

Image reconstruction for robot assisted ultrasound tomography

Fereshteh Aalamifar^a, Haichong K. Zhang^b, Arman Rahmim^{a,c}, Emad M. Bector^{a,b,c}

^aDept. of Electrical & Computer Eng., ^bDept. of Computer Science, ^cDept. of Radiology,
Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

An investigation of several image reconstruction methods for robot-assisted ultrasound (US) tomography setup is presented. In the robot-assisted setup, an expert moves the US probe to the location of interest, and a robotic arm automatically aligns another US probe with it. The two aligned probes can then transmit and receive US signals which are subsequently used for tomographic reconstruction. This study focuses on reconstruction of the speed of sound. In various simulation evaluations as well as in an experiment with a millimeter-range inaccuracy, we demonstrate that the limited data provided by two probes can be used to reconstruct pixel-wise images differentiating between media with different speeds of sound. Combining the results of this investigation with the developed robot-assisted US tomography setup, we envision feasibility of this setup for tomographic imaging in applications beyond breast imaging, with potentially significant efficacy in cancer diagnosis.

1. INTRODUCTION

Ultrasound (US) tomographic images are reconstructed by calculating acoustic properties such as the speed of sound in each pixel or voxel of the image thus opening a new dimension in diagnostic imaging. US tomography is enabled by placing US transducers at two sides of the anatomy of interest. The image is reconstructed by modeling the phenomena happening to the signal traveling through the medium. As an example, the US signal travels with varying speed through different parts of the medium and hence reaches the receiver at various times. The times of arrival can be measured to reconstruct speed of sound image of the medium.

Several configurations can be used to enable US tomographic data collection. One setup that has been developed for breast cancer diagnosis consists of a ring of US transducers [1-6]. The patient lies on a bed and inserts her breast inside a cavity surrounded by the ring of US elements. The elements take turn in transmitting an US signal while all others are listening and recording the waveform for a certain amount of time. These waveforms are then processed to reconstruct reflection, speed of sound, and attenuation images of the breast that has been shown to be superior in terms of lesion detection over conventional B-mode images.

Another recently developed US tomography system contains of two long US probes that are placed parallel to each other [7]. In this study, there are 384 transducers linearly arranged in each probe and the two probes can be translated both axially and elevationally. This system is also used for breast cancer diagnosis. However, this system can be used to image breasts with different sizes and also can scan the axillary region which provides prominent information about early stage cancer [8].

At the same time, the above systems can only be used for breast cancer US tomography because it is hard to encapsulate other organs. An alternative setup that can enable US tomography for other organs consists of two conventional US probes that are aligned. Accurate alignment of these two probes is possible using robotic assistance (robot-assisted US tomography system [9-10]). Unlike the setup involving two long parallel US probes discussed above, in the robot assisted US tomography system, we utilize US systems that are currently approved for use and are available in the clinic. Furthermore, the robotic system enables the alignment in any configuration using 6 degrees of freedom, thus having the potential to be used for more general organs and applications.

In such a setup, the expert can move one of the US probes to any desired location and look for regions of interest using the conventional B-mode US. Subsequent to finding the region of interest, the other robot automatically aligns the second probe to the first one; then the tomographic data can be collected and the reconstructed image be shown to the sonographer. Figure 1 shows the dual robotic arms prototype developed to enable automatic alignment [10].

In this paper, we investigate and test feasibility of several image reconstruction algorithms that can be used with this robotic setup. It should be noted that, in some applications, it might not be possible to acquire a full-angle US tomographic data, and in fact, in most cases, only few views of the object are available (so-called tomosynthesis) due to anatomical

limitations. In addition, the alignment accuracy is always limited by calibrations and robot positioning inaccuracies. Hence, our goal is to investigate how well the image can be reconstructed given such limitations. In this work, we are focusing on reconstructing speed of sound using a ray-based method (versus the so called waveform tomography method [11-12]). This method calculates speed of sound in each pixel of the image using the time of flight information, assuming that the US signal travels in straight paths. We investigate several methods of solving the reconstruction equation including analytical, iterative, and statistical methods. All the presented methods could, to some extent, differentiate among different speeds of sound with a single view data set. We then expand on the simulations and show how incompleteness of data affects the reconstructions. Next, we describe experimental results of a phantom containing lesions with different speeds of sound and then discuss possible future directions.

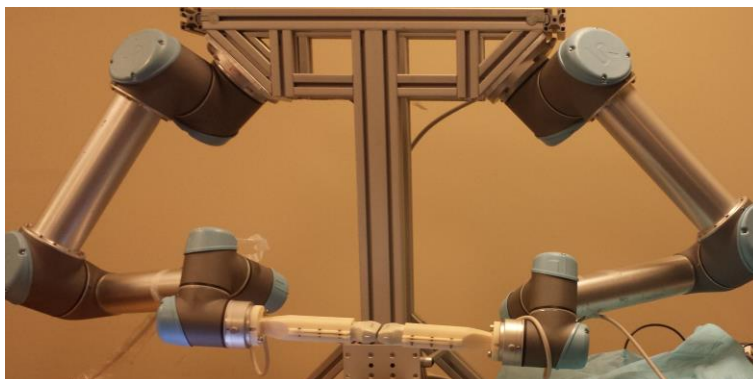


Figure 1. Dual robotic arm setup for automatic alignment of two linear US probes.

