

Article

Efficient Segmentation of a Breast in B-Mode Ultrasound Tomography Using Three-Dimensional GrabCut (GC3D)

Shaode Yu ^{1,2}, Shibin Wu ^{1,2}, Ling Zhuang ³, Xinhua Wei ⁴, Mark Sak ^{3,5}, Duric Neb ^{3,5}, Jiani Hu ^{6,*} and Yaoqin Xie ^{1,*}

¹ Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; sd.yu@siat.ac.cn (S.Y.); sb.wu@siat.ac.cn (S.W.)

² Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China

³ Department of Oncology, the Karmanos Cancer Institute, Wayne State University, Detroit, MI 48201, USA; zhuangling@gmail.com (L.Z.); sakm@karmanos.org (M.S.); duric@karmanos.org (D.N.)

⁴ Department of Radiology, Guangzhou first Hospital, Guangzhou Medical University, Guangzhou 510180, China; weixinhua@aliyun.com

⁵ Delphinus Medical Technologies, Inc., Plymouth, Detroit, MI 46701, USA

⁶ Department of Radiology, Wayne State University, Detroit, MI 48201, USA

* Correspondence: jhu@med.wayne.edu (J.H.); yq.xie@siat.ac.cn (Y.X.); Tel.: +86-755-8639-2281 (Y.X.)

Received: 28 June 2017; Accepted: 4 August 2017; Published: 8 August 2017

Abstract: As an emerging modality for whole breast imaging, ultrasound tomography (UST), has been adopted for diagnostic purposes. Efficient segmentation of an entire breast in UST images plays an important role in quantitative tissue analysis and cancer diagnosis, while major existing methods suffer from considerable time consumption and intensive user interaction. This paper explores three-dimensional GrabCut (GC3D) for breast isolation in thirty reflection (B-mode) UST volumetric images. The algorithm can be conveniently initialized by localizing points to form a polygon, which covers the potential breast region. Moreover, two other variations of GrabCut and an active contour method were compared. Algorithm performance was evaluated from volume overlap ratios (*TO*, target overlap; *MO*, mean overlap; *FP*, false positive; *FN*, false negative) and time consumption. Experimental results indicate that GC3D considerably reduced the work load and achieved good performance ($TO = 0.84$; $MO = 0.91$; $FP = 0.006$; $FN = 0.16$) within an average of 1.2 min per volume. Furthermore, GC3D is not only user friendly, but also robust to various inputs, suggesting its great potential to facilitate clinical applications during whole-breast UST imaging. In the near future, the implemented GC3D can be easily automated to tackle B-mode UST volumetric images acquired from the updated imaging system.

Keywords: ultrasound tomography; whole breast imaging; image segmentation

1. Introduction

More than 1.3 million women worldwide are diagnosed with breast cancer each year, making it the second most common cancer [1,2]. In developed countries, one in eight women might develop this disease in their lifetime [1]; and in underdeveloped countries, the health burden of breast cancer is increasing [2]. It is crucial to screen breast cancer at an earlier stage, since early diagnosis of breast cancer increases treatment options and significantly reduces mortality [3].

As an emerging modality for whole breast imaging, ultrasound tomography (UST) offers many advantages in the screening and diagnosis of breast cancer over commonly used modalities, such as mammography, hand-held ultrasound, computerized tomography (CT) and magnetic resonance

imaging (MRI). Above all, it takes about one minute to scan an average breast using its ring array transducer. From the nipple to the chest wall, UST imaging acquires coronal slices of breast anatomy when the breast is surrounded by a water reservoir. With its contoured curves and recessed lighting, this design reduces human subjectivity and tissue deformation [4,5]. Furthermore, it can simultaneously create three distinct volumes among which the reflection (B-mode) imaging shows tissue structure; the transmission imaging measures changes in sound speed; and the attenuation imaging presents quantitative variance in the sound signal interacting with breast tissue [6,7]. With a universal threshold, breast tissues and tumor regions in UST images are comparably rendered as those in MR images [8]. Moreover, UST aids in tumor differentiation, especially for obscured tumors or tumors located within dense breasts. Importantly, UST images show consistency in breast density estimation as that measured by using mammographic images. This suggests UST may also be used to quantify breast cancer risk factors as well as cancer detection [9–13]. In general, the image acquisition is highly efficient, cost effective and safe, using no ionizing radiation.

Whole breast imaging has highlighted the importance of isolating the breast region from its surrounding water in reflection (B-mode) UST images, since B-mode images present breast anatomy and tissue structure with high in-plane resolution. Intuitively, image segmentation enhances the breast visualization, and accordingly, the observation of suspicious region and clinical diagnosis are easier to perform. Breast segmentation also improves quantitative tissue analysis [9–13] and other follow-up applications [14,15]. It has far reaching consequences in the longitudinal analysis of UST images to facilitate the quantification of breast tissue growth during treatment delivery [16–19], the characterization of physical breast density using intra-patient alignment of UST and MR images [20], and the rendering of breast tumors by co-registration of B-mode images with transmission, attenuation, MR or mammographic images [21,22]. For instance, facilitated by accurate segmentation of the breast in B-mode UST images, parenchymal changes in women undergoing tamoxifen therapy can be quantified and visualized with deformation field in high precision, up to 0.25 mm [16,17,19]. Therefore, women with or without mass lesions and patients receiving treatment, all can benefit from B-mode UST image segmentation to monitor the tissue growth, to detect suspicious regions, or to quantify the treatment outcome.

Efficient segmentation of the breast in B-mode UST images is challenging. During the procedure of image acquisition, the breast is immersed in water. Physically, the skin in UST images encloses the breast tissue and the background is approximately zero due to the surrounding water. In an ideal situation, the breast boundary can be distinguishable from its surrounding water region. However, two major kinds of artifacts exist in the clinic. The first one is caused by the limited distance between the chest wall and the ring transducer, which leads to severe distortion when scanning a large-sized breast. The second one is a circle-like shape with various intensities which mainly occurs around the nipple region. Other artifacts come from breathing, particularly from patients who do not have good control of their breathing. Furthermore, in ultrasound imaging, unavoidable noise is considered to be content related [23,24], which further imposes difficulties on B-mode UST image segmentation. Particularly, breasts have various shapes, sizes and densities across patients, and the imaging position is changeable. As shown in Figure 1, artifacts caused by limited space between the chest wall and the ring transducer severely alter image content, causing annoying stripes and unnatural breast shapes. Note that these distorted images should be tackled by manual segmentation. Artifacts may also cause circle-like shapes to look much like nipples, which might cause misleading image interpretation.

To the best of our knowledge, five methods have been used or developed for UST breast image segmentation. Balic et al. [25] used three-dimensional (3D) active contour [26] for breast boundary detection in attenuation UST images. The method is fast and fully automated. However, this method highly depends on the selection of an initial shape and no systematic evaluation has been used to quantify its performance. Hopp et al. [27] integrated edge detection and surface fitting for breast segmentation in reflection images, and experimental validation has verified its feasibility, while more than 40% of slices required manual correction. Sak et al. [9] took advantage of K-means [28]

and the thresholding method [29] to handle this problem, while proper parameters were needed in these methods for sound speed images. They also carried out a robustness test on attenuation images, and good to great overlap in the volumes of tissue was achieved [30]. Hopp et al. [31] updated the 3D active contour [32] by applying gradient vector flow [33] and encoded aperture characteristics as additional weighting terms. The algorithm is feasible for automated B-mode images and achieves an average surface deviation from manual segmentation of 2.7 mm [31]. Under this situation, manual segmentation of the breast outline has been widely accepted in ongoing studies [10–14]. Apart from poor time consumption, unstable outcome and intensive user interaction, an additional assumption is imposed in [25,27] that the breast boundary in each coronal slice can be modeled with an ellipse.

