A new bias field correction method combining N3 and FCM for improved segmentation of breast density on MRI[a)]

Muqing Lin
Department of Radiological Sciences and Tu and Yuen Center for Functional Onco-Imaging, University of California, Irvine, California 92697

Siwa Chan
Department of Radiology, Taichung Veterans General Hospital, Taichung 404, Taiwan

Jeon-Hor Chen
Department of Radiological Sciences and Tu and Yuen Center for Functional Onco-Imaging, University of California, Irvine, California 92697 and Department of Radiology, China Medical University Hospital, Taichung 404, Taiwan

Daniel Chang and Ke Nie
Department of Radiological Sciences and Tu and Yuen Center for Functional Onco-Imaging, University of California, Irvine, California 92697

Shih-Ting Chen and Cheng-Ju Lin
Department of Radiological Science, China Medical University, Taichung, 404, Taiwan

Tzu-Ching Shih
Department of Radiology, China Medical University Hospital, Taichung 404, Taiwan and Department of Radiological Science, China Medical University, Taichung 404, Taiwan

Orhan Nalcioglu
Department of Radiological Sciences and Tu and Yuen Center for Functional Onco-Imaging, University of California, Irvine, California 92697 and Department of Cogno-Mechatronics Engineering, Pusan National University, Geumjeong-gu, Busan, 609-735, Korea

Min-Ying Su[b)]
Department of Radiological Sciences and Tu and Yuen Center for Functional Onco-Imaging, University of California, Irvine, California 92697

(Received 4 July 2010; revised 14 September 2010; accepted for publication 27 October 2010; published 14 December 2010)

Purpose: Quantitative breast density is known as a strong risk factor associated with the development of breast cancer. Measurement of breast density based on three-dimensional breast MRI may provide very useful information. One important step for quantitative analysis of breast density on MRI is the correction of field inhomogeneity to allow an accurate segmentation of the fibroglandular tissue (dense tissue). A new bias field correction method by combining the nonparametric nonuniformity normalization (N3) algorithm and fuzzy-C-means (FCM)-based inhomogeneity correction algorithm is developed in this work.

Methods: The analysis is performed on non-fat-sat T1-weighted images acquired using a 1.5 T MRI scanner. A total of 60 breasts from 30 healthy volunteers was analyzed. N3 is known as a robust correction method, but it cannot correct a strong bias field on a large area. FCM-based algorithm can correct the bias field on a large area, but it may change the tissue contrast and affect the segmentation quality. The proposed algorithm applies N3 first, followed by FCM, and then the generated bias field is smoothed using Gaussian kernal and B-spline surface fitting to minimize the problem of mistakenly changed tissue contrast. The segmentation results based on the N3+FCM corrected images were compared to the N3 and FCM alone corrected images and another method, coherent local intensity clustering (CLIC), corrected images. The segmentation quality based on different correction methods were evaluated by a radiologist and ranked.

Results: The authors demonstrated that the iterative N3+FCM correction method brightens the signal intensity of fatty tissues and that separates the histogram peaks between the fibroglandular and fatty tissues to allow an accurate segmentation between them. In the first reading session, the radiologist found (N3+FCM>N3>FCM) ranking in 17 breasts, (N3+FCM>N3=FCM) ranking in 7 breasts, (N3+FCM=N3>FCM) in 32 breasts, (N3+FCM=N3=FCM) in 2 breasts, and (N3>N3+FCM>FCM) in 2 breasts. The results of the second reading session were similar. The performance in each pairwise Wilcoxon signed-rank test is significant, showing N3+FCM superior to both N3 and FCM, and N3 superior to FCM. The performance of the new N3+FCM algorithm was comparable to that of CLIC, showing equivalent quality in 57/60 breasts.

