The width and height of the lower neighboring zone are the same as those of the upper zone, as shown in Fig. 3(a) (the neighboring zones should be thin, tall boxes, but they are shortened and widened in Fig. 3(a) for clarity).

Although the segmentation failures shown in Fig. 2 can be caused by different factors, their results are the same: the grayscale and/or derivative near the snake contour is abnormal. Thus, a grayscale constraint energy term, $\mathbf{G_{gc}}$, as well as a derivative constraint energy term, $\mathbf{G_{dc}}$, are added to the external energy to correct the snake positions:

$$\mathbf{G_{ext}}(\mathbf{v}(s)) = w_l\mathbf{G_l} + w_e\mathbf{G_e} + w_t\mathbf{G_t} + w_g\mathbf{G_{gc}} + w_d\mathbf{G_{dc}} \quad (6)$$

The first three terms in $\mathbf{G_{ext}}$ are the same as in the traditional snake. The added $\mathbf{G_{gc}}$/$\mathbf{G_{dc}}$ term is decided by the grayscale/derivative difference between the current frame and the first frame in the sequence. The dimensions of $\mathbf{G_{gc}}$ and $\mathbf{G_{dc}}$ are the same as those of $\mathbf{G_{ext}}$ (and image $\mathbf{I}$). w_g and w_d are the weighting factors of the two constraint terms. For each point (point i) in the snake contour in the current frame (frame k), the mean grayscales of its upper/lower neighboring zones (denoted as $E_{i,k}^{upper}$ and $E_{i,k}^{lower}$) are calculated and compared with the corresponding neighboring zone in the first frame (denoted as $E_{i,1}^{upper}$ and $E_{i,1}^{lower}$). If the difference between these mean grayscales are less than some thresholds, $\mathbf{G_{gc}}$ is $\mathbf{0}$. Similarly, for each point in the snake contour, the derivatives of y with respect to x in the current frame and in the first frame are calculated and denoted as $y'_{i,k}$ and $y'_{i,1}$. If the relationship between $y'_{i,k}$ and $y'_{i,1}$ is acceptable, $\mathbf{G_{dc}}$ is $\mathbf{0}$. If both $\mathbf{G_{gc}}$ and $\mathbf{G_{dc}}$ are $\mathbf{0}$, the grayscale-derivative constraint snake retrieves to the traditional snake. This corresponds to the normal sequences without noise, blurring, and large movements. However, when the snake contour deviates from the correct position and latches onto erroneous edges, the grayscale of its upper and lower neighboring zones could be abnormal. In this case, $\mathbf{G_{gc}}$ and/or $\mathbf{G_{dc}}$ are used to help the snake to return to normal.

(I) Grayscale constraint energy. If the mean grayscale of the neighboring zones in the current frame is too small (namely, $E_{i,1}^{upper} - E_{i,k}^{upper} > tol_m^{upper}$ or $E_{i,1}^{lower} - E_{i,k}^{lower} > tol_m^{lower}$), the neighboring zone is considered to be darker than normal (or its position is probably higher than the correct position). Then $\mathbf{G_{gc}}$ is set to be positive in the upper neighboring zone. The base value of $\mathbf{G_{gc}}$ is set to 200, which means that if w_g=1, the value 200 is added to $\mathbf{G_{ext}}$ for each point in the 3 * 5 area of the upper zone. Conversely, when $E_{i,k}^{upper} - E_{i,1}^{upper} > tol_m^{upper}$ or $E_{i,k}^{lower} - E_{i,1}^{lower} > tol_m^{lower}$, $\mathbf{G_{gc}}$ is set to be positive in the lower neighboring zone. Since the external force field $\frac{\partial \mathbf{G_{ext}}}{\partial u}$ always drives the snake to positions where $\mathbf{G_{ext}}$ is small, the positive $\mathbf{G_{gc}}$ will generate a grayscale constraint force that pushes the snake out of the erroneous edges. The weighting factor w_g is set to be positively correlated to the grayscale difference ($|E_{i,1}^{upper} - E_{i,k}^{upper}|$ or $|E_{i,1}^{lower} - E_{i,k}^{lower}|$), which means that the snake will be pushed back by a larger force if the grayscale is more abnormal.

For example, as shown in Fig. 3(b2), the LI snake (red dashed line) may get stuck in regions of local noise without $\mathbf{G_{gc}}$. However, by comparing the grayscale with that in the benchmark (Fig. 3(a) and (b1)), it is found that the grayscale of the lower zone ($E_{i,k}^{lower}$) is smaller than $E_{i,k}^{lower} - tol_m^{lower}$; thus, $\mathbf{G_{gc}}$ is set to be positive (+200) in the upper neighboring zone of the LI snake (Fig. 3(c2)) and the modified $\mathbf{G_{ext}}$ is calculated (Fig. 3(c3)). The external force is then calculated (Fig. 3(d)), and the LI snake is pushed downwards to the correct position (red solid line in Fig. 3(e)). This also applies to the situations in Fig. 2(b) and (d). The LI snake is pushed down because the grayscale of the lower zone is small in both figures.

To make our method adaptive to sequences with different brightness, tol_m^{upper} and tol_m^{lower} are chosen to be $0.2E_{i,1}^{upper}$ and $0.2E_{i,1}^{lower}$ in most cases. However, there are two exceptions:

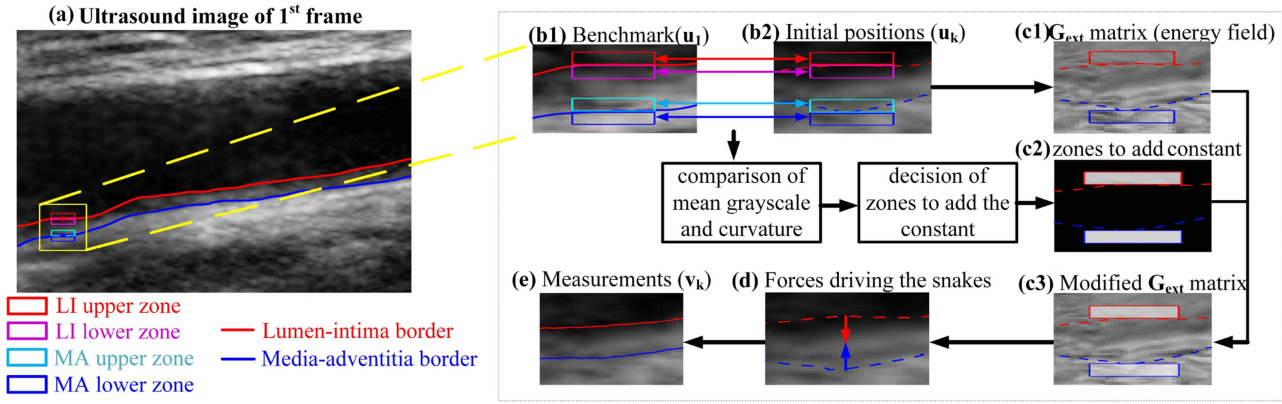

Fig. 3. Illustration of the adaptive snake algorithm. (a): The 1st image is a benchmark. The red, magenta, cyan, and blue rectangles represent the LI upper, LI lower, MA upper, and MA lower neighboring zones of the ith point on the LI/MA border; (b1) and (b2): The mean grayscale and curvature in the neighboring zones in the kth image (b2) are compared with those in the 1st image (b1); (c1)–(c3): The $\mathbf{G_{ext}}$ matrix is calculated using the image features(c1). The zones where the grayscale-derivative constraint energy should be added in the $\mathbf{G_{ext}}$ matrix are identified (c2) and the $\mathbf{G_{ext}}$ matrix is modified (c3); (d): The snakes are next driven to local minima of $\mathbf{G_{ext}}$ (regions that are darker in the $\mathbf{G_{ext}}$ matrix in (c3)); (e): The results of the adaptive snake ($\mathbf{v_k}$) is thus calculated. This approach can automatically adjust the snake parameters according to the image characteristics.

