AUTOMATED SEGMENTATION OF BREAST FAT-WATER MR IMAGES USING EMPIRICAL ANALYSIS

José A. Rosado-Toro[1], Tomoe Barr[2], Jean-Philippe Galons[3], Marilyn T. Marron[4], Alison Stopeck[4], Cynthia Thomson[4], María I. Altbach[2,3] and Jeffrey J. Rodríguez[1,2]

[1]Electrical and Computer Engineering, [2]Biomedical Engineering, [3]Medical Imaging and [4]Arizona Cancer Center
University of Arizona, Tucson, AZ, U.S.A.

ABSTRACT

Breast density (BD) has been advocated as a risk factor for the development of breast cancer. BD is typically measured from mammograms. However for longitudinal studies of patients at risk, BD can be better assessed using MRI due to the lack of ionizing radiation and the 3D capabilities of the technique. A fat-water (FW) imaging technique called RAD-GRASE was developed to acquire images of the entire breast in a few minutes and can generate fat-fraction maps, which can be used to assess BD. The time consuming manual segmentation on ~19 slices per exam can be challenging. In this paper, we present a method to automatically segment the breast tissue in FW images and yield FW profiles of the region of interest (ROIs).

Index Terms— automated segmentation, dynamic programming, k-means++, breast MRI, fat-water MRI

1. INTRODUCTION

Breast density (BD) is routinely estimated by comparing the amount of stromal tissue to fatty tissue in mammograms. Radiologists typically visually categorize BD into Breast Imaging Reporting and Data System (BI-RADS) categories to assess breast density [1]. Since BD has been advocated as a risk factor for the development of breast cancer, it is desirable to use it for following the effect of therapeutics and monitoring subjects at risk. However, the radiation exposure in mammography makes the technique impractical for serial studies of BD. Also mammograms only yield 2D information. Magnetic resonance imaging (MRI) is a non-invasive 3D imaging technique that uses non-ionizing radiation. The capability of MRI to yield information of fat and water content makes the technique attractive for BD measurements. A drawback of MRI is that the technique requires long acquisition times. To that end our group has developed a radial gradient and spin-echo technique (RADGRASE) for fast fat-water (FW) imaging. RADGRASE yields fat, water and fat-fraction (FF) maps for the entire breast from data acquired in only 3 minutes [2]. Our results [3] showed that RADGRASE provides a high correlation with mammography results. Manually drawing ROIs (~19 slices per exam) is time consuming and impractical. The purpose of this paper is to design an automatic segmentation algorithm that takes advantage of the various images and maps yielded by the RADGRASE technique. Section two gives a description of the proposed method and section three compares the results of the automatic breast segmentation with manually drawn ROIs.

2. METHODS

Most algorithms use a two-step approach to automatically detect the breast tissue. The first step is to eliminate the background. Any pixel that is not background will form part of the initial segmentation. The second step is to find the boundary between the breast, and the pectoral muscle, which lies at the posterior border of the breast and remove the pectoral portion of the initial segmentation. Unlike previous methods, which use algorithms and parameters that may not be very intuitive, our algorithm incorporates empirical assumptions based on image characteristics. The novelty of this paper is how empirical analysis can be translated into an effective segmentation algorithm.

Another difference between the method presented here and previous work is that the proposed algorithm separates the left and right breasts into two ROIs by eliminating the chest tissue between the breasts. The splitting is needed since the BD analysis considers each breast separately. Previous work does not offer a splitting method; thus, non-intuitive metrics are used to show the performance when compared to a radiologist trace.

2.1.Removing background pixels

The first assumption in our method is that the water image can be divided in three regions: water (high signal intensity), fat (intermediate signal intensity) and background (low signal intensity). These regions are displayed in Fig. 1a. The goal is to find a threshold (based on signal intensity) that eliminates the background while keeping the other two components. Instead of using an iterative threshold method, as proposed by [4], we use the k-means++ algorithm to cluster the image.

