Optimization of Hidden Markov Model for Density Estimation in Breast Imaging

by

Sylvia Hon

Department of Biomedical Engineering
Duke University

Date: _______________________

Approved:

____________________________
Joseph Y. Lo, Supervisor

____________________________
Joseph A. Izatt

____________________________
Gregg E. Trahey

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Biomedical Engineering in the Graduate School of Duke University

2012
ABSTRACT

Optimization of Hidden Markov Model for Density Estimation in Breast Imaging

by

Sylvia Hon

Department of Biomedical Engineering

Duke University

Date: _______________________

Approved:

____________________________
Joseph Y. Lo, Supervisor

____________________________
Joseph A. Izatt

____________________________
Gregg E. Trahey

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Biomedical Engineering in the Graduate School of Duke University

2012
Abstract

Breast cancer is the second leading cause of cancer-related death in women in the United States; it is fundamentally important to detect risks as accurately and as early as possible for prevention. As breast density often correlates with risk factors, 3D Hidden Markov Model method is one quantitative approach to measuring density in breast tomosynthesis images. The purpose of developing this method was to overcome the poor depth resolution in tomosynthesis images and to get a more accurate density estimation of a woman’s breast through 3D imaging.

Previous study showed that 3D Hidden Markov Model was able to accurately measure breast density for tomosynthesis images that matched closely with the defined ground truth. However qualitatively looking at the segmentation, there were noise and voxels segmented as glandular or adipose that did not reflect the breast anatomy. In addition, the model training process was very time consuming and required many patient cases. To address these limitations, this research was conducted to optimize three of the key model parameters, namely the intensity rescaling of the images, the spatial relation of the neighbors to the current voxel, and the number of training cases. The HMM model was trained using 100 breast MRI volume then validated on 12 breast tomosynthesis volumes.

The HMM model was simplified by cutting the number of grayscale intensity levels in half, while sampling distance between neighboring voxels was doubled and tripled. In spite of these changes, however, there was no significant change to the breast density estimation.
density estimation of the tomo images. We identified new challenges in matching
grayscale histogram intensities between breast MRI and the two breast tomosynthesis
prototype models, and we gained understanding of how the neighbor’s distance brought
in different segmentation characteristics.
# Contents

Abstract .............................................................................................................................. iv  

Contents ............................................................................................................................. vi  

List of Tables .................................................................................................................... vii  

List of Figures .................................................................................................................. viii  

Acknowledgement ............................................................................................................. xi  

Chapter 1. Introduction ....................................................................................................... 1  

1.1  Breast Cancer and Current Imaging Solutions..................................................... 1  
1.2  Breast Density as Risk Biomarker ................................................................. 3  
1.3  3D Hidden Markov Model for Breast Density Quantitation from Tomo Images 5  
1.4  Summary of Study Aims .............................................................................. 6  

Chapter 2. Methodology ..................................................................................................... 8  

2.1 Data collection and processing................................................................................. 8  
2.2 Overview of using 3D Hidden Markov Model for breast density estimation in  
tomosynthesis images.................................................................................................... 14  
2.3 Optimization Parameter 1: Rescaling............................................................ 21  
2.4 Results of Optimization Parameter 1: Rescaling ............................................ 26  
2.5 Size of Training Set for a Robust Probabilistic Model ...................................... 30  
2.6 Optimization Parameter 2: Distance of Neighbors ........................................ 33  
2.7 Results of Optimization Parameter 2: Distance of Neighbors ....................... 35  

Chapter 3. Optimization of HMM method to test Tomo Images................................. 41  

Chapter 4. Discussion ....................................................................................................... 53  

References......................................................................................................................... 56
List of Tables

Table 1. Intensity Range Information for MRI and Tomo Image ........................................ 24

Table 2. Relative density and voxel difference between HMM segmentation and
threshold-applied segmentation for 5 MR cases ............................................................... 28

Table 3. Relative density and voxel difference between HMM segmentation and
threshold-applied segmentation for 5 MR cases to compare the differences in the
neighbors’ distances to current voxel ............................................................................... 35

Table 4. Summary of comparison between ground truth MRI density and tomo density
under different segmentation schemes: mean and standard deviations from difference
(HMM density – MRI density) distributions .................................................................... 46
List of Figures

Figure 1. A central slice representative of most tomo images taken from the original tomo image (left) and its HMM segmentation (right) ................................................................. 6

Figure 2. Original MRI (left) with its threshold-applied segmentation (right) displaying glandular (white=2), adipose (gray=1), and background (black=0) .................................. 9

Figure 3. Average Histograms of MRIs (left) and tomo images (right) after linearly rescaling image down to 300 intensity levels with clip offs at the minimum and maximum for comparison .................................................................................................................. 10

Figure 4. Preprocessed MR case (left) and its histogram-equalized version with 0.5% clip offs at both ends (right) with their respective histogram, both rescaled to 300 intensity levels ........................................................................................................................................ 12

Figure 5. Preprocessed tomo case (left) and its thickness equalized version with 1% clip offs on both ends(right) with their respective histogram, both rescaled to 300 intensity levels ........................................................................................................................................ 13

Figure 6. Diagram of the current voxel and arrangement of its neighbors ...................... 15

Figure 7. Example of the matrix for counting pairings of intensity (rows) vs. neighbor states (columns) .......................................................................................................................... 17

Figure 8. Example of a completed PTM [37] .................................................................. 19

Figure 9. PTM (top) trained on 100 breast MRIs that were rescaled to 300 intensity levels and zoomed in (bottom) to see the fluctuations between intensity levels................. 22

Figure 10. Example of rescaling a MRI that leads to gaps in the histogram ..................... 25

Figure 11. PTMs trained on 20 breast MRIs rescaled to 300 intensity levels, histogram equalized with 10 bins (top), rescaled to 150 intensity levels, histogram equalized with 10 bins (bottom left) and with 15 bins (bottom right) ............................................................... 27
Figure 12. HMM segmentations done by PTM trained on 10 MRIs: threshold-applied (top left), rescaled to 300 intensity levels, histogram equalized with 10 bins (top right); rescaled to 150 intensity levels, histogram equalized with 10 bins (bottom left) or 15 bins (bottom right).................................................................................................................... 29

Figure 13. Diagram showing how the increase in training size decreases change in average absolute density difference results when compared to ground truth ....................... 31

Figure 14. PTM trained from (left to right, top to bottom) 20, 40, 60, 80, 100 breast MRIs ........................................................................................................................................... 32

Figure 15. Alternative neighbor arrangements with respect to the current voxel: two voxels away (left) and three voxels away (right) for in-plane and in-depth ..................... 35

Figure 16. MRI of 21% dense breast: Threshold-segmented, HMM with neighbors one voxel away (top right), two voxels away (bottom left), three voxels away (bottom right); with zoomed in boxes in the same order to evaluate specific regions that match their color borders............................................................................................................................... 36