2. METHODS

Reconstruction Methods

In this section, we describe how to reconstruct the speed of sound in the insonified area between two aligned linear arrays of US transmitters and receivers. Figure 2 depicts two aligned US probes.

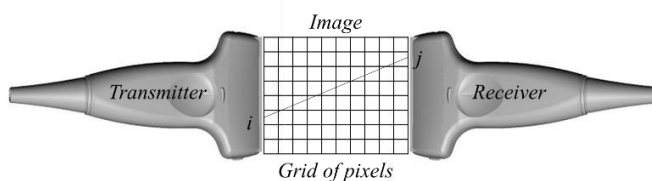


Figure 2. Two aligned linear probes and the grid of pixels representing the image.

Assume the elements on the left probe to act as transmitter while the elements on the right act as receiver. We define a grid of pixels between the two probes that represents the image to be reconstructed. A ray is the straight line that connects one transmitter element to one receiver element. Each ray passes through several grid cells with different crossing lengths. System matrix S , describes how much of each ray travels through each pixel and is constructed using the geometrical relation between transmitters and receivers' locations. This matrix has $N_t \times N_r$ rows and N_g columns, where N_t is the number of transmitting elements, N_r is the number of receiving elements, and N_g is the total number of grid cells. Each row of the system matrix corresponds to one pair of transmitter-receiver, and contains the path lengths corresponding to each grid cell. We constructed system matrix based on the method presented in [13] and utilizing the lateral size of the US probe, number of elements in one probe, and the distance between the two probes. We assumed the location of each transmitter/receiver is in the center, while the element width or height can also be incorporated.

The speed of sound can be reconstructed by time of flight measurements for all rays. Time of flight, T , is represented by a vector of length $N_t \times N_r$ while the image, X , is represented by a vector of length N_g . The speed of sound in each cell of the grid can be estimated. Instead of directly calculating the speed, we calculate its inverse, which is called slowness. The time required for an ultrasound signal to travel along a ray is equal to the sum of the time to pass through all cells; and the

time required to travel through one cell is $t = l * x$, where l is the path length along the cell, and x is the cell's slowness. Hence, $T_j = \sum_{i=1}^n S_{i,j} * x_i$, where T_j is the total time of flight for the j -th ray, $S_{i,j}$ is the path length of j -th ray along i -th cell and x_i is the i -th cell slowness. This leads to solving the following equation for X :

$$SX = T, \tag{1}$$

where S is the system matrix, X is the image vector, and T is the time of flight measurements. Reconstruction of attenuation can also be similarly formulated while the measurement and the image represent different variables.

Rank of $[S|T]$ determines if the above equation is solvable. In the single view case, it is obviously rank deficient (due to incomplete tomographic data) and there are multiple solutions for the equation.

Multiple methods can be used to find a solution to the above equation. We evaluated an analytical method (pseudo-inverse), an iterative method (conjugate gradient) [14], and a statistical method (expectation maximization [15]). The pseudo-inverse method (Pseudo) is most easily implemented and always produces X with minimum norm. Conjugate gradient (CG) is faster but might not converge. The statistical method (EM) is also iterative and thus might not converge to a solution, usually runs fast enough, and is known to perform well in incomplete or noisy data. The following iterative algorithm is used for EM method [15]:

$$x_j^{m+1} = \frac{x_j^m}{\sum_{i=1}^I s_{ij}} \sum_{i=1}^I s_{ij} \left(\frac{t_i}{\sum_{j=1}^J x_j^m s_{ij}} \right), \tag{2}$$

where x_j^m is the j -th pixel value at m -th iteration. s_{ij} is the system matrix array corresponding to i -th ray and j -th pixel. t_i is the time of flight for i -th ray.

We also consider two versions for each algorithm: difference and non-difference methods (Diff). In the difference method, the problem is solving for an X_{obj} that satisfies:

$$S(X_b - X_{obj}) = T_b - T_{obj}, \tag{3}$$

In the difference methods, we assume that a background phantom is scanned without moving the probes. We expect this additional information to reduce inaccuracy in the reconstruction. We evaluated these methods in a simulations study as discussed next.

Simulations

We consider a phantom with three different speeds of sound, namely 1010, 1200, and 1490 m/s (figure 3). Each probe has 128 transducers arranged in 60 mm probe length. Firstly, two linear probes are simulated on top and bottom of the phantom. All the elements on the top probe are transmitters while all on the bottom are receivers (single view case). Then, we simulate a case where the receiver probe moves to three locations: bottom, right, and left (multi view case). To better describe the effect of data incompleteness we also simulate a virtually infinite-length probe and compare its result with the single view case. The (almost) infinite length was created by defining US elements every 1-degree pitch and for -89 to 89 degrees range while the angle corresponding to each element is defined as the angle between two lines: 1. The line connecting center transmitter-receiver elements, 2. The line connecting the element to the farthest element in the other probe.

