Unsupervised clustering method to convert high-resolution magnetic resonance volumes to three-dimensional acoustic models for full-wave ultrasound simulations

Kevin Looby
Carl D. Herickhoff
Christopher Sandino
Tao Zhang
Shreyas Vasanawala
Jeremy J. Dahl

Unsupervised clustering method to convert high-resolution magnetic resonance volumes to three-dimensional acoustic models for full-wave ultrasound simulations

Kevin Looby,^a Carl D. Herickhoff,^b Christopher Sandino,^c Tao Zhang,^d Shreyas Vasanawala,^e and Jeremy J. Dahl^{b,*}

^aStanford University, Department of Electrical Engineering, Palo Alto, California, United States

^bStanford University, Department of Radiology, Palo Alto, California, United States

^cStanford University, Department of Electrical Engineering, Stanford, California, United States

^dSubtle Medical, Menlo Park, California, United States

^eStanford University, Department of Radiology, Palo Alto, California, United States

Abstract. Simulations of acoustic wave propagation, including both the forward and the backward propagations of the wave (also known as full-wave simulations), are increasingly utilized in ultrasound imaging due to their ability to more accurately model important acoustic phenomena. Realistic anatomic models, particularly those of the abdominal wall, are needed to take full advantage of the capabilities of these simulation tools. We describe a method for converting fat–water-separated magnetic resonance imaging (MRI) volumes to anatomical models for ultrasound simulations. These acoustic models are used to map acoustic imaging parameters, such as speed of sound and density, to grid points in an ultrasound simulation. The tissues of these models are segmented from the MRI volumes into five primary classes of tissue in the human abdominal wall (skin, fat, muscle, connective tissue, and nontissue). This segmentation is achieved using an unsupervised machine learning algorithm, fuzzy c-means clustering (FCM), on a multiscale feature representation of the MRI volumes. We describe an automated method for utilizing FCM weights to produce a model that achieves ~90% agreement with manual segmentation. Two-dimensional (2-D) and three-dimensional (3-D) full-wave nonlinear ultrasound simulations are conducted, demonstrating the utility of realistic 3-D abdominal wall models over previously available 2-D abdominal wall models. © 2019 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.6.3.037001](https://doi.org/10.1117/1.JMI.6.3.037001)]

Keywords: ultrasound; simulation; machine learning; modeling.

Paper 18254RR received Nov. 30, 2018; accepted for publication Jul. 2, 2019; published online Jul. 22, 2019.

1 Introduction

Phase aberration and reverberation clutter are ultrasound image-degrading phenomena frequently encountered when imaging overweight, obese, and difficult-to-image patients. Loss in image quality from these mechanisms can impede diagnoses and diminish the quality of care afforded to these patients.¹ Phase aberration and clutter are generated by complex tissue distributions that are not easily modeled, making their image-degrading effects difficult to simulate. The development of complex tissue models [based on *in vivo*, high-resolution three-dimensional (3-D) images] for both linear and nonlinear ultrasound wave propagation simulations is essential for studying these phenomena and developing techniques to mitigate their effects.

Several previous full-wave ultrasound simulations have used two-dimensional (2-D) models of complex tissue that have been generated from histological staining^{2–4} or computational models.⁵ However, these 2-D models are unable to incorporate wave propagation and scattering effects in the elevation dimension and are thus insufficient for modeling the complex 3-D nature of tissue.

Some previous approaches to obtain tissue models for acoustic simulation have employed 3-D magnetic resonance imaging (MRI) or computed tomography (CT) datasets, with voxel size

as small as $1 \times 1 \times 2$ mm, for either slice-by-slice segmentation or a direct mapping to tissue types based on Hounsfield units;^{6,7} other approaches have relied on photographs of cross sections to obtain voxels as small as $0.33 \times 0.33 \times 0.33$ mm, segmented based on manual sampling and mapped to an empirical continuum model of fractional tissue components.^{8,9} In these cases, acoustic simulations assumed linear propagation.

The 3-D acoustic models of complex human tissue previously have been generated from MRI volumes of excised human breast tissue¹⁰ and from MRI and x-ray CT images made available through the National Library of Medicine's The Visible Human Project.^{11–13} However, these models do not represent tissue layout in its natural state. The processes of excising and/or fixing tissue used in histological staining and in The Visible Human Project result in tissue shrinkage and deformations that are not seen *in vivo*.

Creating labeled models by hand through the use of histology and other imaging modalities is difficult and time-consuming. Manual segmentation often requires the effort of an expert, such as a pathologist or radiologist, to identify and label tissues, a task made more difficult and tedious by volumetric data. As such, there is a need for an automated approach to generate 3-D acoustic models that are true to *in vivo* conditions.

*Address all correspondence to Jeremy J. Dahl, E-mail: jjdahl.stanford.edu

Fuzzy c-means clustering (FCM) has previously been used for a variety of image segmentation tasks with MRIs,^{14–19} and Markov random fields (MRFs) in conjunction with FCM have also been used in image segmentation tasks, including image segmentation under noisy conditions²⁰ and for detecting changes in radar images.²¹ A procedure utilizing FCM^{22,23} in conjunction with MRFs is presented here to automatically convert MRIs into 3-D acoustic models.

