Robust estimation of carotid artery wall motion using the elasticity-based state-space approach

Zhifan Gao\textsuperscript{a,b}, Huahua Xiong\textsuperscript{c}, Xin Liu\textsuperscript{a}, Heye Zhang\textsuperscript{a,*}, Dhanjoo Ghista\textsuperscript{d}, Wanqing Wu\textsuperscript{a}, Shuo Li\textsuperscript{e}

\textsuperscript{a} Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
\textsuperscript{b} Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, China
\textsuperscript{c} Department of Ultrasound, The Second People’s Hospital of Shenzhen, China
\textsuperscript{d} University 2020 Foundation, United States
\textsuperscript{e} University of Western Ontario, Canada

A R T I C L E   I N F O

Article history:
Received 1 August 2016
Revised 8 January 2017
Accepted 9 January 2017
Available online 10 January 2017

Keywords:
Carotid artery vessel wall motion
Atherosclerosis
Carotid ultrasound
State-space approach
H\textsubscript{s} filter

A B S T R A C T

The dynamics of the carotid artery wall has been recognized as a valuable indicator to evaluate the status of atherosclerotic disease in the preclinical stage. However, it is still a challenge to accurately measure this dynamics from ultrasound images. This paper aims at developing an elasticity-based state-space approach for accurately measuring the two-dimensional motion of the carotid artery wall from the ultrasound imaging sequences. In our approach, we have employed a linear elasticity model of the carotid artery wall, and converted it into the state space equation. Then, the two-dimensional motion of carotid artery wall is computed by solving this state-space approach using the H\textsubscript{s} filter and the block matching method. In addition, a parameter training strategy is proposed in this study for dealing with the parameter initialization problem. In our experiment, we have also developed an evaluation function to measure the tracking accuracy of the motion of the carotid artery wall by considering the influence of the sizes of the two blocks (acquired by our approach and the manual tracing) containing the same carotid wall tissue and their overlapping degree. Then, we have compared the performance of our approach with the manual traced results drawn by three medical physicians on 37 healthy subjects and 103 unhealthy subjects. The results have showed that our approach was highly correlated (Pearson’s correlation coefficient equals 0.9897 for the radial motion and 0.9536 for the longitudinal motion), and agreed well (width the 95% confidence interval is 89.62 μm for the radial motion and 387.26 μm for the longitudinal motion) with the manual tracing method. We also compared our approach to the three kinds of previous methods, including conventional block matching methods, Kalman-based block matching methods and the optical flow. Altogether, we have been able to successfully demonstrate the efficacy of our elasticity-model based state-space approach (EBS) for more accurate tracking of the 2-dimensional motion of the carotid artery wall, towards more effective assessment of the status of atherosclerotic disease in the preclinical stage.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Background

Atherosclerosis is a condition where the arterial wall becomes narrowed and hardened due to the accumulation of plaque inside the artery (Laurent et al., 2001; Helfand et al., 2009). Without proper treatment, atherosclerosis can trigger fatal cardiovascular events, such as myocardial infarction and stroke (Schaar et al., 2004). A number of studies have reported close relationship between the stiffening degree of carotid artery wall and such that cardiovascular events (van Sloten et al., 2014; Caviezel et al., 2014). Thus, the alteration of the biomechanical dynamics in the carotid artery can be used to evaluate the degree of atherosclerotic disease in the preclinical or subclinical stage (Nichols et al., 2011; Roger et al., 2012). As a convenient tool, the ultrasound imaging technique has been frequently applied to examine the biomechanical dynamics of the carotid artery (Mokhtari-Dizaji et al., 2006; Okimoto et al., 2008).

Previously, the radial motion of the carotid artery wall has been extensively extracted from the ultrasound images for evaluation of atherosclerosis (Stadler et al., 1997; Zhang et al., 2010; Nilsson et al., 2014), such as vasoconstriction and vasodilation.
Fig. 1. The process of BM method. The reference block is firstly compared with every candidate block extracted from the search region (red rectangle) in the current frame of the ultrasound sequence. Then, the best-matched block can be found by considering it as the candidate block (green rectangle) which is most similar with the reference block. The reference block in the next frame is considered equal to the best-matched block. The BM method for the block tracking is implemented by repeating the above process for every frame in the ultrasound sequence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Schwarzacher et al., 1992; lida et al., 2006) and arterial blood pressure (Bortel et al., 2001; Meinders and Hoeks, 2004; Vermeersch et al., 2008). Recently, the longitudinal motion of the carotid artery wall has gradually attracted attention because of its potential to sensitively reflect the status of atherosclerosis (Cinthiso et al., 2006; Zahnd et al., 2012). Several studies have demonstrated that the longitudinal motion can be recognized as a new marker for indicating the early stage of atherosclerosis (Mokhtari-Dizaji et al., 2006; Okimoto et al., 2008; Zahnd et al., 2012; Yli-Ollila et al., 2013; Zahnd et al., 2013). However, it is still very challenging to reliably and quantitatively acquire the longitudinal motion from the ultrasound image sequences because of the homogeneous appearance of the tissues on the carotid artery wall.

1.2. Previous works on tracking the motion of the carotid artery wall

In order to reliably measure the biomechanical performance of the carotid artery wall, many previous attempts have tried to recover the motion of the carotid artery wall from ultrasound image sequences. The block matching (BM) method is one popular approach that has been widely applied in the motion tracking of the carotid artery wall (Persson et al., 2002; Golemati et al., 2003; Cinthiso et al., 2005, 2006). In the BM method, the motion tracking of the target tissue is based on the template of the tissue represented by a block (called reference block). Inside the search region of each frame, many candidate blocks which have the same size as the reference block, are compared to the reference block by using a pre-specified similarity criterion. Then the location of the best-matched candidate block is considered as the location of the target tissue. The motion trajectory through the entire ultrasound sequence is generated by connecting the locations of the best-match blocks between the successive frames in the space-time coordinate system. The flow of the BM method is illustrated in Fig. 1.

However, the BM method has several existing difficulties in tracking the motion of the carotid artery wall from the ultrasound image sequences. The first difficulty comes from the homogeneity of the carotid artery wall tissue texture along the longitudinal direction in the ultrasound images (Zahnd et al., 2013). This homogeneity of tissue texture decreases the image contrast along the longitudinal direction and thus reduces the performance of the BM method in motion tracking, because the pre-specified similarity criterion, such as normalized cross-correlation and sum of absolute differences (Zahnd et al., 2010; Tat et al., 2015), are all generated and based on the high intensity diversity of the pixels within the search region. Also, the relatively low resolution of the ultrasound image, especially in the longitudinal direction, can cause loss of the image details (Zahnd et al., 2013), which will hamper the motion tracking of the carotid artery wall. Finally, the out-of-plane motion and tissue deformation of the carotid artery wall might greatly hinder the performance of the BM method. Therefore, it is necessary to add additional model constraints, which can describe the carotid arterial dynamics, for reducing the disturbance from the above-mentioned difficulties.

The state-space approach is an appropriate strategy to integrate the model constraints for the carotid arterial dynamics:

\[
\begin{align*}
\mathbf{x}_{n+1} &= \mathbf{F}\mathbf{x}_n + \mathbf{w}_n \\
\mathbf{y}_n &= \mathbf{H}\mathbf{x}_n + \mathbf{v}_n \\
\end{align*}
\]

(1)

with the frame index \( n \), the state variable \( \mathbf{x} \), observation variable \( \mathbf{y} \), system noise \( \mathbf{w}_n \), observation noise \( \mathbf{v}_n \) and the coefficient matrices
F and H. The goal of the state-space approach is to use the noisy observation y of the state x to determine its optimal estimate in the next instant. One commonly used algorithm to solve the state space equations is the Kalman filter (Simon, 2006). It provides an efficient and convenient solution of the state-space equations (see Eq. (11)) in the minimum-mean-square-error (MMSE) sense, and ensures that the solution is the best linear unbiased estimate when the system and measurement noise can be assumed to have the Gaussian distribution with known statistics. The Kalman filter was first applied in the motion tracking of the carotid artery wall by Gastounioti et al. (2010). This study combined the Kalman filter and the BM method (KBMM) by extracting the information of the observation variable in the state-space equations from the best-matched block computed by the BM method. Further, Gastounioti et al. (2011) compared the different predefined values of the noise statistics used in the KBMM, and tried another two combinations of the Kalman filter and the BM method for correcting the estimated motion of the carotid artery wall during and after the tracking process. Recently, Zahnd et al. (2012, 2013, 2015) have added a control signal into the state-space equations in order to avoid the divergence of the estimated motion trajectory due to the speckle decorrelation problem in the ultrasound imaging. A comparison of the KBMM methods for the motion tracking methods of the carotid artery wall can be found in Gastounioti et al. (2013).

However, the Gaussian assumption of the system and measurement noise in the state-space equations is typically unrealistic in the ultrasound sequence. Thus, the motion trajectories estimated by KBMM may be biased. Also, it is necessary to know in advance the parameters used in KBMM, such as noise statistics, sizes of the reference block and the search region etc. However, the determination of the appropriate values of these parameters is always a challenge in the previous approaches. Consequently, it is worthwhile to investigate an alternative approach, which is more robust to the noise disturbance and easier to determine the initialization, for more accurate tracking of the motion of the carotid artery wall.

1.3. Objective and contributions of our study

Our preliminary work has developed a robust H∞ filter to obtain the optimal estimate of the state-space equation without the Gaussian assumption of the system and measurement noise (Gao et al., 2015). However, it still used the constant diagonal matrix in the state space equations like the KBM. Hence, in this paper, we have developed an elasticity-model based state-space approach (ESS) for accurately and robustly tracking the motion of the carotid artery wall. This elastic model constraint is mainly inspired by the elastic wall property and the quasi-periodic motion of the carotid artery wall (Zahnd et al., 2013). In the discretization of this state-space approach, the formulation of the coefficient matrices can be derived from the elasticity model, rather than simply set by the constant diagonal matrix as the KBM. Moreover, we have used the H∞ filter to obtain the optimal estimate of the state-space equations because the H∞ filter solves the optimization problem by minimizing the maximum ratio of the estimation error to the disturbance (min-max approach) (Liu and Shi, 2004). This can ensure that if the disturbances are small, the estimation errors will be as small as possible (Hassibi et al., 1996; Liu et al., 2012). This property implies that the H∞ filter is less restricted than the Kalman filter with respect to the noise knowledge. By incorporating this more realistic model constraint for the dynamics of the carotid artery wall, our approach can ensure robust estimation for tracking the motion of the carotid artery wall in the ultrasound image sequence when the noise are unknown and non-Gaussian.

In order to handle the problem of the parameter selection, we have proposed a training strategy to obtain the optimal parameters used in EBS, such as the sizes of reference block and search region, coefficients in the system dynamics equation and the noise covariance, etc. This strategy can facilitate avoiding the artificial errors introduced by the manual parameter selection in the KBM. We have then proposed an evaluation function to measure the similarity between the results from two methods: the EBS and the manual tracing method. The previous BM/KBM methods measured the similarity by only considering the distance between the centers of two blocks (Golemati et al., 2003; Gastounioti et al., 2010; Zahnd et al., 2011, 2012, 2013; Yli-Ollila et al., 2013; Gastounioti et al., 2013). This similarity, however, does not take into account the influence on the similarity measurement based on the sizes of the two blocks and their overlapping degree. This is the motivation behind our proposed evaluation function in this study.

The performance of the EBS was tested on a set of 280 ultrasound sequences (including left and right carotid arteries) acquired from 140 subjects by comparing with the manual tracing results from three medical physicians and three kinds of previous methods.

2. Methodology

The EBS consists of two major steps: updating step and prediction step (explained in Section 2.1). The updating step aims to compute the location of the best-matched block from the reference block in the same frame. The prediction step focuses on the appearance estimation of the reference block in the next frame from the best-matched block and the reference block in the current frame. In order to solve the problem of the parameter determination, we apply a training strategy to automatically compute the optimal values of the parameters used in EBS (explained in Section 2.2). Finally, we develop an evaluation function for measuring the error between the motion tracking results obtained by EBS and by the manual tracing method (explained in Section 2.3). The flowchart of EBS is illustrated in Fig. 2.

