Faster Search Algorithm for Speckle Tracking in Ultrasound Images

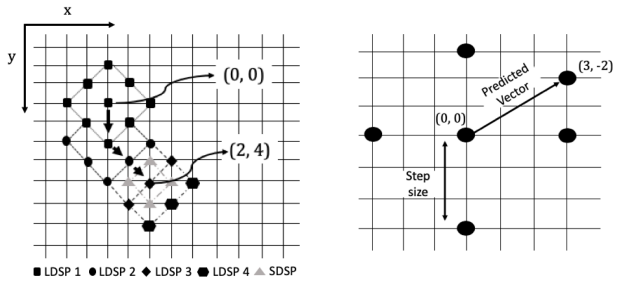
Skanda Bharadwaj[1], Mohamed Almekkawy[1]

***Abstract*— Block matching techniques have been studied exhaustively for motion estimation in Ultrasound (US) images. Exhaustive Search (ES) is the most commonly used search algorithm for block matching in US images. However, ES can be computationally expensive and slow. In this paper, a faster search algorithm called the Adaptive Rood Pattern Search (ARPS) is adopted to US images along with sub-pixel matching to reduce the computational cost and enhance block matching. Both ES and ARPS were applied in the context of block matching based 2D speckle tracking and were compared using Number of Computations per Frame (NCF), Computational Time per Frame (CTF) and Root Mean Squared Error (RMSE) as metrics. Our simulations and experimental results proved that ARPS outperformed ES by a substantial margin. Adaptation of this technique could help improve the performance of real-time motion estimation drastically.**

I. INTRODUCTION

Speckle tracking has wide range of applications in diagnostic ultrasound, such as elasticity imaging [1], elastography [2], and echocaridography [3]. Speckles, in US images, are formed by the combination of constructive and destructive interference of echoes from scatterers in the observed tissue and can be used for motion tracking [4]. One of the most commonly used algorithms for speckle tracking in US images is block matching [5]. In any typical block matching based tracking, a reference block (window/kernel) is defined in the first frame and is tracked in the subsequent frames. Any block in the subsequent frame that is subject to search is called the candidate block. The similarity between the reference block and the candidate block is quantified by defining a cost function, using which the best match is obtained. Different applications of block matching include tracking of carotid artery wall motion [6], shear strain and motion amplitude within the arterial wall [7] and study of motion dynamics of carotid atheromatous plaque [8].

Based on the block matching technique, several adaptive algorithms have been developed to improve motion tracking. Gastounioti et al [9] proposed an adaptive block matching algorithm using a linear Kalman filter. To further enhance motion tracking, Zahnd et al [10] proposed a methodology, which involved temporal update of the reference block using pixel-wise linear Kalman filter. In a very recent study, a non-linear state space approach was proposed by Gao et al [11], where unscented kalman filter with a non-linear model was used to estimate the motion of the carotid artery. Although, all the above-mentioned techniques promise robust and precise tracking, finding a best match in the subsequent frames is still expensive due to ES. ES attempts to find the best match by scanning all possible candidate blocks within a pre-defined search region, thus making it computationally expensive.



(a) Diamond Search (b) ARPS

Fig. 1. (a) represents diamond Search for location $(2, 4)$ using LDSP and SDSP. (b) represents the rood pattern distributed points of ARPS.

Besides ES, few other well-known faster algorithms are new three-step search (NTSS), four-step search(4SS), and diamond search (DS). Although these algorithms are efficient and simple, they lack adaptability. In this paper, we propose to adopt ARPS along with sub-pixel accurate matching. ARPS, an adaptive search algorithm proposed by Nie et al [12], is used to find the best candidate match in any given frame. Making use of the fact that the general motion in a frame is coherent, ARPS exploits the adjustable rood-shaped patterns adaptively. This reduces the computational cost drastically. To further enhance block matching, sub-pixel accurate matching is performed post ARPS matching by interpolating the matched blocked. We also employ a simple linear Kalman filter to mimic the conventional tracking models [9] used in US images.