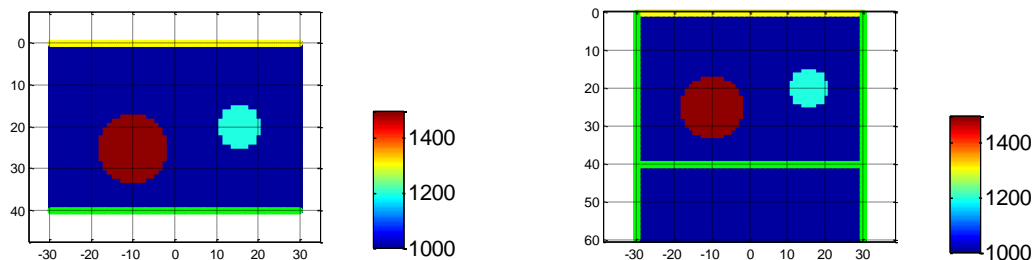


Figure 3. Simulation phantom and transmitter/receiver locations. Green lines show the position of receiver probes while yellow line shows the position of transmitter probe. (a) single view: transmitters on top, receivers in the bottom. (b) three views: transmitters on top, receivers on sides and bottom.

In forward modeling, we used the ground-truth numerical phantom together with the known distances between transmitter-receiver pairs and generated the time of flight. We did not incorporate refraction/scattering effects and assumed that ultrasound wave travels in straight lines. This kind of simulation provides a better understanding of how mathematically solvable the image reconstruction problem is without considering subtleties related to US wave propagation.

Experiment setup

Next, we show the results of reconstruction for an experimental phantom. The goal is to investigate the feasibility of distinguishing between different speeds of sound using US transmission data collected from two conventional US probes that are aligned with millimeter range inaccuracy. It should be noted that the data is collected from single view. To show the proof of concept, we made a preliminary phantom with three different materials. The main body of the phantom is made of silicone ecoflex 10 with 1010 m/s speed of sound. It was made using a 80x40 mm cubic mold (shown in Figure 4(a)) made of acrylic sheets. Using two 3D printed cylinders with diameters 10 mm and 16 mm in the mold, two cylindrical cavities were created in the main body of the phantom and filled with water and alcohol (1480 and 1200 m/s speeds of sound).

Another phantom was also made that only contained the ecoflex material as the background phantom. Two Ultrasonix 128 array, 60 mm US probes were used to scan the phantom from two sides as shown in Figure 4(b). We aligned the probes using the two robotic arms. One probe was connected to an Ultrasonix US touch machine. On the US machine the clinical software was used to set the US probe in transmit mode. The other probe was connected to an Ultrasonix data acquisition device (DAQ).

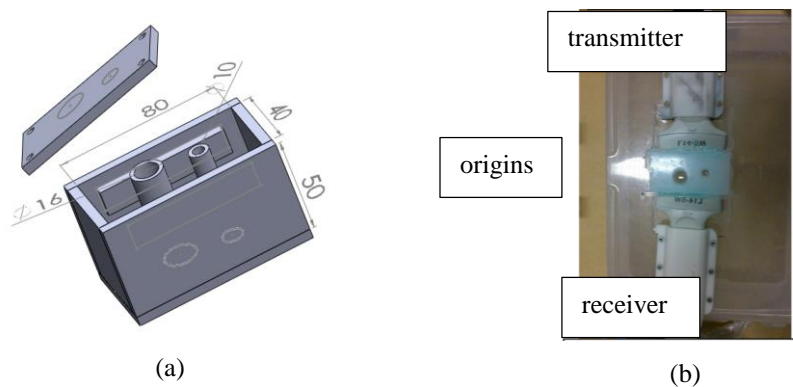


Figure 4. Experimental setup: (a) The phantom mold CAD model. The mold was made using laser cut acrylic pieces. (b) the phantom is placed between two ultrasound probes that are aligned with millimeter range inaccuracy.

Evaluation parameters

We compute bias, noise, and contrast to evaluate the images. The following equations are used to compute these parameters:

$$B = \text{mean}(P_R) - \text{mean}(P_G), \quad (4)$$

$$C = \frac{\text{mean}(P_L)}{\text{mean}(P_{BG})}, \quad (5)$$

$$PC = \frac{C_R}{C_G} \times 100, \quad (6)$$

$$N = \text{std}(P_{BG}), \quad (7)$$

where B , C , PC , N are bias, contrast, percent-contrast, and noise respectively; P_R , and P_G are reconstructed pixels and ground-truth values (for a determined region such as lesion or background) respectively; P_L , and P_{BG} are pixel values inside

lesion and background respectively; and C_R and C_G are contrast values calculated for reconstructed image and ground-truth respectively.

3. RESULTS

Simulation results

Given the time of flight information and the system matrix, we are interested in reconstructing the image, i.e., finding the vector X , that satisfies equation (1). Figure 5 shows the reconstructed images obtained using all the methods mentioned in the previous section without noise or inaccuracy in the system. The goal is to see the effect of incompleteness of tomographic data.