Conclusions: Choosing an appropriate bias field correction method is a very important preprocess-

ing step to allow an accurate segmentation of fibroglandular tissues based on breast MRI for quantitative measurement of breast density. The proposed algorithm combining N3+FCM and CLIC both yield satisfactory results. © *2011 American Association of Physicists in Medicine*. [DOI: 10.1118/1.3519869]

Key words: bias field, field inhomogeneity correction, fibroglandular tissue segmentation

I. INTRODUCTION

Breast density has been proven as an independent risk factor associated with the development of breast cancer.[1–4] With the strong evidence, the Breast Cancer Prevention Collaborative Group (BCPCG) recommended that breast density should be incorporated into the risk prediction model to improve the accuracy of predicting each individual woman's cancer risk.[5] However, all supporting evidence to date was established based on the density analyzed on mammography. The percent density (i.e., percentage of the dense tissue area over the breast area on mammogram) has been shown to be significantly higher in patients who developed breast cancer compared to women who were cancer-free.[1–5] Since the measurement of mammographic density is based on a 2D projection image, it may vary with the different body positions, the level and angle of compressions, and the setting of x-ray source and detector. The variations in the measured mammographic density may lead to different estimates of cancer risk.

Breast MRI acquires three-dimensional (3D) images and provides a strong contrast between fibroglandular tissue (i.e., dense tissue) and fatty tissue for measurement of density. The MRI-based analysis methods have been published by several groups,[6–13] and the results from large series comparing MRI density with mammographic density have started to come out in literature.[14–16] We have published a complete processing method for quantitative measurement of breast density, which includes breast segmentation, intensity inhomogeneity (bias field) correction, and tissue classification using the fuzzy-C-means (FCM) algorithm.[6] Although this method has been successfully applied to analyze the density on non-fat-sat T1-weighted images that were acquired using a 1.5 T MRI scanner with a closed bra-shaped coil,[6–9] the segmentation performance based on images that were acquired using a newer scanner with a flat-bed breast coil is not satisfactory. Due to the strong intensity inhomogeneity near the posterior breast, the fatty tissues near the pectoral muscle have low signal intensities and often are misclassified as dense tissues. That is, the FCM algorithm is not sufficient to perform the bias field correction, and other inhomogeneity correction method is needed.

Intensity inhomogeneity (or bias field herein) often presents as a smooth intensity variation across the image and mainly comes from poor radio frequency (RF) coil uniformity, gradient-driven eddy currents, and patient's anatomy both inside and outside the field of view.[17] In general, inhomogeneity correction can be done prospectively based on adjustments of hardware and acquisition methods, or retrospectively based on postscan image processing. Prospective correction methods include using phantom-based calibration of bias field,[18] multicoil scan,[19] and special imaging sequences.[20] However, prospective methods can only remove the inhomogeneity associated with the scanner, not the scanned subject, and retrospective correction methods are needed to remove the subject dependent inhomogeneity.

Four retrospective correction methods are commonly used.[17] The first method uses filtering, which assumes bias field as a low-frequency artifact;[21,22] but this assumption is only applicable for relatively small structures. The second method is based on surface fitting of intensity[23] or gradient[24] field, and it mainly works for images that have relatively large and distinctive homogeneous areas. The third method is based on histogram, which does not require *priori* knowledge and can be applied to large areas with different anatomical structures.[25,26] The fourth method is based on the combination of bias field estimation and tissue classification as *priori* knowledge of different structures for segmentation (e.g., gray matter and white matter in the brain), so they can simultaneously benefit from each other during the iterative process.[27–30]

In our previous breast density segmentation method based on MRI, the FCM algorithm was used for both homogeneity correction and segmentation.[29] This algorithm enhances the signal intensity of fatty tissues within the low sensitivity region of the coil, but the intensity of some fibroglandular tissues is also further enhanced, and that may change the overall contrast of images leading to wrong segmentation results. The nonparametric nonuniformity normalization (N3) algorithm[25] is a fully automatic histogram-based method and is a popular correction method widely used in literature. The N3 algorithm is able to reduce the bias field while avoiding the problem of generating erroneous contrast. However, the N3 algorithm was originally developed for brain images that have a smaller field of view and generally present much less inhomogeneity than the breast images, and it may not be sufficient for bias field correction of breast images.

