A Combined Regularization Method Using Prior Structural Information for Sound-Speed Image Reconstruction of Ultrasound Computed Tomography

ZISHENG YAO\textsuperscript{1}, SHANGCHUN FAN\textsuperscript{1,2}, YOSHIKAZU NAKAJIMA\textsuperscript{3}, (Member, IEEE), AND XIAOLEI QU\textsuperscript{1,2}

\textsuperscript{1}School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China
\textsuperscript{2}Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing 100191, China
\textsuperscript{3}Biomedical Information Department, Tokyo Medical and Dental University, Tokyo 101-0062, Japan

Corresponding author: Xiaolei Qu (quxiaolei@buaa.edu.cn)

This work was supported in part by the Beihang University under Grant KG12114001 and Grant ZG216S1970, in part by the National Natural Science Foundation of China under Project 61901024, and in part by the Beijing Municipal Natural Science Foundation under Project L192057.

ABSTRACT
Ultrasound computed tomography (USCT) is a promising technique for breast imaging. It provides three modalities: echo image, sound speed image (SSI), and attenuation image. The echo image reveals structural details of the breast with high resolution but is not quantitative. The SSI is quantitative, but its resolution is generally lower than that of the echo image. The SSI reconstruction is an ill-posed problem, thus the Tikhonov or total variation (TV) regularization methods are generally used. Tikhonov regularization is stable, but it tends to smooth the SSI and results in low resolution. TV regularization can preserve the sharp edges and provide higher resolution, but it may result in staircase artifacts. To combine the advantages of Tikhonov and TV regularizations, this article proposes a combined regularization method using prior structural information from the echo image. The proposed method mainly includes three steps. First, the echo image is reconstructed using the synthetic aperture (SA) technique and then segmented into breast and water regions. The segmentation result is used as prior structural information. Second, the USCT sound propagation forward model was built. Finally, with a penalty term according to the prior structural information, the combined regularization method is used to reconstruct the SSI. Both simulation and \textit{ex vivo} experiments were conducted for evaluation. In the simulation, the proposed method has a low root-mean-square-error (13.78 m/s), a high correlation coefficient (0.812), a high structural similarity (0.755), and a low standard deviation of water sound speed (2.33 m/s). In the \textit{ex vivo} experiment, the method has a low standard deviation of water sound speed (1.44 and 3.13 m/s for small and large objects, respectively), a low contrast to noise ratio (3.09 and 5.08), and a high Dice coefficient (0.92 and 0.97). Thus, the proposed combined regularization method using prior structural information can improve reconstruction quality.

INDEX TERMS
Ultrasound computed tomography, sound speed image, combined regularization, Tikhonov regularization, total variation regularization.

I. INTRODUCTION
With the continuous increase in breast cancer incidence worldwide [1], breast cancer imaging is becoming increasingly important for both screening and diagnosis.

Currently, there are three general imaging modalities for breast cancer detection and diagnosis, namely X-ray mammography [2], B-mode ultrasound [3], [4], and magnetic resonance imaging (MRI) [5]. X-ray mammography has been widely used for breast cancer screening. However, mammography exposes women to ionizing radiation, and only provides two-dimensional (2D) projection images of the breast. In the
2D projection images, the overlap between suspicious cancerous tissue and normal glandular tissue makes breast cancer screening difficult, especially for dense breast. B-mode ultrasound imaging avoids exposure to ionizing radiation. Moreover, it provides 2D cross-sectional images, which can prevent the overlap between suspicious cancerous tissue and normal glandular tissue. However, B-mode ultrasound imaging does not provide quantitative images and has the disadvantage of high dependence on the operator. MRI is generally used for breast cancer diagnosis. It automatically obtains three-dimensional (3D) quantitative images with high resolution and contrast. However, it is cumbersome, expensive and time-consuming.

In order to overcome the disadvantages of current imaging techniques, ultrasound computed tomography (USCT) has been proposed as a promising breast imaging tool for both screening and diagnosis purposes [6]–[9]. USCT can automatically and non-invasively obtain 3D breast images in three different modalities, namely echo image, sound speed image (SSI), and attenuation image. The USCT imaging system can be built based on conventional linear transducer array [10]–[16] or the ring transducer array [17], [18]. In this study, the system is based on the ring transducer array, which generally includes a table with an opening on it, a water tank under the opening of the table, and a ring transducer array immersed in the water tank. The imaging procedure is as follows [19]: the patient lies prone on the table with one breast naturally protruding downward through the opening. Under the opening, the breast is immersed in the water tank and surrounded by the ultrasound ring transducer array with hundreds of ultrasound elements. During scanning, the element successively emits ultrasound pulses, while all elements receive both transmitted and echo waves. After scanning, all three modalities of 2D coronal breast images can be reconstructed from the collected raw data. Then, 3D breast images can be obtained by mechanically moving the ring transducer array up and down.