2.1. Tracking the motion of the carotid artery wall

2.1.1. Updating step

In the updating step, we apply the BM method to estimate the location of the best-matched block. Let n denote the time index or frame index, B_{ref} is the reference block, B_{best} is the best-matched block and B_{can} are m different candidate blocks extracted in the search region, where the candidate block has the same size as the reference block; l_i is the value of normalized cross-correlation (Golemati et al., 2003) between the reference block and ith candidate block for any i \in \{1, 2, \ldots, m\}. Then, we can compute l_i by

\[ l_i = \frac{\sum_{a=1}^{p} \sum_{b=1}^{q} (B_{ref}^{a,b} - \mu_{ref}^{a,b}) (B_{can}^{a,b} - \mu_{can}^{a,b})}{\sqrt{\sum_{a=1}^{p} \sum_{b=1}^{q} (B_{ref}^{a,b} - \mu_{ref}^{a,b})^2 \sum_{a=1}^{p} \sum_{b=1}^{q} (B_{can}^{a,b} - \mu_{can}^{a,b})^2}} \]  

(2)

where

\[ \mu_{ref}^{a,b} = \frac{1}{p \times q} \sum_{a=1}^{p} \sum_{b=1}^{q} B_{ref}^{a,b} \]  

(3)

In Eqs. (2) and (3), p is the block height and q is the block width. B_{ref}^{a,b} and B_{can}^{a,b} are the gray values of pixels at the coordinates (a, b) in the reference block and the candidate block respectively. \mu_{ref}^{a,b} and \mu_{can}^{a,b} are the mean values of the pixel
intensity within the reference block and within the ith candidate block respectively. We can then obtain the best-matched block $B^\text{best}_i$ by considering it as the candidate block with the largest value of the normalized cross-correlation, i.e.

$$B^\text{best}_i = B^\text{corr}_j, \quad \text{where } j = \arg \max_{i \in \{1, 2, \ldots, m\}} | \hat{f}_i |$$

(4)

2.1.2. Prediction step

In the prediction step, we estimate the appearance of the reference block in the next frame, based on the appearance of the best-matched block and the reference block in the current frame. In this step, we first describe the brightness change of the target tissue within the reference block by using the linear elastic model, and then represent it as the state-space equations by the state-space approach. In order to avoid the assumption of uncertainty and obtain a better performance to counter the noise disturbance, we use the $H_n$ filter to solve the state-space equations, and obtain the brightness of the reference block in the next frame.

Elasticity model and its state-space representation. In the ultrasound imaging, the brightness of the target tissue within the reference block varies with the motion of the carotid artery wall during the cardiac cycle, and this brightness change may disturb the tracking process. In the EBS, we use the linear elastic model to describe the brightness change. The linear elastic model is formulated in terms of the Hooke's law (Taylor, 2005; Guzman et al., 2014; Wang, 2014; Zhang et al., 2016),

$$F = -\tilde{k} y$$

(5)

where $y$ is the brightness of the reference block shown in Fig. 1, $F$ is the force driving the brightness change and $\tilde{k}$ is the elastic coefficient. We denote $\dot{y}$ and $\ddot{y}$ as the first and second derivatives of $y$ with respect to time $t$ respectively; that is $\dot{y} = \frac{dy}{dt}$ and $\ddot{y} = \frac{d^2y}{dt^2}$. The elastic force $F$ can be formulated by Newton's second law, as

$$F + u = \tilde{m} \ddot{y}$$

(6)

where $u$ is the external influence on the appearance change of the reference block, and $\tilde{m}$ is a parameter to be determined. From Eqs. (5) and (6), we then obtain

$$\ddot{y} + \frac{\tilde{k}}{\tilde{m}} y = \frac{1}{\tilde{m}} u$$

(7)

Eq. (7) is the differential equation for describing the brightness change of the reference block, and it can be reformulated into the state-space formulation as follows (Ogata, 2010),

$$\mathbf{x}_{n+1} = \mathbf{F} \mathbf{x}_n + \mathbf{G} \mathbf{u}$$

$$\mathbf{y}_n = \mathbf{H} \mathbf{x}_n$$

(8)

where

$$\mathbf{x}_n = \begin{bmatrix} x_{n1}^1 & \cdots & x_{n1}^q & x_{n2}^1 & \cdots & x_{n2}^q & \cdots & x_{nq}^1 & \cdots & x_{nq}^q \end{bmatrix}^T$$

$$\mathbf{u} = \begin{bmatrix} u_{n1}^1 & \cdots & u_{n1}^q & u_{n2}^1 & \cdots & u_{n2}^q & \cdots & u_{nq}^1 & \cdots & u_{nq}^q \end{bmatrix}$$

$$\mathbf{y}_n = \begin{bmatrix} y_{n1}^1 & \cdots & y_{n1}^q & y_{n2}^1 & \cdots & y_{n2}^q & \cdots & y_{nq}^1 & \cdots & y_{nq}^q \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} F_1 & F_2 & F_3 & F_4 \end{bmatrix} \mathbf{G} = \begin{bmatrix} G_1 & G_2 \end{bmatrix} \mathbf{H} = \begin{bmatrix} H_1 & H_2 \end{bmatrix}$$

(9)
and

\[
F_1 = \cos \left( \frac{k}{m} \right) I_{p,p} \quad F_2 = \sqrt{m} \sin \left( \frac{k}{m} \right) I_{p,p} \quad F_3 = -\sqrt{m} \sin \left( \frac{k}{m} \right) I_{p,p} \quad F_4 = \cos \left( \frac{k}{m} \right) I_{p,p}
\]

\[
G_1 = \frac{1}{k} \left( 1 - \cos \left( \frac{k}{m} \right) \right) I_{p,p} \quad G_2 = \left( \frac{1}{mk} \sin \left( \frac{k}{m} \right) \right) I_{p,p}
\]

\[
H_1 = I_{p,p} \quad H_2 = O_{p,p}
\]

The derivations from Eq. (7) to Eqs. (8)–(10) can be seen in Appendix A. In Eqs. (8)–(10), \( n \) is the frame index, \( T \) is the sampling period of the ultrasound sequence, \( p \) is the height of the reference block and \( q \) is the width. \( I_{p \times p} \) is the \( p \times p \) identity matrix and \( O_{p \times p} \) is the \( p \times p \) zero matrix. \( x_n \) is the matrix corresponding to the brightness of the reference block \( B_{n}^{ref} \) and \( x_n^{h} \) is the gray value of the pixel at the coordinate \( (i, j) \) in \( B_{n}^{ref} \). Similarly, \( y_n \) corresponds to the best-matched block \( B_{n}^{est} \) and \( y_n^{h} \) is the gray value of the pixel at the coordinate \( (i, j) \) in \( B_{n}^{est} \). \( u \) is the input of the state-space equation independent of time, and is specified to be the reference block \( B_{n}^{ref} \) manually selected in the first frame. Its element \( u_{ij} \) is the gray value of the pixel at the coordinate \( (i, j) \) in \( B_{n}^{ref} \), \( y_n^{h} \) is the first time derivative of \( y_n \).

It is crucial to determine the time to sample the first frame of the ultrasound sequence. Sampling the first frame at a different time can provide the reference block \( B_{n}^{ref} \) with different brightness. It can affect the results of EBS because \( B_{n}^{ref} \) provides the reference value to the brightness of the target tissue. Considering the limitation to the motion of the carotid artery wall due to its tension, the first frame should be sampled at the time that the elastic dynamics is in a relaxed state \( (F = 0) \). This can help to reduce the influence of the tension in the carotid artery wall on the brightness of target tissue as much as possible. However, it is difficult to determine the first frame in the ultrasound sequence corresponding to the relaxed state of the carotid wall, because the gray-level ultrasound sequences do not provide sufficient information to directly evaluate the tension in the carotid artery wall. Hence, we have tagged the ultrasound images with electrocardiography (ECG) gating, and determined the first frame of the ultrasound sequence by using the correspondence between the ECG wave and the dynamic ultrasound image. Because the process of aortic ejection affects the tension in the carotid artery wall through the propagating pulse wave and the corresponding blood pressure, the carotid artery wall is in a relaxed state when the aortic blood pressure is at a low level. Moreover, the aortic blood pressure reaches its minimum value in the QRS duration of the ECG wave (Guyton and Hall, 2006), as illustrated in Fig. 3(a). Therefore, the first frame of the ultrasound sequence should be sampled within the QRS duration. However, the QRS durations extracted from many subjects cannot be clearly observed, as shown in Fig. 3(b), and this may affect the selection of the first frame. In order to avoid this problem, we have sampled the first frame of the ultrasound sequence at a time as close as possible to the R-peak in the ECG wave form.

\[
H_{\infty} \text{ filter. As regards the influence of ultrasound sequence from system noise } w_n \text{ and the observation noise } v_n, \text{ Eq. (8) should be reformulated as follows,}
\]

\[
x_{n+1} = Fx_n + Gu + w_n \quad y_n = Hx_n + v_n
\]

In order to solve Eq. (11), we need to find the optimal estimate \( \hat{x}_n \) of \( x_n \), and this estimation can be implemented by minimizing the following cost function \( J \) by the \( H_{\infty} \) filter (Simon, 2006).

\[
J = \frac{\sum_{n=0}^{N-1} ||x_n - \hat{x}_n||^2_{I^v} + \sum_{n=0}^{N-1} (||w_n - \hat{w}_n||_{Q^w} + ||v_n - \hat{v}_n||^2_{R^v})}{||x_0 - \hat{x}_0||^2_{I^v}}
\]

where \( \hat{x}_1 \) is a priori estimate of \( x_1 \), \( P_0 \) and \( L_0 \) are user-specified symmetric positive definite matrices, \( Q_{n} \) and \( R_{n} \) are the covariance matrices of the noise terms \( w_n \) and \( v_n \), respectively. In the EBS, \( Q_{n} \) and \( L_{n} \) are set to the diagonal matrices independent of the frame index \( n \), that is \( Q_{1} = Q_{2} = \cdots = Q_{N-1} = Q_{N} = 0 \). \( R_{1} = R_{2} = \cdots = R_{N} = R = I \) and \( L_{1} = L_{2} = \cdots = L_{N-1} = L = I \). Because the direct minimization of \( J \) is not tractable, a strategy for generating the optimal estimation of \( \hat{x}_{n+1} \) is developed by making the cost function \( J \) to satisfy a user-specified upper bound (Simon, 2006). By denoting the reciprocal of the upper bound as \( \theta \), we have \( J < 1/\theta \). Then let \( P \) be the covariance of the estimation error, which can be computed by solving the following Ricatti equation,

\[
P = FP(1 - \theta L)P + H^T R^{-1} (1 - \theta L)^{-1} F^T + Q
\]

and the filter gain \( K \) can be defined by the equation

\[
K = P(1 - \theta L)P + H^T R^{-1} H^{-1} - H^T R^{-1} F^T + Q
\]

The optimal state estimate \( \hat{x}_{n+1} \) in the \( n+1 \)th frame can be obtained by

\[
\hat{x}_{n+1} = Fx_n + Gu + FK(y_n - H\hat{x}_n)
\]

The solution of the \( H_{\infty} \) filter shown in Eq. (15) was originally presented by Banavar (1992) and further discussed by Shen (1995) and Shen and Deng (1997). The derivations from Eqs. (11) and (12) to Eqs. (13)–(15) can be found in Simon (2006).

