Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Huanyi Yang

Entitled
Performance Analysis of EM-MPM and K-means Clustering in 3D Ultrasound Breast Image Segmentation

For the degree of Master of Science in Electrical and Computer Engineering

Is approved by the final examining committee:

Lauren Christopher
Chair

Paul Salama

Maher Rizkalla

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Lauren Christopher

Approved by: Brian King 1/18/2013
Head of the Graduate Program Date

Graduate School Form 20
(Revised 9/10)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Performance Analysis of EM-MPM and K-means Clustering in 3D Ultrasound Breast Image Segmentation

For the degree of Master of Science in Electrical and Computer Engineering

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

HUANYI YANG

Printed Name and Signature of Candidate

1/24/2013

Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

PERFORMANCE ANALYSIS OF EM-MPM AND K-MEANS CLUSTERING IN 3D ULTRASOUND BREAST IMAGE SEGMENTATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Huanyi Yang

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Electrical and Computer Engineering

May 2013

Purdue University

Indianapolis, Indiana

To my family.

ACKNOWLEDGMENTS

I express my sincere gratitude to my advisor Professor Dr. Lauren Christopher of the Department of Electrical and Computer Engineering, for the continuous support of my study and research. Her guidance helped me in the research and writing of this thesis.

Besides my advisor, I would like to thank my friends Chao Liu, Yan Sun, Jiang Feng, Cong Shan, Akella Aditya, for their encouragement and help, through the duration of my thesis.

Last but not the least, I would like to thank my family, my parents Yang Min, Cao Fengqing, and my husband Chen Qi, for supporting me spiritually throughout my life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Appendix Figure

Figure Page

ABSTRACT

Yang, Huanyi. M.S.E.C.E. , Purdue University, May 2013. Performance Analysis of EM-MPM and K-means Clustering in 3D Ultrasound Breast Image Segmentation. Major Professor: Lauren Christopher.

Mammographic density is an important risk factor for breast cancer, detecting and screening at an early stage could help save lives. To analyze breast density distribution, a good segmentation algorithm is needed. In this thesis, we compared two popularly used segmentation algorithms, EM-MPM and K-means Clustering. We applied them on twenty cases of synthetic phantom ultrasound tomography (UST), and nine cases of clinical mammogram and UST images. From the synthetic phantom segmentation comparison we found that EM-MPM performs better than K-means Clustering on segmentation accuracy, because the segmentation result fits the ground truth data very well (with superior Tanimoto Coefficient and Parenchyma Percentage). The EM-MPM is able to use a Bayesian prior assumption, which takes advantage of the 3D structure and finds a better localized segmentation. EM-MPM performs significantly better for the highly dense tissue scattered within low density tissue and for volumes with low contrast between high and low density tissues. For the clinical mammogram, image segmentation comparison shows again that EM-MPM outperforms K-means Clustering since it identifies the dense tissue more clearly and accurately than K-means. The superior EM-MPM results shown in this study presents a promising future application to the density proportion and potential cancer risk evaluation.

1. INTRODUCTION

Breast cancer is the most common cancer in women and a frequent cause of death from cancer in most developed countries. Assessing a woman's lifetime risk of breast cancer at an early stage can help save lives by targeting screening and preventive therapy to the at-risk population. Some studies have shown that a woman's breast density proportion is a strong risk factor for breast cancer risk; the higher the proportion of dense tissue, she is 4-5 times more likely to develop breast cancer [1]. Thus, mammographic density can be viewed as an intermediate phenotype for breast cancer. Among many popular methods used for analyzing mammographic density, recent studies have demonstrated the effectiveness of Tomographic Ultrasound (UST) imaging in detecting breast cancer, particularly for women with dense breasts [9] [11]. Since the density is automatically calibrated to water in this imaging mode, UST could provide a more accurate measurement compared with some other detection methods where the result may be affected by operating conditions. In addition, in contrast to the 2-D projection in standard mammography, a 3-D model of the breast density is available in tomographic ultrasound. In order to get a density map of the data, the process of segmenting this tissue density data is needed. Therefore, this study focuses on obtaining a repeatable measure of density proportion using 3-D Ultrasound Tomography. In our research, we apply two methods of segmentation for both synthetic and clinical breast image. Then we compare, analyse and discuss the performance of the segmentation accuracy.

To analyze breast density distribution, a good segmentation algorithm is needed. Commonly used segmentation techniques such as filtering, region growing, thresholding, and non-linear edge operations are not effective enough in UST data due to the noise degradation in ultrasound. Here we choose two advanced robust algorithms, EM-MPM and K-means Clustering to compare. First, the Bayesian algorithm

combining Expectation Maximization with the Maximization of Posterior Marginals (EM-MPM) is considered effective in many difficult segmentation tasks [2] [3] [4]. This algorithm classifies every pixel in an image by assigning a cost to the number of misclassified pixels, and iteratively finds the best probabilistic solution to fit the data. This method has the advantage of using a 3-D neighborhood of pixels as a statistical Bayesian prior, and has the effect of grouping the data similar to the way the tissues are structured. Second, K-means Clustering is another commonly used technique [5] in medical image processing, especially in images with noisy data. This algorithm takes n observations and segments them into k clusters in which each observation belongs to the cluster with the nearest mean. The initialization of the k-means algorithm can be critical, and in this case seeds were automatically placed randomly. In the article written by *Predrag R. Bakic* [5], this K-means Clustering technique is used to analyze the volumetric breast density, based on a set of synthetic images generated with an anthropomorphic software breast phantom, which accurately simulates the arrangement of breast tissues according to the analysis of histological and radiological images. The K-means Clustering segmentation result shows a high correlation with the ground truth information about the simulated breast tissues provided by the phantom, however still not perfect enough from medical perspective. In our work described in this thesis, first we show EM-MPM provides a better result than K-means Clustering on the synthetic UST, with superior Tanimoto Coefficient and Parenchyma Percentage. Second, we apply both methods on clinical data, which again shows EM-MPM has clearer subjective segmentation result than K-means Clustering.

In Chapter 2, the EM-MPM and K-means Clustering algorithms are introduced. A global view and analysis of the algorithm helps on deep understanding of the differences between these two methods. EM and MPM algorithms are discussed here, also the relationship between them is explained. For the K-means Clustering, the algorithm are reviewed step by step, accompanied with equations and examples.

In Chapter 3, the segmentation results from the two algorithms are compared and discussed, using synthetic ultrasound tomography and clinical ultrasound im-

age. First, the data developed from anthropomorphic software breast phantom is reviewed. It is an accurate model containing the simulated information of different density tissues within the breast tissue, which provides a convenient way to compare the accuracy with the ground truth data. Then we follow by using two parameters for measuring the segmentation accuracy, Tanimoto Coefficient and Parenchyma Percentage. Tanimoto Coefficient measures the matching proportion between two images. The Parenchyma Percentage applies to measuring the percentage of the dense tissue compared with the whole tissue. Next we compare and discuss the segmentation results of EM-MPM and K-means Clustering. EM-MPM shows promising results with both superior Tanimoto Coefficient and Parenchyma Percentage. Finally, we introduce the clinical image data. The clinical digital mammogram and clinical ultrasound tomography are explained. Then we apply the two algorithms on the ultrasound tomography images, and compare the result with the currently used standard, the commercial Cumulus software which is based on percent density [5]. Again, we see that EM-MPM has a better segmentation result than K-means Clustering.

In Chapter 4, we conclude with the advantages of EM-MPM segmentation on ultrasound tomography, it showed more accurate breast density analysis compared with K-means Clustering. Also, we point out that the data sets used in our research are limited, especially clinical data cases. More clinical ultrasound images are needed for further analysis of the segmentation accuracy improvement.

2. ALGORITHMS USED IN THE SEGMENTATION

In our study, we apply EM-MPM algorithm on both synthetic and clinical UST images, then compare the EM-MPM segmentation result with the corresponding K-means Clustering segmentation result, which comes from the previous work at Karmanos Cancer Institute. In this chapter a briefly review of the two algorithms is presented.

2.1 EM-MPM

The EM-MPM algorithm consists of two parts: Expectation-Maximization (EM) and Maximization of the Posterior Marginals (MPM) [2] [3]. The EM algorithm finds the estimates for Gaussian mean and variance, while MPM classifies the pixels into N class labels, using the estimated parameters from EM. The basic structure of the image processing is a 3-D neighborhood of pixels. In the 3-D image research field, this forms a mathematical structure called a Markov Random Field (MRF). The MRF is useful because it guarantees local convergence in iterative algorithms which are based on it. The 3-D 6-pixel neighborhood which we use is: right, left, above, below, front, and back around a center pixel.

In the 3-D image, the source image gray level information is considered a 3-D volume of random variables, Y. As medical images, the model assumes that Y contains Gaussian noise due to the image processing, plus the true underlying tissue characteristics. The segmentation result approximates the true tissues, denoted as X, which does not contain noise or distortion. The class label is taken from a set of N labels. Described here is the optimization process by which we classify the pixels into the N labels.

A random class label is initialized into every pixel in X at the beginning of the segmentation process, with an evenly distributed vector of means and variances. Then, the estimate of X is formed by iterating several times through the 3-D data. The probability density function of a mixture of Gaussians, in which the random variable Y is dependent on X, is modeled in following Equation [4]:

$$f_{Y|X}(y|x,\theta) = \prod_{s\epsilon S} \frac{1}{\sqrt{2\pi\sigma_{x_s}^2}} \exp\left\{-\frac{(y_s-\mu_{x_s})^2}{2\sigma_{x_s}^2}\right\} \tag{2.1}$$

Where σ is the variance for each class, μ is the mean for each class, x_s is the center pixel, y_s is the source image, θ is the vector of means and variances of each class, S is the 3-D volume of pixels.

Since we are assuming Bayesian dependence, to find the probability mass function of $X|(Y,\theta)$, we can use Equation 2.2 to iteratively solve for $\hat{x}$:

$$p_{X|Y}(x|y,\theta) = \frac{1}{Z f_Y(y|\theta)} \prod_{s\epsilon S} \frac{1}{\sqrt{2\pi\sigma_{x_s}^2}} exp\{-\frac{(y_s-\mu_{x_s})^2}{2\sigma_{x_s}^2} - \sum_{[r,s]\epsilon C} \beta t(x_s,x_r)\} \tag{2.2}$$

Where C is the neighborhood of X, β is the weighting factor for amount of spatial interaction, and $t\,(x_r,x_s) = 0$ when $x_r = x_s$,or $t\,(x_r,x_s) = 1$ when $x_r \neq x_s$.

We take the $log(p_{X|Y})$ and ignore the terms that do not depend on x, such as $\frac{1}{Z f_Y(y|\theta)}$, then the result is the Equation 2.3 for optimization. It provides the optimized segmentation result by choosing a class label for every pixel in the estimate of X which can maximize the marginal probability mass function:

$$argmax\left\{-log\sigma_{x_s} - \frac{(y_s-\mu_{x_s})^2}{2\sigma_{x_s}^2} - \sum_{\{r,s\}\epsilon C} \beta t\,(x_r,x_s)\right\} \tag{2.3}$$

The $p_{X|Y}(x|y,\theta)$ is the posterior marginal distribution at a pixel location s, so in this equation we are Maximizing the Posterior Marginals (MPM).

Expectation Maximization (EM) is an iterative procedure for estimation of the mean and the variance of each segmentation classes. At each iteration, two steps are performed: the expectation step and the maximization step. First maximization

step is performed, then the segmentation is done in the expectation step, iterating to find the best log-likelihood of the probability that a particular pixel belongs to one of the k classes. The means and variances are represented by the vector: $\theta = (\mu_1, \sigma_1, \ldots, \mu_k, \sigma_k)$. The MPM probability, $p_{x_s|Y}\left(k|y, \theta\left(w-1\right)\right)$, where w is the number of EM iteration, yielded the MPM loop, is directly applied in the EM update Equations for μ_k, σ_k^2:

$$\hat{\mu_k}\left(w\right) = \frac{1}{N_k\left(p\right)} \sum_{s \epsilon S} y_s p_{x_s|Y}\left(k|y, \theta\left(w-1\right)\right) \tag{2.4}$$

$$\hat{\sigma_k}^2 = \frac{1}{N_k(w)} \sum_{s \epsilon S} \left(y_s - \mu_k(w)\right)^2 p_{x_s|Y}\left(k|y, \theta(w-1)\right) \tag{2.5}$$

Where N corresponds to the probability weighted number of pixels in a particular class:

$$N_k\left(w\right) = \sum_{s \epsilon S} p_{x_s|Y}\left(k|y, \theta\left(w-1\right)\right) \tag{2.6}$$

2.2 K-means Clustering

K-means is a simple unsupervised segmentation algorithms that solve various kinds of clustering problem [6]. It is a simple and easy way to classify a given object through a certain number of clusters (assume k clusters), in which each observation is classified to the cluster with the nearest mean, m.

Given a set of observations $(x_1, x_2, ..., x_p)$, the main idea is to define k centroids, one for each cluster, then classifies the n observations into k sets $S = \{S_1, S_2, ..., S_k\}$. Since different centroids location may produce the different result, so it is better to place them as far away as possible from each other. in this study centroids were automatically placed randomly. The next step is to take each point belonging to a given data set and group it to the nearest centroid. When no point is pending, the first step is completed and the first round segmentation is done. At this point we need to re-calculate the k new centroids as centers of the clusters resulting from the

first step. After these k new centroids are calculated, a new segmentation has to be done between the same object data points and the nearest new centroid. In this whole updating loop, the k centroids change their location gradually until no more changes are found. In other words the centroids do not move any more. Finally, this algorithm produces a separation of the objects into groups which can minimize the distance between objects and centroids.

The algorithm is composed of the following steps:

(1) Place k points into the target space composed by the objects that need to be clustered, which represent the initial centroids.

(2) Assign each object to the cluster that has the closest mean with the target centroid, where each x_p goes into exactly one ${S_i}^t$.

$$ {S_i}^t = \{x_p : ||x_p - {m_i}^t|| \leq ||x_p - {m_j}^t||, \vee 1 \leq j \leq k\} \tag{2.7} $$

(3) When all objects have been classified, calculate the new positions of the k centroids.

$$ {m_i}^{t+1} = \frac{1}{{S_i}^t} \sum_{\{x_j\} \epsilon S_i} x_j \tag{2.8} $$

(4) Repeat Steps 2 and 3 until the centroids no longer move, which produces a segmentation of the objects into groups with the minimized distance to the target cluster centroid.

3. RESULT AND DISCUSSIONS

In this section, we compare and analyze the segmentation result from EM-MPM and K-means Clustering. First, we apply the algorithms on twenty sets of 3-D synthetic breast ultrasound tomography (UST) with various densities. Then we compare their volumetric breast density (VBD) result with the ground truth VBD values available from the phantoms. Second, we apply the algorithms on nine sets of 3-D clinical UST images, then compare their segmentation accuracy with the currently used standard, percent density (PD) values which were estimated interactively by a clinical breast radiologist using Cumulus software.

3.1 Result and Disscussions on Synthetic Breast UST

3.1.1 Synthetic Breast UST

Previously, some studies about breast density analysis based on clinical UST images has been conducted at the Karmanos Cancer Institute [5]. They showed a correlation between volumetric breast density estimates from clinical ultrasound tomography (UST) images and the Cumulus based percent density estimates from clinical digital mammogram images. Such studies have a practical limitation since UST images represent a reconstructed 3-D image of the breast, while digital mammography represents a 2-D projection through the breast, in addition, no information is available clinically about the ground truth breast density values. To overcome the limitation of lacking the ground truth, we use the synthetic images generated with an anthropomorphic software breast phantom, developed at the University of Pennsylvania [8]. The phantom is based upon a detailed analysis of breast anatomy visualization by clinical images and sub-gross pathology, it simulates the arrangement of breast tissues

coming from the analysis of histological and radiological images. The ground truth information about the simulated breast tissues provided by the phantom allows the calculation of the absolute error in breast density measurements and potential sources of that error.

Design of the anthropomorphic phantom used in this study allows for simulating multimodality breast images. Examples of clinical digital mammogram, UST image, phantom, simulated mammographic projection, and the corresponding UST images are shown in Figure 3.1 and Figure 3.2. The phantom offers great flexibility in simulating various breast size, glandularity, and internal composition. Starting from a realistic skin surface, the phantom interior includes simulated tissue structures, as adipose compartments, Coopers ligaments, and glandular tissue.

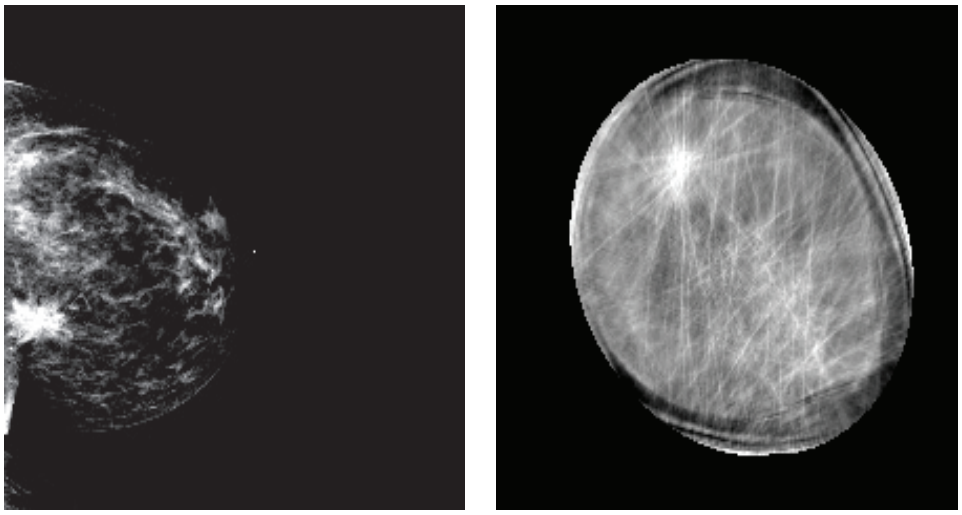
3.1.2 Tanimoto Coefficient

In order to measure the segmentation accuracy, we use the Tanimoto Coefficient to measure of the overlap proportion of two images of a single segmentation class [7]. Given two images A and B (as shown in Figure 3.3), each with n binary pixels. Let's say A is the original ground truth image, with binary pixel value equals 0 if it is not in the class, or equals 1 if in the class. B is the partially correct segmentation image based on A. As shown in Equation 3.1, the Tanimoto Coefficient is defined as the proportion of the overlapping pixels divided by the whole union of total pixel numbers in the target class. The perfect segmentation would be T=1.0. Normally very good segmentations are above T=0.6, and can depend on the noise and distortion presented in the images.

$$T = \frac{N_{11}}{N_{01} + N_{10} + N_{11}} \tag{3.1}$$

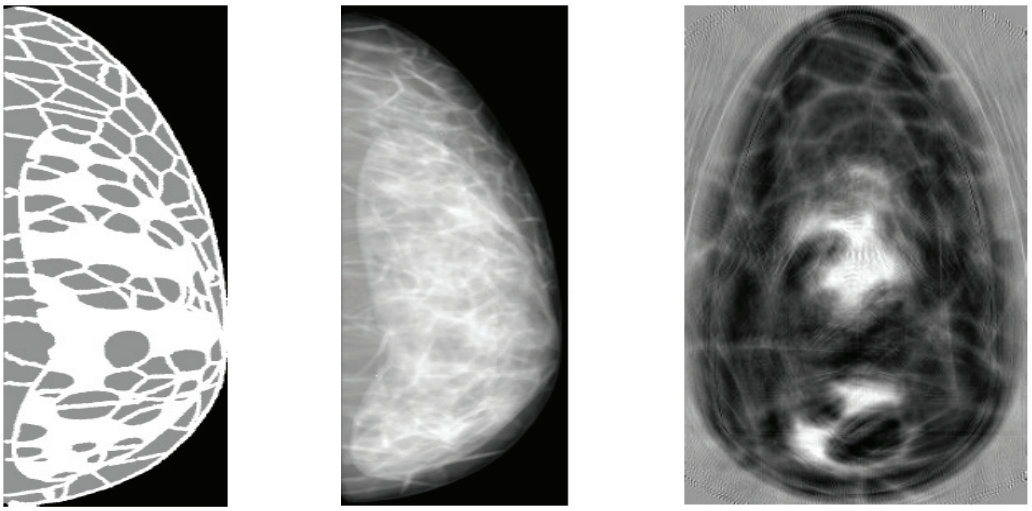
Here we define:

(1) N_{11} represents the total number of pixels where A and B both have a value of 1, which means the correct segmentation part in B.



Clinical Mammogram Clinical UST

Fig. 3.1. Clinical images



Phantom Mammogram from Phantom UST from Phantom

Fig. 3.2. Phantom images

(2) N_{01} represents the total number of pixels where the attribute of A is 0 and the attribute of B is 1, which means the mis-segmentation part in B.
(3) N_{10} represents the total number of attributes where the pixel of A is 1 and the pixel of B is 0, which means the missing segmentation part in B.
(4) N_{00} represents the total number of pixels where A and B both have a value of 0, which means the not-in-class part in both A and B.

3.1.3 Parenchyma Percentage

The second measure of segmentation accuracy is Parenchyma Percentage, which analyzes the volumetric breast density from UST images. We define the parenchyma percentage by dividing the volume of high density tissue by the total volume of the breast tissue (Equation 3.2). From the clinical perspective, the dense tissue means the fibroglandular tissue, Coopers ligaments, and the tumor, which, in our synthetic phantom they are represented by the relatively high brightness pixels compared with the dark back ground, which is shown in Figure 3.4. In the UST slice we can see those brighter areas, and the corresponding dense parts are identified by the segmentation image with the highest gray level(s).

$$Parenchyma\ Percentage = \frac{V_{dense}}{V_{total}} \tag{3.2}$$

3.1.4 Result and Discussions

In this section, the segmentation result from EM-MPM and K-means Clustering are compared and discussed. A total of twenty sets of anthropomorphic phantom UST are selected, which roughly follow the distribution clinically estimated from over 2800 women using digital mammograms and breast CT images.