TABLE I
DIFFERENT ULTRASOUND SYSTEMS, TRANSDUCER FREQUENCY, FRAME TIME, NUMBER OF SEQUENCES OBTAINED FROM EACH CENTER (n), AND NUMBER OF SEQUENCES USED FOR TRAINING ($n_{training}$) FOR EACH CENTER

Center number	Canter name	Ultrasound system	Transducer frequency (MHz)	Frame time (ms)	n	$n_{training}$
1	The Second Hospital of Shenzhen, China	iU22 (Philips Ultrasound, Bothell, WA, USA)	3∼9	25.55∼29.56	62	20
2	Tongji Hospital, Wuhan, China	Vivid E9(GE Vingmed Ultrasound, Fairfield, USA)	4	40.03∼40.95	40	5
3	Tsinghua University Hospital, Beijing, China	iU22 (Philips Ultrasound, Bothell, WA, USA)	3∼9	9.78∼17.55	31	15
4	Renmin Hospital, Shenzhen, China	iU Elite (Philips Ultrasound, Bothell, WA, USA)	3∼9 or 5∼12	26.17∼36.70	20	0

(1)tol_m^{upper} is chosen to be max(10, $E_{i,1}^{upper}$) to push LI up because the mean grayscale in the upper zone of LI is generally less than 10 in sequences without lumen noise [19]. This means that if the lumen is not noisy ($E_{i,1}^{upper} < 10$), the LI snake is pushed up if $E_{i,k}^{upper} - E_{i,1}^{upper} > 10$ (this could confirme that $E_{i,k}^{upper}$ is too large in unnoisy sequences); if the lumen is noisy ($E_{i,1}^{upper} > 10$), the LI snake is pushed up if $E_{i,k}^{upper} > 2E_{i,1}^{upper}$. This ensures that the LI snake is not pushed up too easily because of a large $E_{i,k}^{upper}$ caused by lumen noise. (2) If the lumen is not noisy ($E_{i,1}^{upper} < 10$), the information of $E_{i,k}^{upper}$ is not used for pushing down the LI snake because both $E_{i,1}^{upper}$ and $E_{i,k}^{upper}$ of the LI snake are relatively small, so $E_{i,k}^{upper}$ might be less than $0.8 \times E_{i,1}^{upper}$ even if there is no need to push down the LI snake. Instead, the LI snake is pushed down as long as $E_{i,k}^{lower} < 10$ because the region under the LI snake (inside the IM complex) should never be too dark.

Another reason for using the grayscale constraint energy term is to reduce the influence of LI and MA on each other. Consider a case in which the LI border is not distinct and the MA border is. The LI snake might be attracted to the strong MA border. Conversely, the MA snake could be attracted to the LI border if the LI border is strong and the MA border is weak. However, when the additional grayscale constraint energy is used, the LI snake can automatically avoid the MA borders because the neighboring zones are too bright. Additionally, the MA snake would not be attracted to the LI borders because those locations are dark. In this way, the grayscale constraint energy reduces or eliminates the influence of the mutual effect of the two snakes.

(II) Derivative constraint energy. Similar to the grayscale constraint energy, if the shape of the snake at the ith point in the kth frame is not in agreement with that in the 1st frame, $\mathbf{G_{dc}}$ is used to correct the snake and to maintain a relatively fixed shape of the snake throughout the sequence. In particular, the IM borders are smooth curves, so bumps on the snake contour in the kth frame (Fig. 2(b)) are considered to be not acceptable if there are no bumps in the 1st frame. These bumps are identified by finding local maxima(minima). If y is positive (negative) before and negative (positive) after the ith point, this point is a the local maximum(minimum). Then, the average absolute values of the derivatives of each point in the neighborhood of the ith point (denoted as $\bar{y}'_{i,k}$) are calculated and compared with the mean absolute derivatives of the corresponding region in the first frame (denoted as $\bar{y}'_{i,1}$). If $\bar{y}'_{i,k} - \bar{y}'_{i,1} > tol_c$ ($0.5\bar{y}'_{i,1}$ in our study), it will be confirmed that there does exist an abnormal bump in the snake. $\mathbf{G_{dc}}$ is then set to be positive in the upper

(lower) neighboring zones for local minima (maxima). Similar to $\mathbf{G_{gc}}$, the base value of $\mathbf{G_{dc}}$ is also set to 200, and the weighting factor w_d is set to be positively correlated with the curvature difference ($|\bar{y}'_{i,k} - \bar{y}'_{i,1}|$). With these derivative constraints added to the $\mathbf{G_{ext}}$ matrix, the snakes are pushed to eliminate the abnormal bumps. Actually, the snakes should also be corrected if $\bar{y}'_{i,1} - \bar{y}'_{i,k} > tol_c$, but this never occurs in our experiments. So this case is temporarily not considered in the current version of our algorithm.

For example, as shown in Fig. 3(b2), the MA snake (blue dashed line) may fail to lock onto the correct (but blurred) MA border without additional constraint energy. Unfortunately, the grayscale constraint $\mathbf{G_{gc}}$ cannot be added because neither the upper neighboring zone nor the lower neighboring zone of the MA snake is bright enough in this frame (neither $E^{upper}_{i,k} - E^{upper}_{i,1} > tol^{upper}_m$ nor $E^{lower}_{i,k} - E^{lower}_{i,1} > tol^{lower}_m$ is satisfied). With the help of $\mathbf{G_{dc}}$, a local maximum (the lowest point in the figure) is found. Then $\bar{y}'_{i,k}$ is calculated and compared with $\bar{y}'_{i,1}$. Since the contour near the ith point is nearly horizontal in the 1st frame and slanted in the kth frame, $\bar{y}'_{i,k} - \bar{y}'_{i,1} > tol_c$ is satisfied and $\mathbf{G_{dc}}$ is set to be positive (+200) in the lower neighboring zone of the MA snake (Fig. 3(c2)). Thus, the MA snake is pushed upwards to the correct position (Fig. 3(e)).