Figure 17. MRI of 38% dense breast: Threshold-segmented, HMM with neighbors one voxel away (top right), two voxels away (bottom left), three voxels away (bottom right); with zoomed in boxes in the same order to evaluate specific regions that match their color borders............................................................................................................................... 38

Figure 18. MRI of 51% dense breast: Threshold-segmented, HMM with neighbors one voxel away (top right), two voxels away (bottom left), three voxels away (bottom right); with zoomed in boxes in the same order to evaluate specific regions that match their color borders............................................................................................................................... 39

Figure 19. Alternative neighbor arrangements with respect to the current voxel: two voxels away (left) and three voxels away (right) in all three directions ......................... 42

Figure 20. (top to bottom) Transition of PTM from rescaling from 300 to 150 intensity levels, model based on direct neighbors, neighbors two voxels away in x and y directions, and neighbors two voxels away in x, y, and z directions.................................................. 45
Figure 21. (left to right) Original tomo image with its HMM segmentations looking at neighbors (with respect to current voxel) 1) one voxel away 2) two voxels away in x and y directions 3) three voxels away in x and y directions 4) two voxels away in all three directions and 5) three voxels away in all three direction .................................................. 48

Figure 22. Inspiration tomo case showing a slice with breast-air segmentation artifact (shown by red arrow) not completely removed ................................................................. 50

Figure 23. Histogram comparison of a Novation case (left) and Inspiration case (right) after preprocessing. Histogram comparison of a Novation case and Inspiration case after preprocessing and clipping off 1% of the low and high values to minimum and maximum of 1 and 300 ........................................................................................................ 50

Figure 24. To understand the intensity range that the breast-air artifact encompasses, two slices of a tomo images are shown in 2 columns with all the values below a set threshold value as blue: (top) a lower threshold value set showing the artifact (pointed by red arrow) and (bottom) a higher threshold value set showing the breast values getting clipped off (pointed by green arrow) ........................................................................ 51
Acknowledgement

I would like to thank Dr. Joseph Lo for his patience and endless encouragements throughout my journey in exploring the medical imaging world. Considerably my most respected advisor to date, he never fails to support me in whatever I decide to do and has enlightened me in many aspects of life. His humor has always brightened up a gloomy day.

To my committee members Dr. Trahey and Dr. Izatt for taking time out of their busy research to listen to my defense and to sign off an important chapter of my life.

Without Christy Shafer, I probably would have been very lost in this project and be buried in her MATLAB codes. Thank you Christy for getting me up to speed in this project and teaching me all that you can before you left!

I would like to thank Hyunkoo and Stephanie for keeping me company during the weekly meetings as well as listening to all my troubleshooting panics.

To my family and friends who do not really know what my research is about but still gave endless encouragement and inspiration day in and day, thank you.

Sincere thanks to Yuan who listened intently and patiently critiqued my defense presentation slice by slice.

And most importantly, to my first Portland friend, roommate, baking buddy, and classmate who took 8 out 10 graduate classes with me, Tina Chen, I am sincerely grateful for everything you have done for me. You have taught me so much about image processing and with such patience!
Chapter 1. Introduction

1.1 Breast Cancer and Current Imaging Solutions

Breast cancer is the second leading cause of cancer-related death in women in the United States[1]; hence, it is important to detect risks as accurately and as early as possible for prevention purposes. The current standard imaging modality screening for breast cancer risks in women, mammography [2], provides high contrast and high spatial resolution while requiring low patient dose. However, mammography can only provide radiologists a 2D view of the compressed three-dimensional breasts for qualitative examination. Due to the overlapping of fat and glandular tissue in one flat perspective, it is difficult to discern subtle changes against the complex background [3]. The lack of specificity and sensitivity leads to recall rates of 15% in the U.S.[4]. This can be problematic especially for women whose breasts are extremely dense [5, 6] because the probability of lesions blocked by overlapping tissue will be much higher. To improve diagnostic capabilities, there is a need for 3D breast imaging that allows cross-sectional visualization.

As of now, the modality that exhibits the highest sensitivity of detecting primary and recurrent breast cancer is breast MR. Breast MR utilize the difference in spin-lattice relaxation time (T1) and spin-spin relaxation time (T2) with fat suppression to precisely differentiate fat and adipose tissue in image volume as well as to differentiate scar from recurrent cancer [7-9]. Unfortunately, the modality is not widely available in the clinical
setting and is much more expensive than a mammographic scan. Also, due to the lack of standardization in the technique and interpretation of the images, there has been difficulty in introducing breast MR for routine clinical breast imaging. According to Kuhl et. al’s review, from the plethora of information that can be gathered from MR scans, there are seven different criteria that can be used to classify breast cancer [10].

X-ray projection based volume imaging has been shown to be the most effective in early detection of breast cancer in terms of spatial resolution, low patient dose, and its established presence in the clinical setting. Digital breast tomosynthesis, commonly known as tomo, is a form of limited-angle cone beam CT [11, 12]. Instead of scanning the breast uncompressed in a full 360° for reconstruction of 3D volume [13-16], tomo acquires x-ray projection images of the compressed breast like in conventional full-field digital mammography (FFDM) over a limited number of angles across an arc. With projections acquired over a limited number of angles, the reconstruction of the 3D volume exhibits loss of information due to angular under-sampling.

There are mixed reviews from previous studies on whether tomo improves specificity and sensitivity [17-19]: Hologic found that specificity improves without an increase in detection rate when FFDM was done with two-view tomo, and Gennaro et al. and Teertstra et al. observed no change in specificity or sensitivity when comparing tomo to mammography. However, studies have shown that the availability of cross-sectional information decreases the number or false positive recalls: there was a recall reduction of false positives by 11% when using 2-view tomo [20], and 30% and 40% when using tomo with FFDM compared to using FFDM alone [21, 22]. Also, 2-view tomo
outperforms FFDM in terms of diagnostic accuracy for readers with less than 10 years of experience in the radiology field [4]. Other advantages of using tomo are its similarity to operating FFDM that allows less retraining from hospital staff [23] and its dose comparable to 2-view mammography [24, 25]. The compression of the breast also reduces motion artifacts and ensures proper tissue coverage. Given the many benefits, tomo can become an available and affordable 3D imaging modality for clinical routine breast imaging.

1.2 Breast Density as Risk Biomarker

Breast density is defined as the amount of fibroglandular tissue over the total breast volume. It is often correlated with cancer risks [26-29] because if a woman has high breast density obscuring suspicious findings, she is less likely to receive an accurate diagnosis. Women having breast densities higher than 75% have 4-6 times more risks than women with 0% breast density [30]. Also, since cancerous cells usually form along the epithelial lining of the milk ducts, glandular tissue, having higher density means a woman has more chances of developing cancer in the breast. Relative risk associated with breast density is larger than relative risks associated with familial, menstrual, and reproductive risk factors [31].