K-means++ is a variation of the original k-means algorithm where the centroids are not selected at random, but are initialized to be the farthest possible from each other [6]. This reduces the probability of the k-means optimizing locally rather than globally. Specifically, given an integer k, and a set of pixel intensities I, we wish to compute a set C of k cluster centers (grayscale values) so as to minimize

$$\phi = \sum_{i \in I} \min_{c \epsilon C} |i - c|^2$$

To achieve this, the following steps are taken:

1. Perform the k-means++ initialization to obtain k grayscale values to serve as the initial cluster centers
2. Assign each pixel intensity to the nearest cluster center
3. Recalculate each cluster center to be the mean gray level of the pixels assigned to that cluster
4. Repeat steps 3-4 until the cluster centers no longer change

The k-means++ initialization involves the following steps:

1. Randomly choose a grayscale value as the first cluster center, using a uniform distribution just like in k-means.
2. Assign the probability of choosing the remaining grayscale values according to the distance to the nearest grayscale value that has already been selected as a cluster center.
3. Choose the next cluster based on the above probabilities
4. Repeat steps 2 and 3 until all the centers have been selected

For this application cluster the water image using a k = 3. The cluster with the lowest intensity will contain the background pixels that we wish to eliminate, and the other two labels will contain pixels with mainly fat and water components respectively, Fig. 2 shows an example of the clustering result, where many of the voxels have been correctly labeled as (water, fat, background) = (white, gray, black). Since there is no a priori knowledge of the breast composition, the initial segmentation is defined as any voxel that is not background. An example is shown in Fig. 4b. But clustering alone is not sufficient to outline the breasts because this step joins both breast and pectoral voxels.

2.2.Finding the pectoral boundary

To remove the pectoral region we a second assumption; that the fat shows a well-defined boundary between the breast and pectoral muscle (dark rim in Fig. 3). We exploit this assumption by computing the gradient of the fat image in the vertical direction, forming a gradient image.

This gradient image allows us to outline the pectoral boundary. Previous algorithms try to fit models [7], look for lines that are within a given slope [4] or use a series of thresholds in the Hessian image [5]. Our method is along the lines of [7], but rather than looking for lines at certain angles we proceed to find a contour that minimizes the gradient image. For this task we use dynamic programming.

Dynamic programming is a method for solving complex problems by breaking them down into simpler sub-problems. Avidan & Shamir [8] used dynamic programming in order to estimate seams in an image with the intent of removing the paths of least resistance. They used this approach to reduce the image size. We take this idea and apply it to outlining the pectoral boundary. In dynamic programming the *cumulative minimum energy* M is calculated using

$$\mathrm{M(m,n)} = \mathrm{e(m,n)} + \min\left(\begin{bmatrix} \mathrm{M(m\text{-}1,n\text{-}1)} \\ \mathrm{M(m,n\text{-}1)} \\ \mathrm{M(m,n+1)} \end{bmatrix}\right)$$

where e(m,n) is the gradient image at location m,n.

Instead of starting the contour on the right side of M as in [8], we start the contour the lowest gradient pixel in the chest region. The chest region is defined as the region in the middle of the image (assumption based on the fixed position of the breast receiver coil used in MRI scans) and has a width of approximately 28 pixels (based on anatomical features). Once we have the starting point and the energy image we use dynamic programming to obtain the optimal contour forming the pectoral boundary (see Fig. 4c).

Since we know that the pixels posterior to pectoral boundary are pectoral muscle, we label those pixels as background in the segmentation (see Fig. 4d).

2.3. Removing the chest tissue between the breast

To remove the chest tissue connecting the breasts we use a morphological opening operation. We find the smallest structuring element that can remove the biggest amount of tissue in the estimated chest region (Fig. 4d). The goal is to label as background region the chest tissue (see Fig. 4e).

The result of the opening operation will be two breast ROIs that are not connected. Figure 4 shows the all the steps required to generate the automatic ROIs.

3. RESULTS

Data was acquired on 3 patients using a 1.5T GE Signal NV-CV/i scanner with RADGRASE technique using a phased-array breast RF coil. Four gradient echoes per spin echo were acquired with receiver bandwidth of ±125 kHz, ETL = 12, matrix size = 256×256, TR = 1 s, NEX = 1, FOV = 34 cm and slice thickness= 7 mm. The acquisition of 19 slices takes about 3 minutes.

After segmenting the breast images into two ROIs we used similarity metrics to determine the accuracy of the automatic ROIs compared to the manual ROIs. The following similarity metrics were used:

Dice [9]:

$$\frac{2|X \cap Y|}{|X| + |Y|}$$

Recall:

$$\frac{|X \cap Y|}{|X|}$$

Precision:

$$\frac{|X \cap Y|}{|Y|}$$

For all these metrics, X is an ROI in the manual segmentation (gold standard) and Y is the corresponding ROI in our segmentation. Table I shows the performance analysis using the Dice coefficient, recall and precision. Based on these results, the segmentation accuracy is sufficiently high to encourage further investigation of RADGRASE imaging for use in serial studios of BD for at-risk patients.