First we apply our algorithms, EM-MPM on each slice separately of the 3-D reconstructed phantom ultrasound tomography. Here we use three class option. For K-means Clustering the two class option is used in the previous work at Karmonos

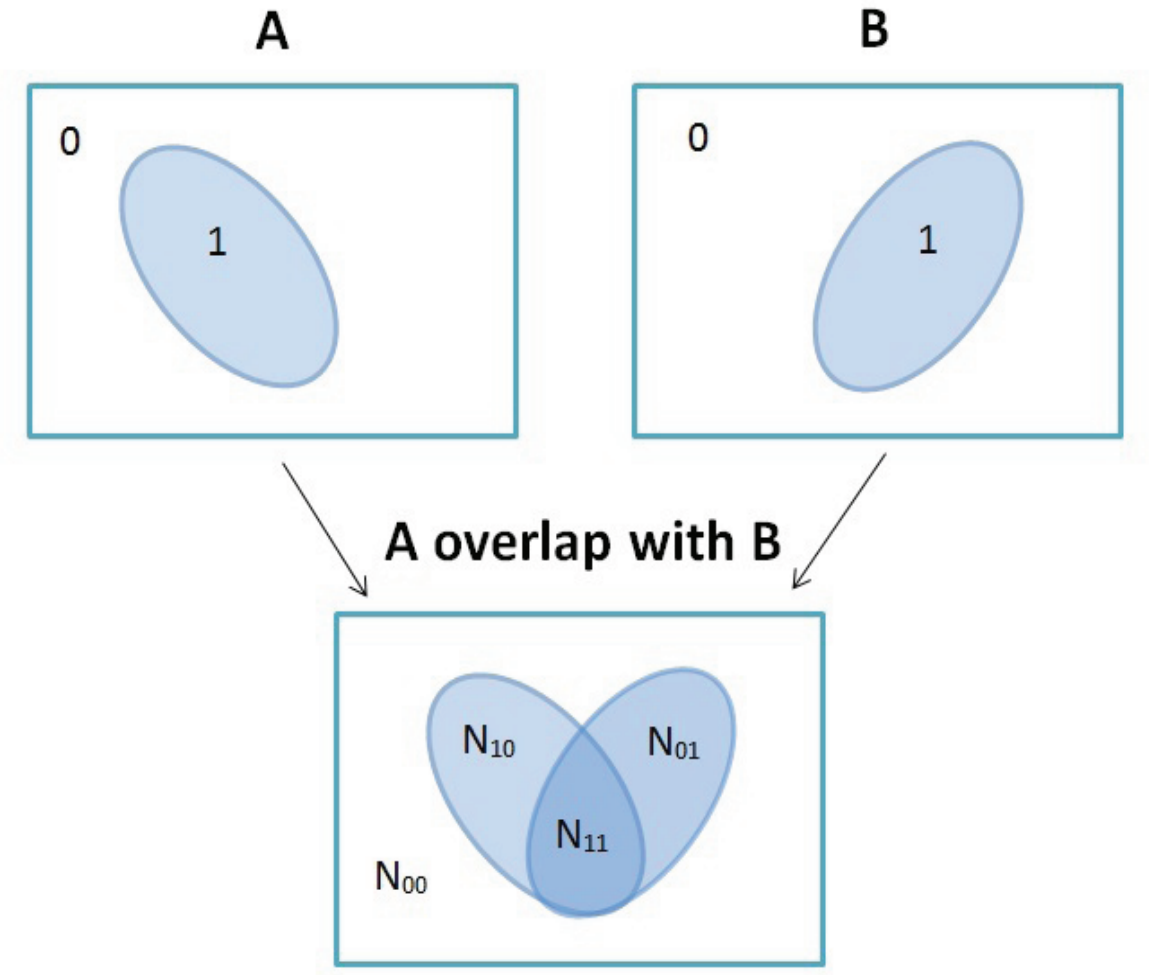
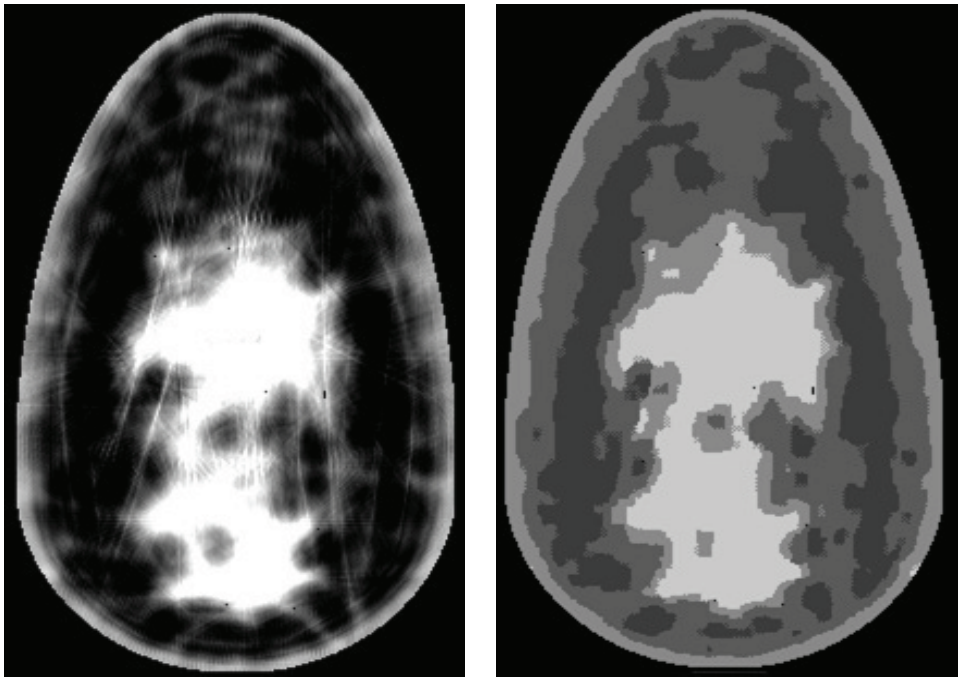


Fig. 3.3. Tanimoto Coefficient



Software-generated UST Slice EM-MPM Segmentation Image

Fig. 3.4. Software-generated UST slice and the corresponding EM-MPM segmentation image

Cancer Institute [5]. We calculate and compare the two methods' segmentation accuracy using 3-D volumetric Tanimoto Coefficient and Parenchyma Percentage. Tanimoto Coefficient is calculated by comparing the overlapping dense part between the segmentation image and the corresponding ground truth image. The dense part defined in the ground truth image are represented by higher gray levels 191, 202 and 223, corresponding to the edge (skin), the intersection lines (connective tissue) and bright areas (dense tissues) within breast. Parenchyma Percentage is calculated as dividing the dense tissue over the whole volume of the phantom. The dense tissue is represented by the top class in segmentation image, or top three brightness pixels in ground truth image.

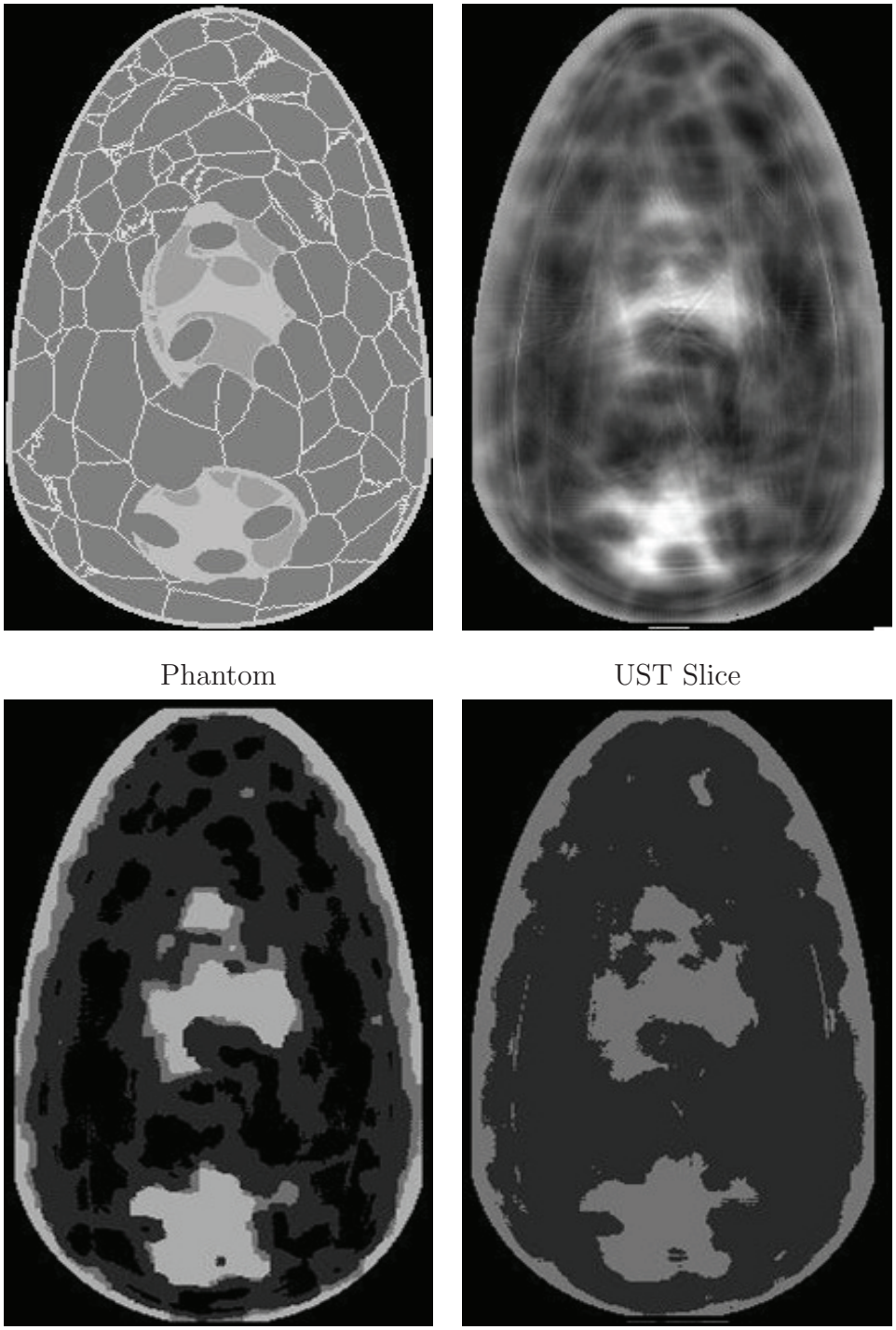
A selection of phantom cases are shown in Figure 3.5 to Figure 3.10 (with remaining cases in Appendix A), comparing the segmentation result between EM-MPM and K-means Clustering. Phantom images presented here are some specific slices chosen

out of the 3-D UST with various densities. In Figure 3.5 and Figure 3.6, where the dense tissues clustered tightly and could be clearly identified from the less dense part, EM-MPM and K-means Clustering performs almost the same, where both of them can find the dense part out of the normal tissue. In addition, EM-MPM works a little better on the edge (skin) detection, which picks thinner skin area compared with K-means Clustering.

In Figure 3.7 and Figure 3.8, there are two cases of low density phantom slices, which the dense tissue is small and sparsely scattered. Here we can see EM-MPM performs better than K-means Clustering, it finds the tiny dense parts clearly, and matches very well with the ground truth image. For K-means Clustering, instead of concentrating on the small dense area, it tends to evenly spread the dense part within the whole tissue, which is due to the updating centroids principle.

Figure 3.9 and Figure 3.10 show two cases of high density phantom slices, where the dense tissue took over more than 50 percent of the whole tissue. Again EM-MPM out performs K-means Clustering, which accurately segmented the dense part from the background (compared with the ground truth data). On the contrary, K-means Clustering tends to omit some dense parts, average it, and make it less dense.

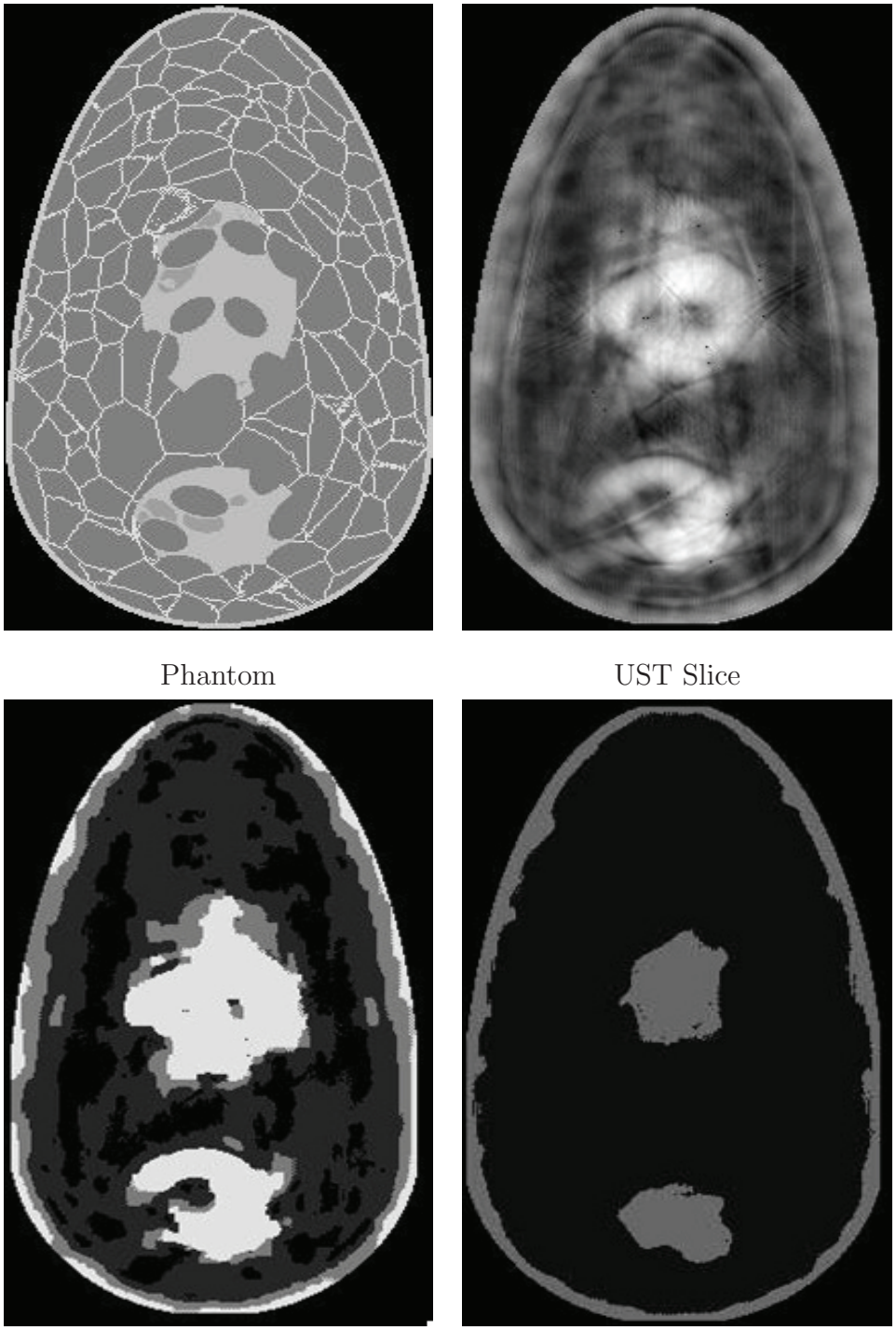
Figure 3.11 illustrates the Tanimoto Coefficient comparison between EM-MPM and K-means Clustering. As mentioned in previous section, Tanimoto Coefficient is used to judge the similarity between two images, the higher value of Tanimoto Coefficient, the higher similarity of the two images. In this study, the overlapping points were accumulated slice by slice and the results were added up to yield a volume based Tanimoto Coefficient. To show things more clearly, the twenty UST phantom cases are arranged by ascending sequence of the density (according to the ground truth data). In the chart we can see, from low density to high density case, EM-MPM always has higher Tanimoto Coefficient than K-means Clustering (with improvement between 0.05-0.3), which means EM-MPM out performs K-means Clustering on segmentation accuracy.



Phantom UST Slice

EM-MPM K-means Clustering

Fig. 3.5. Comparison between EM-MPM and K-means Clustering on phantom UST, clearly clustered case #1. (Both EM-MPM and K-means Clustering match with the ground truth image)



Phantom UST Slice

EM-MPM K-means Clustering

Fig. 3.6. Comparison between EM-MPM and K-means Clustering on phantom UST, clearly clustered case #3. (Both EM-MPM and K-means Clustering match with the ground truth image)

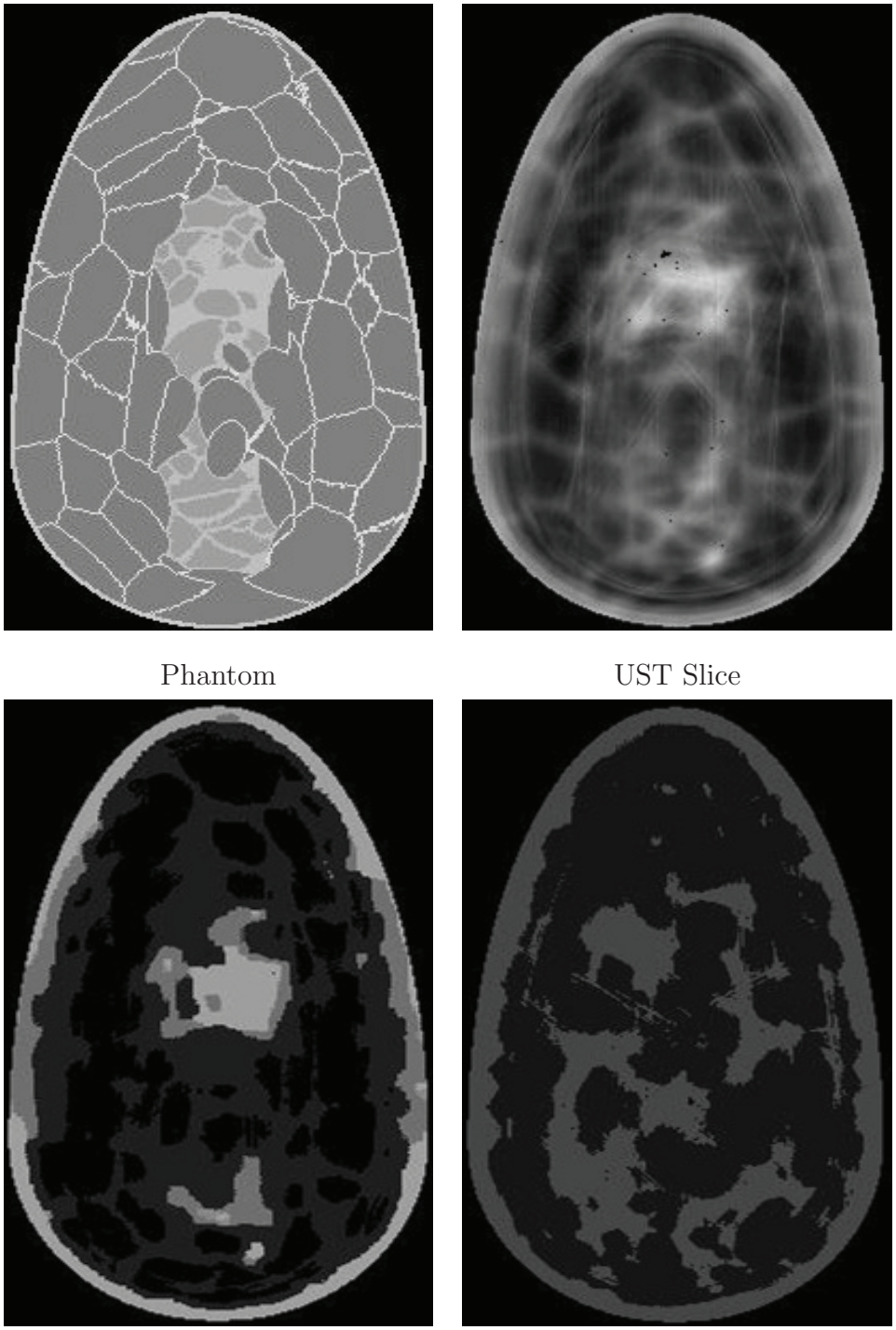


Fig. 3.7. Comparison between EM-MPM and K-means Clustering on phantom UST, low density case #2. (EM-MPM performs better than K-means Clustering)

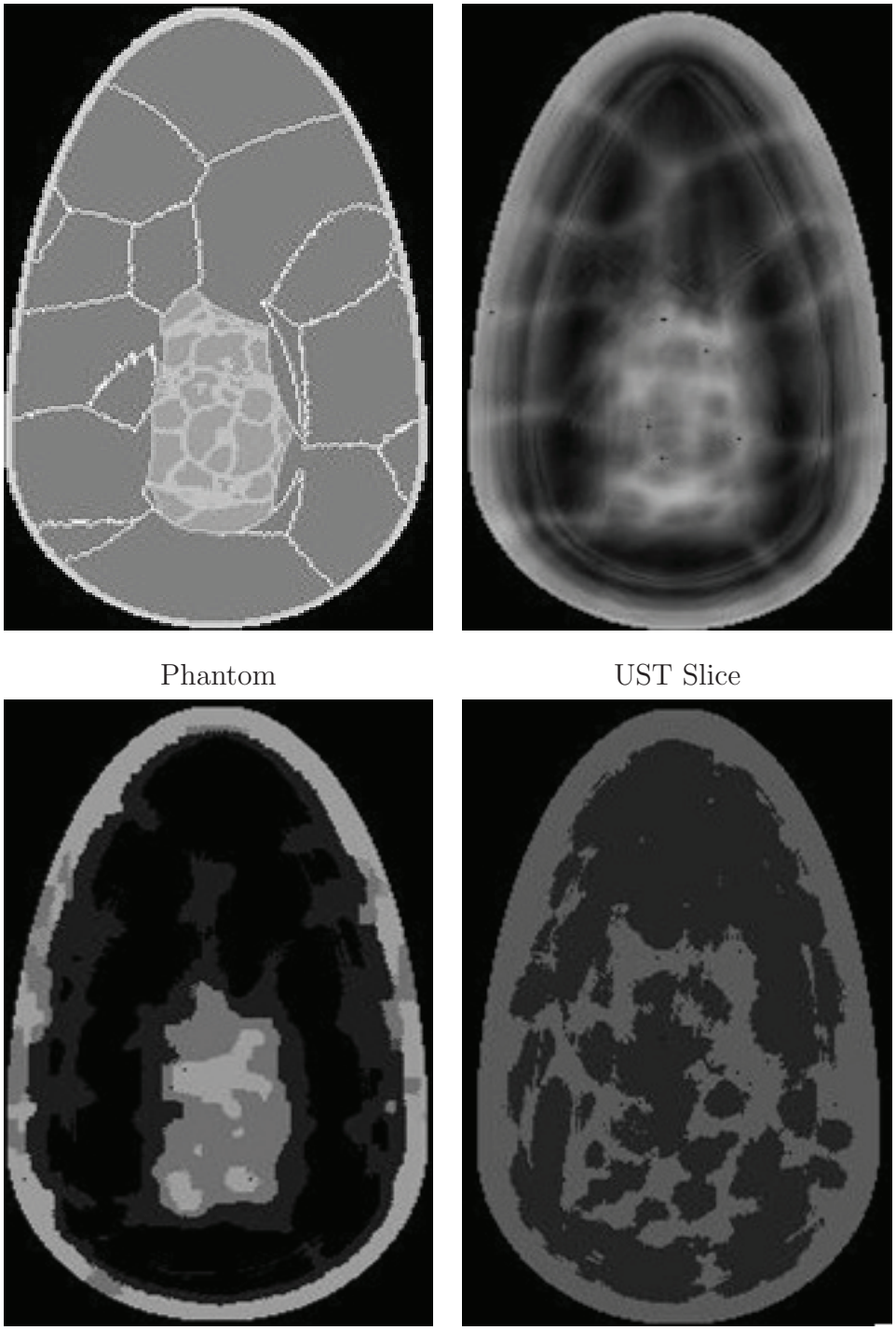
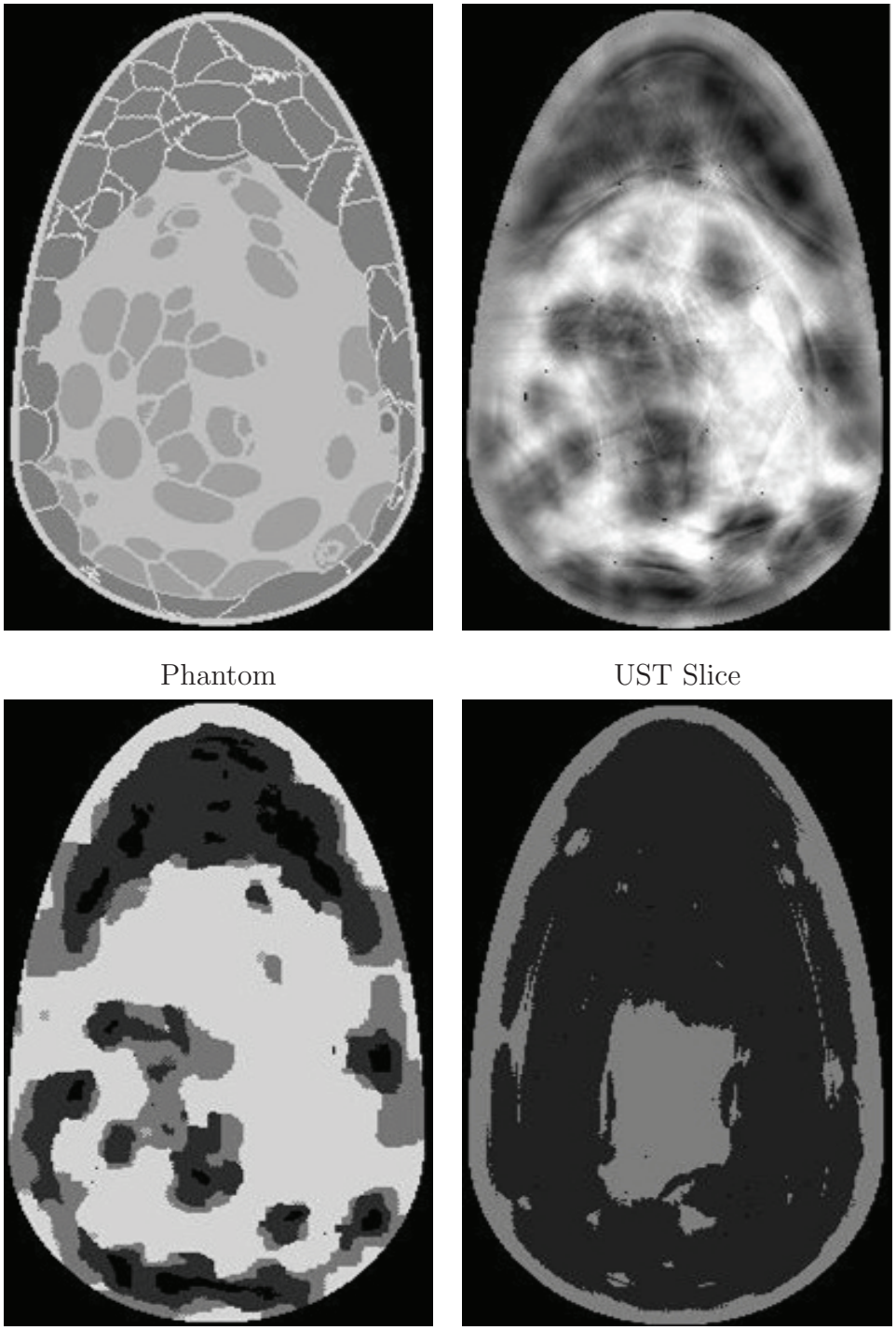


