Acoustic Shadow Detection: Study and Statistics of B-Mode and Radiofrequency Data

Ricky Hu[a,*], Rohit Singla[a], Farah Deeba[a], Robert N. Rohling[a,b]

[a] *Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada*
[b] *Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada*

Abstract

An acoustic shadow is an ultrasound artifact occurring at boundaries between significantly different materials, resulting in a continuous dark region on the image. Shadow detection is important as shadows can identify anatomical features or obscure regions of interest. A study was performed to scan human subjects (N=35) specifically to explore the statistical characteristics of various shadows from different anatomy and with different transducers. Differences in shadow statistics were observed and used for shadow detection methods with radiofrequency (RF) or brightness-mode (B-mode). Both methods utilized adaptive thresholding, needing only the transducer pulse width as an input parameter for easy utilization by different operators or equipment. Mean Dice coefficients ($\pm$ standard deviation) of 0.90$\pm$0.07 and 0.87$\pm$0.08 were obtained for the RF and B-mode methods, which is within range of manual annotators. The results indicate that the methods are able to detect shadows with high versatility in different imaging scenarios. The understanding of shadow statistics can be used for more specialized methods can be developed for specific applications in the future. The detection methods can also potentially aid interpretation of ultrasound images or serve as an important pre-processing step for machine learning methods.

Keywords: Acoustic Shadow, Ultrasound, Speckle, Radiofrequency, Segmentation

Introduction

Ultrasound devices have become increasingly affordable and portable, encouraging applications such as point-of-care ultrasound (Bouhemad et al., 2011), novice usage (Sippel et al., 2011), and analysis by machine learning (Ghose et al., 2013). However, ultrasound is susceptible to unique artifacts that increase the difficult of interpretation and processing of images. One artifact is an acoustic shadow, which occurs when an ultrasound wave crosses a boundary of two materials with high impedance differences (Kremkau and Taylor, 1986). The wave is almost completely reflected and depicted beyond the boundary is a continuous dark region and a loss of anatomical features. Shadows occur in air-tissue, tissue-bone, and tissue-lesion interfaces. Shadows can aid interpretation, such as identifying gall stones (Good et al., 1979) or spinal levels (Galiano et al., 2005). However, shadows, such as from poor transducer contact, can lead to misinterpretation of anatomy, particularly by novice users and automated processing algorithms. Thus, the identification of shadows is an important preprocessing step in many applications.

Several methods have been used in literature to detect shadows and illustrative examples are discussed. Geometric techniques model the path of an ultrasound signal for an expected image along the scanline using a random walk (Karamalis et al., 2012). Pixels are then flagged as a shadow if it is below a confidence threshold. However, geometric techniques require knowledge of ultrasound transducer properties to parameterize random walk weights, such as the focal length, radius of curvature, and thickness. The technique is more cumbersome to implement across different ultrasound machines. This also reduces applicability for machine learning applications as accurate transducer parameter labels are required for each image.

Pixel intensity methods ignore the transducer properties and analyze only the graphical properties of animage (Hellier et al., 2010). Shadows have been detected on brain images by analyzing the entropy along a scanline to flag pixels of sudden low entropy as a potential shadow. The technique achieved a comparable Dice similarity coefficient as geometric methods but require specific thresholding, window sizing, filtering, and image mask parameterization for different anatomy and transducers. The drawback is that parameterization requires image processing expertise, infeasible in novice point-of-care applications.

Machine learning methods have gained significant interest in medical imaging analysis. To our knowledge, no machine learning method has demonstrated capability of general shadow detection from multiple anatomy. Deep learning methods have identified features in a specific image sets that contain shadows, such as neuroanatomical regions in cranial scan (Milletari et al., 2017) or spinal

*Corresponding Author: Ricky Hu, Robotics and Control Laboratory, University of British Columbia, Room 3090, 2332 Main Mall, Vancouver, BC, Canada V6T 1Z4. Email: rhu@ece.ubc.ca

levels in a posterior scan (Hetherington et al., 2017). Although machine learning has the potential of providing automated feature recognition in multiple applications, a large data set is required for an algorithm to recognize certain features. Ultrasound imaging is highly variable due to unique artifacts, operator technique, and equipment. In addition, shadows are a very general feature that occur in various imaging scenarios. Previous techniques focus on a single anatomical region and training data was from a consistent imaging scenario. However, it is difficult to construct a training data set with the generality required to recognize shadows in different scenarios usable for a variety of ultrasound applications.