II. MATERIALS AND METHODS

A. Basic Principles of Adaptive Rood Pattern Search

ARPS relies on the fact that neither large search patterns (window/kernel size) nor small search patterns alone can efficiently estimate the motion vectors [12]. While the small search patterns are useful in detecting small motions, they tend to get stuck in local minimum. On the other hand, large search patterns contain sparsely spaced search points and incur unnecessary search for small motion vectors. Therefore, ARPS exploits the use of both large and small search patterns. It makes use of the fact that, if a given frame has a particular set of motion vectors, then it is highly likely that the subsequent frame also has a similar set of motion vectors.

*This work was not supported by any organization

[1]Skanda Bharadwaj and Mohamed Almekkawy are with the School of Electrical Engineering and Computer Science, Penn State University, University Park, USA {ssb248, mka9}@psu.edu

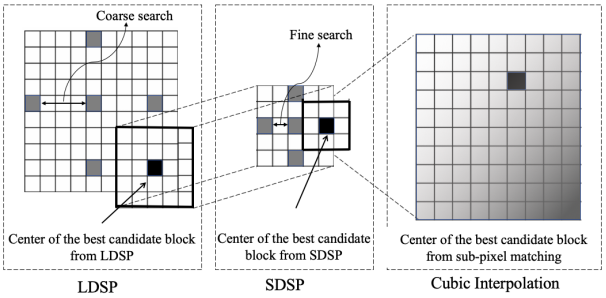


Fig. 2. Pipeline of block matching

ARPS has two diamond search patterns − Large Diamond Search Pattern (LDSP) and Small Diamond Search Pattern (SDSP)− much like the conventional diamond search. This is represented in Fig. 1a with an example of path to the motion vector (2, 4). In this example, ARPS uses four LDSP and one SDSP (it should be noted that the center of the reference block is considered $(0, 0)$). But, as opposed to diamond search, ARPS is adaptive in that it uses the predicted motion vector from the previous frame to check for the candidate block. In addition, it also searches at rood pattern distributed points as represented in Fig. 1b. ARPS first performs an LDSP and when a best match is obtained, it performs a fine tuning using SDSP. Once the SDSP selects the best candidate block, sub-pixel matching is performed by interpolating the candidate block. Fig. 2 represents the pipeline of the proposed block matching technique.

B. Kalman Filter

A simple linear Kalman filter is used to update the reference block. Prediction of the reference block is obtained from a pre-defined motion model. The best-matched candidate block obtained from the preceding frame is used as measurement. The state and measurement equations are given by

$$x_k = Fx_{k-1} + Bu_k + n_k \tag{1}$$

$$z_k = Hx_{k-1} + v_k \tag{2}$$

where x_k is the current state at time k, F is the state transition model, B is the control matrix, u is the control vector, H represents the measurement (or observation) model, n and v are the process and measurement noise respectively. Two phases of the Kalman filter are the *prediction* and *update* phases. Priori state estimate $(\hat{x}_k^-)$ and priori state covariance estimate (P_k^-) are calculated in the prediction phase using equations 3 and 4

$$\hat{x}_k^- = F\hat{x}_{k-1} + Bu_{k-1} \tag{3}$$

$$P_k^- = FP_{k-1}F^T + Q \tag{4}$$

where, Q represents the process covariance matrix. In the *update* phase, Kalman gain (K) is calculated to appropriately weight the error between observation and prediction using equation 5. Using Kalman gain and the available new measurement, the posterior estimate of the state $\hat{x}_k$ is obtained using equation 6. Finally, posterior estimate of the state covariance P_k is updated using equation 7.

$$K = P_k^- H^T (HP_k^- H^T + C)^{-1} \tag{5}$$

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \tag{6}$$

$$P_k = (I - KH)P_k^- \tag{7}$$

where, C in equation 5 represents the measurement covariance matrix. Since we do not consider any tissue-type in particular, we designed a generic simple linear Kalman filter to update the reference block. The pixel coordinates are governed by the kinematic equation of a free moving object. Displacement (x and y) and velocities ($\dot{x}$ and $\dot{y}$) of the pixels in lateral and axial directions are considered to be the states of the system and are together represented by $x_k = [x, y, \dot{x}, \dot{y}]^T$. Displacement S is governed by equation 8 and velocity v is governed by equation 9.