2 Background

2.1 Fuzzy C-Means

The FCM algorithm is an unsupervised learning method that seeks to find an optimal grouping of data into i clusters by minimizing the objective function $J_i(u, v)$, given as

$$J_i(u, v) = \sum_{k=1}^N \arg \min_i (u_{ik})^m \|y_k - v_i\|^2. \quad (1)$$

More precisely, FCM minimizes intracluster variation by assigning the k 'th data point (y_k) weighted memberships (u_{ik}) to each cluster (c_i), where a “cluster” is one of the several partitions in the data that the algorithm seeks to define. The weights, u_{ik} , may be interpreted as probabilities of membership; that is, the FCM algorithm can be viewed as a method for approximating the likelihood of a data point belonging to each of the i learned clusters. In the context of modeling abdominal tissue layers, u_{ik} is considered to be the probability that voxel y_k belongs to cluster $c_i \in \{\text{skin, fat, muscle, connective tissue, no tissue}\}$.

The membership weight for the k 'th data point and cluster c_i is given by

$$u_{ik} := \frac{1}{\sum_{n=1}^C \left(\frac{\|y_k - v_i\|}{\|y_k - v_n\|} \right)^{\frac{2}{m-1}}}, \quad (2)$$

where v_i is the centroid, or center of mass, of cluster c_i and is given by

$$v_i := \frac{\sum_{k=1}^N (u_{ik})^m y_k}{\sum_{k=1}^N (u_{ik})^m}, \quad (3)$$

with the constraint $m > 1$.

The weights, u_{ik} , in Eq. (2) are typically initialized randomly, and the centroids, v_i , are calculated via Eq. (3) using the initialized weights. The algorithm then proceeds iteratively, alternately updating the weights and centroids until the objective function converges. The parameter m in Eqs. (2) and (3) determines the degree to which data may belong to a single cluster (“hard” clustering, $m \approx 1$) or multiple clusters (as m increases, membership weights approach a uniform distribution). “Hard” cluster assignments may be computed from the fuzzy weights as

$$c_i := \arg \min_j (u_{ik})^m \|y_k - v_j\|^2. \quad (4)$$

2.2 Markov Random Fields

An MRF is a natural prior in many image segmentation tasks and serves as a model for the ways in which the assigned labels (i.e., random variables of the MRF) are able to interact with each other. The MRF assigns higher probabilities to labels in

which neighboring pixels are assigned matching labels, thereby encouraging neighboring pixels to take on the same label. In doing so, spatial consistency is enforced, resulting in less noise and more connectivity in tissue structures.

If the probability distribution, $P(x)$, of all possible pixel label assignments is positive [i.e., $P(x) > 0, \forall x \in X$], then an MRF can be modeled as a Gibbs probability distribution, according to the Hammersley–Clifford theorem, as

$$P(x) = \frac{e^{-U(x)}}{\sum_{x \in X} e^{-U(x)}}. \quad (5)$$

Equation (5) describes a set of probabilities over the possible label assignments $x \in X$ for a given Markov blanket. A label assignment x refers to a possible labeling of a random variable (pixel label assignment) and all other random variables in its Markov blanket, and X is the set of all such possible label assignments. Here, U is the energy function for the Gibbs distribution and is calculated by summing the clique potentials, V_q , across the Markov blanket for a given label assignment x , as follows:

$$U(x) = \sum_{q \in M} V_q(x_i, x_{i'}). \quad (6)$$

The clique potential is a value representing the likelihood of the assignments found in a given clique, where V_q equals $+\delta$ for $x_i \neq x_{i'}$ and $-\delta$ otherwise. The parameter δ is an empirically chosen constant, determined by the size of the Markov blanket and the degree of certainty desired from the prior.

2.3 Reformulated Fuzzy C-Means

The FCM algorithm is reformulated here to incorporate the MRF as a prior, similar to Honglei et al.¹⁴ Eq. (5) is used to calculate $P(x)$, the $N \times C$ set of the MRF's estimates of the probabilities that each of the N data points belongs the C clusters. We let p_{ik} denote the MRF's estimation that data point k belongs to cluster i and include this term in the updated objective function, given by

$$J(u, v, p) = \sum_{k=1}^N \sum_{i=1}^C (1 - p_{ik}) (u_{ik})^m \|y_k - v_i\|^2, \quad (7)$$

given the constraint

$$\sum_{k=1}^N u_{ik} = 1. \quad (8)$$

The constraint that u_{ik} must sum to 1 across the different clusters ensures that u_i represents a probability distribution over the clusters k .