For decades, breast density was assessed subjectively by radiologists using subjective criteria such as Wolfe’s parenchymal patterns, Tablar classification, and the modern BI-RADS (Breast Imaging Reporting and Data System) density ratings [32]. There are
software that aid in measuring breast density for mammography that require user’s interactive and manual input for segmentation of the glandular and adipose tissue [30]. All these methods can give varying results depending on the doctor’s or user’s level of expertise, so an automated method is needed to limit these measurement inconsistencies. Commercial products now exist for fully automated measurements of breast density [33]. Quantitative measurements of breast density are necessary for studying epidemiology, predicting an individual’s breast cancer risks, and for monitoring efficacy of various therapies such as hormone replacement or chemoprevention [32].

It has been shown that with FFDM, the breast density is often overestimated [32] due to the 2D projection nature of the image. Breast density measurement is more accurate using a 3D imaging modality such as breast MR, but breast MR is not widely available and is expensive to use. Therefore, a mammography-based technique such as tomo will be the ideal imaging modality to use for more accurate density measurements. However, due to the angular under-sampling, geometric variations and misalignments from projection to projection [34], and the ramp filter in the reconstruction of filtered back projection [35], the spatial frequency in the depth direction is limited. To overcome this limitation, our lab developed a method to segment the glandular tissue from the adipose tissue to quantitatively measure breast density in tomo images.
1.3 3D Hidden Markov Model for Breast Density Quantitation from Tomo Images

To make quantitative measurement of breast density possible in tomosynthesis images, our lab previously developed a method using 3D hidden Markov model (HMM) to distinguish glandular from adipose tissue, the two main types of tissues within the breast, in 3D clinical tomo images. The assumption made was that each voxel of the breast volume was either 100% adipose or glandular [36, 37].

This approach built a statistical model off of information from volumes of breast MRIs where adipose and glandular tissues are easily differentiable. It paired the voxel’s probability of being glandular to the occurrence frequency of specific combination of its intensity level and its surrounding neighbors (glandular or adipose). For example, looking at the breast anatomy and understanding that glandular voxels had higher intensity values, if a voxel’s intensity level was low and/or was surrounded by all adipose voxels, the likelihood of the voxel being adipose would be high. The established probabilistic model was then used to segment the breasts in tomo volumes into adipose and glandular for breast density measurement.

Since there was no gold standard for human subject data, our lab utilized the density results from corresponding MR cases within 1.5 years of tomo scan dates as ground truth and evaluated the performance of the HMM method. The study compared the tomo breast densities achieved by HMM and by two other simpler threshold segmentation methods and showed that HMM outperformed the threshold segmentation methods [36, 37]. Although the HMM method captured majority of the glandular tissue within the breast in tomo images, evaluating the segmentations qualitatively showed small groups of
glandular tissue spread throughout the breast that did not reflect true anatomy as shown in Figure 1. Despite having density results that correlated closely with the MRIs, this limitation should not be overlooked. Moreover, the HMM method required a larger number of cases to train a very complex model. There is a need for simpler approaches that may be easier to train or generalize better.

Figure 1. A central slice representative of most tomo images taken from the original tomo image (left) and its HMM segmentation (right)

1.4 Summary of Study Aims

In this study, our goal is to optimize the probabilistic model in the 3D HMM method so that the segmentation process is less sensitive to image noise and reflects actual glandular anatomy in tomo images. The three aspects of the model that we will explore
are the intensity level distributions in breast MRIs and tomo images, the distance of the neighbors, and the number of MRIs used for training.
Chapter 2. Methodology

2.1 Data collection and processing

In previous work from our lab, we established the feasibility of a cross-modality scheme to develop a 3D breast density estimation model. Breast MRIs were used to train the model because they provide excellent differentiation between adipose and glandular tissue. The model was then evaluated on the smaller number of breast tomo scans available, with ground truth provided by breast MRI from those same subjects. 100 breast MRIs from 54 patients were used for training the probabilistic model to test on 12 breast tomo images from seven patients. There were 22 breast tomo images available, but 10 cases were omitted due to artifacts from bad initial segmentation. Note that each breast MRI or tomo image consists of either the right or left single breast. The MRI data were captured using the following devices with dedicated breast coil: 3.0T Siemens Magnetom Trio, 1.5T Siemens Magnetom Avanto, GE Signa HDxt 3.0T, and GE Signa HDxt 1.5T Optima. The MRIs were T1-weighted and fat suppressed with a resolution of 1-2mm slice thickness of 448mm x 448mm or 512mm x 512mm matrix. The 12 used tomo volumes were captured using the Siemens Mammmomat Novation\(^{\text{TOMO}}\) prototype, and 10 unused were captured using Siemens Mammmomat Inspiration\(^{\text{TOMO}}\). 25 projections were first acquired over a 45° angular span (24cm x 30cm amorphous selenium detector, 0.085mm pixels) and filtered back projection with Siemens TomoEngine version 9 software reconstructed the images into volume of 1mm thick slices each with 0.085mm x 0.085mm pixels [37].
In order to obtain adipose and glandular information from MRIs for the model, the breasts in MRIs were segmented from the background, isolated from the chest wall, and had their skin (5 voxels from breast-air boundary toward chest) removed. The background, chest, and skin were all redefined with a background value of 0. Because the histograms of MRIs often exhibited two modes to clearly distinguish adipose and glandular values, a simple iterative bimodal threshold (equivalent to K-means clustering with K=2) was used to segment the breast voxels into glandular and adipose values of 2 and 1 as seen in Figure 2. Tomo images used for testing were similarly pre-processed: the breast region was segmented from the background, and the skin and associated artifacts were removed (5 voxels at the nipple increasing to 20 voxels at maximum radial distance from the nipple). The skin for both type of image data had to be removed due to its brightness; the adipose-glandular segmentation process would have mistaken the skin as glandular and overestimated the density.

Figure 2. Original MRI (left) with its threshold-applied segmentation (right) displaying glandular (white=2), adipose (gray=1), and background (black=0)
With the breast segmented from the background and skin removed, the average gray value histograms for both image data types can be seen in Figure 3. The mode for MRIs is mainly shifted to the left, whereas for tomo images, it is shifted to the right on the same grayscale range. Since the HMM method utilizes voxel’s intensity information from MRIs to test on tomo images, several image processing steps were taken to match the MRIs closer to the tomo images.