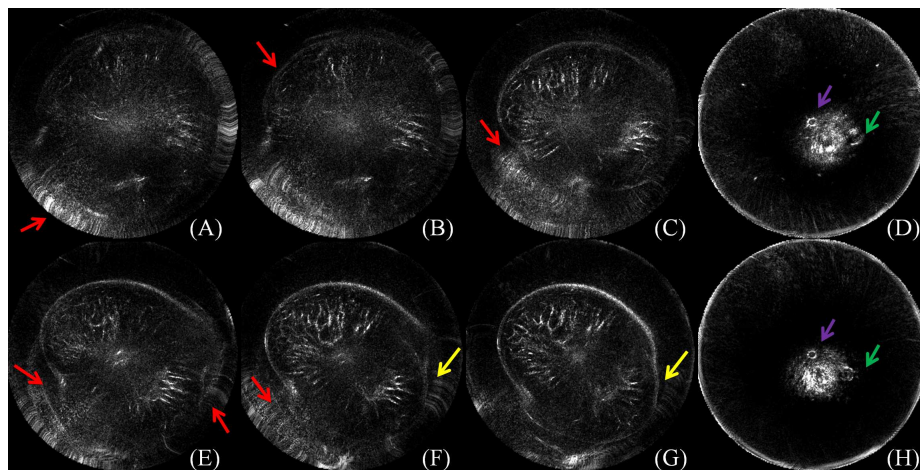


Figure 1. (A–H) The challenges of breast segmentation in B-mode ultrasound tomography (UST) images. Red arrows indicate the major artifacts due to the limited space between the breast boundary and the ring transducer; yellow arrows indicate unnatural distortion of the breast shapes; purple arrows indicate the second kind of circle-like artifacts; and green arrows indicate the real nipple regions. Annoying stripes (A–C, E–G) and unnatural breast shapes (C, E–G) are observed. Moreover, artifacts with circle-like shapes (indicated by purple arrows) look like nipples (indicated by green arrows) (D, H). The figure can be enlarged to view the details.

Massive algorithms have been proposed for image segmentation [34–38], while few methods are explored to fulfill the necessities of UST image segmentation in terms of accuracy and efficiency. Both the K-means [28] and the thresholding [29] methods are relatively straightforward and only pixel intensities are taken into consideration. Watershed methods classify pixels into regions using gradient descent and analyzes weak points along region boundaries [39,40]. SNAKE (also known as active contour) utilizes splines to fit image structures of lines and edges [26,32,37]. Watershed and SNAKE methods benefit from the distinguishable intensity difference between the breast boundary and its surrounding regions. Notably, for SNAKE, the initialization shape plays an important role in the successful segmentation of the UST images [25]. Whole breast segmentation can also be fulfilled by the graph cuts methods [41–45]. Among these methods, GrabCut [42] is more preferable for B-mode UST images. At first GrabCut integrates texture and edge information, and it models the foreground and the background with Gaussian mixture models (GMMs) [43]. Besides, it substantially simplifies the manual interaction by drawing a rectangle for algorithm initialization that saves time. In addition, GrabCut allows users to incorporate prior knowledge, to validate results and to correct errors in the process of segmentation.

This paper explores the performance of three variations of the GrabCut method in whole breast segmentation using B-mode UST images. One variation is the modified GrabCut (GC) for slice-wise image segmentation and the other two variations are 3D implementations. The first 3D variation takes 6-connected neighboring voxels (GC3D) and the second variation takes 26-connected voxels

(GC3D_26) into foreground/background modeling. Additionally, this study compares these results to a previously simplified SNAKE (sSNAKE) [46]. Altogether, the four semi-automated methods are investigated in this paper, and the performance of breast UST image segmentation is evaluated in terms of accuracy, time consumption and ease-of-use.

2. Materials and Methods

2.1. Data Collection

Thirty reflection (B-mode) UST volumetric images were collected and analyzed (SoftVue™, Delphinus Medical Technologies, Plymouth, MI, USA). The reflection imaging reveals the internal anatomical structures of breasts, and therefore, it is much easier to interpret than those images from sound speed imaging and attenuation imaging. The physical resolution of B-mode UST images is $[0.5, 0.5, 2.0] \text{ mm}^3$, and the matrix size in the coronal plane is $[512, 512]$. The imaging system, SoftVue™, has received FDA clearance for diagnostic imaging purposes and is not intended for use as a replacement for screening mammography (<http://www.delphinusmt.com/>).

2.2. Image Preprocessing

A physician with more than fifteen years of work experience was asked to define the starting and the ending slice by using the software package ImageJ [47]. It is observed that slices near the chest wall are distorted by artifacts, while slices around the mammilla could be extensively blurred [13]. How to select the starting and the ending slice is illustrated in Figure 2 with a representative image. The image is composed of 36 slices. The top row shows the 12th–15th tomogram near the chest wall, and the bottom row is the 29th–32nd tomogram near the nipple position. Under the context of selecting the starting and the ending slice, adjacent tomograms should be fully considered. In this case, (B) was chosen to be the first slice because of its visible gap between artifact distortion and the breast boundary, and (G) was chosen as the last slice with the present nipple. Therefore, the index of remaining slices ranges from the 13th–31st in this case. In other words, 19 slices are left for follow-up image segmentation.

To the dataset in this study, the number of remaining slices ranged from 11 to 21 with an average number of 16 slices. That means 22 mm–42 mm, an average of 32 mm, of the breast was analyzed. In general, 13–18 slices were removed from each image stack as a result of this process.

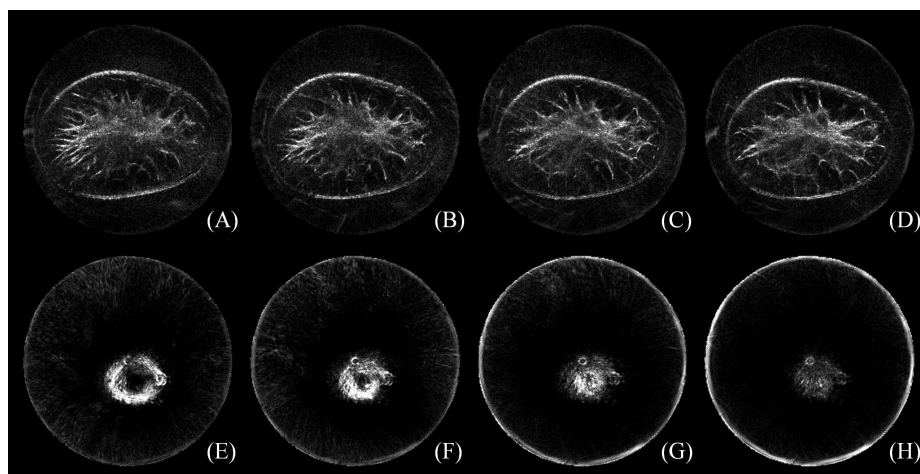


Figure 2. (A–H) The process of how to select the starting and the ending slice with a representative image. Under the context of four tomograms in the top row, (B) was chosen to be the first slice because of its visible gap between artifact distortion and the breast boundary; in the bottom row, (G) was chosen as the last slice with the nipple present. The figure can be enlarged to view the details.

2.3. Benchmark Building

The benchmark (also known as gold standard data) is created by a senior physician with the help of a free-form curve fitting method for slice-wise image processing [46]. The method consists of point localization, Hermite cubic curve interpolation and active contour for boundary segmentation. While different from the original method, the first two steps were employed to outline the breast region, because the points were elaborately located. During the process of placing landmarks by the physician, the number of points can be as many as desired localized on the breast boundary. Then, nine points are uniformly generated between an adjacent point pair using the Hermite cubic curve interpolation [48]. The closed region is masked with the function *poly2mask* in MATLAB (MathWorks, Natick, MA, USA). After that, outlined regions were further verified by two other physicians to minimize the possible bias. A house-built MATLAB program was implemented to display the original image and its corresponding image with the outlined breast region side by side, and the two physicians worked together to browse the image pair slice by slice. If the breast region was not appropriate, the slice would be re-outlined. Finally, these slice with outlined breast region are stacked to form the ground truth data for the validation of segmentation algorithms.

Figure 3 presents the process of the manual outline of the breast region in one slice. A physician first localized points on the breast boundary. Then, the red points in (B) were interpolated by using Hermit curve fitting as shown in (C) with the green line. Finally, the region in the closed green line was taken as the breast region. The pixels in the breast region were marked as 1, and those outside the breast region were marked as 0 for binary storage. In this slice, 46 points were manually localized.