By using the wave signals collected from the scanning, the echo image, SSI, and attenuation image, can be reconstructed. The echo image can be obtained by the synthetic aperture (SA) technique using the collected echo wave signal. It generally has the highest resolution among these three modalities. In addition, it can also provide more structural details of the breast tissue with high resolution, but it does not provide quantitative information of the breast [19]. Both the attenuation image and SSI can be reconstructed using the collected transmitted wave signals, and they can provide quantitative information (attenuation and sound speed) of the breast tissue, which is useful for breast cancer diagnosis. Previous studies [20] have reported that both attenuation and sound speed vary between the tumor and normal tissue. Particularly, the sound speed in malignant tissue is about two percent higher than that in normal tissue [9]. Accordingly, it is believed that the SSI of USCT can be relatively useful in breast imaging. Therefore, this study mainly focuses on the reconstruction of the SSI.

At present, there are two kinds of SSI reconstruction methods: waveform-based methods [21] and ray-based methods [22], [23]. Waveform-based methods are generally more sophisticated. They build a forward model of a USCT based on a full-wave equation in a frequency or space domain, then iteratively solve the forward model and reconstruct the SSI. Waveform-based methods are commercially available, and can provide high quality image, but their computational costs are high [24]–[26], because they need to solve the full-wave equation. On the other hand, ray-based methods are faster and more stable [27], and the reconstructed SSIs can be used as the initial condition of iterations for waveform-based methods. Ray-based methods generally assume that the sound wave travels along a straight or a bent ray path from an emitter to a receiver. According to whether a ray follows a straight or bent path, ray-based methods can be classified into two categories: straight-ray method and bent-ray method. The reconstruction quality of the bent-ray method is higher than that of straight-ray method, as the bent-ray method considers the effects of refraction. This study focuses on improving the bent-ray method.

Since the SSI may change drastically due to a small error in received data, the reconstruction of an SSI is an ill-posed problem. The Tikhonov and TV regularization methods [28]–[31] have been generally used to solve the ill-posed problem when using bent-ray methods. Tikhonov regularization is robust to noise, but it tends to smooth the reconstructed SSIs, resulting in low resolution. It is suitable for reconstructing smooth regions, such as the water region in a USCT breast image. TV regularization can preserve the edges in the reconstructed SSI, giving it a higher resolution. It is good at reconstructing breast regions with complex glandular and adipose distribution, but it may result in staircase artifacts [32]. To combine the advantage of the Tikhonov and TV regularization methods, a combined regularization method is proposed in this study to exploit the advantages from both regularization methods.

Two previous studies [33], [34] have proposed improvements to the SSI quality by prior structural information. Carson et al. proposed a post-processing method to improve the SSI by using prior information for the ring transducer array [33]. Sanabria et al. proposed to assign target region of interest in a prior structure with a single sound speed value for the linear transducer array, which can significantly reduce the number of degrees of freedom in the reconstruction inverse problem [34].

In this article, we combine Tikhonov and TV regularization using the prior structural information from the echo image segmentation to improve the bent-ray SSI reconstruction. The proposed method is applicable to both in vivo and ex vivo experiments. But we only perform the preliminary validation of our proposed method by simulation and ex vivo experiments, due to the limitation of equipment.
II. METHODS

The schematic of the proposed method is shown in Figure 1, which mainly consists of seven steps. First, ultrasound raw data are acquired by the ring transducer array. Second, echo image is reconstructed by the SA method. Third, the echo image is segmented into breast and water regions, which is used as the prior structural information for reconstruction. Fourth, the propagation time between each pair of emitter and receiver is obtained by the first arrival picking method. Fifth, the forward model is built for sound wave propagation. Sixth, current SSI is reconstructed using the prior structural information (obtained by the third step) by the proposed combined regularization method. Lastly, if the iteration converges (the difference between current and previous SSIs is small enough), the reconstruction process is over. Otherwise, it goes back to the fifth step. After SSI reconstruction, the quality of the SSI is also quantitatively evaluated.

A. DATA ACQUISITION

The data acquisition procedure is shown in Figure 2 [35]. First, a spherical wave is emitted by one ultrasound element in the ring transducer array. The element can be viewed as a one-point source. The wave propagates into the tissue inside the ring, thereby undergoing a complex combination of reflection, refraction, scattering, and attenuation. Then, ultrasound signals are received by all the elements in the ring transducer array. This process is repeated for each element. After scanning, a 3D dataset is obtained, which can be expressed as

\[ D(t,p,m,n) \]

where \( D \) is the sampling point, and \( m \) and \( n \) represent the indices of the receiving elements and emitting elements, respectively.

B. ECHO IMAGING BY SA METHOD

The schematic diagram of echo imaging using the SA method is illustrated in Figure 3. The SA method can implement both dynamic-transmit and dynamic-receive focusing and provide high quality echo images [18], [36].