![Average Histogram of 100-breast MRIs](image1.png) ![Average Histogram of 22-breast TOMOs](image2.png)

**Figure 3. Average Histograms of MRIs (left) and tomo images (right) after linearly rescaling image down to 300 intensity levels with clip offs at the minimum and maximum for comparison**

It was impossible to match each individual MRI histogram to the tomo image histograms, so the MRI histograms were made more general through histogram equalization with 10 bins to represent the typical appearance range of tomo images. Prior to histogram equalization, the dynamic range of the image was restricted to the main part of the histogram. Specifically, highest and lowest 0.5% of the original intensity values within the breast were set to the minimum and maximum grayscale of 1 and 300, and the rest were linearly rescaled from 2 to 299. For the tomo images, the volume was
downsampled discretely by a factor of 5 in both directions of the transverse slices to match the MRI resolution and to save memory. The depth resolution was left alone as it was already limited due to the compression during the scanning process. Then a thickness equalization scheme was used to correct for the uneven signal detection across the breast since the inner part of the breast was compressed by the paddle but the periphery of the breast was not in direct contact to the paddle. This equalization method increased the brightness of the voxels along the breast’s periphery to be consistent with the brightness in the inner part of the breast. Since tomo images had a larger intensity range due to the longer tails at both ends of the distribution, instead of 0.5% cutoff, the highest and lowest 1% of the intensity values within the breast were set to the minimum and maximum grayscale of 1 and 300, and the rest were linearly rescaled from 2 to 299. The changes in the respective image of one case of MRI and tomo can be seen in Figure 4 and Figure 5, where the mode of the distribution for both the processed MRI and tomo are more centralized leaning slightly to the right after processing.
Figure 4. Preprocessed MR case (left) and its histogram-equalized version with 0.5\% clip offs at both ends (right) with their respective histogram, both rescaled to 300 intensity levels
Figure 5. Preprocessed tomo case (left) and its thickness equalized version with 1% clip offs on both ends (right) with their respective histogram, both rescaled to 300 intensity levels
2.2 Overview of using 3D Hidden Markov Model for breast density estimation in tomosynthesis images

The basis behind the usage of 3D Hidden Markov Model (HMM) as a volumetric approach to quantitation was that voxel values are dependent on the surrounding voxel values, particularly due to the inter-slice blur caused by the limited acquisition angles in tomosynthesis imaging. Within the breast anatomy, similar tissues, mainly glandular and adipose, naturally connect together in contiguous structures. Briefly, Markov model consists of a chain of events where the next state depends on the current state and not on the sequence of events that preceded it, whereas, in hidden Markov Model, the state is not visible, but the output that is dependent on the state is visible [38]. In the case of applying HMM to tomo images where we are uncertain of the voxel’s tissue type, the hidden states are either glandular or adipose producing a certain known voxel value output. Hence, from training MRIs where adipose and glandular voxels are easily differentiated and depicted, a model is established to determine a voxel’s probability of being glandular or adipose tissue based on its voxel value and surrounding voxel tissue type.

Now with the knowledge of the processed image data and general understanding of how HMM works, details of 1) training MRIs to establish the probabilistic model and 2) using the probabilistic model to segment tomo images will be shown below:
**Training on MRIs**

The probability of a voxel being “glandular” (G) or “adipose” (A) tissue will depend on the current voxel’s intensity value and the state (glandular or adipose) of its seven neighbors in the arrangement shown below (3 in-plane and 4 in-depth) in Figure 6. The neighbors arranged in all directions from the current voxel constitute the three-dimensional structure of the model.

![Diagram of the current voxel and arrangement of its neighbors](image)

**Figure 6. Diagram of the current voxel and arrangement of its neighbors**

1. Due to more noise near the top and bottom of the breasts in MRIs, the range of usable slices is manually selected for each individual MRI case for training. Traversing through each line in the MRI:
   a. Read the intensity value (1 to 300) from current voxel of the rescaled image:
b. Note the pairing of the states of the current voxel’s neighbor (with a binary representation syntax of A for adipose and G for glandular, such as AAAAAAAA if all neighbors are adipose) versus the previously established ground truth tissue type of the current voxel:

Seeing that there are seven neighbors and voxels have a possibility of being one of the two states (A or G), the total possible number of neighbor arrangement is $2^7 = 128$. 
2. After going through each voxel within the breast, the number of occurrence and
the number of glandular occurrence for each pairing of voxel value and neighbor
states are recorded in separate matrices as shown in Figure 7.

![Figure 7. Example of the matrix for counting pairings of intensity (rows) vs.
neighbor states (columns)](image)

3. With the total occurrence matrix and glandular occurrence matrix filled in after
training through all the MRIs, the probability for each entry pairing of the voxel
value and neighbor states can be calculated by:

\[
P(Glandular) = \frac{\sum_{i=1}^{N} GlandularOccurrence}{\sum_{i=1}^{N} TotalOccurrence}, N = NumberofMRIs
\]

As a result, a probability transition matrix (PTM) is formed. For example, after
traversing through all the breast voxels through \( N \) amount of the MRIs, if there
are 10 occurrences for the 300-GGGGGGG entry where 9 occurrences are
glandular, the probability of a voxel with intensity value of 300 and surrounded
by glandular neighbors will be $9/10 = 90\%$.

4. To create a PTM of sufficient confidence, the entries that have less than two
occurrences are disregarded and later filled in by using linear discriminant
analysis (LDA) model. LDA trains on the PTM entries with two or more
occurrences and classifies the empty entries into one of the corresponding group
in the training based on voxel value and neighbor states. Prior to applying LDA,
the entries with probabilities are rounded off to the nearest whole percent to
minimize the number of training groups. Afterward, the empty entries are filled in
with a probability estimated from the LDA. An example of a complete PTM is
shown below:
Figure 8. Example of a completed PTM [37]

This PTM has probabilities that range from 0 to 1, with high values (red color) indicating a high probability for the voxel to be glandular. The intensity level ranges from 1 to 300 for this PTM. The neighbor states are relabeled with numbers from 1 to 128 where 1 is all adipose (AAAAAAA) and 128 is all glandular neighbors (GGGGGGG). As expected, the probability of a voxel being glandular increases for higher intensity levels (lower rows) as well as great numbers of glandular neighbors (right columns). More detail will be presented in subsequent sections focusing on the optimization of these PTMs.
Testing HMM on Tomo Images

With the PTM established, it can be now used to determine the state of the breast voxel within tomo images knowing the voxel value and the states of the voxel’s neighbors. However, since none of voxel’s states has been established yet to be used to retrieve probability from the PTM, the tomo volume was initialized to be completely adipose. The reasoning behind this initialization is because as the algorithm traverse through each row of the tomo volume, it touches base with the periphery of the skin-removed breast first. By looking at the anatomy of the breast, the periphery is usually padded with adipose tissue. As the algorithm continues on, it will begin picking up glandular voxels due to the high intensity value. Due to how the neighbors’ states are needed for testing and how the neighbors are arranged around the current voxel, 3 walls of voxels around the tomo volume’s edge during initialization will not be traversed through and are redefined to background values upon completion of testing.

1. Traversing through each row of the tomo image:

   a) For non-background voxels with non-background neighbors, the voxel value and its neighbors’ states are noted. The voxel is then assigned a probability from the established PTM from training on MRIs based on the voxel value and neighbors’ states. Then the state (glandular or adipose) of the current voxel is drawn from a Bernoulli distribution with probability of success.

   b) For non-background voxels with background neighbors, the state of the voxel remains its initialized state, adipose.
c) Background voxels remain background.

2. With all the voxels’ states within the breast determined, the density of the breast is calculated:

\[
BreastDensity = \frac{\sum GlandularVoxels}{\sum TotalBreastVoxels}
\]

The following sections will cover the optimization done by this project to improve upon the established 3D HMM method. Optimization of the PTM was done using three key parameters: intensity level range, size of training set, and distance of neighbors from current voxel.

