Initialization of active contours for segmentation of breast cancer via fusion of ultrasound, Doppler, and elasticity images

Chadaporn Keatmanee a,c, Utairat Chaumrattanakul b, Kazunori Kotani c, Stanislav S. Makhanov a,n

a Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, Thailand
b Department of Radiology, Thammasat University, Pathum Thani, Thailand
c Japan Advanced Institute of Science and Technology, Ishikawa, Japan

ARTICLE INFO
Article history:
Received 3 August 2017
Received in revised form 15 December 2017
Accepted 19 December 2017
Available online 24 December 2017

Keywords:
Segmentation
Breast cancer
Active contours
Initialization
Ultrasound
Elastography
Doppler image

ABSTRACT
Active contours (snakes) are an efficient method for segmentation of ultrasound (US) images of breast cancer. However, the method produces inaccurate results if the seeds are initialized improperly (far from the true boundaries and close to the false boundaries). Therefore, we propose a novel initialization method based on the fusion of a conventional US image with elasticity and Doppler images. The proposed fusion method (FM) has been tested against four state-of-the-art initialization methods on 90 ultrasound images from a database collected by the Thammasat University Hospital of Thailand. The ground truth was hand-drawn by three leading radiologists of the hospital. The reference methods are: center of divergence (CoD), force field segmentation (FFS), Poisson Inverse Gradient Vector Flow (PIG), and quasi-automated initialization (QAI).

A variety of numerical tests proves the advantages of the FM. For the raw US images, the percentage of correctly initialized contours is: FM-94.2%, CoD-0%, FFS-0%, PIG-26.7%, QAI-42.2%. The percentage of correctly segmented tumors is FM-84.4%, CoD-0%, FFS-0%, PIG-16.67%, QAI-22.44%. For reduced field of view US images, the percentage of correctly initialized contours is: FM-94.2%, CoD-0%, FFS-0%, PIG-65.6%, QAI-67.8%. The correctly segmented tumors are FM-88.9%, CoD-0%, FFS-0%, PIG-48.9%, QAI-44.5%. The accuracy, in terms of the average Hausdorff distance, is respectively 2.29 pixels, 33.81, 34.71, 7.7, and 8.4, whereas in terms of the Jaccard index, it is 0.9, 0.18, 0.19, 0.63, and 0.48.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

An annual check for breast cancer includes a mammogram, and ultrasound examination of the breast. The mammogram is considered a primary tool for women who display no symptoms of the disease [1]. However, if the findings are uncertain, a woman may be called for further tests, which include extra mammographic views and breast ultrasound. Therefore, in practice, ultrasound is an additional tool, which is important in cases of dense breasts (young women).

Along with the conventional US imagery, modern US machines produce Elastography and Doppler images, which in many cases improve the quality of the diagnosis [21]. Elastography is used as an adjunct technique to help in discrimination between benign and malignant masses, based on their stiffness [84]. Power Doppler is another non-invasive US modality which supplements conventional US. Doppler images visualize the appearance and morphology of blood vessels associated with a mass. The analysis uses the fact that a benign mass has little or no vascular flow, whereas a malignancy increases the blood flow in the vicinity of the tumor [15].

Since the Elastography and Doppler images are usually available in cases when cancer is suspected, we propose to use them, along with a conventional US image to improve the quality of automatic segmentation of breast tumors. In particular, we focus on the fusion of the US, elasticity, and Doppler images in the framework of active contours (snakes) [43]. Active contours are one of the most popular segmentation techniques applied to many image processing problems, originating from different applications. The most successful modifications of the active contours are the gradient vector flow (GVF) snakes [90], generalized gradient vector flow snakes (GGVF) [89], multidirectional GGVF snakes [76], Vector Field Convolution snakes (VFC) [48], and the recent Adaptive Diffusion Flow snakes (ADF) [87].

However, the accuracy and computational time of the above mentioned models depend on the initial location (seed snakes). Unfortunately, if the seeds are far from the boundary of the object,
the snake can attach itself to false boundaries. Modern US/sharewave machines generate three types of the images mentioned above. Each of them helps radiologists to localize and classify a tumor i.e. the tumors characterized by the low intensity of the gray level in the US image, high stiffness in the elasticity image, and by an increased vascularity in the Doppler images. Therefore, we introduce a novel initialization method based on the fusion of the conventional US, elasticity and Doppler images.

The proposed algorithm combines the images by means of a distance transform and a low intensity mask to generate a suitable initial contour and a supplementary balloon force. A video demo of the algorithm is at http://drive.google.com/file/d/13DWNiBnMPC_P8nKn6TGfVoQD0n9Z5Ko2.

2. Related work

Segmentation of breast abnormalities has received considerable attention in the literature. An extensive review by Noble and Boukerroui [65] mentions that such segmentation can be treated as a general image processing problem and includes a priori information of ultrasound. The segmentation algorithms include numerous modifications of the conventional thresholding, neural networks (see a concise survey in [72]), watershed techniques [35], statistical methods [8,54], active contours (see extended surveys [13,30], and a large list of references compiled in [68]), level set method (see recent surveys in [28,29]), and graph-based segmentation refined by active contours [34]. Excellent results have been obtained by combining a modified watershed model and all tissue classification for segmentation of 3D US images [26].

However, among these segmentation techniques, neural networks and other AI based methods require feature selection and training. Initial seeds are required for the watershed and graph-based methods, as well as for the active contours and the level-set methods. Finally, an unsupervised Fuzzy C-means method (FCM) is a good alternative to techniques requiring prior information of ultrasound. The algorithm differentiates the initial snakes selected from the edge map are not robust and may not be applicable to multiple objects.