$$S = ut + \frac{1}{2}at^2 \tag{8}$$

$$v = u + at \tag{9}$$

where, u represents the initial velocity, t represents time and a represents acceleration. The process covariance matix Q is modeled as

$$Q = \begin{bmatrix} \frac{dt^2}{4} & 0 & \frac{dt^3}{2} & 0 \\ 0 & \frac{dt^2}{4} & 0 & \frac{dt^3}{2} \\ \frac{dt^3}{2} & 0 & dt^2 & 0 \\ 0 & \frac{dt^3}{2} & 0 & dt^2 \end{bmatrix}$$

The measurement covariance matrix C is considered to be an identity matrix multiplied by a constant term ($C = cI$). For our experiments, $dt = 0.5$, $a = 0.004m/s^2$ (arbitrarily low chosen values to mimick tissue movement) and for simplicity we set $c = 1$.

III. Experiments

Fig. 3 represents the schematic of our experiment. ARPS and ES were applied to six different carotid artery datasets (including cross-section and longitudinal-section view). Among the six datasets, one of them (referred to as CA1 dataset) was a B-mode image sequence of the perpendicular cuts of the carotid artery (cross-section view) obtained from [13], [14], [15]. The ground truth of the arterial movement for this dataset was known a priori. Fig. 4a represents one of the frames of CA1 dataset along with the region of interest (ROI). The red circle in the image represents the location of the blood vessel. This ROI was tracked within a search window defined by the parameter p, which is shown in Fig. 4b. For CA1 dataset, whole blood

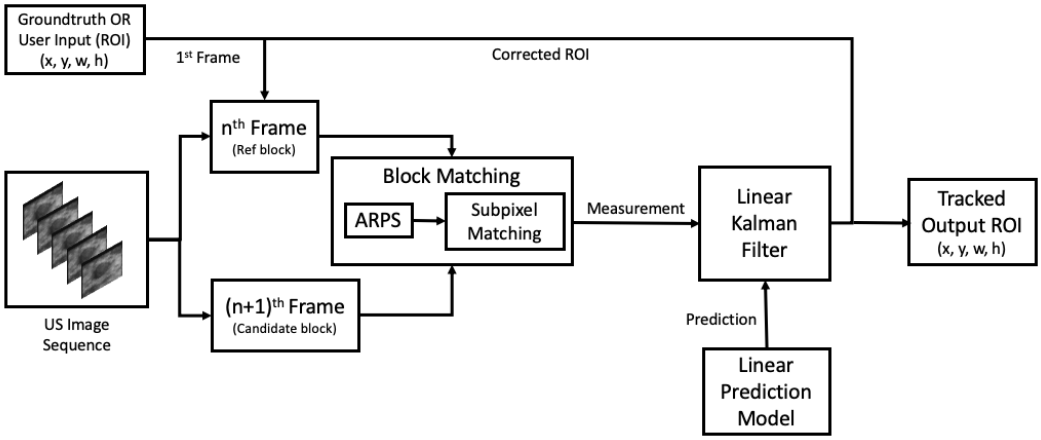
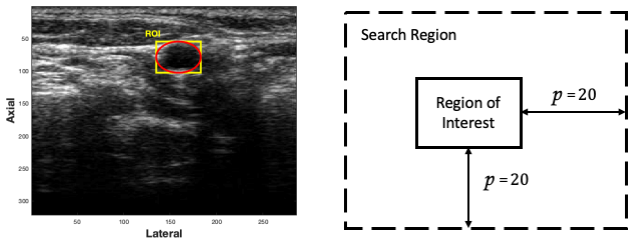


Fig. 3. Schematic of our proposed approach. Input ROIs are tracked using ARPS and ES with sub-pixel accurate matching.

vessel was tracked using the two search algorithms and was compared against the ground truth.