### 2.3 Optimization Parameter 1: Rescaling

The PTM shown in Figure 8 was established previously by our lab. The PTM was trained on MRIs rescaled to 300 intensity levels. 300 intensity levels and 7 neighbors gave the PTM 38,400 entries. With more entries to fill in, more MRIs may be required for the training set to fill in the entries. On the other hand, by decreasing the number of grayscale levels, each PTM entry may have more occurrences making each entry, ultimately all entries, more robust. Also, having a smaller PTM size may expedite the training and testing process.
Figure 9. PTM (top) trained on 100 breast MRIs that were rescaled to 300 intensity levels and zoomed in (bottom) to see the fluctuations between intensity levels.

The PTM was reproduced to the best of our ability based on explained methods to establish a starting point for future comparison. The PTM is shown in Figure 9. Looking
at the PTM in Figure 9, the trend shows that the probability for a voxel to be glandular becomes higher when the voxel value is high and when the voxel is surrounded by more glandular neighbors. On the other hand, when the voxel’s intensity value is low and its neighboring voxels are mostly adipose, then the probability of being glandular is low. However, regardless of the states of the neighbors, it seems that if the voxel value is above 275, the voxel’s probability of being glandular is very high, and if the voxel value is below 80, the probability is very low. Beyond the general trend from the top left corner to the bottom right corner, the PTM itself is not very smooth. When taking a closer look as shown in the zoomed in box in Figure 9, there are big changes between two neighboring intensity levels with the same neighbors. For example, around the 105th neighbor states arrangement, amongst the column of red voxels (P~1.00), there are red-orange voxels (P~0.83) that skip almost every few intensity levels. As this could have introduced noise into the segmentations of the tomo images since the determination of the current voxel relied on the neighbors (established states for previous voxels), the intensity distributions of the MRIs and tomo images were studied more in-depth.

Previously, the intensity range for each MRI and tomo image was rescaled from 1 to 300 to establish uniformity working with different images from two different imaging modalities as well as to capture most of the MR images’ bit-depth information. To completely deem whether rescaling down to 300 intensity levels was the best choice, histograms of all the available MRIs and tomo images were examined with intensity range information displayed in Table 1.
Table 1. Intensity Range Information for MRI and Tomo Image

<table>
<thead>
<tr>
<th>Type of Imaging</th>
<th>Smallest Intensity Range</th>
<th>Highest Intensity Range</th>
<th>Average Intensity Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>169</td>
<td>8196</td>
<td>1906</td>
</tr>
<tr>
<td>Tomo Image</td>
<td>1641</td>
<td>4095</td>
<td>3447</td>
</tr>
</tbody>
</table>

The intensity range of values within the 100 breasts over the 54 MR cases, prior to rescaling and histogram equalizing, varies by an order of magnitude from 169 to 8,196. Shrinking the histogram by rescaling intensity values down can lead to posterization or loss of information especially for images with details in many shades. On the other hand, rescaling up to cover a larger gray scale range leads to gaps within the histogram because the histogram is forced to spread the discrete intensity levels over a broader range which can be seen in Figure 10. Since the HMM depends on the voxel value, the latter situation of histogram manipulation should be avoided if possible. Seeing the smallest MRI intensity range from Table 1 was 169 and taking note of the noise that existed between every other or every few intensity levels in the PTM, rescaling the images down to 150 gray scale levels instead of 300 was attempted.
Figure 10. Example of rescaling a MRI that leads to gaps in the histogram
To evaluate how rescaling the image to 150 intensity levels could affect the PTM and ultimately the density estimation of the breast, the PTM was trained on 20 breast MRIs that were linearly rescaled from 1 to 150 after clipping off 0.5% of the lower and higher intensity values to the minimum and maximum. In addition, to keep the intensity distribution of MRIs similar to the distribution of a tomo image after histogram shrinking, a side study was conducted to observe whether histogram equalizing with 15 bins instead of 10 bins would make a difference. The newly made PTMs were used to test five separate MR cases (10 breasts) for comparison purposes.

2.4 Results of Optimization Parameter 1: Rescaling

Compared to the PTM trained on MRIs rescaled to 300 intensity levels, the PTMs trained on MRIs rescaled to 150 intensity levels seems to be less crowded and less noisy as seen on Figure 11. These characteristics may be the result of intensity level reduction by a factor of two which means 50% less PTM entries. Also, the reduction in intensity levels may have contributed an averaging effect along the individual columns of the PTM. Five MR cases were tested to see which PTM produced the breast density estimations closest to the threshold-segmented densities (ground truth) and with the least ‘false voxel’ difference when comparing the HMM segmentation and threshold-applied segmentation. Here, ‘false voxels’ refers to the voxels in the image volume that were segmented to be a tissue state that it should not be. Analysis of the number of false voxels can only be made with MRIs as we have the ground truth for every voxel.
Figure 11. PTMs trained on 20 breast MRIs rescaled to 300 intensity levels, histogram equalized with 10 bins (top), rescaled to 150 intensity levels, histogram equalized with 10 bins (bottom left) and with 15 bins (bottom right)
The relative differences in density estimation and voxel were calculated using equation 3 and 4 for the five randomly selected MR cases that were not used in training the PTMs. The results are displayed in Table 2.

\[
\text{Density Difference} \% = \text{abs}\left(\frac{\text{Density}_{\text{HMM}} - \text{Density}_{\text{Threshold}}}{\text{Density}_{\text{Threshold}}}\right) \times 100\% \quad (3)
\]

\[
\text{Voxel Difference} \% = \frac{\text{sum}((\text{Segmentation}_{\text{HMM}} - \text{Segmentation}_{\text{Threshold}}) \neq 0)}{\text{NumberOfBreastVoxels}} \times 100\% \quad (4)
\]

Table 2. Relative density and voxel difference between HMM segmentation and threshold-applied segmentation for 5 MR cases

<table>
<thead>
<tr>
<th>PTM</th>
<th>Relative Density Difference (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td>Case 4</td>
<td>Case 5</td>
<td></td>
</tr>
<tr>
<td>300IL, 10 Bins</td>
<td>10.90</td>
<td>43.21</td>
<td>44.34</td>
<td>49.27</td>
<td>33.50</td>
<td>36</td>
</tr>
<tr>
<td>150IL, 10 Bins</td>
<td>9.81</td>
<td>40.52</td>
<td>42.51</td>
<td>47.16</td>
<td>33.58</td>
<td>35</td>
</tr>
<tr>
<td>150IL, 15 Bins</td>
<td>11.02</td>
<td>41.37</td>
<td>43.46</td>
<td>49.69</td>
<td>34.43</td>
<td>36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PTM</th>
<th>Relative Voxel Difference (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td>Case 4</td>
<td>Case 5</td>
<td>Average</td>
</tr>
<tr>
<td>300IL, 10 Bins</td>
<td>7.63</td>
<td>19.68</td>
<td>22.37</td>
<td>27.09</td>
<td>14.28</td>
<td>18</td>
</tr>
<tr>
<td>150IL, 10 Bins</td>
<td>7.77</td>
<td>19.13</td>
<td>21.73</td>
<td>25.89</td>
<td>14.36</td>
<td>18</td>
</tr>
<tr>
<td>150IL, 15 Bins</td>
<td>8.17</td>
<td>19.43</td>
<td>22.21</td>
<td>27.02</td>
<td>14.65</td>
<td>18</td>
</tr>
</tbody>
</table>