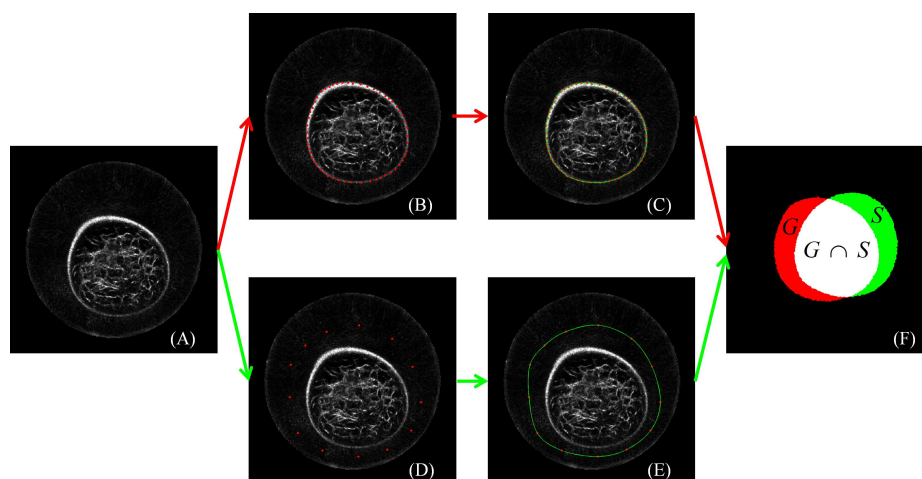


Figure 3. The semantic description of benchmark building, algorithm initialization and performance evaluation with one slice as an example. (A) The source image. (B) Elaborately manual localization of points on the breast boundary. (C) Automated closed curve generation. (D) User interaction for algorithm initialization with points placed between the ring transducers and the breast boundary. (E) Automated polygon generation. (F) Performance quantification by comparing the ground truth G and the segmentation result S . Note that (A, B, C), (A, D, E) and (F) correspond to the process of benchmark building, algorithm initialization and performance evaluation, respectively. The figure can be enlarged to view the details.

2.4. GC3D

2.4.1. Image Segmentation

Given an image, a trimap T is manually outlined ($T = T_B \cup T_F \cup T_U$). T_B , T_F and T_U correspond to the set of pixels in the background, the foreground and the unknown region. The pixel intensity is one of the values $\mathbf{z} = \{z_1, \dots, z_i, \dots, z_n\}$. Image segmentation is to group each pixel in T_U to T_B ($\alpha = 0$) or T_F ($\alpha = 1$). Note that α indicates the label of voxels. In GrabCut, α evolves to be either 0 or 1 for each

voxel at the end of the iterations, i.e., $\alpha = \{0, 1\}$. A parameter $\underline{\theta}$ describes the grey-level distributions, $\underline{\theta} = \{h(\mathbf{z}; \alpha)\}$. The task of image segmentation becomes inferring the unknown α for pixels in T_U from the learned model $\underline{\theta}$ based on the given data \mathbf{z} .

2.4.2. s - t Graph

An image can be described as an undirected graph, $g = \langle v, e \rangle$, where v is the vertex or node set and e is the edge set [49–51]. Two special vertices are added in the graph as shown in Figure 4. One is the source point (s), and the other is the sink point (t). There are two kinds of edges, solid lines (N -links) and dotted lines (T -links). The former represents the edges between adjacent ordinary vertices, and the latter indicates the edges between s (or t) and the ordinary vertices. In addition, the edge is weighted based on the similarity between the vertices on it. Shown in Figure 4, any given pixel is strictly and loosely connected to 4 and 8 neighboring pixels, respectively. One can imagine that in 3D space, there are 6 and 26 neighboring voxels connected to the central one.

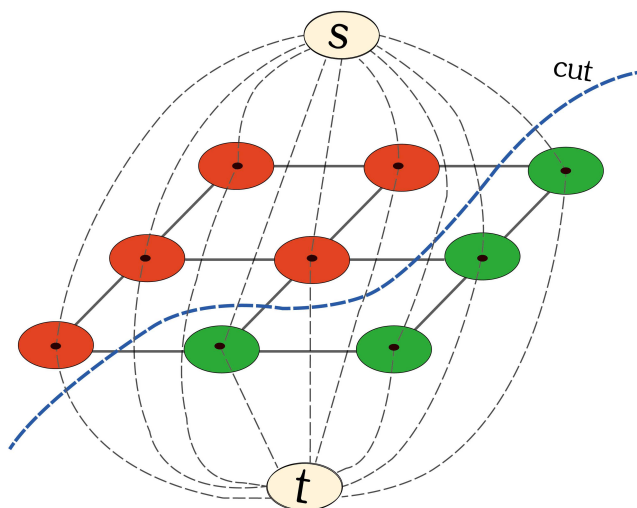


Figure 4. s - t graph. An image can be expressed as an undirected graph with two additional vertices (s and t). Image segmentation is to find a cut (blue dotted line) that separates the foreground (red rounds) from the background (green rounds). The figure can be enlarged to view the details.

2.4.3. Segmentation by Iterative Energy Minimization

GrabCut fulfills image segmentation by energy minimization. The energy function \mathbf{E} is defined in Equation 1, where $U(\alpha, \mathbf{k}, \underline{\theta}, \mathbf{z})$ concerns the weights in N -links (the region energy) and $V(\alpha, \mathbf{z})$ concerns the weights in T -links (the boundary energy). The vector $\mathbf{k} = \{k_1, \dots, k_l, \dots, k_N\}$ distinguishes GrabCut from other graph cuts methods in the optimization framework, since GrabCut makes use of Gaussian mixture models (GMMs) to model the background and the foreground pixels. GMM is a probabilistic model. It assumes all the data points are obtained from a mixture of a finite number of Gaussian distributions with unknown parameters. To each GMM in the GrabCut method, it contains K components, and each pixel is assigned a unique component ($k_l \in \{1, \dots, K\}$).

$$\mathbf{E}(\alpha, \mathbf{k}, \underline{\theta}, \mathbf{z}) = U(\alpha, \mathbf{k}, \underline{\theta}, \mathbf{z}) + V(\alpha, \mathbf{z}), \quad (1)$$

GrabCut consists of incomplete initialization and iterative energy minimization. The required user interaction is to draw a rectangle. The region outside and inside the rectangle corresponds to T_B and T_U . Then, the background and the foreground GMM components are preliminarily initialized with pixels in T_B ($\alpha = 0$) and in T_U ($\alpha = 1$), respectively. Next, a s - t graph is built, and a coarse

segmentation is obtained by using the max-flow min-cut algorithm [52,53]. With updated T_B and T_F , new GMMs are learned, and a new graph is created; thereby, an intermediate segmentation is achieved. Iteratively, GrabCut runs until the energy function E converges. If the segmentation result is unsatisfactory, further user editing can be utilized as a remedy for image segmentation.

2.4.4. Implementation of GC3D

Instead of drawing a rectangle, GC3D generates a polygon for algorithm initialization. It requires the points to be localized between the ring transducer and the breast boundary in the coronal slice nearest to the chest wall. If the polygon covers the whole potential breast region, the initialization is correct. The polygon is then duplicated and propagated to the other slices to save time and energy. At last, a volume mask is created for incomplete labeling.

For medical image modeling, such as in GC3D, each component (θ) contains three parameters, one weight (π), one mean (μ) and one covariance (δ). How to calculate the boundary energy $U(\alpha, \mathbf{k}, \underline{\theta}, \mathbf{z})$ is shown as follows:

$$U(\alpha, \mathbf{k}, \underline{\theta}, \mathbf{z}) = \sum_n D(\alpha_n, k_n, \underline{\theta}, z_n) = \sum_n \{-\log p(z_n | \alpha_n, k_n, \underline{\theta}) - \log \pi(\alpha_n, k_n)\}, \quad (2)$$

where $p()$ is a Gaussian probability distribution and $\pi()$ are a mixture of weight coefficients. Here we define T -links between nodes and s are noted as T_s , and T -links between each node and t are as T_t . If one pixel belongs to T_F , $T_s = W$ and $T_t = 0$, otherwise $T_s = 0$ and $T_t = W$. Note that W is the largest, constant edge weight in the s - t graph. However, if one pixel x is in T_U , its probability is formulated in Equation 3 where the symbol $'$ denotes the transpose operation.

$$D(x) = -\log \sum_{i=1}^K \left\{ \frac{\pi(\alpha_x, i)}{\sqrt{\det(\delta(\alpha_x, i))}} \times \exp\left(\frac{1}{2}(z_x - \mu(\alpha_x, i))' \delta(\alpha_x, i)^{-1} [z_x - \mu(\alpha_x, i)]\right) \right\}. \quad (3)$$

The calculation of the region energy $V(\alpha, \mathbf{z})$ is related to N -links, and the major difference comes from how to weigh the edges between these connected nodes in the graph in 3D space. Assuming $N(x, y)$ is the edge weight between any connected nodes of x and y , $V(\alpha, \mathbf{z})$ is computed as follows,

$$V(\alpha, \mathbf{z}) = \sum_{(x,y) \in e} [\alpha_x \neq \alpha_y] N(x, y), \quad (4)$$

and the inequality operator (\neq) indicates that it yields 0 when α_x and α_y are equal and 1 if otherwise. Defining Euclidean distance as $d(x, y)$, the edge weight $N(x, y)$ is

$$N(x, y) \times d(x, y) = w \times e^{-\beta \|I_x - I_y\|} \quad (5)$$

in which $\|I_x - I_y\|$ represents the intensity difference and w is a constant. In GC3D, 6-connected voxels are considered, and $d(x, y) = 1$. As for β , it is defined as $\frac{1}{\beta} = \frac{2}{P} \sum_{i=1}^P \sum_{j=1}^Q \|I_i - I_j\|^2$, where P is the voxel number in the volume and Q indicates the number of the connected neighbors to the central voxel.