The dynamic-receive focusing is conducted in two steps, as follows. First, the ultrasound propagation time is calculated from the emitter to the focal point and then to the receiver. It can be expressed as:

\[ t(p,m,n) = \frac{|p - e_n| + |r_m - p|}{c} \]  

where vector \( p \) denotes the receive focusing position; vector \( e_n \) and \( r_m \) represent the position of the emitter and the receiver, respectively; \( c \) is the average sound speed, assumed to be a constant value. Second, a low-resolution image is reconstructed for each emission using echo signals with the following equation:

\[ I_n(p) = \sum_{m=1}^{M} a(m,n) D(t(p,m,n),m,n) \]

where \( M \) is the total number of receivers used, \( a(m,n) \) is a weighting function derived from the Hann function. It should be noted that since the elements on the opposite of emitters have strong transmission signals, the signals from only half of the elements are used, as indicated in Figure 3.

The dynamic-transmit focusing is performed by summing up all the low-resolution images with the following equation:

\[ I(p) = \sum_{i=1}^{N} I_i(p) \]

where \( I(p) \) is the reconstructed high-resolution echo image; \( N \) is the number of low-resolution images, which is equal to the number of emitters.
C. SEGMENTATION BY ACTIVE CONTOUR MODEL

After acquiring the high-resolution echo image, we use the active contour model [37], [38] to segment and separate the breast and water regions. The active contour model depicts a deformable parametric curve and a corresponding energy function. The energy function is defined as follows:

\[ E_{\text{total}} = \int_s k_1 \left| \frac{\partial v(s)}{\partial s} \right|^2 + k_2 \left| \frac{\partial^2 v(s)}{\partial s^2} \right|^2 + E_{\text{ext}}(v(s)) \, ds \]  

(4)

where \( s \) is an independent variable describing the boundary in the form of a Fourier transform, and \( v(s) \) represents the point on the initial boundary. In this equation, there are three integral terms. The first term is the elastic energy, described by the first order derivative of \( v(s) \). The second term is the bending energy, described by the second order derivative of \( v(s) \). The third term is the external energy. The sum of the elastic energy and the bending energy is called internal energy, which controls the elastic deformation of a contour. External energy denotes the degree to which the deformation curve conforms to the local features of the image.

In our study, segmentation of the echo image by the active contour model is performed as follows. First, in the echo image, several points are manually selected to form a close-loop geometry as the initial boundary. The only principle for the selection is to ensure that the initial boundary encloses the real boundary between the breast and water. Second, the energy function is obtained according to equation (4). Finally, the energy function \( E_{\text{total}} \) is minimized using an iterative method, until the energy function converges to an acceptable value.

After segmentation, the imaging area can be separated into two regions: the inner and the external regions. The inner region is the breast region and the external is the water region. The locations of both regions are the prior structural information.

D. FORWARD MODELING FOR SOUND SPEED RECONSTRUCTION

The transmitted signals are used to reconstruct the SSIs. First, the sound propagation time between each pair of emitter and receiver is picked up by the Akaike information criterion (AIC) method [19] using the transmitted signals. After that, \( \Delta t_1 \) is calculated. It represents the propagation time difference between ultrasound propagation through water with the imaging object inside and through pure water. Then, the forward model can be described as follows:

\[ \sum_{j=1}^{M} l_{ij} \Delta x_j = \Delta t_1 \]  

(5)

where \( \Delta x_j \) is the slowness perturbation for the \( j \)th grid cell in the imaging region, and \( l_{ij} \) represents the ray length of the \( i \)th ray within the \( j \)th grid cell. Combining all pairs of receiver and emitter, equation (5) can be written in matrix form:

\[ LS = T \]  

(6)

where the slowness matrix \( S \) can be converted to the sound speed distribution.

E. SOLVING INVERSE PROBLEM BY REGULARIZATION METHODS

Generally, regularization methods should be used to solve equation (6), as it is an ill-posed problem.

When Tikhonov regularization is used, the SSI is reconstructed using the following equation [39]:

\[ f = \arg \min_S \left( \|LS - T\|_2^2 + \alpha \|S\|_2^2 \right) \]  

(7)

where \( \alpha \) is the regularization parameter which can be manually adjusted. Previous studies have verified that this method improves the stability of the solution for ill-posed problems. Moreover, the position and size of the inclusion are quite accurate [40]. However, this method tends to provide a smooth solution with relatively low sharpness of edges, which results in low resolution SSIs.

When TV regularization is used, the SSI is reconstructed using the following equation [9], [31]:

\[ f = \arg \min_S \left( \|LS - T\|_2^2 + \beta TV(S) \right) \]  

(8)

\[ TV(S) = \iint \sqrt{|\nabla S|^2 + \epsilon^2} \, dx \, dy \]  

(9)

where \( \beta \) is the regularization parameter which can be manually adjusted, and \( \epsilon \) is a small constant value, which makes the TV term differentiable at the origin point. Previous studies have demonstrated that TV regularization can detect sharp interfaces [40], but the form of the penalty term is more complicated than that in Tikhonov regularization.