There are two objectives to this paper. First, to address the need for understanding general characteristics of shadows, a study was conducted to scan multiple anatomy and transducers specifically to analyze the statistics of different types of shadows. Second, to address existing needs for versatile detection with minimal parameterization, previous methods were then extended utilizing statistical thresholding of radiofrequency (RF) or brightness-mode (B-mode) data to detect shadows from various imaging scenarios.

Materials and Methods

Data Collection

Ultrasound RF and B-mode data was acquired by scanning 37 adult participants with informed written consent, approved by the University of British Columbia Research Ethics Board (Study ID: H18-01199). The scans included a forearm scan near the distal end of the pronator quadratus, an elbow scan near the cubital fossa, and a rib scan on the anterior surface of right ribs 11-12. Each scan was taken with both a curvilinear (C5-2/60, Ultrasonix, Canada) and linear (L14-5/38, Ultrasonix, Canada) transducer. Different transducer settings were used for each anatomical region and transducer, summarized in Table 1. Shadows are expected to occur due to superficial and deep bones and from an air gap created by the lateral edges of the transducer not being in flush contact with the skin. The experiment was designed to generate a dataset from various imaging scenarios to explore general shadow characteristics and to validate the versatility of the shadow detection methods.

Radiofrequency Speckle Analysis

To analyze shadows, windows of speckle were analyzed on the RF signal. Speckle occurs due to multiplicative scattering of acoustic waves in a material, resulting in a granular appearance on the image. The benefit of RF analysis is that B-mode image processing commonly attempts to remove speckle, but speckle contains information of the acoustic interactions in tissue (Burckhardt, 1978). Speckle can then characterize different regions, such as a region of tissue or a region of signal loss in a shadow. In addition, B-mode image formation can be manipulated by an operator to visually enhance an image, such as adjusting time-gain compensation or dynamic range. Thus, the underlying speckle analysis can provide shadow detection usable across different machines and operators.

One of the first models for speckle is with a one parameter Rayleigh distribution to model the probability density of a random walk (Burckhardt, 1978). The Rayleigh distribution is capable for modeling fully developed speckle, which does not occur in limited scattering (Tuthill et al., 1988). More generalized models have been applied such as the Rician, Homodyned-K, and Nakagami distributions to characterize speckle (Destrempes and Cloutier, 2010). The utility of speckle has been demonstrated in the literature to classify tumorigenicity of breast lesions (Byra et al., 2016) or levels of liver fibrosis (Ho et al., 2012) by categorizing image regions based on the speckle pattern. Shadow characterization presents a simpler problem as a shadow and non-shadow region contain significantly different speckle patterns. Thus, the Nakagami distribution expressed in Eq. 1 was chosen to model speckle. The Nakagami distribution provides greater generality than the Rayleigh distribution while being more computationally efficient than the Rician or Homodyned K distributions (Destrempes and Cloutier, 2010).

$$\Phi(x,\mu,\omega) = 2(\frac{\mu}{\omega})^{\mu}\frac{1}{\Gamma(\mu)}x^{(2\mu-1)}e^{\frac{-\mu}{\omega}x^2} \tag{1}$$

where x is RF intensity, μ is a shape parameter, ω is a scale parameter and $\Gamma(\mu)$ is the gamma distribution.

To characterize shadows, the raw RF data was first processed by computing the echo envelope of each scanline with a Hilbert transform. This creates a pre-scan converted image, visually similar to B-mode but without filtering to remove speckle. Next, the RF image was divided into overlapped windows with a width of a single RF data point and a length of three times the pulse width. This patch size was demonstrated in literature to be sufficiently large to capture multiple wavelengths and scattering events while being small enough to be useful in differentiating different regions on the millimeter scale (Byra et al., 2016). Next, each window was fit to a Nakagami distribution using a maximum likelihood estimate to compute a map of Nakagami parameters μ and ω, as shown in Fig. 1.