In summary, two parameters are tunable. One is K , the number of Gaussian components in each GMM; and the other is w , the constant weighting value in Equation (5). Off-line experiments were carried out to find the optimal K ($K = 2, 3, \dots, 9$) and w ($w = 30, 40, \dots, 90$). The results indicated that GC3D is consistent regarding to the change of the tunable parameters ($\pm 1\%$ changes in MO). Therefore, we use the suggested values, i.e., $K = 5$ [42] and $w = 50$ [43] in this study.

2.5. Experiment Design

The reliability of gold standard data is estimated with the intra- and the inter-observer variability analysis. Intra-observer analysis is based on two sessions of outlining the breast regions by a senior

physician (15-year work experience), while inter-observer comparison is between the senior physician and a junior physician (2-year work experience). Three months later, the re-test was performed, following the same test procedure in the benchmark building. Note that the first and the second session from the senior physician is marked as **S** and **S'**, and the session from the junior physician is marked as **J**.

Four semi-automatic methods were evaluated. Besides GC3D, two other implementations of GrabCut are involved: GC with monochrome slices as its input, and GC3D_26 taking 26-connected neighboring voxels into the computation. In addition, sSNAKE was compared which allows for control point localization [46]. Different from the original method, control points were placed near but not on the breast boundary before contour propagation.

Two more experiments were further conducted. One was to investigate whether localizing points to form a polygon is more user-friendly than drawing a rectangle in the algorithm initialization. The other was to study the robustness when different users (two non-physicians and a junior physician) manipulate GC3D.

2.6. Performance Evaluation

To evaluate the reliability of gold standard data, Dice similarity index was used [54]. It is defined as below,

$$DICE = 2 \frac{|X \cap Y|}{|X| + |Y|}, \quad (6)$$

where X and Y are segmentation results, and $|\cdot|$ denotes the number of elements in one volume.

Algorithm performance was estimated from two volume overlap agreement measures and two overlap error measures. The former two measures are target overlap (TO) and mean overlap (MO), and the latter two are false positive (FP) and false negative (FN) errors [55]. Given the segmentation result (S) and the ground truth (G), TO , MO , FP and FN are correspondingly defined as below,

$$TO = \frac{|G \cap S|}{|G|}, \quad (7)$$

$$MO = 2 \frac{|G \cap S|}{|G| + |S|}, \quad (8)$$

$$FP = \frac{|S| - |G \cap S|}{|G|}, \quad (9)$$

$$FN = \frac{|G| - |G \cap S|}{|G|}. \quad (10)$$

The average time consumption (TC) for each volume was also concerned. It accounts for the time spent on the manual initialization to the end of whole volume segmentation. Note that the user was aware of the time consumption evaluation when performing the task. In Equation 11, tc_i stands for the time cost for the i th volume, and $n = 30$.

$$TC = \frac{1}{n} \sum_{i=1}^n tc_i. \quad (11)$$

2.7. Software Platform

Three variations of the GrabCut method (GC, GC3D and GC3D_26) were implemented with C++ (Visual Studio 2010, <https://www.visualstudio.com/>). The computer was equipped with 8 Intel (R) Cores (TM) of 3.70 GHz and 8 GB DDR RAM. The algorithm implementation did not use any optimizations or strategies for algorithm acceleration. Other involved software include OpenCV (<http://opencv.org/>) and ITK (<http://www.itk.org/>). In addition, sSNAKE [46] was utilized with MATLAB (MathWorks, Natick, MA, USA).

To evaluate the reliability of ground truth building, intra-class correlation (ICC) analysis was employed. Two-way mixed effects model (ICC3,1) was used to assess the intra-observer test-retest reliability, and two-way random-effects model (ICC2, k) was used to quantify the inter-observer reliability [56,57]. ICC coefficient values (r) ranging from 0.81 to 1.00 indicate excellent reliability; 0.61 to 0.80 good reliability; 0.41 to 0.60 moderate reliability; 0.21 to 0.40 fair reliability and below 0.2 poor reliability [58]. The paired sample t test was used to compare the differences in measurements between observers and a significance level of 0.05 was set. Statistical analysis was performed using the R software, Version 3.0.1 (<http://www.R-project.org>).

3. Results

3.1. Reliability of the Ground Truth Data

The segmentation results are in good overlap and $DICE$ values are generally greater than 0.96. Furthermore, excellent intra- and inter-observer reliability are observed and ICC coefficient values are all greater than 0.95. In particular, the paired sample t test indicates that no significant difference is found between two sessions of breast segmentation by the senior physician ($p = 0.4359$) and between the segmentation results from the senior and the junior physicians ($p \geq 0.3934$) in Table 1.

Table 1. Reliability of the ground truth data. p values are from the paired t test of the segmentation results. The (*) indicates the intra-observer correlation, and (+) denotes the inter-observer correlation.

	SS'	SJ	S'J
$DICE$	0.9762 ± 0.0417	0.9641 ± 0.0432	0.9657 ± 0.0556
p value	0.4359	0.3934	0.7649
ICC	0.9517 (*)	0.9588 (+)	0.9757 (+)

3.2. Perceived Evaluation

The perceived evaluation of eight cases is demonstrated in Figure 5. In each case, lines in red, green, blue, yellow and pink color correspond to segmentation results from manual delineation, GC, GC3D, GC3D_26 and sSNAKE, respectively. It shows that GC, GC3D and sSNAKE successfully separate the breast from the water region, while GC3D_26 causes over-segmentation. By enlarging the figure, we find that the outlines from sSNAKE are slightly outside of the breast boundary (A, E), while GC and GC3D result in tight outlines.

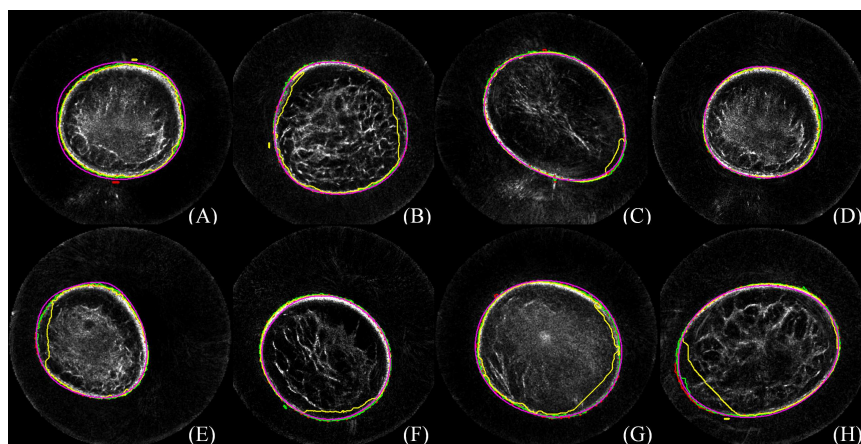


Figure 5. Perceived evaluation of breast segmentation based on eight cases. In each case, lines in red, green, blue, yellow and pink color correspond to segmentation results from manual delineation, GrabCut (GC), Three-dimensional GrabCut (GC3D), GC3D_26 and simplified SNAKE (sSNAKE), respectively. The figure can be enlarged to view the details.