The idea to initialize the snakes at the CoD of the GVF-type vector field was first mentioned in [83]. Further, Ge and Tan [88] define the CoD by analyzing relative directions of the vector field in a sliding $2 \times 2$ window (a generalization to larger windows is not available). He et al. [95] uses Phase Portrait Analysis (PPA) to detect the critical points of the vector field and a rule that “the initial contours should be set to contain all of the node points in the object area and none of the others”. Although PPA has been used in a variety of image processing applications, e.g. [16,51,52,74,80,93], the standard PPA classifier based on “if then” rules often cannot be adapted to the case of snake initializations characterized by irregular nodes corrupted by noise.

The similarity of the GGVF and the Navier-Stokes equations makes it possible to use the analogy of a flow through a porous medium. Consequently, Ray et al. [67] treats the initial snakes as sources of flow, emitting normal unit vectors into the image domain. The authors also noticed critical points of the flow and proposed to merge multi-snakes, initialized around those points for segmentation of the MRI images of lungs.

A competing idea is placing the seed points uniformly or randomly over the entire image, evolving them from each seed point, and analyzing the resulting configuration [70]. However, the required classifier to validate the final configuration must be trained, which makes the model image-dependent.

A partial solution to the problem is the above mentioned QAI method by Tauber et al. [78,79]. The method employs the CoDs combined with a tracing procedure to create a “skeleton” of the object, consisting of centers of strong and weak divergence. The centers of weak divergence are the points where the vectors of the GVF diverge in one (either horizontal or vertical) direction. The centers of strong divergence feature both horizontal and vertical divergences. The initial snake is generated around the skeleton. However, the initialization is not entirely automatic. The algorithm still requires at least one manually generated point inside the object. Moreover, in some cases the skeleton can evolve outside the boundary of the object.

The above mentioned PIG model [49] establishes the relationship between the external force field and the underlying external energy field via the solution of the corresponding Poisson equation. The model has been applied to 2D and 3D cases for a variety of medical images. The isoline of the minimum energy is selected as the initial contour. However, the model may suffer from incomplete isolines, as well as from over segmentation.

An automatic initialization method has been proposed in [32] for PET images of the liver. The candidate contours are generated by Canny edge detection and subsequently classified by a genetic algorithm. The algorithm has been applied to segmentation of face contours in video files [33]. A similar idea was introduced in [81] for detection of the synovial boundaries in US images. However, the initial snakes selected from the edge map are not robust and may not be applicable to multiple objects.

The idea of trial snakes (TS), combined with PPA, was applied to US images of breast cancer in [44]. The PPA makes it possible to detect the centers of convergence and divergence, as well as the attracting and repelling nodes. The algorithm differentiates between the internal and external seeds by running multiple TS from the critical points and checking their intersections with the boundary of the image. The most serious drawback of TS is that they require a considerable amount of computational time.

The initialization algorithms for the US images often rely on gray levels and textures, to place the seed points inside the tumor [19,41,58]. Saliency and feature maps have been proposed in [71]. Fergani et al. [24] introduces a special vector field to hybridize the GVF and the texture. A Chan-Vese type model is proposed in [55]. A few papers related to a specific medical image processing task use the typical position of a human organ in the US images (see, for instance, Akgul et al. [31]). However, these models are image dependent and may not work if strong noise is present.

Therefore, this paper proposes a new fast algorithm for automatic initialization, which combines the conventional grayscale US image with the corresponding elasticity and Doppler images.
The technique makes it possible to locate the initial contour inside the tumor, close enough to the true boundaries to ensure convergence of the active contour. The paper also introduces a modification of the balloon type active contour, based on a combination of the radial force derived from the fusion image and the GVF-type force.

3. Background

3.1. Ultrasonography for breast cancer screening

The World Health Organization reports that breast cancer is common in women. The cases of breast cancers in developing countries are increasing due to the increase in life expectancy, urbanization, and western lifestyles [86]. The mammogram remains the primary screening tool for women who display no symptoms of the disease [1]. However, high-resolution US is considered one of the most appropriate adjunct tools [73,69] due to a number of advantages, such as no exposure to radiation, simplicity, and low cost. Besides, in dense breasts, mammography has limited sensitivity, whereas US is useful to examine dense breast tissue. Note that women with dense parenchyma have an increased risk of breast cancer. Recent studies have shown that the detection of small cancers with high-resolution US has increased by 3–4 cancers per 1000 women without clinical or mammographic abnormalities [59]. Apart from screening, US is used in daily practice, in order to improve lesion detection and characterization. Several large studies have shown a possible role of US for cancer detection and differentiation e.g. [45]. Thus, computer aided diagnosis, including US imaging, is likely to improve survival rates [62,63].

3.1.1. Elastography

Elastography (Fig. 1(b)) is a US imaging modality to classify breast masses based on their stiffness [37,91]. A US machine displays a color elastography image as follows [39]:

- Score 1: blue indicates soft and loose structures.
- Score 2: a combination of blue and green indicates soft-rigid structures.
- Score 3: red and dark red at the center of the mass, and green at the periphery, indicate a hard to soft mass.
- Score 4: red and dark red indicates a hard and tight mass.
- Score 5: red and dark red covering the mass and the surrounding tissue refer to a hard expanding mass. Scores 1, 2, and 3 represent benign features, whereas masses scoring 4 and 5 are likely to be malignant [45,91].