In practice, the appearance of the reference block \( B_{n}^{ref} \) corresponds to a matrix composed of the left \( p \) column vectors of \( \hat{x}_n \), and the appearance of the best-matched block \( B_{n}^{est} \) corresponds to \( y_n \). Thus Eq. (15) shows that the reference block in the next frame can be computed by the weighted summation between the best-matched block and the reference block in the current frame. In addition, it is worth noting that Eq. (15) begins to work at the third frame in the ultrasound sequence. The reason is that \( y_n \) in Eq. (9) cannot be computed when estimating \( B_{n}^{ref} \). Instead, we make \( B_{n}^{ref} \) equal to \( B_{n}^{est} \). The block diagram of tracking the motion of the carotid artery wall between successive frames is shown in Fig. 4.

2.2. Parameters determination

The parameters in the BM-based method for tracking the motion of the carotid artery wall need to be decided and determined beforehand. In the previous studies, the parameters were usually set manually by ultrasound physician experts (Cinthio et al., 2005; Gastounioti et al., 2010; Yli-Ollila et al., 2013; Zahn et al., 2013). The manual parameter determination, however, highly relies on the professional experience and may lead to error due to the potential expert’s biases. In order to deal with the problem of the manual parameter setting, we have proposed a strategy to automatically determine the parameters based on the Powell method (Powell, 1964; Antoniou and Lu, 2007). In our strategy, the optimal parameter set is acquired through testing different configurations of the parameter set in the parameter space. The parameter space is characterized by a high-dimensional rectangular space. Every dimension of this rectangular space corresponds to one parameter, and the length of this dimension equals to a predefined interval of the corresponding parameter.

In EBS, the following ten parameters need to be determined in advance: the block height \( h_b \), the ratio \( \tilde{r}_b \) of the block width to the block height, the ratio \( \tilde{r}_c \) of the search region height to the block height, the ratio \( \tilde{r}_w \) of the search region width to the
block width, the coefficients $\tilde{m}$ and $\tilde{k}$ in Eq. (10), the reciprocal of the upper bound $\tilde{\theta}$ of the cost function $J$, and the coefficients $\tilde{q}, \tilde{r}, l$ of the matrix $Q, R, L$. The predefined intervals of these parameters are shown in Table 1. These intervals were determined by setting their values slightly larger than those specified in the previous studies (Golemati et al., 2003, 2012; Cinthio et al., 2005, 2006; Gastounioti et al., 2010, 2011, 2013; Zahnd et al., 2011, 2012, Yli-Ollila et al., 2013). Then we normalize all the parameters into the interval $[0, 1]$, and denote the normalized parameter set by $s$, where $s = [s_1, s_2, \ldots, s_N]$ ($N = 10$) corresponds to the set $[\tilde{h}_b, \tilde{T}_b, \tilde{T}_w, \tilde{m}, \tilde{k}, \tilde{q}, \tilde{r}, \tilde{l}]$. The training strategy is now outlined below.

In the initialization step, we initialize the normalized parameter set $s$ by $s_0$. Every element in $s_0$ is assigned by a random value...
Table 1
The predefined parameter intervals for the parameter training process.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
<td>Block height</td>
<td>([0.3])</td>
</tr>
<tr>
<td>( r_b )</td>
<td>Ratio of the block width to the block height</td>
<td>([1.3])</td>
</tr>
<tr>
<td>( r_s )</td>
<td>Ratio of the search region height to the block height</td>
<td>([1.25])</td>
</tr>
<tr>
<td>( r_w )</td>
<td>Ratio of the search region width to the block width</td>
<td>([1.25])</td>
</tr>
<tr>
<td>( \vec{m} )</td>
<td>Parameter in Eq. (5)</td>
<td>([0.1])</td>
</tr>
<tr>
<td>( \vec{r} )</td>
<td>Parameter in Eq. (6)</td>
<td>([0.1])</td>
</tr>
<tr>
<td>( \theta )</td>
<td>Reciprocal of the upper bound of the cost function ( J )</td>
<td>([2.10])</td>
</tr>
<tr>
<td>( \eta )</td>
<td>Coefficient of the matrix ( Q )</td>
<td>([0.1])</td>
</tr>
<tr>
<td>( \vec{\tau} )</td>
<td>Coefficient of the matrix ( \mathbf{R} )</td>
<td>([0.1])</td>
</tr>
<tr>
<td>( \vec{\Theta} )</td>
<td>Coefficient of the matrix ( \mathbf{L} )</td>
<td>([0.1])</td>
</tr>
</tbody>
</table>

distributed by the uniform distribution in the interval \([0, 1]\). Then we define the direction matrix \( \mathbf{d} \) and initialize it by

\[
d = \text{diag}(\mathbf{s}) = \begin{bmatrix}
s_1 & 0 & \cdots & 0 \\
0 & s_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & s_N
\end{bmatrix}\]

(16)

where \( \mathbf{d} \) is a diagonal matrix and \( d_i \) denotes the \( i \)th row vector of \( \mathbf{d} \).

Then in the iteration step, we compute the optimal learning rate \( \alpha \) in the \( i \)th iteration by minimizing the evaluation function \( f \) by

\[
\alpha_i = \arg \min_{\alpha \in \mathbb{R}} f(s_{i-1} + \alpha \mathbf{d}_i)
\]

(17)

where the evaluation function \( f \) measures the error between the results from EBS and the manual tracing method, and is defined in Section 2.3. \( R_1 \) is the range of the learning rate \( \alpha \), defined by

\[
R_1 = \left[ \min_{k \in [1, N]} (z_k), \max_{k \in [1, N]} (z_k) \right] \quad z_k = \frac{s_{i-1, k}}{d_{i, k}}
\]

(18)

where \( s_{i-1, k} \) is the \( k \)th element of the vector \( s_{i-1} \), and \( d_{i, k} \) is the \( k \)th element of \( \mathbf{d}_i \). Then the parameter set \( \mathbf{s} \) can be updated by the equation

\[
s_i = s_{i-1} + \alpha_id_i
\]

(19)

When \( i = N \), we let

\[
\mathbf{s}_{N+1} = \mathbf{s}_0 + \alpha_i \mathbf{d}_{N+1}
\]

(20)

where

\[
\alpha_i = \arg \min_{\alpha \in \mathbb{R}} f(s_0 + \alpha \mathbf{d}_{N+1})
\]

(21)

and \( R_2 \) is defined by

\[
R_2 = \left[ \min_{k \in [1, N]} (z_k), \max_{k \in [1, N]} (z_k) \right] \quad z_k = \frac{s_{0, k}}{d_{(N+1), k}}
\]

(22)

where \( s_{0, k} \) is the \( k \)th element of \( \mathbf{s}_0 \), and \( d_{(N+1), k} \) is the \( k \)th element of \( \mathbf{d}_{N+1} \).

In the convergence step, the error \( e \) is defined as follows:

\[
e = \| \alpha_i \mathbf{d}_{N+1} \|
\]

(23)

If \( e \) is smaller than a predefined threshold \( \varepsilon = 0.1 \) (\( \varepsilon \) is denoted by \( \varepsilon^* \) at this time), then the iteration will stop and the optimal parameter vector \( \mathbf{s}^* = \mathbf{s}_{N+1} \). If not, the matrix \( \mathbf{d} \) is updated by

\[
\mathbf{d} = \mathbf{Z} \mathbf{d}
\]

(24)

and then the iteration from Eq. (17) to Eq. (23) can be continued.

Algorithm 1. Parameter training.

\textbf{Input:} \( (1) \ I_1, I_2, \ldots, I_M \) - initialization of the parameter set \( \mathbf{s} \) in \( M \) training processes.

\textbf{(2) } \varepsilon \ - \text{threshold.}

\textbf{Output:} \( \mathbf{s}^* \) - the optimal parameter set.

\textbf{for} \( m = 0 \) to \( M \) \textbf{do}

\textbf{Initialize} the parameter set and the direction matrix in Eq. (16)

\[
\mathbf{s}_0^* \leftarrow I_m;
\]

\[
\left[ \begin{array}{c}
\mathbf{d}_1^0 \\
\mathbf{d}_2^0 \\
\vdots \\
\mathbf{d}_M^0
\end{array} \right] \left[ \begin{array}{c}
\mathbf{d}_1^0 \\
\mathbf{d}_2^0 \\
\vdots \\
\mathbf{d}_M^0
\end{array} \right]^T \leftarrow \text{diag}(\mathbf{s}_0^*)
\]

\( j = 0 \);

while \( \varepsilon^* \geq \varepsilon \) \textbf{do}

\textbf{for} \( i = 1 \) to \( N \) \textbf{do}

\[
R_1 \leftarrow \left[ \min_{k \in [1, N]} \alpha_k \left( \frac{\text{min}_{k \in [1, N]} \alpha_k \left( \frac{\text{min}_{k \in [1, N]} \alpha_k} {4\alpha_k} \right) \right) \right] \\
\alpha_i \leftarrow \arg \min_{\alpha \in \mathbb{R}} f(s_i^* + \alpha d_i^0);
\]

\[
s_i^* \leftarrow s_{i-1}^* + \alpha_i d_i^0;
\]

\( j = j + 1; \)

end

\( \varepsilon_m^* \leftarrow \varepsilon^*; \)

\( s_m^* \leftarrow s_{N+1}^*; \)

end

\( m^* \leftarrow \arg \min_{m \in [1, M]} \varepsilon_m^*; \)

\( \mathbf{s}^* \leftarrow s_{m^*}^*; \)

end

In order to reduce the random error from the initialization of the parameter set \( \mathbf{s} \), the above training process is repeated \( M \) times \((M = 10)\), which means the parameter set \( \mathbf{s} \) should be re-initialized in each training process. Thus for any \( m \in \{1, 2, \ldots, M\}, \) we can compute an optimal parameter set \( \mathbf{s}_m^* \) and the corresponding error \( \varepsilon_m^* \) in the \( m \)th training. Then we can compute the optimal parameter set \( \mathbf{s}^* \) in \( M \) training processes by means of the equations,

\[
\mathbf{s}^* = \mathbf{s}_{m^*}^*; \quad m^* = \arg \min_{m \in [1, M]} \varepsilon_m^*
\]

(25)

After the training process, every parameter in \( \mathbf{s}^* \) will be transformed from \([0, 1]\) into their original intervals shown in Table 1. The pseudo-code of the parameter training process is shown in Algorithm 1.

2.3. The evaluation function f-distance

For the performance evaluation of the tracking method, we need to measure the error or similarity between results separately obtained from the computer-aided method and the manual tracing method. We denote \( \mathbf{O}_1 \) as the block tracked by the computer-aided method, and \( \mathbf{O}_2 \) as the block obtained by the manual tracing method. In the previous studies, the error was measured by the Euclidean distance between the centers of \( \mathbf{O}_1 \) and \( \mathbf{O}_2 \) (center-to-center Euclidean distance) \((\text{Golemati et al., 2003, 2012; Cinthio et al., 2006, 2005; Gastouioti et al., 2010, 2011, 2013; Zahnd et al., 2011, 2012, 2013; Yli-Ollila et al., 2013})\). In contrast, we have proposed an evaluation function to measure the error (called \( f \)-distance) between \( \mathbf{O}_1 \) and \( \mathbf{O}_2 \), based on the below mentioned basis or motivations.

In the parameter training process, we need an evaluation function to measure the performance of EBS in the current parameter set, shown in Eqs. (17) and (21). Nevertheless, the four parameters \((h_b, \bar{h}_b, \bar{t}_h, \bar{f}_w)\) in the parameter set related to the block size
have no influence on the performance of EBS, if we consider the center-to-center Euclidean distance as the evaluation function. This means that the four parameters cannot be updated in the training process when using the center-to-center Euclidean distance. Thus, we need to apply an evaluation function in the parameter training process in order to enable every parameter able to affect the performance of EBS.

In the second motivation, we consider that not only the center-to-center distance, but also other information of the block, such as the block size and the overlapping degree between $O_1$ and $O_2$, can also facilitate the performance measurement. For the $f$-distance, the increase of the block size and the overlapping degree lead to the decrease of the error between $O_1$ and $O_2$, when the center-to-center Euclidean distance is unchanged. For example, the relative position between the two blocks shown in Fig. 5(a) equals to that shown in Fig. 5(b). If we only consider the center-to-center Euclidean distance, the errors measured by the distance in Fig. 5(a) and (b) are the same. However, the two blocks in Fig. 5(b) may be more similar due to the increase of the block size and overlapping degree.