Figure 6 illustrates a case from three views. It shows that algorithms successfully separate the breast from the water region and no major difference is observed among algorithms. However, minor differences can be found. As indicated by the red arrow, GC3D_26 generates the least smooth breast boundary as seen in the transverse view, and GC leads to jagged edges at the nipple region in the coronal view.

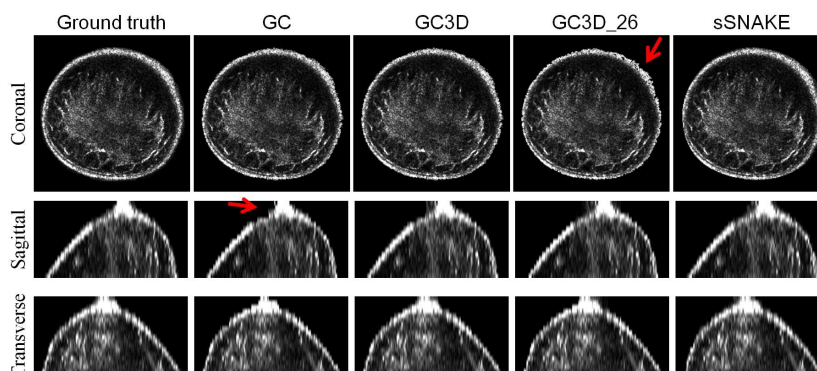


Figure 6. Perceived evaluation of a case from three views. Algorithms successfully separate the breast region from the background. Minor defects are indicated with red arrows; GC3D_26 generates the least smooth breast boundary, and GC leads to jagged edges at the nipple region. Images are interpolated in the sagittal and coronal view and cropped in three views for display purposes. The figure can be enlarged to view the details.

3.3. Quantitative Comparison

The quantitative comparison is shown in Table 2. It was found that sSNAKE achieved the top accuracy (*TO* and *MO*), followed by GC3D and GC, while it also had the highest *FN* value. On the other hand, GC3D is the most convenient, because it requires the least number of points (≈ 12 points per volume) to be placed for algorithm initialization. Compared to GC3D, GC, sSNAKE and benchmark building takes 8, 19 and 35 times more points to a volumetric image, respectively. In addition, *TC* values show that GC3D could process a UST volume within about 1.2 min, followed by GC3D_26 and GC.

Table 2. Quantitative comparison of breast segmentation. The (\ddagger) indicates that it is not comparable between GC3D and sSNAKE because of different implementation languages.

	TO	MO	FN	FP	TC (Min)	No. of Points Localized
Benchmark					11.8 ± 4.82	423 ± 87
GC	0.82	0.90	0.005	0.18	2.37 ± 0.84	98 ± 21
GC3D	0.84	0.91	0.006	0.16	1.23 ± 0.62	12 ± 3
GC3D_26	0.65	0.75	0.023	0.35	1.48 ± 0.65	12 ± 3
sSNAKE (\ddagger)	0.89	0.93	0.029	0.11	36.8 ± 5.16	226 ± 32

3.4. Ease-Of-Use

Figure 7 shows coronal images with different inputs (A, C, E, G) and corresponding results of the segmented breast regions (B, D, F, H). Note that only the slice with the maximum breast region is illustrated. The red arrows in (B, D) point to the regions with incorrect segmentation. It indicates that polygons lead to better segmentation (F, H) compared to results with rectangles as the input (B, D), respectively.

Performance estimation of GC3D in breast segmentation with different initialization approaches is shown in Table 3. In comparison to the polygon input, a rectangle input results in lower accuracy (*TO* and *MO*) and higher error (*FN* and *FP*). It requires 3–4 tries to draw an appropriate rectangle and thereby prolongs the time ($\approx 50\%$ increase). Note that an appropriate rectangle requires the rectangle

covering the breast region while not interacting with the ring transducer. Due to several attempts at selecting an appropriate rectangle, the segmentation time increases. However, when the rectangle contains voxels on the ring transducer, GMMs would be misleading, and thereby, the performance in UST segmentation decreases.

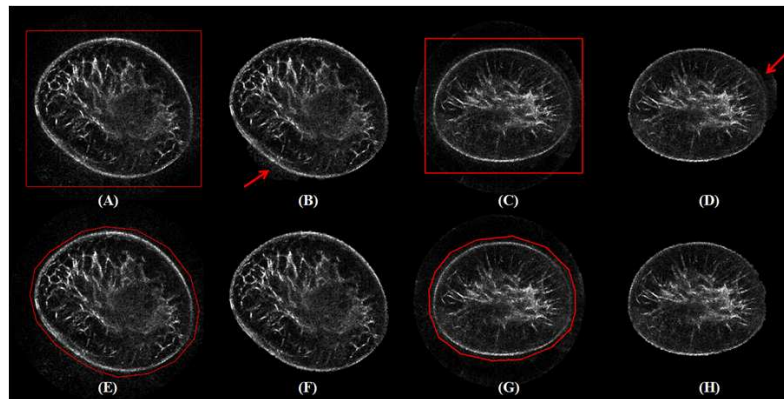


Figure 7. Algorithm initialization of GC3D by drawing a rectangle versus localizing points to form a polygon. Two cases are illustrated from the coronal view and red arrows point to over-segmented regions. The figure can be enlarged to view the details.

Table 3. Comparison of breast segmentation using GC3D with different initialization approaches.

	TO	MO	FN	FP	TC (Min)
Rectangle	0.83	0.88	0.021	0.17	1.85 ± 1.12
Polygon	0.84	0.91	0.006	0.16	1.23 ± 0.62

3.5. Robustness

The robustness analysis is carried out by manipulating GC3D and points are localized for algorithm initialization. Quantitative results from three users (two non-physicians and a physician) are demonstrated in Figure 8. It indicates that the metric values are very close to each other (#01, non-physician: $TO = 0.85$, $MO = 0.91$, $FP = 0.006$, and $FN = 0.154$; #02, non-physician: $TO = 0.84$, $MO = 0.91$, $FP = 0.007$, and $FN = 0.155$; and #03, physician: $TO = 0.85$, $MO = 0.91$, $FP = 0.007$, and $FN = 0.155$). One-way ANOVA indicates no significant difference among each metric (p -value > 0.99). The ICC coefficients revealed the measurement agreement among users ($r > 0.98$).

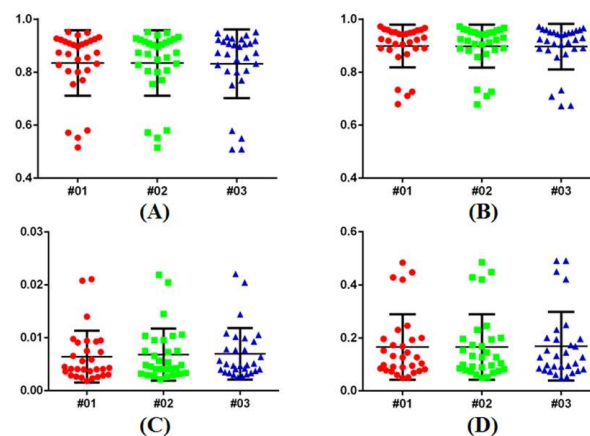


Figure 8. Robustness analysis of GC3D manipulated by different users on UST image segmentation. (A) TO ; (B) MO ; (C) FP ; (D) FN . No significant difference is found among observers in each metric.

3.6. Failure Case Analysis

Figure 8 reveals that there are four cases in which GC3D failed to segment the major breast regions ($TO < 0.65$). Subsequently, these failure cases were represented and further edited with user interaction. Results before and after user editing are illustrated in Figure 9. The editing was fulfilled with pink dots on the breast boundary, and these voxels were labeled into the foreground (T_F). Note that the pink dots are visible after the figure is enlarged.

With labels on the breast boundary, the segmentation after editing (D, H, L, P) was more accurate than those before editing (B, F, J, N), respectively. Interestingly, the failure cases (A, E, I, M) shared similar tissue contrast, i.e., black boundary and bright inner tissue in perception. It was also found that with additional labeling, improved breast segmentation was achieved, and major breast regions were separated.