From the above examples, we observe that the use of an appropriate grayscale constraint energy and derivative constraint energy complement each other and yield reasonable segmentation results even if the sequence has a noisy lumen, blurred adventitia, and large movements between consecutive frames. With these accurate measurement, the H_∞ filter is able to robustly obtain the optimal estimation by minimizing the worst-case error.

III. Experiments and Results

A. Population Study, Data Collection, and Parameter Training

A total of 3 synthetic models and 153 subjects were enrolled in our study. The synthetic sequences were generated from the toolbox for *in silico* evaluation of motion estimators for the arterial wall, developed by the BioSim Laboratory of the National Technical University of Athens [29]. For the real subjects, four ultrasound physicians from four different medical centers collected carotid ultrasound sequences using ultrasound systems and linear array transducers. The ultrasound system name, frequency of the transducers, frame time, number of sequences (n) obtained from each center, and number of sequences used for training the optimal parameter settings ($n_{training}$) are presented in Table I. The ultrasound sequences were recorded through 3∼6 consecutive full cardiac cycles with a recognizable R peak in the ECG wave. During collection, the subjects were examined in the supine position, with their heads turned 45° away from the examined side. All imaging data were saved in DICOM format for off-line analysis. We ran our program on Matlab 2015B on a desktop computer with an Intel(R) Core(TM) i5-5200U CPU (2.20 GHz) and 8 GB RAM. First, the sequences from the four centers were divided into a training set and a testing set. The

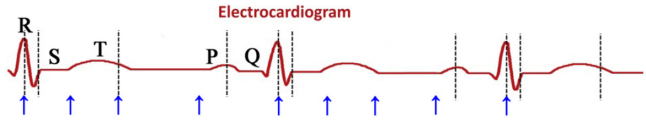

Fig. 4. The frames that are manually delineated in a sequence (marked by the blue arrows).

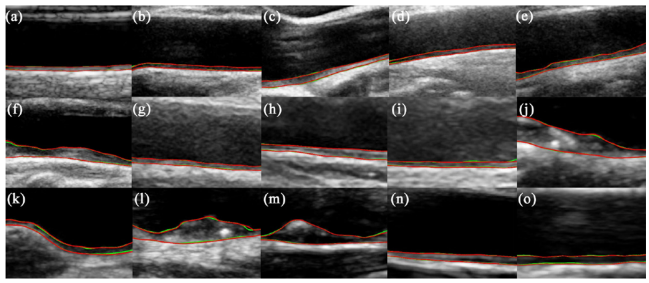

Fig. 5. The segmentation results of some sample images in the sequences. Though the image characteristics of these sequences are significantly different, the segmentation results of all the sequences are close to the manual delineations, which means our method performs well in different data sets.

parameters of our H-inf grayscale-derivative constraint snake were determined using the training set. Then, the testing set was used to evaluate our proposed method. All the sequences from Renmin Hospital were used for testing to evaluate the robustness of our method to sequences with unknown morphologies and/or dynamics.

After the training process, the optimal parameter settings were empirically determined and kept unchanged throughout the testing process. The kernel size and standard deviation of the Gaussian filter used for pre-processing were 5×5 and 1px; the user-specified performance bound of H_∞ filter $\theta = 0.1$; the weight factors of elastic energy $\alpha(s) = 4$ and bending energy $\beta(s)=2$; the time step length $\delta t = 0.0025$; the number of iterations for solving u_i using (5) was 20; the spatial sampling distance $\delta s = 1$; the weight factors for $\mathbf{G_l}$, $\mathbf{G_e}$, and $\mathbf{G_t}$ were 0.9, 10, and 2, respectively; the weight factor of the grayscale constraint energy term w_g was $(1 + 0.05 \times |E^{upper}_{i,1} - E^{upper}_{i,k}|)$ or $(1 + 0.05 \times |E^{lower}_{i,1} - E^{lower}_{i,k}|)$; and the weight factor of the derivative constraint term w_d was $10 \times |\bar{y}'_{i,k} - \bar{y}'_{i,1}|$.

Three ultrasound physicians (P1∼P3) were asked to manually delineate the IM borders. The manual delineation was performed at the end of the R and T peaks, as well as for the frames in the middle of the two peaks, as shown by the blue arrows in Fig. 4. The reason for this decision was that the R peak represents the end of diastole (corresponding to the largest IMT), and the end of the T peak represents the end of systole (corresponding to the smallest IMT). A total of 1404 (156 sequences $\times$ 9 = 1404) ultrasound frames in the sequences were delineated. P1, P2 and P3 manually segmented all the abovementioned frames blinded to the other results. The differences between these segmentations were used to assess the inter-operator variability. Then, P1 segmented the IM borders again one month after his first segmentation without seeing his

TABLE II
SEGMENTATION ERROR DISPLAYED AS $MAE \pm$ STANDARD DEVIATION

Center number	Pixel size	Border name	Error in pixel	Error in μm	Inter obs error 1	Inter obs error 2	Intra obs error
0	–	LI	0.58 ± 0.69	–	–	–	–
MA	0.57 ± 0.74	–	–	–	–		
1	52	LI	0.62 ± 0.72	32 ± 37	39 ± 41	33 ± 44	27 ± 38
MA	0.67 ± 0.79	35 ± 41	40 ± 39	44 ± 50	31 ± 41		
2	111	LI	0.62 ± 0.80	68 ± 89	57 ± 80	64 ± 81	39 ± 52
MA	0.64 ± 0.68	71 ± 75	64 ± 77	81 ± 79	59 ± 71		
3	53	LI	1.25 ± 1.37	67 ± 73	77 ± 86	71 ± 76	51 ± 52
MA	1.29 ± 1.13	69 ± 60	82 ± 68	77 ± 70	61 ± 46		
4	84	LI	0.81 ± 0.82	69 ± 70	73 ± 53	84 ± 60	58 ± 72
MA	0.95 ± 0.93	81 ± 79	103 ± 61	105 ± 72	63 ± 100		
average	72	LI	0.77 ± 0.94	53 ± 67	56 ± 66	51 ± 61	43 ± 58
MA	0.82 ± 0.89	57 ± 63	63 ± 63	65 ± 64	52 ± 67		

The center numbers correspond to those in Table I, and 0 indicates the synthetic models. Except the fourth row (error in pixels), all the errors are in μm. The denotation inter obs error 1 means the segmentation error between P1 and P2, inter obs error 2 means the segmentation error between P1 and P3, and intra obs error means the segmentation error between the two delineations of P1. In the 3rd ~ 8th rows, the number in the upper cell of each column represents the LI segmentation error, and that in the lower cell represents the MA segmentation error.

initial segmentation. This information was used to assess intra-operator variability.