The purpose of this work is to develop a correction scheme utilizing the combination of N3 and FCM-based algorithms (denoted as N3+FCM). The field inhomogeneity was first removed by using the N3 algorithm,[25] followed by iterative FCM-based algorithm and B-spline surface fitting until the correction is completed. The corrected images were then segmented to differentiate between fibroglandular and fatty tissues. Furthermore, we also implemented a method recently reported by Li *et al.*[30] based on a coherent local intensity clustering (CLIC) criterion function. This method is based on FCM but ensures the smoothness of the estimated bias field. It has shown promising results in removing strong

intensity inhomogeneity on brain MR images but has not been tested in breast images. The quality of the segmentation results obtained using the proposed N3+FCM correction method and that obtained using the FCM alone and the N3 algorithm alone, as well as CLIC, were compared.

II. MATERIALS AND METHODS

II.A. Subjects

This study recruited 30 healthy subjects (age: 23–61; mean: 35), including 25 premenopausal (mean: 30 yr old) and 5 postmenopausal women (mean: 58 yr old). Since both breasts in each subject were analyzed separately, a total of 60 breasts was studied. All subjects did not have any symptom of breast discomfort or palpable breast mass. None of the subjects received hormonal replacement therapy, took oral contraceptives, or had prior history of breast disease or treatment. This study was approved by the Institutional Review Board and all participants gave written informed consent.

Breast MRI was performed using a 1.5 T scanner (Siemens, Erlangen, Germany). One set of 3D T1-weighted images without fat saturation was acquired using the FLASH sequence. The imaging parameters were TR/TE =11/4.7 ms, flip angle=20°, FOV=35 cm, slice thickness =2 mm, and matrix=256×256. The total imaging time for this sequence is 2.5 min.

II.B. Breast segmentation and inhomogeneity correction

The first analysis step is to segment the breast region from the body. An initial cut was performed by manually drawing a horizontal line through the posterior boundary of the sternum, and all tissues below this line were removed. Then, the procedures described in a previous publication,[6] including using the semiautomatic FCM-based method to extract the breast region and using the B-spline fitting to exclude the chest wall muscle, were applied to segment out the breast. For exclusion of skin and nipples, we have developed an improved method, which will be done after the inhomogeneity correction.

We adopted a widely used concept to model the image with intensity inhomogeneity as a smooth multiplicative field.[17,25] The image formation can be written as

$$V(x) = U(x)b(x) + n(x), \tag{1}$$

where V is the measured intensity at location x, U is the true intensity emitted by the tissue, b is an unknown smoothly varying bias field, and n is the white Gaussian noise assumed to be independent of U. To avoid the difficulty of additive and multiplicative interference, we simplify the image model as a noise-free case, while transforming it into log space to make the bias field additive. Using the notation $\hat{I}(x) = \log[I(x)]$, the image model is rewritten as

$$\hat{V}(x) = \hat{U}(x) + \hat{b}(x). \tag{2}$$

The purpose of the inhomogeneity correction is to estimate the bias field b. The flowchart of the proposed correction method is shown in Fig. 1. The first step is to apply the N3 algorithm to the original image I to correct the inhomogeneity. However, it can be seen that after the N3 correction, the fatty tissues close to chest wall still appear dark. Then, the FCM algorithm is applied to the N3 corrected image N to further reduce the inhomogeneity, denoted as F. Using the simplified model in Eq. (2), the additional bias field B after the FCM correction F compared to after the N3 correction N is calculated by

$$B(x) = \exp(\hat{N}(x) - \lambda\hat{F}(x)), \tag{3}$$

where λ is an adjustable constant factor that is used to increase the contrast of the bias field, and was defined empirically. For the 1.5 T breast MR images analyzed in this study, we found λ=0.8 generated the best result. Initially when λ was not introduced (or, set as one), the correction effect could not be visually appreciated. After this parameter was added, the value from 0.5 to 0.9 was tested in selected cases. Generally, when the intensity inhomogeneity is stronger, a smaller λ may increase the contrast of the estimated bias field. However, it was found that λ=0.8 gave a satisfactory result for most cases analyzed in this study. In future studies,