The $f$-distance between $O_1$ and $O_2$ is defined by

$$f(O_1, O_2) = E_1 + E_2$$

(26)

The low value of the $f$-distance indicates the low level of the error between $O_1$ and $O_2$. In Eq. (26), $E_1$ is the weighted Euclidean distance between the centers of $O_1$ and $O_2$.

$$E_1(c_1, c_2) = (1 - \beta_1 + \beta_2)||c_1 - c_2||_2$$

(27)

where $c_1$ and $c_2$ are the coordinates of center points of $O_1$ and $O_2$, respectively. Then, $E_2$ is the minimum distance between $O_1$ and $O_2$, defined by

$$E_2 = \min_{a_1, a_2} ||a_1 - a_2||_2, \ \forall a_1 \in O_1, \forall a_2 \in O_2$$

(28)

$\beta_1$ is the overlapping rate between $O_1$ and $O_2$, i.e.

$$\beta_1 = \frac{2|O_1 \cap O_2|}{|O_1| + |O_2|}$$

(29)

where the operator $|.|$ is the cardinal number of the set. Finally, $\beta_2$ describes the computational cost spent by the BM method with the NCC criterion. Let $h$ and $w$ be the height and width of the search region, and $h_m$ and $w_m$ be the height and width of the ultrasound image. Then we define $\beta_2$ by

$$\beta_2 = \frac{\kappa p_1}{p_2}$$

(30)

In Eq. (30), $\kappa$ is weight value set 50 in order to make $\beta_1$ and $\beta_2$ to be of the same order of magnitude, and $p_1$ corresponds to the current computational cost of the BM method. In the BM method, the sizes of the reference block and the search region can be obtained from the current parameter set $s_i$ in the iteration of the parameter training, or from the optimal parameter set used in the performance evaluation of EBS. Considering the NCC computation in Eq. (2), $p_1$ can be formulated as

$$p_1 = 2hhw_h(w_h - 1)(w_h - 1)$$

(31)

where $h_h$ is the height of the reference block and $w_h$ is the width. In Eq. (31), the term $2hhw_h$ corresponds to the cost of NCC computation between the reference block and one candidate block, and the coefficient “2” considers both the addition and the multiplication. The term $(w_h - 1)(w_h - 1)$ shows the number of NCC computations implemented in the search region. Due to $w_h = \tilde{r}_h \bar{h}_h$, Eq. (31) can be reformulated as

$$p_1 = 2\tilde{r}_h \bar{h}_h^2 (w_h - 1)(w_h - 1)$$

(32)

Then, we have $p_2$ to correspond to the maximum computational cost of the BM method. In the computation of $p_2$, the search region is the entire area of the ultrasound image. Thus we define $p_2$ by

$$p_2 = \max_l 2\tilde{r}_l \bar{l}_l^2 (h_m - l + 1)(w_m - l)$$

(33)

where $l$ is the height of the reference block with any size, and thus its width is $\tilde{r}_l l$. The solution of Eq. (33) can be seen in Appendix B.

The $f$-distance has different properties when the two blocks $O_1$ and $O_2$ are overlapped or separated. When the two blocks are overlapped, the $f$-distance is dominated by the weighted Euclidean distance $E_1$, because the minimum distance $E_2$ equals to zero. Then, its coefficient $1 - \beta_1$ is used to decrease the value of the $f$-distance when the block size increases. This property of the $f$-distance impels the choice of the relatively large block, because it can lead to reduction of the error between $O_1$ and $O_2$. Nevertheless, it has a side effect that the error between $O_1$ and $O_2$ may be considered to stay at a low level when their centers are distant, because the two blocks have large size at this time. In addition, it also leads to raising the computational cost since the amount of calculation of NCC substantially increases. In order to avoid this problem, we propose the coefficient $\beta_2$ to adjust the value of $E_1$. The coefficient $\beta_2$ can prevent the excessive increase.
of the block size by increasing the value of the $f$-distance when thecomputational cost grows. Considering the case of $O_1$ and $O_2$ being separated, the coefficient $1 - \beta_1$ is unable to affect the $f$-distance because $1 - \beta_1$ is identical to one, and thus the formulation of $E_1$ degenerates into $E_1 = (1 + \beta_2)||Q_2 - C_1||$. Therefore, $E_2$ is used to decrease the value of the $f$-distance when the block size increases, and the coefficient $\beta_2$ is applied as a penalty for the excessive increase of the block size. The comparison between the $f$-distance and the center-to-center Euclidean distance are shown in our experiments (see Section 3.4.3).

3. Experiments and results

3.1. Population study

A total of 140 subjects were involved in this study. The healthy subjects included 18 males (59 $\pm$ 12 years old) and 19 females (59 $\pm$ 8 years old). The unhealthy population included 60 males (58 $\pm$ 12 years old) and 43 females (62 $\pm$ 11 years old). The inclusion criterion for the unhealthy subjects was the presence of one of the three diagnosed diseases: heart disease, type 1 or 2 diabetes and hypertension. Fig. 6 displays the population distribution across the three diseases types. Each participant was informed of the purpose and procedure of this study. Informed consent was obtained from each participant. This study was approved by the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (China), and the Second People’s Hospital of Shenzhen (China).

3.2. Data collection

The data collection was implemented by an ultrasound physician with more than 10-year experiences by using a high-resolution ultrasound system (iU22, Philips Ultrasound, Bothell, WA, USA) and a 7.5 MHz liner array transducer. For each subject, two ultrasound sequences were separately acquired from bilateral carotid arteries (left and right carotid arteries), and thus 280 ultrasound sequences were used in this study. The ultrasound sequences were recorded through six consecutive full cardiac cycles with recognizable R-peak in the ECG wave. When the R-peak of ECG wave was unrecognizable, the ultrasound sequence and its corresponding ECG signal should be recollected until the R-peak was recognized. Then all imaging data were saved as DICOM format into CD for off-line analysis. During the collection, the subjects were examined in the supine position, with the head turned 45° away from the examined side. The following settings of the ultrasound machine were used for all acquisitions: the dynamic range was 60 dB, the sequence frame rate was 24–51 frames/second, the pixel size in both radial and longitudinal directions was 19.2 pixels/mm.

In addition, the manual tracing results were performed by three experienced ultrasound physicians, and considered as the ground truth. For each ultrasound sequence, every expert traced one motion trajectory of the carotid artery wall twice, and these two tracings occurred one month apart. Thus six sets of manual tracing results were obtained from the three experts for all the subjects.

3.3. Parameter training

In the training process, we divided the dataset into the training set and the test set. In order to reduce the influence on the training process from the differences between healthy and unhealthy populations, the training set and the test set contained the data from both healthy and unhealthy subjects. Thus, we randomly and equally divided the set of all the healthy subjects into three subsets (denoted by $DH_1$, $DH_2$, $DH_3$), and divided the set of all unhealthy subjects into three subsets (denoted by $DU_1$, $DU_2$, $DU_3$). By this partition of the healthy and unhealthy subjects, we could obtain three training sets $Tr_1$, $Tr_2$ and $Tr_3$ by

$$Tr_1 = DH_1 \cup DU_1, \quad Tr_2 = DH_2 \cup DU_2, \quad Tr_3 = DH_3 \cup DU_3$$

(34)

Corresponding to $Tr_1$, $Tr_2$ and $Tr_3$, we could obtain three test sets $Te_1$, $Te_2$ and $Te_3$ by

$$Te_1 = DH_3 \cup DH_2 \cup DU_3 \cup DU_2$$

$$Te_2 = DH_1 \cup DH_3 \cup DU_1 \cup DU_3$$

$$Te_3 = DH_1 \cup DH_2 \cup DU_1 \cup DU_2$$

(35)

In addition, for the parameter training, the manual tracing results were required to compute the evaluation function $f$ for optimizing the learning rate $\alpha$ in Eqs. (17) and (21). Thus, different manual tracing results might lead to different values of the parameters in the training process. In order to investigate the influence of the different manual tracings on the tracking accuracy, we repeated the training process on each of the three training sets and each of the six sets of manual tracing results. Therefore, totally 18 training processes were performed in our approach, and then 18 corresponding test processes (denoted by $D_1$, $D_2$, …, $D_{18}$) could be used to evaluate the performance of EBS.

Fig. 7 shows the change of the parameters listed in Table 1 during the training process. In each subfigure of Fig. 7, the box “0” (see the x-axis) shows the initial parameter distribution before the training. The boxes “1”–“18” correspond to the parameter distribution after the eighteen training processes respectively. For each training process, we then select an optimal parameter set (including ten parameters) corresponding to the smallest value of $e$.
shown in Eqs. (23) and (25). The dispersion of optimal parameters in the eighteen training processes is shown as the box “19”. For example, the box “19” of the left upper subfigure represents the dispersion of the optimal values of the parameter $\tilde{h}$ in each of the boxes “1”-“18”. The correspondence between the subfigures and the parameters are shown in the lower right corner of Fig. 7.

Fig. 8(a) gives an example of ten repeated trainings in a training process. This figure shows that the variation of $e$ defined in Eq. (23) against the iteration number. Each marker corresponds to the training process with a configuration of the initialization, and the arrow on each marker shows the increase (dark green arrow) or decrease (blue arrow) of $e$ between the two neighboring iterations. The green dotted line shows the level of the convergence threshold $\varepsilon = 0.1$. When $e < \varepsilon$ or the training has been iterated by ten times, the training process is terminated (without the arrow). It means that maximum iteration number is ten. Then, the parameters with the lowest value of $e$ are considered as the optimal parameters in this training process. In this example, the training process corresponding to the optimal parameter values finally obtained, is showed as the red dashed curve. Considering all eighteen training processes, Fig. 8(b) shows the change of $e$ in Eq. (23) against the iteration number on each training process. The iteration corresponds to the process of the parameter update from Eq. (17) to Eq. (23). The results show that all the training processes can converge within a small number of iterations, and the values of $e$ are $0.0024 \sim 0.0204$ when all the training processes have been converged.

### 3.4. Performance evaluation

The performance of EBS was evaluated by comparison with the manual tracing method and with the three kinds of previous methods. Six different measurements of the best-matched block were proposed to demonstrate the results of EBS: the radial position of motion trajectory (RP) and the longitudinal position of motion trajectory (LP), the Euclidean distance between the motion trajectories along radial direction $(\tilde{e}_1)$, longitudinal direction $(\tilde{e}_3)$ and on the 2D plane $(\tilde{e}_4)$ and the $f$-distance $(\tilde{e}_5)$. The first two measurements were analyzed by the linear regression analysis and the Bland–Altman analysis, and the last four measurements were investigated by the distance analysis. The linear regression analysis was applied to measure the correlation between the results of the EBS and the manual tracing method. Then the Bland-Altman analysis was used to measure the agreement between the two methods (the EBS and the manual tracing method). It used a scatterplot (Bland–Altman plot) of the difference between the same measurements calculated by using the two methods against their average. We denote $d$ and $s$ to be the mean value and the standard deviation of the difference between the two methods. Then the 95% confidence interval CI is $(d - 1.96s, d + 1.96s)$. In the distance analysis, the distance was used to measure the spatial error between the motion trajectories of the carotid artery wall obtained separately by the EBS and the manual tracing method.