Equations (7) and (8) form an optimization problem that can be solved by the method of Lagrange multipliers. The appropriate Lagrangian for this solution is

$$J_i(u, v, p) = \sum_{k=1}^N \sum_{i=1}^C (1 - p_{ik}) (u_{ik})^m \|y_k - v_i\|^2 - \sum_{i=1}^C \lambda_i \left(\sum_{k=1}^N u_{ik} - 1 \right), \quad (9)$$

where the Lagrange multiplier λ_i is introduced to incorporate the constraint in Eq. (8). Solving Eq. (9) by the method of Lagrange multipliers provides the new update rules for the cluster centroids and weights

$$v_i := \frac{\sum_{k=1}^N (1 - p_{ik})(u_{ik})^m y_k}{\sum_{k=1}^N (1 - p_{ik})(u_{ik})^m}, \quad (10)$$

$$u_{ik} := \left[\sum_{l=1}^N \frac{(1 - p_{ik})^{\frac{1}{m-1}} \|y_k - v_i\|^{\frac{2}{m-1}}}{(1 - p_{il})^{\frac{1}{m-1}} \|y_l - v_i\|^{\frac{2}{m-1}}} \right]^{-1}. \quad (11)$$

2.4 Application to Magnetic Resonance Imaging Volumes

The modified FCM algorithm outlined in Sec. 2.3 is used to segment MRI volumes in a way that facilitates the production of “hard” and “soft” acoustic models. After an appropriate MRI volume has been acquired, a feature space is produced for the clustering algorithm. The feature space may be as simple as taking the MRI signal intensity for each voxel; however, more complex feature spaces may be used to emphasize other aspects of the data.

Once the feature space has been produced, the modified FCM algorithm is applied to the data. First, the weights u_{ik} are randomly initialized, and one iteration of the update rules [Eqs. (10) and (11)] is completed. After initial estimates are made for v_i and u_{ik} , the MRF is evaluated with Eqs. (5) and (6).

The algorithm then proceeds by alternately updating v_i [Eq. (10)], u_{ik} [Eq. (11)], and the MRF prior [Eqs. (5) and (6)]. After each set of updates, the objective function is evaluated for convergence. If the change in the objective function between two iterations is less than the threshold γ , the algorithm is terminated.

Once convergence is achieved, the weight matrix is used to produce an acoustic model. A hard acoustic model, wherein every physical location in the model is identified as exactly one of the tissue classes, is obtained by taking an argmax across the calculated weights. A soft model is produced by making use of the tissue parameter mixing equation²⁴ (see Sec. 3.4).

3 Methods

3.1 Data Acquisition

In an *in-vivo* human patient, a fat–water-separated MRI volume was acquired using the dual-echo Dixon imaging technique.²⁵ This method makes use of the chemical shift that occurs between water molecules and fat molecules to separate water- and fat-based tissues. This is achieved by acquiring “in-phase” images, where the water and fat signals are completely in phase, and “out-of-phase” images, where the two signals are 180 deg out of phase. Fat–water-separated images were extracted from the in-phase and out-of-phase images using the projected power method for resolving phase ambiguity, as described by Zhang et al.²⁶

The fat–water-separated volume was acquired at a resolution of 312 μm in-plane and 500 μm through-plane (i.e., between each 2-D imaging plane). Sample images from the acquired dataset are provided in Fig. 1. Imaging was performed on a 3T GE MR750 MRI scanner with a 32-channel cardiac coil at the Lucille Packard Children’s Hospital at Stanford, in Palo Alto, California. A dual-echo 3-D spoiled gradient echo sequence

Fig. 1 (a) Representative axial slice from the water signal volume and (b) corresponding axial slice from the fat signal volume.

was used with TE1/TE2 of 3.3/4.4 ms, TR of 8 ms, and flip angle of 15 deg. Informed consent was obtained under an approved Institutional Review Board protocol (ID# 22454).

3.2 Preprocessing

The MRI volumes were first registered using intensity-based registration (imregister, MATLAB, Natick, Massachusetts) to ensure proper alignment between corresponding structures in each volume. Because imregister relies on correlation between the volumes to effectively register them, additional preprocessing is necessary to ensure that fat and water signals are represented consistently across volumes (i.e., fat signals are dark and water signals are bright in both volumes).

The preprocessing step essentially inverts the fat and water signals in the fat volume while keeping the background signal the same. Briefly, a binary mask is created by summing the fat and water volumes and thresholding the summed volume. The fat signal volume is then subtracted from the mask to replace tissue signal voxels with one minus their original values. Thereby, the water-based tissues are represented by large values in the volume, and fatty regions take on small values in both volumes. These transformed volumes are used as input to the registration algorithm. After registration, the images are cropped to remove any regions of nonoverlap between the volumes.

3.3 Feature Space

A total of 19 features comprise the feature space used here. This space includes three primary features, the first two of which are the registered water and fat volumes, where both are normalized to the range [0,1] [Figs. 2(a) and 2(b), respectively], and a third feature made by adding the normalized water volume to the modified fat volume that has each normalized pixel value subtracted from 1 [Fig. 2(c)]. This third feature is renormalized to the range [0,1] and was included because it allowed for all five tissue classes to be visualized in one feature.

