

# Classification of Abdominal Tissues by $k$ -Means Clustering for 3D Acoustic and Shear-Wave Modeling

Kevin T. Looby  
klooby@stanford.edu

## I. ABSTRACT

Clutter is an effect that degrades the quality of medical ultrasound imaging. To better understand and subsequently reduce the effects of clutter, the underlying mechanisms must be studied. Clutter is thought to be highly reliant on the reverberations of acoustic signal due to microscopic subcutaneous tissues.

This paper presents a method for producing a four-tissue acoustic map to be used in three-dimensional nonlinear ultrasound simulations. A full volumetric map was successfully produced through the application of the  $k$ -means clustering algorithm to a fat-water separated MRI volume.

## II. MOTIVATION

### A. Clutter

In ultrasonic imaging, clutter, a phenomenon associated with poor-quality images, produces a temporally-stable obstruction resulting in decreased image contrast and a reduced-ability to discern imaging targets. Clutter is frequently a factor when imaging overweight, obese, and difficult-to-image patients. The effect of clutter is attributed to the interactions of acoustic waves with the subcutaneous tissue layers. Particularly, acoustic reverberation, off-axis scattering, and phase aberrations from the subcutaneous layers are significant sources of clutter [1] [2].

### B. Tissue Classification for Modeling

The imaging difficulties presented by clutter are a major obstacle for clinical ultrasound imaging. As such, modeling the acoustic wave-tissue interactions that produce clutter is of importance to ultrasound researchers. While much previous work has been done in characterizing this phenomenon using two-dimensional models, the problem must be examined on a three-dimensional scale in order to produce simulation results of significant value. Presently, the literature lacks three-dimensional models of these interactions.

This project, which forms a piece of an ongoing project that I am conducting in Jeremy Dahl's ultrasound research lab, seeks to produce acoustic maps to be used in three-dimensional ultrasound simulations. An acoustic map, in the context of this project, is a

mapping of tissue characteristics to spatial locations. These maps will be produced through the processing of 3D magnetic resonance imaging (MRI) scans.

At this time, the primary aim is to take a set of MRI slices and assign each image pixel a label corresponding to the tissue displayed in that pixel. The primary tissues of interest are skin, fat, muscle, and connective tissues. Any pixels corresponding to regions without signal are classified together as background pixels in a fifth cluster.

## III. MATERIALS

The MRI datasets used in this project were acquired at the Lucas Center at Stanford specifically for this project. In order to achieve high signal-to-noise ratio (SNR) at a very high resolution (100 micron isotropic), prohibitively long scan times are required. As such, a slab of pork belly is used as an ideal ex-vivo model. It is desired that the results achieved on this model be replicated when applied to a non-ideal in-vivo human scan, which will show a significant degradation in SNR.

A tailored MRI pulse was used to generate two datasets. One dataset contains the signals from protons within water-based tissues; the other contains signal from protons within fat-based tissue. Although out of the scope of this project, as brief explanation, this was achieved by making use of the chemical shift seen in the resonant frequencies of shielded protons. This allows for the signals from different chemical species (here, water and fat) to be distinguished based on the phase of the received signal.

For illustrative purposes, a two-dimensional slice of the dataset is provided in Fig. 1 and Fig. 2. Before running the algorithm, some preprocessing was required. First, the datasets were separately normalized to have a minimum voxel intensity of zero and a maximum voxel intensity of one. The datasets were then registered to each other, as it is vital that the voxel correspondences between the two datasets be accurate.

## IV. METHODOLOGY

### A. $k$ -Means Algorithm

The  $k$ -means clustering algorithm is fairly straightforward. A set of centroids is initially chosen, either at

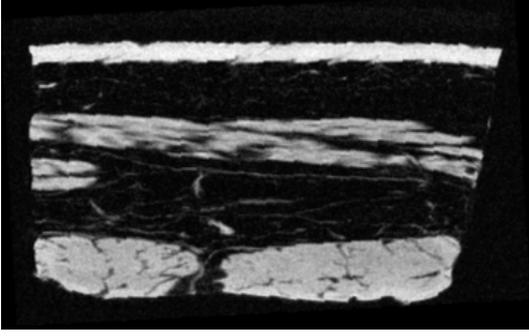


Fig. 1. Selected slice from the water-only dataset. Gray levels closer to white indicate a stronger water signal at that pixel.