#### 3.4.1. Accuracy and robustness

As regards the accuracy, Fig. 9 and Table 2 illustrate the results of linear regression analysis over all the tests. The correlation coefficient $r$ is 0.9897 for RP and 0.9536 for LP, and the root mean square error (RMSE) is 25.9803 μm for RP and 142.8288 μm for LP. Then, Fig. 10 and Table 3 illustrate the results of Bland–Altman analysis over all the tests. The values of $s$ are $-11.35$ μm for RP and $-10.18$ μm for LP, and the confidence interval width are $89.62$ μm for RP and $387.26$ μm for LP. Finally, Table 4 shows the results of the distance analysis for all the tests $D_1$, $D_2$, ..., $D_{18}$. The values of Euclidean distance in the radial direction, longitudinal direction, and 2D plane are $33.99 \pm 19.23$ μm, $135.76 \pm 87.83$ μm and $148.40 \pm 84.86$ μm, respectively; the values of $f$-distance is $3.22 \pm 3.77$ μm.

As regards the robustness, Table 2 shows the results of the linear regression analysis for each test. The standard deviations of $r$ and RMSE are $0.0014$ and $4.4822$ μm for RP, and $0.0047$
Fig. 8. (a) Example of one training process repeated M times (M = 10) with different parameter initialization. This figure shows that the variation of e (defined in Eq. (23)) against the iteration number. Each marker corresponds to the training process with a kind of initialization, and the arrow on each marker shows the increase (dark green arrow) or decrease (blue arrow) of e between the two neighboring iterations. The green dotted line shows the level of the convergence threshold $\varepsilon = 0.1$. When $e < \varepsilon$ or the iteration reaches M times, the corresponding training process is terminated (without the arrow). Then, the parameters with the lowest value of e are considered as the optimal parameters in this training process. In this example, the training process which finally reaches the optimal parameter values is showed by the red dashed curve. (b) The variation of e in all eighteen training processes. The boxes (left y-axis) show the dispersion of e of all training processes. One iteration shown in x-axis denotes the process from Eq. (17) to Eq. (23). The green dotted line (left y-axis) shows the level of the convergence threshold $\varepsilon = 0.1$. The magenta curve (right y-axis) shows the variation of number of the training processes without convergence, with the iteration number. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 9. Results of the linear regression. In the figure, (a) and (b) show the scatter plots of our approach against the manual tracing method for all test processes for the radial motion and the longitudinal motion, respectively. In order to clearly present the distribution of scatters in (a) and (b), the density or the frequency of these scatters in the two-dimensional plane are displayed as the density images in (c) and (d), respectively. The green lines are the linearly fitting lines of the scatter points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Table 2
Results of the linear regression analysis. In this table, b is the slope of the fitting straight line and a is the intercept. r is the Pearson’s correlation coefficient. RMSE is the root-mean-square error.

<table>
<thead>
<tr>
<th>test</th>
<th>RP</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>1</td>
<td>0.9965</td>
<td>17.5657</td>
</tr>
<tr>
<td>2</td>
<td>1.0109</td>
<td>13.5823</td>
</tr>
<tr>
<td>3</td>
<td>1.0003</td>
<td>6.8971</td>
</tr>
<tr>
<td>4</td>
<td>0.9941</td>
<td>17.5493</td>
</tr>
<tr>
<td>5</td>
<td>0.9964</td>
<td>17.9671</td>
</tr>
<tr>
<td>6</td>
<td>1.0005</td>
<td>16.3005</td>
</tr>
<tr>
<td>7</td>
<td>0.9900</td>
<td>12.5900</td>
</tr>
<tr>
<td>8</td>
<td>0.9923</td>
<td>11.6905</td>
</tr>
<tr>
<td>9</td>
<td>1.0100</td>
<td>4.0609</td>
</tr>
<tr>
<td>10</td>
<td>0.9906</td>
<td>15.5701</td>
</tr>
<tr>
<td>11</td>
<td>0.9943</td>
<td>14.7409</td>
</tr>
<tr>
<td>12</td>
<td>1.0069</td>
<td>9.2494</td>
</tr>
<tr>
<td>13</td>
<td>1.0107</td>
<td>5.4623</td>
</tr>
<tr>
<td>14</td>
<td>1.0133</td>
<td>8.0244</td>
</tr>
<tr>
<td>15</td>
<td>1.0182</td>
<td>6.6137</td>
</tr>
<tr>
<td>16</td>
<td>0.9870</td>
<td>13.3857</td>
</tr>
<tr>
<td>17</td>
<td>1.0065</td>
<td>11.9144</td>
</tr>
<tr>
<td>18</td>
<td>1.0111</td>
<td>11.5441</td>
</tr>
<tr>
<td>Total</td>
<td>1.002</td>
<td>11.8612</td>
</tr>
</tbody>
</table>

* p-value < 0.01.

Fig. 10. The results of the Bland–Altman analysis of the average between our approach and the manual tracing method against their difference for all test processes. The red lines show the level of the mean value d and the green lines show the level of 1.96× standard deviation s. Fig (a) and (b) show the results for the radial motion and the longitudinal motion, respectively. Fig (c) and (d) show the density or the frequency of the scatters in (a) and (b) respectively, for the clear observation of the scatter distribution in the two-dimensional plane. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and 9.3100 μm for LP. Then Table 3 shows the results of the Bland–Altman analysis for each test. The standard deviation of d and the confidence interval width are 3.59 μm and 6.40 μm for RP, and 9.74 μm and 32.28 μm for LP. Next, Table 4 shows the results of distance analysis for each test. The standard deviation of the mean Euclidean distance in the radial direction, longitudinal direction, and 2D plane are 0.47, 6.86 and 7.41 μm, respectively; the standard deviation of mean f-distance is 0.20 μm. In the next, we tested the performance variation of our approach on the ultrasound sequence with different frame rates (24–51 frames/second) by the f-distance. The results show that the maximum variation of the performance of our approach is 2.13 μm. This implies that
the different frame rates little influence our approach. Then, we artificially added the different levels (5 dB to 25 dB signal-to-noise ratio) of Rayleigh noise to our ultrasound dataset in order to test the noise resistance of EBS and retested EBS. Fig. 11 indicates that the EBS performs well on the noise-corrupted ultrasound sequences. In addition, we used the Mann-Whitney U test to analyse the performance difference of the EBS on the left and right carotid arteries. The null hypothesis is that the mean results of EBS separately tested on the left and right carotid arteries have no significant difference. The results in Table 5 show that the p-value of the Mann-Whitney U test was greater than 0.82 for LP and 0.63 for RP. Finally, we have retested our approach on healthy subjects, unhealthy subjects and the subjects with plaque in the unhealthy subjects. The results show that the error of our approach (measured by f-distance) tested in the healthy subjects, unhealthy subjects and the subjects with plaque in the unhealthy subjects are 3.28 ± 3.99 μm, 3.19 ± 3.66 μm and 3.26 ± 3.83 μm respectively, and there is no significant difference among the three error values (significance level is 0.5). This result can demonstrate the generalization of our approach in different subjects.

In addition, we have also tested EBS with extreme parameter settings. In each parameter setting, we set an extreme value to one parameter and keep other parameters unchanged, and then executed EBS in order to obtain the error of motion trajectory (measured by f-distance). For the two parameters \( r_{lp} \) and \( r_{rw} \), only the extreme large value was considered, because the size of the search region should be larger than the block size. For each of the remaining parameters, extreme large value and extreme small value were both considered. Table 6 shows the performance of EBS with extreme parameter settings. The results show that the extreme values of parameters can lead to very poor performance of EBS. This implies the sensitive of EBS to the extreme parameter values, and moreover indicates the robustness of EBS to the variation of the initial parameters due to our parameter training strategy.
Table 5
Statistic significant difference of our approach applied in the left and right carotid artery walls. The null hypothesis is that the mean results of our approach separately tested on the left and right carotid artery walls are same. The difference is considered not significant when the p-value is greater than 0.05.

<table>
<thead>
<tr>
<th>test</th>
<th>statistical significance (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP</td>
<td>LP</td>
</tr>
<tr>
<td>1</td>
<td>0.95 0.97</td>
</tr>
<tr>
<td>2</td>
<td>0.88 0.64</td>
</tr>
<tr>
<td>3</td>
<td>0.83 0.91</td>
</tr>
<tr>
<td>4</td>
<td>0.88 0.87</td>
</tr>
<tr>
<td>5</td>
<td>0.99 0.63</td>
</tr>
<tr>
<td>6</td>
<td>0.84 0.94</td>
</tr>
<tr>
<td>7</td>
<td>0.83 0.96</td>
</tr>
<tr>
<td>8</td>
<td>0.85 0.89</td>
</tr>
<tr>
<td>9</td>
<td>0.98 0.89</td>
</tr>
<tr>
<td>10</td>
<td>0.88 0.99</td>
</tr>
<tr>
<td>11</td>
<td>0.90 0.73</td>
</tr>
<tr>
<td>12</td>
<td>0.91 0.95</td>
</tr>
<tr>
<td>13</td>
<td>0.92 0.94</td>
</tr>
<tr>
<td>14</td>
<td>0.82 0.94</td>
</tr>
<tr>
<td>15</td>
<td>0.93 0.86</td>
</tr>
<tr>
<td>16</td>
<td>0.97 0.93</td>
</tr>
<tr>
<td>17</td>
<td>0.87 0.89</td>
</tr>
<tr>
<td>18</td>
<td>0.92 0.90</td>
</tr>
<tr>
<td>Total</td>
<td>0.86 0.77</td>
</tr>
</tbody>
</table>

Table 6
Errors of our approach measured by f-distance (μm) when using the extreme parameter values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>f-distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>0.1</td>
<td>579.87</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>700.24</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>314.61</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>261.31</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>68.67</td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td>48.88</td>
</tr>
<tr>
<td>h</td>
<td>0.001</td>
<td>95.41</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>76.24</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>132.51</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>67.54</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>63.91</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>46.89</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>86.23</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>47.54</td>
</tr>
<tr>
<td>h</td>
<td>0.1</td>
<td>149.76</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>54.55</td>
</tr>
<tr>
<td>h</td>
<td>10</td>
<td>366.02</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>433.51</td>
</tr>
</tbody>
</table>

3.4.2. Comparison with the state-of-the-art methods

In order to perform a reliable comparison with the state-of-the-art methods, we implemented EBS, our preliminary work (denoted by HBM) (Gao et al., 2015) and the three kinds of popular methods by using Matlab R2012a on a desktop computer with Intel(R) Xeon(R) CPU E5-2650(2 GHz) and 32GB DDR2 memory:

Conventional block matching method: In the BM method, the prediction of the reference block was implemented by considering the reference block in the next frame as the best-matched block in the current frame. We compared EBS with the conventional block matching method (CBM) (Golemati et al., 2003), the echo tracking method (ETM1) (Cinthio et al., 2005) and ETM2 (Cinthio et al., 2006), and the guided speckle tracking method (GST) (Zahnd et al., 2011).

Kalman-based block matching method: Kalman filter was combined with the BM method in three different ways. The first one (denoted by K0) was to use it to solve the state-space equations of the motion of the carotid artery wall, in order to estimate the position of the reference block. In the other two ways, the results of the motion tracking were corrected during the tracking process (denoted by K1) and after the tracking process (denoted by K2) by the Kalman filter. By combining the three ways, seven block matching methods with Kalman filter could be applied (Gastounioti et al., 2011): the BM methods with K0 (KBM1), with K0 and K1 (KBM2), with K0 and K2 (KBM3), with K0, K1 and K2 (KBM4), with K1 (KBM5), with K2 (KBM6), with K1 and K2 (KBM7). In addition, the Kalman filter with control signal (KBM8 (Zahnd et al., 2013) and KBM9 (Tat et al., 2015)) were also applied to be combined with the BM method for tracking the motion of the carotid artery wall.

Optical flow: Optical flow is a method for tracking the instantaneous motion of intensity points in a sequence of images. One optical flow method was used to compare with our approach: Horn–Schunck method (OP) (Horn and Schunck, 1981; Golemati et al., 2012).

Table 7 presents the comparative results between EBS and the state-of-the-art methods. The results show that EBS performs better than these other methods.