Comparing the MRIs histogram equalized with 10 bins versus 15 bins, Table 2 shows that using 10 bins gives slightly better density estimation as well as accuracy of voxel states for all five cases. Using 15 bins, the intensity histograms of the MRIs appear more equalized, and according to Figure 11, the PTM shows a slight downward shift of the probabilities where the value of a voxel regardless of its neighbor states now needs to
be lower in order to have a high probability of being adipose. This can be seen visually where the various colors in the PTM break off into a sea of blue.

All in all, the differences between rescaling images to 150 intensity levels over 300 intensity levels or histogram equalizing with 10 bins over 15 bins did not exceed 2%. Out of the five MR cases, the PTM that trained on images rescaled to 150 intensity levels and histogram equalized with 10 bins overall provided the best density estimation and least false voxels. Qualitatively checking the HMM segmentations of the five MR cases, there was no significant differences in the segmentation characteristics as shown by a representative case in Figure 12. The rest of this study went forth with training MRIs rescaled to 150 intensity levels and histogram equalized with 10 bins.

Figure 12. HMM segmentations done by PTM trained on 10 MRIs: threshold-applied (top left), rescaled to 300 intensity levels, histogram equalized with 10 bins (top right); rescaled to 150 intensity levels, histogram equalized with 10 bins (bottom left) or 15 bins (bottom right)
2.5 Size of Training Set for a Robust Probabilistic Model

Women have different breast densities and different anatomical characteristics, so to build a probabilistic model representative of the population, the size of the training set matters. Comparing the PTMs trained from 20 breast MRIs as seen in Figure 11 and PTMs trained from 100 breast MRIs as seen in Figure 9, they are remarkably different in terms of the clear distinction of a trend. Unlike the 20 breast PTM, for the particular arrangement of neighbors around the current voxel, the 100 breast PTM has a distinct linear trend of increasing probability of being glandular from the top left (voxel value 0 and neighbor states AAAAAAAA) to bottom right (voxel value 300 and neighbor states GGGGGGGG). It also exhibits these bands of columns that are less distinct in the PTM that was trained with 20 breasts. To ensure a PTM is robust enough to reproduce density results consistently within one percent and to minimize training time, PTMs were trained on 20, 40, 60, 80, and 100 breast MRIs, building up incrementally with cases selected at random. The PTMs were then used to test on the 12 tomo cases that have a range of small to large breast densities. With reference to the ground truth, the differences in densities were measured to observe the change with increase in training size.
Figure 13. Diagram showing how the increase in training size decreases change in average absolute density difference results when compared to ground truth.
Figure 14. PTM trained from (left to right, top to bottom) 20, 40, 60, 80, 100 breast MRIs
As the size of the training set increased, the probability trend in the PTMs became clearer, smoother, and less noisy as seen in Figure 14. Beginning around with 60 breast MRIs, the PTMs stopped changing significantly beside the upper right area, where the probabilities in that area was mainly filled in by the LDA model. The chance of those combinations of voxel value and neighbor states occurring was minimal and should not have affected the density measurements much. Also, with a larger training set, the LDA had less PTM entries to fill. With 20 breast MRIs, the LDA had to fill in approximately 14% of the entries; whereas, with 100 breast MRIs, the LDA had to fill in only 6%. The line diagram in Figure 13 shows that beginning somewhere between 40 to 60 breast MRIs, the changes in the average density differences begins to hit a plateau. However, keeping in mind that the PTMs will be tested on only 12 tomo cases, we chose to continue using all 100 MRIs available for training the PTMs to minimize error unless stated otherwise.

2.6 Optimization Parameter 2: Distance of Neighbors

The PTM up until now had been built from information from the neighbor arrangement as shown in Figure 6. Neighbors immediately next to the current voxel were chosen to fulfill the assumption that similar breast tissues clump together. Hence, it was natural to build a probabilistic model that looked at a current voxel’s direct neighbors. However, looking at Figure 1, one of the limitations of the HMM method from previous study was that there would be small groups of glandular tissue stranded or spread over
the segmentation that did not reflect true anatomy. Also, the apparent big glandular
structures often appeared quite disconnected in islands. To understand what caused these
limitations, this study explored how the spatial relation of the neighbors affected the state
of the voxel, and ultimately, the overall segmentation.

To minimize breast density offset, only the spatial relationship of the neighbors in
the transverse plane (x and y directions) would be explored initially. Due to tomo images’
poor depth resolution, it was thought that there would not be a strong tissue relationship
between the current voxel’s plane and its out-of-plane neighbor much further in depth.
Hence, the spatial relationship of the neighbors in the z direction was studied later when
testing the tomo images.

In order to study how the neighbors’ spatial relation to the current voxel affects
the HMM segmentation process, PTMs were trained on 20 breast MRIs picking up
information in two new arrangements: neighbors two voxels away and neighbors three
voxels away for both in-plane and in-depth to the current voxel as shown in Figure 15.
The in-depth plane did not change in terms of distance. The PTMs were then tested on
the same five MR cases previously in the optimization of the rescaling factor to see if
there was any improvement in density results and number of false voxels using equation
3 and 4. Also, segmentations were qualitatively examined to determine if the new
arrangement in the neighbors created PTMs that brought forth new patterns and
characteristics.
Figure 15. Alternative neighbor arrangements with respect to the current voxel: two voxels away (left) and three voxels away (right) for in-plane and in-depth

2.7 Results of Optimization Parameter 2: Distance of Neighbors

Table 3. Relative density and voxel difference between HMM segmentation and threshold-applied segmentation for 5 MR cases to compare the differences in the neighbors’ distances to current voxel

<table>
<thead>
<tr>
<th>Neighbor Distance</th>
<th>Relative Density Difference (%)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
</tr>
<tr>
<td>1</td>
<td>-9.81</td>
<td>-40.52</td>
</tr>
<tr>
<td>2</td>
<td>-6.30</td>
<td>-39.56</td>
</tr>
<tr>
<td>3</td>
<td>-5.18</td>
<td>-39.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neighbor Distance</th>
<th>Relative Voxel Difference (%)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
</tr>
<tr>
<td>1</td>
<td>7.77</td>
<td>19.13</td>
</tr>
<tr>
<td>2</td>
<td>7.44</td>
<td>18.77</td>
</tr>
<tr>
<td>3</td>
<td>7.59</td>
<td>19.06</td>
</tr>
</tbody>
</table>

Using the 20 breast MRIs trained PTM to test validate the two new neighbor arrangements on five separate MR cases, there were no changes in breast density estimations or number of false voxels, according to Table 3, when processing two or
three voxels away. Three cases were also examined qualitatively as they covered a range of breast density from 20% (little-medium dense) to 50% (very dense).