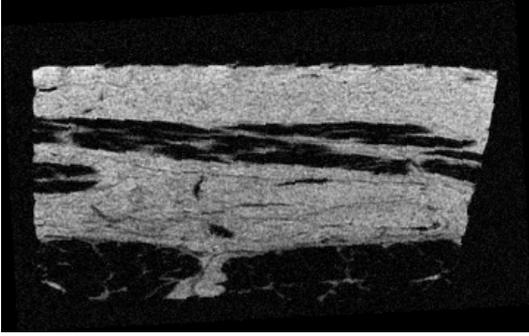


Fig. 2. Selected slice from the fat-only dataset. Gray levels closer to white indicate a stronger fat signal at that pixel.

random or with some initial estimate of the likely final centroids. During each iteration, every point (in this application, pixel) is assigned the label corresponding to the centroid closest in Euclidean distance in the feature space (the feature space will be described in the next section). Once every point has been assigned a label, the locations of the centroids are updated to the means of the locations of all points that were assigned the corresponding label in the previous step. This process is repeated until a convergence criterion has been met.

Mathematically, this can be described (as presented in the course lecture notes) by the following: repeat steps 1 and 2 until convergence is achieved.

$$c^{(i)} := \arg \min_j \left\| x^{(i)} - \mu_j \right\|^2 \quad (1)$$

$$\mu_j := \frac{\sum_{i=1}^m \mathbf{1} \{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^m \mathbf{1} \{c^{(i)} = j\}} \quad (2)$$

In these equations,  $c^{(i)}$  denotes the label assigned to data point  $x^{(i)}$ ,  $\mu_j$  the location of centroid  $j$ , and  $\mathbf{1}\{\}$  is the indicator function yielding 1 when its argument is true. The algorithm is implemented using the built-in function provided with MATLAB for computational efficiency.

## B. Feature Space

In the k-means clustering method, the feature space can be thought of as a space with dimensions defined by the features chosen to describe the data to be categorized. The centroids take on values within this space, and distances within the space can be calculated between the values taken on by the centroids and the values held by the data.

A substantial piece of this project involves devising and properly weighting the features that are used to describe each pixel. The dataset can be considered as a three-dimensional matrix whose dimensions correspond to the three spatial dimensions  $x$ ,  $y$ , and  $z$ . Each entry in the matrix is a scalar representing the pixel intensity at each spatial location. As such, the only information available is a pixel intensity and its location in Cartesian space.

The results that will be presented in the next section were achieved using a weighting of nine features. The features were chosen and weighted in order to best accommodate the clustering of the data into the five desired clusters.

1) *Known Signal*: As the data has already been grouped into fat signal and water signal, it is very desirable to make use of this information. The voxels corresponding to the water-based tissues (skin, muscle, and connective tissue) should be represented by high signal values in the water signal dataset, while the pixels corresponding to the fat signal should of course be represented by high signals in the fat signal dataset. Furthermore, the regions corresponding to a lack of signal should have very low pixel values in both datasets.

This information was included as three binary features. Representative slices of these features are provided in Fig. 3, 4, and 5. As the remainder of the difficulty rests in distinguishing the three water-based tissues, the following features will only be applied to the water signal volume.



Fig. 3. Thresholding the normalized water signal dataset gives a binary mask corresponding to voxels to be associated with water-based tissues.

2) *Connected Regions*: Intuitively, neighboring voxels of similar intensity are likely to belong to the same tissue class. This is a very natural assumption and can be implemented by using a region-labeling



Fig. 4. Thresholding the normalized fat signal dataset gives a binary mask corresponding to voxels to be associated with fat-based tissues.



Fig. 5. The water and fat features were combined to provide a binary mask to encourage the separation of tissue voxels from background noise.

algorithm. The algorithm implemented in this work used the MATLAB function `bwlabel()`. This function takes a binarized volume and outputs a label for each detected connected region.

It was found that in three-dimensional space, the connective tissue fibers link the larger tissue structures to each other, as anticipated by prior knowledge of this anatomy. As such, running the region-labeling algorithm in three dimensions did not provide useful results. However, by labeling regions in two-dimensional space, it is possible to separate the connective tissue regions from the larger skin and muscle masses. A representative slice of this feature is shown in Fig 6.



Fig. 6. Each connected region in the volume was assigned a unique integer label. The labels were then normalized to a maximum value of one.