B. Examples of IM Segmentation in Ultrasound Sequences With Variant Morphologies and Dynamics

The IM borders were successfully segmented for all the sequences enrolled, although the morphologies and dynamics of the sequences varied widely. The segmentations of the figures that were wrongly segmented by traditional snake method are presented in Fig. 2(f) and (h). Moreover, representative segmentations of the synthetic models and real images are illustrated in Fig. 5. The red lines represent the reference contours segmented by the physicians (the mean value of the four independent delineations by P1∼P3), and the green lines represent the results of our segmentation method. Fig. 5 represents image sequences with different morphologies from the 4 centers. Fig. 5(a) is the synthetic sequence, and its IM borders are clear and horizontal; Fig. 5(b)–(f) are selected from five sequences from center 1, Fig. 5(b) is characterized by relatively clear IM borders, Fig. 5(c) is characterized by noise just above the LI border, Fig. 5(d) is characterized by severe noise in the lumen, Fig. 5(e) is characterized by a blurred MA border and noise just above the LI border, and Fig. 5(f) is characterized by plague and a rather thick IMT. Fig. 5(g)–(i) are from center 2, Fig. 5(g) is affected by severe lumen noise, Fig. 5(h) has a significant horizontal edge inside the IM complex, Fig. 5(i) has a blurred flocculent structure on the left half of the LI border. Fig. 5(j)–(m) are from center 3, Fig. 5(j) is characterized a plaque with some very bright points inside (which form erroneous edges), and the blurred flocculent structure also exists on the right half of the LI border; Fig. 5(k) has twisting IM borders (horizontal on the left side, bending in the middle, and horizontal on the right side) with a blurred MA border, and Fig. 5(l) has a significant plague whose shape is different from that in Fig. 5(f), a bright spot also occurs inside the plague, Fig. 5(m) also has a significant plague, but several places inside it are very dark. Fig. 5(n) and (o) are from center 4, Fig. 5(n) has a similar morphology to those from center 2, but it suffers from large radial movement of the IM borders in consecutive frames, Fig. 5(o) is characterized by a discontinuous LI border and a dark IM complex. It should also be noted that some of these IM borders are horizontal, some are slanted, while others have plagues with arc-shaped structures.

The accuracy of our method is quantitatively evaluated using the mean absolute average error (MAE) [10] in both pixels and μm, as shown in Table II ($MAE \pm$ standard deviation). All the unitary segmentation errors of each point are aggregated into a single vector; then, MAE and standard deviation are computed. It should be noted that the pixel size of the synthetic models is not given in [29]; thus, the error of the synthetic models is only given in pixels. Moreover, the average error in pixels is the weighted average (the weight factor of each center is the number of sequences) of all the sequences, while the average error in μm is the weighted average of the real sequences (the synthetic models are excluded). For the 156 sequences involved, the overall accuracy of our method is approximately less than 1 pixel, and the absolute error in μm is ∼55 μm. This error is very close to the inter- and intra-analyst variability.

The computation time is proportional to the number of points in the snake and inversely proportional to the number of iterations used to solve u_i. When there are 360 pixels in the snake, the processing time is ∼0.24 s for a single frame.

The Bland-Altman plots in Fig. 6 show that the IM border positions estimated by our method are in agreement with the manual delineations. The plot consists of more than 50000 points from the 156 sequences. The different colors represent data from different centers (blue: synthetic, red: center 1, yellow: center 2, magenta: center 3, black: center 4, the data from the synthetic sequences are not shown in Fig. 6(c) and (d)). Fig. 6 shows that the signed bias between our approach and the ground truth is 1.44 μm (0.020 pixel) for LI and −3.38 μm (0.047 pixel) for MA. The 95 % confidence intervals (dark green lines) are 227 μm for LI and 239 μm for MA. Furthermore, the scattered points are linearly fitted, and the absolute values of the slopes of the fitted lines (black dashed lines) are less than 0.001 for both LI and MA. The average of our approach and the ground truth has little effect on their difference.

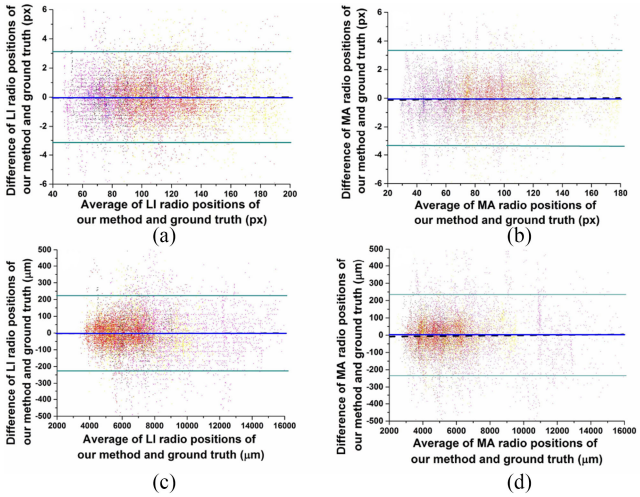

Fig. 6. Bland-Altman plot for segmentation ((a): LI, in pixels; (b): MA, in pixels; (c): LI, in μm; (d): LI, in μm). The blue lines represent the difference between our method and the ground truth. The intervals between the dark green lines are the 95% confidence intervals. The black dashed lines are the linearly fitted results.

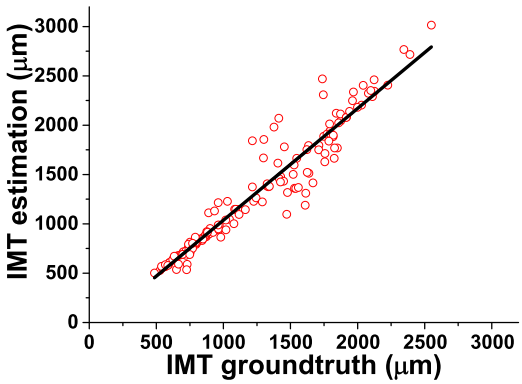

Fig. 7. Linear regression of the IMT obtained from our method and the manual delineations.

In addition to the radial positions, the IMT is also evaluated to further assess the results. Fig. 7 shows the results of a linear regression analysis of all the sequences. Each circle in Fig. 7 corresponds to the IMT value of one sequence. Again, the liner regression shows that the IMTs obtained using our proposed method are consistent with the segmentations performed by P1∼P3. The R^2 value of the linear regression is 0.92.

Our approach is compared with four other methods: (1) Kalman snake method; (2) snake method; (3) dynamic programming [15]; (4) level set method using Chan-Vese energy functional [27]. We attempted to re-implement these methods, performed the evaluations using the same sequences from the 4 medical centers, and calculated the mean absolute errors (MAE s). Fig. 8 compares the average MAE of these methods. The error of our method is generally smaller than that of the other methods for the sequences from the four centers. The average MAE using our method is 53 μm for LI and 57 μm for MA, whereas those using Kalman snake, snake, dynamic programming, and level set method are, respectively, 66 μm for LI and 68 μm for MA, 65 μm for LI and 68 μm for MA,

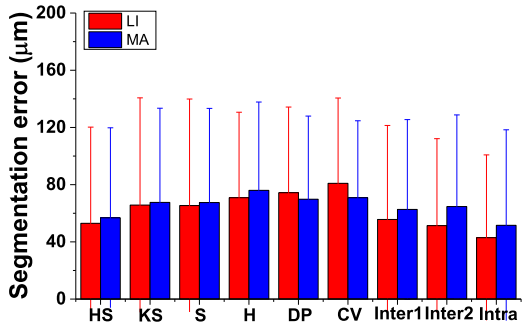

Fig. 8. Comparison of the MAE of our method with those of other methods (HS: H_∞ snake, our proposed method; KS: Kalman snake method; S: snake method; H: H_∞ filter used with snake (not modified by the grayscale-derivative constraints); DP: dynamic programming; CV: level set method using Chan-Vese energy functional; inter1: inter-observer variability of P1 and P2; inter2: inter-observer variability of P1 and P3; intra: intra-observer variability). The red/blue bars represent the LI/MA segmentation error.