The next 15 features were constructed by applying varying amounts of Gaussian blurring to the three primary features. Blurring was applied using Gaussian kernels with standard deviations of 0.04, 0.11, 0.17, 0.20, and 0.22 mm. These values were

Fig. 2 Selected features shown here include (a) the raw water volume, (b) the raw fat volume, (c) the combined volume, and (d) the skin isolation feature. The blurred features have been omitted for brevity.

selected by performing hyperparameter tuning, a process by which the number and magnitudes of the standard deviations were swept over a broad range of values to select those that produced the highest segmentation accuracy when compared with a manually segmented image (Sec. 4.1).

A final feature was included to enable the algorithm to discriminate between voxels representing skin from those representing muscle, because these two tissue types produced similar MRI signal intensities. Otsu's method²⁷ was applied to the water volume to create a binary mask. Each 3-D connected region was assigned a separate label and the region corresponding to the skin was extracted, resulting in a binary mask identifying regions containing skin. Blurring was applied to this feature to reduce the certainty near edges of the isolated skin structures. This feature is shown in Fig. 2(d).

3.4 Acoustic Map Production

Using the feature space and the methods described in Sec. 2, label weights (i.e., probabilities) were computed for each pixel in the MRI volume. The neighborhood for the MRF was defined to be the six voxels that share faces with the voxel in question. The energy function parameter [δ in Eq. (6)] was chosen to be 0.05. This value was chosen via a randomized hyperparameter search.

3.4.1 Hard segmentation

To create hard segmentation acoustic models of tissue, an arg-max operation was taken across the five clusters, and each voxel was assigned a label according to its highest weighting value u_{ik} (i.e., highest probability of membership to a particular cluster). Each cluster (i.e., set of labeled voxels) was then manually assigned a specific tissue type and set of acoustic parameters (speed of sound c_0 , density ρ_0 , attenuation α , and nonlinearity parameter B/A), as taken from Tabei et al.³ and given in Table 1. In ultrasound imaging, it is necessary to use a coupling agent (e.g., ultrasound gel or water), to fill the space between the imaging transducer and the tissue sample. For this reason, the "no tissue" cluster was treated as pure water at 37°C.

Table 1 Literature values for sound speed and density in relevant tissues.

Tissue	c_0 (m/s)	ρ_0 (kg/m ³)	α (dB/MHz/m)	B/A
Skin	1613	1120	2.1	8
Fat	1478	950	0.4	9.6
Muscle	1547	1050	0.15	9
Connective tissue	1613	1120	0.5	8
Water (37°C)	1524	993	0.005	5

3.4.2 Soft segmentation

The soft acoustic model was created by using the soft clustering weights directly as estimates of the tissue composition of each imaged voxel. That is, the tissue label of a given voxel is represented as a weighted mixture of the five possible tissue classes.

The speed of sound for a mixture of tissues is determined by Eq. (12), which is of the form given by Sehgal et al.²⁴

$$\frac{1}{c_{\text{mix}}} = \sum_{n=1}^C \frac{x_n}{c_n}. \quad (12)$$

The values for the remaining parameters were produced by a simple weighted average, as given in Eq. (13)

$$p_{\text{mix}} = \sum_{n=1}^C x_n p_n. \quad (13)$$

In these equations, x_n denotes the weight value computed for the n 'th cluster, c_n is the speed of sound for the tissue type associated with cluster n , and p_n represents any of the other required parameters (density, attenuation, and the nonlinearity parameters). By definition, the percentage compositions sum to unity.

3.5 Simulations

Ultrasound imaging simulations were performed to demonstrate the utility of the 3-D acoustic models using the acoustic simulation suite k-Wave,²⁸ to model nonlinear full-wave propagation using a mapped k-space pseudospectral method, as described by Treeby.²⁹ In each simulation, a 120-element linear array transducer (pitch = 0.370 mm, element height = 3.7 mm) was defined just above the skin layer of the acoustic model; a 1.5-cycle, 1 MPa amplitude pulse was transmitted at 1 MHz from a 64-element active aperture within this array. Scatterers were introduced by randomly modifying the sound speed of a subset of voxels by $\leq \pm 5\%$. A 7-mm anechoic lesion was placed at a depth of 3.75 cm. All scan lines were focused at a depth of 3.75 cm for transmit and receive (delay-and-sum) beamforming.

The simulated medium was padded on all sides by a perfectly matched layer with an attenuation of 2 dB/cm/MHz, in order to prevent signal from reflecting back from the edges of the simulated region.

To perform 2-D simulations, a single axial imaging plane was selected from the acoustic model and duplicated in the elevation dimension to create a surface large enough to accommodate the simulated transducer. Because the volume contained no variations in the elevation dimension, the result is an approximation to a true 2-D simulation with no elevation extent. Matched 3-D simulations were performed by centering the transducer above the axial imaging plane used in the 2-D simulations. In the 3-D simulations, the full acoustic model was incorporated, including the natural variations in the elevation dimension.