Based on the characteristics of the widely used Tikhonov regularization and TV regularization methods, we propose a combined regularization method using the prior structural information from echo image. The combined regularization method can exploit the advantages of two currently available methods for sound speed reconstruction. It can be expressed as:

\[ f = \arg \min_S \left( \|LS - T\|_2^2 + \gamma_1 \|S_{\text{ext}}\|_2^2 + \gamma_2 TV(S) \right) \]  

(10)

where \( \gamma_1 \) and \( \gamma_2 \) are two separated regularization parameters, which can be manually selected; \( S_{\text{ext}} \) corresponds to the slowness perturbation in the external water region. As shown in equation (10), two penalty terms are used in the combined regularization method. The TV term, which helps preserve the sharpness of the edge, is identical to that in the TV regularization method. The other term correlates to the echo image and can be viewed as an additional term, which aims to control the sound speed reconstruction error in the external region.

To solve equations (7), (8), and (10), the Limited-Memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization algorithm [41] is employed. It uses an estimation of the inverse Hessian matrix to steer the search through a variable
space, and only stores a few vectors of the inverse Hessian matrix, which can significantly reduce the memory cost.

**F. SIMULATION, EXPERIMENT AND EVALUATION METHODS**

In this section, the simulation experiment settings, the ultrasound CT prototype, and the *ex vivo* experiment are introduced. Then, four quantitative metrics are employed for quality evaluation of the reconstructed SSIs.

1) **SIMULATION**

The simulation is carried out by using the k-wave simulation toolbox [42], which is an efficient toolbox designed for time domain acoustic and ultrasound simulations in complex and tissue-realistic media.

In order to reduce computational cost for simulation, a 2D simulation is conducted. In the simulation, 128 equispaced elements are placed on a ring with a diameter of 100 mm. The central frequency of the emitted sound wave is 1.6 MHz, and the sample frequency for the received signal is 10 MHz. To mimic the real breast sound speed distribution, 2D MRI images (from an open-source dataset [43]) are segmented into skin, glandular, and adipose regions, then sound speeds of 1600, 1570, and 1470 m/s, respectively, are assigned to them [44]. The speed of sound in water is set as 1540 m/s. The grid size in the simulation is 0.195 mm x 0.195 mm. The time step is 39.88 ns, which is computed automatically according to the grid size and the sound speed distribution by the k-wave toolbox.

2) **EX VIVO EXPERIMENTS**

An ultrasound CT prototype [19] has been built for *ex vivo* experiments. It is composed of a 100-mm ring transducer (Figure 4(a)), on which 1024 elements are mounted with a 0.325 mm pitch. The sampling frequency is set as 10 MHz, and the central frequency of these elements is 1.6 MHz. For each emission, four neighboring elements emit at the same time as one-point source to increase the signal-to-noise ratio of the sound wave. In our experiment, the total scanning included 128 emitters and receivers at the same interval, as our ultrasound platform, which is shown in Figure 4(b), is a 128-channel system (Vantage 128, Verasonics, Inc., Kirkland, WA, USA) [45]. Two pieces of chicken breast meat with different volumes were selected as *ex vivo* imaging objects, one of which is shown, during scanning, in Figure 4(c). Waveforms without and with meat are shown in Figure 4(d) and (e), whose signal-to-noise ratio (SNR) is 43 dB and 59 dB, respectively.

3) **METRICS FOR EVALUATION OF RECONSTRUCTED SSIS**

In this study, six quantitative metrics are used for evaluation, including the root-mean-square error (RMSE), correlation coefficient (CC), structural similarity (SS), contrast noise ratio (CNR) [46]–[48], Dice coefficient, and the standard deviation of the water sound speed.

The RMSE can be defined as follows:

$$RMSE(x, y) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (x_i - y_i)^2}$$  \hspace{1cm} (11)

where $x$ is the simulated sound speed distribution map written in a column vector, $y$ is the reconstructed SSI written in a column vector, and $M$ is the number of pixels in the SSI. The RMSE is a non-negative real number. The smaller it is, the higher the image quality is. Particularly, when the RMSE equals to 0, it means that the reconstructed SSI and the simulation pattern are identical.

The CC can be defined as follows:

$$CC(x, y) = \frac{\sum_{i=1}^{M} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{M} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{M} (y_i - \bar{y})^2}}$$  \hspace{1cm} (12)

where $\bar{x}$ and $\bar{y}$ are the average sound speed values for the sound speed map for simulation and the SSI, respectively. The CC ranges from $-1$ to $1$. The value of the CC indicates the linear correlation between the reconstructed SSI and the simulated sound speed distribution map. Accordingly, the quality of reconstructed SSI is high when the CC is close to 1. Particularly, when the CC reaches 1, it means that the reconstructed SSI and the simulation pattern are identical.

