Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method

Shandong Wu, Susan P. Weinstein, Emily F. Conant, and Despina Kontos^{a)}

Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 17 June 2013; revised 19 September 2013; accepted for publication 15 October 2013; published 13 November 2013)

Purpose: Breast magnetic resonance imaging (MRI) plays an important role in the clinical management of breast cancer. Studies suggest that the relative amount of fibroglandular (i.e., dense) tissue in the breast as quantified in MR images can be predictive of the risk for developing breast cancer, especially for high-risk women. Automated segmentation of the fibroglandular tissue and volumetric density estimation in breast MRI could therefore be useful for breast cancer risk assessment.

Methods: In this work the authors develop and validate a fully automated segmentation algorithm, namely, an atlas-aided fuzzy C-means (FCM-Atlas) method, to estimate the volumetric amount of fibroglandular tissue in breast MRI. The FCM-Atlas is a 2D segmentation method working on a slice-by-slice basis. FCM clustering is first applied to the intensity space of each 2D MR slice to produce an initial voxelwise likelihood map of fibroglandular tissue. Then a prior learned fibroglandular tissue likelihood atlas is incorporated to refine the initial FCM likelihood map to achieve enhanced segmentation, from which the absolute volume of the fibroglandular tissue (IFGTI) and the relative amount (i.e., percentage) of the IFGTI relative to the whole breast volume (FGT%) are computed. The authors' method is evaluated by a representative dataset of 60 3D bilateral breast MRI scans (120 breasts) that span the full breast density range of the American College of Radiology Breast Imaging Reporting and Data System. The automated segmentation is compared to manual segmentation obtained by two experienced breast imaging radiologists. Segmentation performance is assessed by linear regression, Pearson's correlation coefficients, Student's paired *t*-test, and Dice's similarity coefficients (DSC).

Results: The inter-reader correlation is 0.97 for FGT% and 0.95 for IFGTI. When compared to the average of the two readers' manual segmentation, the proposed FCM-Atlas method achieves a correlation of $r = 0.92$ for FGT% and $r = 0.93$ for IFGTI, and the automated segmentation is not statistically significantly different ($p = 0.46$ for FGT% and $p = 0.55$ for IFGTI). The bilateral correlation between left breasts and right breasts for the FGT% is 0.94, 0.92, and 0.95 for reader 1, reader 2, and the FCM-Atlas, respectively; likewise, for the IFGTI, it is 0.92, 0.92, and 0.93, respectively. For the spatial segmentation agreement, the automated algorithm achieves a DSC of 0.69 ± 0.1 when compared to reader 1 and 0.61 ± 0.1 for reader 2, respectively, while the DSC between the two readers' manual segmentation is 0.67 ± 0.15 . Additional robustness analysis shows that the segmentation performance of the authors' method is stable both with respect to selecting different cases and to varying the number of cases needed to construct the prior probability atlas. The authors' results also show that the proposed FCM-Atlas method outperforms the commonly used two-cluster FCM-alone method. The authors' method runs at ~ 5 min for each 3D bilateral MR scan (56 slices) for computing the FGT% and IFGTI, compared to ~ 55 min needed for manual segmentation for the same purpose.

Conclusions: The authors' method achieves robust segmentation and can serve as an efficient tool for processing large clinical datasets for quantifying the fibroglandular tissue content in breast MRI. It holds a great potential to support clinical applications in the future including breast cancer risk assessment. © 2013 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4829496>]

Key words: magnetic resonance imaging (MRI), breast, fibroglandular tissue, segmentation, atlas, fuzzy C-means (FCM)

1. INTRODUCTION

Mammographic percentage density (PD%) has been reported as a known imaging-derived risk factor for breast cancer.^{1–3} The PD% estimated in mammography, which is a 2D projection image of the 3D structures in the breast, suffers from the tissue superimposition problem and may not provide an accurate assessment of the true amount of fibroglandular tissue.⁴

Breast magnetic resonance imaging (MRI) (Refs. 5–7) provides 3D scanning and studies have suggested that the volumetric percentage of fibroglandular tissue (FGT%) computed in breast MRI, similarly in concept to the mammographic PD%, is also related to breast cancer risk, especially for the subpopulation of high-risk women.^{8,9}

In order to quantifying the FGT% in breast MRI, accurate segmentation of both the whole-breast (i.e., the region

circumscribed by the air-breast interface and the chest wall line) and the fibroglandular tissue is needed. Automation of this segmentation task is challenging in several aspects. First, segmentation of the whole-breast as an organ from the remaining parts of the MR image is a difficult preprocessing step because of the low intensity contrast between the breast and the chest wall.^{10,11} Second, as the fibroglandular tissue may be present with varying amounts, shapes, and patterns, it is challenging to model and distinguish the true fibroglandular tissue from other tissues (e.g., fat) by simple geometric/shape descriptors. In addition, the bias field is common in breast MRI where the intensity inhomogeneity may considerably affect the appearance of breast tissue properties,¹² causing additional difficulty in discriminating the fibroglandular tissue from other tissues and artifacts in breast MR images.¹³

To date fibroglandular tissue content has been mostly estimated qualitatively by visual assessment in breast MR.⁹ Several semiautomated segmentation methods have been used in previous studies such as interactive thresholding⁵ and user-assisted fuzzy C-means (FCM) clustering.^{8,14,15} While useful, nonautomated methods are subjective and can introduce inter- and intrareader variability. In addition, they can be time-consuming, therefore being less practical for processing large datasets and clinical translation. Some attempts have been made for automated segmentation of fibroglandular tissue in breast MRI such as the adaptive thresholding,¹⁶ hierarchical support vector machines (SVM),¹⁷ Gaussian distribution curve-fitting algorithm,¹⁸ and multiorgan-atlas methods.¹⁰ While these methods are useful, they are limited. Only ten cases¹⁶ and four cases¹⁷ were tested for validation. Several methods still rely on manual or user-assisted delineation for segmenting the whole-breast,^{16,18} which is an essential step for estimating the total volume of the breast, which is needed (i.e., the denominator) for computing the FGT%. Customized MR protocols (i.e., three-point Dixon images) (Refs. 17 and 18) have also been explored for FGT estimation; yet, these methods may not directly apply to standard clinical breast MR imaging protocols. Finally, the fibroglandular tissue was previously estimated through a multiorgan-atlas method¹⁰ which simultaneously segments the heart, pectoral muscle, lungs, and breasts; however, the performance for fibroglandular tissue estimation by this method is less favorable as compared to other organs as the authors indicated.