Simulations were performed using both hard and soft acoustic models. Each set of hard and soft model simulations contained four pairs of matched 2-D and 3-D simulations. Three imaging planes were selected from the abdominal layer acoustic model, and for each selected plane, a 2-D simulation and a 3-D simulation were performed. In addition, a 2-D and a 3-D homogeneous (no abdominal layer) simulation was performed as a reference case.

Isovelocity and isoimpedance simulations of nonlinear full-wave propagation as described in Pinton et al.¹ were also performed using the 3-D acoustic models, in order to analyze the reverberation (isovelocity) and phase aberration (isoimpedance) components individually. Briefly, isovelocity simulations were made by modifying the density and speed of sound of the modeled medium, such that the speed of sound was homogeneous throughout the medium and the density was altered to maintain the original acoustic impedance while accommodating the homogeneous speed of sound. An isoimpedance simulation was performed by subtracting the beamformed RF data of a simulation containing only the abdominal layer from a full simulation of the medium.

For each simulated image, the image brightness magnitude was normalized (note: because radiologists, sonographers, and other ultrasound users typically observe the magnitude of B-mode images, the image magnitude is typically the signal of interest; normalization yields unitless values) and the contrast-to-noise ratio (CNR—also referred to as the contrast-to-speckle ratio, as defined by Patterson et al.³⁰) was computed. This was done by first finding the average normalized-magnitude signal and the signal variance within the lesion (S_i and σ_i^2), and similarly finding the average signal and signal variance in the background outside the lesion (S_o and σ_o^2) at the same depth. The ratio of the contrast to the speckle noise may then be calculated as

$$\text{CNR} = \frac{S_o - S_i}{(\sigma_o^2 + \sigma_i^2)^{1/2}}. \quad (14)$$

4 Results

4.1 Segmentation

One cross section of the MRI volume was manually segmented to provide a measure of the algorithm's segmentation accuracy. The manually segmented image is shown in Fig. 3(b) and is compared to the algorithm's corresponding segmented image Fig. 3(a). The segmentation produced by the c-means algorithm is 89.49% in agreement with the hard manual segmentation over all tissue types. Per-class true-positive accuracies as well as intersection over union (IoU) values are given in Table 2.

4.2 Acoustic Models

Both hard and soft acoustic models were produced using the MRI volume and the methods described previously. Visualization of the weights computed by soft clustering are shown in Fig. 4. Cross sections of the sound speed and density

Fig. 3 (a) Manually segmented cross section and (b) corresponding automated hard segmentation. Each gray level indicates one of the five tissue classes, which are, in order of increasing brightness, skin, connective tissue, background, muscle, and fat.

Table 2 Clustering accuracy for each class.

Label	IoU	Accuracy (%)
Skin	0.847	84.26
Fat	0.827	90.84
Muscle	0.892	88.57
Connective tissue	0.549	70.60
No tissue	0.946	99.85

Fig. 4 Visualization of the weights calculated by fuzzy clustering. Images are labeled with the names of the corresponding tissue type.

Fig. 5 (a) Hard speed of sound map and (b) hard density map generated by taking an argmax over the soft weights. Together, the speed of sound model and the density model comprise an acoustic map. Color bar values are in units of m/s for the speed of sound and kg/m^3 for density.

components of the hard model are shown in Fig. 5, whereas the same cross sections of the soft model are shown in Fig. 6.

4.3 Simulations

4.3.1 Hard model simulations

The results of the four pairs of simulations using hard maps are shown in Fig. 7. The top row of Fig. 7 shows the center 2-D slice of the acoustic model that was imaged in each simulation. The bottom row displays the B-mode ultrasound images that were produced. Each imaging plane in the top row is accompanied by a matched pair of 2-D [Fig. 7(a)] and 3-D [Fig. 7(b)] simulations. All simulation images are displayed with a dynamic range of 60 dB. Because the clutter magnitude from the hard models can be stronger than what is observed *in vivo*, a spatial Gaussian blur with standard deviation of 3 pixels was applied across the hard models prior to the simulations to decrease clutter magnitude.

For reference, the leftmost pair of images in Fig. 7 are simulations of homogeneous tissue and did not include an abdominal-layer model. The remaining three pairs of images use the hard abdominal-layer model, with the center plane of the acoustic model (represented by the speed of sound values) shown above the pair of ultrasound images. Owing to propagation through and scattering from the inhomogeneous tissue

Fig. 6 (a) Soft speed of sound map and (b) soft density map generated by parameter mixing using the fuzzy clustering weights. Color bar values are in units of m/s for the speed of sound and kg/m^3 for density.

structures in the hard abdominal-layer model, significant clutter and aberration are apparent from the signal within the anechoic lesion and distortion of the shape of the lesion. This is in contrast to the homogeneous simulations, wherein dark circular lesions are clearly visible. In each matched pair of simulations, the 2-D and 3-D simulations both show overall image degradation but the amount and type of degradation differ.