74 μm for LI and 70 μm for MA, and 81 μm for LI and 71 μm for MA. Specifically, the p-values are 0.027 and 0.043 for, respectively, LI and MA if we use the mean error of each sequence as the samples to perform a t-test to compare the proposed method and the Kalman snake method. For other methods, the p-values are less than 0.05 (snake: 0.027 for LI and 0.04 for MA; Chan-Vese: 0.018 for LI and 0.026 for MA; Dynamic programming: 0.022 for LI and 0.038 for MA). These results shows that the superiority of our method is statistically significant.

To evaluate the sensitivity of our method to noise, six different types of noise: (I) Gaussian white noise, (II) Pepper-salt noise, (III) Uniform distribution noise, (IV) Poisson noise, (V) Rayleigh noise, and (VI) Speckle noise with different signal-to-noise ratios (SNR = 30 dB∼3 dB) generated by Matlab software are added to the ultrasound sequences. Because there are only 3 synthetic sequences, and the noise characteristics can not be exactly the same even for the same subject due to phase deviation caused by the operation of the physicians [30], some unknown noise may occur in ultrasound sequences, even if the image features of the IM complex are almost identical. To test the robustness of our method to this unknown noise, we also add these six types of noise to the real ultrasound sequences. These six types of noises are of different characteristics, such as additive and multiplicative, as well as zero mean and none-zero mean. Fig. 9 shows that the segmentation error of our approach remains small even if the signal-to-noise ratio becomes small. Overall, the largest effect of additional noise at the smallest SNR is 21 μm for LI and 22 μm for MA, as shown in Fig. 9(a) and (b).

To test the effect of noise in the initial contour, the initial contours are: (1) 1 pixel above the ground truth; (2) 1 pixel below the ground truth; (3) added with uniformly distributed noise between -1 and 1; (4) 2 pixels above the ground truth; (5) 2 pixels below the ground truth; and (6) added with uniformly distributed noise between -2 and 2. As shown in Fig. 10, the mean effect of the initial contour on the segmentation error is less than 1 pixel ($\approx$ 72 μm on average) when the error of initial

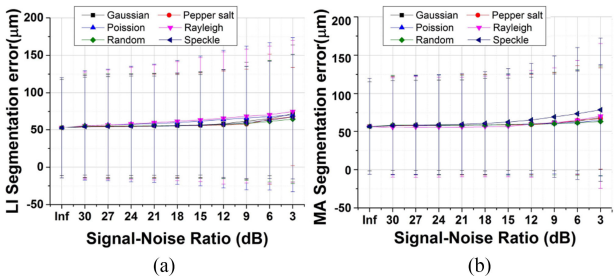

Fig. 9. Errors of our approach for segmentation at different noise levels. The denotation Inf represents the case without additional noise. Fig. 9(a) and (b) represent the average LI/MA segmentation error at different noise levels. Each plot in these figures corresponds to one type of noise.

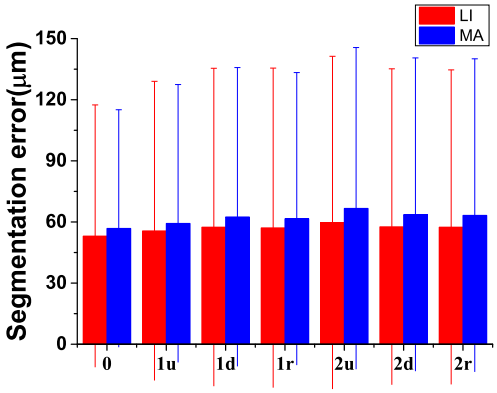

Fig. 10. The effect of initial contour error on segmentation error. The denotations 0, 1u, 1d, 1r, 2u, 2d, 2r respectively represents the initial contour is at the ground truth, 1 pixel above the ground truth, 1 pixel below the ground truth, added with a uniformly distributed noise between -1 pixel and 1 pixel, 2 pixels above the ground truth, 2 pixels below the ground truth, added with a uniformly distributed noise between -2 pixels and 2 pixels. The red/blue bars represent the LI/MA segmentation error.

TABLE III
SEGMENTATION ERROR MAE IN μm WHEN SEQUENCES FROM DIFFERENT CENTER IS USED FOR TESTING

experiment name	training centers	testing center	MAE in μm (LI)	MAE in μm (MA)
EXP1	1 2 3	4	53 ± 67	57 ± 63
EXP2	1 2 4	3	61 ± 89	62 ± 76
EXP3	1 3 4	2	56 ± 72	59 ± 68
EXP4	2 3 4	1	65 ± 89	66 ± 74

4 experiments are carried out using respectively sequences from 1,2,3 for training and 4 for testing, 1 2 4 for training and 3 for testing, 1 3 4 for training and 2 for testing, 2 3 4 for training and 1 for testing (named as EXP1∼4).

contour is less than 2 pixels, regardless of the type of noise added to the initial contour.

To ensure that our method is not over-fitted to the training data shown in Table I, we conducted a set of experiments using sequences from different centers as training sets. In each experiment, sequences from one center are left out for testing, and the sequences from the other three centers are used for training. The results are shown in Table III. The standard deviation of the 4 MAEs is small (LI: 5.79 μm, MA: 4.31 μm), which demonstrates that our method is not over-fitted to any of these data sets.

IV. DISCUSSIONS

1) The H_∞ State-Space Framework: The H_∞ filter is used to guarantee robustness to discrepancies in the system models. In the time updating process of the framework, the IM border positions in the current frame are initiallized with $\hat{\mathbf{u}}_{k-1}$. This can eliminate the effect of erroneous strong edges far from the correct IM border positions in consecutive frames. Similar work using $\hat{\mathbf{u}}_{k-1}$ as the *a priori* estimate can also be seen in many works on segmentation [14], [20] and/or tracking [20], [31], [32], where different segmentation/tracking methods (dynamic programming, Kalman snake, block matching, ultrasound tagging) are used. In the Kalman filter, the system model ($\mathbf{F_k}$, $\mathbf{H_k}$, $\mathbf{Q_k}$, and $\mathbf{R_k}$) is assumed to be precisely known, and the trace of the error covariance matrix $\mathrm{Tr}(\mathbf{P_k})$ is minimized as cost function J.