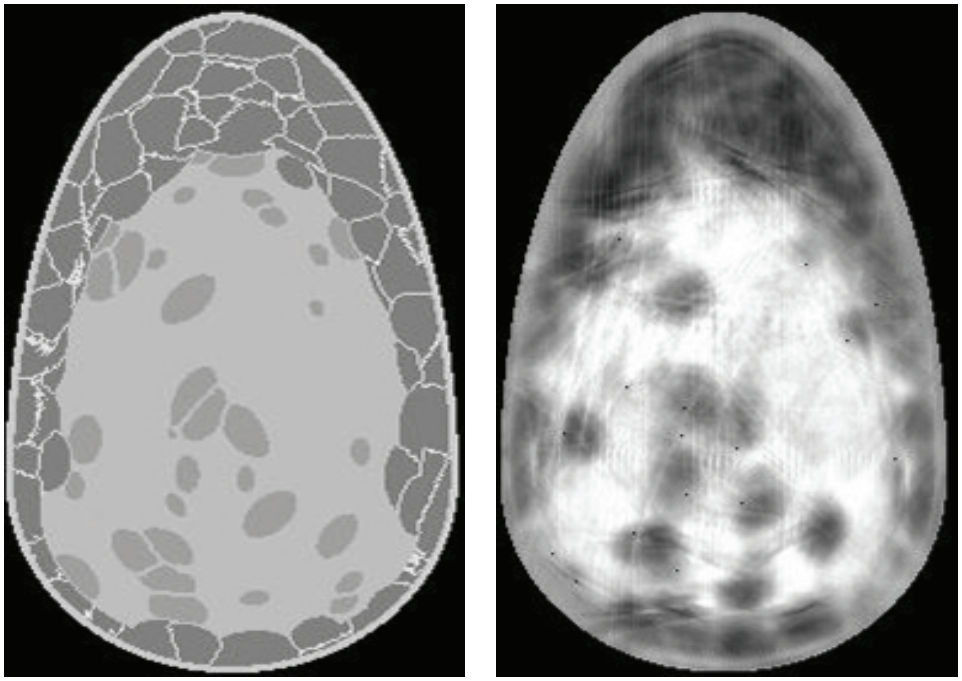
Fig. 3.8. Comparison between EM-MPM and K-means Clustering on phantom UST, low density case #20. (EM-MPM performs better than K-means Clustering)



Phantom UST Slice

EM-MPM K-means Clustering

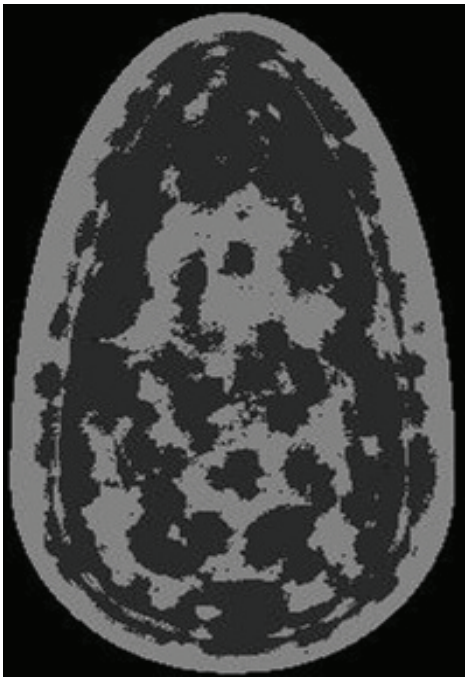
Fig. 3.9. Comparison between EM-MPM and K-means Clustering on phantom UST, high density case #4. (EM-MPM performs better than K-means Clustering)



Phantom

UST Slice

EM-MPM



K-means Clustering

Fig. 3.10. Comparison between EM-MPM and K-means Clustering on phantom UST, high density case #19. (EM-MPM performs better than K-means Clustering)

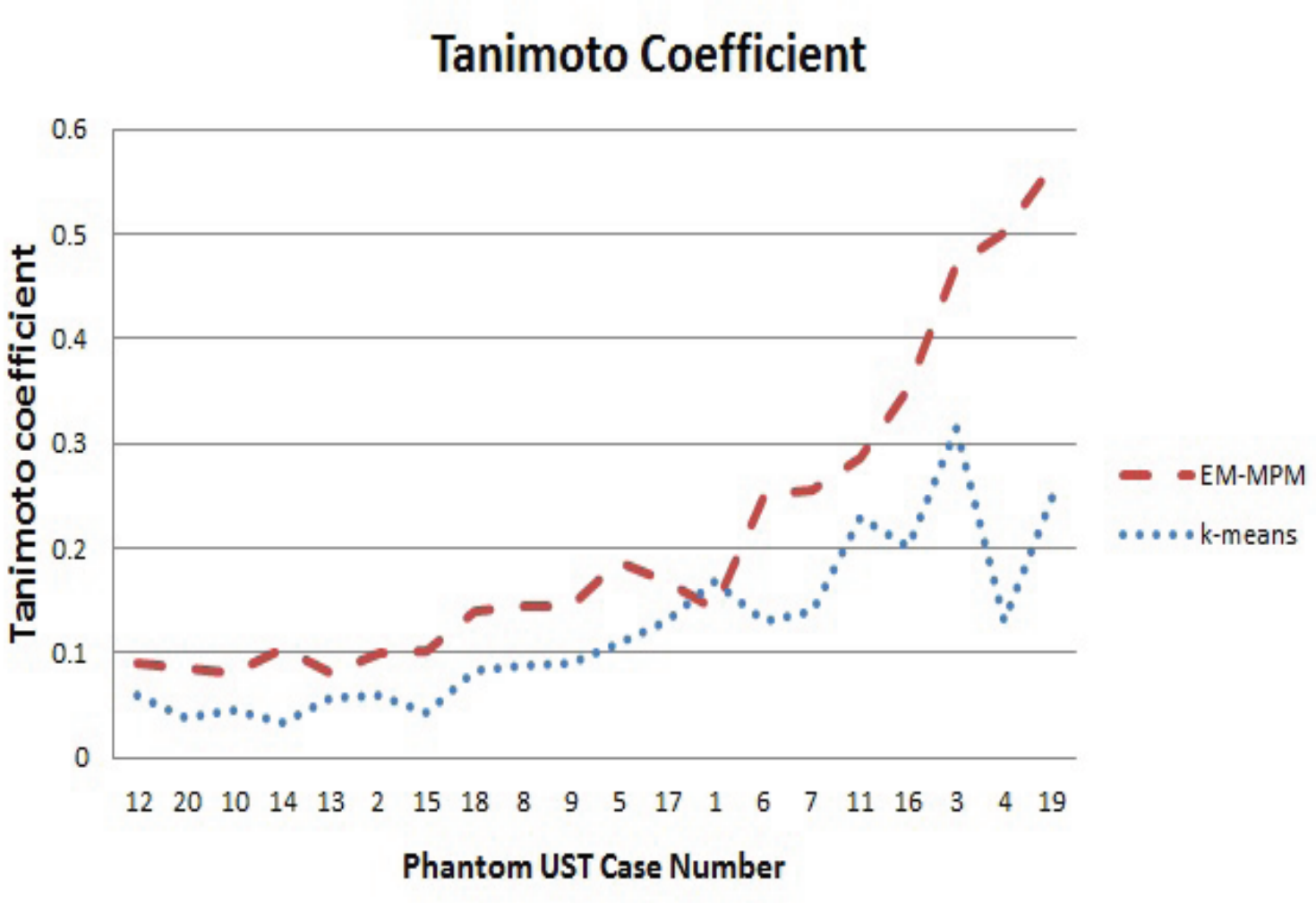


Fig. 3.11. Tanimoto Coefficient Comparison of EM-MPM and K-means Clustering

Figure 3.11 shows that in low density cases, the EM-MPM Tanimoto Coefficients are relatively low for both methods (around 0.1-0.2), it is because the the edge (skin) and the intersection lines (connective tissue) parts are distorted in the UST slices. In the EM-MPM segmentation result, it shows the edge (skin) part is thicker than original phantom, but for the intersection lines (connective tissue) they are grouped to the less dense class. So when the phantom has low density, these edge and intersection lines mis-classification will greatly impact the Tanimoto Coefficient. However, for the edge (skin) issue, its high gray level effect is primarily due to the finite data coverage by the anthropomorphic simulation software mentioned in Section 3.1.1, which could be avoided by simulating larger transducer array in the future [8]. The intersection lines (connective tissue) should not be regarded as an issue, since our segmentation is based on the software simulated UST slice, which has already averaged the intersection lines due to the adding noise and limited resolution. In high density cases, the Tanimoto Coefficient of EM-MPM is much superior to K-means Clustering, and reaches 0.5 to 0.6, representing a very good matching between the segmentation and the ground truth. Therefore EM-MPM has high segmentation accuracy, especially for the important high density cases. This could greatly help the dense proportion and potential cancer risk evaluation in the clinical cases.

Parenchyma Percentage comparison is shown in Figure 3.12. As in the Tanimoto Coefficient chart, the twenty cases are arranged by ascending sequence of the density. The dense parts (the first class in the segmentation result) were accumulated slice by slice then divided by the 3-D phantom volume to calculated the volume based percent density. It clearly shows that from low dense to high dense cases, EM-MPM segmentation result matches very well with the phantom percentage density, which again shows that it accurately finds the dense tissue proportion as defined in the original phantom picture (the three higher gray level parts). However, K-means Clustering doesn't track with the ground truth data, but tends to average the percentage around 0.3 to 0.4 in all cases. This is because it tends to overestimate

the low density parts in low dense cases but discard some dense part in high density cases (as shown in Figure 3.7 to Figure 3.10).

3.2 Result and Discussions on Clinical Images

Previously we mentioned that breast density is an important risk factor for breast cancer, early detection can help prevent the cancer at an early stage. In the previous section we demonstrated that EM-MPM has higher segmentation accuracy than K-means Clustering on the synthetic breast UST phantom. In the this section, we will apply these two algorithms on the clinical images, 2-D digital mammogram and 3-D ultrasound tomography.

3.2.1 Clinical Images

As a commonly used screening and diagnostic tool, mammography is the process of using low-energy X-rays to examine the human breast. Like all X-rays, mammograms use doses of ionizing radiation to create images. Radiologists then analyze the images for any abnormal findings. The radiological appearance of breast tissue differs between individuals because of variations in breast tissue composition, and differences in the X-ray attenuation properties of fat, epithelium, and stroma. Fat appears dark on a mammogram, whereas epithelium and stroma appear light or white, an appearance that we refer to as mammographic density.

As shown in Figure 3.13, for the denser or fibro-glandular breast, the mammogram will look bright or cloudy; for the not dense or fatty breast, most of the fibrous tissue is replaced with fatty tissue, so the mammogram tends to look black or gray.

Compared with the 2-D view of digital mammography image, the 3-D ultrasound tomography could provide more detail and accurate map about the tissue architecture. Images shown below (Figure 3.14) is the operator-independent whole-breast ultrasound tomography prototype which is under clinical trails in Karmanos Cancer Institute [9] [10]. A patient exam begins with the patient lying prone on the scan-

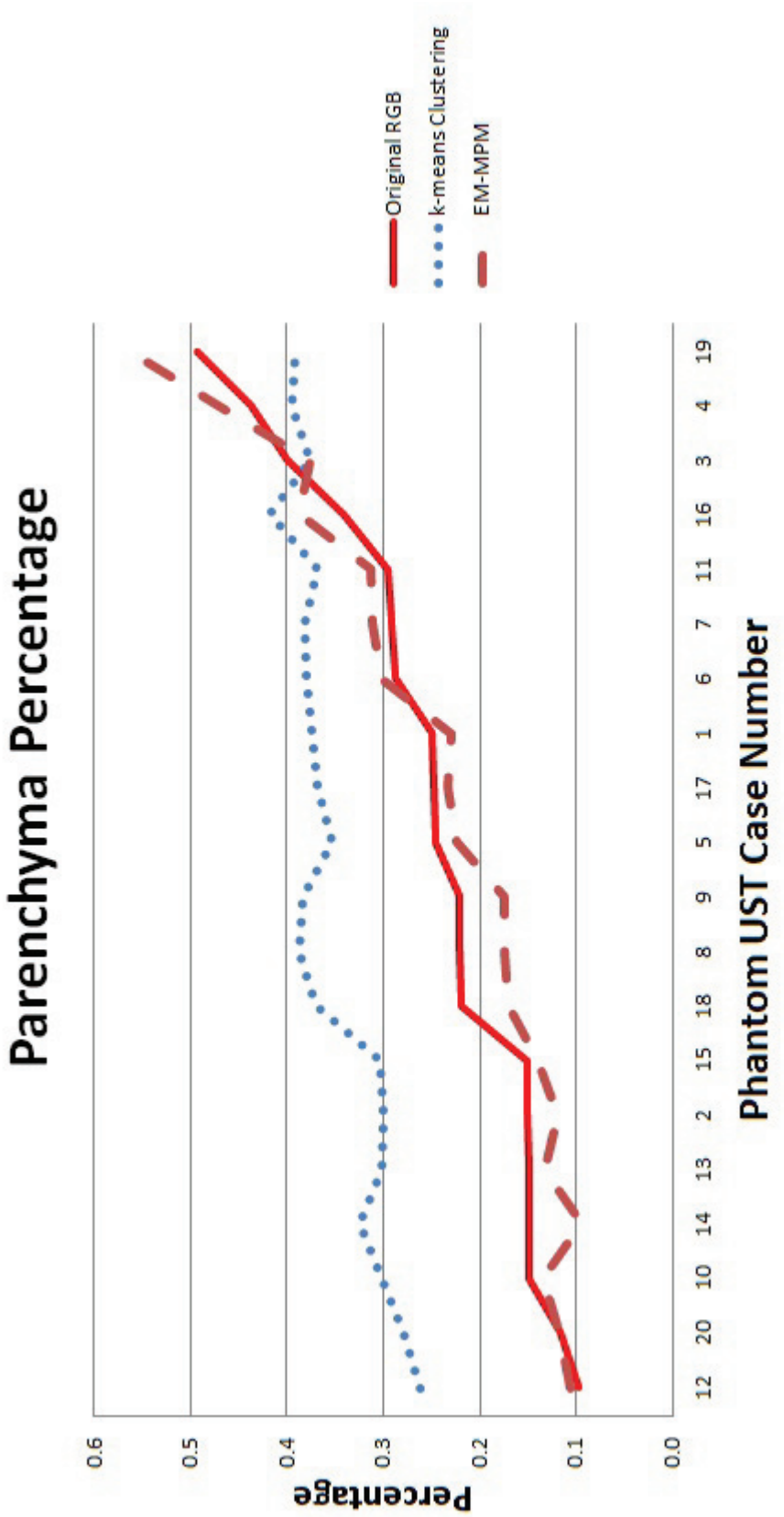
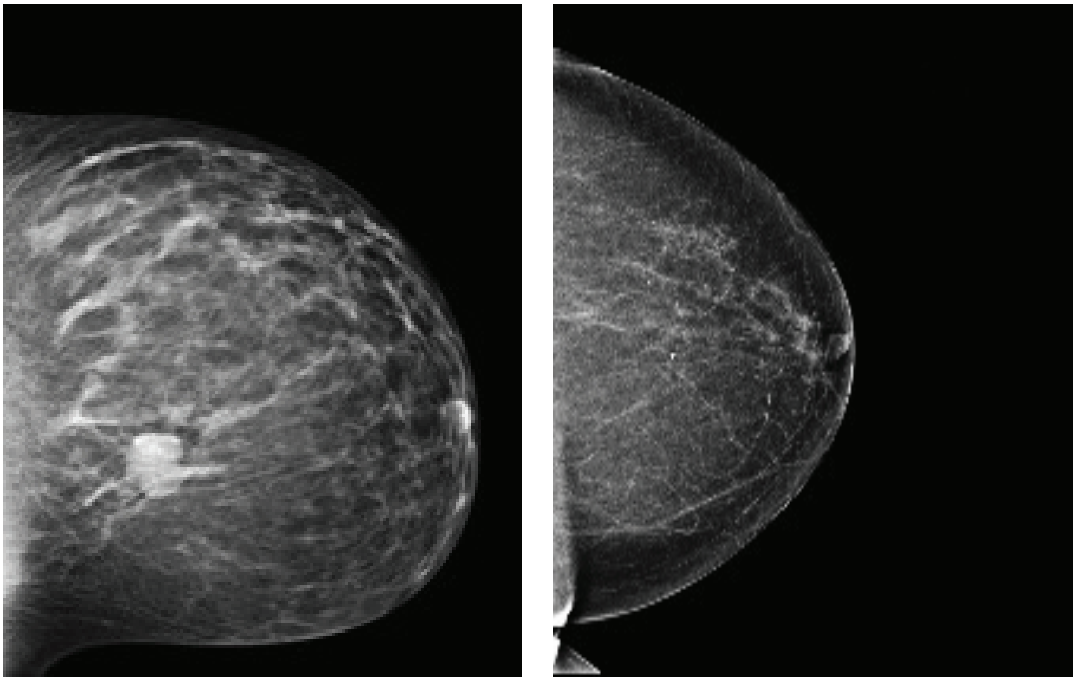


Fig. 3.12. Parenchyma Percentage Comparison of EM-MPM and K-means Clustering



Dense Not Dense

Fig. 3.13. Mammogram images of dense and not dense breast

ner table. The table consists of flexible sailcloth, which contours to the patient's body, thereby increasing access to the axillary regions of the breast and increasing patient comfort. The breast is suspended in the imaging tank that lies below the table, through a hole in the table. The imaging tank is filled with warm, clean water. The ultrasound sensor (Figure 3.15), in the shape of a ring, surrounds the breast and moves from the chest wall to the nipple region of the breast on a motorized gantry, gathering data along the way.

As we know, sound speed images are based on the arrival times of acoustic signals. Cancerous tumors have enhanced sound speed relative to normal breast tissue, a characteristic which is used on differentiation of masses, normal tissue, and fat [11]. Figure 3.16 shows two UST image slice, comparing the dense breast with the not dense one. In the first picture we can see the not dense breast looks smoothly dark gray

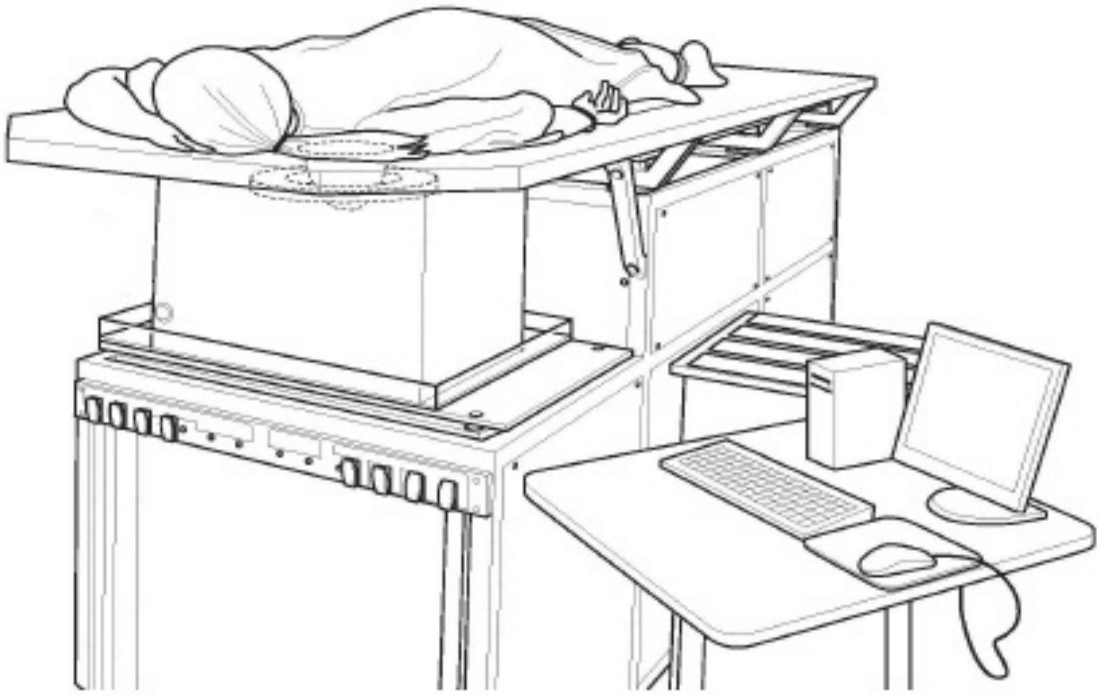


Fig. 3.14. The breast UST clinical prototype [9]

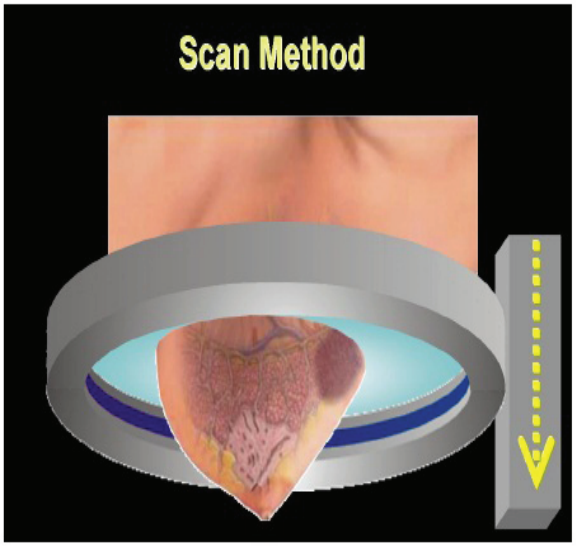
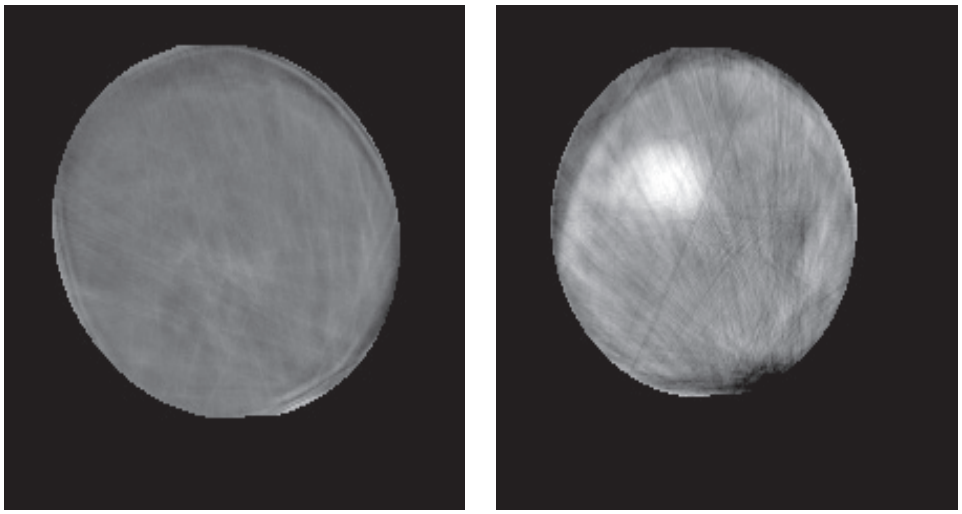


Fig. 3.15. The UST ring detection [9]

(fatty structure) in the tissue, only some sparse light spots (fibrous stroma) appears. In the dense UST image, we can see light gray (parenchyma) or white (tumor) area.