To detect shadows, Otsu's method was applied on the entire image to automatically compute a threshold for the ω parameter. This was sufficient as the ω parameter is significantly different for shadow regions with abundant speckle and non-shadow regions with minimal speckle, Then, for each scanline, the axially deepest data point that is above the threshold is labeled as the shadow boundary and all data points below are labeled as a shadow.

B-mode Scanline Analysis

Many ultrasound machines do not provide access to RF data for speckle analysis. Thus, a previous pixel-intensity shadow detection method on B-mode images was modified and extended. Scanline entropy was investigated on B-mode images to characterize different types of shadows, but with the addition of adaptive thresholding of entropy to address the need for usability with minimum configuration. First, the cumulative scanline entropy is computed for each pixel, similar to the "Rupture Criterion" (Hellier et al., 2010), with the window size fixed as three times the pulse width, η, as defined in Eq. 2. For the case of curvilinear images, radial scanlines were linearly interpolated between the two symmetric lateral edges of the image.

$$S_{i,j} = \sum_{i=1}^{3\eta} I(i-1) log_2 \frac{I(i-1)}{I(i+1)} + I(i+1) log_2 \frac{I(i+1)}{I(i-1)} \quad (2)$$

where $S_{i,j}$ is the cumulative entropy at pixel i on scanline j, η is the pulse width, and $I(i)$ is the pixel intensity of i.

Next, Otsu's method is applied similarly to compute a threshold entropy value. The intuition of the threshold is different than in RF analysis. In RF analysis, the threshold separates patches of intense and minimal speckle. In B-mode analysis, the threshold separates pixels of a shadow boundary, which has high entropy, and pixels away from shadow boundary, which include shadow and non-shadow regions. Thus, shadows can be identified by finding the last pixel on a scanline with an entropy higher than the threshold, representing a bright shadow boundary.

Validation

A trained annotator (RH) manually outlined the boundary of the shadow regions on B-mode images. The manual regions were used as a gold standard, as manual identification is common in clinical practice and has been used in previous literature for comparison (Hellier et al., 2010). A Dice coefficient was computed to compare similarity of manual and automated shadow detection. The manual outline was used to define four regions for classification of statistical parameters: a non-shadow region above the boundary, a shadow region below the boundary, a "transition region", which is a window defined as three pulse widths long axially below the boundary, and a "deep shadow region", which is the data below the transition region.

Results

Examples of detected shadows from both methods are highlighted in gray in Fig. 2 in different imaging scenarios. The Dice coefficients for both methods for different anatomy and transducers are shown in Table 2. The mean Dice coefficients ($\pm$ standard deviation) were 0.90$\pm$0.07 and 0.87$\pm$0.08 for RF and B-mode methods. Manual annotation was repeated five times with a mean Dice coefficient of 0.92$\pm$0.02 for all images and transducers.

With the benefit of a varied dataset, general statistics of shadows can be analyzed, as summarized in Table 3 and Table 4. For shadow detection, the parameters between a shadow and non-shadow are of particular interest. Shadows were observed to have a mean Nakagami ω parameter of 4.14 $\pm$ 0.40 and a mean entropy of 1.03 $\pm$ 0.29 whereas non-shadows were observed to have a mean ω of 6.24 $\pm$ 0.92 and 2.20 $\pm$ 0.81.

Discussion

The RF and B-mode shadow detection developed achieved a high Dice similarity coefficient for all anatomy and transducer types. Previous studies reported a mean Dice coefficient between manual annotators of 0.91$\pm$0.07 (Hellier et al., 2010). Every scenario detected from both methods achieved a was within the range of manual detection within operator variability. An important feature of shadow detection is being able to differentiate between a shadow and simply high attenuation of the signal. Both scenarios result in an eventual loss of signal. Shadow detection, however, has a characteristic high intensity shadow boundary before a significant loss in signal, compared to gradually signal losses in attenuation. Both methods were capable of this distinction. The high accuracy supports the versatility of the detection method as both methods are able to identify shadows across different anatomy and transducers with minimum configuration.