The SS can be defined as follows:

$$SS(x, y) = \frac{4\bar{x}\bar{y}\sigma_{xy}}{\bar{x}^2 + \bar{y}^2}\left(\frac{\sigma_{x}^2 + \sigma_{y}^2}{\sigma_{x}^2 + \sigma_{y}^2}\right)$$  \hspace{1cm} (13)
where $\sigma_x$ and $\sigma_y$ are the standard deviation of the sound speed map for simulation and SSI, respectively; $\sigma_{xy}$ is the covariance between the simulated sound speed distribution map and SSI. The derivation of the SS considers the characteristics of the human visual system. It ranges from 0 to 1. Like the CC, the quality of the reconstructed SSI is high when the SS is close to 1. In particular, when the SS reaches 1, it means that the reconstructed SSI and the simulation pattern are identical.

To evaluate the contrast of the reconstructed SSI, the CNR is introduced, which is defined as follows:

$$\text{CNR} = \sqrt{\frac{2}{\sigma_1^2 + \sigma_2^2}} \left| \mu_1 - \mu_2 \right|$$

where $\mu_1$ and $\mu_2$ are the average sound speeds of the background region and imaging object region, respectively; $\sigma_1$ and $\sigma_2$ are the variance of the sound speed in the background region and imaging object region, respectively. In this study, we are interested in the contrast between internal breast region and external water region. A higher CNR value generally indicates a better image contrast.

The geometric fidelity [34] of the reconstructed SSI is evaluated by introducing the Dice coefficient, which is defined as follows:

$$\text{Dice} = \frac{2 \left| c_{\text{inc}}^* \cap c_{\text{inc}} \right|}{\left| c_{\text{inc}}^* \right| + \left| c_{\text{inc}} \right|}$$

where $c_{\text{inc}}^*$ and $c_{\text{inc}}$ are the regions of the reconstructed and ground-truth inclusions, respectively.

The noises and artifacts in the reconstructed SSI are evaluated by calculating the standard deviation of the sound speed in the water region. A lower standard deviation in the water region generally indicates lower noises and fewer artifacts.

For the simulation experiment, the reconstructed SSI and the simulation pattern can be compared by calculating the RMSE, CC, SS and the standard deviation of the sound speed in the water region. However, for ex vivo experiments, RMSE, CC and SS cannot be calculated as there are no reference images. Therefore, the CNR, Dice coefficient and the standard deviation of the sound speed in the water region are calculated.

**III. RESULTS**

**A. SIMULATION EXPERIMENT RESULTS**

The sound speed distribution maps obtained by segmenting the MRI breast images are shown in Figure 5. In order to ensure that the k-wave simulations work well, the sound speed distribution maps are interpolated to $512 \times 512$ pixels, so that the grid size of the simulation (0.195 mm) is significantly smaller than the wavelength (0.981 mm). Four regions, namely water, adipose, glandular, and skin regions, whose sound speeds were set to be 1540, 1470, 1570 and 1600 m/s, respectively, are observed in each sound speed distribution map.

The echo images reconstructed by the SA method using echo signal are shown in Figure 6. These echo images have high resolutions and sharp boundaries between water and skin, which are useful for segmenting water and breast region with high accuracy. The active contour model is used for segmentation. The initial boundary between water and breast is shown in the left part of Figure 7, and the detected boundary is indicated in the right part of Figure 7. The magnified details in Figure 7 reveal that the boundary successfully shrinks and
Z. Yao et al.: Combined Regularization Method Using Prior Structural Information for SSI Reconstruction

FIGURE 8. Segmentation of each simulated sound speed distribution map. In (a)-(e), the white region corresponds to the breast region, and the black region corresponds to the water region.

closely covers the outside of the breast skin. After establishing the boundary between water and breast, the water and breast regions can be easily separated and segmented. The segmentation results of the echo image in Figure 6 are presented in Figure 8, in which the white region corresponds to the breast, and the black background region corresponds to water. The Dice coefficient for these five segmentations are 0.996, 0.994, 0.995, 0.994, 0.995, respectively.

Figure 9 shows the SSIs reconstructed by the combined regularization method with different regularization parameters. The reconstructed SSIs with variation of $\gamma_2$ (parameter for TV term in Equation 10) are shown in Figure 9(a)-(e), and the reconstructed SSIs with variation of $\gamma_1$ (parameter for Tikhonov term in Equation 10) are shown in Figure 9(f)-(j). We selected $\gamma_1 = 10$ and $\gamma_2 = 0.1$ for all subsequent reconstructions in both the simulation and experiment studies.

The reconstructed SSIs for five sound speed distribution maps in simulation (Figure 5) are shown in Figure 10. The images in Figure 10 (a)-(e) are reconstructed by the TV regularization method, those in Figure 10 (f)-(j) are reconstructed by the Tikhonov regularization method and those in Figure 10 (k)-(o) are reconstructed by the proposed combined regularization. The size of each reconstructed SSI is $128 \times 128$ pixels. Visually, the quality of the SSIs reconstructed by Tikhonov regularization is lower than that of those reconstructed by TV regularization. Compared with TV regularization, the combined regularization method achieves better quality in the exterior water region, which can be observed in the inset showing magnified details on the right of Figure 10.