3.4.3. Comparison between f-distance and Euclidean distance

Fig. 12 compares the f-distance computed by the evaluation function f proposed in Section 2.3 with the Euclidean distance used in the previous methods. Fig. 12(a) shows the example of two blocks with the same block size. Fig. 12(b) displays the change of f-distance (blue curve) and Euclidean distance (red curve) when the block size changes. Fig. 12(c) illustrates the change of f-distance (blue curve), weight Euclidean distance (green star) and minimum distance (magenta curve) against the center-to-center Euclidean distance (red curve) when the block size is unchanged.

3.4.4. Computational cost

We computed the variation of the computational cost against the performance of EBS (measured by f-distance) in different degrees of the image interpolation (1 ~ 10). Table 8 shows that when the interpolation degree increases, the computational cost increases but the performance of EBS decreases. In addition, we have compared the computational cost between EBS and the previous methods for different degrees of the image interpolation. Fig. 13(a) shows the change of the average computational cost per frame (second) against the increase of the interpolation scale. Fig. 13(b) shows the change of the standard deviation of the computational cost per frame (second) against the increase of the scale of the interpolation. The results show that our approach is highly efficient in the motion tracking of the carotid artery wall.

4. Discussion

In this section, we discuss the accuracy and robustness of our EBS approach, as well as the initialization issues, computational costs and the comparison with other methods.

4.1. Accuracy and robustness

We have evaluated the accuracy of EBS from 280 ultrasound sequences of 140 subjects by comparing the measurements (RP and LP) separately computed by the two methods: EBS and the manual tracing method performed by three experienced ultrasound physicians (ground truth). The comparative results of RP and LP show that EBS is highly correlated (correlation coefficient is 0.9897 for RP and 0.9536 for LP), is well agreed (average bias is 16.43 μm for RP and 27.47 μm for LP) with the manual tracing method, as well as having low-level error measured by both the Euclidean distance.
Table 7
Comparison with the previous methods. In the table, $e_r$, $e_l$ and $e_t$ represent the errors for the radial motion, longitudinal motion and 2D motion measured by the center-to-center Euclidean distance. $f_j$ is the error measured by the $f$-distance.

<table>
<thead>
<tr>
<th>Method</th>
<th>Euclidean distance ($\mu m$)</th>
<th>$e_r$</th>
<th>$e_l$</th>
<th>$e_t$</th>
<th>$f_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBM</td>
<td>54.74 ± 76.10</td>
<td>216.67 ± 219.08</td>
<td>234.88 ± 220.37</td>
<td>7.70 ± 9.24</td>
<td></td>
</tr>
<tr>
<td>ETM$_1$</td>
<td>107.56 ± 231.97</td>
<td>463.12 ± 323.79</td>
<td>485.44 ± 383.02</td>
<td>392.53 ± 385.23</td>
<td></td>
</tr>
<tr>
<td>ETM$_2$</td>
<td>113.86 ± 272.76</td>
<td>473.64 ± 383.49</td>
<td>490.57 ± 370.85</td>
<td>379.47 ± 401.38</td>
<td></td>
</tr>
<tr>
<td>GST</td>
<td>71.05 ± 103.60</td>
<td>238.13 ± 248.61</td>
<td>264.57 ± 253.58</td>
<td>6.22 ± 7.66</td>
<td></td>
</tr>
<tr>
<td>KBM$_1$</td>
<td>69.72 ± 102.56</td>
<td>237.19 ± 244.10</td>
<td>263.20 ± 248.90</td>
<td>5.11 ± 4.45</td>
<td></td>
</tr>
<tr>
<td>KBM$_2$</td>
<td>82.20 ± 114.01</td>
<td>243.85 ± 259.71</td>
<td>277.51 ± 263.93</td>
<td>5.03 ± 4.36</td>
<td></td>
</tr>
<tr>
<td>KBM$_3$</td>
<td>56.16 ± 78.43</td>
<td>224.14 ± 222.05</td>
<td>242.90 ± 223.28</td>
<td>4.61 ± 4.63</td>
<td></td>
</tr>
<tr>
<td>KBM$_4$</td>
<td>53.22 ± 73.60</td>
<td>223.87 ± 239.36</td>
<td>240.37 ± 240.59</td>
<td>5.81 ± 8.43</td>
<td></td>
</tr>
<tr>
<td>KBM$_5$</td>
<td>56.14 ± 78.02</td>
<td>223.01 ± 220.64</td>
<td>241.76 ± 221.83</td>
<td>5.58 ± 8.30</td>
<td></td>
</tr>
<tr>
<td>KBM$_6$</td>
<td>53.76 ± 76.52</td>
<td>213.30 ± 233.15</td>
<td>230.40 ± 235.63</td>
<td>4.68 ± 6.90</td>
<td></td>
</tr>
<tr>
<td>KBM$_7$</td>
<td>59.11 ± 84.24</td>
<td>253.10 ± 270.63</td>
<td>271.95 ± 271.90</td>
<td>5.65 ± 5.39</td>
<td></td>
</tr>
<tr>
<td>KBM$_8$</td>
<td>36.41 ± 45.88</td>
<td>178.02 ± 318.96</td>
<td>188.87 ± 318.15</td>
<td>4.60 ± 4.64</td>
<td></td>
</tr>
<tr>
<td>KBM$_9$</td>
<td>42.87 ± 61.89</td>
<td>227.62 ± 506.72</td>
<td>237.32 ± 506.24</td>
<td>5.91 ± 8.29</td>
<td></td>
</tr>
<tr>
<td>OP</td>
<td>61.97 ± 88.26</td>
<td>242.90 ± 259.78</td>
<td>263.52 ± 262.05</td>
<td>5.50 ± 4.44</td>
<td></td>
</tr>
<tr>
<td>HBM</td>
<td>35.84 ± 43.76</td>
<td>166.56 ± 210.45</td>
<td>172.11 ± 233.96</td>
<td>4.41 ± 3.98</td>
<td></td>
</tr>
<tr>
<td>Our approach</td>
<td>33.99 ± 19.23</td>
<td>135.76 ± 87.83</td>
<td>148.40 ± 84.86</td>
<td>3.22 ± 3.77</td>
<td></td>
</tr>
<tr>
<td>Intra-observer</td>
<td>24.16 ± 15.58</td>
<td>107.44 ± 87.33</td>
<td>114.80 ± 87.79</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Inter-observers</td>
<td>42.39 ± 24.95</td>
<td>159.30 ± 135.63</td>
<td>172.92 ± 135.91</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 12. An example of the comparison between the center-to-center Euclidean distance and the $f$-distance: (a) Two blocks (solid block and dashed block) with the same size for the distance measurement. (b) The change of the Euclidean distance (red curve) and $f$-distance (blue curve) when the block size changes but the center-to-center Euclidean distance between the two blocks is unchanged. (c) The change of the Euclidean distance (red curve), $f$-distance (blue curve), weight Euclidean distance (green star) and minimum distance (magenta curve) when the center-to-center Euclidean distance between the two blocks changes but their size is unchanged. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 8
Computational cost (second) against the performance of EBS measured by $f$-distance ($\mu m$) in different degrees of the image interpolation (1 ~ 10).

<table>
<thead>
<tr>
<th>Interpolation degree</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$-distance</td>
<td>–</td>
<td>0.03</td>
<td>0.08</td>
<td>0.13</td>
<td>0.19</td>
<td>0.28</td>
<td>0.37</td>
<td>0.50</td>
<td>0.64</td>
<td>0.90</td>
</tr>
<tr>
<td>Distortion</td>
<td>–</td>
<td>9.24</td>
<td>5.85</td>
<td>5.76</td>
<td>4.20</td>
<td>3.87</td>
<td>3.64</td>
<td>3.49</td>
<td>3.37</td>
<td>3.29</td>
</tr>
</tbody>
</table>
block during the cardiac cycle. This characteristic mainly results from the carotid movement and the out-of-plane motion, and may cause the target tissue to have the significantly low contrast, or even temporarily disappear in the ultrasound sequence. From the viewpoint of the state-space equations, the brightness change and the low contrast it causes can improve the error of the observation (the variable y shown in Eq. (8)), and thus increase the estimation error of the tissue appearance (the variable x). In our approach, the use of state-space equations with the linear elastic model can deal with the problem of brightness change. This can lead the target block to be prevented from heavily deviating from its initial state due to the brightness change. In addition, the NCC in the updating step can also be robust to the brightness change and the low contrast it causes, because of its normalization process.

### 4.2. Initialization issues

The initialization of EBS includes the location of the first frame sampled in the ultrasound sequence, and the determination of the parameters used in EBS. The sample of the first frame in the ultrasound sequence is crucial, because the first frame contains the reference of the target tissue. It provides the information of the block’s initial state, and the elastic model can drive the block toward the initial state when the block is deviated from its initial state. For the elastic model, it is required to sample the first frame when the carotid artery wall is relaxed. Because of the difficulty to measure the tension in the carotid artery wall in the gray-scale information, recognizing the frame corresponding to the relaxed state of the carotid artery wall is difficult. To overcome this problem, we collected the ECG gating signal when sampling the ultrasound data. Because the tension in the carotid artery wall is less influenced by the blood pressure and pulse wave at a low level of aortic blood pressure, we can consider that carotid artery wall stays in the relaxed state during the QRS duration in the ECG signal. However, the QRS duration may be uncertainly observed in the ultrasound, as shown in Fig. 3. Therefore, we sampled the first frame of the ultrasound sequence at the time around the R-peak in the ECG wave. From another viewpoint, this selection of the first frame also meets the clinical protocol of obtaining atherosclerotic predictors (suggested by the Rotterdam study (Bots et al., 1997)), such as carotid intima-media thickness and flow-mediated dilation (Kemp et al., 2016; Shaikh et al., 2016).

As regards the parameters determination, the previous works initialized the parameters manually by experts based on their professional experiences. This may cause errors due to the expert’s biases, and thus led to the acquirement of automatic parameter determination. A simple and direct strategy to automatically determine the parameter is the exhaustive search method (Ross, 2005). It tests the tracking performance by using every configuration in the parameter space, and then determining the optimal configuration. However, the exhaustive search method is often not adopted, especially for the large-size database, because it can involve enormous computational costs and thus make the parameter determination intractable. In contrast, the EBS applies a greedy strategy to training the parameters. We denote n as the number of samples along each dimension in the parameter space, and m as the number of parameters. Compared to the exhaustive search method, our strategy decreases the computational complexity from $O(n^m)$ to $O(nm)$. Compared to the gradient-based greedy search method (such as steepest-descent method and Newton method), this training strategy can ensure that the rate of convergence is reasonable even when the initial approximation or the co-ordinate directions are poor, and moreover significantly reduce the computational costs due to the avoidance of the derivative computation. However, the parameter optimization in our strategy may be trapped in the local optimum, rather than end up in the global optimum of the parameters. To address this problem, we repeated the training process ten times with different initial values, and considered the optimal results in these training processes as the final parameters used in the EBS. From Fig. 7, we can see that the dispersion degrees of all the ten parameters are reduced after the training process. This means that all parameters can converge to similar numbers. Although these similar configurations of the parameters can still be considered as the different local optimums, our experiments show that these local optimums have little influence the performance of our approach with respect to correlation, agreement, Euclidean distance and f-distance (see Tables 2–4). This can indicate that the local optimum issue is not severe after using our parameter strategy.

### 4.3. Computational costs

The main factor affecting the computational cost of the tracking method is the scale of interpolation. Many previous works set

![Fig. 13](image_url) Comparison of the computational costs between our approach (EBIM) and the previous methods: (a) The change of the average computational cost per frame [second] against the increase of the interpolation scale. (b) The change of the standard deviation of the computational cost per frame [second] against the increase of the interpolation scale.
the scale to be greater than one, in order to measure the sub-pixel motion of the carotid artery wall for acquiring a higher tracking accuracy. However, higher computation cost will incur when the scale of interpolation increases. Fig. 13 illustrates the scale of interpolation against the computational cost. The results show that the computational cost of EBS and all the previous methods increase 8.24 ~ 44.43 times when the scale of interpolation increases from one to ten. In addition, the computational cost of EBS (0.90 ± 0.02 second per frame) is moderate compared with all other methods. It is however a little higher than most Kalman-based methods (KBM1 ~ KBM3) and our previous method (HBM) due to the large-size state matrix used in the dynamic equations shown in Eq. (9) and (11). This leads to the increase of the computational costs in the matrix multiplication and in the numerical solution of the Riccati equation shown in Eq. (13). Nevertheless, EBS considerably reduces the computational cost in comparison with the manual tracing method. Moreover, the computational efficiency can be further improved if we re-implement the overall algorithm of EBS by using C/C++ programming language and GPU acceleration technology.