The sizes of each of these connected regions can also be usefully incorporated into the feature space. As the connective tissue regions (the smallest connected regions) were found to be the most difficult to classify, the region-size feature was chosen to emphasize clustering these small regions together. This was achieved by summing the number of pixels in each region and then thresholding the labeled volume so that only the regions significantly smaller than the mean region size were kept (Fig. 7).



Fig. 7. Thresholding out the largest of the regions leaves a fair estimate of voxel regions expected to relate to the connective tissues class.

3) *Intensity*: Voxel intensity naturally makes up one feature. Alone, this is in no way sufficient for characterization, as the different water-based tissue types produce similar signals (see Fig. 1). Towards the edges of tissue regions, the MRI signal decays, and as such there is a broad overlapping of intensities between the tissue types. These variations were accounted for by assigning the median voxel intensity over each connected region to every voxel within that region.



Fig. 8. The median intensity value over each region in Fig. 6 was assigned to every voxel within that region to promote uniform classification by region and voxel intensity.

4) *Spatial Distribution*: Again, referring to Fig 1, it can be seen that there is a significant macroscopic distribution of tissues in the  $y$ -direction (vertical). It is possible to take advantage of this distribution by using the  $y$ -position of each voxel as a feature.

Experimentation was done with using polynomial weightings of the voxel locations to try to enforce prior knowledge about the expected spatial distributions of

the water-based tissues. However, a simple logarithm of the vertical position (with respect to the orientations of the slices as presented in the figures in this work) was found to effectively enforce the expected spatial distributions of tissues while preserving the ability to generalize to future datasets.

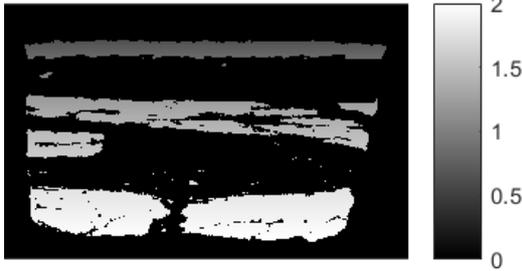


Fig. 9. Taking a weighted logarithm of the vertical voxel locations enforces the prior expectation regarding the spatial distribution of abdominal tissues.

5) *Image Gradients*: Additionally, there are spatial variations in pixel intensity throughout each tissue. As a particular example, the connective tissues (thin tendrils as seen in Fig. 1) display a bright signal of a width on the order of 1-2 pixels. In an attempt to characterize such spatial variation, local gradients in the  $x$ - and  $y$ - directions are used as features. These gradients are calculated with a two-dimensional convolution of the image data with appropriate kernels, defined as:

$$G_x = \text{Im} * \begin{bmatrix} -1 & 1 \end{bmatrix} \quad (3)$$

$$G_y = \text{Im} * \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (4)$$

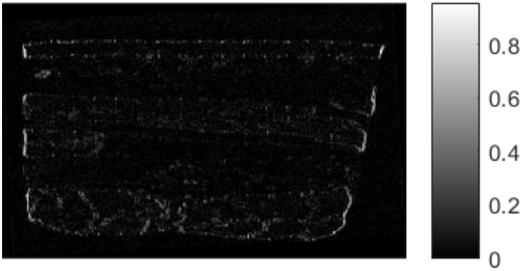


Fig. 10. The convolution described in Eqn. 3 produces a volume with local horizontal gradients represented by voxel intensity.

In equations 3 and 4,  $G_x$  and  $G_y$  are the local gradients in the  $x$ - and  $y$ - directions respectively, and  $*$  is the convolution operator.

6) *Feature Space Weighting*: The complete feature space can be represented as in Eqn. 5.  $S_{\text{total}}$  is the complete signal mask,  $S_{\text{water}}$  and  $S_{\text{fat}}$  the water and fat signal masks,  $R_{\text{label}}$  the connected region labels,

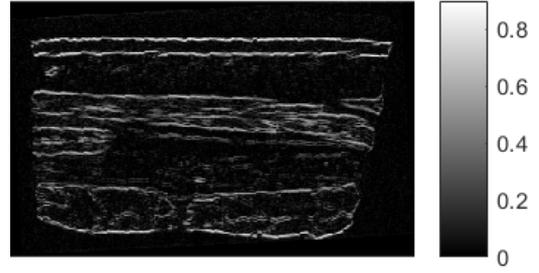


Fig. 11. The convolution described in Eqn. 4 produces a volume with local vertical gradients represented by voxel intensity.