For quantitative evaluation, Figure 11 shows the RMSE, CC, SS and the standard deviation of sound speed in the water region. Regarding the RMSE, Figure 11(a) shows that the mean value of RMSE for the Tikhonov, TV and proposed methods are 15.58, 14.04 and 13.78 m/s, respectively. Regarding the CC, Figure 11(b) reveals that the mean value of CC for Tikhonov, TV, and the proposed methods are 0.776, 0.800 and 0.812, respectively. Regarding the SS, Figure 11(c) shows that the mean value of the SS for the Tikhonov, TV and the proposed methods are 0.662, 0.747 and 0.755, respectively. Regarding the standard deviation, Figure 11(d) shows that the mean value of the standard deviation for the Tikhonov, TV and the proposed methods are 3.64, 3.51, 2.33 m/s, respectively.

B. EX VIVO EXPERIMENTAL RESULTS

The reconstructed SSIs for a small piece of chicken breast meat and a large piece of chicken breast meat are shown in Figure 12. Specifically, Figure 12(a) and (d) show SSIs reconstructed by the TV regularization method; Figure 12(b) and (e) show SSIs reconstructed by the Tikhonov regularization method; Figure 12(c) and (f) show SSIs reconstructed by the proposed combined regularization method. The ground truth sound speeds of the chicken meat were measured with five repeats by using two wideband ultrasound transducers, a pulse receiver, and a Vernier caliper. The measured sound speed was $1572.0 \pm 3.0$ m/s. Compared with the measured sound speed, the average sound speed in the reconstructed SSI by Tikhonov, TV, and the proposed methods were 1558.3,
1567.8, 1567.7 m/s for small chicken meat and 1543.9, 1566.0, 1566.6 m/s for large chicken meat, respectively.

The standard deviation of sound speed in water regions, the CNR of the reconstructed SSIs, and the Dice coefficient between chicken meat regions in the echo image and SSIs, are shown in Figure 13 (a), (b) and (c), respectively. The proposed method has a Dice coefficient close to that of the TV method, which is higher than that of the Tikhonov method.

**IV. DISCUSSION**

In this article, a combined regularization method is proposed for reconstruction of SSI using prior structural information from an echo image. Both simulation and ex vivo
In this study, we propose a combined regularization method to improve the SSI quality using structural information. Currently, there are also two alternatives that utilize structural information. First, the speed of sound in the water region can be assumed to be constant, so that only the breast region needs to be reconstructed by TV regularization. Second, the water region can be reconstructed only by Tikhonov regularization, then the breast region only by TV regularization. We have tested these two alternatives. The reconstructed SSIs using the proposed method, the first alternative, and the second, are displayed in Figure 14 (a), (b), and (c), respectively. They achieved similar reconstruction quality for the breast region. In this study, we do not set the speed of sound in water as a constant value, because the temperature and speed of sound in water may vary slightly in different regions in real experiments.

Compared with the currently available Tikhonov and TV regularization methods, the proposed combined regularization method improves the quality of the reconstructed SSI. The proposed method is quantitatively evaluated by both simulation and ex vivo experiments. However, our study still has several limitations and considerable potentially interesting future work still needs to be done. First, in the proposed reconstruction method, we only use the boundary between breast and water regions as prior structural information. However, the echo images contain much more information, and it will be an interesting future work to improve the reconstruction quality using more prior structural information with multiple layers (water, glandular, and adipose layers), which can be obtained by fuzzy binarization segmentation method [49]. Second, the reconstruction quality for different imaging objects is highly variable (Figure 11 and 13), when using the proposed or the currently available methods. One possible reason is that the low resolution of SSI causes serious image blur, which further results in large error of speed of sound reconstruction in small or narrow regions. Therefore, the adipose region of pattern 1 and the glandular region of pattern 5 have lower reconstruction quality and worse reconstruction metrics than other regions (Figure 11). Finally, the proposed combined regularization method may also be available for USCT based on the linear transducer array. The prior structural information, including skin layer and muscle layer, can be obtained from echo images, then slight variation of the speed of sound in the homogenous tissue region can be assumed, and the combined regularization may improve the smoothness of the homogenous tissue regions and preserve the high frequency variation in other regions.

V. CONCLUSION

In this study, a combined regularization method was proposed for SSI reconstruction in USCT. In the proposed method, an extra term was added to the currently available TV regularization method by using the prior structural information.
from echo images. Both \textit{ex vivo} and simulation experiments showed that the proposed method helps to improve the quality of the reconstructed SSI.

\section*{Acknowledgment}
The authors would like to acknowledge Prof. Azuma at The University of Tokyo for providing experimental equipment.