Figure 5 shows examples of ROIs generated by the proposed segmentation (green outline) and the manual ROIs (red outline). Note that the automatic method does an excellent job of finding the breast-air and breast-pectoral muscle boundary; in these boundaries the automatic and manual ROIs match well. The main area of difference is on the sides of the breast, as can be seen in the middle and right images in Fig. 5.

Other differences between the manual and automatic ROIs come from the fact that the former was performed using the fat image while our algorithm uses the water image to find the breast-air boundary. The water image is preferred because the fat image does not define well the breast-air boundary in subjects with higher BD (i.e., subjects with a high water component in the images).

4. CONCLUSION

We showed that empirical properties of the breast could be translated to an algorithm for automatic breast image segmentation. This reduces the necessity of using parameters that have no clear correlation with the images.

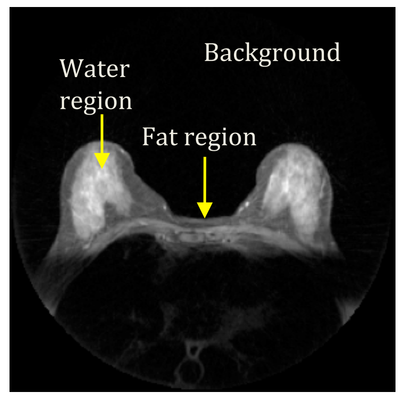
For future work we wish to test this algorithm using more general imaging methods as well as compare our results with other algorithms to assess the performance of our algorithm for other imaging modalities.

Acknowledgements: Work was supported in part by NIH grants HL085385 and CA149417.

5. FIGURES AND TABLES

TABLE I. PERFORMANCE ANALYSIS

Segmentation Metric	Mean	Std. dev.
Dice Coefficient	0.9161	0.0261
Recall	0.9505	0.0283
Precision	0.8852	0.0393



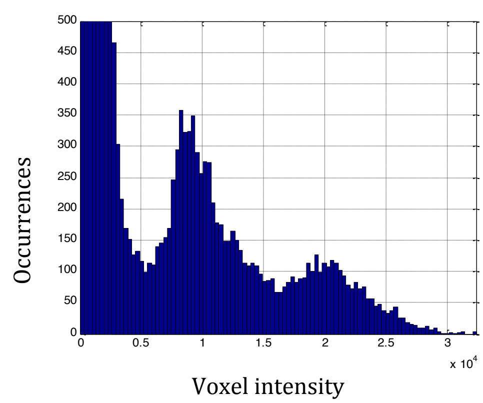


Figure. 1. Water image component with associated histogram (clipped at 500).

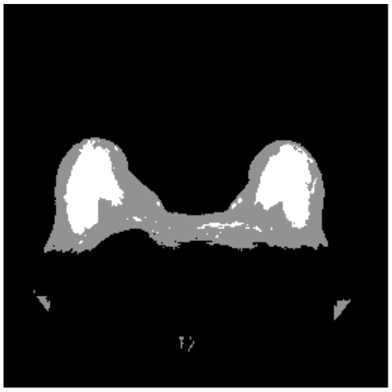


Figure. 2. K-means++ clustering. K = 3

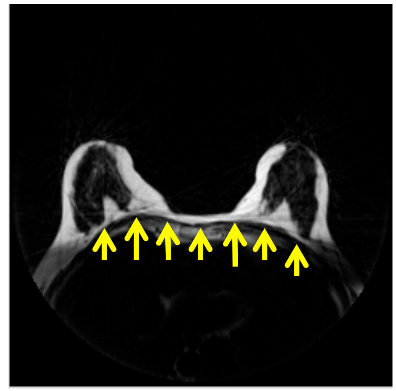


Figure. 3. Fat image showing pectoral boundary; (arrows) clearly delineated by a dark rim