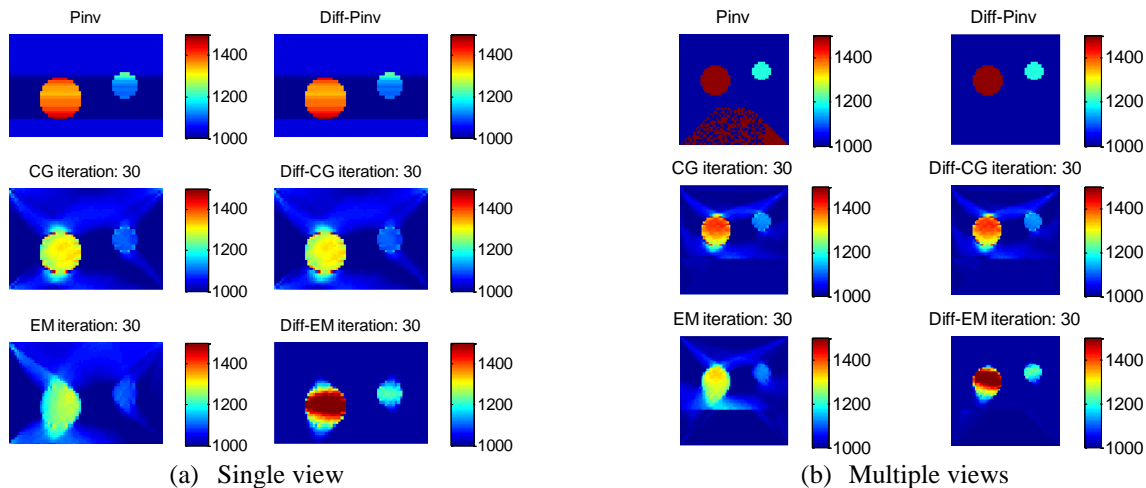


Figure 5. Comparison of different methods to solve reconstruction. Top row shows pseudo-inverse and difference-pseudo-inverse methods. Middle row shows conjugate gradient and difference-conjugate-gradient methods. Bottom row shows expectation maximization and difference-expectation-maximization methods.

Figure 5(b) shows reconstruction results when there is a transmitting probe on top and the receiver probe moves to three locations. It can be seen that multi-view approach results in improvement over single view. In this case, Pseudo methods could superiorly reconstruct the image because of relative completeness of data. The red pixels on the lower part of the Pseudo method is due to lack of any information in that region; in other words, no ray passes through the pixels in that area. This phenomenon does not appear in the iterative methods because of initial values were set for those pixels. Table 1 compares bias, noise, and contrast for different methods.

Table 1. Bias, noise, contrast, and percent-contrast for different methods of reconstruction.

			Pseudo	Diff-Pseudo	CG	Diff-CG	EM	Diff-EM
Single view	Bias (m/s)	Water	-126.96	-126.96	-179.05	-179.19	-230.75	86.18
		Alcohol	-76.73	-76.73	-95.33	-95.20	-98.19	18.79
		Background	19.29	19.29	15.62	16.00	26.85	0.025
	Percent Contrast	Water	68.23	68.23	58.54	58.51	45.13	117.95
		Alcohol	48.54	48.54	40.97	41.05	33.30	109.87
	Noise		24.71	24.71	25.91	25.85	38.63	0.05
Multi view	Bias (m/s)	Water	0	0	-95.30	-93.63	-178.97	73.25
		Alcohol	0	0	-70.7282	-70.11	-76.19	25.55
		Background	0	0	6.66	6.25	26.12	0.17
	Percent Contrast	Water	100	100	78.24	78.71	55.83	115.20
		Alcohol	100	100	58.88	59.44	44.99	113.34
	Noise		0	0	23.20	23.22	19.66	0.69

An observation inferred from Table 1 is that even though it provided the best results in single view condition, the Diff-EM method can lead to higher than %100 contrast. This extra contrast is also a phenomenon caused by the incompleteness of available tomographic data.

In addition, increasing the length of the probes, while keeping them in the one-view configuration, can increase the amount of available data and lead to better results. Figures 6 and 7 demonstrate these facts by comparing results using a pair of 6cm versus infinitely long probes. The results show bias, contrast, and noise for the single view case as a function of iterations, when reconstructing using the EM method.