\section*{References}

[1] R. L. Siegel and K. D. A. M. Jemal, “Cancer statistics.”, \textit{CA A, Cancer J. Clinicians}, vol. 69, no. 1, pp. 7–34, 2019.

[2] P. T. Weatherall, G. F. Evans, G. J. Metzger, M. H. Saborrian, and A. M. Leitch, “MRI vs. Histologic measurement of breast cancer following chemotherapy: Comparison with X-ray mammography and palpation.”, \textit{J. Magn. Reson. Imag.}, vol. 13, no. 6, pp. 868–875, 2001.

[3] G. J. Whitman, J. D. Blume, J. B. Cormack, and E. B. Mendelson, “Operator dependence of physician-performed whole-breast US: Lesion detection and characterization.”, \textit{Radiology}, vol. 18, no. 3, pp. 252–253, 2006.

[4] C. Bamber and C. R. Hill, “Acoustic properties of normal and cancerous human liver—I. Dependence on pathological condition.”, \textit{Ultrasound Med. Biol.}, vol. 7, no. 2, pp. 121–133, Jan. 1981.

[5] S. C. Rankin, “MRI of the breast.”, \textit{Brit. J. Radiol.}, vol. 73, no. 73, pp. 806–818, 2000.

[6] N. Duric, P. Littrup, L. Poulo, A. Babkin, R. Pevzner, E. Holzapfel, O. Rama, and C. Glide, “Detection of breast cancer with ultrasound tomography: First results with the computed ultrasound risk evaluation (CURE) prototype.”, \textit{Med. Phys.}, vol. 34, no. 2, pp. 773–785, 2007.

[7] L. Huang and N. Q. Nguyen, “Ultrasound bent-ray tomography using both transmission and reflection data.”, \textit{SPIE}, Int. Soc. Opt. Eng., vol. 9040, no. 3, 2014, Art. no. 90400R.

[8] L. Huang, J. Shin, T. Chen, Y. Lin, K. Gao, M. Intrator, and K. Hanson, “Breast ultrasound tomography with two parallel transducer arrays.”, \textit{SPIE}, Med. Imag., \textit{Ultrason. Imag. Tomogr.}, vol. 9738, Mar. 2016, Art. no. 97380C.

[9] C. Li, N. Duric, P. Littrup, and L. Huang, “\textit{In vivo} breast sound-speed imaging with ultrasound tomography.”, \textit{Ultrasound Med. Biol.}, vol. 35, no. 10, pp. 1615–1628, Oct. 2009.

[10] S.-W. Huang and P.-C. Li, “Ultrasound computed tomography reconstruction of the attenuation coefficient using a linear array.”, \textit{IEEE Trans. Ultrason., Ferroelec., Freq. Control}, vol. 52, no. 11, pp. 2011–2022, Nov. 2005.

[11] C. H. Chang, S. W. Huang, H. C. Yang, Y. H. Chou, and P.-C. Li, “PG2G-9 reconstruction of ultrasonic sound velocity and attenuation coefficient using linear arrays: Clinical assessment.”, \textit{Ultrasound Med. Biol.}, vol. 33, no. 11, pp. 1681–1687, 2007.

[12] S. J. Sanabria and M. B. O. R. Gorske, “Speed-of-sound imaging based on reflector delineation.”, \textit{IEEE Trans. Biomed. Eng.}, vol. 66, no. 7, pp. 1949–1962, Jul. 2019.

[13] S. J. Sanabria and O. Gorske, “Hand-held sound-speed imaging based on ultrasound reflector delineation.”, \textit{Proc. MICCAI}, 2016, pp. 568–576.

[14] S. J. Sanabria, O. Gorske, K. Martini, S. Forte, T. Frauenfelder, R. A. Kubik-Huch, and M. B. Rominger, “Breast-dedensity assessment with hand-held ultrasound: A novel biomarker to assess breast cancer risk and to tailor screening.”, \textit{Eur. Radiol.}, vol. 28, no. 8, pp. 3165–3175, Aug. 2018.

[15] S. J. Sanabria, K. Martini, G. Freyöttäker, L. Ruby, O. Gorske, T. Frauenfelder, and M. B. Rominger, “Speed of sound ultrasound: A pilot study on a novel technique to identify sarcopenia in seniors.”, \textit{Eur. Radiol.}, vol. 29, no. 1, pp. 3–12, Jan. 2019.

[16] M. Jaeger, G. Held, S. Peeters, S. Preisser, M. Grüning, and M. Frenz, “Computed ultrasound tomography in echo mode for imaging speed of sound using pulse-echo sonography: Proof of principle.”, \textit{Ultrasound Med. Biol.}, vol. 41, no. 1, pp. 235–250, Jan. 2015.

[17] X. Qu, T. Azuma, H. Nakamura, H. Imoto, S. Tamano, S. Takagi, S.-I. Umemura, I. Sakuma, and Y. Matsumoto, “Bent ray ultrasound tomography reconstruction using virtual receivers for reducing time cost.”, \textit{Proc. SPIE}, vol. 9419, Mar. 2015, Art. no. 94190F.