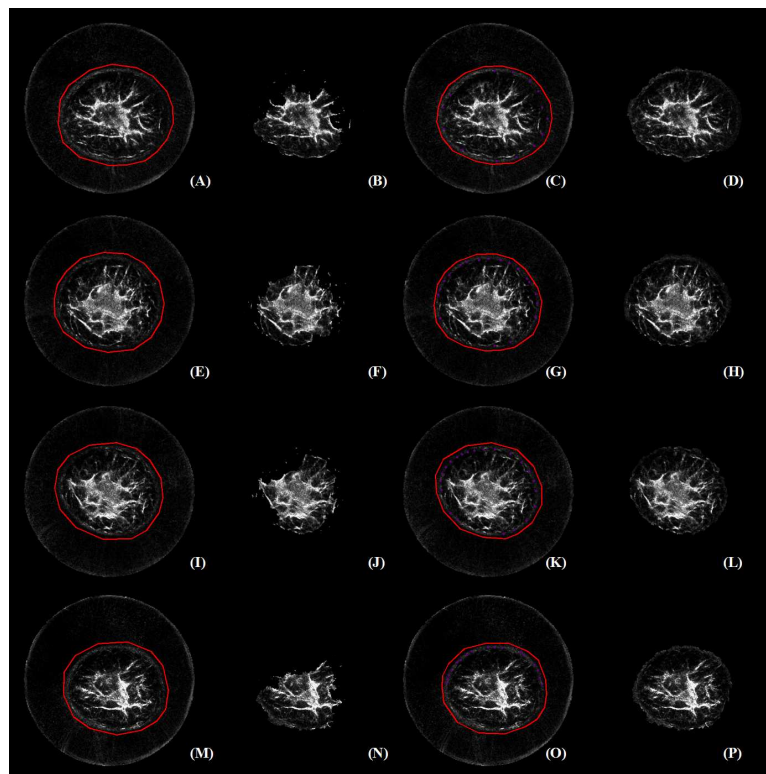


Figure 9. Failure case analysis and additional user editing for image segmentation. It is observed that failed cases share similar image contrast, i.e., black boundary and bright inner tissue. In addition, user editing benefits accurate breast segmentation. The figure can be enlarged to view the details.

4. Discussion

Efficient segmentation of the entire breast in B-mode UST images plays an important role in visualization, analysis and diagnosis of breast disease. This paper explored various implementations of GrabCut (GC, GC3D and GC3D_26) and the simplified SNAKE (sSNAKE) for semi-automatic breast segmentation. Quantitative comparison based on thirty breast volumes indicated that GC3D achieves a good balance of segmentation accuracy, time investment, ease-of-use and robustness. In particular, GC3D only requires users to place several points for algorithm initialization, which improves usability. Therefore, to trade-off between efficiency and accuracy, GC3D is recommended for the breast segmentation task with respect to UST image applications.

The reliability of benchmark building was investigated based on the intra- and the inter-observer variability analysis (Table 1). Statistical analysis showed that there were no significant differences between two sessions of the senior physician and between the session of the senior and the junior

physicians. Furthermore, the accuracy of built data sets as gold standard was validated and the segmentation results were in good coincidence ($DICE > 0.96$). One reason for the excellent correlation in breast segmentation is that the breast regions were outlined by physicians with professional background. The other reason is that two other physicians additionally checked the segmentation results and the benchmark was built with three observers by consensus. It should be noted that the segmentation result of the first session by the senior physician was taken as the gold standard data for further algorithm comparison.

Four semi-automated algorithms were evaluated and GC3D achieved good segmentation within an acceptable time (Table 2). It outperformed GC3D_26. The difference between the algorithms was that GC3D_26 took 20 more connected voxels into consideration. Therefore, we can conclude that additional voxel connections have a negative effect on the image segmentation. Voxels at the corner might induce inaccurate estimation of the region energy $V(\alpha, \mathbf{z})$ (Equation (1)) because of identical $d(x, y)$ (Equation (5)) and anisotropic image resolution. On the other hand, GC3D was better than GC on TO , MO and FN . In GC3D, a s - t graph was built with voxels in a volume; and in GC, it was based on voxels in a slice. Consequently, GC3D benefits from accurate estimation of GMMs with a large number of sample voxels. It was also observed that GC achieved slightly better FP than GC3D which might be due to the dedicated point localization in each slice when performing GC. Compared to GC3D with one coarse initialization, slice-wise GC refined the localization of points in each slice, and thereby, less background voxels (T_B) were grouped into the unknown region (T_U). However, GC required an average of 98 points to be placed for each volumetric image, which increased time and user interaction. If the points were localized once and then propagated to other slices, the time was reduced (1.52 ± 0.46 min), while the segmentation accuracy decreased ($TO = 0.79$, $MO = 0.84$, $FP = 0.012$ and $FN = 0.21$). It should be noted that GC is easy for deployment on multi-core computers, which will cut down the segmentation time with parallel programming [59,60]. Although sSNAKE suffers from intensive user interaction in addition to complex parameter tuning, it showed promising results in slice-wise breast segmentation in UST images.

It is easy for users to perform GC3D for B-mode UST image segmentation. What a user needs to do is localize several points between the breast boundary and the ring transducer. Compared to GC and sSNAKE, GC3D requires far fewer points to be localized, which is especially beneficial in a setting with large-scale UST image analysis (Table 2). In addition, the point localization is more flexible than drawing a rectangle, regardless of the breast size, shape and position (Table 3). The latter operation increases time cost by 50% because 3–4 attempts are needed in practice. Particularly, drawing a rectangle results in less accurate GMMs, since more voxels in the background (T_B) are taken into the unknown region (T_U). B-mode UST images with lower contrast between ambiguous breast boundaries and bright inner tissues (Figure 9) may need additional touch-ups on breast boundaries. Notably, in addition to the ease of use, its robustness to different users is another main advantage of GC3D in clinical applications. In other words, GC3D can be used by non-physicians, which facilitates large-scale image studies (Figure 8). At present, the manual outline of the breast boundary in UST images is accepted, even though it is laborious and tedious. With the help of GC3D, non-physicians may be trusted to outline the breast boundary after distorted slices are removed, and thereby, save time and energy.

Several issues were identified and the most concerned issues were from image preprocessing and benchmark building. First, removing distorted images reduces the challenges in B-mode UST image segmentation. It should be recognized that removing the distorted images is necessary (Figure 1). It is known to us that artifact removal is extremely difficult in medical imaging [61], and the most fundamental solution comes from the upgrading of imaging devices [62,63]. Therefore, in this paper, we followed previous studies [9,12,13] to define the starting and the ending slice. After distorted images were removed, the image contrast is enhanced, and a certain degree of artifacts and noise become invisible in the background region. Note that these artifacts and noise were tackled by the implemented GC3D. Following the selection of the starting and ending slice, an average breast of

about 32 mm remained. Since all slices provide useful information for tissue analysis and disease diagnosis, these distorted slices were then processed manually. This made the reproducibility of the selection of the starting and ending slice not as critical.

A few of the other concerns were technical issues. Above all, it should be recognized that a matting tool would likely perform as well as GrabCut in this task. These tools are not limited to graph cut methods [41–43] and active contours [26,32,33]; but also for level sets [40], multilevel thresholding [64] and neural networks [65,66]. However, at the preliminary stage of B-mode UST image segmentation, it is helpful to implement a user-friendly tool to facilitate clinical applications, and consequently, this study fills the gap in existing research. This study investigated three variations of GrabCut, since GrabCut has shown superiority among various matting methods [42], and their performance was validated in the task of reflection UST image segmentation. Furthermore, in order to establish the robustness and the speed advantage of GC3D, it is quite meaningful to compare the performance between GC3D and other matting methods [67], such as 3D active contours. However, a previous study [25] indicated that the success of the active contour for attenuation mode UST image segmentation is prone to the suitability of the initialization shape. Thus, a simplified active contour (sSNAKE) [46] was evaluated. It resulted in accurate segmentation, while intensive user interaction hampered its wide application. Moreover, only a handful of literatures are related to UST image segmentation [9,25,27,31]. Besides aforementioned challenges, one probable reason is that UST images are not publicly available, since UST imaging is an emerging modality. In such a situation, the major progress of UST image analysis is made by Delphinus Medical Technologies, Inc. [68] and the team from Karlsruhe Institute of Technology [69]. The manual segmentation of UST images has been currently preferred, particularly because the imaging system is being upgraded. Thus, only a few of articles have been published on this topic. Last but not the least, automated image segmentation is always desirable [25,30,31]. In this study, due to the fixed location of the ring transducer in the B-mode UST images, GC3D has the full potential to be automated. UST imaging device, SoftVueTM, is continually upgraded, and high-quality UST images can be acquired in the near future. At that time, it will be possible to isolate the whole breast in B-mode UST images by using GC3D in a fully-automated manner. By applying fully-automated methods, reproducible results can be obtained, and diagnostic time can be significantly reduced. As a result, radiologists and physicians alike can be released from this kind of tedious and time-consuming work.