In the measurement updating process, the superiority of the H_∞ filter is due to its robustness to unknown errors in the system and noise model because it minimizes the worst-case error. This is exactly the case in sequence segmentation because the IM border positions are not the same in consecutive frames, and we have no *a priori* idea of the precise movement for each point. Thus, we have to assume that the state transition matrix $\mathbf{F_{k-1}}$ in (1) is $\mathbf{I}$, which unfortunately does not always represent the reality. Similar problems are encountered when determining the $\mathbf{Q}$ and $\mathbf{R}$ matrices. This means that there may be errors in the state-space model. Thus, we should give more attention to the measurements $\mathbf{v_k}$ by increasing $\mathbf{K_k}$. However, $\mathbf{v_k}$ may also be inaccurate; thus, blindly placing more emphasis on $\mathbf{v_k}$ may also result in poor results. One way to solve this problem is to obtain more precise $\mathbf{v_k}$ (which will be discussed next); another way is to find an optimal solution to minimize the largest error. This is done in the H_∞ filter by assuming that there is an adversary that maximizes J to degrade our state estimate by choosing $\mathbf{u_0}$, $\mathbf{q_k}$, and $\mathbf{r_k}$. Then, we minimize J (2) under the worst case caused by the adversary with respect to $\hat{\mathbf{u}}_k^+$ and $\mathbf{v_k}$. Thus, it is derived that (as shown in (3)) the H_∞ filter gain $\mathbf{K_k}$ is larger than the Kalman gain because the term $\theta \mathbf{L_k P_k}$ is subtracted. In this way, the largest error is restricted to a relatively small scale ($< 1/\theta$). This provides the optimal way to increase $\mathbf{K_k}$. In our experiments, Fig. 8 shows that the segmentation error is smallest when using H_∞ filter. Additionally, the first and third bars in Fig. 8 show that the H_∞ filter results in a decrease of segmentation error by $\sim$10 μm, which demonstrates the role of the H_∞ filter. Additionally, for sequences from the synthetic models and center 1 (which is the most fully trained center), the errors of the H_∞ filter and that of the Kalman filter are similar (Fig. 8(a)); however, for the sequences that are not fully trained, the superiority of the H_∞ filter to the Kalman filter can be more clearly demonstrated. (The segmentation errors in pixels using Kalman are: synthetic models: 0.60(LI) and 0.64(MA); center 1: 0.68(LI) and 0.73(MA); center 2: 0.74(LI) and 0.74(MA); center 3: 1.72(LI) and 1.58(MA); center 4: 0.93(LI) and 1.09(MA). These results can be compared with the results in the fourth row of Table II.)

2) The Grayscale-Derivative Constraint Snake: Under the H_∞ state-space framework, the grayscale-derivative constraint

snake algorithm is used to acquire accurate measurements. In previous work, [6], [33], [34] use traditional snakes or Williams and Shah snakes to segment the IM borders. However, traditional snakes may fail in the cases shown in Fig. 2(b) and (d). In [16], the snakes are modified by an attractive term to segment single ultrasound frames. The snake is pushed to the correct border by an additional constraint energy that is proportional to the distance of the snake from the edge if it is in image regions with a uniform intensity. This algorithm also maintains a constrained distance between the two snakes. This modification can solve the problem shown in Fig. 2(d); however, when there are strong erroneous edges, or when the shape of the LI/MA border is irregular (such as those in Fig. 5(d)–(f) and (j)–(l)), constraining the snakes only by attracting them to the edges may be not sufficient.

Except for the image derivatives, our grayscale-derivative constraint snake makes full use of the local grayscale and derivative information. By comparing the local grayscales of the neighboring zones of each point in the current frame with its benchmark, the grayscale constraint snake can determine whether the snake position is correct (if not, $\mathbf{G_{gc}}$ will be automatically used to modify the snake). This determination is not affected by blurred borders, so our method can segment sequences with a blurred flocculent structure near the LI border (as in some sequences in center 3). Moreover, the derivative constraints prefer a relatively fixed and smooth shape and prevent abrupt variation in radial position. The weighting factor w_d is relatively large and positively correlated with the curvature difference, so a strong force is formed to counteract the erroneous force caused by the bright or dark spots inside the IM complex (as in some sequences in center 3). Thus, our method produces reasonable results for sequences with large discrepancies in morphology, intensity, and dynamics (Figs. 2(e)–(h) and 5). The third and fourth bars in Fig. 8 indicate that the error using the H_∞ filter and snake without grayscale-derivative constraints is larger than that using grayscale-derivative constraint snake without the H_∞ filter. Thus, accurate measurements for the H_∞ filter is important in the state estimation. Nevertheless, we should not ignore the contribution made by the robustness of the H_∞ filter. Actually, the accuracy of the grayscale-derivative snake and the robustness of the H_∞ filter are two complementary strategies for accurate segmentation. The H_∞ filter gives more attention to the measurements, and the grayscale-derivative snake helps the measurements to be more precise.

The GVF snake [35] is much newer than the traditional snake. However, we find that the GVF snake does not statistically outperform the traditional snake in our application. This may be because the main advantage of the GVF snake is that it is able to attract snakes into boundary concavities from relatively farther positions in homogeneous regions [35]. However, in our application, there are almost no concavities in the IM borders. Also, the grayscale is not uniform in most regions of the ultrasound images because IM borders and non-Gaussian image noises can produce boundaries. Although the parameter μ can be tuned to address image noise, we find that if μ is large, the results can be poor (the condition of convergence may be violated); if μ is small, the role of the GVF force is not significant (the result is almost the same as that of the traditional snake). Thus, we designed our algorithm basing on the traditional snake to reduce the computational cost.

It is also interesting that the base value of $\mathbf{G_{gc}}$ (which is set to 200) need not be changed throughout our experiments. It works well with sequences with different morphologies and dynamics. This can be explained by the fact that a positive $\mathbf{G_{gc}}$ will generate a grayscale constraint force and change the force acting at each point of the snake ($F_{i,t-1}$ in (5), which drives the snake to the correct positions). For different sequences, $F_{i,t-1}$ differs for the same $\mathbf{G_{gc}}$. However, the movement of the snake is determined by iteratively solving (5), so the snake position is decided by the combined effects of the force $F_{i,t-1}$, time step length δt, and number of iterations. Thus, the difference in $F_{i,t-1}$ can be compensated by changing the number of iterations automatically (if $F_{i,t-1}$ is larger, fewer iterations are needed and the snake stops iteration automatically at the image boundaries; if $F_{i,t-1}$ is smaller, more iterations are needed, but the snake can still arrive at the minimum in 20 iterations). Therefore, when the force $F_{i,t-1}$ changes within a certain limit, the results of the snake do not change, so the same $\mathbf{G_{gc}}$ can be used for different sequences.