Note that some machines display the stiffness in the reverse palette, i.e. blue indicates the hardest tissue, whereas red indicates the soft tissue.

3.1.2. Power Doppler imaging

Power Doppler (Fig. 1(c)) is another non-invasive supplement to the conventional US. The sound waves bounce off solid objects, including blood cells. The movement of the blood cells causes a change in the pitch of the reflected sound waves as a result of the Doppler effect. Typically, a benign mass has little or no appearance of vascular flow, whereas a malignant mass is often characterized by an increased blood flow. The advantages of Power Doppler are high sensitivity, low angle dependency, and no aliasing [64,92]. Therefore, it is also used as an adjunct image modality for breast cancer diagnosis [64,38]. When the cells move towards the transducer, the frequency of returning ultrasound waves is greater than that of the emitted waves, and the blood flow is depicted in red. When the cells move away the flow is depicted in blue. The intensity of the color is proportional to the flow velocity [66]. However, in this paper, we do not analyze the Doppler palette. We only register the presence or absence of the Doppler spots.

3.1.3. Combination of the imaging modalities in clinical practice

The efficiency of combined US, Doppler, and elasticity imaging in diagnosing breast malignancy is still controversial. For instance, Davoudi et al. [18] reports that “using the Doppler image alone has little value in differentiating between malignant and benign breast lesions”. The research conducted by [25] concludes that Doppler imaging does not contribute to categorization of solid breast masses. However, there is growing evidence that analysis of a combination of images improves the characterization of breast lesions. Thomas et al. [79] reports that sensitivity/specificity was 96%/98% for US, 100%/40% for US and mammography, and 96%/80% for the combined mode, including Doppler. Cho et al. [15] characterizes the results obtained from 5 readers by the area under the receiver operated characteristic curve AUC as follows: “the AUC of the US mode, elastography, and Doppler US (average, 0.844; range, 0.797–0.876) was greater than that of the US mode alone (average, 0.771; range, 0.738–0.798) for all readers”. Li et al. [53] reports “the specificity of making the decision for biopsy increased from 6.5% to 38.7% when US was combined with color Doppler and elasticity without a statistically significant change in sensitivity”. Elkharbhotli and Farouk [22] shows that a combined use of US, elasticity, and color Doppler achieved an NPV of 95% “thus allowing sparing of unnecessary invasive diagnostic procedures”. In summary, there is enough evidence that a combination of the conventional US, elastography, and Doppler images improves the accuracy of diagnosis. Therefore, the above imaging modalities will be increasingly used in clinical conditions to allow for computerized segmentation and classification of tumors.

3.2. Combination of the imaging modalities for classification

Combinations of the different US modalities have been used in several classification algorithms. For instance, the conventional
US and Elastography are used by [7,96], whereas a combination of US and Doppler images is used by [36,31,12,47]. However, the idea of using the three modalities for initialization, and evolving the active contours, seems to be overlooked.

### 3.3. Active contour model

An active contour (snake) is a parametric curve \( \nu(s) = (x(s), y(s)) \) which grows or contracts inside an image to attach itself to the boundary of the desired object. The evolution of the snake is simulated by Euler equations for the energy functional, given by

\[
E = \int_0^1 \left[ \frac{1}{2} (\alpha \cdot |\nu_s(s)|^2 + \beta \cdot |\nu_y(s)|^2) + E_{\text{ext}}(\nu(s)) \right] \, ds,
\]

where the subscripts denote partial derivatives, \( \alpha \) and \( \beta \) are weighting parameters to control the snake's tension and rigidity, and \( E_{\text{ext}}(\nu(s)) \) is the external force. The most successful modifications of the active contours are Gradient Flow (GVF) snakes [90], Generalized Gradient Vector Flow (GGVF) snakes [89], multidirectional GGVF snakes [76], and the non-linear diffusion model [85]. Recent GVF-type models are Normal Gradient Vector Flow [40], Infinity Laplacian [27], Harmonic Gradient Vector Flow [82], Convolution Vector Flow [48], Dynamic Directional Gradient Vector Flow [14], Adaptive Diffusion Flow [87], and Multi Feature Gradient Vector Flow [68].

### 4. Method

This section presents a new initialization method, based on the fusion of the US, elasticity, and Doppler images.

#### 4.1. Preprocessing

The proposed algorithm preprocesses the three types of images introduced above as follows:

**US image** (Fig. 1(a)):

\[
U_{\text{mask}} = \text{Binarize}(\text{Gaussian}(U_{\text{raw}})).
\]  

After Gaussian smoothing, the US image is binarized, creating a low intensity mask (see Fig. 2(b)). Alternatively, one can generate the low intensity mask by (Fig. 2(c)):

\[
U_{\text{mask}} = \text{RegionGrowing}((\text{Binarize}(U_{\text{raw}}))).
\]

The mask applies to an image \( I \) as follows:

\[
\text{Mask}(I_{ij}) = \begin{cases} I_{ij}, & \text{if } U_{\text{mask},ij} = 1, \\ 0, & \text{otherwise}. \end{cases}
\]

The edge map is generated using Fuzzy C-means clustering [6] (see Fig. 2(a)).