and  $R_{\text{size}}$  the sizes of the connected regions.  $I_i$  denotes the intensity of the  $i$ th pixel,  $y_i$  denotes the normalized and weighted  $y$ -position (vertical) of the  $i$ th pixel, and, as before,  $G_{x,i}$  and  $G_{y,i}$  are the local gradients calculated for each pixel.

$$x^{(i)} = \begin{bmatrix} S_{\text{total}} \\ S_{\text{water}} \\ S_{\text{fat}} \\ R_{\text{label}} \\ 10R_{\text{size}} \\ I_i \\ 2y_i \\ G_{x,i} \\ G_{y,i} \end{bmatrix} \quad (5)$$

Multiplicative weightings were added to the region size and  $y$ -location features in order to encourage the algorithm to emphasize these features over the rest, as these were found to produce better results when weighted.

## V. RESULTS

Clustering was run on the water-only and fat-only volumes (Fig. 1 and Fig. 2, respectively) using the features defined in the previous section. By running  $k$ -means on this feature set, a complete acoustic map is obtained (Fig. 12). For the sake of providing a visualization of the data, each of the images displayed throughout this work are of the same two-dimensional slice of the larger three-dimensional volume. The clustering was successfully applied to the entire volume.

## VI. DISCUSSION AND CONCLUSIONS

Overall, excellent qualitative results were achieved. As is to be expected, some misclassifications must be accepted, as any MRI acquired at this resolution will have low SNR and contrast. The feature space was specifically defined with this in mind in order to produce benign misclassifications whenever possible.

In medical ultrasound, images are constructed based on acoustic reflections. Such reflections occur when a propagating acoustic wave encounters an acoustic impedance mismatch. Larger mismatches produce

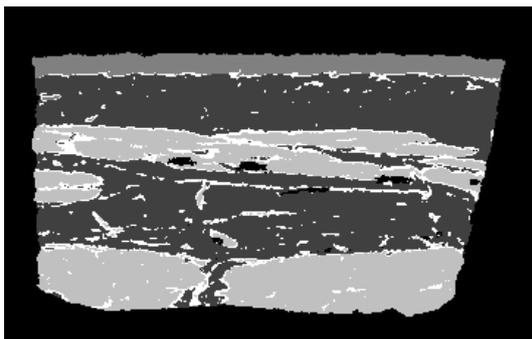


Fig. 12. The  $k$ -means algorithm produces an integer-labeled mapping of the dataset. A 2D slice of the labeled tissues with each tissue class represented by a different grayscale value is provided. In order of increasing "whiteness," the classes are background (black), fat, muscle, skin, and connective tissue (white).

larger reflections. This aspect of ultrasound physics was used to guide the definition of "benign" and "malignant" misclassifications.

A benign misclassification was defined to be a classification of either skin or muscle voxels into the connective tissue class. As long as kept to small numbers, these misclassifications will have negligible effect on simulation results.

There are two types of misclassifications defined as malignant for this application. The first was defined to be any deep classification of a fat voxel as water-based tissue or water-based voxel as fat. Here, I use deep in the anatomical sense (that is, an example of this misclassification would be an erroneous voxel in the middle of a piece of muscle labeled as fat). This type of error was heavily influenced by the Rician noise inherent in MRI scan data and was nearly eliminated by the addition of the connected regions feature (Fig 6).

The second malignant misclassification type is the classification of any tissue voxel as "background" or "empty space." Such erroneous voxels would produce extremely large imaging artifacts, the impedance mismatch between any tissue and empty space is large. The preliminary results presented in the Milestone suffered greatly from these misclassifications.

Much of the subsequent work involved incorporating features and preprocessing steps to eliminate these errors. The inclusion of a more sophisticated image registration transform to ensure alignment between the fat and water volumes along with inclusion of the connected region and binary signal features allowed for the almost complete removal of these misclassifications.

## VII. FUTURE WORK

The results presented in Fig. 12, are more than satisfactory to serve as a dataset for preliminary *ex vivo* simulations. As mentioned, it is desired to translate this process to usage with *in vivo* scan data, which will suffer a further reduction in SNR and image contrast.

Investigations into the robustness of this algorithm to the presence of these less ideal conditions will need to be performed.

## REFERENCES

- [1] J. J. Dahl and N. M. Sheth, "Reverberation clutter from subcutaneous tissue layers: Simulation and in vivo demonstrations," *Ultrasound Med Biol*, 2014.
- [2] 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 Ferroelect Freq Contr*, vol. 26, pp. 1272–1283, 2011.