### 4.4. Comparison with other methods

We compared the EBS with three kinds of state-of-the-art methods for tracking the motion of the carotid artery wall. Due to absence of public database for testing the accuracy of the motion tracking method, these methods were re-implemented and tested by using the same programming language and our ultrasound study data for this comparison. The EBS mainly differs from these previous methods in three aspects: the determination of the block size, the strategy to estimate the tissue location in the next frame, the measurement method of the error between the computed-aided tracking method and the manual tracing method.

The determination of the block size (reference block and search region) should consider two factors: the ratio of the block width to the block height $\tilde{r}_b$ and their values. In the carotid ultrasound imaging, the longitudinal resolution is usually not greater than the radial resolution. Thus, the block should include more pixels along the longitudinal direction in order to obtain the same amount of tissue information as that along the radial direction. Furthermore, the carotid artery wall has more homogeneity along the longitudinal direction than in the radial direction. This indicates that the tracking method can more easily capture the radial motion of the carotid artery wall. Therefore, in most of the previous methods, the block width was set larger than the block height in order to obtain sufficient longitudinal information for improving the tracking performance ($\tilde{r}_b$: 3.88 ± 3.82). Then considering the amount of block size (height and width), there is a tradeoff between the accuracy and robustness of the tracking method. Due to the point spread effect of the ultrasound scanner, a relatively small-size block can facilitate improving the tracking accuracy (Cinthio et al., 2005). Nevertheless, the tracking method with small-size blocks is sensitive to image noise, the out-of-plane motion and the movements of the subjects and the ultrasound scanner during examination. Whereas, increasing the block size can improve the robustness, it reduces the accuracy and increases the computational cost. Therefore, it is significant to choose an appropriate block size for achieving the tradeoff balance between accuracy, robustness and the computational cost. In the previous methods, the determination of the ratio of the block width to the block height and their values were performed manually by the physician expert. Combining with the expert’s experience, the tracking method was in general tested many times with different block sizes, with the size leading to the best tracking performance being considered as the final determination. This however highly relies on the expert’s experience, and thus may reduce the performance of EBS due to the potential expert’s bias. Differing from the previous studies, our EBS has employed a training strategy to automatically compute the block size in order to overcome the disadvantages of the manual determination. Especially, in order to make the block width greater than the block height, we have forced that the ratio $\tilde{r}_b$ cannot get a value smaller than one in the parameter updating of the proposed training strategy.

As regards the strategy to estimate the tissue location in the next frame, the EBS was compared with three kinds of methods. The conventional BM method only used the BM method to predict the motion of the target tissue within the block. Its performance may be reduced by the disturbance such as image noise and out-of-plane motion. In order to improve the disturbance resistance, the Kalman-based BM method applied the state-space equations of the block motion. It represented the block’s information (location and/or gray-scale value) by the state variable in each frame, and solved the state-space equations by the Kalman filter. Based on the conventional BM method and the Kalman-based method, EBS applied the state-space equation produced by the linear elastic model, in order to tracking the motion of the carotid artery wall. The use of the linear elastic model can reduce the disturbance from the bright change of the target tissue on the carotid artery wall. Furthermore, EBS has adopted the $H_\infty$ filter to obtain the optimal estimate of the state of the target block in each ultrasound frame. The $H_\infty$ filter does not require knowing the noise distribution and statistics in advance, and has moreover performed better than the Kalman filter on the noise resistance due to its properties of the $H_\infty$ norm. In addition, the optical flow-based method is another widely used motion tracking approach which does not rely on the BM method, and it tracks the motion of the carotid artery wall by computing the frame-to-frame optical flow.

As regards measuring the error between the two best-matched blocks separately acquired from the computer-aided tracking method and manual tracing method, the previous works applied the Euclidean distance between the centers of the two blocks (center-to-center Euclidean distance). In contrast, the EBS employed one measurement of the error (called $f$-distance). The $f$-distance is not only related to the center-to-center Euclidean distance, but is also influenced by the block size and the overlapping degree between the two blocks. There are the two good properties of the $f$-distance. Also, we have employed an experiment to validate these properties by comparison with the center-to-center Euclidean distance, as shown in Fig. 12. Firstly, the $f$-distance can keep the block size in the appropriate interval. The use of the block with very small size may make the tracking method highly sensitive to the image noise, and the spatial distance between such two blocks can be considered to be distant with respect to the block size. Further, the use of the block with very large size may increase the computational cost of the tracking method. In Fig. 12(b), we can see that the appropriate interval of the block size varies within 15 and 125 range, because the $f$-distance within this interval is smaller than the center-to-center Euclidean distance. In contrast, the error measured by the $f$-distance will increase when the block size is outside this interval. Thus, the $f$-distance facilitates decreasing the influence of predicting the block motion from the image noise by avoiding the small-size block, and then reducing the computational cost of the tracking method due to the use of a large-size block. From the aspect of parameter updating in the training process, the values of the block size outside the appropriate interval in certain sub-regions in the parameter space, can increase the output value of the evaluation function $f$, as shown in Eqs. (17) and (21). Therefore, using the $f$-distance can keep the block size within the appropriate interval, as well as minimize the error measured by $f$-distance, as shown in Fig. 12(b). The other property of the $f$-distance is that it is more sensitive to the location change of the blocks than to the center-to-center Euclidean distance. Because the $f$-distance is the summation of the weight Euclidean distance and the minimum
distance [see Eq. (26)], we have displayed the $f$-distance (blue curve), weight Euclidean distance (green star) and minimum distance (black curve) against the center-to-center Euclidean distance (red curve) in Fig. 12(c). When the two blocks are overlapped, the minimum distance becomes zero and the weight Euclidean distance decreases faster than the center-to-center distance. Thus, the $f$-distance decreases faster than the center-to-center Euclidean distance. When the two blocks are separated, the minimum distance increases with the interval between the two blocks, and the weight Euclidean distance equals to the center-to-center Euclidean distance. Thus the $f$-distance increases at a higher speed than the center-to-center Euclidean distance. We can thereby consider the $f$-distance to have a better performance for the measurement of the location variance between the two blocks.

**Conclusion**

The motion of the carotid artery wall has been increasingly recognized to facilitate the evaluation of atherosclerotic disease in its early stage. In this study, we have proposed a useful method to track the two-dimensional motion of the carotid artery wall, by applying the state-space equations to describe the motion of the target tissue within the block. Also, based on the elasticity model, an extra constraint was developed to decrease the influence of brightness change of the target tissue during the processing. Moreover, the state-space equations were solved by combining the block matching method and the $H_{\infty}$ filter. In comparison with the Kalman filter used in the previous studies, the use of the $H_{\infty}$ filter makes the solution of the state-space equations more robust to the noise disturbance.

Finally, in order to reduce the reliance of the parameter determination on the expert’s experiences, we have proposed a training strategy to automatically determine the optimal values of the parameters used in our approach. The performance of our approach was evaluated on the ultrasound sequences of bilateral carotid arteries from 140 subjects. The results show that our approach has the same order of accuracy as the manual tracing method performed by the medical physicians, and can hence be deemed to be effective in the motion tracking of the carotid artery wall.

**Acknowledgement**

This work was supported in part by the Shenzhen Research and Innovation Funding (JCYJ20151030151431727, SGLH201502131432079) and the National Key Research and Development Program of China (2016YFC130302, 2016YFC1301702). We gratefully acknowledge the volunteers who participated in our study.

**Appendix A. Derivation from Eq. (7) to Eq. (8)**

We firstly rewrite Eq. (7) as follows,

$$\ddot{y} + \frac{k}{m} y = \frac{1}{m} u.$$  \hspace{1cm} (A.1)

This equation can describe the brightness change of the target tissue within the rectangle block on the carotid artery wall. Then we consider the block as a matrix with size of $p \times q$, and replace the $y$ and $u$ by bold letters $\mathbf{y}$ and $\mathbf{u}$ in order to indicate that the two variables are matrices. We denote the elements of $\mathbf{y}$ as $y_{11}, \ldots, y_{pq}$ and the elements of $\mathbf{u}$ as $u_{11}, \ldots, u_{pq}$. We thus have

$$\mathbf{y} = \begin{bmatrix} y_{11} & \cdots & y_{1q} \\ \vdots & \ddots & \vdots \\ y_{p1} & \cdots & y_{pq} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_{11} & \cdots & u_{1q} \\ \vdots & \ddots & \vdots \\ u_{p1} & \cdots & u_{pq} \end{bmatrix}.$$  \hspace{1cm} (A.2)

Hence, Eq. (A.1) can be reformulated by

$$\begin{pmatrix} y_{11} \\ \vdots \\ y_{p1} \end{pmatrix} = \begin{pmatrix} k \\ \vdots \\ m \end{pmatrix} \begin{pmatrix} y_{11} \\ \vdots \\ y_{pq} \end{pmatrix} + \begin{pmatrix} \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} y_{11} \\ \vdots \\ y_{pq} \end{pmatrix}$$

$$= \frac{1}{m} \begin{pmatrix} u_{11} \\ \vdots \\ u_{pq} \end{pmatrix}.$$  \hspace{1cm} (A.3)

Then for $\forall i \in \{1, 2, \ldots, p\}$, $\forall j \in \{1, 2, \ldots, q\}$, let $x_{ij} = y_{ij}$ and $\dot{e}_{ij} = \dot{y}_{ij}$, we can get the following equations from Eq. (A.3),

$$\dot{x}_{ij} = e_{ij},$$

$$\dot{e}_{ij} = -\frac{k}{m} x_{ij} + \frac{1}{m} u_{ij},$$

$$y_{ij} = x_{ij}. $$  \hspace{1cm} (A.4)

Eq. (A.4) can be rewritten into the following state-space representation,

$$\dot{\mathbf{x}} = A \mathbf{x} + \mathbf{b} u.$$  \hspace{1cm} (A.5)

and

$$\mathbf{x} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{q1} & \cdots & x_{qp} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_{11} & \cdots & y_{1q} \\ \vdots & \ddots & \vdots \\ y_{pq} \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} u_{11} & \cdots & u_{1q} \\ \vdots & \ddots & \vdots \\ u_{pq} \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathbf{O}_{p \times p} & \mathbf{I}_{p \times p} \\ -\frac{1}{m} \mathbf{I}_{p \times p} & \mathbf{O}_{p \times p} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{O}_{p \times p} \\ -\frac{1}{m} \mathbf{I}_{p \times p} \end{bmatrix}.$$  \hspace{1cm} (A.7)

where $\mathbf{I}_{p \times p}$ is the $p \times p$ identity matrix and $\mathbf{O}_{p \times p}$ is the $p \times p$ zero matrix.