3) Evaluation of our Method: As shown in Table I, the manufacturers model name, the transducer frequency, and the frame time are not the same in these datasets. Thus, these datasets provide sequences with different characteristics and image quality: sequences from the synthetic models have the best image quality (Fig. 5(a)) with well-visible IM borders; sequences from center 1 are the most fully trained, because they have reasonable image quality and a large variety of morphologies, noise, blurring, intensity, and dynamics (Fig. 5(b)–(f)); sequences from center 2 are mostly horizontal with relatively clear MA borders, but the lumen is always noisy and the LI border is sometimes not clearly detailed. Additionally, large movements of IM borders are observed in consecutive frames in several sequences (Fig. 5(g)–(i)); sequences from center 3 are the most challenging because they have irregular, unexpected morphologies and poorly visible IM borders, and in several sequences, there are high-intensity spots inside the IM complex, while in some others, there are blurred flocculent structure on the LI border (Fig. 5(j)–(m)); sequences from center 4 are similar to those from center 2, but they are more challenging because the movements are larger. Moreover, the IM borders are sometimes discontinuous and not sufficiently clear (Fig. 5(n) and (o)). All 156 sequences with different characteristics are properly segmented without changing the parameters in our method.

The accuracy of the method is of the same order of magnitude as the inter-/intra-observer variabilities, as shown in Table II and the first bars in Fig. 8. The segmentation errors in pixels of centers 1 and 2 and the synthetic models are small. The segmentation error in pixels of center 3 is the largest, but the error is still close to the inter-/intra-observer variabilities. The error in pixels of center 4 is relatively large, probably because the grayscale and/or morphology of the IM complex are not the same in the 1st and kth frames in several sequences

(which is a limitation of our method that will be discussed next). However, the error increment is not significant (0.2 pixels for LI and 0.3 pixels for MA compared with centers 1 and 2).

Table III confirms that our method is not over-fitted to the training datasets. The mean absolute average errors for both LI and MA are all small ($\sim$60 μm or $\sim$0.83 pixel) regardless of which center is left out as the testing set. The standard deviation of the 4 MAE s is small (LI: 5.79 μm, MA: 4.31 μm). The variance of the model (measured by the variance of the four MAE s) is also small (LI: 33.52 μm^2, MA: 18.58 μm^2). Additionally, we perform a supplementary experiment using another 20, 5, and 15 sequences (which means none of these sequences are used for training in the original experiment mentioned in Table I and II) from center 1, 2 and 3 for training. The results of this experiment are MAE for LI: 59 $\pm$ 70 μm and MAE for MA: 60 $\pm$ 71 μm, which are very close to those in Table II. This shows that our model is stable when the training data set is perturbed. These results demonstrates that our method is not over-fitted to any of these data sets in our experiment. Also, the fact that our model trained using center 1, 2, and 3 performs well for center 4 (which is unknown to the model during the training process) proves that our model is able to accurately segment sequences from medical centers whose sequences are not encountered in the training process.

In comparison with other methods, Fig. 8 shows that the accuracy of our method is higher. Also, as reported in the review papers [3], [10], the accuracy of our method is comparable to the most novel methods. Notwithstanding, these methods should be compared with care because the evaluations are not the same in different papers [14]. On one hand, the sequences used in different studies are not the same. On the other hand, although we attempted to re-implement the methods used in other works [15], [27], the technical details may not be exactly the same. To genuinely compare different segmentation approaches, evaluation should be performed using the same data sets for all methods with the same evaluation metrics, and a larger-scale study is required [11], [14].

4) Limitations and Perspectives: Our work is intended for segmenting ultrasound sequences with different morphologies and/or dynamics. However, there are some limitations in our proposed method. First, although the H_∞ filter is robust to system error, it is still beneficial to the segmentation process if the system matrices ($\mathbf{F_k}$, $\mathbf{H_k}$, $\mathbf{Q_k}$, and $\mathbf{R_k}$) are precisely defined. Although the contours are successfully extracted using our method, these matrices are simply set to be $\mathbf{I}$, which does not always represent the reality. If a preliminary estimation of the IM border positions can be obtained using edge detectors, this estimation could be used to tune the system matrices [31] and accordingly contribute to improvement of the method.

Secondly, the parameters for the H_∞ filter and the grayscale-derivative constraint snake are selected empirically based on the training set. We tune the parameters to their optimal values by a grid search with some granularity. For example, for the threshold $E_{i,1}^{upper}$, the search is in $[0.1, 0.4]E_{i,1}^{upper}$ and the granularity is 0.1 $E_{i,1}^{upper}$. However, the optical threshold may be between the grids. The use of state-of-the-art methods, such as deep learning, to determine the optimal parameters could further improve the accuracy.

Third, our method maintains a relatively fixed shape of the IM borders, so it might be sub-optimal if the appearance of the IM complex varies greatly during the sequence. For example, if there exist well-visible interfaces in the first frame that become blurred (or even disappear) in later frames, the error of these frames may be large. Actually, this occurs in 5 sequences from center 4, which results in an increase in MAE of $\sim$0.1 pixel at this center.

V. Conclusion

In this paper, we have developed an accurate approach to segment the intima-media layer during the cardiac cycle. In our approach, a worst-case estimation error minimization H_∞ filter is used in combination with a grayscale-derivative constraint snake to address the noise, morphologies and dynamics of the ultrasound sequences. Our proposed approach is evaluated using 3 synthetic models and 156 sequences from 4 medical centers, and the results are compared with manually traced results as well as the results of 4 other methods: snake, Kalman snake, dynamic programming, and level set method using the Chan-Vese energy functional. The results adequately demonstrate the accuracy and robustness of our method, and its potential to segment IM borders for clinical diagnosis. Our work is implemented in a Matlab graphical user interface and will be publicly available shortly.

References

[1] J. A. Schaar *et al.*, "Terminology for high-risk and vulnerable coronary artery plaque," *Eur. Heart J.*, vol. 25, pp. 1077–1082, 2004.

[2] J. F. Polak, M. J. Pencina, K. M. Pencina, C. J. O'Donnell, P. A. Wolf, and R. B. D'Agostino, "Carotid-wall intima-media thickness and cardiovascular events," *New England J. Med.*, vol. 365, pp. 213–221, 2011.

[3] T. Z. Naqvi and M.-S. Lee, "Carotid intima-media thickness and plaque in cardiovascular risk assessment," *JACC Cardiovascular Imag.*, vol. 7, pp. 1025–1038, 2014.

[4] M. W. Lorenz, H. S. Markus, M. L. Bots, M. Rosvall, and M. Sitzer, "Prediction of clinical cardiovascular events with carotid intima-media thickness: A systematic review and meta-analysis," *Circulation*, vol. 115, pp. 459–467, 2007.

[5] H. Okimoto *et al.* "A novel method for evaluating human carotid artery elasticity: Possible detection of early stage atherosclerosis in subjects with type 2 diabetes," *Atherosclerosis*, vol. 196, pp. 391–397, 2008.