**Elasticity image** is binarized (Fig. 3(a)). The resulting output image (Fig. 3(b)) is obtained by applying the low intensity mask (see Figs. 1(b) and 2(b)):

\[
E = \text{Mask}(\text{Binarize}(\text{Red}(E_{\text{raw}}))).
\]

The areas characterized by an increased blood flow are represented by colored spots superimposed on \( U_{\text{raw}} \). The Doppler image (Fig. 1(c)) is converted into grayscale and binarized, using an automatic threshold (see Fig. 4(a)). The low intensity mask is then applied to the resulting image (see Fig. 4(b)):

\[
D = \text{Mask}(\text{Binarize}(\text{GrayScale}(D_{\text{raw}}))).
\]

Elimination of outliers is performed by following [10,11] using the Mahalanobis distance [60] and 97.5%-quantile of the Chi-square distribution (Fig. 4(c)).

#### 4.2. Distance transform. Soft intersection

The distance transform is defined with regard to a prescribed set of points \( S \) as follows:

\[
d_s(P) = \min_{S} \| P - S \|,
\]

where \( \| \cdot \| \) denotes the Euclidean distance and \( P \) is an arbitrary point. In image processing, the set \( S \) usually is an object (and possibly noise), and \( P \) is a point in the binarized image. Consider a normalized \( d_s(P) \), so that \( \max d_s(P) = 1 \). In the forthcoming figures, \( d_s(P) \) is represented by a grayscale image, which indicates how far a particular point \( P \) is from set \( S \). Clearly, \( d_s(S) = 0 \). An introductory example is shown in Fig. 5.

Define the fusion image by a soft intersection of \( d_u, d_e, \) and \( d_D \):

\[
d_f(P) = w_u d_u(P) + w_e d_e(P) + w_D d_D(P),
\]

where \( d_u(P), d_e(P), \) and \( d_D(P) \) are the distance transforms, corresponding to \( U_{\text{edge}}, E \), and \( D \), respectively, and \( w_u, w_e, w_D \) are the weighting coefficients: \( w_u + w_e + w_D = 1 \) (Fig. 6(a)-(d)). In practice we use a simple average \( w_u = w_e = w_D = \frac{1}{3} \). The contribution of each image to the fusion image is illustrated in Fig. 7. As a variant, \( d_u, d_e, \) and \( d_D \) can be regarded as the RGB colors, and \( d_f(P) \) as a function to convert RGB image to grayscale. Using this analogy, \( d_f(P) \) can also be defined similarly to the lightness conversion method [42]:

\[
d_f(P) = \left( \max(d_u(P), d_e(P), d_D(P)) + \min(d_u(P), d_e(P), d_D(P)) \right)/2.
\]

The numerical experiments show that the above fusion formulas (9) and (10) produce very close results. Our forthcoming section “Results and discussion” illustrates the proposed FM applied with the soft intersection (9). Fig. 8 demonstrates the advantages of

---

**Fig. 2.** (a) Edge map (U-image), (b) \( U_{\text{mask}} \) (Gaussian), (d) \( U_{\text{mask}} \) (Region Growing).
the soft intersection for the image in Fig. 1 with reference to the binary intersection. Clearly, the soft intersection in Fig. 8(f) produces an appropriate grayscale image, which can subsequently be used to initialize an active contour, whereas the binary intersections in Fig. 8(d) and (e) are not suitable for initialization.

4.3. Automatic initialization

Consider a sequence of thresholds $T = T_1, T_2, \ldots, T_N$, obtained by an iterative binarization applied to $d_t$. We apply the Otsu algorithm, employing multiple automatically evaluated thresholds
Another option is a quantized version of the fusion image, thresholded at \( N \) equally spaced levels (Fig. 9):

\[
T = \left\{ \frac{\max(d_F)}{N}, \frac{2\max(d_F)}{N}, \ldots, \frac{(N - 1)\max(d_F)}{N} \right\}.
\]

Consider a fusion image \( d_F(T) \), binarized at the level \( T \). Denote the corresponding edge map by \( E_F(T) \) and a set of all continuous contours from \( E_F(T) \) by \( C_F(T) \). Our basic idea is that the best threshold is the one that generates the closest contour \( C_F(T) \) to the convex hull of the Doppler points in the \( D \)-image (Fig. 10). Additionally, the contour is verified by a decision tree generated by the CART algorithm [9].

A pseudo-code of the algorithm is given below:

**input**: \( U_{raw}, E_{raw}, D_{raw} \)

\( U, E, D \) ← preprocess the input images using (5)–(7), respectively

\( d_F \) ← distance transform (9)

\( T \) ← sequence of thresholds (11)

\( B_{o, All} \) ← NULL

\( B_D \) ← convex hull of \( D \)

\( c_{bd} \) ← centroid of \( B_D \)

**for each** \( T_k \) **in** \( T \)

\( d_{F_k} \) ← threshold \( d_F \)

\( E_{F_k} \) ← edgemap of \( d_{F_k} \)

\( C_{F_k} \) ← all continuous contours of \( E_{F_k} \) (see Remark 1)

\( B_o \) ← contour \( C_{F_k} \) closest to \( B_D \) in terms of the Hausdorff distance \( dist_1 \)

YesNo ← evaluate the resulting contour \( B_o \) by CART algorithm (decision tree)

if YesNo \( B_{o, All} \) ← \( B_{o, All} + B_o \)

if \( B_{o, All} \) = NULL

break

else

\( B_{init} \) ← contour \( B_{init, o} \) closest to \( B_D \), in terms of \( dist_1 \)

\( c_o \) ← centroid of \( B_{init} \)

\( B_{init} \) ← contour \( B_{init} \) scaled by factor \( \gamma \), with \( c_o \) being the origin