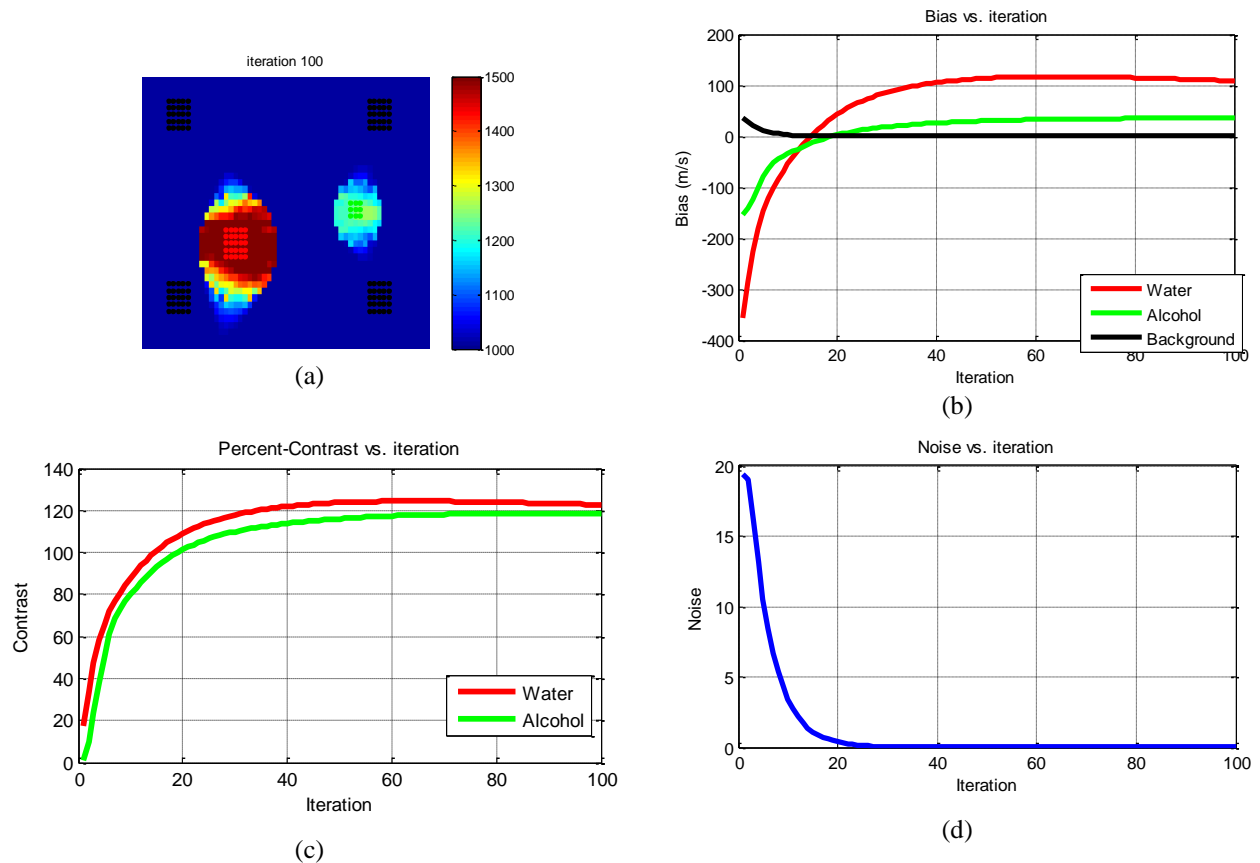


Figure 6. Bias, contrast, and noise as a function of iterations, using a pair of 6cm probes in single view configuration for Diff-EM method. (a) reconstructed image and the regions used to calculate the three parameters. Red dots show pixels used as inside water, green dots show inside alcohol region, and black dots show sample pixels selected as background. (b) bias as a function of iterations. Bias for water and alcohol hits zero around 18th iteration and then increases and levels off. (c) percent-contrast versus iteration. Likewise, at around 18th iteration the contrast hits 100% and then increases. (d) Background noise decreases as a function of iteration and stays at zero.