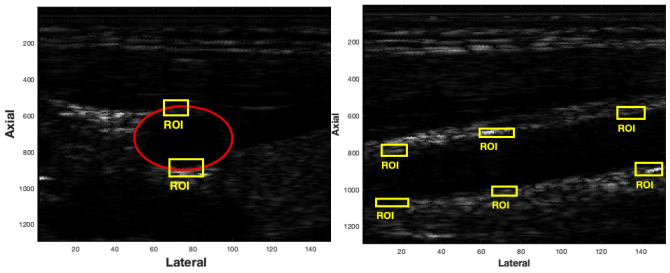
For CA1 dataset, the ROI available from the ground truth for the first frame was given as input to ARPS and ES and was tracked in the subsequent frames. The best match from the subsequent frames were obtained by performing similarity matching using normalized cross-correlation. Further, sub-pixel accurate matching was performed using cubic interpolation. After tracking through all the frames of the image sequence, the centroid locations of the tracked ROIs were compared against the centroid locations of the ground truth ROIs. For each frame, NCF and CTF were calculated. NCF is defined as the number of times the cost function is calculated for a given frame and CTF is defined as the total time taken by the block matching algorithm to process an entire frame. Once the entire sequence of images was processed, average NCF and CTF were compared between the two search algorithms. Finally, RMSE between centroids of the search algorithm and centroids of the ground truth were compared.

Rest of the five datasets (referred to as CA2 through CA6) were acquired using a Sonix RP (Ultrasonix, Canada) ultrasound scanner for high frame-rate data collection [16], [17] (the experimental procedures involved in collecting the data sets used in this paper are the same as [16], [17] and were approved by University of Minnesota. Please check the Acknowledgments section.) A linear array probe (LA14-5/38) with a center frequency of 7.5 MHz was used to acquire the data. Among the five datasets, two datasets (CA2 and CA3) were cross-section view, and the other three datasets (CA4, CA5 and CA6) were longitudinal-section view of the carotid artery. In each of the datasets, multiple ROIs containing intima-media complex and lumen are considered for tracking. Fig. 5a and 5b represent sample frames from datasets CA2 (cross-section view) and CA4 (longitudinal-section view), respectively, with the considered ROIs for tracking. It should be noted that all the image units are in pixels instead of mm (this helps to compare other figures and metrics such as centroid locations and RMSE introduced



(a) (b)

Fig. 4. (a) Sample frame from the CA1 dataset. The red circle represents the blood vessel and the yellow box represents the ROI. (b) represents the search region within which the algorithm operates to find the best match.



(a) (b)

Fig. 5. (a) represents a sample frame of CA2 dataset (cross-sectional view) and (b) represents a sample frame of CA4 dataset (longitudinal view) with chosen ROIs to be tracked. The red circle in (a) represents the circumference of the carotid artery.

in subsequent sections). NCF and CTF were compared between the two search algorithms. Since the ground truth for these datasets were unavailable, the RMSE of centroids is calculated between ARPS and ES. This metric signifies how close the two algorithms are to each other when the same ROI is tracked.

IV. Results and Discussion

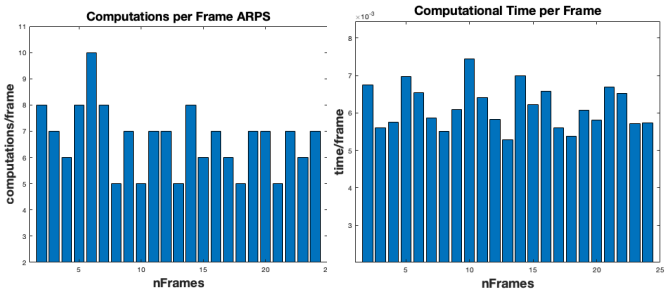
A. ARPS and ES on CA1 dataset

For the experiments on CA1 dataset, a search region with $p = 20$ pixels was defined for both ARPS and ES. Images were of the size (322×288) pixels. Due to the exhaustive