5. Conclusions

Three-dimensional GrabCut (GC3D) can be utilized for efficient segmentation of the breast in reflection (B-mode) UST volumetric images. This method has many advantages. It achieves good performance in an acceptable time; and it is user friendly and enables the process to be used for large scale studies. It has the potential to be fully automated and thus save time to release physicians and radiologists from manual breast segmentation. Our future work will focus on the automation of GC3D by incorporating the results from this study.

Acknowledgments: The authors would like to thank the editor and reviewers for their valuable advices that have helped to improve the paper quality. This work is supported by the grants of the National Key Research Program of China (Grant No. 2016YFC0105102), the Union of Production, Study and Research Project of Guangdong Province (Grant No. 2015B090901039), the Technological Breakthrough Project of Shenzhen City (Grant No. JSGG20160229203812944), the Shenzhen Fundamental Research Program (JCYJ201500731154850923), the Natural Science Foundation of Guangdong Province (Grant No. 2014A030312006) and the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology.

Author Contributions: S.Y., J.H. and Y.X. conceived of and designed the experiments. S.Y., S.W., X.W. and Y.X. performed the experiments. S.Y. and S.W. analyzed the data. M.S., N.D. and J.H. contributed reagents/materials/analysis tools. S.Y. wrote the paper. L.Z., M.S., N.D. J.H. and Y.X. discussed and proof-read the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analysis or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

UST	Ultrasound tomography
CT	Computerized tomography
MRI	Magnetic resonance imaging
GC3D	Three-dimensional GrabCut that takes six-connected neighboring voxels into computing
GC	GrabCut
GC3D_26	Three-dimensional GrabCut that takes 26-connected neighboring voxels into computing
sSNAKE	A simplified SNAKE or active contour method
GMM	Gaussian mixture model
TO	Target overlap
MO	Mean overlap
FP	False positive
FN	False negative
TC	Time consumption

References

1. Siegel, R.; Miller, K.; Jemal, A. Cancer Statistics 2015. *CA Cancer J. Clin.* **2015**, *65*, 5–29.
2. Fan, L.; Strasser-Weippl, K.; Li, J.; St Louis, J.; Finkelstein, D.M.; Yu, K.; Chen, W.; Shao, Z.; Goss, P.E. Breast cancer in China. *Lancet Oncol.* **2014**, *15*, e279–e289.
3. Sak, M.A.; Littrup, P.J.; Duric, N.; Mullooly, M.; Sherman, M.E.; Gierach, G.L. Current and future methods for measuring breast density: A brief comparative review. *Breast Cancer Manag.* **2015**, *4*, 209–221.
4. Duric, N.; Littrup, P.; Schmidt, S.; Li, C.; Roy, O.; Bay-Knight, L.; Janer, R.; Kunz, D.; Chen, X.; Goll, J.; et al. Breast imaging with the softvue imaging system: First results. *Med. Imaging Ultrason. Imaging Tomogr. Ther.* **2013**, *8675*, 185–187.
5. Duric, N.; Littrup, P.; Li, C.; Roy, O.; Schmidt, S.; Cheng, X.; Seamans, J.; Wallen, A.; Bay-Knight, L. Breast imaging with softvue: Initial clinical evaluation. *SPIE Med. Imaging* **2014**, *2014*, 252–260.
6. Duric, N.; Littrup, P.; Poulo, L.; Babkin, A.; Pevzner, R.; Holsapple, E.; Rama, O.; Glide, C. Detection of breast cancer with ultrasound tomography: First results with the computed ultrasound risk evaluation (CURE) prototype. *Med. Phys.* **2007**, *34*, 773–785.
7. Li, C.; Duric, N.; Littrup, P.; Huang, L. In vivo breast sound-speed imaging with ultrasound tomography. *Ultrasound Med. Biol.* **2009**, *35*, 1615–1628.
8. Ranger, B.; Littrup, P.; Duric, N.; Chandiwala-Mody, P.; Li, C.; Schmidt, S.; Lupinacci, J. Breast ultrasound tomography versus magnetic resonance imaging for clinical display of anatomy and tumor rendering: Preliminary results. *Am. J. Roentgenol.* **2012**, *198*, 233–239.
9. Sak, M.; Duric, N.; Littrup, P.; Ali, H.; Vallieres, P.; Sherman, M.; Gierach, G. Using speed of sound imaging to characterize breast density. *Ultrasound Med. Biol.* **2017**, *43*, 91–103.
10. Glide-Hurst, C.; Duric-Weippl, N. Novel approach to evaluating breast density utilizing ultrasound tomography. *Med. Phys.* **2007**, *24*, 744–753.
11. Glide-Hurst, C.; Duric-Weippl, N.; Littrup, P. Volumetric breast density evaluation from ultrasound tomography images. *Med. Phys.* **2008**, *35*, 3988–3997.
12. Duric, N.; Norman, B.; Littrup, P.; Sak, M.; Myc, C.; Li, C.; West, E.; Minkin, S.; Martin, L.; Yaffe, M.; et al. Breast density measurements with ultrasound tomography: A comparison with film and digital mammography. *Med. Phys.* **2013**, *40*, 013501.
13. Khodr, Z.; Sak, M.; Pfeiffer, R.; Duric, N.; Littrup, P.; Bey-Knight, L.; Ali, H.; Vallieres, P.; Sherman, M.; Gierach, G. Determinants of the reliability of ultrasound tomography sound speed estimates as a surrogate for volumetric breast density. *Med. Phys.* **2015**, *42*, 5671–5678.
14. O'Flynn, E.; Fromageau, J.; Ledger, J.; Messa, M.; D'Aquino, A.; Schoemaker, A. Breast density measurements with ultrasound tomography: A comparison with non-contrast MRI. *Breast Cancer Res.* **2015**, *17* (Suppl. S1), O3.
15. Ruiter, N.V.; Hopp, T.; Zapf, M.; Kretzek, E.; Gemmeke, A. Analysis of patient movement during 3D USCT data acquisition. *SPIE Med. Imaging* **2016**, *2016*, 979009.