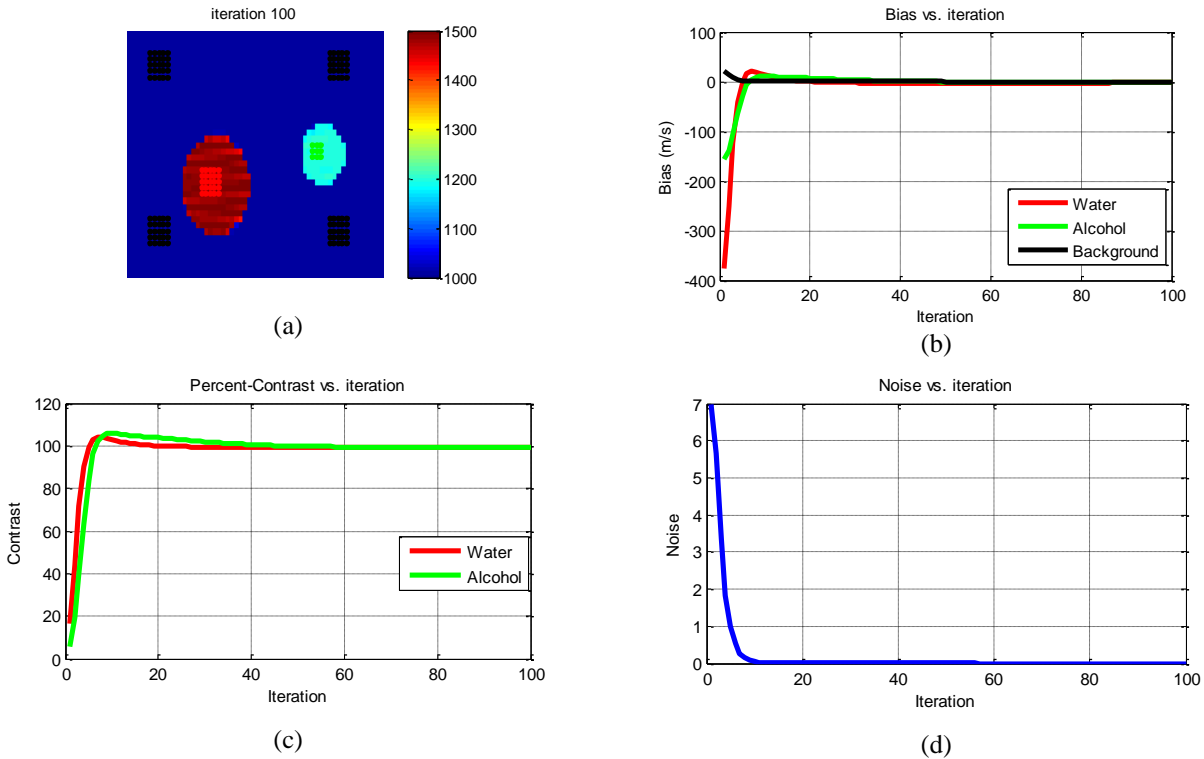


Figure 7. Bias, contrast, and noise as functions of iteration, using a pair of infinite length probes in single view configuration for Diff-EM method. (a) reconstructed image and the regions used to calculate the three parameters. Red dots show pixels used as inside water, green dots show inside alcohol region, and black dots show sample pixels selected as background. (b) bias as a function of iterations. Bias for water and alcohol hits zero around 18th iteration and stays at zero. (c) percent-contrast also reaches 100 around 18th and levels off. (d) Background noise decreases as a function of iteration and stays at zero.

Experimental results

First, we scanned the main phantom using the first probe; then the robotic arm was moved to the aligned position and tomographic data was collected. Without moving the probes, we removed the phantom, inserted the background phantom and repeated the data collection process. Since the robots were not calibrated very well, we used the time of flight in background phantom to estimate the distance between two probes which was estimated as 36 mm. After collecting all the waveforms, the time of flight was manually picked. The system matrix was also constructed using the probes' geometrical information and distance between the probes. The Bmode image as seen by the transmitter probe is shown in Figure 8(a).

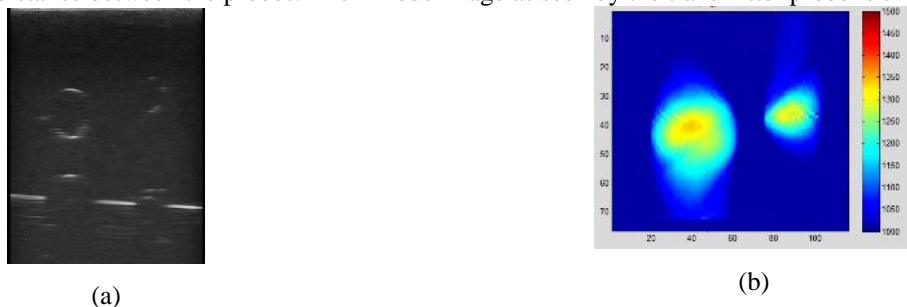


Figure 8. (a) Bmode image from transmitter probe. (b) Diff-EM reconstruction.

Since the Diff-EM method performed best for the single view simulation, we reconstructed the image using this method. The resulting image is shown in Figure 8(b). Figure 9 shows the bias, percent contrast, and noise versus number of iterations.