Not Dense Dense

Fig. 3.16. UST clinical images of dense and not dense breast

3.2.2 Result and Discussions

To test the performance of the segmentation algorithms, the breast density is calculated out of the 3-D UST segmentation result, then compared with the 2-D percent density value from digital mammogram. This percent density value from digital mammogram images is estimated by a clinical radiologist (with over 5 years of experience in mammography) using Cumulus 4.0, an interactive software package developed at the University of Toronto and validated in many studies [5]. Cumulus is based on manual exclusion of the pectoral muscle and interactive selection of thresholds for segmenting the breast outline and the regions of dense tissue. In our study, both volumetric breast density from UST and percentage density from Cumulus estimation, are computed as the ratio of the area corresponding to the dense tissue and the total mammographic breast tissue.

First, to see the algorithm performance on clinical data, we apply EM-MPM on nine digital mammogram cases, and directly compare the 2-D mammogram percentage density between EM-MPM and Cumulus estimation. To provide more density information within the breast, here we use four classes to describe the different density structures. The breast image cases involved in this study will vary based on volume and thickness, so the same tissues may appear as different brightness in the X-ray mammogram images. From the radiologist's perspective, the dense tissue is manually chosen and defined differently for breast mammogram images with different background brightnesses. For example, if the whole breast looks dark in the mammogram image (or say, it is small or thin), then any light brands or light gray area (fibrous stroma or parenchyma) will appear with higher contrast with the normal fatty tissues, and will be regarded as 'dense tissue'; If the breast is large and thick, it will appear comparatively brighter even for the fat tissue, so only structures which looking much brighter will be chosen as 'dense part'. Based on this consideration, our contribution is to compensate the percentage density calculation with different choice of the 'dense tissue' for the 2-D mammogram images:

(1) For low brightness mammograms, whose fatty tissue brightness is less than 35, we choose the top three EM-MPM bright classes as dense tissue.

(2) For middle brightness mammograms, whose fatty tissue brightness range within 35 and 50, we choose the top two EM-MPM bright classes as dense tissue.

(3) For high brightness mammograms, whose fatty tissue brightness is higher than 50, we choose the top one EM-MPM bright class as dense tissue.

Figure 3.17 illustrate a selection of mammogram images with different brightness and their corresponding EM-MPM segmentation result (with remaining cases in Appendix B). Here we can see, the EM-MPM segmentation image groups structures clearly according to their brightness (absolute density), which provide us clear and adequate information about the scattered fibrous stroma, parenchyma, and the tumor.

Figure 3.18 shows the 2-D mammogram percentage density comparison result from EM-MPM and Cumulus estimation. The nine mammogram cases are arranged by descending sequence of the density (according to Cumulus estimation, the currently used standard). Here we can see, EM-MPM taces the Cumulus estimation very well in most of the cases, which means EM-MPM could provide good information about the mammogram density.

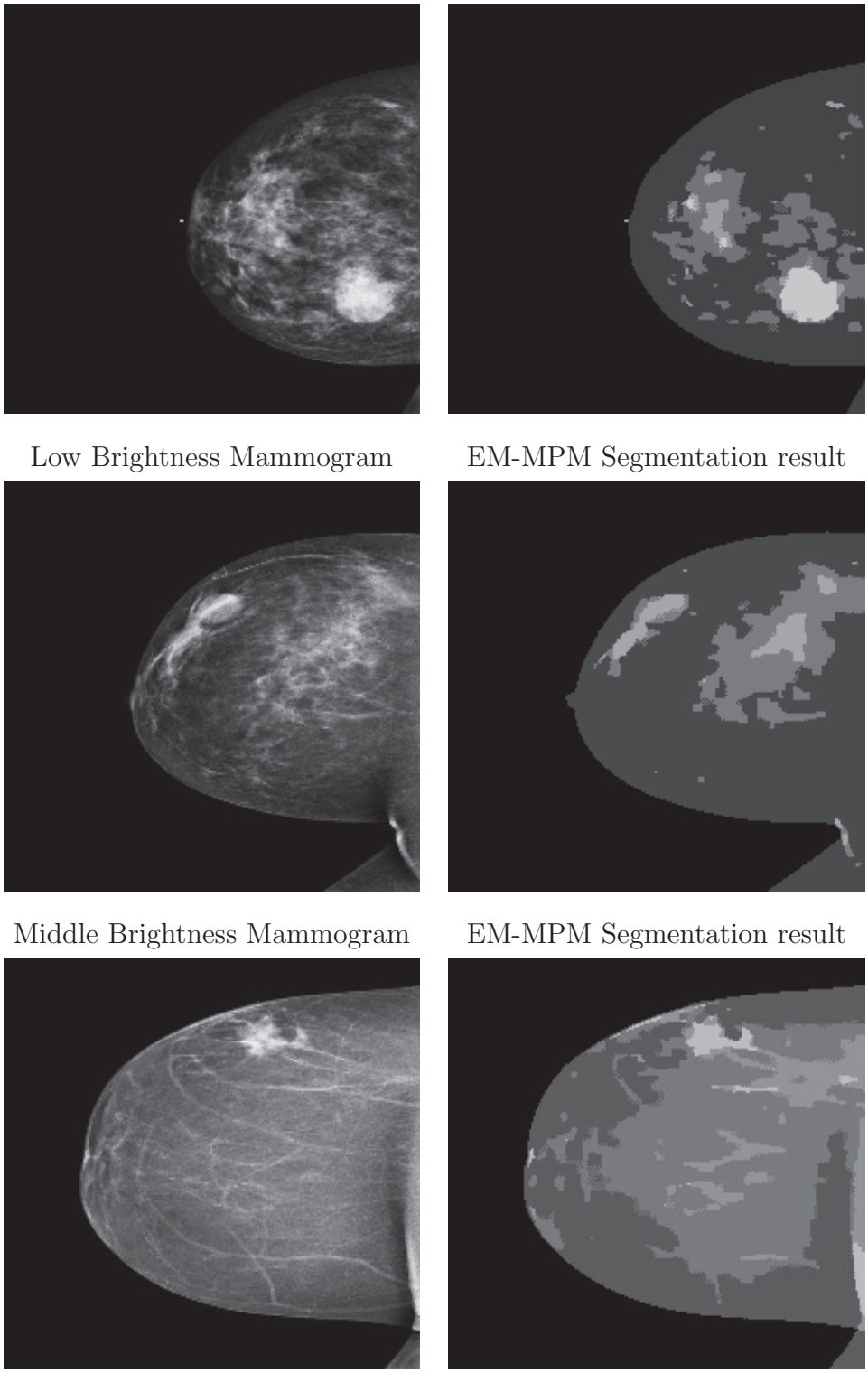
Next, we apply both EM-MPM and K-means Clustering on the nine 3-D breast UST cases, then compare the volumetric breast density calculated from the segmentation result with the percentage density Cumulus estimation which is performed on the corresponding 2-D mammogram images. In UST images the object density under detection is automatically calibrated to water density, it means all cases are calibrated with a common density standard. Therefore different breast size is no longer an issue. The brightness is directly proportional to the true volume density.

Similar to the synthetic phantom Parenchyma Percentage calculation, here the clinical UST dense part is accumulated slice by slice then divided by the 3-D breast volume to yeild the volume based percent density. For K-means Clustering, since only two classes segmentation is performed, so the top class was chosen as the dense part. For EM-MPM segmentation, except two very dense cases (number 6 and number 7 only the top class is chosen) we choose the top two classes as the dense part.

Figure 3.19 presents the total nine cases percentage density comparison between EM-MPM, K-means Clustering and Cumulus estimation. The cases are arranged by descending sequence according to Cumulus estimation result. From the illustration we can see, both EM-MPM and K-means have some points deviate away from the Cumulus standard, which represents an inconclusive comparison result.

To shown things more clearly, we choose five cases out of nine (Figure 3.20 to Figure 3.24) to show the detailed segmentation result comparison, with remaining cases in Appendix C. Case #5 (Figure 3.20) and and #9 (Figure 3.21) are two cases that both EM-MPM and K-means Clustering matches very well with the Cumulus standard (judeged from Figure 3.19). But we can see EM-MPM clusters the dense

tissue (tumor and the surrounding parenchyma structure) more clearly than K-means Clustering. Figure 3.22 shows the case #1 which EM-MPM fits well with Cumulus but K-means Clustering failed. We can see both of them find where the tumor appears, but EM-MPM identified larger high density area compared with K-means. Figure 3.23 and Figure 3.24 are two cases, #6 and #7, who have whole-breast higher density compared with others, which could be recognized by the brighter looked UST images. In these two cases, the top two classes (tumor part and parenchyma structure) in EM-MPM took over very large part of the tissue, but in K-means Clustering, since only two classes are chose for segmentation, so only the tumor part is marked as the dense tissue.



Low Brightness Mammogram EM-MPM Segmentation result

Middle Brightness Mammogram EM-MPM Segmentation result

High Brightness Mammogram EM-MPM Segmentation result

Fig. 3.17. Mammogram images and the corresponding EM-MPM segmentation result

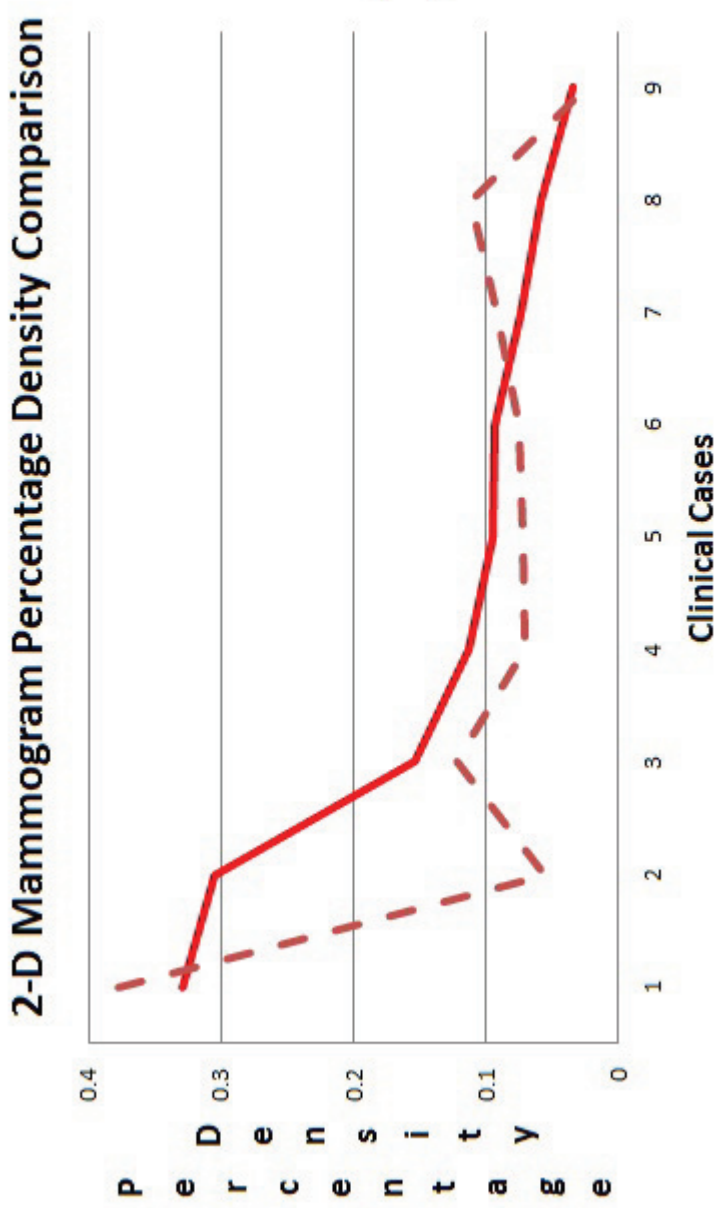


Fig. 3.18. Mammogram Percentage Density Comparison of EM-MPM and Cumulus Estimation

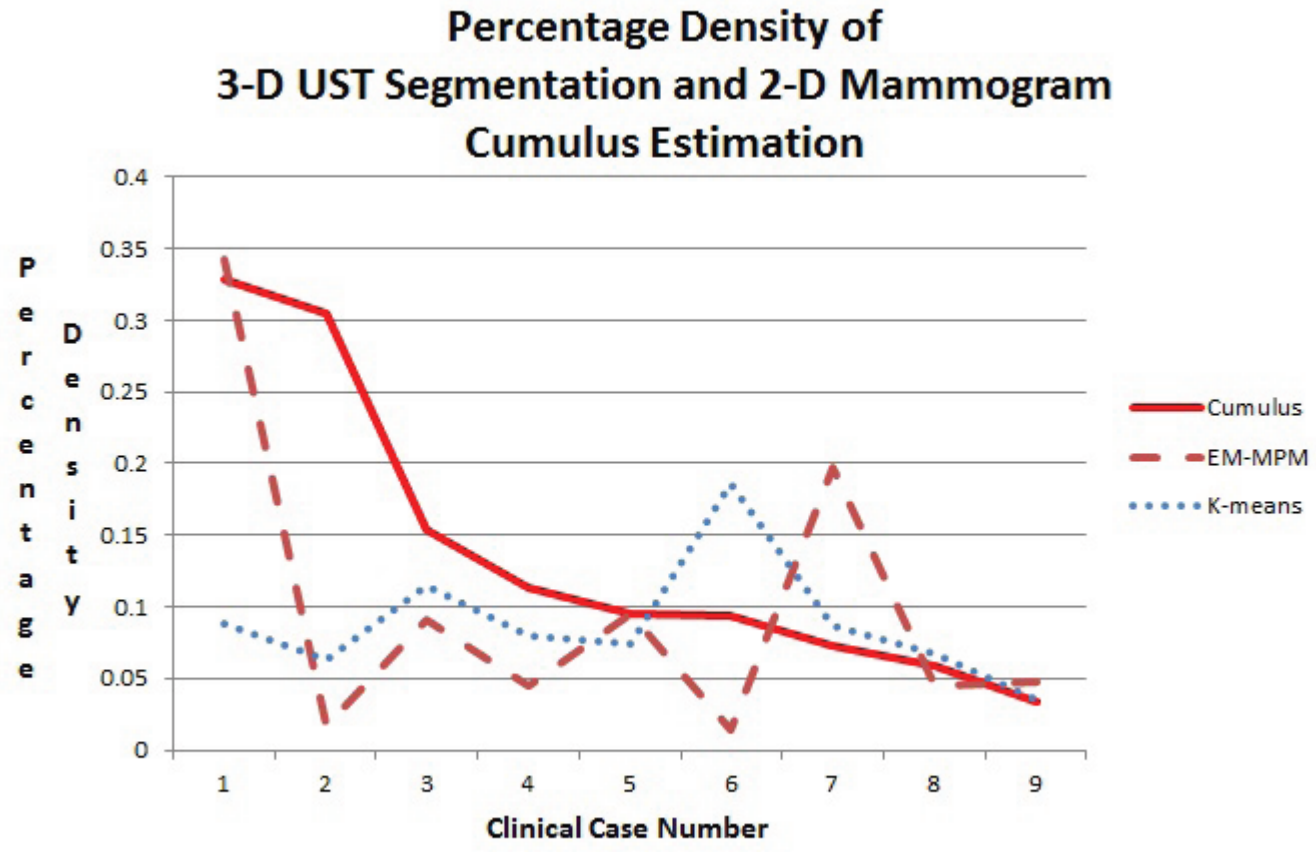


Fig. 3.19. Percentage Density Comparison of EM-MPM, K-means Clustering and Cumulus Estimation

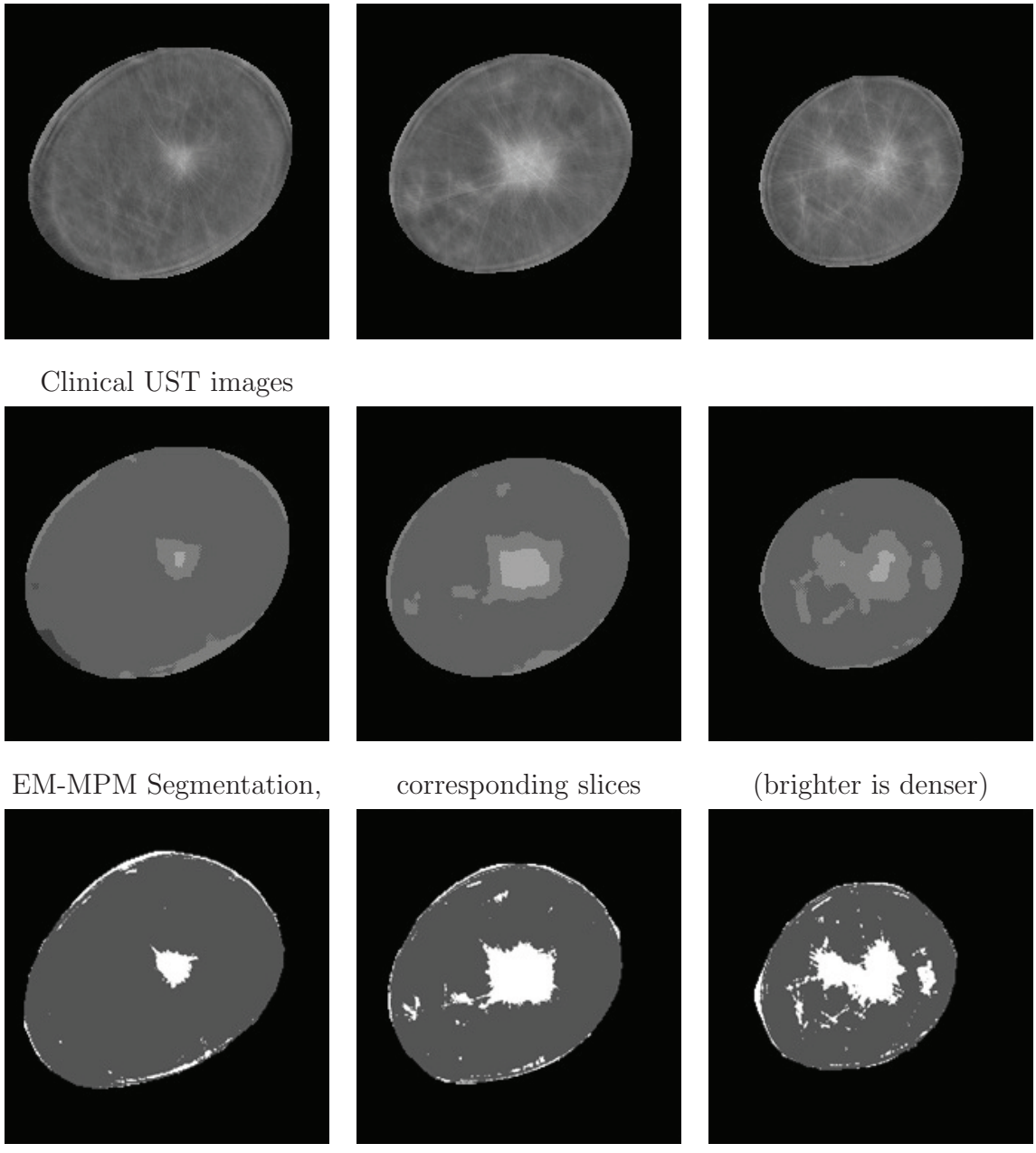


Fig. 3.20. Comparison between EM-MPM and K-means Clustering on clinical UST, case #5. (In this case both EM-MPM and K-means Clustering match the Cumulus estimation)

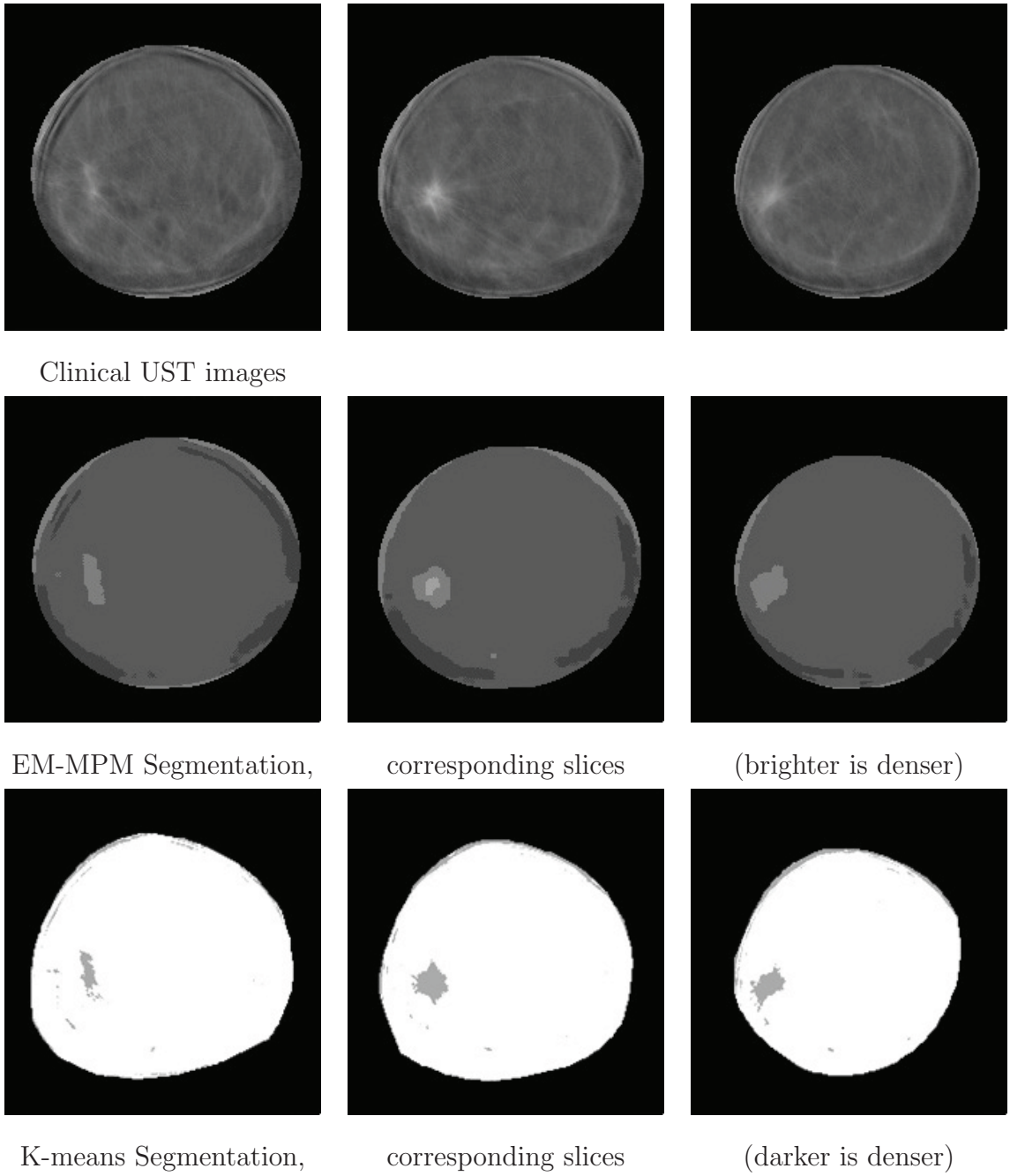
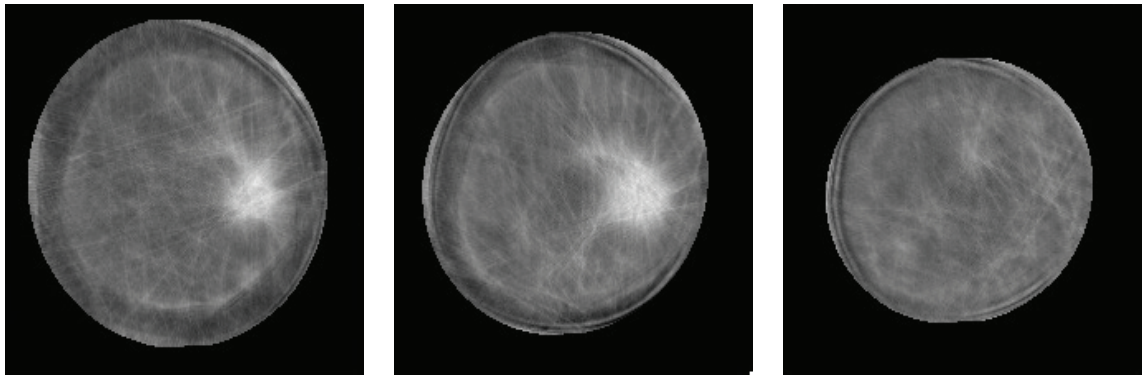
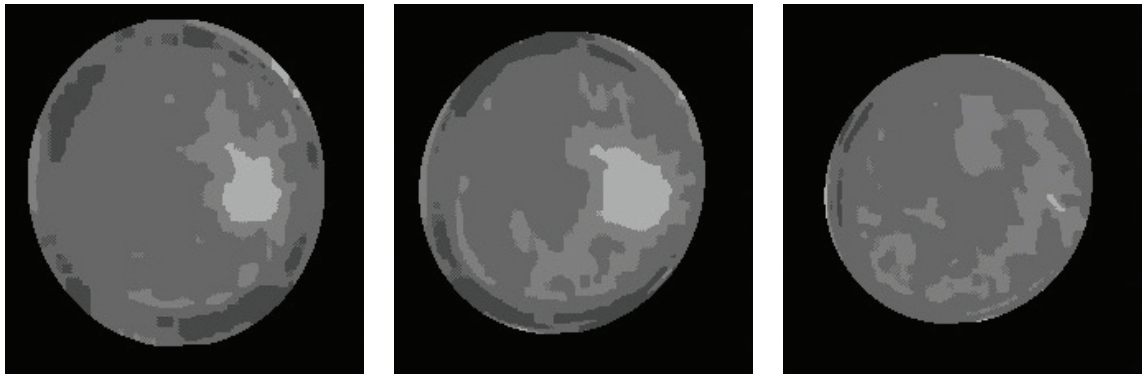