**return** \( B_{init} \)

**Remark 1.** The procedure to trace the contours is similar to the routine findContours of openCV [46].

We construct the decision tree (DT) using 30 additional US images submitted to the MatLab function \texttt{fitctree} [61]. The function is designed to implement a conventional CART DT [9, 17, 56, 57] using the Gini index. The DT is based on the following features:

- Average gradient of the gray level along \( B_o \) relative to the max gradient: \( l(T) \)
- Ratio of intersection \( B_o \) and \( B_D \) relative to the \( B_D : A_{B_o,B_D} = \text{Area}(B_o \cap B_D)/\text{Area}(B_D) \)
- Distance between \( c_o \) and \( c_{bd} : d_{c_o,c_{bd}} \)
- Boolean variable \( L = B_D \in B_o \)
The resulting impurity measures are:

$$I_{\text{Gini}} = 0.0223, \quad A_{\text{B}, B_0, \text{Gini}} = 0.0199, \quad d_{\text{G}, E_\text{out}, \text{Gini}} = 0.0239, \quad I_{\text{Gini}} = 0.0221.$$ 

The above measures clearly indicate that the resulting DT is a good quality classifier. The corresponding DT and the thresholds are shown in Fig. 11. Note that in order to ensure that $B_{\text{out}}$ is inside the tumor, $B_0$ is scaled by factor $\gamma$. In principle, $\gamma$ can be taken sufficiently small so that the snake evolves from the centroid. The FM-balloon force (see the next section) delivers the snake to the boundary, even from a single point. However, in order to improve the computational time, in practice, we use $\gamma = 0.4$.

### 4.4. Fusion radial force

Recall that the conventional active contour is represented by Eq. (1) [43]. Our proposed external force $E_{\text{ext}}(s(t))$ is:

$$E_{\text{ext}}(s(t)) = E_1^\text{ext}(s(t)) + E_2^\text{ext}(x,y),$$

where $E_1^\text{ext}(s(t))$ is the traditional gradient based force, and $E_2^\text{ext}(x,y)$ is the balloon- type fusion radial force (FRF). The FRF is proportional to the distance between $(x,y)$ and $B_0$ so that $E_2^\text{ext}(x,y) = 0$ if $(x,y) \in B_0$, and $E_2^\text{ext}(x,y) = E_2^\text{max}$ if $(x,y) = c_0$, where $E_2^\text{max}$ is evaluated experimentally (see Fig. 12).

### 5. Performance measures

In order to compare the fusion method with the conventional algorithms, we introduce the following performance measures.

#### 5.1. Contour based accuracy measures

The Hausdorff distance given by

$$\text{dist}_{H}(X,Y) = \max \left\{ \max_{a \in X} \min_{b \in Y} \| a - b \|, \max_{b \in Y} \min_{a \in X} \| a - b \| \right\}$$

Fig. 7. Generation of the fusion image.

Fig. 8. Soft (top) and hard (bottom) (binary) intersection: (a) $U_{edge}$, (b) $E$, (c) $D$, (d) hard intersection of $U_{edge}$ and $E$, and $D$, (e) hard intersection of $E$ and $D$, (f) soft intersection $d_F$.

Fig. 9. (a) Quantized $d_F$-image (b) the corresponding contours.
where \( \| \| \) denotes the Euclidean distance, \( X \) the ground truth contour, and \( Y \) the snake contour.

The averaged Hausdorff distance is defined by

\[
\text{dist}_{\bar{H}}(X, Y) = \max \left\{ \sum_{a \in X} \min_{b \in Y} \| a - b \|, \sum_{b \in Y} \min_{a \in X} \| a - b \| \right\},
\]

where \( L_X, L_Y \) is the length of the true contour, and the resulting contour, respectively.

The relative Hausdorff distance is given by

\[
\text{dist}_{H}(X, Y) = \frac{\text{dist}_{\bar{H}}(X, Y)}{L_X},
\]

where \( \xi = 1000 \) is the normalizing coefficient. The distance evaluates the relative importance of the difference between the two curves. For instance, if \( \text{dist}_{H}(X, Y) = 10 \), and \( L_X = 100 \) pixels, the error is unacceptable, however, if for instance, \( L_X = 10,000 \), then \( \text{dist}_{H}(X, Y) \) is appropriate. The importance of the Hausdorff distance in comparing planar curves is parametrization invariance. Although \( \text{dist}_{H} \) is not a distance in a rigorous mathematical sense (it does not satisfy the triangle inequality), Dubuisson and Jain [20] shows that it is the best for matching curved objects.

The contour-based true positive rate is:

\[
\text{TP}_c = \frac{\text{TP}_Y}{N_Y},
\]

where \( \text{TP}_Y \) is the number of true positive pixels, and \( N_Y \) is the total number of pixels belonging to the resulting active contour (in practice we consider \( L_X = N_X \) and \( L_Y = N_Y \)).

5.2. Region based accuracy measures

The most used metric in validating medical segmentations [75] is the Dice coefficient given by
where $R_x$ and $R_y$ are the regions corresponding to the contours $X$ and $Y$, respectively. $A_{R_x}$ and $A_{R_y}$ are the areas of $R_x$ and $R_y$, respectively.

5.3. Performance of the initialization procedure

The performance of the initialization is evaluated for the entire series of images by $N_{corr}$, defined as the percentage of images for which the internal and external seeds were correctly differentiated, $S_{corr}$, the percentage of images for which the contour was correctly segmented (the final snake is considered correct if $dist_{frf}(X, Y) \leq 3$), and the computational time $T_{comp}$.