Looking at how the spatial relation of the neighbors affect the segmentation product, models built based on closer neighbors and further neighbors had their strong points as well as accompanied undesirable effects.

**Figure 16. MRI of 21% dense breast (left to right):** Threshold-segmented, HMM with neighbors one voxel away, two voxels away, three voxels away; with zoomed in boxes in the same order to evaluate specific regions that match their color borders

Here in Figure 16’s zoomed in red boxes, none of the models preserved the structure of the small enclosed adipose pockets (gray) within the larger glandular
structures (white). This is consistent with the HMM design which tends to rule out small, isolated structures. As shown in the last row, the model that looked at neighbors further away seems to follow the actual trend of the glandular structures (white), but since the neighbors are further out from the current voxel, the model had difficulty in connecting immediate neighbors together. This can partially explain the hazy sampled effect in the segmentations that looked at neighbors two or three voxels away.

The hazy effect can be seen throughout the entire segmentations where the model looked at neighbors two to three voxels away. In a way, it made the segmentations more noisy and the glandular tissues less connected. At the same time, an advantage to looking at neighbors further in distance, the model built a better visual representation of glandular tissues where one could see the contour outline of the structures including areas that require fine details. This is evident in Figure 16 and Figure 17.
Figure 17. MRI of 38% dense breast (top to bottom): Threshold-segmented, HMM with neighbors one voxel away, two voxels away, three voxels away; with zoomed in boxes in the same order to evaluate specific regions that match their color borders
Also, looking at neighbors at a further distance may be advantageous in cases where the breast density is high like in Figure 18. The small pockets of adipose closely packed together within the very dense tissue must have thrown off the model that looked at neighbors directly next to them causing the model to think that the area was surrounded by adipose. The model that looked at neighbors two or three voxels away may realize that the overall surrounding area is very dense, and that the area was actually not surrounded by all adipose but rather small pockets of them.

From this experiment with modeling the PTM looking at neighbors at different distances, it was shown that there are disadvantages and advantages for each model. Further work remains to minimize the hazy effect such as by subsampling the image.
instead of increasing the HMM neighbor distance. Although both approaches are functionally similar in providing anatomical information from further away, image subsampling should provide contiguous segmentation results that are less prone to high frequency noise. This study will follow up with testing the tomo images with the optimized parameters.
Chapter 3. Optimization of HMM method to test Tomo Images

The previous optimization studies were trained and validated on subsets of MRIs. There were no significant differences in density results and false voxel numbers. However, from the two studies, we were able to minimize the size of the PTM and made each PTM entry more robust. Also, from experimenting with models that looked for neighbors at different distances, we were able to observe the outcome of different segmentation characteristics and patterns. To truly see if the optimization changes would affect tomo images similarly, PTMs of the three neighbor configurations were trained on all 100 breast MRIs, rescaled to 150 intensity levels and histogram-equalized with 10 bins and used to test the 12 tomo images. In addition, two extra PTMs were trained at looking at neighbor distances of 2 or 3 voxels away in all three directions isotropically, e.g. \( dx=dy=dz=2 \) and \( dx=dy=dz=3 \). This exploration of neighbors’ distance in the \( z \) direction allowed us to explore more about the limitations in the depth resolution. These two new neighbor configurations can be seen in Figure 19.
Another interesting item to cover is the transition and changes in the PTM appearances itself as shown in Figure 20. From the PTM with 300 intensity levels, the PTM with 150 intensity levels looks a lot less noisy with the PTM entries more blurred and averaged out. When the model looks at the neighbors’ states immediately next to the current voxel, the PTM shows a general trend from blue to yellow to red (probability of a voxel being a glandular state) from the top left hand corner to the bottom right hand corner. However, when the model begins looking at neighbors at a further distance in the x and y direction, the PTM begins showing much more distinct pattern and trend within each 16 neighbor arrangement columns (or 4 bit) along the neighbor state’s axes. When the model looks into neighbors further down in all directions including the z direction,
the trend within each 4-bit column is a lot flattened down. This can potentially mean that
the model is finding a closer connection between the tissue structural and connectivity
pattern within the MRIs.
Neighbor States

Intensity Level

100-breast MRI PTM: rescaled to 300 intensity levels

100-breast MRI PTM: rescaled to 150 intensity levels, direct neighbor arrangement
Figure 20. (top to bottom) Transition of PTM from rescaling from 300 to 150 intensity levels, model based on direct neighbors, neighbors two voxels away in x and y directions, and neighbors two voxels away in x, y, and z directions.
As there was currently no way to compare tomo segmentation with voxel by voxel ground truth like the way with MRIs, the HMM segmentation results were compared with the results prior to optimization and results of single-threshold and double-threshold method, all with reference to the ground truth densities of corresponding MRIs. The single-threshold is the same method used for segmenting the MRIs. The double-threshold is applied after the single-threshold method to separate the mode on the higher end with values above defined as glandular and below adipose [37].

The mean of difference distribution for each segmentation scheme as well as the standard deviation are shown in Table 4. None of the differences between tomo density estimates and MRI ground truth were significant (Student’s t-test, two tailed p>0.05).

<table>
<thead>
<tr>
<th>Tomo Segmentation Scheme</th>
<th>Mean</th>
<th>Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Threshold</td>
<td>28.56%</td>
<td>18.13%</td>
</tr>
<tr>
<td>Double Threshold</td>
<td>-2.94%</td>
<td>11.98%</td>
</tr>
<tr>
<td>300 Intensity Levels, dx=1 dy=1 dz=1</td>
<td>1.87%</td>
<td>12.44%</td>
</tr>
<tr>
<td>150 Intensity Levels, dx=1 dy=1 dz=1</td>
<td>2.24%</td>
<td>12.51%</td>
</tr>
<tr>
<td>150 Intensity Levels, dx=2 dy=2 dz=1</td>
<td>2.81%</td>
<td>12.76%</td>
</tr>
<tr>
<td>150 Intensity Levels, dx=3 dy=3 dz=1</td>
<td>2.69%</td>
<td>12.81%</td>
</tr>
<tr>
<td>150 Intensity Levels, dx=2 dy=2 dz=2</td>
<td>2.66%</td>
<td>12.78%</td>
</tr>
<tr>
<td>150 Intensity Levels, dx=3 dy=3 dz=3</td>
<td>1.97%</td>
<td>12.70%</td>
</tr>
</tbody>
</table>

Surprisingly, the density results from the double threshold segmentation method were not far off from the different HMM segmentation methods. This may be the result of testing on a small sample size of tomo images. Looking carefully amongst the 12 tomo
cases, there was one particular highly dense (65%) case where the double threshold did poorly (38%) while HMM schemes gave a better density estimation (60-62%).