For a general observation for shadows, the computed Nakagami ω parameter of all manually outlined shadows indicate that there is a statistically significant difference between shadow and non-shadow regions, regardless of anatomy and transducer and even with the error in the transition regions considered. The speckle and its statistics from shadows is thus distinct from the speckle created by tissue, muscle, or fat. This observation can be utilized in the future for further analyze of shadows.

In RF detection, both false positive and false negative errors most frequently occurred immediately below a shadow boundary as opposed to B-mode detection where errors were in various regions. To study the frequent areas of error further, the "transition region" immediately below a manually annotated shadow boundary and a "deep shadow region" below the transition region was investigated. The Nakagami ω parameter of transition regions of all anatomy and transducers were within a standard deviation of both shadow and non-shadow regions. The deeper shadow regions were observed to have a lower Nakagami ω parameter than shadow regions and with a lower standard deviation as summarized in Table 3. The spread of the speckle also significantly decreases after the transition region. This indicates that the transition region cannot be fully distinguished from either a shadow or non-shadow and presents as it is statistically similar to the two. This is likely the cause of the errors, as the speckle distribution is much more consistent in the deep shadow regions compared to any other region. Physically, speckle interactions appear to gradually lessen after a brightest point on a scanline, possibly due to incomplete total reflection at a boundary. The boundary is thus is not an instantaneous division between non-shadow and shadow, rather, there is a transition region with statistics between a shadow and non-shadow before the speckle fully resembles a shadow.

In the transition region of B-mode images, the entropy values were similar but consistently higher than non-shadow values. This is expected as entropy is expect to be the highest when there is the greatest change in pixel intensity, which occurs at a shadow boundary, even with the a non-instantaneous non-shadow to shadow transition. However, the averaged entropy of all non-shadow regions have a greater spread than the Nakagami parameters, likely due to the differing operator settings used. Thus, B-mode detection may not be as consistent as RF detection.

In previous literature, shadows were defined qualitatively (Kremkau and Taylor, 1986) as a sudden loss of signal and brightness. The observed transition region in this study suggests that the qualitative definition of a shadow may be insufficient for accurate detection. One algorithm may detect the shadow starting immediately after the brightest location, or another may use a convention such as a full width at half maximum to define where the signal has sufficiently low intensity to resemble the start of a shadow. There is a decision point required for a clear definition for where a shadow begins to improve shadow detection accuracy, both from a signaling perspective for image processing and a visual perspective for manual inspection.

The findings in this study result in several implications. First, the statistics of acoustic shadows have been investigated on a dataset with shadows occurring from multiple scenarios as opposed to specific cases where shadows are observed. This provided a more generalizable observation that shadows can be characterized by distinctive speckle distributions regardless of anatomy and equipment and that there exists a transition region before the loss of speckle in a shadow. Second, the shadow detection methods demonstrated high accuracy, indicating that accurate shadow detection is possible regardless of transducer or imaging location. In future studies, the speckle statistics observed can be used to develop further models for anatomical features containing shadows, such as training a machine learning algorithm to measure the size of gall stones. Future studies would also have to take into consideration the most frequent source of error of shadow detection as the shadow boundary. For instance, a convention can be develop for one algorithm to detect the surface of the gallstone as the brightest point on a scanline to be comparable with other images.

Conclusions

Acoustic shadows from different imaging scenarios were investigated. RF and B-mode methods were developed for acoustic shadow detection requiring only the transducer pulse width as the input parameter. When comparing to manual detection, the methods achieved a Dice similarity coefficient within range of manual observers. The work focused on applying shadow detection and statistical analysis to a varied dataset of three different anatomical locations and two different transducer to provide a representative understanding of general acoustic shadows. The statistics of acoustic shadow indicate that shadows contain a distinct speckle distribution compared to non-shadows and the speckle characteristics transition at the shadow boundary. The statistical findings of shadows can aid interpretation of ultrasound images in the future using speckle analysis. The versatility of the shadow detection method has the potential to improve the interpretation of ultrasound images with shadow artifacts or to serve as a pre-processing step for machine learning methods.