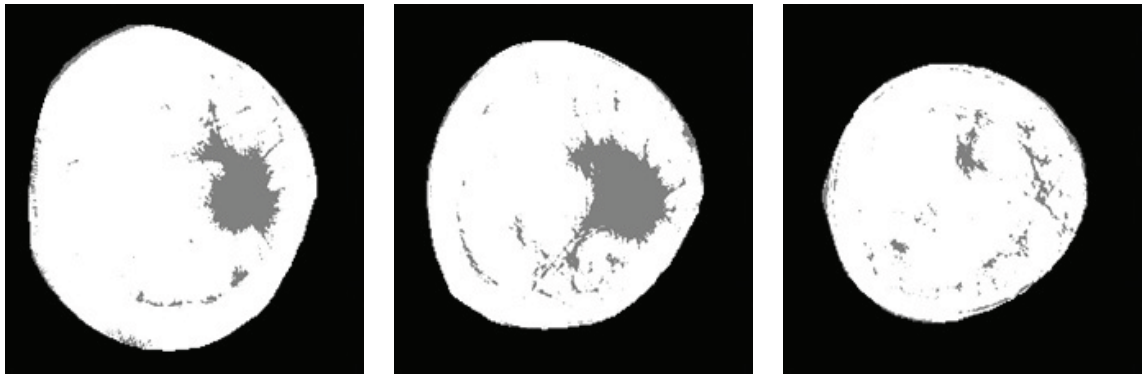
Fig. 3.21. Comparison between EM-MPM and K-means Clustering on clinical UST, case #9. (In this case both EM-MPM and K-means Clustering match the Cumulus estimation)



Clinical UST images

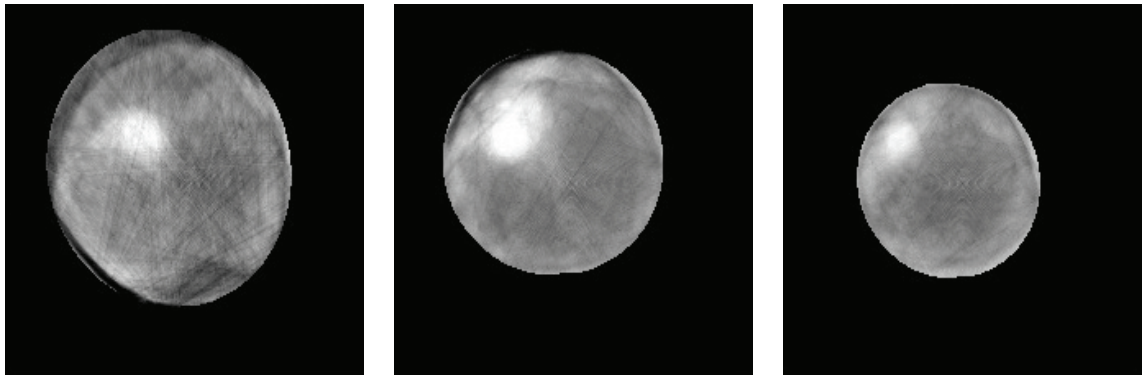


EM-MPM Segmentation, corresponding slices (brighter is denser)

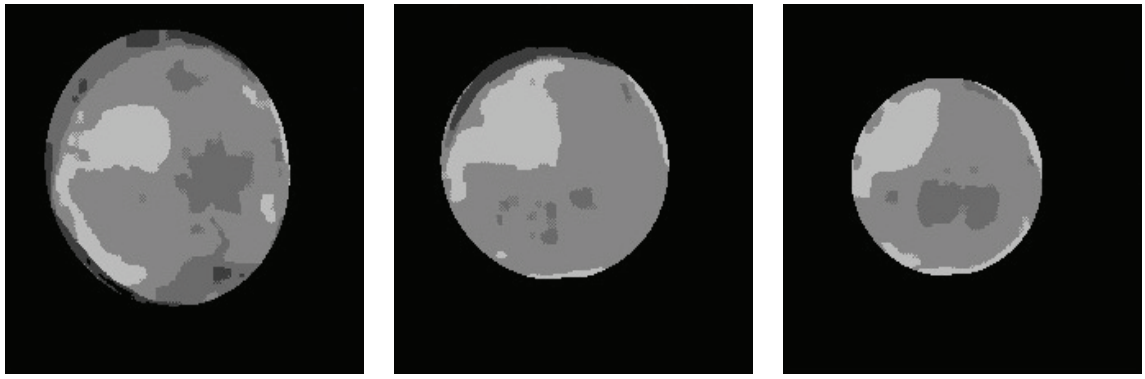


K-means Segmentation, corresponding slices (darker is denser)

Fig. 3.22. Comparison between EM-MPM and K-means Clustering on clinical UST, case #1. (In this case EM-MPM matches very well with Cumulus but K-means Clustering failed)



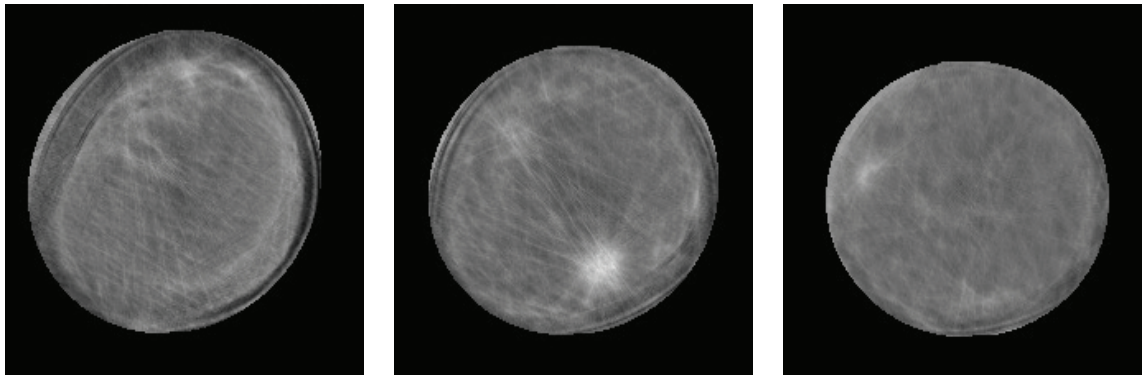
Clinical UST images



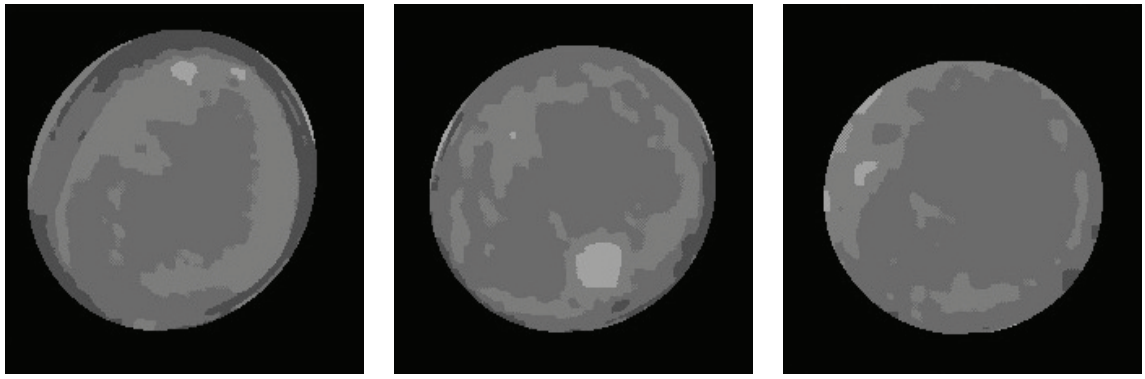
EM-MPM Segmentation, corresponding slices (brighter is denser)

K-means Segmentation, corresponding slices (brighter is denser)

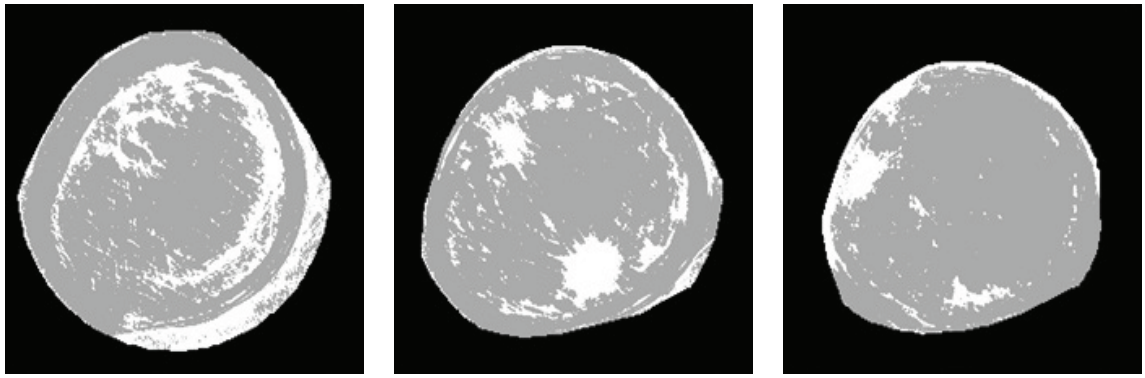
Fig. 3.23. Comparison between EM-MPM and K-means Clustering on clinical UST, case #7. (High density case, EM-MPM failed to match with Cumulus, but K-means Clustering does)



Clinical UST images



EM-MPM Segmentation, corresponding slices (brighter is denser)



K-means Segmentation, corresponding slices (brighter is denser)

Fig. 3.24. Comparison between EM-MPM and K-means Clustering on clinical UST, case #6. (High density case, both EM-MPM and K-means Clustering failed to match with Cumulus)

4. CONCLUSION AND FUTURE WORK

4.1 Conclusion

In this thesis, we compared two segmentation algorithms, EM-MPM and K-means Clustering. They are tested on synthetic phantom UST data, clinical mammogram data and UST image data. From the comparison of segmentation pictures we can see, EM-MPM finds the dense tissues accurately regardless of the limited resolution and scattered dense parts, which are clearly shown in the superior matching result compared with the ground truth phantom and Cumulus estimation. However, K-means Clustering could only handle the low density, clearly clustered cases as shown in the phantom UST segmentation result, but failed on the cases with scattered high dense or low dense cases. This is because the EM-MPM algorithm pays more attention to the local neighborhood choices than K-means. EM-MPM classifies every pixel in the image by assigning a cost to the number of misclassified pixels, and iteratively finds the best probabilistic solution to fit the data. This has the advantage of using a 3-D neighborhood of pixels as a statistical Bayesian prior, and it has the effect of grouping the data similar to the way the tissues are structured. In contrast, K-means focus on the cluster centroids. When the target classes are sparsely scattered instead of tightly clustered, K-means tends to group pixels according to the geometric centroid distances, without consideration about the local neighbors.

In Chapter 1, we introduced that breast density proportion is a strong risk factor for breast cancer risk, the early detection of breast density would help prevent breast cancer at an early stage. Among many popular methods used for analyzing mammographic density, we see that UST is an effective method which could provide accurate 3-D breast density map. In order to obtain a credible measure of density proportion from Ultrasound Tomography, we choose two robust segmentation algorithms, EM-

MPM and K-means Clustering, which, from previous studies, have been verified with stable performance on medical images containing added noise and limited resolution.

In Chapter 2, we briefly reviewed the EM-MPM and K-means Clustering algorithms. It presented a briefly analysis of the algorithms, and provides an understanding of the differences between these two methods. EM-MPM algorithms are discussed here, also the relationship between them is explained. For the K-means Clustering, the algorithm are reviewed step by step, accompanied with equations and examples.

In Chapter 3, the segmentation results from the two algorithms are compared and discussed. First we introduced the anthropomorphic software breast phantom which is obtained from taking 3-D tomographic ultrasound. Then we compared the phantom UST segmentation result of EM-MPM and K-means Clustering. EM-MPM performs very well on various kinds of density cases, and the segmentation result matches with the ground truth phantom (with Tanimoto Coefficient around 0.5 in the high density cases, and Parenchyma Percentage fits very well with the ground truth data). However K-means Clustering works on clustered dense cases, but does not perform well on scattered dense cases. After testing algorithm performance on the synthetic data, we applied them to clinical mammogram data. The clinical application of the X-ray mammography and ultrasound tomography prototype are explained. Then we compared the 3-D volumetric breast density from UST segmentation with the 2-D digital mammogram percentage from Cumulus estimation. Both of them have some mis-matched points compared to Cumulus. For this small sample, the comparison is inconclusive. However, subjectively judged from the segmentation images, EM-MPM shows more accurate dense tissue identification than K-means.

4.2 Future Work

Only twenty synthetic phantoms and nine clinical images have been tested. More cases are needed to verify the segmentation performance, especially for the clinical cases.

At present the EM-MPM could only provide the segmentation information based on the absolute gray level, but from radiologists' perspective the dense part might need to be considered according to the whole breast density. So our next step is to work with Karmanos Cancer Institute on more clinical cases. The initial results are promising, and the future work will be performing a statical distribution analysis to see the internal relationship between breast density, 2-D mammogram Cumulus estimation and 3-D UST EM-MPM segmentation. EM-MPM shows significant promise in providing accurate volumetric percentage density.

LIST OF REFERENCES

LIST OF REFERENCES

[1] N. F. Boyd, J. M. Rommens, K. Vogt, V. Lee, and J. L. Hopper, "Mammographic breast density as an intermediate phenotype for breast cancer," *The Lancet Oncology*, vol. 6, pp. 798–808, 2005.

[2] L. A. Christopher, E. J. Delp, and P. L. Carson, "3-D bayesian ultrasound breast image segmentation using the EM-MPM algorithmy," *Proceedings of the IEEE Symposium on Biomedical Imaging*, 2002.

[3] L. A. Christopher and E. J. Delp, "New approaches in 3D ultrasound segmentation," *Proceedings SPIE and IST Electronic Imaging and Technology Conference*, 2003.

[4] M. L. Comer and E. J. Delp, "The EM-MPM algorithm for segmentation of textured images: Analysis and further experimental results," *IEEE Transactions on Image Processing*, 2000.

[5] P. R. Bakic, C. Li, E. West, and M. Sak, "Comparison of 3D and 2D breast density estimation from synthetic ultrasound tomography images and digital mammograms of anthropomorphic software breast phantoms," *Medical Physics*, 2008.

[6] G. Q. Wu, X. G. Hu, J. Zhang, L. Li, X. Wu, and H. G. Li, "K-means clustering with bagging and map reduce," *IEEE Conference on System Sciences*, 2001.

[7] D. J. Rogers and T. T. Tanimoto, "A computer program for classifying plants," *Science*, vol. 132, pp. 1115–1118, 1960.

[8] P. R. Bakic, C. Zhang, and A. Maidment, "Development and characterization of an anthropomorphic breast software phantom based upon reagon-growing algorithm," *Medical Physics*, vol. 38, pp. 3165–3176, 2011.

[9] N. Duric, P. Littrup, C. Li, and S. Schmidt, "In-vivo imaging results with ultrasound tomography: report on an ongoing study at the karmanos cancer institute," *Medical imaging 2010 ultrasonic imaging tomography and therapy*, vol. 7629, 2010.

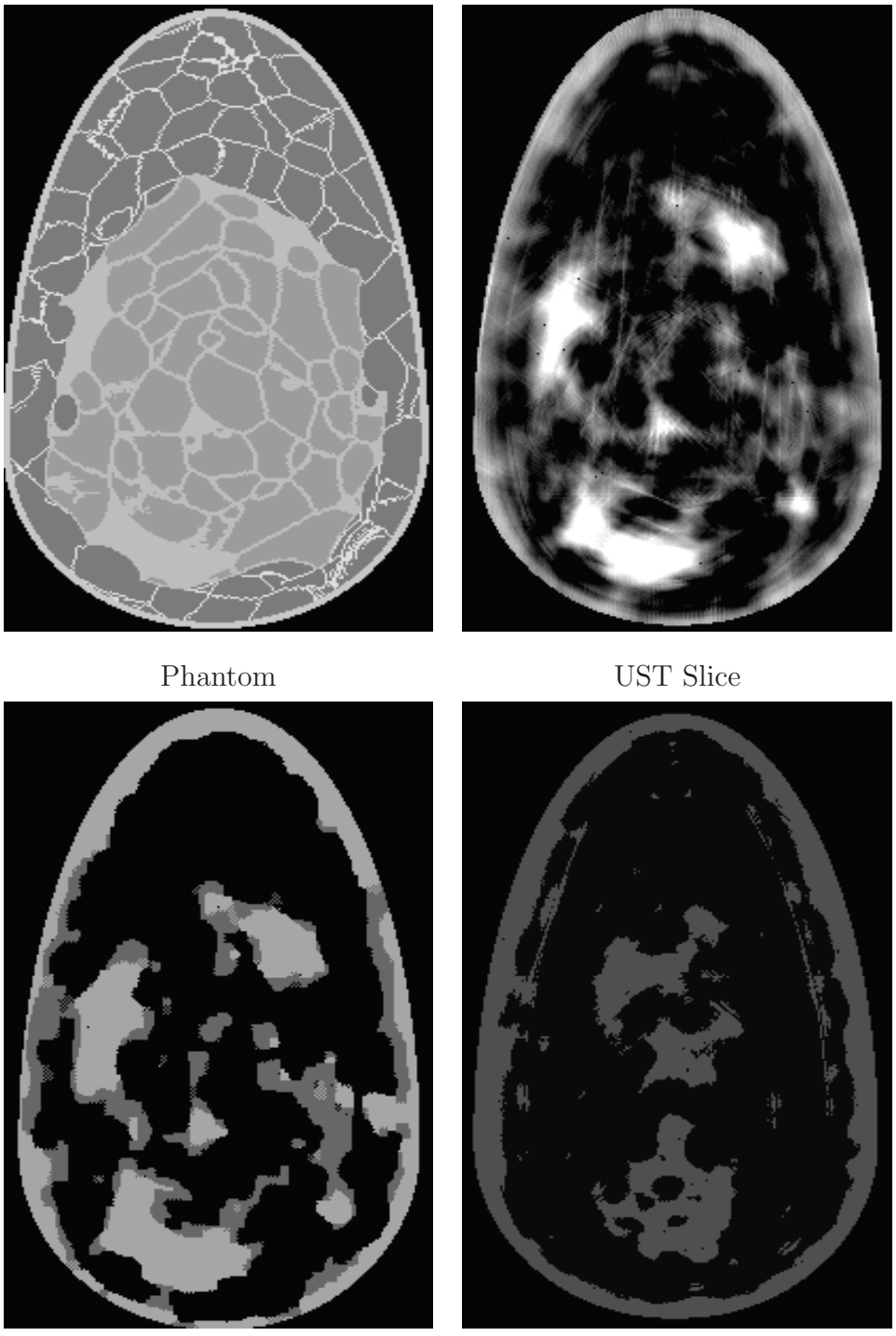
[10] C. K. G. Hurst, N. Duric, and P. Littrup, "Volumetric breast density evaluation from ultrasound tomography images," *Medical Physics*, vol. 35, 2008.

[11] C. Glide, N. Duric, and P. Littrup, "Novel approach to evaluating breast density utilizing ultrasound tomography," *Medical Physics*, vol. 34, 2007.

APPENDICES

A. SEGMENTATION COMPARISON ON PHANTOM UST IMAGES

In this section we present the remaining phantom UST cases segmentation between EM-MPM and K-means Clustering. We use three class option for EM-MPM, and two class option is used for K-means Clustering in the work of Karmonos Cancer Institute. The top brightness class in the segmentation image is regarded as the dense tissue, then we compare it with the original ground truth phantom. It shows clearly that EM-MPM has better matching result than K-means Clustering. In various kind of density cases, EM-MPM accurately finds the dense part within the phantom, and match very well with the ground truth phantom. However, as mentioned in previous Section 3.1.4, K-means Clustering result has higher density percentage in low density cases, but lower density percentage in high density cases.



Phantom | UST Slice

EM-MPM | K-means Clustering

Fig. A.1. Comparison between EM-MPM and K-means Clustering on phantom UST, case #5.

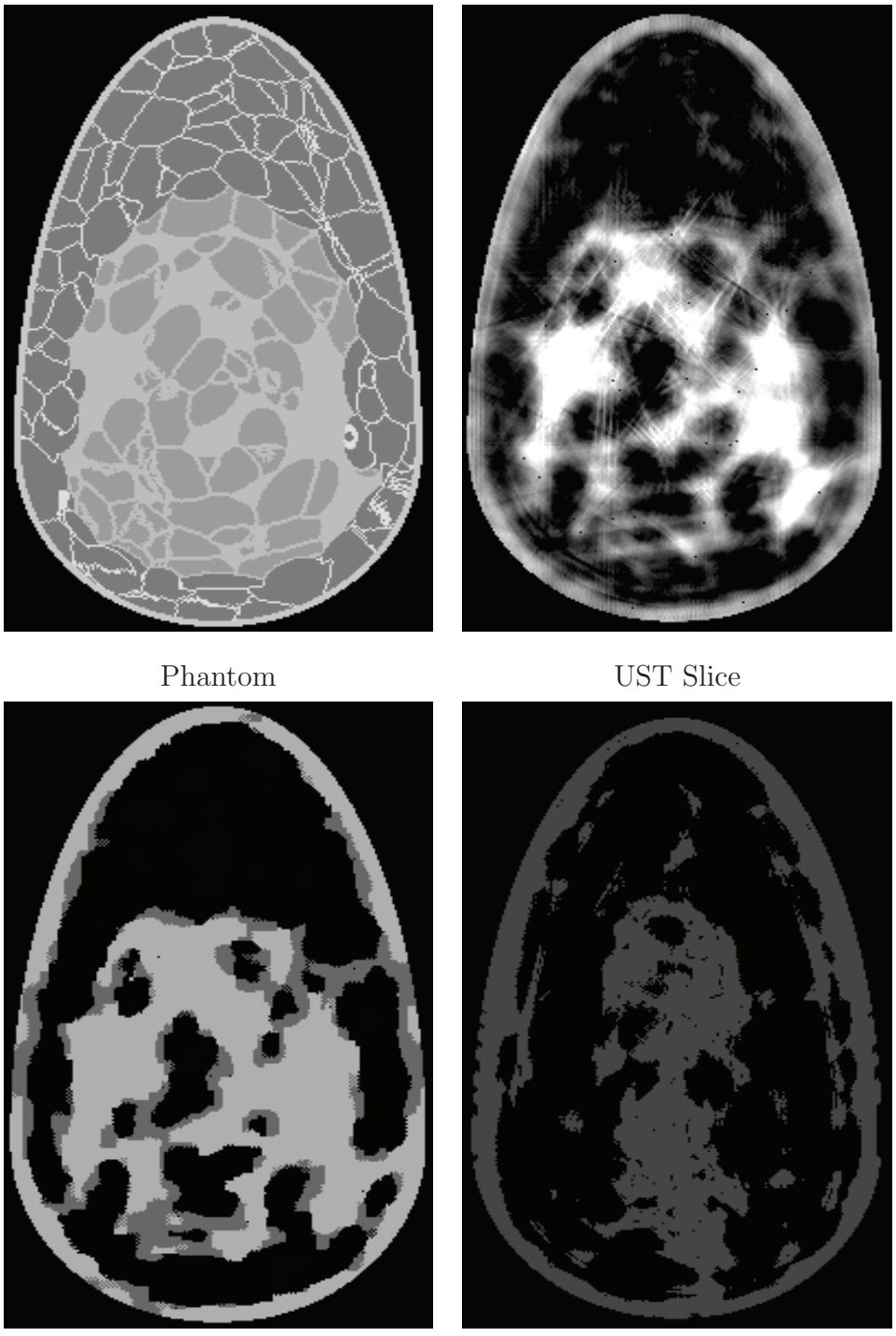


Fig. A.2. Comparison between EM-MPM and K-means Clustering on phantom UST, case #6.