16. Sak, M.; Duric, N.; Littrup, P.; Li, C.; Bey-Knight, L.; Sherman, M.; Boyd, N.; Gierach, G. Breast density measurements using ultrasound tomography for patients undergoing tamoxifen treatment. *SPIE Med. Imaging* **2013**, *2013*, 86751.
17. Ou, Y.; Weinstein, S.P.; Conant, E.F.; Englander, S.; Da, X.; Gaonkar, B.; Hsieh, M.K.; Rosen, M.; DeMichele, A.; Davatzikos, C.; et al. Deformable registration for quantifying longitudinal tumor changes during neoadjuvant chemotherapy. *Magn. Reson. Med.* **2015**, *73*, 2343–2356.
18. Hopp, T.; Zapf, M.; Kretzek, E.; Henrich, J.; Tukalo, A.; Gemmeke, H.; Kaiser, C.; Knaudt, J.; Ruiter, N.V. 3D Ultrasound Computer Tomography: Update from a clinical study. *SPIE Med. Imaging* **2016**, *2016*, 97900A.
19. Sak, M.; Duric, N.; Littrup, P.; Sherman, M.; Gierach, G. Ultrasound tomography imaging with waveform sound speed: Parenchymal changes in women undergoing tamoxifen therapy. *SPIE Med. Imaging* **2017**, *2017*, 101390W.
20. Hopp, T.; Dapp, R.; Zapf, M.; Kretzek, E.; Gemmeke, H.; Ruiter, N.V. Registration of 3D ultrasound computer tomography and MRI for evaluation of tissue correspondences. *SPIE Med. Imaging* **2015**, *2015*, 94190Q.
21. Hopp, T.; Duric, N.; Ruiter, N.V. Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization. *Med. Image Anal.* **2013**, *17*, 209–218.
22. Hopp, T.; Duric, N.; Ruiter, N.V. Image fusion of ultrasound computer tomography volumes with X-ray mammograms using a biomechanical model based 2D/3D registration. *Comput. Med. Imaging Gr.* **2015**, *40*, 170–181.
23. Majdouline, Y.; Ohayon, J.; Keshavarz-Motamed, Z.; Cardinal, M.H.C.; Garcia, D.; Allard, L.; Lerouge, S.; Arsenault, F.; Soulez, G. Endovascular shear strain elastography for the detection and characterization of the severity of atherosclerotic plaques: In vitro validation and in vivo evaluation. *Ultrasound Med. Biol.* **2014**, *40*, 890–903.
24. Feng, X.; Guo, X.; Huang, Q. Systematic evaluation on speckle suppression methods in examination of ultrasound breast images. *Appl. Sci.* **2016**, *7*, 37.
25. Balic, I.; Goyal, P.; Roy, O.; Duric, N. Breast boundary detection with active contours. *SPIE Med. Imaging* **2014**, *2014*, 90400D.
26. Delgado-Gonzalo, D.; Chenouard, N.; Unser, M. Spline-based deforming ellipsoids for interactive 3D bioimage segmentation. *IEEE Trans. Image Process.* **2013**, *22*, 3926–3940.
27. Hopp, T.; Zapf, M.; Ruiter, N.V. Segmentation of 3D ultrasound computer tomography reflection images using edge detection and surface fitting. *SPIE Med. Imaging* **2014**, *2014*, 90401R.
28. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recog. Lett.* **2010**, *31*, 651–666.
29. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* **1975**, *11*, 23–27.
30. Sak, M.; Duric, N.; Littrup, P.; Westerberg, K. A comparison of automated versus manual segmentation of breast UST transmission images to measure breast volume and sound speed. *SPIE Med. Imaging* **2017**, *2017*, 101391H.
31. Hopp, T.; You, W.; Zapf, M.; Tan, W.Y.; Gemmeke, H.; Ruiter, N.V. Automated breast segmentation in Ultrasound Computer Tomography SAFT images. *SPIE Med. Imaging* **2017**, *2017*, 101390G.
32. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vision* **1988**, *1*, 321–331.
33. Xu, C.; Prince, J.L. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* **1998**, *7*, 359–369.
34. Cabezas, M.; Oliver, A.; Llado, X.; Freixenet, J.; Cuadra, M.B. A review of atlas-based segmentation for magnetic resonance brain images. *Comput. Methods Progr. Biomed.* **2011**, *104*, e158–e177.
35. Heimann, T.; Meinzer, H.P. Statistical shape models for 3D medical image segmentation: A review. *Med. Image Anal.* **2009**, *13*, 543–563.
36. Peng, B.; Zhang, L.; Zhang, D. A survey of graph theoretical approaches to image segmentation. *Pattern Recog.* **2013**, *46*, 1020–1038.
37. Zhang, R.; Zhu, S.; Zhou, Q. A novel gradient vector flow Snake model based on convex function for infrared image segmentation. *Sensors* **2016**, *16*, 1756.
38. Liu, J.; Tang, Z.; Cui, Y.; Wu, G. Local competition-based superpixel segmentation algorithm in remote sensing. *Sensors* **2017**, *17*, 1364.
39. Grau, V.; Mewes, A.U.J.; Alcaniz, M.; Kikinis, R.; Warfield, S.K. Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Med. Imaging* **2004**, *23*, 447–458.
40. Wang, L.; Shi, F.; Li, G.; Gao, Y.; Lin, W.; Gilmore, J.H.; Shen, D. Segmentation of neonatal brain MR images using patch-driven level sets. *NeuroImage* **2014**, *84*, 141–158.
41. Boykov, Y.; Funka-Lea, G. Graph cuts and efficient N-D image segmentation. *Int. J. Comput. Vision* **2006**, *70*, 109–131.

42. Rother, C.; Kolmogorov, V.; Blake, A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Gr.* **2004**, *23*, 309–314.
43. Blake, A.; Rother, C.; Brown, M.; Perez, P.; Torr, P. Interactive image segmentation using an adaptive GMMRF model. *ECCV 2004, 2004*, 428–441.
44. Lee, G.; Lee, S.; Kim, G.; Park, J.; Park, Y. A modified GrabCut using a clustering technique to reduce image noise. *Symmetry* **2016**, *8*, 64.
45. Chen, D.; Li, G.; Sun, Y.; Kong, J.; Jiang, G.; Tang, H.; Ju, Z.; Yu, H.; Liu, H. An interactive image segmentation method in hand gesture recognition. *Sensors* **2017**, *17*, 253.
46. Zhou, W.; Xie, Y. Interactive contour delineation and refinement in treatment planning of image-guided radiation therapy. *J. Appl. Clin. Med. Phys.* **2014**, *15*, 4499.
47. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671.
48. Ciccone, L.; Guay, M.; Sumner, R. Flow Curves: An Intuitive Interface for Coherent Scene Deformation. *Comput. Gr. Forum* **2016**, *35*, 247–256.
49. Boykov, Y.; Veksler, O.; Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 1222–1239.
50. Carreira, J.; Sminchisescu, C. CPMC: Automatic object segmentation using constrained parametric min-cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 1312–1328.
51. Talbot, J.F.; Xu, X. Implementing Grabcut. *Brigh. Young Univ.* **2006**, *3*, 1–4.
52. Lawler, E.L. *Combinatorial Optimization: Networks and Matroids*; Courier Corporation: Mineola, NY, USA, 2001.
53. Boykov, Y.; Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 1124–1137.
54. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29.
55. Klein, A.; Andersson, J.; Ardekani, B.A.; Ashburner, J.; Avants, B.; Chiang, M.C.; Christensen, G.E.; Collins, D.L.; Gee, J.; Hellier, P.; et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* **2009**, *46*, 786–802.
56. Shrout, P.E.; Fleiss, J.L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bulletin* **1979**, *86*, 420–428.
57. McGraw, K.O.; Wong, S.P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1996**, *1*, 30–46.
58. Lin, H.S.; Chen, Y.J.; Lu, H.L.; Lu, T.W.; Chen, C.C. Test–retest reliability of mandibular morphology measurements on cone-beam computed tomography-synthesized cephalograms with random head positioning errors. *Biomed. Eng. Online* **2017**, *16*, 62.
59. Ramirez, J.; Temoche, P.; Carmona, R. A volume segmentation approach based on GrabCut. *CLEI Electron. J.* **2013**, *16*, 4–14.
60. Chen, L.; Wu, S.; Zhang, Z.; Yu, S.; Xie, Y.; Zhang, H. *Real-Time Patient Table Removal in CT Images*; Springer HIS: Cham, Switzerland, 2016; pp. 1–8.
61. Liang, X.; Zhang, Z.; Niu, T.; Yu, S.; Wu, S.; Li, Z.; Zhang, H.; Xie, Y. Iterative image-domain ring artifact removal in cone-beam CT. *Phys. Med. Biol.* **2017**, *62*, 5276–5292.
62. Sandhu, G.; Li, C.; Roy, O.; Schmidt, S.; Duric, N. Frequency domain ultrasound waveform tomography: Breast imaging using a ring transducer. *Phys. Med. Biol.* **2015**, *60*, 5381.
63. Wang, K.; Matthews, T.; Anis, F.; Li, C.; Duric, N.; Anastasio, M. Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2015**, *62*, 475–493.
64. Li, Y.; Jiao, L.; Shang R.; Stolkin, R. Dynamic-context cooperative quantum-behaved particle swarm optimization based on multilevel thresholding applied to medical image segmentation. *Inf. Sci.* **2015**, *294*, 408–422.
65. Torbati, N.; Ayatollahi, A.; Kermani, A. An efficient neural network based method for medical image segmentation. *Compute. Biol. Med.* **2014**, *44*, 76–87.
66. Zhang, W.; Li, R.; Deng, H.; Wang, L.; Lin, W.; Ji, S.; Shen, D. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage* **2015**, *108*, 214–224.
67. Zhu, Q.; Shao, L.; Li, X.; Wang, L. Targeting accurate object extraction from an image: A comprehensive study of natural image matting. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 185–207.

68. Matthews, T.P.; Wang, K.; Li, C.; Duric, N.; Anastasio, M.A. Regularized dual averaging image reconstruction for full-wave ultrasound computed tomography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2017**, *64*, 811–825.
69. Birk, M.; Kretzek, E.; Figuli, P.; Weber, M.; Bacjer, J.; Ruiter, N.V. High-speed medical imaging in 3D ultrasound computer tomography. *IEEE Trans. Parallel Distrib. Syst.* **2016**, *27*, 455–467.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).