[6] D.-c. Cheng, A. Schmidt-Trucksass, K.-s. Cheng, and H. Burkhardt, "Using snakes to detect the intimal and adventitial layers of the common carotid artery wall in sonographic images," *Comput. Methods Progr. Biomed.*, vol. 67, pp. 27–37, 2002.

[7] F. Faita, V. Gemignani, E. Bianchini, C. Giannarelli, L. Ghiadoni, and M. Demi, "Real-time measurement system for evaluation of the carotid intima-media thickness with a robust edge operator," *J. Ultrasound Med.*, vol. 27, pp. 1353–1361, 2008.

[8] Y.-B. Lee, Y.-J. Choi, and M.-H. Kim, "Boundary detection in carotid ultrasound images using dynamic programming and a directional Haar-like filter," *Comput. Biol. Med.*, vol. 40, pp. 687–697, 2010.

[9] J. F. Polak *et al.*, "Changes in carotid intima-media thickness during the cardiac cycle the multi-ethnic study of atherosclerosis: The multi-ethnic study of atherosclerosis," *J. Amer. Heart Assoc.*, vol. 1, pp. 301–306, 2012.

[10] F. Molinari, G. Zeng, and J. S. Suri, "A state of the art review on intima-media thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound," *Comput. Methods Progr. Biomed.*, vol. 100, pp. 201–221, 2010.

[11] C. P. Loizou, "A review of ultrasound common carotid artery image and video segmentation techniques," *Med. Biol. Eng. Comput.*, vol. 52, pp. 1073–1093, 2014.
[12] P. J. Touboul *et al.*, "Mannheim carotid intima-media thickness and plaque consensus (2004–2006–2011). An update on behalf of the advisory board of the 3rd, 4th and 5th watching the risk symposia, at the 13th, 15th and 20th European Stroke Conferences, Mannheim, Germany, 2004, Brussels, Belgium, 2006, and Hamburg, Germany, 2011," *Cerebrovascular Diseases*, vol. 34, pp. 290–296, 2012.
[13] R. H. Selzer, W. J. Mack, P. L. Lee, H. Kwong-Fu, and H. N. Hodis, "Improved common carotid elasticity and intima-media thickness measurements from computer analysis of sequential ultrasound frames," *Atherosclerosis*, vol. 154, pp. 185–193, 2001.
[14] G. Zahnd *et al.*, "A fully-automatic method to segment the carotid artery layers in ultrasound imaging: Application to quantify the compression-decompression pattern of the intima-media complex during the cardiac cycle," *Ultrasound Med. Biol.*, vol. 43, pp. 239–257, 2017.
[15] Y. Zhou, X. Cheng, X. Xu, and E. Song, "Dynamic programming in parallel boundary detection with application to ultrasound intima-media segmentation," *Med. Image Anal.*, vol. 17, pp. 892–906, 2013.
[16] F. Molinari *et al.*, "Constrained snake vs. conventional snake for carotid ultrasound automated IMT measurements on multi-center data sets," *Ultrasonics*, vol. 52, pp. 949–961, 2012.
[17] C. P. Loizou, S. Petroudi, M. Pantziaris, A. N. Nicolaides, and C. S. Pattichis, "An integrated system for the segmentation of atherosclerotic carotid plaque ultrasound video," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 61, no. 1, pp. 86–101, Jan. 2014.
[18] Q. Liang, I. Wendelhag, J. Wikstrand, and T. Gustavsson, and A Multiscale, "Dynamic programming procedure for boundary detection in ultrasonic artery images," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 127–142, Feb. 2000.
[19] C. P. Loizou, C. S. Pattichis, and M. Pantziaris, "Snakes based segmentation of the common carotid artery intima media," *Med. Biol. Eng. Comput.*, vol. 45, pp. 35–49, 2007.
[20] S.-H. Lee and S. Lee, "Adaptive Kalman snake for semi-autonomous 3D vessel tracking," *Comput. Methods Progr. Biomed.*, vol. 122, pp. 56–75, 2015.
[21] S. Delsanto, F. Molinari, P. Giustetto, W. Liboni, S. Badalamenti, and J. S. Suri, "Characterization of a completely user-independent algorithm for carotid artery segmentation in 2-D ultrasound images," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 4, pp. 1265–1274, Aug. 2007.
[22] F. Molinari, G. Zeng, and J. S. Suri, "Intima-media Thickness: Setting a standard for a completely automated method of ultrasound measurement," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 57, no. 5, pp. 1112–1124, May 2010.
[23] R.-M. Menchn-Lara, M.-C. Bastida-Jumilla, J. Morales-Snchez, and J.-L. Sancho-Gmez, "Automatic detection of the intima-media thickness in ultrasound images of the common carotid artery using neural networks," *Med. Biol. Eng. Comput.*, vol. 52, pp. 169–181, 2013.
[24] J. Ivins and J. Porrill, "Everything you always wanted to know about snakes (but were afraid to ask)," *Artif. Intell.*, pp. 1–35, 1995.
[25] A. Gastounioti, S. Golemati, J. Stoitsis, and K. S. Nikita, "Kalman-filter-based block matching for arterial wall motion estimation from B-mode ultrasound," *Proc. IEEE Int. Conf. Imag. Syst. Tech.*, 2010, pp. 234–239.
[26] D. Simon, *Optimal State Estimation*, Hoboken, NJ, USA: Wiley, 2006.
[27] A. M. F. Santos, R. M. d. Santos, P. M. A. C. Castro, E. Azevedo, L. Sousa, and J. M. R. S. Tavares, "A novel automatic algorithm for the segmentation of the lumen of the carotid artery in ultrasound B-mode images," *Expert Syst. Appl.*, vol. 40, pp. 6570–6579, 2013.
[28] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.
[29] G. Aimilia, D. Sofia, K. Vasileios, and N. Konstantina, "A toolbox for in silico evaluation of motion estimators for the arterial wall," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014.
[30] K. Z. Abd-Elmoniem, A.-B. M. Youssef, and Y. M. Kadah, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 9, pp. 997–1014, Sep. 2002.
[31] Z. Gao *et al.*, "Robust estimation of carotid artery wall motion using the elasticity-based state-space approach," *Med. Image Anal.*, vol. 37, pp. 1–21, 2017.
[32] G. Zahnd, M. Orkisz, A. Srusclat, P. Moulin, and D. Vray, "Evaluation of a Kalman-based block matching method to assess the bi-dimensional motion of the carotid artery wall in B-mode ultrasound sequences," *Med. Image Anal.*, vol. 17, pp. 573–585, 2013.
[33] C. P. Loizou, M. Pantziaris, M. S. Pattichis, E. Kyriacou, and C. S. Pattichis, "Ultrasound image texture analysis of the intima and media layers of the common carotid artery and its correlation with age and gender," *Comput. Med. Imag. Graph.*, vol. 33, pp. 317–324, 2009.
[34] C. P. Loizou, C. S. Pattichis, A. N. Nicolaides, and M. Pantziaris, "Manual and automated media and intima thickness measurements of the common carotid artery," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 56, no. 5, pp. 983–994, May 2009.
[35] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.