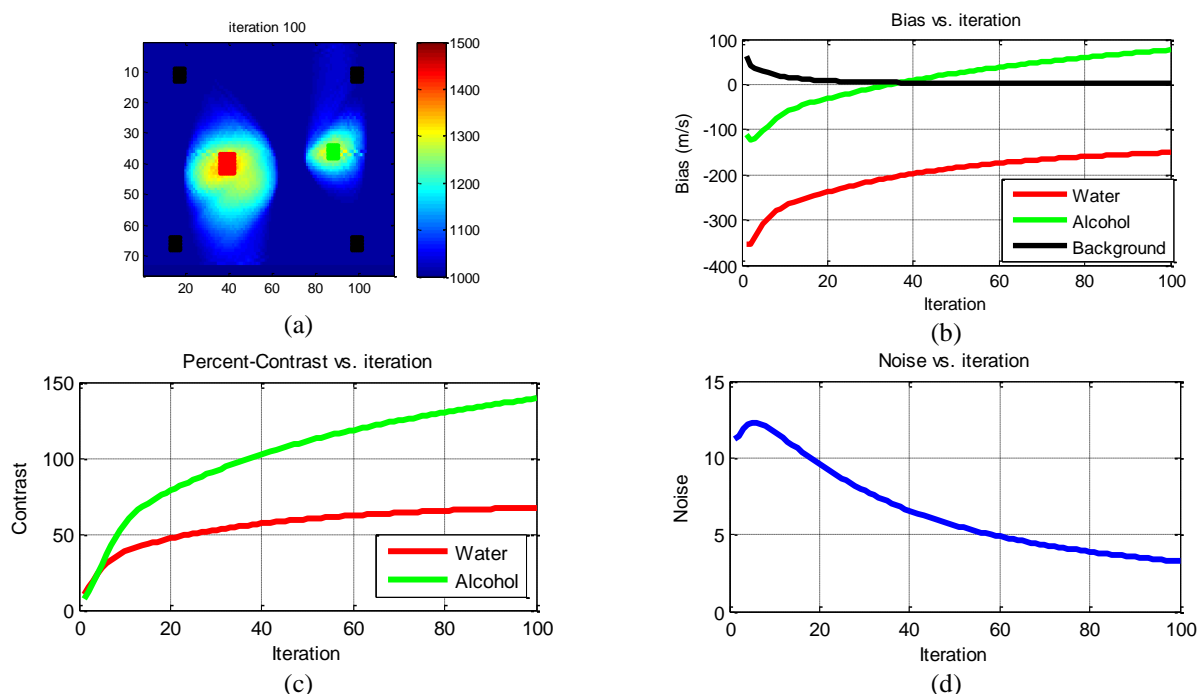


Figure 9. Experimental results of single view reconstruction using Diff-EM method. (a) The regions used to calculate the three parameters. Red dots show inside water region, green dots depict inside alcohol, and black dots show the background regions. (b) The bias in calculating water speed of sound hits zero at iteration 40th and then increases while the one for alcohol never reaches zero. (c) Similar to simulation, at 40th iteration, water contrast increases artificially; percent contrast in alcohol region does not reach 100% after 100 iterations. (d) Noise decreases versus iteration.

The experimental results, even though showing a similar trend to the simulation results, depict larger bias and noise and lower percent contrast. The factors that have contributed to this include error in probes alignment, and noise in time of flight data. In addition, time of flight picking is a challenging task. Another important contributing factor is the effect of refraction, i.e. the US wave does not travel a straight path. Several possible methods are proposed in the next section to address these issues.

4. DISCUSSION

As mentioned before, this paper reports preliminary image reconstruction results for robot-assisted US tomography. In this section, we present an overview of limitations and our vision for future work to address those limitations.

Ray-based approximation: In the simulations provided in this paper, we have not considered refraction/scattering effects. An US simulator can be used to incorporate US wave propagation characteristics. Along this line, it would be interesting to investigate how waveform tomography methods [11-12] can improve the results.

Incomplete data: To address incomplete tomographic data issue, one way is to use reflection data during reconstruction. Another solution is to use reflection data is to create image priors. One piece of information available when having two conventional US probes is the Bmode image. Sonographers are usually able to label different regions of this mode of image. Such information can be used to create an image prior by, for example, constraining the labeled areas with the known speed of sound.

Reconstruction: Another interesting parameter that can be reconstructed is attenuation coefficient. The same reconstruction algorithms can be used in a similar fashion; however, the original transmitter signal should be known. In

this case, the signal intensity is the segmented data and used in the equation. Accurate segmentation is an important step in US tomography process and automating this process would make real-time tomographic images available; an improved automated time of flight picker is provided in [17].

Alignment inaccuracy: Use of background phantom is helpful in overcoming the effect of alignment inaccuracy, and also time of flight picking bias on reconstruction. It might be feasible to insert such background phantom in applications where the patient can be removed without a need to move the probes. However, we speculate that improving on the alignment accuracy, using several other tracking loops such as using the US signals, has the potential to eliminate the need for a background phantom. The misalignment also affects the location of the lesion appearing in the image. This should be considered in applications where accurate localization of the suspicious regions is important.

5. CONCLUSIONS

US tomography can be used to quantitatively measure acoustic properties of the scanned region. In this paper, we provided proof of concept, and investigated multiple reconstruction techniques, for robot-assisted US tomography. We provided simulation and experimental results of several ray-based tomographic reconstruction approaches that can be used for the robot-assisted US tomography setup as a first step to illustrate potentials and limitations. We showed that, even though the transmission data acquired from one pose of two aligned probes is not enough to reconstruct a perfectly accurate map of speed of sound, it can still distinguish between different speeds of sound. We also showed that having data from another or several other views, if possible in the specific application, can provide more valuable information and lead to improved reconstructions. We also demonstrated in a phantom experiment that by using data from one pose of the two probes, and even encountering inaccuracy in the millimeter range, it is still possible to reconstruct an image that captures the regions with different speeds of sound. Even though straight ray-based methods are fast, they are unable to model refraction phenomenon in transmission ultrasound; in addition, the same setup can be used to collect reflection data. The recently proposed method of reflection-transmission waveform tomography may address these shortcomings. Investigating capabilities of this method for the robot-assisted US tomography setup is a next step of this work and may produce images of higher quality and quantitative accuracy at the expense of slower run times. Hence, one vs. the other approach may be desired depending on the application.

ACKNOWLEDGEMENTS

Financial support was provided by Johns Hopkins University internal funds and NSF grant IIS-1162095. We would like to thank Bill Bauer and Gina Grdina for the helpful discussions. We would like to also thank Dr. Hani Bakhshae and Dr. Rajat Mittal for their technical support.