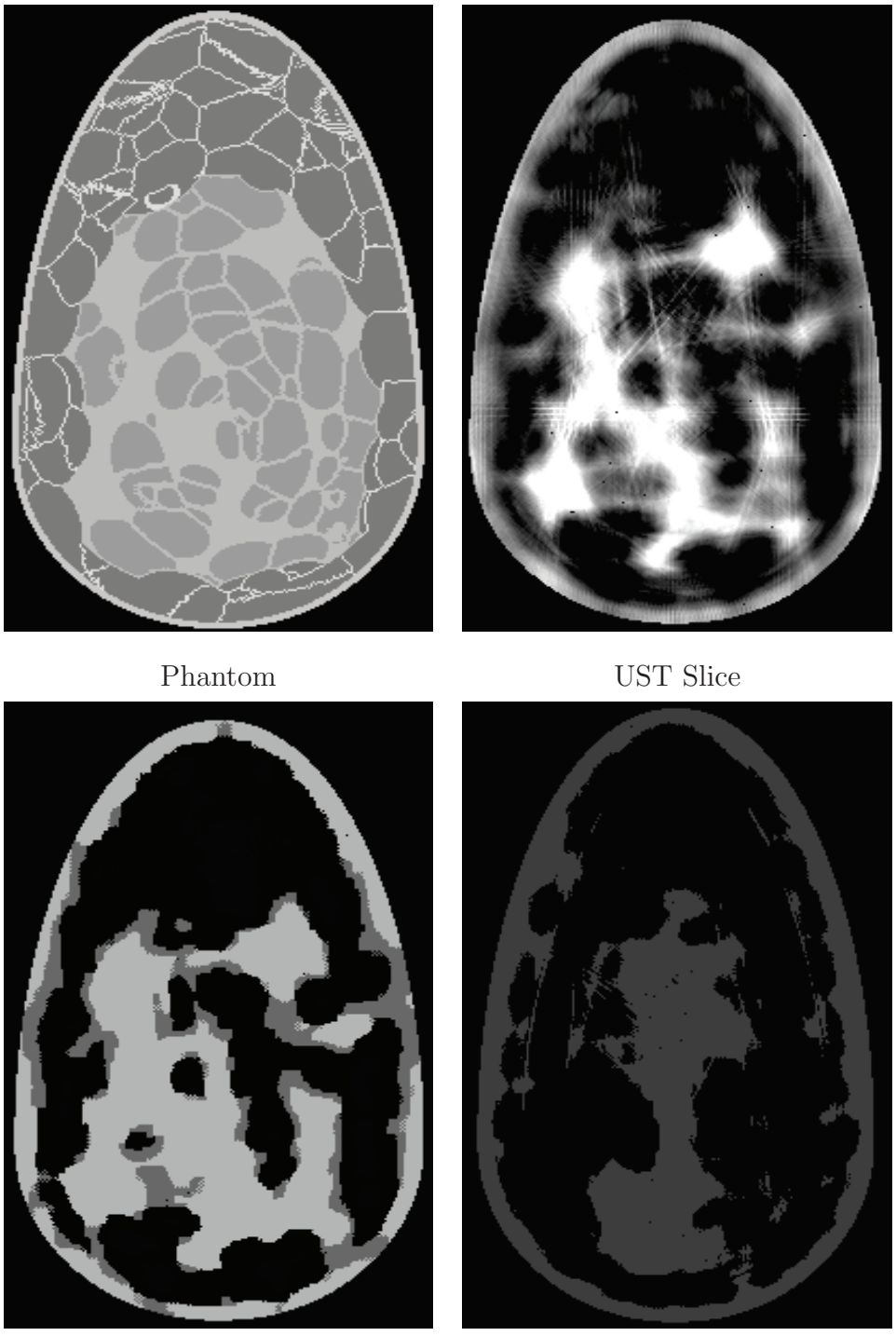
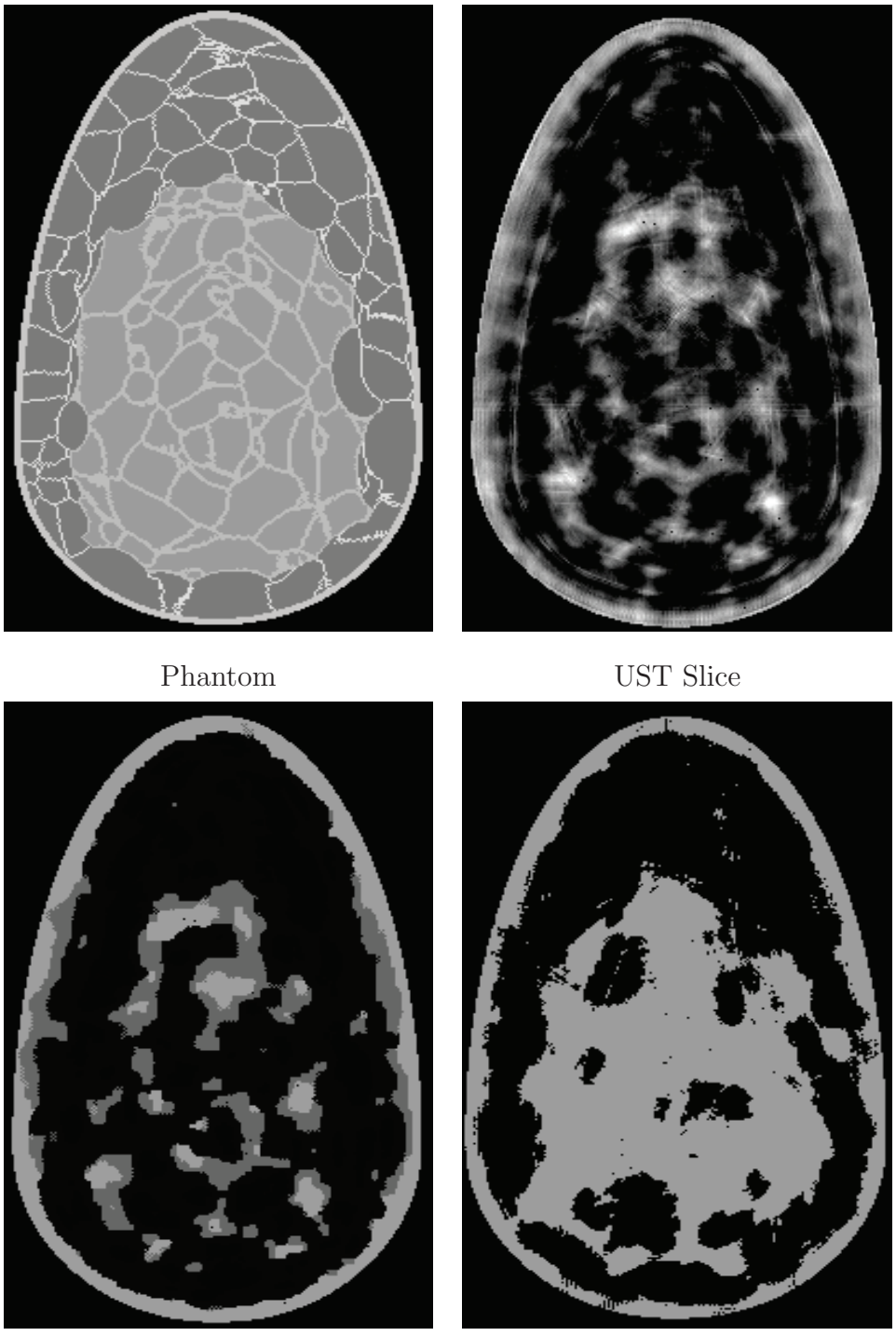


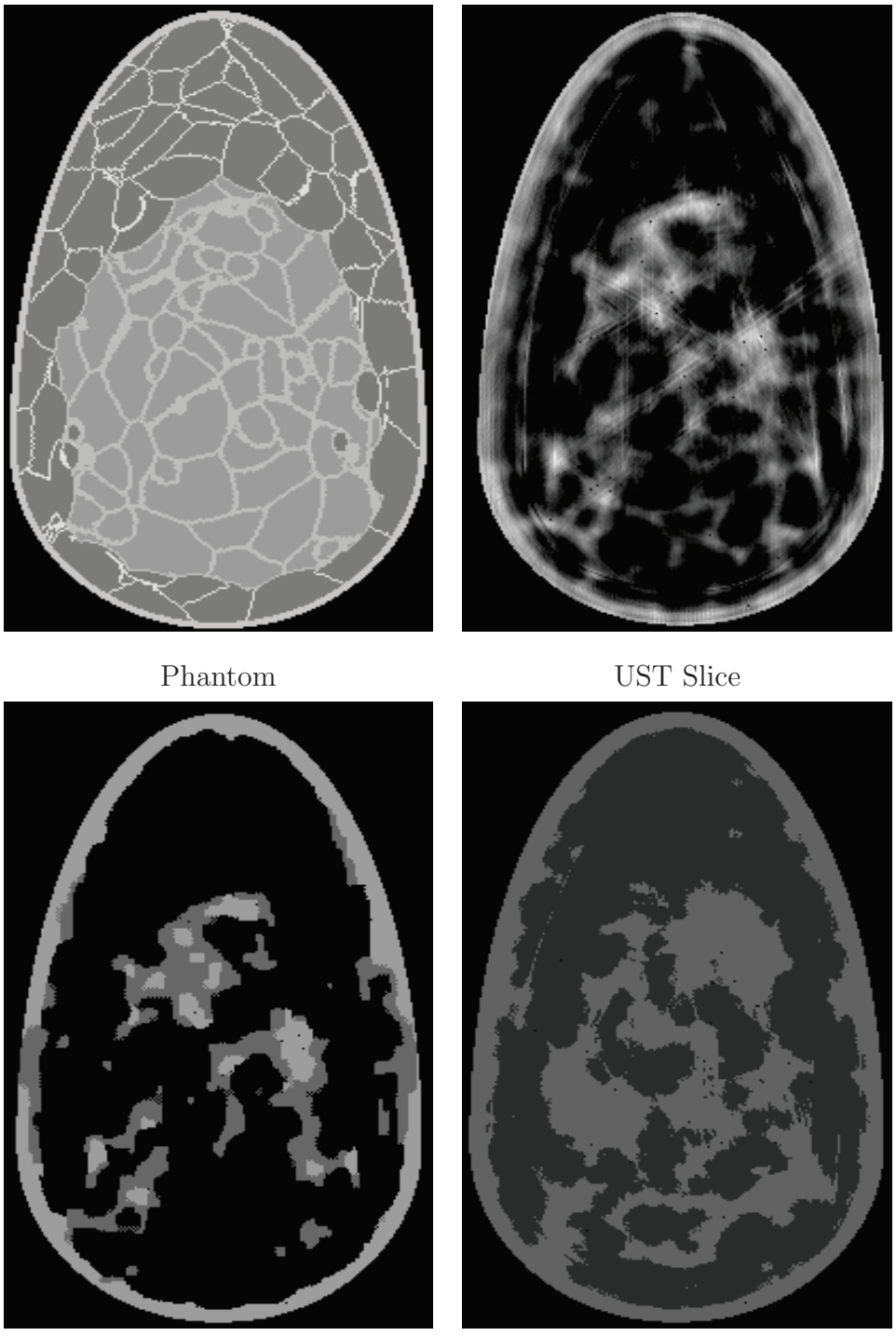
Fig. A.3. Comparison between EM-MPM and K-means Clustering on phantom UST, case #7.



Phantom | UST Slice

EM-MPM | K-means Clustering

Fig. A.4. Comparison between EM-MPM and K-means Clustering on phantom UST, case #8.



Phantom

UST Slice

EM-MPM

K-means Clustering

Fig. A.5. Comparison between EM-MPM and K-means Clustering on phantom UST, case #9.

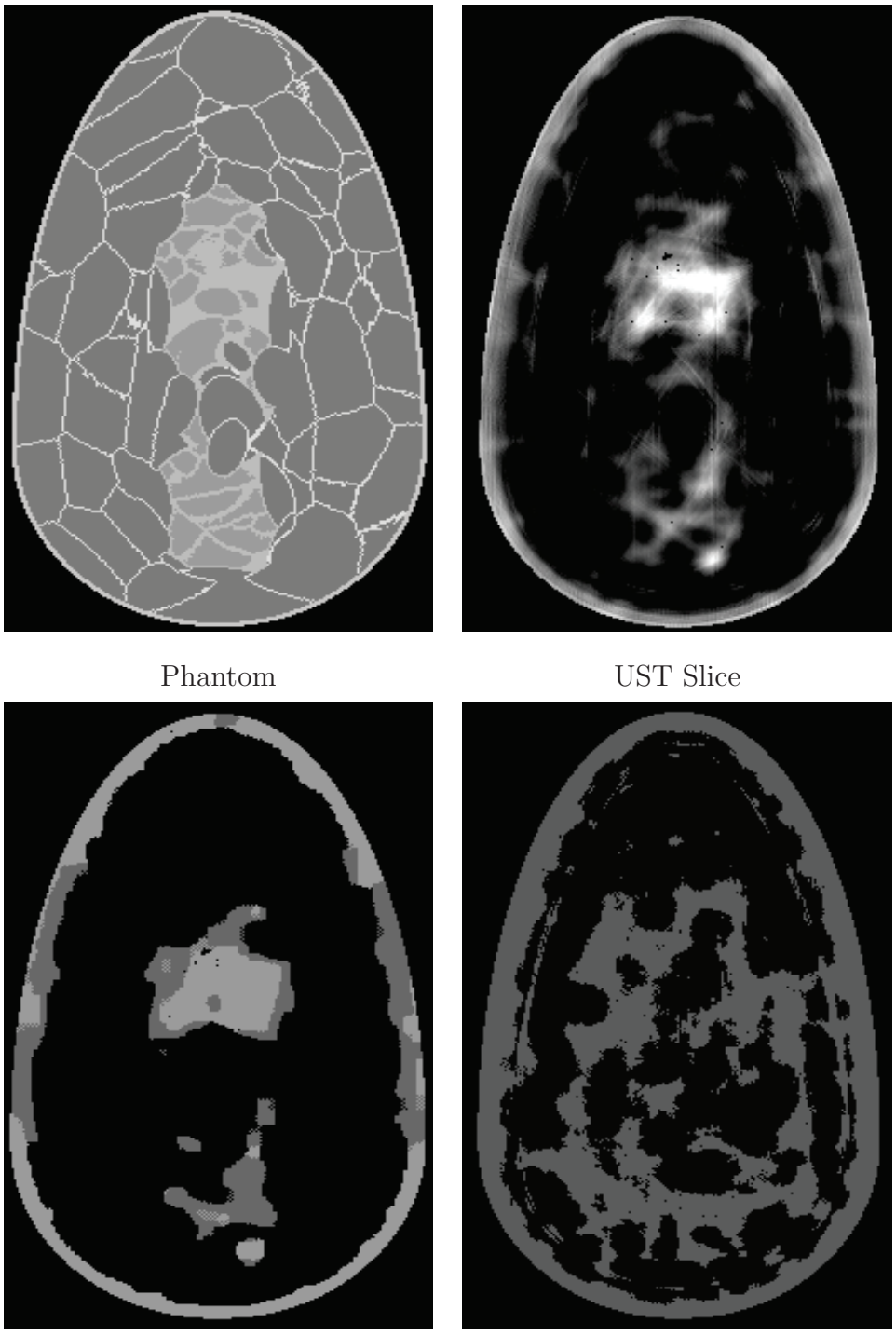
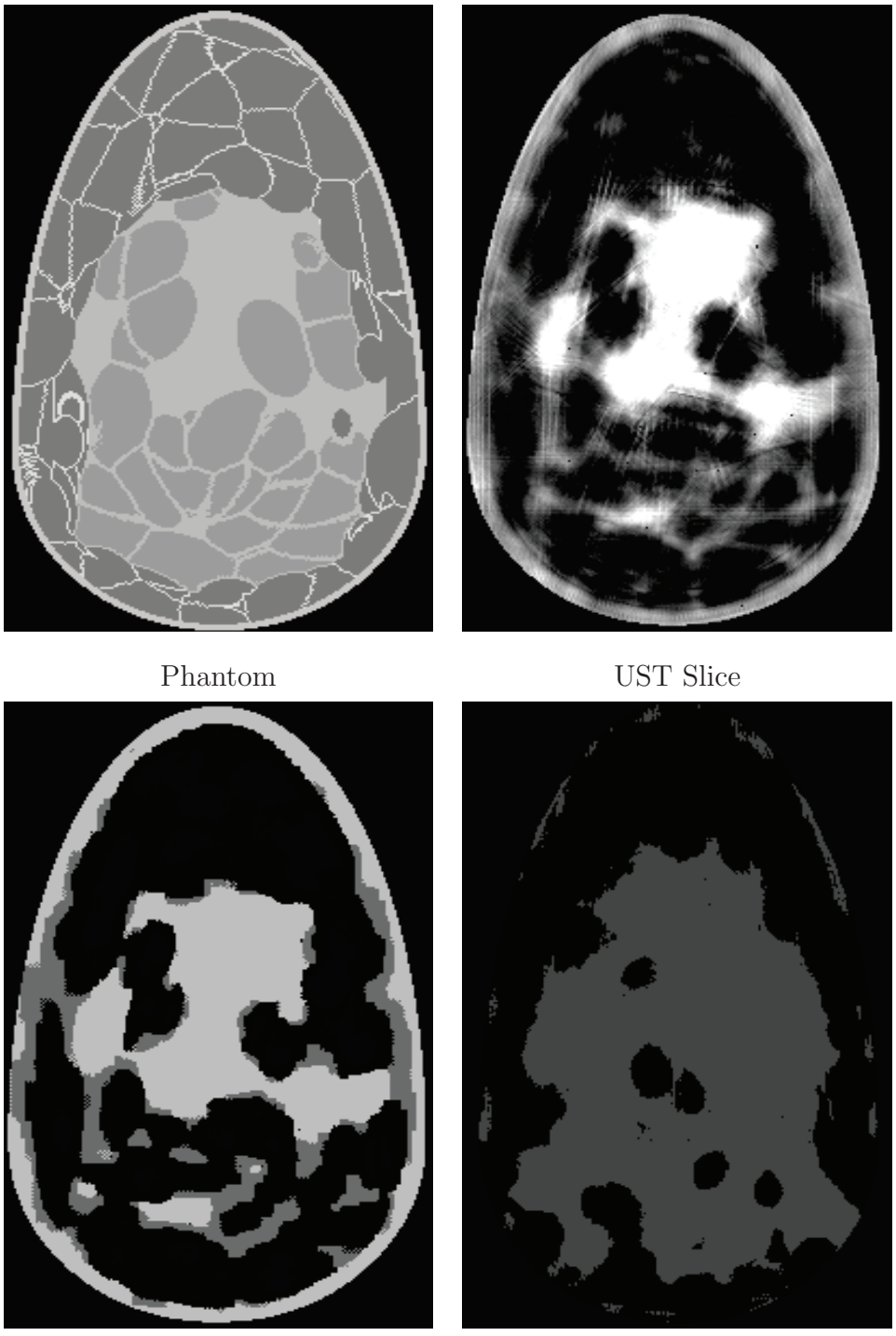


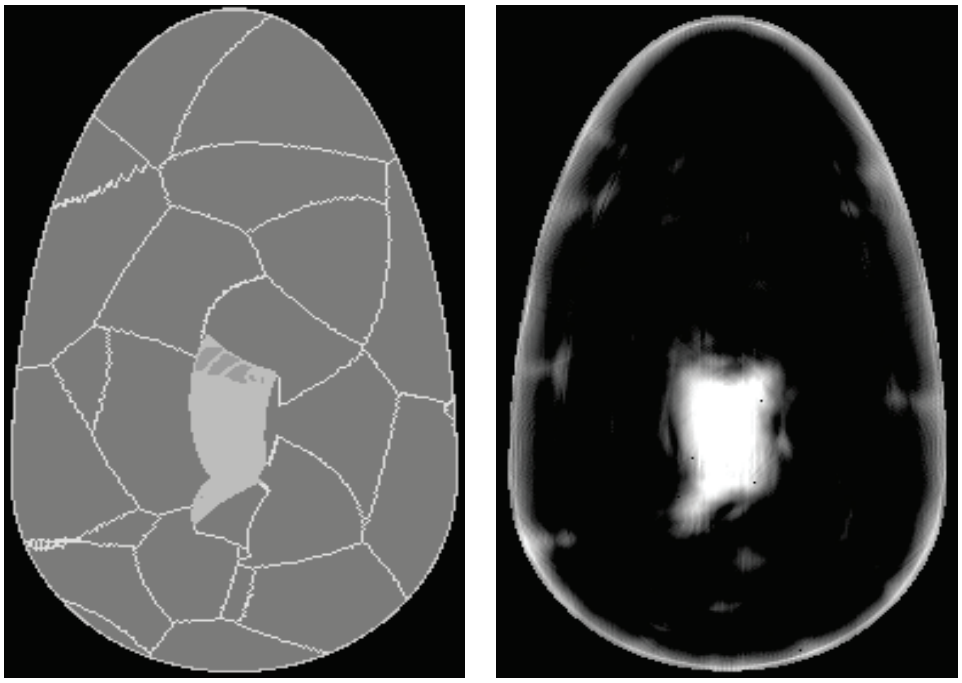
Fig. A.6. Comparison between EM-MPM and K-means Clustering on phantom UST, case #10.



Phantom | UST Slice

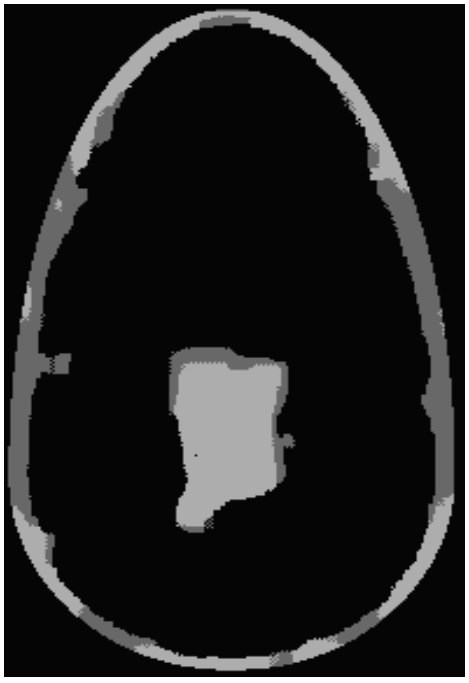
EM-MPM | K-means Clustering

Fig. A.7. Comparison between EM-MPM and K-means Clustering on phantom UST, case #11.