TABLE I

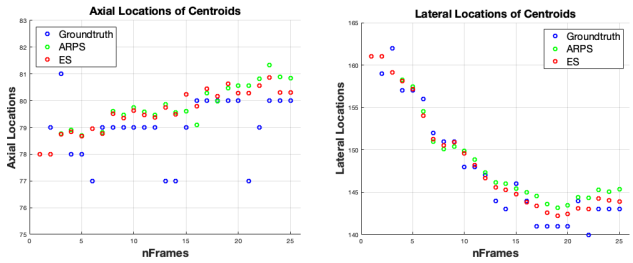
AVERAGE NCF AND CTF VALUES FOR CA1 DATASET.

	Average NCF	Average CTF
ES	1681	$650ms$
ARPS	5.67	$10ms$



(a) NCF (b) CTF

Fig. 6. NCF and CTF of ARPS for CA1 dataset.



(a) Axial direction (b) Lateral direction

Fig. 7. Centroid locations of the ground truth, ARPS and ES through all the frames of the CA1 dataset for axial and lateral directions.

TABLE II

RMSE VALUES IN PIXELS FOR ES AND ARPS WITH RESPECT TO THE GROUND TRUTH ON CA1 DATASET.

RMSE	ES	ARPS with subpixel	ARPS w/o subpixel
Axial	2.37	2.76	2.95
Lateral	2.39	2.32	5.24

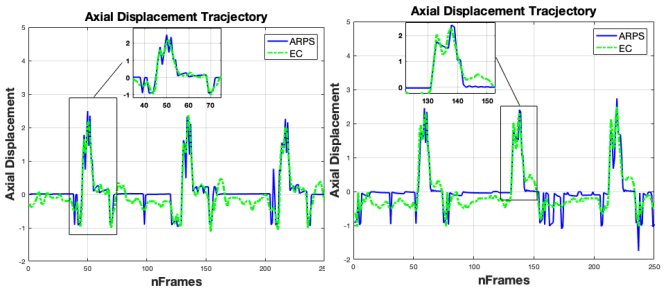
search strategy, NCF and CTF for ES are exactly the same for every frame. On the contrary, NCF and CTF reduce drastically when ARPS is used. Fig. 6a and 6b represent NCF and CTF for ARPS. Average NCF and CTF were calculated for both ARPS and ES and are summarized in Table 1.

The ROIs tracked by both algorithms were compared against the available ground truth. Fig. 7a and 7b represent the centroid locations of the ground truth, ARPS and ES in the axial and lateral directions, repectively. The RMSE for ES and ARPS with respect to the ground truth are summarized in Table 2. It can be seen that ARPS applied with sub-pixel accurate matching substantially decreases the error.

B. ARPS and ES on Datasets CA2 through CA6

Similar to above-mentioned experiments, both ARPS and ES were applied on datasets CA2 through CA6. Multiple ROIs from the same frame were selected in order to check the robustness of ARPS against ES and also to examine the movement of different locations axially and laterally of the same blood vessel. For instance, six regions were selected in CA4 dataset image as shown in Fig. 5b. Three regions along the upper wall and three regions along the lower wall of the carotid artery. These ROIs contained both intima-media complex and the lumen. Images in this sequence were (1296×152) pixels in size. Selected regions were tracked over all 271 frames using both ARPS and ES. Fig. 8a represents the axial displacement trajectory obtained using ARPS and ES. A similar output for CA2 (cross-section view) dataset is also represented in Fig. 8b (graphs for lateral displacement trajectory are not shown due to negligible lateral movement). It is clear that both algorithms follow similar trends with extremely negligible error. For the CA4 dataset, NCF and CTF averaged over all the bounding boxes for both algorithms are summarized in Table 3. Since no ground truth data was available for these images, the RMSE of the centroids tracked by ARPS and ES were compared against each other. RMSE between ARPS and ES averaged for all six bounding boxes was found to be 0.0157 pixels in the axial direction and 0.0027 pixels in the lateral direction. NCF and CTF for CA2 dataset is represented in Table 4. RMSE between ARPS and ES for CA2 dataset was found to be 0.0774 pixels in the axial direction and 0.0239 pixels in the lateral direction. Similar experiments were carried out for all the remaining four datasets. Due to limited space, outputs of all datasets are not shown. However, NCF and CTF calculated for ES and ARPS averaged over all datasets is shown in Table 5. The averaged RMSE of displacement between ARPS and ES over all datasets was found to be 0.061 pixels along the axial direction and 0.023 pixels along the lateral direction. ARPS was $\sim$221 times faster than ES in terms of NCF, and $\sim$72 times faster in terms of CTF. Therefore, evidently, ARPS outperforms ES with only a marginal trade-off for accuracy.