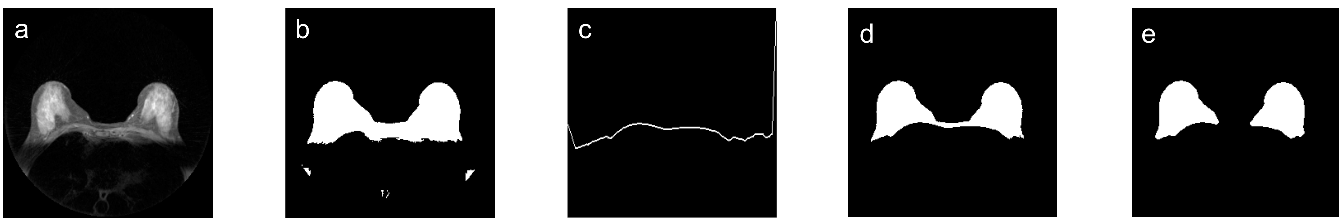


Figure. 4. Result of proposed method: a) water component image, b) initial segmentation, c) pectoral boundary d) refined segmentation e) binary mask after chest removal

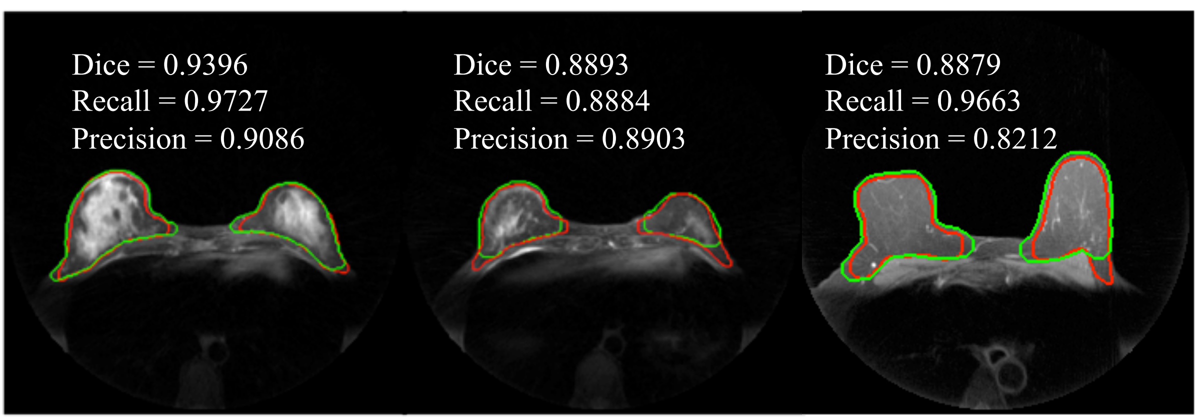


Figure 5. Breast segmentation. Red contour is the manual segmentation; green contour is the proposed method.

6. REFERENCES

[1] E.S. Burnside, E.A. Sickles L.W. Bassett, et al., "The ACR BI-RADS experience: learning from history," *Journal American College of Radiology*, pp. 851-860, Dec. 2009.

[2] Z. Li, C. Graff, A.F. Gmitro, et al., "Rapid water and lipid imaging with T2 mapping using a radial IDEAL-GREASE technique", *Magnetic Resonance in Medicine*, pp. 1415-1424, Jun. 2009.

[3] T.P. Trouard, P. Thompson, C. Huang, et al. " Fat water ratio and diffusion-weighted MRI applied to the measure of breast density as a cancer risk biomarker," *Proceedings of the International Society for Magnetic Resonance in Medicine*, p. 4749, May 2010.

[4] V. Giannini, A. Vignati, L. Morra, et al., "A fully automatic algorithm for segmentation of the breast in DCE-MR images," *Proceeding of Engineering in Medicine and Biology Society*, pp. 3146-3149, Jan. 2010.

[5] L. Wang, B. Platel, T. Ivanovskaya, et al., "Fully automatic breast segmentation in 3D breast MRI," *International Symposium on Biomedical Imaging*, pp.1024-1027, Apr. 2012.

[6] D. Arthur and S. Vassilvitskii, "K-means++: the advantages of careful seeding," *Proceeding Eighteenth Annual Association Computing Machinery Society for Industrial and Applied Mathematics, Symposium on Discrete Algorithms*, pp. 1027-1035, Jan. 2007.

[7] M. Koenig, H. Laue, T. Boehler et al., "Automatic segmentation of relevant structure in DCE MR mammograms," *Proceeding SPIE*, pp. 65141S-1-65141S-6, Mar. 2007.

[8] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing,"*Association for Computer Machinery Transactions on Graphics*, p. 10, Jul. 2007.

[9] L. Dice, "Measures of the amount of ecologic association between species," *Ecology Journal,* pp. 297-302, Jul. 1945.