4.3.2 Soft model simulations

Images showing the four pairs of 2-D/3-D simulations performed with the soft model are shown in Fig. 8. As before, the top row shows the central 2-D slice of the soft acoustic model, and the bottom row displays the B-mode images with a 60-dB dynamic range. The left image in each pair shows the image from the 2-D simulation, and the right image is the result from the 3-D simulation. As with the hard models, there are marked differences between the image degradation seen in each 2-D simulation compared to the corresponding 3-D simulation. The lesion CNR is displayed at the bottom of the image.

4.3.3 Isoimpedance and isovelocity simulations

Figure 9 shows the images from the 3-D simulations, including an image using the 3-D soft abdominal-layer model, a corresponding isoimpedance image, and a corresponding isovelocity

Fig. 7 Simulation results using the hard acoustic model. The top row of images shows the speed of sound values in the imaging plane; the bottom row shows the resulting simulations. Two simulations were run for each acoustic model location: the left is a 2-D simulation, and the right is the corresponding 3-D simulation. Lesion CNR values are provided at the bottom of each simulation.

image. Visually, the isoimpedance simulation is identical to the original simulation, and the measured CNR from these two images agree (both images have a CNR of 0.32). The reverberation clutter magnitude from the isoimpedance simulation was measured to be ~ 40 dB below the backscattered signal from the tissue region surrounding the anechoic lesion, indicating that this acoustic model produced a small amount of reverberation clutter.³¹

In the isovelocity simulation, the lesion is nearly perfectly reconstructed. The image appears highly similar in structure to the homogeneous 3-D simulation image shown in Fig. 8. The CNR of the lesion in the isovelocity simulation is 0.83, which is nearly identical to the CNR of the lesion in the 3-D homogeneous simulation. Because the isovelocity simulation removes phase aberration while preserving reverberation clutter, the near-perfect reconstruction of the lesion in the isovelocity case implies a significant amount of image degradation due to phase aberration. The combination of the results of both the isoimpedance and the isovelocity cases indicates that phase aberration is the primary source of noise in this particular abdominal model, with little impact from reverberation.

5 Discussion

The unsupervised learning method of reformulated FCM clustering presented here is used to segment *in-vivo* MRI volumes and yield realistic 3-D acoustic models of tissue; this method is less invasive and less labor-intensive than previous methods using excised or fixed tissue. *In-vivo* MRI volumes can provide the highest spatial resolution of tissue structure in the body,

but MRI voxel values are typically uncalibrated and unitless (unlike CT voxels, which are typically in Hounsfield units), so segmentation generally cannot be automated based on predetermined thresholds, and mapping to acoustic parameters will require some amount of manual examination. Compared to previous approaches, this method uses higher-resolution *in-vivo* 3-D MRI data (voxel volume $>40\times$ smaller than those in 3-D datasets used by Aroyan et al.⁷ and Mougenot et al.⁶), and, by grouping voxels into a specified number of clusters, provides effective autosegmentation and easy assignment of tissue type and acoustic properties for use in simulation. These 3-D models more accurately represent the medium that ultrasound imaging systems are applied to, and the results presented here demonstrate that 2-D simulations are not guaranteed to be an accurate representation of the true 3-D tissue.

The clustering segmentation method used here agreed well with the manually segmented cross section of tissue, with nearly 90% of the labels assigned by the clustering method matching the manually assigned labels. Because no ground truth is available for this segmentation process, comparison with a manual segmentation was used only as a qualitative evaluation metric. Misclassifications that arise in the comparison are not unexpected, and the majority of these errors occur in distinguishing connective tissues from fat. The connective tissue structures are the least well-preserved structure in the automated segmentation. This is attributed to the absence of a positive MRI signal for these tissues in the water volume. In Fig. 2, the connective tissues are nearly absent in the water volume, but they are visualized by an absence of signal in the fat volume. The connective

Fig. 8 Simulation results using the soft acoustic model. The top row of images shows the speed of sound mapping at the imaging plane; the bottom row shows the resulting simulations. Two simulations were run for each acoustic model location: the left is a 2-D simulation, and the right is the corresponding 3-D simulation. Lesion CNR values are provided at the bottom of each simulation.

tissue structures are also the smallest structures targeted by this method, and therefore improved resolution and signal intensity in the raw MRI volumes would substantially improve the algorithm's ability to distinguish these tissues.

While useful for general evaluation, the segmentation accuracy need not be a perfect match to the MRI images in this application. Given the lack of 3-D acoustic models to date, and especially those that represent *in-vivo* anatomy, it is necessary only to produce a model that reasonably approximates the true tissue structure in the MRI volume. In our evaluation here, we have captured a sufficiently detailed and accurate model of the actual tissue structure to yield expected phase aberration and reverberation clutter effects. The segmentation quality is limited by the signal-to-noise ratio and contrast of the input MRI volume. In addition, the method presented here relies on image-processing techniques to distinguish between the skin and the muscle classes. If the skin class is either not presented or not oriented properly, a failure case may result.