In this work, we propose a fully automated segmentation method, namely, the FCM-Atlas, for fibroglandular tissue segmentation and volumetric density estimation in breast MRI. Our method is a 2D segmentation algorithm working on a slice-by-slice basis. First, the whole-breast is segmented by using our previously validated fully automated method.¹⁹ Then, the fibroglandular tissue is segmented based on a voxel-wise likelihood map that is generated by integrating the cluster membership values yielded by FCM and a prior learned fibroglandular tissue likelihood atlas. The concept of incorporating a prior likelihood atlas in FGT segmentation has been shown to be effective in our preliminary evaluations,²⁰ which motivated the approach proposed here, to incorporate the atlas into the FCM clustering algorithm for introducing additional knowledge on the spatial distribution of the fibroglan-

dular tissue to enhance the segmentation. Our method is evaluated by a representative dataset of 60 3D bilateral breast MRI scans (in total 120 breasts, 3360 2D MR slices) (Ref. 21) that span the full American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) breast density ranges, achieving promising segmentation performance.

2. METHODS

2.A. Whole-breast segmentation

The FGT% is the relative amount of the fibroglandular tissue volume contained within the volume of the breast. Hence, the first step of our method is to segment the whole-breast, for which we need to identify the air-breast interface curve and the chest wall line that separates the breast from the chest cavity. We used our previously validated method of an integrated algorithm based on synergistic edge extraction from multiple image inputs to segment the whole-breast.^{11,19} It is within the segmented whole-breast region that the fibroglandular tissue is subsequently estimated. In order to reduce the effect of the bias field, the nonparametric nonuniform intensity normalization (N3) algorithm²² is applied to the segmented breast to alleviate the intensity inhomogeneity. In Secs. 2.B and 2.C we first introduce the process of the atlas construction (Sec. 2.B) and then present the proposed FCM-Atlas algorithm (Sec. 2.C).

2.B. Fibroglandular likelihood atlas construction

In our method the atlas represents a statistical likelihood of being fibroglandular tissue for each voxel on a standard breast template. It is derived by deformable breast image warping and kernel density estimation techniques, based on “ground truth” segmentation of the fibroglandular tissue provided by expert human readers (i.e., radiologists), from a set of training samples of 3D breast MRI scans. The construction of the likelihood atlas is performed first at the intrascan level and then followed by an interscan level processing.²⁰

2.B.1. Intrascan learning

For a given 3D scan of a segmented whole-breast, the first step is to create a case-specific breast template in the form of a closed contour of the breast. To do so, all slices of the given scan are aligned via rigid registration²³ to the central slice. Then, the breast contour for each slice is extracted and interpolated into equidistantly distributed points of a fixed number (200 in our settings, according to previous studies²⁰) [Figs. 1(a) and 1(b)]. The points of breast contours across the different slices are then matched by the nonlinear dynamic time warping (DTW) algorithm²⁴ to establish point-correspondence [Fig. 1(c)], from which a breast contour is produced by averaging the corresponding matched points across the different slices. The averaged contour delineates a 2D (i.e., projected) mean shape of the breast defined as the case-specific 2D breast template for the given scan [Fig. 1(c)]. Next, voxelwise likelihood of fibroglandular tissue

FIG. 1. Intrascan fibroglandular likelihood atlas construction. (a) and (b) Two representative slices of a 3D MRI scan with the breast contours plotted. (c) The DTW-based matching of the breast contours between (b) and the generated case-specific 2D breast template. (d) Based on the built point-correspondence in (c), the breast in (b) is deformed to the breast template. (e) The learned case-specific 2D likelihood atlas, where the colorbar indicates voxelwise likelihood of being fibroglandular tissue within the breast template.

is learned for each point in the case-specific 2D breast template. For each of the slice in the training scans (i.e., where “ground truth” is labeled via manual segmentation), we take its breast contour points as original landmarks and warp the whole slice to align it with the case-specific breast template whose breast contour points are taken as the corresponding target landmarks [Fig. 1(d)]. Here the warping is implemented in a deformable manner by the Delaunay triangulation and cubic interpolation.²⁵ In this way we transfer all slices and the associated segmented fibroglandular tissues into the same coordinate space. Subsequently, we project the segmented fibroglandular tissues for each slice into the case-specific breast template space where a 2D spatial distribution of the accumulated projected points (i.e., the transformed fibroglandular tissue labels) is formed. Let $\{Z_i\}_{i=1}^M$ represent the M accumulated points, where $Z_i = [x_i, y_i]$ encodes the 2D coordinates for point i , we apply the following 2D kernel density estimation²⁶ to derive a continuous distribution function of the projected points,

$$f_H(Z) = \frac{1}{M} \sum_{i=1}^M G_H(Z - Z_i), \quad (1)$$

where Gaussian kernel is selected for G_H and the parameter of bandwidth H is automatically determined according to previously reported methods.²⁶ Function $f_H(Z)$ represents a voxelwise likelihood of being fibroglandular tissue in the breast template, which yields the case-specific fibroglandular likelihood atlas for the given MRI scan [Fig. 1(e)].