Acknowledgements

This work is supported by the National Sciences and Engineering Research Council of Canada.

References

Bouhemad B, Brisson H, Le-Guen M, Arbelot C, Lu Q, Rouby JJ. Bedside ultrasound assessment of positive end-expiratory pressure-induced lung recruitment. American Journal of Respiratory and Critical Care Medicine, 2011;183:341–347.

Burckhardt CB. Speckle in ultrasound B-mode scans, 1978.

Byra M, Nowicki A, Wróblewska-Piotrzkowska H, Dobruch-Sobczak K. Classification of breast lesions using segmented quantitative ultrasound maps of homodyned K distribution parameters. Med. Phys., 2016;43:5561–5569.

Destrempes F, Cloutier G. A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope. Ultrasound Med. Biol., 2010;36:1037–1051.

Galiano K, Obwegeser AA, Bodner G, Freund M, Maurer H, Kamelger FS, Schatzer R, Ploner F. Ultrasound guidance for facet joint injections in the lumbar spine: A computed tomography-controlled feasibility study. Anesthesia and Analgesia, 2005;101:579–583.

Ghose S, Oliver A, Mitra J, Martí R, Lladó X, Freixenet J, Sidibé D, Vilanova JC, Comet J, Meriaudeau F. A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images. Medical Image Analysis, 2013;17:587–600.

Good LI, Edell SL, Soloway RD, Trotman BW, Mulhern C, Arger Pa. Ultrasonic properties of gallstones. Effect of stone size and composition. Gastroenterology, 1979;77:258–263.

Hellier P, Coupé P, Morandi X, Collins DL. An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. Medical Image Analysis, 2010;14:195–204.

Hetherington J, Lessoway V, Gunka V, Abolmaesumi P, Rohling R. SLIDE: automatic spine level identification system using a deep convolutional neural network. International Journal of Computer Assisted Radiology and Surgery, 2017;12:1189–1198.

Ho MC, Lin JJ, Shu YC, Chen CN, Chang KJ, Chang CC, Tsui PH. Using ultrasound Nakagami imaging to assess liver fibrosis in rats. Ultrasonics, 2012;52:215–222.

Karamalis A, Wein W, Klein T, Navab N. Ultrasound confidence maps using random walks. Medical Image Analysis, 2012;16:1101–1112.

Kremkau FW, Taylor KJ. Artifacts in ultrasound imaging. Journal of Ultrasound in Medicine, 1986;5:227–237.

Milletari F, Ahmadi SA, Kroll C, Plate A, Rozanski V, Maiostre J, Levin J, Dietrich O, Ertl-Wagner B, Bötzel K, Navab N. Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. Computer Vision and Image Understanding, 2017;164:92–102.

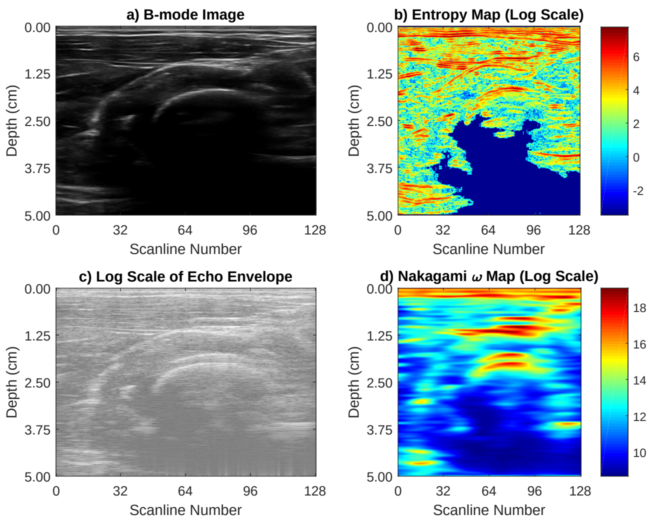
Sippel S, Muruganandan K, Levine A, Shah S. Review article: Use of ultrasound in the developing world. Int. J. Emerg. Med., 2011;4:72.