[18] X. Qu, T. Azuma, T. Yogi, S. Azuma, H. Takeuchi, S. Tamano, and S. Takagi, “Synthetic aperture ultrasound imaging with a ring transducer array: Preliminary \textit{ex vivo} results.”, \textit{J. Med. Ultrason.}, vol. 43, no. 4, pp. 1–11, 2016.

[19] X. Qu, T. Azuma, H. Imoto, R. Raufy, H. Lin, H. Nakamura, S. Tamano, S. Takagi, S.-I. Umemura, I. Sakuma, and Y. Matsumoto, “Novel automatic first-arrival picking method for ultrasound sound-speed tomography.”, \textit{Jpn. J. Appl. Phys.}, vol. 54, no. 7, Art. no. 07FH10.

[20] T. Chen, J. Shin, and L. Huang, “Ultrasound transmission attenuation tomography using energy-scaled amplitude ratios.”, \textit{Int. Soc. Opt. Photon.}, vol. 9790, Apr. 2016, Art. no. 979016.

[21] W. Kun, T. Matthews, F. Anis, C. Li, N. Duric, and M. A. Anastasio, “Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography.”, \textit{IEEE Trans. Ultrason., Ferroelectr., Freq. Control}, vol. 62, no. 3, pp. 93–147, Mar. 2015.

[22] A. Andersen, “A ray tracing approach to restoration and resolution enhancement in experimental ultrasound tomography.”, \textit{Ultrasound. Imag.}, vol. 12, no. 4, pp. 268–291, Oct. 1990.

[23] X. Huang, J. Sun, and Z. Sun, “Local algorithm for computing complex travel time based on the complex eikonal equation.”, \textit{Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.}, vol. 93, no. 4, Apr. 2016, Art. no. 045307.
Z. Yao et al.: Combined Regularization Method Using Prior Structural Information for SSI Reconstruction

ZISHENG YAO was born in Fuzhou, Fujian, China, in 1995. He received the B.E. degree in measurement and control technology and instrumentation from Beihang University, Beijing, China, in 2018, where he is currently pursuing the master’s degree with the School of Instrumentation and Optoelectronic Engineering.

Since 2018, he has been a Trainee Engineer of instrument control and measurement with the China Instrument and Control Society. His research interests include ultrasound imaging, non-destructive testing technology, and optimization methods for inverse problems.

XIAOLEI QU was born in Shaanxi, China, in 1984. He received the B.E. degree in software engineering from Xi’an Jiaotong University, Xi’an, China, in 2007, the M.S. degree in pattern recognition from the Huazhong University of Science and Technology, Wuhan, China, in 2009, and the Ph.D. degree in bioengineering from The University of Tokyo, Tokyo, Japan, in 2012.

From 2012 to 2014, he was a Senior Engineer with the TechnoWin Research Center (Samsung Group), Fukuoka, Japan. From 2014 to 2017, he was a Project Researcher with The University of Tokyo. Since 2017, he has been an Associate Professor with the School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing, China. He is the author of more than 20 articles and holds 4 patents. His research interests include medical ultrasound imaging, image processing, and recognition.

SHANGCHUN FAN was born in 1962. He received the B.E. degree in measuring and testing technologies and instruments, the M.S. degree in precision instruments and mechanical engineering, and the Ph.D. degree from Beihang University, Beijing, China, in 1980, 1986, and 1990, respectively.

Since 1990, he has been a Professor with the School of Instrumentation and Optoelectronic Engineering, Beihang University. From 1996 to 1997, he was a Visiting Scholar with Pusan National University, Busan, South Korea. He is the author of more than 100 articles, more than 50 granted patents, and 11 textbooks. His research interests include sensor technology, signal detection, and processing technology, especially resonant sensor. He received one Second Class Prize of the State Technological Innovation Awards, in 2013, one Second Class Prize of the State Scientific and Technology Progress Awards, in 2008, and one Second Class Prize of the State Defense Science and Technology, in 2006. He is the Director of the China Instrument and Control Society.

YOSHIKAZU NAKAJIMA (Member, IEEE) received the Ph.D. degree in computer science from Osaka University, Osaka, Japan, in 1997. He was with the Wellness Business Development Center in the Head Office and the Information Technology Research and Development Center, Mitsubishi Electric Corporation, from 1997 to 2000. He joined the Division of Functional Diagnostic Imaging, Graduate School of Medicine, Osaka University, where he was an Assistant Professor, from 2001 to 2004. He was an Associate Professor with the Department of Bioengineering, The University of Tokyo, from 2005 to 2017.

He is currently a Professor and the Director of the Biomedical Information Department, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Tokyo. He has been working in computer science, medical image processing, and surgical robotics fields. His research interest includes medical imaging and informatics.