2.B.2. Interscan learning

By repeating the above intrascan processing for multiple 3D scans we obtain multiple case-specific likelihood atlases [Figs. 2(a)–2(d)]. At the interscan level, we apply similar processes as described above to generate an overall likelihood atlas over the entire training sample set. First, the contours of multiple case-specific 2D atlases are extracted and matched by the DTW algorithm to build point-correspondences, from which the corresponding matched points are averaged to generate an average breast contour across the training set, which is defined as an overall 2D breast template. Then all the case-specific atlases are deformed by the Delaunay triangulation and cubic interpolation²⁵ to the same standard space circumscribed by the overall breast template; in this standard space, all the case-specific atlases are then averaged to generate an overall fibroglandular likelihood atlas, denoted by A [Fig. 2(e)]. This overall likelihood atlas is then used to aid the fibroglandular tissue segmentation as described in Sec. 2.C.

2.C. Segmentation by FCM-Atlas

The proposed FCM-Atlas method goes through two steps, first FCM clustering, which is followed by the incorporation of the learned fibroglandular likelihood atlas (Sec. 2.B), working on a slice-by-slice basis.

The FCM performs a soft clustering step to assign each voxel a degree of membership to each of a predefined number of clusters.²⁷ For each slice of the segmented whole-breast

FIG. 2. Interscan fibroglandular likelihood atlas construction. (a)–(d) Four selected examples of case-specific atlases derived from four different breast MRI cases that have varying fibroglandular tissue amounts, and (e) an overall atlas (i.e., A) learned over the scans of the full training cases (here, 12 cases). The contours delineate the shape of a breast and the contour in (e) represents a generalized standard breast template. The colorbar indicates voxelwise likelihood of fibroglandular tissue.

of a breast MRI scan we denote the voxel set of the slice by $\{X_i | i \in [1, N]\}$ (N is the number of voxels) and the number of clusters by C . The standard FCM algorithm works toward the minimization of the following objective function F (assuming m is a constant and $m \geq 1$),²⁸

$$F = \sum_i^N \sum_{j=1}^C u_{ij}^m \|X_i - r_j\|^2, \quad (2)$$

where u_{ij} is the degree of membership of voxel X_i belonging to cluster j , and r_j represents the center of cluster j . An iterative optimization of the function F is implemented by fuzzy partitioning through updating the cluster center r_j and the membership u_{ij} by

$$u_{ij} = \frac{1}{\sum_{k=1}^C (\|X_i - r_j\| / \|X_i - r_k\|)^{2/(m-1)}}. \quad (3)$$

The final output of the membership (denoted by u_{ij}^*) represents a voxelwise likelihood map that labels the probability of each voxel being fibroglandular tissue. An initial fibroglandular tissue segmentation may be obtained by directly thresholding the likelihood map u_{ij}^* , as in previous work.²⁹

In our implementation, we use an adaptive method to determine the optimal cluster number C for (2). In general the segmented breast can be divided into two broad intensity-based clusters corresponding to fibroglandular tissue and fat; however, artifacts and residual of the reduced bias field may impact the intensity profile of the breast, particularly for very fatty or very dense breasts. Previous work has indicated that an adaptive multicluster FCM approach can achieve improved performance for dense tissue segmentation in mammograms.³⁰ Therefore, we also follow a multicluster FCM approach, where the optimal cluster number for a given breast MRI scan is adaptively determined. A lookup table is first built based on the training set to associate an initial coarse FGT% estimate with the optimal cluster number for each training sample (in our setup the training sets are the cases used to construct the atlas and they are selected so that they span the four BI-RADS breast density categories). This initial coarse FGT% estimate is estimated using standard two-cluster FCM, and is linked to the optimal cluster number as follows: for a varying range of the cluster number C , the FCM-Atlas generated segmentation is compared to the corresponding expert's manual segmentation and the optimal cluster number is identified as the one that leads to the highest agreement with the manual segmentation. Here we apply a bounding constraint on the cluster number allowing it to range from 2 (as there will always be a fibroglandular tissue and fat cluster) up to 5 (for speed and memory considerations). For a testing breast MRI scan, we first estimate its coarse FGT% based on two-cluster FCM and then search the lookup table (built based on the training set) to pick the optimal cluster number of the entry whose recorded coarse FGT% estimate is closest to the coarse FGT% estimate of the specific testing scan. Once the optimal number of clusters is identified, we then run FCM clustering as proposed by our FCM-Atlas method by assigning the parameter C in (2) with the identified optimal cluster number.

As indicated in previous work,²⁹ the intensity information alone, which is the feature that the FCM works on, is not sufficient to fully exclude those regions/voxels that may have similar intensity but are not necessarily fibroglandular tissue. Hence, following the first step of the adaptive FCM clustering, we proceed to further refine the likelihood map u_{ij}^* by incorporating the *a priori* learned fibroglandular tissue likelihood atlas (i.e., A constructed in Sec. 2.B), which introduces additional prior knowledge about the statistical spatial distribution of the fibroglandular tissue, to update the FCM-generated segmentation. To do so, we compute an updated likelihood map u_{ij}^{*r} as follows:

$$u_{ij}^{*r} = u_{ij}^* \cdot (W \cdot A), \quad (4)$$

where W represents the warping transformation that deforms the standard atlas to the shape of the specific breast being processed. The warping transformation is necessary here since the atlas is learned over a standard breast template [e.g., the breast contours shown in Fig. 2(e)] and it needs to be transformed to match the specific shape of an unseen (i.e., new) test slice in order to multiply with the FCM-generated likelihood map u_{ij}^* . The warping transformation W is implemented by deforming the atlas again using the Delaunay triangulation and cubic interpolation techniques,²⁵ similar to the interscan processing for atlas construction. Finally, the fibroglandular tissue segmentation is achieved by thresholding the refined likelihood map u_{ij}^{*r} , where the threshold is automatically determined by using the Otsu's method³¹ and the voxels that have a greater value than the Otsu threshold are identified as the fibroglandular tissue. The entire workflow of the proposed FCM-Atlas method is demonstrated by the example shown in Fig. 3.