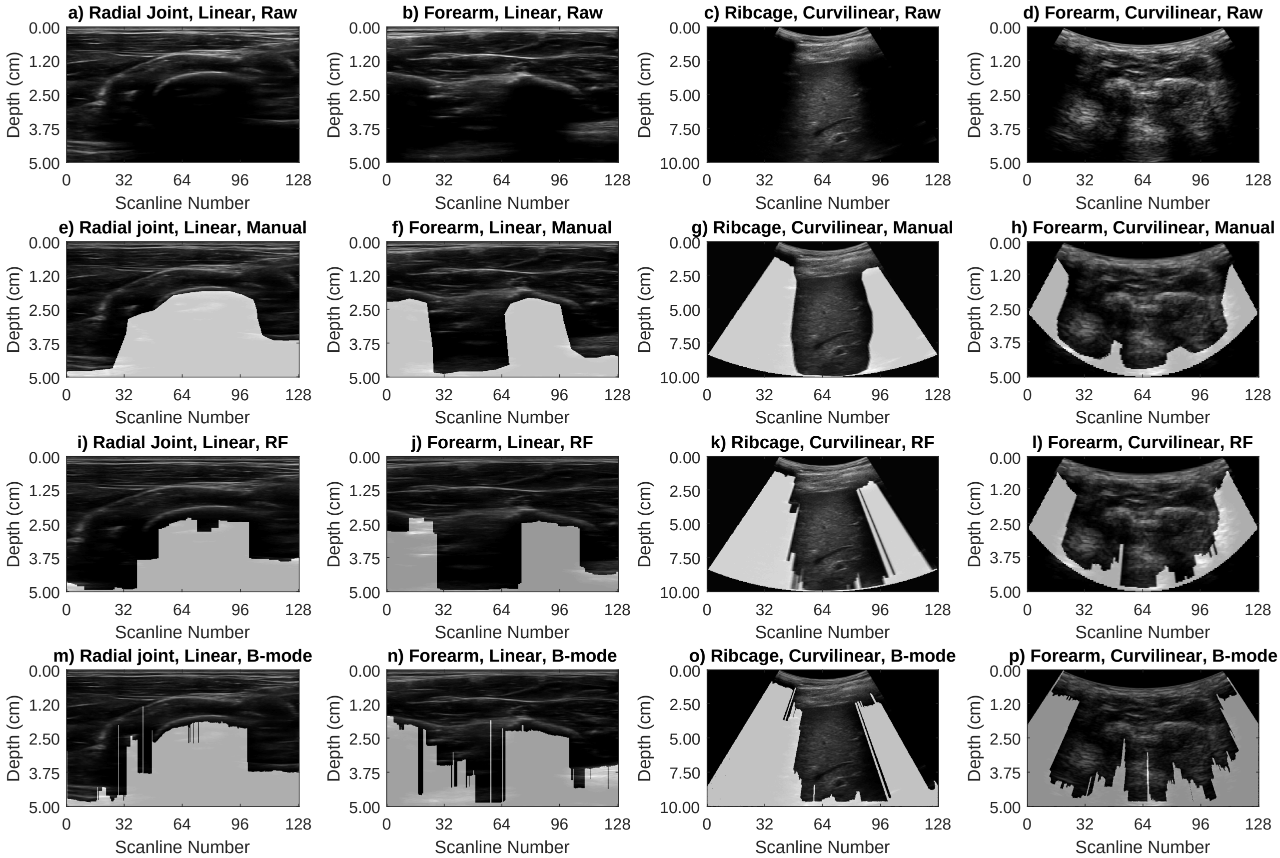
Tuthill TA, Sperry RH, Parker KJ. Deviations from rayleigh statistics in ultrasonic speckle. Ultrasonic Imaging, 1988;10:81–89.