REFERENCES

- [1] Duric, Nebojsa, Peter Littrup, Lou Poulou, Alex Babkin, Roman Pevzner, Earle Holsapple, Olsi Rama, and Carri Glide. "Detection of breast cancer with ultrasound tomography: First results with the Computed Ultrasound Risk Evaluation (CURE) prototype." *Medical physics* 34, no. 2 (2007): 773-785.
- [2] Delphinus Medical Technologies, "<http://www.delphinusmt.com/our-technology/softvue-system>." Accessed 1/11/2015.
- [3] Duric, Nebojsa, Peter Littrup, Alex Babkin, David Chambers, Stephen Azevedo, Arkady Kalinin, Roman Pevzner, Mikhail Tokarev, and Earle Holsapple. "Development of ultrasound tomography for breast imaging: Technical assessment." *Medical Physics* 32.5 (2005): 1375-1386.
- [4] Li, Cuiping, Nebojsa Duric, Peter Littrup, and Lianjie Huang. "In vivo Breast Sound-Speed Imaging with Ultrasound Tomography." *Ultrasound in medicine & biology* 35.10 (2009): 1615-1628.
- [5] Glide, Carri, Nebojsa Duric, and Peter Littrup. "Novel approach to evaluating breast density utilizing ultrasound tomography." *Medical physics* 34.2 (2007): 744-753.
- [6] Littrup, Peter J., Neb Duric, Stephen Azevedo, David Chambers, James V. Candy, Stephen Johnson, Gregory Auner, John Rather, and Earle T. Holsapple. "Computerized ultrasound risk evaluation (CURE) system: Development of combined transmission and reflection ultrasound with new reconstruction algorithms for breast imaging." *Acoustical Imaging*. Springer US, 2002. 175-182.

- [7] Huang, Lianjie, Junseob Shin, Ting Chen, Youzuo Lin, Miranda Intrator, Kenneth Hanson, Katherine Epstein, Daniel Sandoval, and Michael Williamson. "Breast ultrasound tomography with two parallel transducer arrays: preliminary clinical results." In *SPIE Medical Imaging*, pp. 941916-941916. International Society for Optics and Photonics, 2015.
- [8] Huang, Lianjie, Youzuo Lin, Zhigang Zhang, Yassin Labyed, Sirui Tan, Nghia Nguyen, Kenneth Hanson, Daniel Sandoval, and Michael Williamson. "Breast ultrasound waveform tomography: Using both transmission and reflection data, and numerical virtual point sources." In *SPIE Medical Imaging*, pp. 90400T-90400T. International Society for Optics and Photonics, 2014.
- [9] Aalamifar, Fereshteh, Rishabh Khurana, Alexis Cheng, Russell H. Taylor, Iulian Iordachita, and Emad M. Boctor. "Enabling technologies for robot assisted ultrasound tomography: system setup and calibration." In *SPIE Medical Imaging*, pp. 90401X-90401X. International Society for Optics and Photonics, 2014.
- [10] Aalamifar, Fereshteh, Dengrong Jiang, Haichong K. Zhang, Alexis Cheng, Xiaoyu Guo, Rishabh Khurana, Iulian Iordachita, and Emad M. Boctor. "Co-robotic ultrasound tomography: dual arm setup and error analysis." In *SPIE Medical Imaging*, pp. 94190N-94190N. International Society for Optics and Photonics, 2015.
- [11] Zhang, Zhigang, Lianjie Huang, and Youzuo Lin. "Efficient implementation of ultrasound waveform tomography using source encoding." In *SPIE Medical Imaging*, pp. 832003-832003. International Society for Optics and Photonics, 2012.
- [12] Lin, Youzuo, Lianjie Huang, and Zhigang Zhang. "Ultrasound waveform tomography with the total-variation regularization for detection of small breast tumors." In *SPIE Medical Imaging*, pp. 832002-832002. International Society for Optics and Photonics, 2012.
- [13] Siddon, Robert L. "Fast calculation of the exact radiological path for a three - dimensional CT array." *Medical physics* 12, no. 2 (1985): 252-255.
- [14] Shewchuk, Jonathan Richard. "An introduction to the conjugate gradient method without the agonizing pain." (1994).
- [15] Rahmim, Arman. "Statistical list-mode image reconstruction and motion compensation techniques in high-resolution positron emission tomography (PET)." PhD diss., University of British Columbia, 2005.
- [16] Aalamifar, Fereshteh, Alexis Cheng, Younsu Kim, Xiao Hu, Haichong K. Zhang, Xiaoyu Guo, and Emad M. Boctor. "Robot-assisted automatic ultrasound calibration." *International journal of computer assisted radiology and surgery* (2016): 1-9.
- [17] Li, Cuiping, Lianjie Huang, Nebojsa Duric, Haijiang Zhang, and Charlotte Rowe. "An improved automatic time-of-flight picker for medical ultrasound tomography." *Ultrasonics* 49, no. 1 (2009): 61-72.