Based on the whole-breast segmentation and the fibroglandular tissue segmentation for each breast MRI scan (i.e., two bilateral breasts), we compute the absolute total volume of the breast, denoted by $|Breast|$, and the total fibroglandular tissue volume ($|FGT|$), from which the FGT% is derived by

$$FGT\% = |FGT|/|Breast| \times 100. \quad (5)$$

2.D. Validation

The automated fibroglandular tissue segmentation is compared with manual segmentation performed by human experts. Two fellowship trained and board certified breast imaging radiologists (S.P.W. and E.F.C., each with more than 10 years of experience) manually segmented the fibroglandular tissue for all cases of the dataset by using an inhouse developed interactive breast MRI segmentation software, the "MR-wizard V. 1.0." MRwizard works similar to the concept of the "Cumulus" software³² which has been used extensively for mammographic density estimation.¹ Briefly, after a 3D MRI scan of the segmented whole-breast is loaded to the MRwizard software, the user can go through every slice of the scan to segment the fibroglandular tissue by first outlining a closed contour of the general region of the fibroglandular tissue, and then applying a local thresholding on the image intensity values (thresholding has been used in previous studies⁵ for

FIG. 3. An example demonstrating the workflow of the proposed FCM-Atlas segmentation method. (a) A representative slice. (b) The adaptive FCM-generated likelihood map (i.e., u_{ij}^*). (c) The initial segmentation by thresholding u_{ij}^* . (d) An overall fibroglandular tissue likelihood atlas (i.e., A). (e) The warped likelihood atlas (i.e., $W \cdot A$) that deforms the standard atlas A to align with the specific slice being processed. (f) The refined likelihood map (i.e., u_{ij}^{*r}). (g) The final fibroglandular tissue segmentation by thresholding the refined likelihood map u_{ij}^{*r} . (h) The segmentation contours superimposed on the initial breast MR slice.

similar purposes) within the outlined region. In our study, the manual segmentations were done independently by the two readers and they are considered here as our gold standard for the validation of our method.

The primary metrics for assessing the segmentation performance includes linear regression, Pearson's correlation coefficients, and Student's two-sided paired t -test, all evaluated on measures by case. Specifically, we show the linear regression and compute the Pearson's correlation coefficients between the FCM-Atlas generated segmentation and the manual segmentation of reader 1, reader 2, and the average of the two readers. The Student's t -test is used to measure the statistical significance of difference in the mean of the different segmentation. Finally, we also use the Dice's similarity coefficients (DSC) (Ref. 33) as a secondary metric to measure the spatial agreement between the automated and manual segmentation.

3. EXPERIMENTS AND RESULTS

3.A. Dataset

Our method is evaluated with a representative set of 60 3D bilateral sagittal breast MRI scans (120 breasts; 56 slices per scan; $56 \times 60 = 3360$ 2D slices in total), randomly selected from cancer-unaffected women in our high-risk screening population.²¹ The age of the 60 women ranges from 26 to 63 years with an average of 44.5 years. Women were imaged prone in either a 1.5T scanner (GE LX echo speed, GE Health, Nutley, NJ, or Siemens Sonata, Siemens Medical Solutions, Malvern, PA) or a 3T scanner (Siemens Trio, Siemens Medical Solutions, Malvern, PA) using a dedicated surface breast coil array and a varying range of clinical imaging parameters;

matrix size: 256×256 (except 192×192 for one case); slice thickness: 2.3–4 mm; slice spacing: 2.3–4 mm; field of view: 18–22 cm; and flip angle: 15° or 20° . The T1-weighted, nonfat-suppressed MR series is used to apply our method because of the high intensity contrast between the fibroglandular tissue and fat. The 60 cases were selected so that they span the full range (i.e., all four categories) of the ACR BI-RADS breast density categories.³⁴ There are 17, 15, 14, and 14 cases for BI-RADS density categories I, II, III, and IV available, respectively. FGT% and lFGTl are computed per each woman (i.e., for both breasts).

Out of the 60 cases, we randomly selected 12 cases, i.e., three for each of the four BI-RADS density categories, to construct the atlas; these 12 cases are also used as training set to learn the optimal parameter values of the cluster number for the multicluster FCM (Sec. 2.C). Both of the two readers' manual segmentations of the 12 cases are used in the atlas construction in order to make the atlas more representative. The rest (i.e., 48 cases) of the dataset (i.e., testing cases) are used to test the segmentation method. As a robustness analysis, we also evaluated the variation of the segmentation performance with respect to using different and a varying number of cases to construct the atlas. Note that in our approach, there is no overlap between the training and testing cases; therefore, any testing scan is considered as a truly previously "unseen" sample.

3.B. Results

The Pearson's correlation coefficient between the two readers is 0.97 for FGT% and 0.95 for lFGTl (Fig. 4).

FIG. 4. Linear regression and correlation between the two readers for the FGT% ($r = 0.97$) (a) and IFGTI ($r = 0.95$) (b).