Although the 3-D models and simulations incorporate the elevation dimension, a consistent decrease in image degradation was not observed compared to the 2-D simulations, as might be expected due to the 2-D nature of phase aberration and reverberation clutter.^{11,32–34} The 3-D simulations show clutter and phase aberration in both greater and lesser extents to their 2-D simulation counterparts, meaning that the 2-D simulations are not a simple prediction of the 3-D tissue structure.

The interfaces between tissue types in the hard model have very high impedance mismatches and sharp boundaries, which result in strong acoustic reflections, reverberation, and phase

aberration that can often appear stronger than that observed *in vivo*. Methods for “softening” the hard acoustic model (i.e., reducing the strength of these image degrading effects) either requires adjusting the acoustic properties of the tissues to create smaller impedance mismatch or requires blurring the model to reduce sharp transitions between tissues. Adjusting the acoustic properties often requires values that are not well supported by measurements found in the literature. Blurring the hard model may be a better solution, but more sophisticated blurring methods that account for the geometry of the tissue structures would be more appropriate than the spatial Gaussian blurring applied here. The soft model is an alternative solution that preserves tissue structure and allows the MRI signal intensities to define a more natural transition between tissue types, which may be preferred to the uniformly blurred hard model. The CNR values of the soft acoustic models are larger than that of the blurred hard acoustic models, highlighting the key difference in these models. In the isovelocity and isoimpedance analyses of the soft model in Fig. 9, it is apparent that phase aberration was the dominating image-degradation mechanism for this particular model (i.e., soft segmentation and tissue structure). The mixing equation described by Sehgal and Greenleaf²⁴ provides an approximate framework for combining tissue parameters based on the tissue probabilities of the clustering algorithm; however, it may be possible to achieve more realistic mixing of acoustic parameters by using quantitative MRI techniques that better derive the actual fat and water concentrations of tissues³⁵ and perform mixing based on the concentrations of these tissue rather than cluster probabilities.

Fig. 9 (a) A 3-D B-mode simulation, (b) corresponding isoimpedance simulation, and (c) corresponding isovelocity simulation. Lesion CNR values are provided below each simulation.

6 Conclusion

A complete, automated method for generating high-resolution, 3-D acoustic models has been presented. This method takes, as input, a 3-D fat–water-separated MRI volume from dual-echo Dixon imaging and produces an output that is a volume where each spatial location in the input volumes has been assigned a label corresponding to one of the following tissue types: skin, fat, muscle, connective tissue, and “no tissue.” The acoustic models were utilized in acoustic-wave propagation simulations by mapping tissue-specific and acoustically relevant parameters to the segmented volume. The results and analyses demonstrated differences in ultrasound images formed from 2-D and 3-D simulations derived from the same acoustic models. In addition, simulations were compared using both hard and soft segmentation techniques to form the acoustic models. The soft model produced images with less degradation to image quality compared to the hard model without requiring modification to the acoustic parameters or uniform blurring of the acoustic properties.

Disclosures

Funding and in-kind support for this project was provided to Dr. Dahl by GE Healthcare through the Blue Sky Collaboration with the Department of Radiology at Stanford University. No conflicts of interest, financial or otherwise, are declared by the other authors.

References

1. G. F. Pinton, G. E. Trahey, and J. J. Dahl, “Sources of image degradation in fundamental and harmonic ultrasound imaging: a nonlinear fullwave simulation study,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58**(6), 1272–1283 (2011).
2. L. M. Hinkelman et al., “Modified Gomori trichrome stain for macroscopic tissue slices,” *J. Histotechnol.* **19**(4), 321–323 (1996).
3. T. D. Mast et al., “Simulation of ultrasonic pulse propagation through the abdominal wall,” *J. Acoust. Soc. Am.* **102**(2), 1177–1190 (1997). [Erratum: *J. Acoust. Soc. Am.* **104**, 1124–1128 (1998)].
4. M. Tabei, T. D. Mast, and R. C. Waag, “Simulation of ultrasonic focus aberration and correction through human tissue,” *J. Acoust. Soc. Am.* **113**(2), 1166–1176 (2003).