In the next step, we discretize the above continuous-time representation of the state-space equations. By multiplying Eq. (A.5) on both sides by $e^{-A_t}$, we have

$$e^{-A_t} (\dot{\mathbf{x}} - A \mathbf{x}) = e^{-A_t} \mathbf{b} u$$  \hspace{1cm} (A.8)

In order to indicate that the variable $\mathbf{y}$ relies on the time $t$, $\mathbf{y}$ is replaced by $\mathbf{x}(t)$ in the following equations. The left-hand side of Eq. (A.8) can be reformulated into a derivative format as shown below,

$$\frac{d}{dt} (e^{-A_x} \dot{x}(t)) = e^{-A_x} \mathbf{b} u$$  \hspace{1cm} (A.9)

Then by following equation by integrating Eq. (A.9) in the range $[t_1, t_2]$, we can obtain the following equation,

$$\int_{t_1}^{t_2} \frac{d}{dt} (e^{-A_x} \dot{x}(t)) dt = \int_{t_1}^{t_2} e^{-A_x} \mathbf{b} u(t) dt.$$  \hspace{1cm} (A.10)

By Newton-Leibniz formula (Larson and Edwards, 2009), the left-hand side of Eq. (A.10) can be transformed into

$$e^{-A_x} \dot{x}(t)|_{t_1}^{t_2} = \int_{t_1}^{t_2} e^{-A_x} \mathbf{b} u(\eta) d\eta.$$  \hspace{1cm} (A.11)
and we can derive the following equation from Eq. (A.11).

$$\mathbf{x}(t_2) = e^{\mathbf{A}(t_2-t_1)}\mathbf{x}(t_1) + \int_{t_1}^{t_2} e^{\mathbf{A}(t_2-\eta)}\mathbf{Bu}(\eta)d\eta$$  \hspace{1cm} (A.12)

We denote $T$ to be the sampling period and $n$ as the frame index in the ultrasound sequence. Let $t_1 = nT$ and $t_2 = (n+1)T$, and thus $t_1$ and $t_2$ represent the time sampling of the $n$th and $n+1$th images, respectively. We assume that $\mathbf{u}(\eta)$ is a constant value with $\mathbf{u}(\eta) = \mathbf{u}(nT)$ when $\eta \in [t_1, t_2]$. By substituting $t_1 = nT$ and $t_2 = (n+1)T$ into Eq. (A.12), we can obtain

$$\mathbf{x}((n+1)T) = e^{\mathbf{A}T}\mathbf{x}(nT) + \int_{0}^{(n+1)T} e^{\mathbf{A}(n+1)T-\eta}\mathbf{Bu}(nT)d\eta$$  \hspace{1cm} (A.13)

Let $t = (n+1)T - \eta$, then Eq. (A.13) can be rewritten as

$$\mathbf{x}((n+1)T) = e^{\mathbf{A}T}\mathbf{x}(nT) + \int_{0}^{T} e^{\mathbf{A}\eta}\mathbf{Bu}(nT)d\eta$$  \hspace{1cm} (A.14)

Because $\mathbf{x}(nT)$ corresponds to the block's appearance in the $n$th frame in the ultrasound sequence, we replace $\mathbf{x}(nT)$ by $\mathbf{x}_n$, and then Eq. (A.14) can be reformulated into

$$\mathbf{x}_{n+1} = e^{\mathbf{A}T}\mathbf{x}_n + \left(\int_{0}^{T} e^{\mathbf{A}\eta}\mathbf{Bu}\right)\mathbf{u}_n$$  \hspace{1cm} (A.15)

Then, Eq. (A.6) can be discretized with respect to time into the equation

$$\mathbf{y}_n = \mathbf{C}\mathbf{x}_n$$  \hspace{1cm} (A.16)

Therefore, the discrete-time formulations of Eqs. (A.5) and (A.6) can be represented as follows,

$$\mathbf{x}_{n+1} = \mathbf{F}\mathbf{x}_n + \mathbf{G}\mathbf{u}_n$$  \hspace{1cm} (A.17)

where $\mathbf{F} = e^{\mathbf{A}T}$, $\mathbf{G} = \int_{0}^{T} e^{\mathbf{A}\eta}\mathbf{Bu}$ and $\mathbf{H} = \mathbf{C}$.

Finally, we need to derive the analytical form of $\mathbf{F}$ and $\mathbf{G}$, and they can be reformulated as follows,

$$\mathbf{F}(t) = \mathbf{F}(t)|_{t=T} = e^{\mathbf{A}T}$$  \hspace{1cm} (A.18)

$$\mathbf{G}(t) = \mathbf{G}(t)|_{t=T} = \int_{0}^{T} e^{\mathbf{A}\eta}\mathbf{Bu} |_{t=T} = \mathbf{B}$$  \hspace{1cm} (A.19)

Considering $\mathbf{F}(t)$, we can obtain the following equation by the inverse Laplace transform $\mathcal{L}^{-1}$.

$$\mathbf{F}(t) = \mathcal{L}^{-1}\{(s\mathbf{I} - \mathbf{A})^{-1}\}$$  \hspace{1cm} (A.20)

Then substituting Eq. (A.7) into Eq. (A.20), we have

$$\mathbf{F}(t) = \mathcal{L}^{-1}\left(\begin{bmatrix} s\mathbf{I}_{p \times p} & -\mathbf{I}_{p \times p} \\ \mathbf{I}_{p \times p} & s\mathbf{I}_{p \times p} \end{bmatrix}\right) = \mathcal{L}^{-1}\{\mathbf{V}\}$$  \hspace{1cm} (A.21)

where $s$ is the complex variable. Because the matrix $\mathbf{V}^{-1}$ is full rank, we can compute its inverse matrix by the Gaussian elimination method (Meyer, 2000), as follows,

$$\mathbf{V} = \begin{bmatrix} \frac{s}{s^2 + \frac{k}{m}} & \frac{1}{s^2 + \frac{k}{m}} \\ -\frac{k}{m} & \frac{s}{s^2 + \frac{k}{m}} \end{bmatrix}$$  \hspace{1cm} (A.22)

Then $\mathbf{F}(t)$ can be computed by the inverse Laplace transform of $\mathbf{V}$ as follows,

$$\mathbf{F}(t) = \begin{bmatrix} \cos\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p} & \frac{m}{k} \sin\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p} \\ -\frac{k}{m} & \frac{s}{s^2 + \frac{k}{m}} \end{bmatrix}$$  \hspace{1cm} (A.23)

Then by substituting Eq. (A.23) into Eq. (A.18), the analytical form of $\mathbf{F}$ is obtained as

$$\mathbf{F} = \begin{bmatrix} \cos\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p} & \frac{m}{k} \sin\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p} \\ -\frac{k}{m} & \frac{s}{s^2 + \frac{k}{m}} \end{bmatrix}$$  \hspace{1cm} (A.24)

Regarding $\mathbf{G}(t)$, we have

$$\mathbf{G}(t) = \int_{0}^{t} e^{\mathbf{A}\eta}\mathbf{Bu} |_{t=T} = \int_{0}^{T} \mathbf{Bu} = \mathbf{B}$$  \hspace{1cm} (A.25)

Then by substituting Eq. (A.25) into Eq. (A.19), the analytical form of $\mathbf{G}$ can be derived as,

$$\mathbf{G} = \begin{bmatrix} \frac{1}{k} \left(1 - \cos\left(\frac{\sqrt{m}}{\sqrt{k}}t\right)\right) \mathbf{I}_{p \times p} & \frac{\sqrt{\frac{m}{k}} \sin\left(\frac{\sqrt{m}}{\sqrt{k}}t\right)}{\sqrt{k}} \mathbf{I}_{p \times p} \end{bmatrix}$$  \hspace{1cm} (A.26)

In summary, based on Eqs. (A.5), (A.6), (A.7), (A.17), (A.24) and (A.26), the discrete-time formulation of state-space equations can be represented as follows,

$$\begin{align*}
\mathbf{x}_{n+1} &= \mathbf{F}\mathbf{x}_n + \mathbf{G}\mathbf{u}_n \\
\mathbf{y}_n &= \mathbf{H}\mathbf{x}_n
\end{align*}$$  \hspace{1cm} (A.27)

and

$$\begin{bmatrix} \mathbf{x}^{11}_n & \cdots & \mathbf{x}^{11}_n & \mathbf{x}^{12}_n & \cdots & \mathbf{x}^{1q}_n \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{x}^{11}_n & \cdots & \mathbf{x}^{n1}_n & \mathbf{x}^{12}_n & \cdots & \mathbf{x}^{nq}_n \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{x}^{11}_n & \cdots & \cdots & \cdots & \cdots & \mathbf{x}^{n1}_n \\
\mathbf{u}^{11}_n & \cdots & \cdots & \mathbf{u}^{11}_n & \cdots & \mathbf{u}^{1q}_n \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
\mathbf{u}^{11}_n & \cdots & \cdots & \cdots & \cdots & \mathbf{u}^{1q}_n \\
\mathbf{y}^{11}_n & \cdots & \cdots & \cdots & \cdots & \mathbf{y}^{1q}_n \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
\mathbf{y}^{11}_n & \cdots & \cdots & \cdots & \cdots & \mathbf{y}^{1q}_n \\
\mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 & \mathbf{F}_4 & \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix}^T & \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \end{bmatrix} \\
\mathbf{F}_1 &= \cos\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p}, \mathbf{F}_2 = \frac{\sqrt{m}}{k} \sin\left(\frac{\sqrt{m}}{\sqrt{k}}t\right) \mathbf{I}_{p \times p}
\[ F_3 = -\sqrt{\frac{k}{m}} \sin \left( \sqrt{\frac{k}{m}} \frac{t}{p_x,p_y} \right) \]

\[ F_4 = \cos \left( \sqrt{\frac{k}{m}} \frac{t}{p_x,p_y} \right) \]

\[ G_1 = \frac{1}{k} \left( 1 - \cos \left( \sqrt{\frac{k}{m}} \frac{t}{p_x,p_y} \right) \right) \]

\[ G_2 = \left( \frac{1}{mk} \sin \left( \sqrt{\frac{k}{m}} \frac{t}{p_x,p_y} \right) \right) \]

\[ H_1 = p_x,p_y \quad H_2 = O_{p_x,p_y} \]

where \( n \) is the frame index and \( T \) is the sampling period in the ultrasound sequence. Eqs. (8), (9) and (10) are obtained by means of Eqs. (A.27) and (A.28).

### Appendix B. Solution of Eq. (33)

We rewrite Eq. (33) as

\[ p_2 = m \left( \frac{d g(l)}{dl} \right) \]

and

\[ g(l) = 2\tilde{r}_l^2 \left( h_m - l + 1 \right) \left( w_m - \tilde{r}_l \right) + 1 \]

where \( h_l \) and \( w_l \) are the height and width of the ultrasound image, respectively. Expanding Eq. (B.2), we have

\[ g(l) = 2\tilde{r}_l^2 - 2\tilde{r}_l \tilde{h}_l h_m + h_m + w_m + 1 \]

\[ + 2\tilde{r}_l \left( h_m w_m + h_m + w_m + 1 \right) \]

In order to find the maximum value of the function \( g \) with respect to the variable \( l \), we make the first derivative of \( g(l) \) equal to zero, i.e.

\[ \frac{dg(l)}{dl} = 8\tilde{r}_l^2 \tilde{h}_l \tilde{w}_l - 2\tilde{r}_l \tilde{h}_l h_m + h_m + w_m + 1 = 0 \]

Since \( \tilde{r}_l \neq 0 \), we divide both sides of Eq. (B.4) by \( 2\tilde{r}_l \tilde{h}_l \), and get a quadratic equation

\[ 4\tilde{r}_l^2 \left( h_m w_m + h_m + w_m + 1 \right) + 2 \left( h_m w_m + h_m + w_m + 1 \right) = 0 \]

Eq. (B.5) has two solutions \( l_1, l_2 \) as follows:

\[ l_1 = \frac{3 \left( h_m w_m + h_m + w_m + 1 \right) + \sqrt{9 \left( h_m w_m + h_m + w_m + 1 \right)^2 - 32 \left( h_m w_m + h_m + w_m + 1 \right)}}{8 \tilde{r}_l} \]

\[ l_2 = \frac{3 \left( h_m w_m + h_m + w_m + 1 \right) - \sqrt{9 \left( h_m w_m + h_m + w_m + 1 \right)^2 - 32 \left( h_m w_m + h_m + w_m + 1 \right)}}{8 \tilde{r}_l} \]

Then the value of \( p_2 \) in Eq. (B.1) can be computed by substituting Eq. (B.6) into the following equation

\[ p_2 = \max \{ g(l) \} \]

### References