Student's t -test shows that the manual segmentation of the two readers is statistically significantly different for both FGT% (difference: $4.9 \pm 2.4\%$, $p < 0.01$) and IFGTI (difference: $63.0 \pm 27.7 \text{ cm}^3$, $p < 0.01$). When comparing the automated segmentation of the proposed FCM-Atlas method to the readers' manual segmentation, the correlation coefficient for the FGT% is 0.92 for reader 1, 0.91 for reader 2, and 0.92 for the average of the two readers; likewise, the correlation coefficient for the IFGTI is 0.92 for reader 1, 0.92 for reader 2, and 0.93 for the average of the two readers (Fig. 5). Student's t -test shows that the FCM-Atlas-based segmentation is significantly different for both FGT% and IFGTI compared to the manual segmentation of both reader 1 (FGT% difference: $1.4 \pm 1.4\%$, IFGTI difference: $24.8 \pm 32.7 \text{ cm}^3$, all $p < 0.01$) and reader 2 (FGT% difference: $3.5 \pm 1.0\%$, IFGTI difference: $38.2 \pm 5.0 \text{ cm}^3$, all $p < 0.01$); however, it is not significantly different when compared to the average of the two readers for both FGT% (difference: $1.0 \pm 0.1\%$, $p = 0.46$) and IFGTI (difference: $6.7 \pm 17.2 \text{ cm}^3$, $p = 0.55$). In terms of the spatial agreement of the automated segmentation to the manual segmentation, the averaged DSC (mean $\pm$ std) is 0.69 ± 0.1 for reader 1 and 0.61 ± 0.1 for reader 2, while the averaged DSC between the two readers' manual segmentation is 0.67 ± 0.15 . In addition, as a metric of algorithm consistency, we computed the left vs right breast correlation for each woman for FGT% and IFGTI, due to the natural expectation of a strong bilateral similarity for the FGT segmentation (Fig. 6). The bilateral correlation coefficient for the FGT% is 0.94, 0.92, and 0.95 for reader 1, reader 2, and FCM-Atlas, respectively; likewise, for the IFGTI, it is 0.92, 0.92, and 0.93, respectively.

For the robustness analysis of our atlas-based segmentation method, we repeated the experiments 50 times by each time randomly selecting a different set of 12 cases (i.e., three cases for each BI-RADS density category) to build the atlas and using the rest cases for testing. The averaged correlation coefficient (mean $\pm$ std) between the proposed automated FGT segmentation and manual FGT segmentation of the testing cases for the FGT% is 0.92 ± 0.01 for reader 1, 0.90 ± 0.01 for reader 2, and 0.92 ± 0.01 for the average of the two readers; likewise, the averaged correlation coefficient for the IFGTI is

0.90 ± 0.03 for reader 1, 0.89 ± 0.03 for reader 2, and 0.91 ± 0.03 for the average of the two readers. The averaged DSC is 0.67 ± 0.01 for reader 1 and 0.60 ± 0.01 for reader 2, respectively. We further tested the segmentation performance variation with respect to using a different number of cases to construct the atlas. We varied the total number of cases for constructing the atlas ranging from 4, 8, 12, 16, to 20 (where each setting uses equal number of cases for each BI-RADS density category) and repeated 50 times by randomly selecting different cases each time and then computed the averaged correlation coefficients between the automated FGT segmentation and manual FGT segmentation of the testing cases (Fig. 7). We also computed the coefficient of variation (CV) of the averaged correlations across the range of the case numbers (i.e., 4–20); we found that for all of the six measures shown in the legend of Fig. 7, CV is always less than 0.01, which indicates that our method is stable in segmentation performance with respect to the selected range of the case numbers for constructing the atlas.

Selected segmentation examples are shown in Fig. 8 with the comparison to the corresponding manual segmentations. The average segmentation time of our automated method (including both the whole-breast and FGT segmentation) is ~ 5 min for each 3D scan (56 slices, inplane resolution 256×256) in comparison to ~ 55 min needed for the manual segmentation of the whole-breast¹⁹ and FGT by our study radiologists.

We also compared our method to the commonly used standard FCM method without incorporating the atlas. In previous work the cluster number for standard FCM is either fixed (e.g., 2) (Ref. 29) or interactively determined by users.^{8,14} We implemented the automated two-cluster FCM method;²⁹ the correlation coefficient for the FGT% is 0.92 for reader 1, 0.88 for reader 2, and 0.91 for the average of the two readers, which is slightly lower than the proposed FCM-Atlas method. For the IFGTI, the correlation coefficient for the IFGTI is 0.76 for reader 1, 0.65 for reader 2, and 0.72 for the average of the two readers, which is much lower than FCM-Atlas. Student's t -test shows that the two-cluster standard FCM segmentation is statistically significantly different for both FGT% and

FIG. 5. Linear regression between the proposed FCM-Atlas method and the readers' segmentation for FGT% (left column) and IFGTI (right column). (a) and (d) FCM-Atlas vs reader 1. (b) and (e) FCM-Atlas vs reader 2. (c) and (f) FCM-Atlas vs average of the two readers.

IFGTI when compared to the manual segmentation of both reader 1 (FGT% difference: $17.3 \pm 1.1\%$, IFGTI difference: $236.9 \pm 70.5 \text{ cm}^3$, all $p < 0.001$), and reader 2 (FGT% difference: $22.2 \pm 1.3\%$, IFGTI difference: $299.9 \pm 98.2 \text{ cm}^3$, all $p < 0.001$), as well as the average of the two readers (FGT% difference: $19.7 \pm 0.2\%$, IFGTI difference: $268.4 \pm 86.0 \text{ cm}^3$, all $p < 0.001$). In terms of the spatial agreement, the standard two-cluster FCM-alone method resulted in an averaged DSC of 0.53 ± 0.18 for reader 1 and 0.41 ± 0.18 for reader 2, respectively, which is much lower than our proposed FCM-Atlas method. As can be seen from the linear regression plots shown in Fig. 9, the shift in the vertical direction of the re-

gression lines indicates considerable oversegmentation by the two-cluster FCM-alone method (compared to our FCM-Atlas method in Fig. 5), which is also reflected in the segmentation examples shown in Figs. 3(c) and 8(e).