One representative tomo case was studied qualitatively to see if similar segmentation characteristics observed in the prior tests with increasing the distance between the neighbors and the current voxel would appear in the tomo images as well.
Figure 21. (left to right) Original tomo image with its HMM segmentations looking at neighbors (with respect to current voxel) 1) one voxel away 2) two voxels away in x and y directions 3) three voxels away in x and y directions 4) two voxels away in all three directions and 5) three voxels away in all three directions
The segmentation characteristics from the HMM results were comparable to the results from testing the MRIs, looking at Figure 21. When looking at the neighbors directly next to the current voxel (such as $dx=dy=dz=1$), the glandular tissue were more intact and together; whereas when looking at neighbors further away, similar tissue voxels appeared every other voxel in the segmentation giving an overall hazy effect. However, despite the hazy effect, the overall contour and structure of the main glandular tissue seemed more pronounced and connected like how breast tissues are supposed to be. With the additional study of neighbors’ distance in the depth (z) direction, there were surprising discoveries. Some of the out-of-plane artifacts were minimized, and the noise that seemed randomly scattered all over the breast became more connected like the strands extending off of the main glandular structure as can be seen in Figure 21(second to last). As described previously, it may be possible to minimize this hazy effect by subsampling the image itself rather than increasing the neighborhood distance of the HMM.

For this study, one of the biggest limiting factors was the number of test subjects. The ten cases from the newer Inspiration tomo prototype that were omitted had breast-air segmentation artifacts in more than 30% of the slices as shown in Figure 22. The Inspiration cases often resulted in extremely overestimated breast density because the artifacts mainly took up lower gray intensity values, thus shifting the intensity profile with the image information to the high end of the range. Figure 23 shows the histogram of a Novation case in comparison to an Inspiration case after the tomo image had been pre-processed, and 1% of its low and high values were clipped off to the arbitrarily
selected minimum and maximum value of 1 and 300. Hence, the clip off percentage of the low and high end for the Inspiration cases had to be adjusted differently.

Figure 22. Inspiration tomo case showing a slice with breast-air segmentation artifact (shown by red arrow) not completely removed

Figure 23. Histogram comparison of a Novation case (left) and Inspiration case (right) after preprocessing. Histogram comparison of a Novation case and Inspiration case after preprocessing and clipping off 1% of the low and high values to minimum and maximum of 1 and 300
Specifically looking at one Inspiration case, it was determined that different clip off percentages had to be applied to the low and high values. Also, even though the breast-air segmentation artifacts mainly encompassed the lower gray intensity levels, there was not a simple threshold value or clip off percentage in the low end to completely disregard the artifact without compromising true breast voxels. The difficulty of selecting a threshold value can be seen in Figure 24 for the same two slices within the tomo image.

Figure 24. To understand the intensity range that the breast-air artifact encompasses, two slices of a tomo images are shown in 2 columns with all the values below a set threshold value as blue: (top) a lower threshold value set showing the artifact (pointed by red arrow) and (bottom) a higher threshold value set showing the breast values getting clipped off (pointed by green arrow)
To demonstrate the sensitivity of the results to the clipping thresholds, this Inspiration tomo case was rescaled with two manually selected thresholds at the lower end of the intensity scale. The first (lower threshold) was chosen to minimize the breast-air artifact values while preserving all the breast values, while the second (slightly higher) threshold was chosen to completely remove the breast-air artifact values but compromised some breast values. The clip off percentage on the higher end was set to 1% for both experiments. These processed images were then tested by the PTM that was trained on MRIs rescaled to 300 intensity levels. The results showed that with the new manually selected clip off percentage, the density estimation was much closer to the ground truth (21%). Previously when applying the same clip off percentage (1%) that was use for the Novation tomo cases, the HMM density estimation for this tomo case was 88%. With the manually selected clip off percentages, the density estimations were 27% and 9% for the first threshold and second threshold respectively. The clip off percentage that did not compromise the breast values did much better in density estimation. Due to the complexity of the problem and how the clip off percentages could substantially affect density estimation, the Inspiration cases were excluded from this study.
Chapter 4. Discussion

The 3D Hidden Markov Model method was previously established by our lab to accurately determine the breast density from using tomosynthesis, an inexpensive 3D breast imaging modality compared to breast MR. The segmentation method using the HMM was necessary because unlike breast MR, the fibroglandular and adipose tissues are not easily distinguishable.

First and foremost, to establish a baseline to start from and to later be used for comparison purposes, the PTM that was used to determine the results in the previous study was reproduced to the best of our ability. The difference between the mean density differences was 1.7% in spite of the fact that the number of tomo cases used was reduced by almost a half.

From this study, the density estimations gathered from the optimized parameters showed no significant changes. The rescaling study showed that a PTM that trained on MRIs rescaled to 150 intensity levels worked just as well as the PTM that trained on MRIs rescaled to 300 intensity levels. The benefits of a smaller PTM are that it requires less training time and a smaller amount of MRIs to make it robust enough to use as a model. Also, the qualitative results of the tomo segmentations from studying the neighbors’ distance parameter provided great insight on how the HMM model contributed differently to the segmentation process while providing relatively the same density estimation.
One of the main intentions of the study was to minimize the noise in the HMM segmentations, so the structure of glandular and adipose tissues can reflect more closely to true breast anatomy. By using a model that reached out to grab information from further neighbors, the outcome of the segmentations was noisier due to the hazy effect. It may be possible to minimize this artifact by subsampling the image resolution instead of reaching out to a further neighbor. However, on the bright side, the previously interpreted isolated noise was more established to look like glandular structures, the bigger glandular structures had a more contiguous structure to it, and out-of-plane artifacts were reduced.

Also, from this study we see the benefits of the model looking at both the nearby and further neighbors. When using further neighbors, the PTMs showed less noise and more regular patterns, suggesting a better correspondence to the continuous relationship that predicts higher density given increasing intensity value and/or glandular neighbors. As another way to make the model more robust and increasing the specificity in the testing process, the number of neighbors can be increased. However, adding an extra neighbor would be 256 arrangements and adding three neighbors would be 1,024 arrangements. A fully populated 3x3x3 cube would be 26 neighbors (27 total, minus 1 for the current voxel). Although spanning the same space as dx=dy=dz=3, this bigger model would have $2^{26}$ arrangements! Adding arrangements to the PTM will not only affect memory and computational time but also the number of MRIs required for a robust probabilistic model.

Last but not least, before trying other ways to optimize the HMM method, more tomo cases have to be tested to confirm with the results in this study. The Inspiration
tomo cases were not used in this study due to poor breast-air segmentation artifacts; it would not be time efficient and consistent to manually select different clip off percentages for each case. To make the 3D HMM method adaptable to all tomo cases, an automatic algorithm that can preprocess the tomo images to have similar intensity profile prior to testing needs to be made. With more optimization with the neighbor parameter and more efficient image processing, 3D Hidden Markov Model has even more potential to measure breast density accurately from tomosynthesis images.
References