Figure Captions

Figure 1: A visualization of the B-mode and RF parameter maps. The b) Entropy Map was computed from processing of the a) original B-mode image and the d) Nakagami ω map was computed from the c) echo envelope. Note that the echo envelope contains noticeable speckle, which has been used to fit a Nakagami distribution to characterize shadow. The region at depth 2.50cm and scanlines 32-40 is attenuation and not a shadow. This is an important distinction in shadow detection and both maps show the region as below a threshold to flag a shadow boundary.

Figure 2: A comparison of the original B-mode images, the detected shadows manual detection, RF detection, and B-mode detection. Both detection methods perform similar to manual detection. Both methods perform slightly less accurately on curvilinear images, likely due to the reduced resolution from interpolating the scanlines. Most errors of RF detection occur near the shadow boundary, likely due to the transitioning speckle from non-shadow to shadow.





Tables

Table 1: Transducer properties for different imaging scenarios.

	Anatomy	Frequency	Depth	Gain
Linear Transducer (L14-5/38)	Forearm	11.0MHz	5.0cm	50%
	Elbow	11.0MHz	5.0cm	40%
	Ribcage	5.0MHz	10.0cm	30%
Curvilinear Transducer (C5-2/60)	Forearm	4.0MHz	5.0cm	50%
	Elbow	4.0MHz	5.0cm	40%
	Ribcage	3.3MHz	10.0cm	30%

Table 2: Mean Dice coefficients for different imaging scenarios ± standard deviation.

		RF	**B-Mode**
Linear (L14-5/38)	Forearm	0.91±0.05	0.89±0.06
	Elbow	0.94±0.06	0.90±0.07
	Ribcage	0.87±0.09	0.84±0.06
Curvilinear (C5-2/60)	Forearm	0.89±0.05	0.86±0.08
	Elbow	0.93±0.04	0.90±0.09
	Ribcage	0.83±0.08	0.83±0.10
Mean	**All Anatomy**	**0.90±0.07**	**0.87±0.08**

Table 3 : The mean Nakagami ω and Entropy values of different anatomy, transducer, and shadowing region ± standard deviation.

	Linear (L14-5/38)			**Curvilinear (C5-2/60)**		
	Forearm	Elbow	Ribcage	Forearm	Elbow	Ribcage
Nakagami ω (Log Scale)						
Shadow	4.15 ± 0.45	4.18 ± 0.45	4.04 ± 0.42	4.22 ± 0.32	4.19 ± 0.40	4.08 ± 0.37
Non-Shadow	6.19 ± 0.96	6.49 ± 0.97	6.29 ± 0.95	6.54 ± 0.88	6.29 ± 1.04	5.64 ± 0.71
Transition	4.94 ± 0.62	5.36 ± 0.62	4.96 ± 0.38	5.26 ± 1.02	5.37 ± 0.99	4.59 ± 0.92
Deep Shadow	4.13 ± 0.43	4.16 ± 0.43	4.03 ± 0.41	3.93 ± 0.20	4.09 ± 0.30	4.03 ± 0.26
Entropy (Log Scale)						
Shadow	0.92 ± 0.22	1.10 ± 0.36	1.04 ± 0.27	1.06 ± 0.28	0.96 ± 0.21	1.10 ± 0.37
Non-Shadow	2.34 ± 0.96	2.34 ± 0.80	2.14 ± 0.82	1.67 ± 0.82	1.75 ± 1.14	1.88 ± 0.42
Transition	2.45 ± 0.62	2.56 ± 0.53	2.15 ± 0.51	2.18 ± 1.21	1.93 ± 1.10	1.99 ± 1.10
Deep Shadow	0.71 ± 0.43	0.89 ± 0.26	0.92 ± 0.40	0.98 ± 0.21	0.82 ± 0.19	1.04 ± 0.26

Table 4 : The mean Nakagami ω and Entropy values of all anatomy and transducers for different shadowing regions ± standard deviation.

	Mean Nakagami ω (Log Scale)	**Mean Entropy (Log Scale)**
Shadow	4.14 ± 0.40	1.03 ± 0.29
Non-Shadow	6.24 ± 0.92	2.02 ± 0.81
Transition	5.08 ± 0.77	2.21 ± 0.84
Deep Shadow	4.06 ± 0.34	0.89 ± 0.27