4. DISCUSSION

In this paper we present a new fully automated fibroglandular tissue segmentation method for breast MRI, which combines a model-based component (e.g., FCM) and a learning-based component (e.g., atlas). While FCM clustering works in the image intensity space, constructing and integrating the

FIG. 6. Linear regression and bilateral correlation on the FGT% (left column) and |FGT| measures (right column) between each woman's left and right breasts for reader 1 (a) and (d), reader 2 (b) and (e), and the proposed FCM-Atlas method (c) and (f).

likelihood atlas introduces additional prior knowledge about the spatial distribution characteristics of the fibroglandular tissue within the breast, which can refine the FCM-generated likelihood map to enable enhanced segmentation.

While the correlation between the manual segmentation of the two readers is high for both the FGT% ($r = 0.97$) and |FGT| ($r = 0.95$), the inter-reader variation of the manual segment is also considerable: the slope of the linear regres-

sion is about 0.76 (Fig. 4), the spatial agreement in terms of DSC is about 0.67, and the two readers' segmentations are statistically significantly different ($p < 0.01$). These results indicate that each reader has a fairly consistent manner in perceiving the amount of the fibroglandular tissue (i.e., high correlation) but their agreement in the spatial locations of the tissue is moderate (i.e., low DSC), which reflects the challenges of the fibroglandular tissue perception and subsequent

FIG. 7. Segmentation performance for our FCM-Atlas method with respect to using a different number of cases to construct the atlas. Note that the horizontal axis is the total case number (i.e., with equal number of cases selected for each BI-RADS density category), and the vertical axis represents the averaged correlation coefficients between the proposed automated FGT segmentation and the manual FGT segmentations (reader 1, reader 2, and their average) of the testing cases.

segmentation, even when performed by experienced breast imaging radiologists. The proposed fully automated FCM-Atlas method was evaluated by a representative and diverse set of breast MRI cases (60 bilateral scans, 120 breasts in total) that span the entire range of the ACR BI-RADS breast

density categories, which reflects clinical reality and varying segmentation difficulty. When compared to the average of the two readers, our method achieves a comparable performance to that of the readers, and more importantly, the automated segmentation is not significantly different to the average of the two readers' manual segmentation. The difference in both FGT% and |FGT| of our automated method to either of the two readers' manual segmentation is also less than how much the two readers are different with each other. Our method can therefore serve as a quantitative reader providing objective and reproducible estimates of the fibroglandular tissue. In addition, the high bilateral correlation provides additional evidence for our method's consistency and validation.

The experimental results also show that our proposed FCM-Atlas method outperforms the commonly used standard two-cluster FCM-alone method. While the FCM-alone method is close in terms of its correlation to the expert readers as the FCM-Atlas for FGT%, its performance is much lower for |FGT| in the corresponding correlation and considerably lower in the DSC measures than our FCM-Atlas. In addition, the difference in both FGT% and |FGT| of the FCM-alone method is considerably larger than that of the proposed FCM-Atlas method, in the comparison to either of the readers or their average. There are two main reasons that our FCM-Atlas method performs better. First, we use an adaptive method to determine the optimal cluster number for the FCM rather than *a priori* fixing the cluster number as 2. The adaptive multicluster method can better deal with the large variation in

FIG. 8. Segmentation examples for our fully automated FCM-Atlas algorithm with comparison to manual segmentation. Each row shows one case example. (a) Segmented breast. (b) Manual segmentation of reader 1. (c) Manual segmentation of reader 2. (d) The proposed FCM-Atlas based segmentation. (e) Two-cluster FCM-alone based segmentation.

FIG. 9. Linear regression and correlation between the automated segmentation of the standard two-cluster FCM-alone method and the readers' manual segmentation for FGT% (left column) and |FGT| (right column). (a) and (d) Two-cluster FCM-alone vs reader 1. (b) and (e) Two-cluster FCM-alone vs reader 2. (c) and (f) Two-cluster FCM-alone vs average of the two readers.

the intensity of the breast MR images, which is in line with the use of multicluster FCM for improving mammographic density segmentation³⁰ as well as for semiautomated fibroglandular tissue segmentation in MR images where the cluster number is interactively optimized by users.^{8,14} The second reason is due to the incorporation of the fibroglandular tissue likelihood atlas. As illustrated in Fig. 8, the atlas can help to considerably reduce false positive segmentation in the FCM-generated results.

One of the main parameters for our method is how many cases are needed to construct the atlas and subsequently how

much the proposed method is robust with respect to using different and a varying number of cases to construct the atlas. In the robustness analysis, we tested the segmentation performance variation in this regard and the experiment results indicate that our method is stable with respect to random permutations of different cases and a varying range (i.e., 4–20) of case number for constructing the atlas. This is an encouraging merit as we demonstrate that our method is not very much sensitive to the selection of the cases for constructing the atlas. Furthermore, as can be seen from Fig. 7, when even as low as four cases are used to construct the atlas, the

segmentation performance remains almost comparable to that of 12 cases, which implies that only a small set of cases may be needed for manual segmentation by experts in order to construct a fibroglandular tissue atlas. This minimum requirement for constructing the atlas could make our method relatively adaptable and able to translate to other studies. As mentioned above, incorporating the atlas is to mainly reduce false positive segmentation in the output of the multicluster FCM, for which the atlas is essentially expected to convey a general outline of the fibroglandular tissue's spatial distribution properties. Since our atlas is constructed by using cases with varying levels of breast density (i.e., including both fatty and dense breasts), we speculate that it may be possible to generate a generic atlas and use it for different breast MRI datasets so that avoiding having to construct the atlas with each new dataset. In fact, the atlas that was built in this work may be directly applicable/transferable to other datasets with similar breast MR imaging protocols, which we will seek to investigate in our future work. It is worth acknowledging that we observed a small decrease in our algorithm performance when as many as 20 cases were used to construct the atlas (Fig. 7), which may indicate overfitting and that the optimal number of cases for the atlas may be less than 20 for our dataset. We plan to further explore these aspects of atlas construction in future larger studies.