5. Y.-T. Shen, M. I. Daoud, and J. C. Lacefield, “Computational models of distributed aberration in ultrasound breast imaging,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **57**(12), 2627–2636 (2010).
6. C. Mougenot et al., “High intensity focused ultrasound with large aperture transducers: a MRI based focal point correction for tissue heterogeneity,” *Med. Phys.* **39**(4), 1936–1945 (2012).
7. J. L. Aroyan, “Three-dimensional modeling of hearing in *Delphinus delphis*,” *J. Acoust. Soc. Am.* **110**(6), 3305–3318 (2001).
8. T. D. Mast, “Two- and three-dimensional simulations of ultrasonic propagation through human breast tissue,” *Acoust. Res. Lett. Online* **3**(2), 53–58 (2002).
9. O. M. Al-Bataineh et al., “Utilization of the k-space method in the design of a ferroelectric hyperthermia phased array,” *Ferroelectrics* **331**(1), 103–120 (2006).
10. G. Salahura et al., “Large-scale propagation of ultrasound in a 3-D breast model based on high-resolution MRI data,” *IEEE Trans. Biomed. Eng.* **57**, 1273–1284 (2010).
11. M. Jakovljevic et al., “Blocked elements in 1-D and 2-D arrays-part I: detection and basic compensation on simulated and in vivo targets,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **64**(6), 910–921 (2017).
12. N. Bottenus, G. Pinton, and G. Trahey, “Large coherent apertures: improvements in deep abdominal imaging and fundamental limits imposed by clutter,” in *Proc 2016 IEEE Int. Ultrason. Symp.* (2016).
13. G. F. Pinton, “Reverberation clutter and sources of image degradation in transcostal imaging,” in *Proc 2016 IEEE Int. Ultrason. Symp.* (2016).
14. Y. HongLei et al., “Remote sensing classification using fuzzy C-means clustering with spatial constraints based on Markov random field,” *Eur. J. Remote Sens.* **46**(1), 305–316 (2013).
15. M. N. Ahmed et al., “A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data,” *IEEE Trans. Med. Imaging* **21**, 193–199 (2002).
16. K.-S. Chuang et al., “Fuzzy C-means clustering with spatial information for image segmentation,” *Comput. Med. Imaging Graphics* **30**(1), 9–15 (2006).
17. D. L. Pham and J. L. Prince, “Adaptive fuzzy segmentation of magnetic resonance images,” *IEEE Trans. Med. Imaging* **18**, 737–752 (1999).
18. D. C. Park, “Intuitive fuzzy C-means algorithm for MRI segmentation,” in *2010 Int. Conf. Inf. Sci. Appl.*, pp. 1–7 (2010).
19. S. K. Adhikari et al., “A spatial fuzzy C-means algorithm with application to MRI image segmentation,” in *2015 Eighth Int. Conf. Adv. Pattern Recognit. (ICAPR)*, pp. 1–6 (2015).
20. S. P. Chatzis and T. A. Varvarigou, “A fuzzy clustering approach toward hidden Markov random field models for enhanced spatially constrained image segmentation,” *IEEE Trans. Fuzzy Syst.* **16**, 1351–1361 (2008).
21. M. Gong et al., “Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images,” *IEEE Trans. Fuzzy Syst.* **22**, 98–109 (2014).
22. J. C. Bezdek, R. Ehrlich, and W. Full, “FCM: the fuzzy C-means clustering algorithm,” *Comput. Geosci.* **10**(2), 191–203 (1984).
23. K. Looby et al., “K-means clustering for high-resolution, realistic acoustic maps,” *Proc. SPIE* **10580**, 1058014 (2018).
24. C. M. Sehgal et al., “Measurement and use of acoustic nonlinearity and sound speed to estimate composition of excised livers,” *Ultrasound Med. Biol.* **12**(11), 865–874 (1986).
25. H. Eggers et al., “Dual-echo Dixon imaging with flexible choice of echo times,” *Magn. Res. Med.* **65**(1), 96–107 (2011).
26. T. Zhang et al., “Resolving phase ambiguity in dual-echo Dixon imaging using a projected power method,” *Magn. Res. Med.* **77**(5), 2066–2076 (2017).
27. N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
28. B. E. Treeby and B. T. Cox, “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields,” *J. Biomed. Opt.* **15**(2), 021314 (2010).
29. B. E. Treeby, “Modeling nonlinear wave propagation on nonuniform grids using a mapped k-space pseudospectral method,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **60**(10), 2208–2213 (2013).
30. M. Patterson and F. Foster, “The improvement and quantitative assessment of b-mode images produced by an annular array/cone hybrid,” *Ultrason. Imaging* **5**(3), 195–213 (1983).

31. M. Lediju et al., “Quantitative assessment of the magnitude, impact, and spatial extent of ultrasonic clutter,” *Ultrason. Imaging* **30**(3), 151–168 (2008).
32. P. D. Freiburger et al., “Two dimensional ultrasonic beam distortion in the breast: in vivo measurements and effects,” *Ultrason. Imaging* **14**(4), 398–414 (1992).
33. L. M. Hinkelman, T. L. Szabo, and R. C. Waag, “Measurements of ultrasonic pulse distortion produced by human chest wall,” *J. Acoust. Soc. Am.* **101**(4), 2365–2373 (1997).
34. L. M. Hinkelman et al., “The effect of abdominal wall morphology on ultrasonic pulse distortion. part I. measurements,” *J. Acoust. Soc. Am.* **104**(6), 3635–3649 (1998).
35. A. Grimm et al., “Evaluation of 2-point, 3-point, and 6-point Dixon magnetic resonance imaging with flexible echo timing for muscle fat quantification,” *Eur. J. Radiol.* **103**, 57–64 (2018).

Biographies of the authors are not available.