Phantom

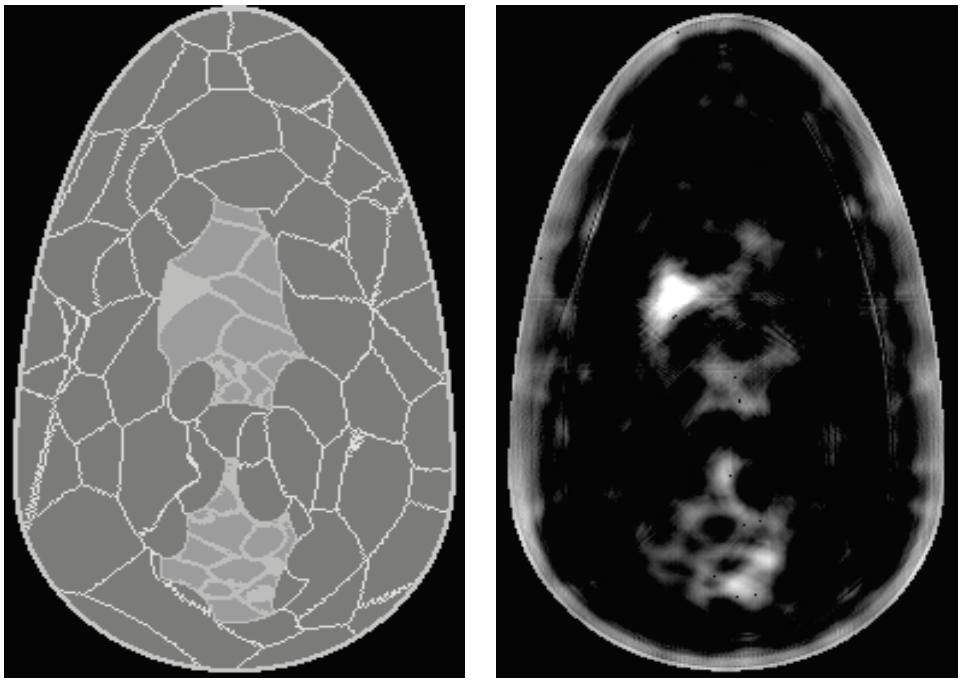
UST Slice



EM-MPM

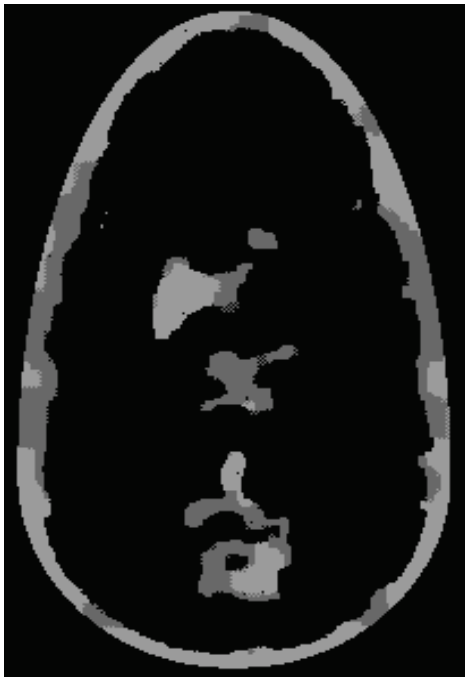
K-means Clustering

Fig. A.8. Comparison between EM-MPM and K-means Clustering on phantom UST, case #12.



Phantom

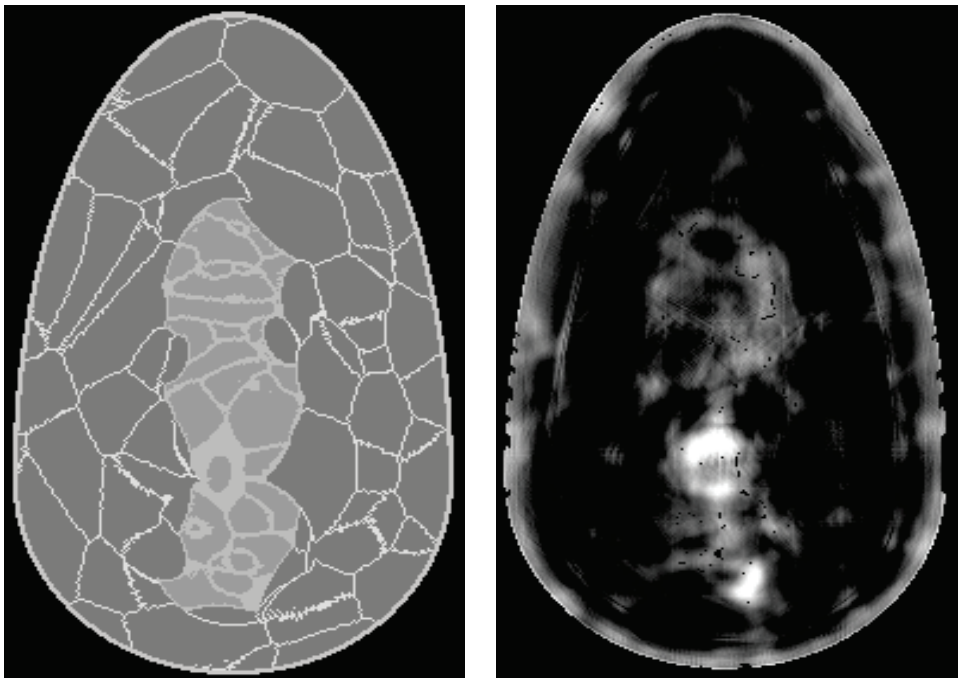
UST Slice



EM-MPM

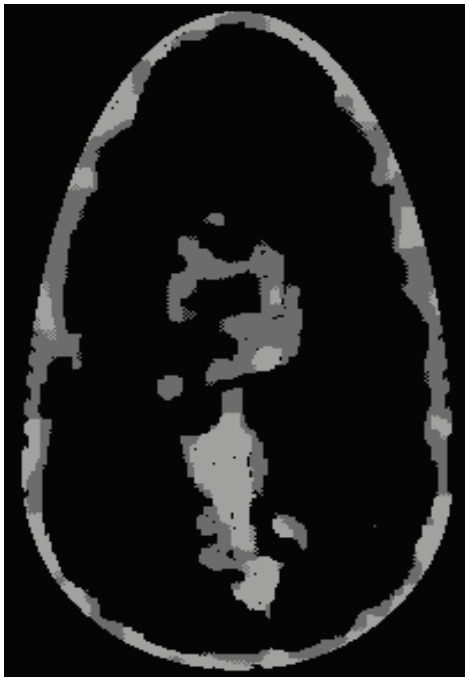
K-means Clustering

Fig. A.9. Comparison between EM-MPM and K-means Clustering on phantom UST, case #13.

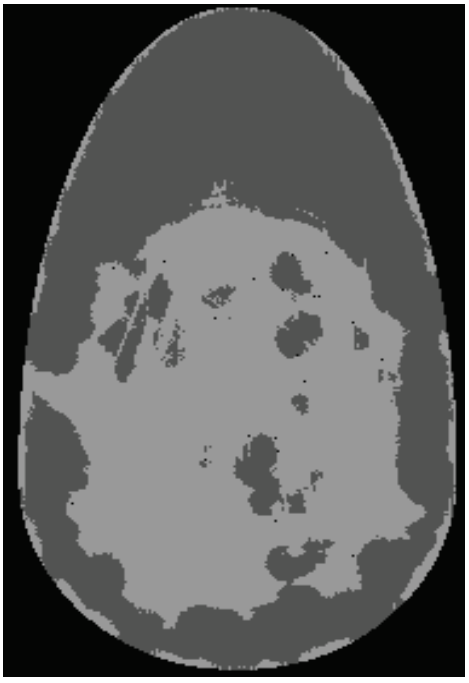


Phantom

UST Slice



EM-MPM



K-means Clustering

Fig. A.10. Comparison between EM-MPM and K-means Clustering on phantom UST, case #14.

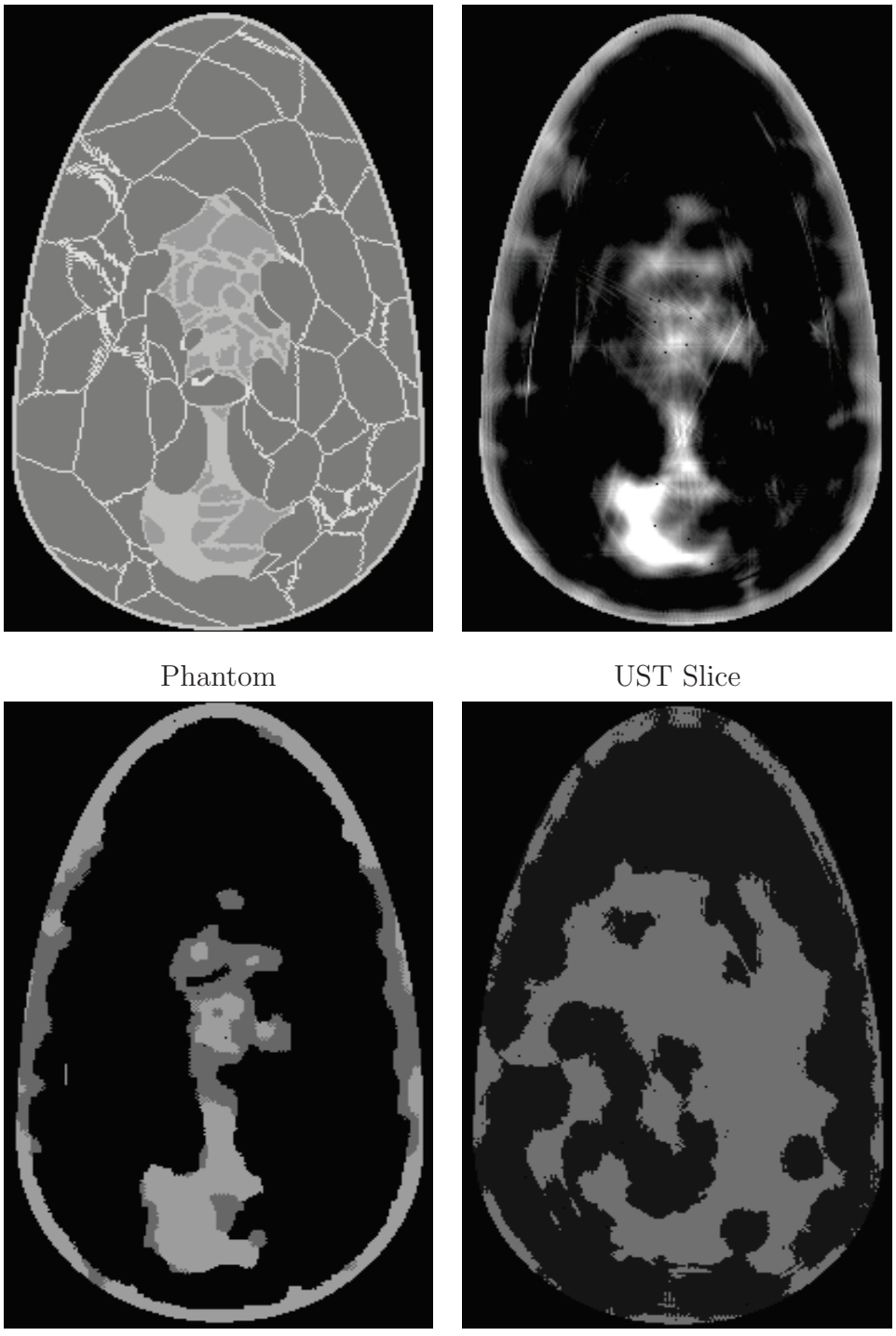
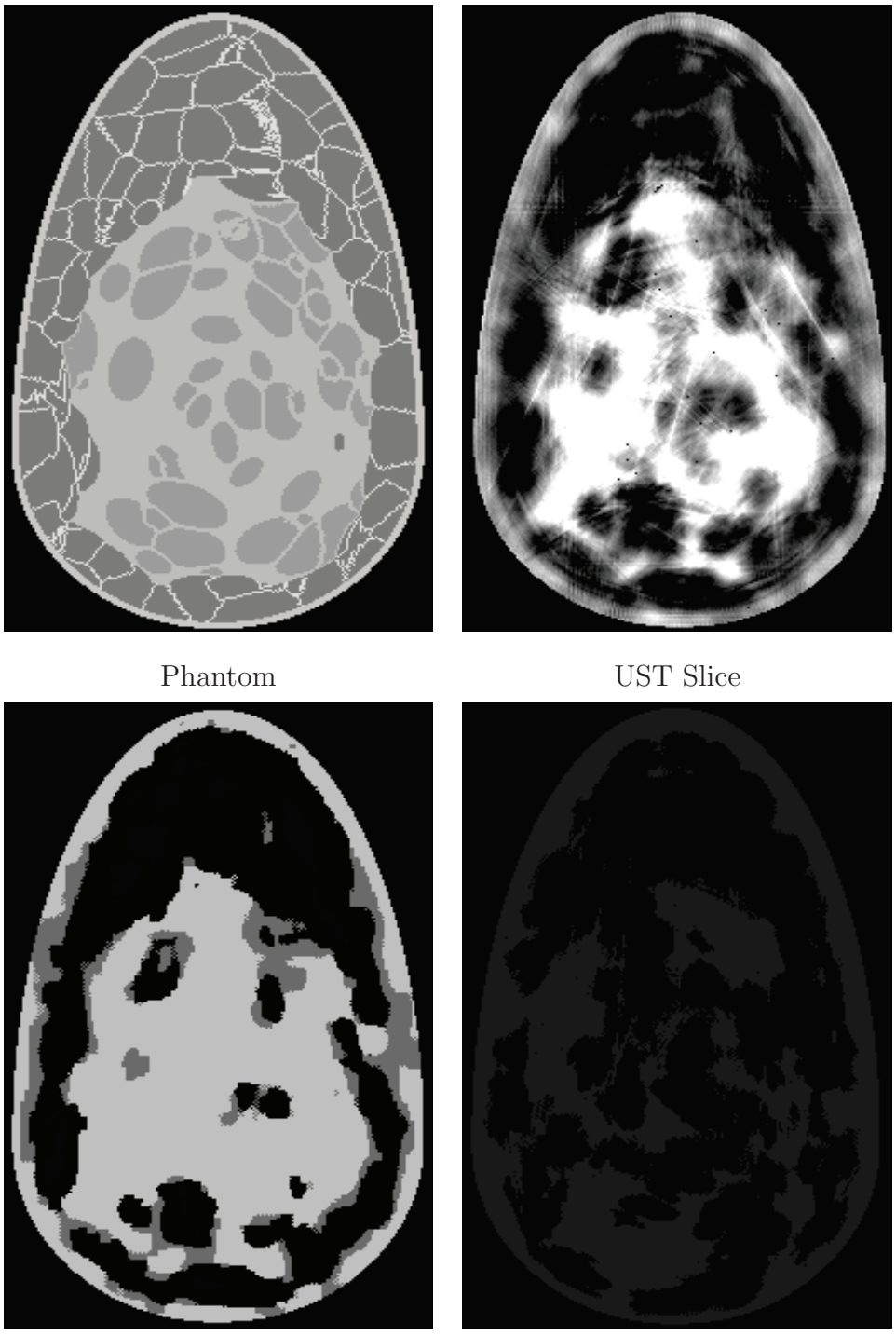


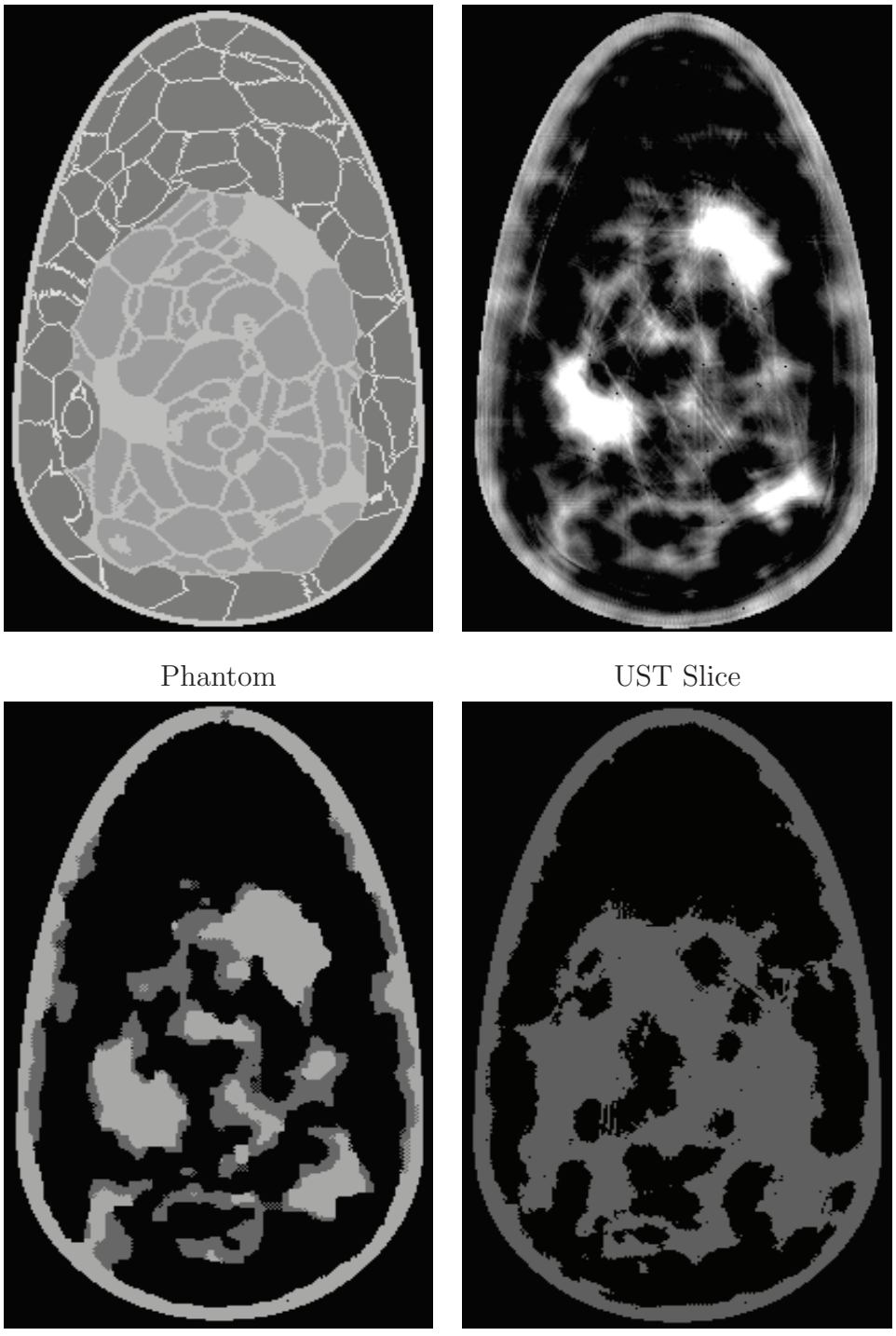
Fig. A.11. Comparison between EM-MPM and K-means Clustering on phantom UST, case #15.



Phantom UST Slice

EM-MPM K-means Clustering

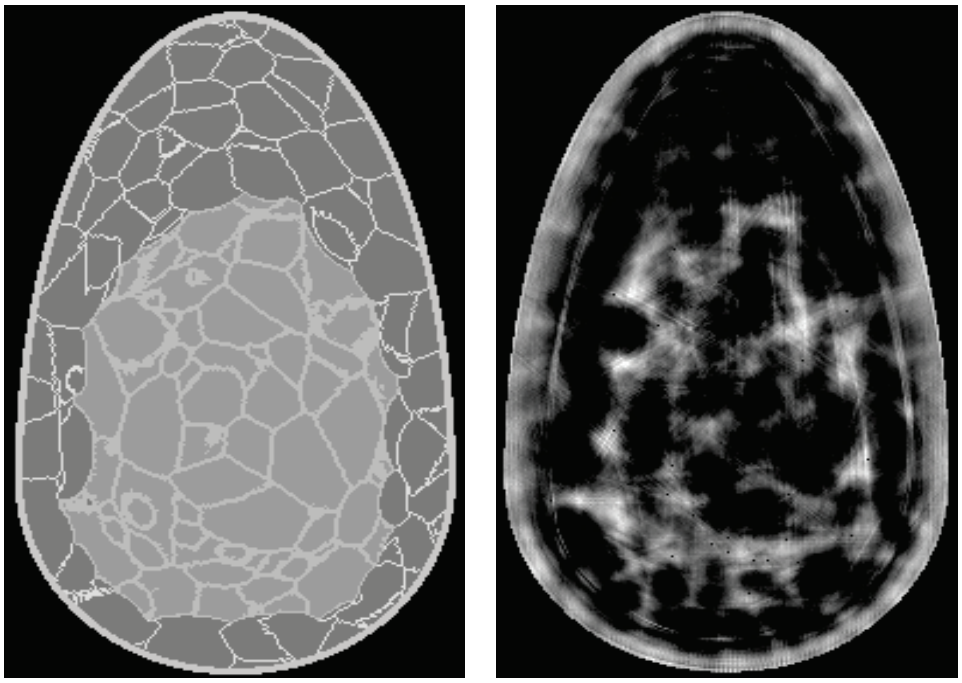
Fig. A.12. Comparison between EM-MPM and K-means Clustering on phantom UST, case #16.



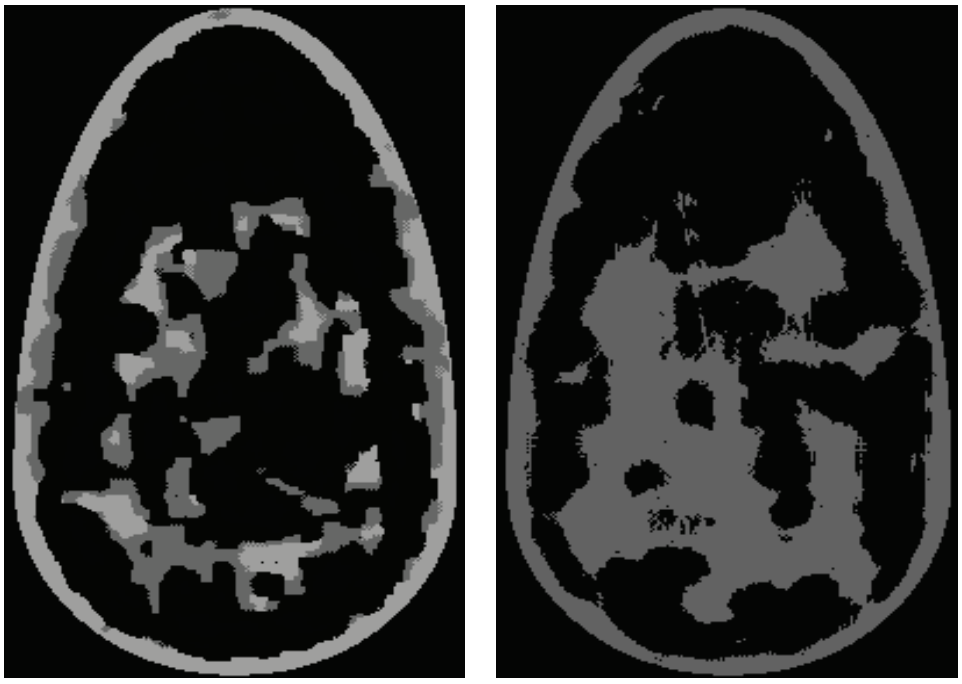
Phantom UST Slice

EM-MPM K-means Clustering

Fig. A.13. Comparison between EM-MPM and K-means Clustering on phantom UST, case #17.