Medical image atlases have been used to model the anatomy/structure of various shapes of organs.^{10,35} Our atlas has a similar notion to that of a statistical atlas, but a different implementation technique. In consideration of the fact that the shape, distribution, and amount of the fibroglandular tissue may vary largely among cases, we did not choose to treat the fibroglandular tissue as a well-defined structure in the MR images, as commonly done with other organs.^{10,35} Rather, we investigated the spatial distribution properties of the fibroglandular tissue through constructing a likelihood atlas to aid the subsequent segmentation. In general, our method is advantageous in comparison to previous related work in several aspects: it is fully automated for both the whole-breast and fibroglandular tissue segmentation in the process of computing FGT% and lFGTl, compared to previous work that requires manual segmentation or user assistance for delineating the whole-breast;^{16,18} it is evaluated by a large dataset of 60 bilateral cases with a varying range of MR acquisition parameters, which provides a more extensive validation compared to related previous studies;^{16,17} finally, our method is developed based on standard clinical breast MR imaging protocols so that it may have broader applicability than methods based on custom MRI protocols.^{17,18}

Some limitations of our study must be noted. Currently our segmentation method works in the 2D space. We chose to first implement our approach in 2D, rather than in 3D, due to computational complexity considerations and to demonstrate proof-of-concept that incorporating a prior likelihood atlas can improve the segmentation. In general, a 2D segmentation method is relatively straightforward to implement, more computationally efficient, and closer to the radiologists' practice of reading 2D MR slices in the clinical setting. While current performance is very promising, we also plan to extend

our method to 3D, as part of our future work, by developing the related techniques needed for constructing a 3D fibroglandular tissue atlas. The construction of the atlas and the FCM-Atlas method will also be further tested by independent datasets as well as by using a broader range of breast MR imaging parameters and protocols, field strengths, and MR scanner vendors to evaluate our algorithm's dependence on such parameters and generalizability. In addition, while our method currently runs on sagittal view breast MR images, for images acquired in other views (i.e., axial), we plan to experiment with obtaining the corresponding sagittal view images by reslicing the images in order to run our proposed segmentation method. This will be tested in suitable datasets. Future work will also include assessment of intrareader segmentation variation and additional readers for the inter-reader analysis. Finally, we will also seek to test the clinical utility of our method to automatically quantify the fibroglandular tissue in large clinical breast MRI datasets for breast cancer risk assessment in high-risk women⁹ and for supporting other clinical applications.

5. CONCLUSION

We have developed a fully automated fibroglandular tissue segmentation method for breast MRI that integrates adaptive FCM clustering and a prior learned likelihood atlas. Our method achieves high segmentation performance in a time-efficient manner, while automation enables reproducible segmentation. It can serve as an effective tool for processing large scale clinical datasets to support clinical applications such as breast cancer risk assessment and may be useful to be integrated into the clinical workflow in the future.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIH)/National Cancer Institute (NCI) 1R21CA155906-01A1 grant and the University of Pennsylvania Institute for Translational Medicine and Therapeutics (ITMAT) through a UL1RR024134 grant from the National Center for Research Resources. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources or the National Institutes of Health. The authors would like to thank Ms. Kathleen Thomas, who was the research coordinator of the clinical trial from where the MRI data originated, and Dr. Johnny Kuo for developing and maintaining the image database.

^{a)} Author to whom correspondence should be addressed. Electronic mail: despina.kontos@uphs.upenn.edu; Telephone: 215-615-0827.

¹ N. F. Boyd, H. Guo, L. J. Martin, L. Sun, J. Stone, E. Fishell, R. A. Jong, G. Hislop, A. Chiarelli, S. Minkin, and M. J. Yaffe, "Mammographic density and the risk and detection of breast cancer," *N. Engl. J. Med.* **356**(3), 227–236 (2007).

² J. N. Wolfe, A. F. Saftlas, and M. Salane, "Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: A case-control study," *Am. J. Roentgenol.* **148**(6), 1087–1092 (1987).