As noted above, the segmentation accuracy depends, not only on initialization, but on the segmentation model as well. For instance, the level set method, clustering, watershed segmentation, region growing, and edgeless active contours may benefit from the proposed FM. However, this is out of the scope of this paper. At present, the model is focused on the parametric active contours.

6. Results and discussion

6.1. Experimental dataset

The algorithm has been tested on 90 US images of breast cancer from 90 different patients obtained by a Philips iU22 ultrasound machine at the Thammasat University Hospital. The resolution ranges from $200 \times 200$ to $300 \times 400$ pixels. The ground truth contours have been drawn by three leading experts with the Department of Radiology of Thammasat University using an electronic pen and Samsung Galaxy Tablet computer. The final ground truth was obtained by the majority voting rule (two out of three).

6.2. Numerical experiments and discussion

The FM has been tested against four state-of-the-art initialization models, namely, center of divergence (CoD) [88], force field segmentation (FFS) [50], Poisson inverse gradient (PIG) [49], and quasi automatic initialization (QAI) [77,78], using the performance measures Eqs. (13)–(20). In order to prove the efficiency of the FRF, we compare it with the Vector Field Convolution (VFC) snake [48] and the recent Adaptive Diffusion Flow (ADF) [87] methods, which have been proven to be superior to GVF [90], Normal Gradient Vector Flow [40], Infinity Laplacian GVF [27], and Harmonic Gradient Vector Flow [82]. Fig. 13 is an example, comparing the initialization and the resulting snake produced by the FM/FRF with CoD/VFC, FFS/VFC, PIG/VFC, and QAI/VFC. Fig. 13(a) and (b) shows a US image with a “false” tumor on the right side of the image and a shadow at the lower left corner, characterized by grayscale comparable with the gray level of a true tumor. The resulting edge map in Fig. 13(c) shows multiple irregular contours. Clearly, if a contracting snake is initialized at the boundary of the image, it will attach itself to a wrong object and produce a totally inappropriate contour. Therefore, this US image requires a high quality initial snake, preferably expanding from the inside of the true contour. The CoD and FFS produce multiple seeds at the CoDs (Fig. 13(e) and (f)). However, due to noise, the corresponding multiple snakes are unable to merge (Fig. 13(j) and (k)). In turn, the PIG misses the true tumor and generates the initial contour inside a false object (Fig. 13(g)). Therefore, the resulting snake is inappropriate (Fig. 13(l)). Finally, QAI requires one user-defined point inside the actual object. Due to this, QAI generates the initial contour around a correct location of the tumor. Since QAI is based on a “skeleton” of the object, which connects the CoDs, the method wrongly includes a CoD located outside the object (Fig. 13(h)). Consequently, the expanding snake grows outside the tumor and partially attaches...
to a false boundary (Fig. 13(m)). The FM initialization method outperforms the above techniques because it has more information about the location of the tumor. Although the low intensity mask in Fig. 2(b) cannot localize the tumor, the combination of the Doppler and elasticity images excludes the artifact (false tumor), unwanted shadows, and produces an appropriate initial contour \( B_{\text{ini}} \). Finally, the proposed DT verifies the candidate contour using supplementary features.

Tables 1–3 illustrate numerical tests of the proposed method vs. the above mentioned techniques, using the initialization and accuracy measures (13)–(20) introduced in Section 5. For every measure, we calculate the mean \( \mu \) and the standard deviation \( \sigma \). For all evaluation measures related to FM \( \sigma / \mu < 1 \). It indicates a low spread of the error.

Table 1 demonstrates the advantages of the proposed initialization method \( (N_{\text{corr}} = 92.2, S_{\text{corr}} = 84.4) \). The CoD and FFS failed

<table>
<thead>
<tr>
<th>Model</th>
<th>Initialization measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comp. time ( T_{\text{comp}} ), sec</td>
</tr>
<tr>
<td></td>
<td>( \mu )</td>
</tr>
<tr>
<td>FM</td>
<td>11.39</td>
</tr>
<tr>
<td>CoD</td>
<td>17.72</td>
</tr>
<tr>
<td>FFS</td>
<td>19.53</td>
</tr>
<tr>
<td>PIG</td>
<td>9.17</td>
</tr>
<tr>
<td>QAI</td>
<td>145.23</td>
</tr>
</tbody>
</table>

Tables 1–3 illustrate numerical tests of the proposed method vs. the above mentioned techniques, using the initialization and accuracy measures (13)–(20) introduced in Section 5. For every measure, we calculate the mean \( \mu \) and the standard deviation \( \sigma \). For all evaluation measures related to FM \( \sigma / \mu < 1 \). It indicates a low spread of the error.

Table 1 demonstrates the advantages of the proposed initialization method \( (N_{\text{corr}} = 92.2, S_{\text{corr}} = 84.4) \). The CoD and FFS failed
\( N_{corr} = S_{corr} = 0 \), whereas PIG and QAI yield a lower performance \((N_{corr} = 26.7, S_{corr} = 16.7, 22.4, \text{respectively})\). All tested methods have been implemented on the MatLab platform, using an AMD PRO A8-8600B R6 CPU, 1.6 GHz, with 8 GB RAM and 64-bit OS. The FM is the second best by the average speed (11.4 s), but significantly better in all other categories, including correctly initialized and correctly segmented images. Tables 2 and 3 show the performance of FM/FRF, CoD/VFC, FFS/VFC, PIG/VFC, and QAI/VFC. The results produced by the reference methods are characterized by low accuracy. For instance, in terms of \( \text{dist}_{H^2} \), the average accuracy of the FM is 2.4 pixels, whereas for CoD, FFS, PIG, and QAI it is 321.6, 486.57, 58.5, and 82.23, respectively.