From the established results, it is easy to see that ES can be computationally very expensive. The intrinsic characteristic of ES is the brute-force search. A search region with p pixels will take $(2p+1)^2$ number of computations. Therefore, the computations increase as a function of a second order polynomial $(O(p^2))$. This can drastically reduce the computational time of the tracking algorithm. On the other hand, ARPS does only a constant number of computations $(O(1))$. In order to substantiate the claim, the search region was modelled as a function of the user input ROI $(p = max(ROI\_height, ROI\_width)/2)$. When an ROI of size (200×50) pixels was chosen, the NCF of ES was $\sim 4 \times 10^4$ and that of ARPS was still 6.9. This proves that ARPS is not computationally expensive even with very high resolution images. This ensures that ARPS can be used to track motion in real-time on portable devices. The above-mentioned results were obtained by running the code on MATLAB 2018 (The MathWorks, Natick, Massachusetts, USA), on a MacBook Air with 1.8GHz dual-Core Intel Core i5 processor and 8GB RAM.



(a) CA4 dataset (b) CA2 dataset

Fig. 8. Axial displacement trajectory for CA2 and CA4 dataset.

TABLE III

NCF AND CTF VALUES AVERAGED FOR SIX ROIS IN CA4 DATASET.

	Average NCF	Average CTF
ES	1330.12	$408.6ms$
ARPS	6.88	$7.23ms$

An inherent limitation of ARPS is that it only searches at certain given points within the pre-defined search region. It can be seen from Fig. 8 that ES has a better frame-to-frame resolution compared to ARPS. But it is evident from the above-mentioned results that this limitation is negligible considering the advantages of using ARPS.

V. CONCLUSION AND FUTURE WORK

The performance of ARPS and ES were tested on six different carotid artery datasets. In all the cases, ARPS outperformed ES in terms of computations by a substantial margin. ARPS was three order of magnitude less than ES in terms of NCF and two orders of magnitude less than ES in terms of CTF with only a marginal depreciation in RMSE. This indicates that the usage of ARPS as a search technique within the block matching paradigm can help improve the performance of real-time motion estimation in very high resolution images.

VI. ACKNOWLEDGMENTS

We would like to thank Professor Emmad Ebbini from UISPL at the University of Minnesota for sharing the data with our group that was used for the analysis.

TABLE IV

NCF AND CTF VALUES AVERAGED FOR TWO ROIS IN CA2 DATASET.

	Average NCF	Average CTF
ES	1681	$712.9ms$
ARPS	6.99	$8.15ms$

TABLE V

NCF AND CTF VALUES AVERAGED OVER ALL DATASETS.

	Average NCF	Average CTF
ES	1526.5	$541.3ms$
ARPS	6.91	$7.55ms$

REFERENCES

[1] Jingfeng Jiang and Timothy Hall, "A robust real-time speckle tracking algorithm for ultrasonic elasticity imaging," in *2009 IEEE International Ultrasonics Symposium*. IEEE, 2009, pp. 451–454.

[2] Jonathan Ophir, Ignacio Cespedes, Hm Ponnekanti, Youseph Yazdi, and Xin Li, "Elastography: a quantitative method for imaging the elasticity of biological tissues," *Ultrasonic imaging*, vol. 13, no. 2, pp. 111–134, 1991.

[3] Yong Yue, John W Clark Jr, and Dirar S Khoury, "Speckle tracking in intracardiac echocardiography for the assessment of myocardial deformation," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 2, pp. 416–425, 2008.