- ³N. F. Boyd, J. W. Byng, R. A. Jong, E. K. Fishell, L. E. Little, A. B. Miller, G. A. Lockwood, D. L. Titchler, and M. J. Yaffe, "Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian national breast screening study," *J. Natl. Cancer Inst.* **87**(9), 670–675 (1995).
- ⁴D. B. Kopans, "Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk," *Radiology* **246**(2), 348–353 (2008).
- ⁵M. Khazen, R. M. Warren, C. R. Boggis, E. C. Bryant, S. Reed, I. Warsi, L. J. Pointon, G. E. Kwan-Lim, D. Thompson, R. Eeles, D. Easton, D. G. Evans, and M. O. Leach, "A pilot study of compositional analysis of the breast and estimation of breast mammographic density using three-dimensional T1-weighted magnetic resonance imaging," *Cancer Epidemiol. Biomarkers Prev.* **17**(9), 2268–2274 (2008).
- ⁶D. Saslow, C. Boetes, W. Burke, S. Harms, M. O. Leach, C. D. Lehman, E. A. Morris, E. Pisano, M. Schnall, S. Sener, R. A. Smith, E. Warner, M. Yaffe, K. S. Andrews, and C. A. Russell, "American cancer society guidelines for breast screening with MRI as an adjunct to mammography," *CA-Cancer J. Clin.* **57**(2), 75–89 (2007).
- ⁷E. A. Morris and L. Liberman, *Breast MRI: Diagnosis and Intervention* (Springer-Verlag, New York, 2005).
- ⁸C. Klifa, J. Carballido-Gamio, L. Wilmes, A. Laprie, J. Shepherd, J. Gibbs, B. Fan, S. Noworolski, and N. Hylton, "Magnetic resonance imaging for secondary assessment of breast density in a high-risk cohort," *Magn. Reson. Imaging* **28**(1), 8–15 (2010).
- ⁹V. King, J. D. Brooks, J. L. Bernstein, A. S. Reiner, M. C. Pike, and E. A. Morris, "Background parenchymal enhancement at breast MR imaging and breast cancer risk," *Radiology* **260**(1), 50–60 (2011).
- ¹⁰A. G. Mérida, M. Kallenberg, R. Martí, and N. Karssemeijer, "Fully automatic fibroglandular tissue segmentation in breast MRI: Atlas-based approach," in *Workshop in Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Toronto, Canada, 2011.
- ¹¹S. Wu, S. P. Weinstein, E. F. Conant, A. R. Localio, M. D. Schnall, and D. Kontos, "Fully automated chest wall line segmentation in breast MRI by using context information," *Proc. SPIE* **8315**, 831507 (2012).
- ¹²U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Trans. Med. Imaging* **26**(3), 405–421 (2007).
- ¹³M. Lin, S. Chan, J. H. Chen, D. Chang, K. Nie, S. T. Chen, C. J. Lin, T. C. Shih, O. Nalcioglu, and M. Y. Su, "A new bias field correction method combining N3 and FCM for improved segmentation of breast density on MRI," *Med. Phys.* **38**(1), 5–14 (2011).
- ¹⁴K. Nie, J. H. Chen, S. Chan, M. K. Chau, H. J. Yu, S. Bahri, T. Tseng, O. Nalcioglu, and M. Y. Su, "Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI," *Med. Phys.* **35**(12), 5253–5262 (2008).
- ¹⁵T. V. Clendenen, A. Zeleniuch-Jacquotte, L. Moy, M. C. Pike, H. Rusinek, and S. Kim, "Comparison of 3-point dixon imaging and fuzzy C-means clustering methods for breast density measurement," *J. Magn. Reson. Imaging* **38**, 474–481 (2013).
- ¹⁶C. H. Wei, Y. Li, P. J. Huang, C. Y. Gwo, and S. E. Harms, "Estimation of breast density: An adaptive moment preserving method for segmentation of fibroglandular tissue in breast magnetic resonance images," *Eur. J. Radiol.* **81**(4), e618–e624 (2012).
- ¹⁷Y. Wang, G. Morrell, M. E. Heibrun, A. Payne, and D. L. Parker, "3D multi-parametric breast MRI segmentation using hierarchical support vector machine with coil sensitivity correction," *Acad. Radiol.* **20**(2), 137–47 (2013).
- ¹⁸L. J. Lu, T. K. Nishino, R. F. Johnson, F. Nayeem, D. G. Brunder, H. Ju, M. H. Leonard, J. J. Grady, and T. Khamapirad, "Comparison of breast tissue measurements using magnetic resonance imaging, digital mammography and a mathematical algorithm," *Phys. Med. Biol.* **57**(21), 6903–6927 (2012).
- ¹⁹S. Wu, S. P. Weinstein, E. F. Conant, M. D. Schnall, and D. Kontos, "Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images," *Med. Phys.* **40**(4), 042301 (12pp.) (2013).
- ²⁰S. Wu, S. P. Weinstein, and D. Kontos, "Atlas-based probabilistic fibroglandular tissue segmentation in breast MRI," in *Proceedings of the Medical Image Computing & Computer-Assisted Intervention (MICCAI), Part II*, Lecture Notes in Computer Science (LNCS) Vol. 7511 (Springer-Verlag, Berlin, 2012), pp. 437–445.
- ²¹S. P. Weinstein, A. R. Localio, E. F. Conant, M. Rosen, K. M. Thomas, and M. D. Schnall, "Multimodality screening of high-risk women: A prospective cohort study," *J. Clin. Oncol.* **27**(36), 6124–6128 (2009).
- ²²J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imaging* **17**(1), 87–97 (1998).
- ²³J. R. Bergen, P. Anandan, T. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proceedings of the Second European Conference on Computer Vision*, Lecture Notes in Computer Science Vol. 588, edited by G. Sandini (Springer, Berlin, 1992), pp. 237–252.
- ²⁴S. Wu, and Y. F. Li, "Flexible signature descriptions for adaptive motion trajectory representation, perception and recognition," *Pattern Recognit.* **42**(1), 194–214 (2009).
- ²⁵D. F. Watson, *Contouring: A Guide to the Analysis and Display of Spatial Data* (Pergamon, Oxford, 1994).
- ²⁶Z. I. Botev, J. F. Grotowski, and D. P. Kroese, "Kernel density estimation via diffusion," *Ann. Stat.* **38**(5), 2916–2957 (2010).
- ²⁷J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.* **3**(3), 32–57 (1973).
- ²⁸J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms* (Plenum, New York, 1981).
- ²⁹S. Wu, S. P. Weinstein, B. Keller, E. F. Conant, and D. Kontos, "Fully-automated fibroglandular tissue segmentation in breast MRI," in *Proceedings of the International Workshop on Breast Imaging (IWDM)*, Lecture Notes in Computer Science (LNCS) Vol. 7361 (Springer-Verlag, Berlin, 2012), pp. 244–251.
- ³⁰B. M. Keller, D. L. Nathan, Y. Wang, Y. Zheng, J. C. Gee, E. F. Conant, and D. Kontos, "Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy C-means clustering and support vector machine segmentation," *Med. Phys.* **39**(8), 4903–4917 (2012).
- ³¹N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
- ³²M. J. Yaffe, "Mammographic density, measurement of mammographic density," *Breast Cancer Res.* **10**(3), 209–218 (2008).
- ³³L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).
- ³⁴American College of Radiology, *Breast Imaging Reporting and Data System (BI-RADS)*, www.arc.org: American College of Radiology, 2003.
- ³⁵G. Chintalapani, L. M. Ellingsen, O. Sadowsky, and R. H. Taylor, "Statistical atlases of bone anatomy: Construction, iterative improvement and validation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Part I* (Springer-Verlag, Berlin, 2007), pp. 499–506.