### 6.3. Impact of the radial force

The next important question is the impact of the FRF. Fig. 14 compares segmentation produced by the proposed FRF with that produced by VFC and ADF. Note that the initial contour is obtained by FM. Tables 4–6 clearly demonstrate that the FM-based initialization improves ADF and VFC. For instance, FM/VFC segments

#### Table 3

<table>
<thead>
<tr>
<th>Model</th>
<th>Jaccard</th>
<th>Dice</th>
<th>SEN</th>
<th>( H_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM</td>
<td>0.89</td>
<td>0.07</td>
<td>0.78</td>
<td>0.90</td>
</tr>
<tr>
<td>COD</td>
<td>0.17</td>
<td>0.09</td>
<td>0.59</td>
<td>0.28</td>
</tr>
<tr>
<td>FFS</td>
<td>0.16</td>
<td>0.12</td>
<td>0.75</td>
<td>0.27</td>
</tr>
<tr>
<td>PIG</td>
<td>0.1</td>
<td>0.24</td>
<td>2.40</td>
<td>0.15</td>
</tr>
<tr>
<td>QAI</td>
<td>0.34</td>
<td>0.36</td>
<td>1.06</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Fig. 14. Force field: (a) FRF, (b) VFC, (c) ADF; snake evolution: (d) FRF, (e) VFC, (f) ADF; segmentation results: (g) FRF, (h) VFC, (i) ADF.
56% of the images, whereas the best initialization using QAI/VFC detects only about 22%. However, the performance is still substantially lower than the 84% produced by FM/FRF. The accuracy of VFC has been improved (see a decrease of \(H_2\) from 39 pixels (PIG/VFC) to 9.58 (FM/VFC) in Table 5). However, the proposed FM/FRF shows \(H_2 = 2.4\) and the smallest standard deviation of about 0.95. As a matter of fact, for the contour based measures, FM has the smallest ratio of \(\sigma / \mu\) in all categories.

### 6.4. Reduced field of view

In clinical practice, the radiologist often defines a reduced field of view (RFOV), which can be used by a computerized segmentation procedure (see Fig. 15). Since a part of our image set does not have an RFOV, we analyzed 30 images with an RFOV defined by a radiologist. The ratio of the area of the tumor’s minimum bounding rectangle (MBR) and the RFOV was approximately 1/2. Therefore for the remaining images we generated an RFOV automatically by increasing the sides of the MBR by \(\sqrt{2}\). Fig. 16 displays the results obtained by FM/FRF, CoD/VFC, FFS/VFC, PIG/VFC, and QAI/VFC for a sample RFOV-image. Furthermore, Tables 7–9 compare the numerical results obtained with the RFOV. Clearly, the RFOV improved the performance of PIG and QAI to 48 and 44% respectively. However, FM also improved to 89%. Let us also compare the improvement in the accuracy taking \(H_2\) as the reference: FM from 2.4 to 2.29, PIG from 58 to 7.7, and QAI from 82 to 8.4. Hence, RFOV has a great impact on the two reference methods, improving their accuracy by 8–10 times. However, FM still remains the best method in all categories.

### 6.5. Relative impact of different modalities

An important question is whether the FM/FRF requires all three image modalities, and which modality is the most important. Tables 10–12 show the accuracy of the proposed method applied to combinations \((U, E), (U, D)\), etc. Clearly, combining the three types of images produces the best accuracy. Interestingly enough, \((E, D)\) is the second best in accuracy \(S_{\text{corr}} = 73.3\%\). However, the segmentation procedure uses a mask produced by the US image. Since this combination implicitly uses the US-image, it is incorrect to say that the algorithm is based solely on \((E, D)\). This complies

![Fig. 15. Minimum bounding rectangle and the RFOV.](image-url)
Fig. 16. Segmentation on the RFOV, (a) US image, (b) ground truth (c), U_{edge}, (d) FM-initialization, (e) CoD initialization, (f) FFS-initialization, (g) PIG-initialization (h) QAI-initialization. Segmentation results: (i) FM/FRF, (j) CoD/VFC, (k) FFS/VFC, (l) PIG/VFC, (m) QAI/VFC.

Table 7
Initialization for RFOV images.

<table>
<thead>
<tr>
<th>Model</th>
<th>Comp. Time $T_{com}$, sec</th>
<th>Correctly initialized $N_{corr}$, %</th>
<th>Correctly segmented $S_{corr}$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu/\sigma$</td>
</tr>
<tr>
<td>FM</td>
<td>6.58</td>
<td>1.75</td>
<td>0.26</td>
</tr>
<tr>
<td>CoD</td>
<td>1.54</td>
<td>1.37</td>
<td>0.89</td>
</tr>
<tr>
<td>FFS</td>
<td>1.62</td>
<td>1.65</td>
<td>1.02</td>
</tr>
<tr>
<td>PIG</td>
<td>1.47</td>
<td>1.17</td>
<td>0.79</td>
</tr>
<tr>
<td>QAI</td>
<td>1.17</td>
<td>4.79</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table 8
RFOV, contour based measures.