[4] Mohamed Almekkawy and Emad Ebbini, "Two-dimensional speckle tracking using parabolic polynomial expansion with riesz transform," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 201–205.

[5] Andrea Giachetti, "Matching techniques to compute image motion," *Image and Vision Computing*, vol. 18, no. 3, pp. 247–260, 2000.

[6] Mohamed Khaled Almekkawy, Yasaman Adibi, Fei Zheng, Emad Ebbini, and Mohan Chirala, "Two-dimensional speckle tracking using zero phase crossing with riesz transform," in *Proceedings of Meetings on Acoustics 168ASA*. Acoustical Society of America, 2014, vol. 22, p. 020004.

[7] Magnus Cinthio, Asa Ryden Ahlgren, Jonas Bergkvist, Tomas Jansson, Hans W Persson, and Kjell Lindstrom, "Longitudinal movements and resulting shear strain of the arterial wall," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 291, no. 1, pp. H394–H402, 2006.

[8] Jon Bang, Torbjørn Dahl, Annemarieke Bruinsma, Jon Harald Kaspersen, Toril A Nagelhus Hernes, and Hans Olav Myhre, "A new method for analysis of motion of carotid plaques from rf ultrasound images," *Ultrasound in medicine & biology*, vol. 29, no. 7, pp. 967–976, 2003.

[9] Aimilia Gastounioti, S Golemati, J Stoitsis, and KS Nikita, "Kalman-filter-based block matching for arterial wall motion estimation from b-mode ultrasound," in *2010 IEEE International Conference on Imaging Systems and Techniques*. IEEE, 2010, pp. 234–239.

[10] Guillaume Zahnd, Maciej Orkisz, André Sérusclat, Philippe Moulin, and Didier Vray, "Evaluation of a kalman-based block matching method to assess the bi-dimensional motion of the carotid artery wall in b-mode ultrasound sequences," *Medical Image Analysis*, vol. 17, no. 5, pp. 573–585, 2013.

[11] Zhifan Gao, Yanjie Li, Yuanyuan Sun, Jiayuan Yang, Huahua Xiong, Heye Zhang, Xin Liu, Wanqing Wu, Dong Liang, and Shuo Li, "Motion tracking of the carotid artery wall from ultrasound image sequences: a nonlinear state-space approach," *IEEE transactions on medical imaging*, vol. 37, no. 1, pp. 273–283, 2017.

[12] Yao Nie and Kai-Kuang Ma, "Adaptive rood pattern search for fast block-matching motion estimation," *IEEE Transactions on Image processing*, vol. 11, no. 12, pp. 1442–1449, 2002.

[13] Radek Benes, Jan Karasek, Radim Burget, and Kamil Riha, "Automatically designed machine vision system for the localization of cca transverse section in ultrasound images," *Comput. Methods Prog. Biomed.*, vol. 109, no. 1, pp. 92–103, Jan. 2013.

[14] Kamil Říha and Igor Potúček, "The sequential detection of artery sectional area using optical flow technique," in *Proceedings of the 8th WSEAS International Conference on Circuits, systems, electronics, control & signal processing*. World Scientific and Engineering Academy and Society (WSEAS), 2009, pp. 222–226.

[15] Kamil Říha, Jan Mašek, Radim Burget, Radek Beneš, and Eva Závodná, "Novel method for localization of common carotid artery transverse section in ultrasound images using modified viola-jones detector," *Ultrasound in medicine & biology*, vol. 39, no. 10, pp. 1887–1902, 2013.

[16] Yayun Wan, Dalong Liu, and Emad S Ebbini, "Simultaneous imaging of tissue motion and flow velocity using 2d phase-coupled speckle tracking," in *2010 IEEE International Ultrasonics Symposium*. IEEE, 2010, pp. 487–490.

[17] Yayun Wan, Dalong Liu, and Emad S Ebbini, "Imaging vascular mechanics using ultrasound: Phantom and in vivo results," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2010, pp. 980–983.