Phantom UST Slice



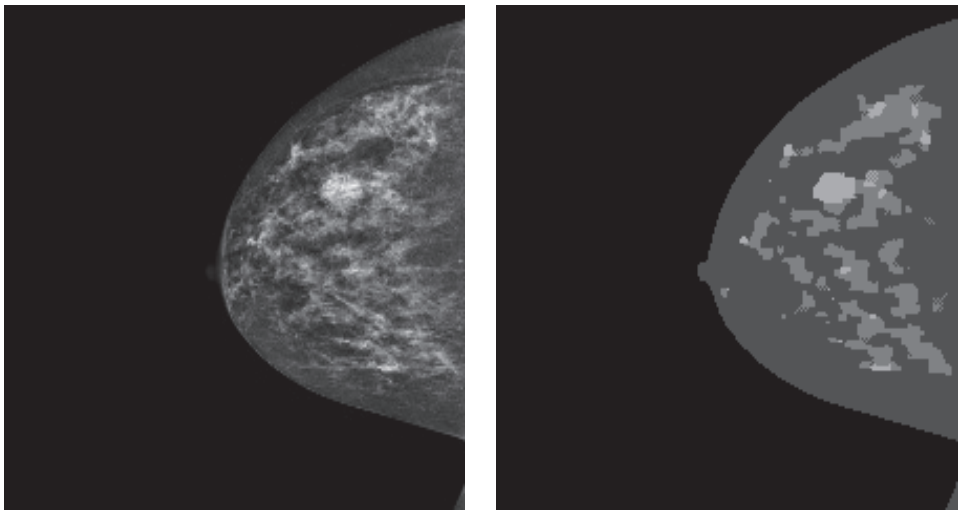
EM-MPM K-means Clustering

Fig. A.14. Comparison between EM-MPM and K-means Clustering on phantom UST, case #18.

B. SEGMENTATION COMPARISON ON CLINICAL UST IMAGES

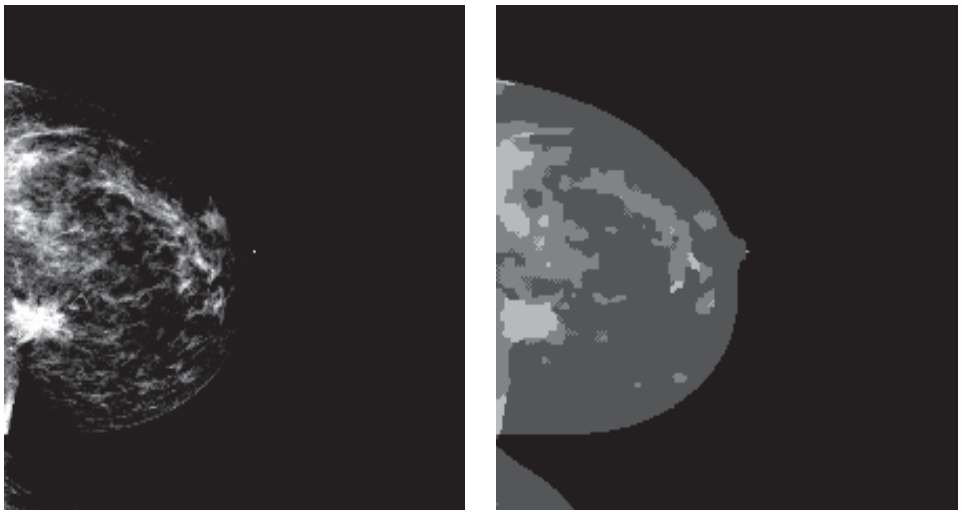
In this section we present the remaining clinical images segmentation comparison. Figure B.1 to Figure B.6 illustrate the 2-D mammogram images segmentation using EM-MPM algorithm. We use four class option for EM-MPM. Within the breast tissue, the brighter gray level areas (compared with the background fat tissue) are regarded as dense tissue, then we compare it with the original mammogram image. It shows clearly that EM-MPM correctly figures out the dense part based on the absolute brightness value, and matches very well with the original mammogram image.

Figure B.7 to Figure B.10 present the remaining clinical 3-D UST cases segmentation comparison between EM-MPM and K-means Clustering. We use four class option for EM-MPM, and two class option is chosen for K-means Clustering in the work of Karmonos Cancer Institute. The top brightness class (or darker class in several cases of K-means Clustering) in the segmentation image is regarded as the dense tissue. Then we compare the segmentation result with the original ground truth clinical images. Again EM-MPM shows better segmentation result than K-means Clustering, with clearer dense tissue segmentation and better match with the original images. It clearly shows that EM-MPM has significant promise in providing accurate volumetric percentage density.



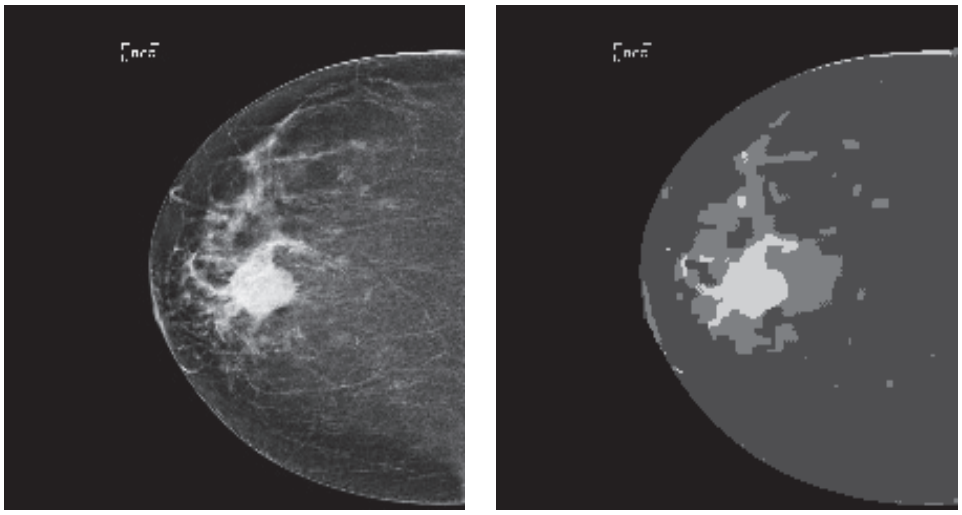
Mammogram EM-MPM Segmentation Result

Fig. B.1. Mammogram images and the EM-MPM segmentation result, case #2



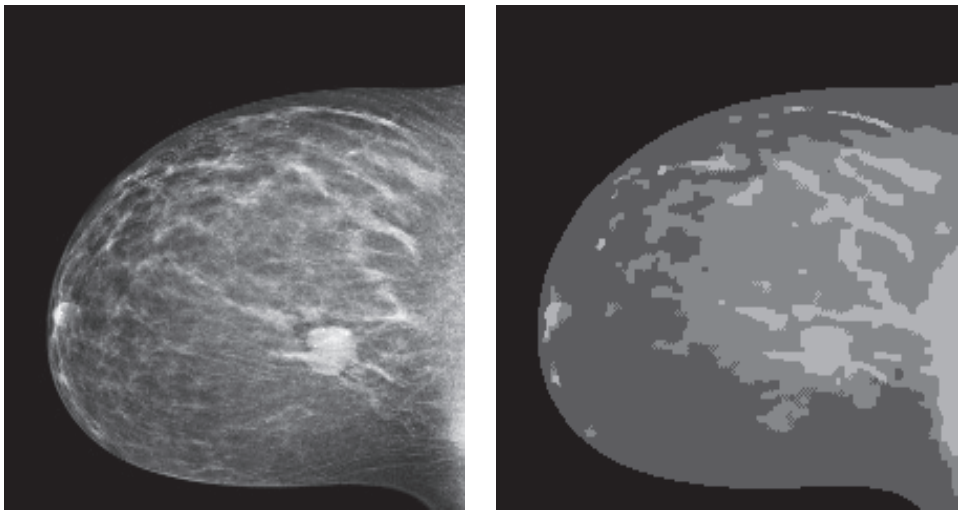
Mammogram EM-MPM Segmentation Result

Fig. B.2. Mammogram images and the EM-MPM segmentation result, case #3



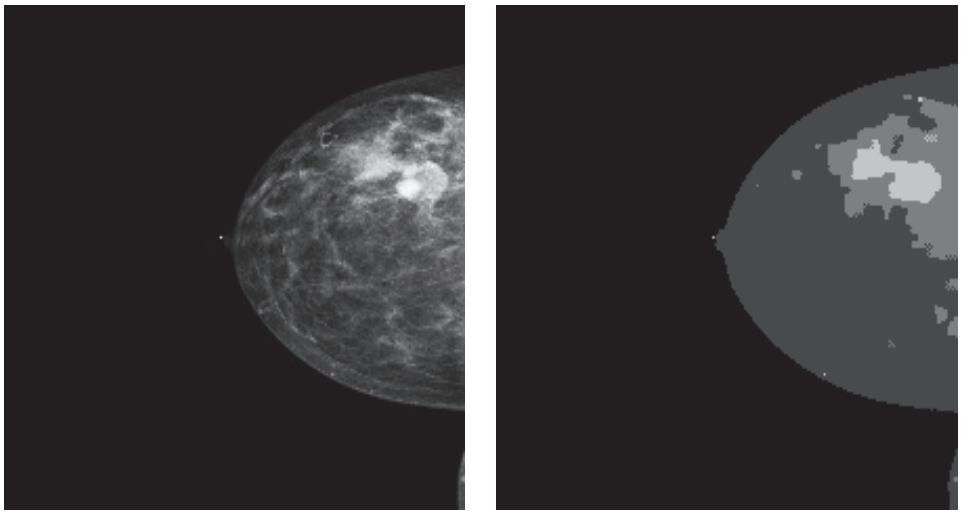
Mammogram EM-MPM Segmentation Result

Fig. B.3. Mammogram images and the EM-MPM segmentation result, case #5



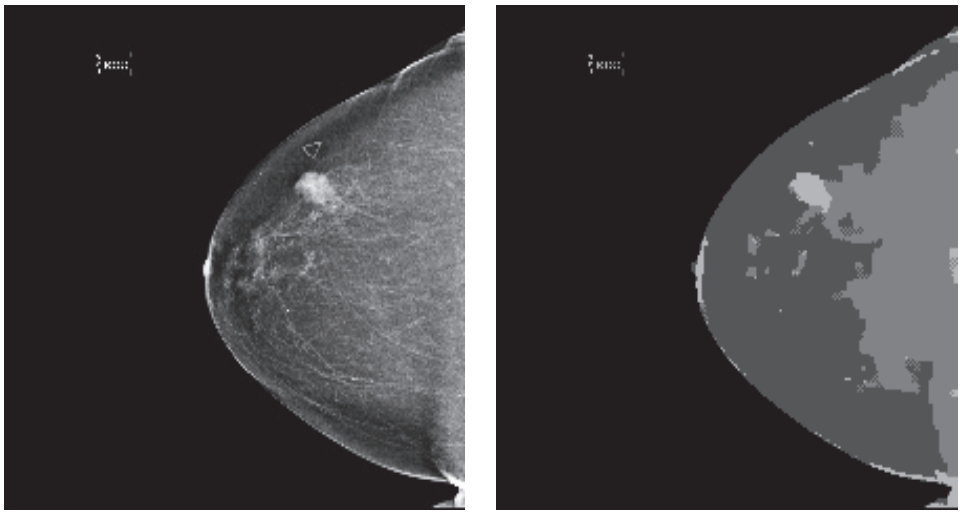
Mammogram EM-MPM Segmentation Result

Fig. B.4. Mammogram images and the EM-MPM segmentation result, case #6



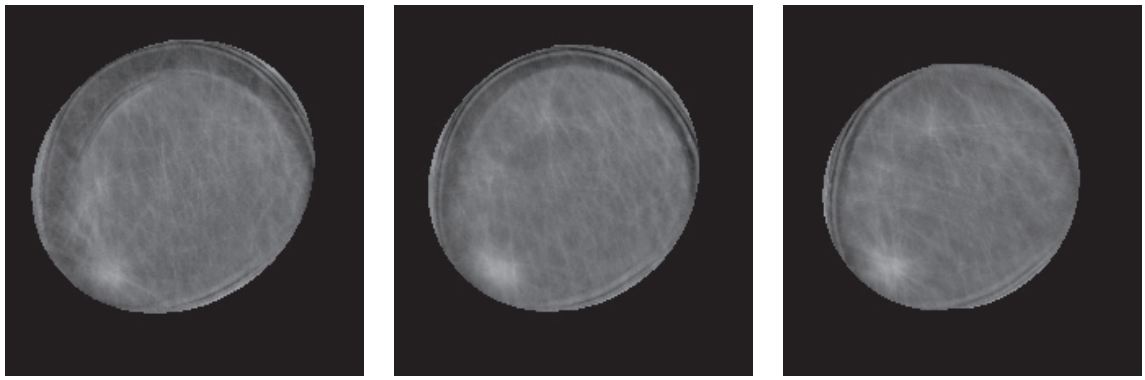
Mammogram EM-MPM Segmentation Result

Fig. B.5. Mammogram images and the EM-MPM segmentation result, case #7



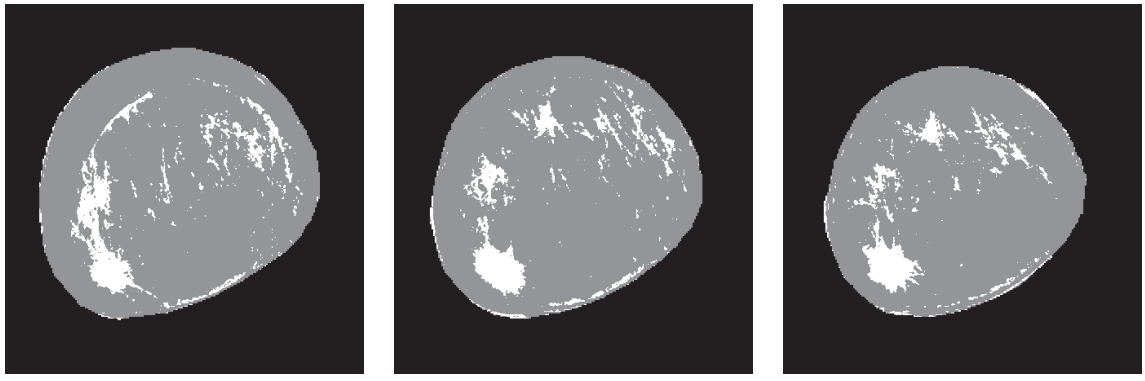
Mammogram EM-MPM Segmentation Result

Fig. B.6. Mammogram images and the EM-MPM segmentation result, case #8



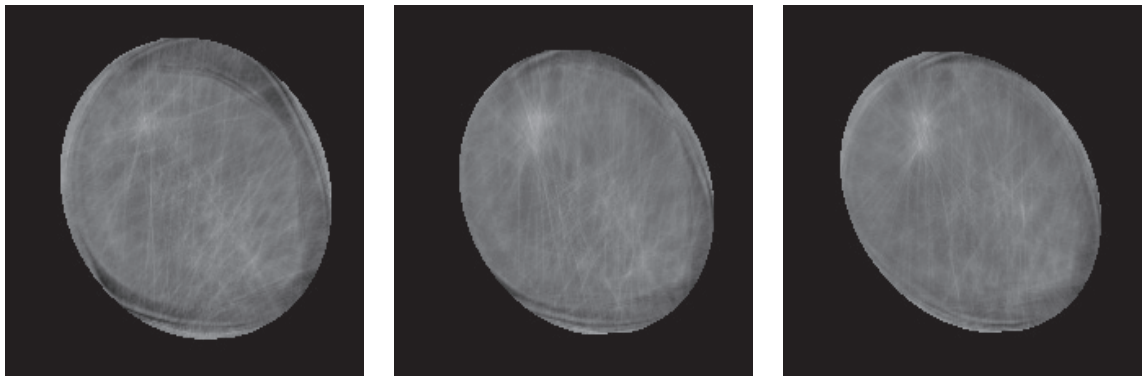
Clinical UST images

EM-MPM Segmentation, corresponding slices (brighter is denser)



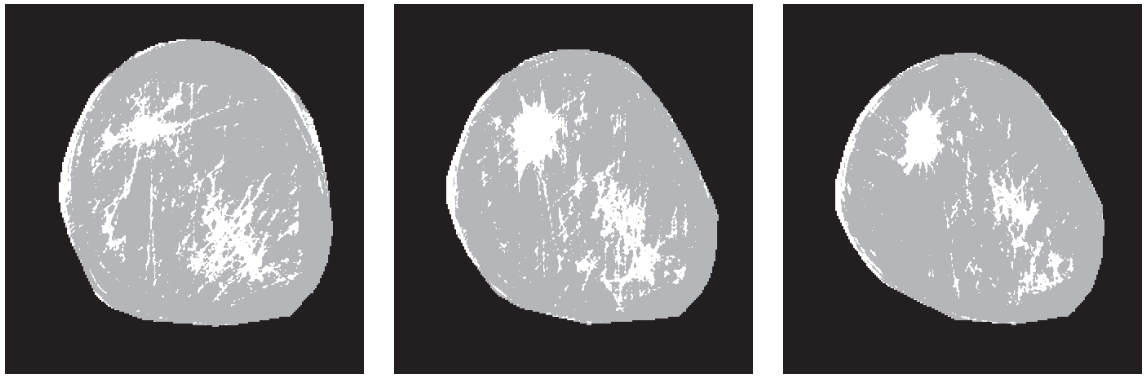
K-means Segmentation, corresponding slices (brighter is denser)

Fig. B.7. Comparison between EM-MPM and K-means Clustering on clinical UST, case #2.



Clinical UST images

EM-MPM Segmentation, corresponding slices (brighter is denser)



K-means Segmentation, corresponding slices (brighter is denser)

Fig. B.8. Comparison between EM-MPM and K-means Clustering on clinical UST, case #3.

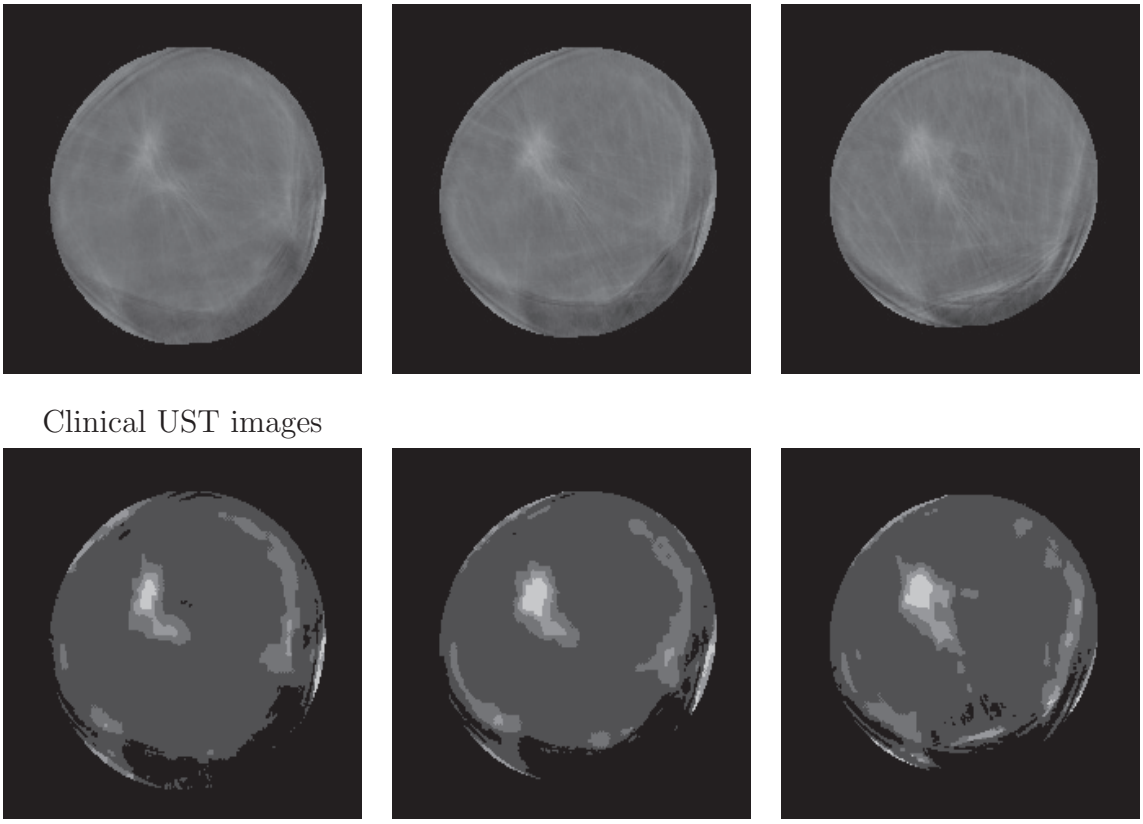


Fig. B.9. Comparison between EM-MPM and K-means Clustering on clinical UST, case #4.

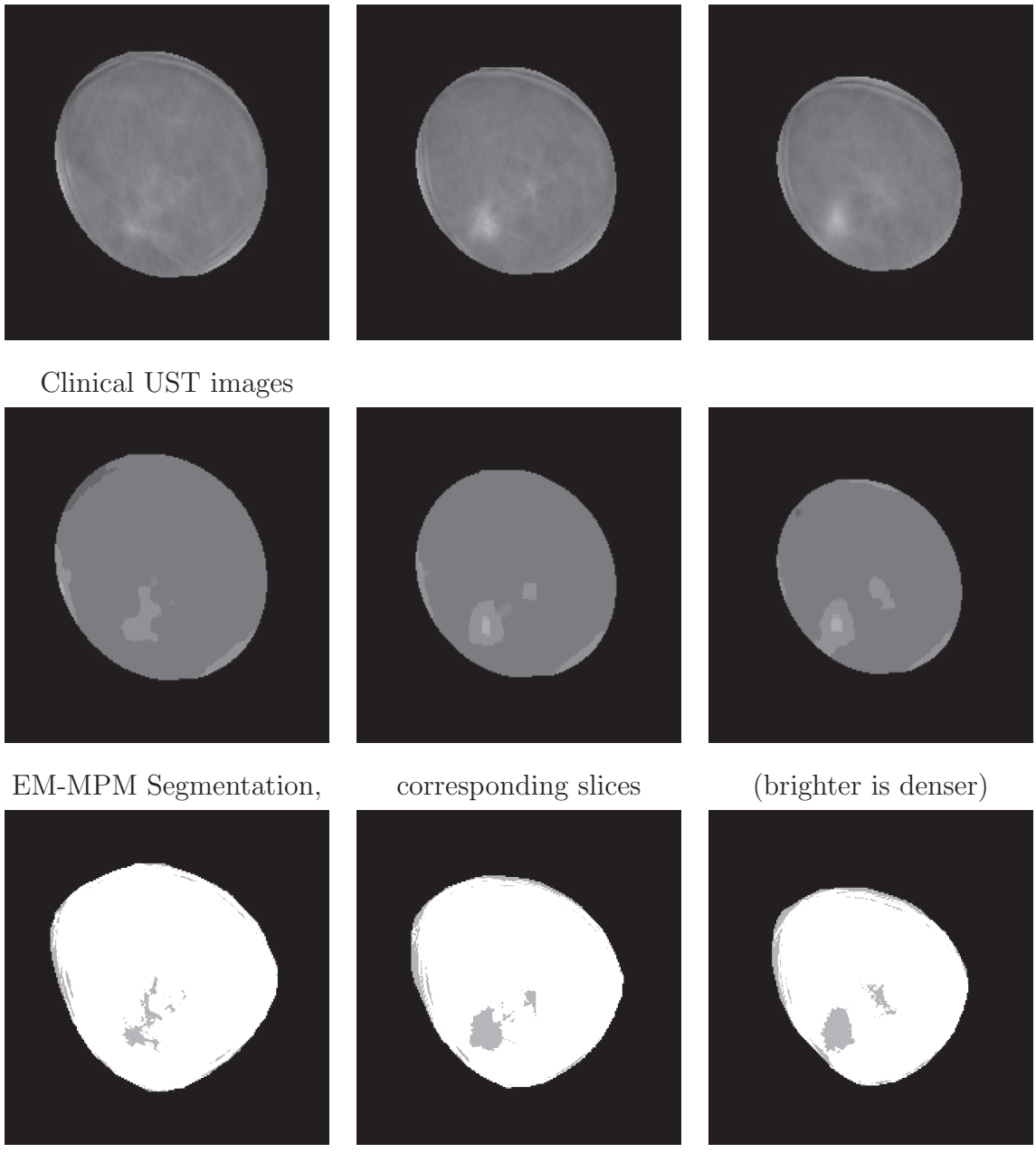


Fig. B.10. Comparison between EM-MPM and K-means Clustering on clinical UST, case #8.