<table>
<thead>
<tr>
<th>Model</th>
<th>$H_1$</th>
<th>$H_2$</th>
<th>$H_3$</th>
<th>$TP_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu/\sigma$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>FM</td>
<td>7.13</td>
<td>2.17</td>
<td>0.30</td>
<td>2.29</td>
</tr>
<tr>
<td>CoD</td>
<td>43.49</td>
<td>8.85</td>
<td>0.20</td>
<td>33.81</td>
</tr>
<tr>
<td>FFS</td>
<td>42.97</td>
<td>9.54</td>
<td>0.22</td>
<td>34.81</td>
</tr>
<tr>
<td>PIG</td>
<td>25.37</td>
<td>21.67</td>
<td>0.85</td>
<td>7.7</td>
</tr>
<tr>
<td>QAI</td>
<td>26.44</td>
<td>21.07</td>
<td>0.80</td>
<td>8.43</td>
</tr>
</tbody>
</table>

Table 9
RFOV, region based measures.

<table>
<thead>
<tr>
<th>Model</th>
<th>$Jaccard$</th>
<th>Dice</th>
<th>$SEN$</th>
<th>$H_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\mu/\sigma$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>FM</td>
<td>0.90</td>
<td>0.05</td>
<td>0.55</td>
<td>0.91</td>
</tr>
<tr>
<td>CoD</td>
<td>0.18</td>
<td>0.04</td>
<td>0.22</td>
<td>0.31</td>
</tr>
<tr>
<td>FFS</td>
<td>0.19</td>
<td>0.06</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td>PIG</td>
<td>0.63</td>
<td>0.24</td>
<td>0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>QAI</td>
<td>0.58</td>
<td>0.25</td>
<td>0.43</td>
<td>0.66</td>
</tr>
</tbody>
</table>
with recent clinical research [18,25], which reports that the Doppler image alone does not significantly contribute to categorization of solid masses.

6.6. Limitations of the method

The method requires a good quality Doppler image $D_{raw}$. If $D_{raw}$ does not present a well-defined cluster of Doppler spots, the procedure works in the $U-E$ mode, which may reduce the accuracy of segmentation (see Table 10). Some low grade cancers may not appear in the Doppler images. For instance, Adler et al. [2] reports that “four percent of the cancers had no detectable [Doppler] flow”.

7. Conclusions

The proposed new automatic procedure for initialization of active contours, applied to the segmentation of ultrasound images of breast cancer, outperforms preceding algorithms. The procedure includes FM initialization and a radial force based on the fusion of the conventional US, Doppler, and elasticity images. Although it requires training a decision tree, the procedure is also automatic. We conjecture that the proposed algorithm is applicable to similar US images without any modifications.

Acknowledgement

We wish to thank the Associate Editor of the journal Dr.Spyretta Golemati and anonymous reviewers for their insightful comments. This research is sponsored by the Thailand Research Fund, grant BRG5780012, and the Center of Excellence in Biomedicine Engineering, Thammasat University.

References


Table 10

<table>
<thead>
<tr>
<th>Model</th>
<th>Comp. Time $T_{comp}$, sec</th>
<th>Correctly initialized $N_{corr}$, %</th>
<th>Correctly segmented $S_{corr}$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM(U. E. D)</td>
<td>11.39</td>
<td>2.07</td>
<td>0.18</td>
</tr>
<tr>
<td>FM(E. D)</td>
<td>11.29</td>
<td>1.78</td>
<td>0.16</td>
</tr>
<tr>
<td>FM(U. E)</td>
<td>11.26</td>
<td>1.91</td>
<td>0.17</td>
</tr>
<tr>
<td>FM(U. D)</td>
<td>11.30</td>
<td>1.79</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 11

<table>
<thead>
<tr>
<th>Model</th>
<th>$H_1$ Contour based measures</th>
<th>$H_2$ Contour based measures</th>
<th>$H_3$ Contour based measures</th>
<th>TP$_C$ Contour based measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM(U. E. D)</td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\sigma/\mu$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>FM(E. D)</td>
<td>7.54</td>
<td>3.40</td>
<td>0.45</td>
<td>2.40</td>
</tr>
<tr>
<td>FM(U. E)</td>
<td>12.72</td>
<td>9.49</td>
<td>0.75</td>
<td>3.78</td>
</tr>
<tr>
<td>FM(U. D)</td>
<td>42.95</td>
<td>41.50</td>
<td>0.97</td>
<td>30.61</td>
</tr>
<tr>
<td>FM(U. E. D)</td>
<td>17.89</td>
<td>10.00</td>
<td>0.56</td>
<td>5.58</td>
</tr>
</tbody>
</table>

Table 12

<table>
<thead>
<tr>
<th>Model</th>
<th>Jaccard</th>
<th>Dice</th>
<th>SEN</th>
<th>$H_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM(U. E. D)</td>
<td>$\mu$</td>
<td>$\sigma$</td>
<td>$\sigma/\mu$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>FM(E. D)</td>
<td>0.89</td>
<td>0.07</td>
<td>0.78</td>
<td>0.90</td>
</tr>
<tr>
<td>FM(U. E)</td>
<td>0.81</td>
<td>0.10</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>FM(U. D)</td>
<td>0.36</td>
<td>0.29</td>
<td>0.80</td>
<td>0.47</td>
</tr>
<tr>
<td>FM(U. E. D)</td>
<td>0.60</td>
<td>0.21</td>
<td>0.35</td>
<td>0.73</td>
</tr>
</tbody>
</table>