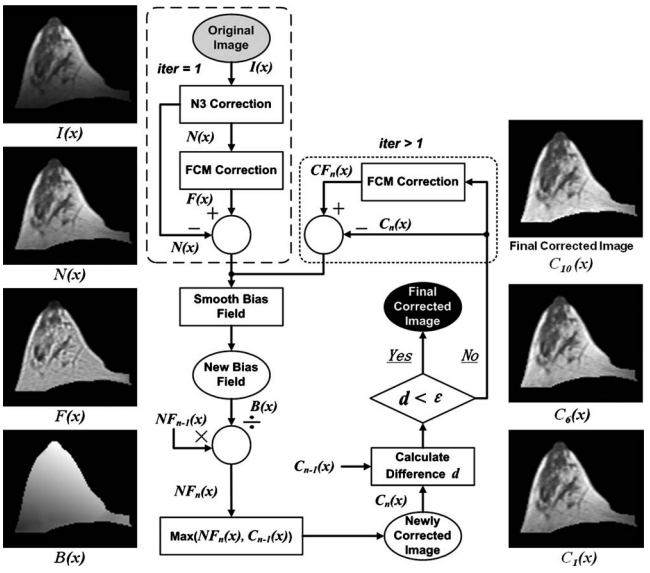

FIG. 1. The iterative process of the N3+FCM bias field correction algorithm. The original image is denoted as I. The N3 corrected image is denoted as N, which still shows inhomogeneous intensity in the posterior breast. Then, the FCM is applied to N to generate the corrected image, denoted as F. Although the bias field shown in N is removed in F, but the intensity of some anterior fibroglandular tissues is brightened too much and appears as the fatty tissue. The bias field B is estimated by calculating the difference between N and F in log space (illustrated in Fig. 2). Then, the Gaussian kernel and the B-spline surface fitting is used to smooth the bias field, so the problem of erroneously changing contrast in the anterior breast can be minimized. By the deconvolution of new smoothed bias field B from N, the bias field corrected image is generated, denoted as NF. In order to keep the dynamic range in the entire intensity spectrum increasing, the intensity of each pixel on the NF is compared to the corrected image in the previous iteration (i.e., the original image I for the first iteration), and the higher intensity is used to form the corrected image C_1 after the first iteration. This process is repeated until the stopping criteria are met. For this example, ten iterations are needed. The corrected image after the first, sixth, and tenth iterations, C_1, C_6, and C_{10} are shown.

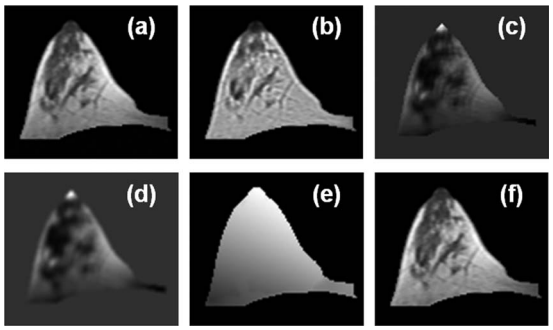

FIG. 2. The calculation of the bias field for the example shown in Fig. 1 during the first iteration. (a) The N3 corrected image N. (b) The FCM corrected image F using the N3 corrected image N as the input. (c) The calculated bias field by taking exponential of $(\hat{N}-\lambda\hat{F})$ using Eq. (3), with $\lambda=0.8$. The dark area in the anterior breast indicates the erroneous change in contrast that makes fibroglandular tissues appear as fatty tissues. (d) Smoothing of (c) using a Gaussian kernel. (e) Smoothing of (d) using B-spline surface fitting to obtain the estimated bias field B. The dark area in the anterior breast shown in (c) and (d) is removed after the B-spline surface fitting, and the homogeneity correction is mainly seen in the posterior breast. (f) The corrected image NF_1 calculated by the deconvolution of B from N. Note that all displayed images are in a relative scale.

this parameter may need to be adjusted based on the quality of the original images.

The intensity ranges of both N and F after the N3 and FCM correction are normalized to the range of the original image I slice by slice for further processing. In order to ensure that the bias field varies smoothly, smoothing is necessary. Figure 2 shows the process of generating the estimated bias field and how this bias field is used to obtain a new corrected image. The obtained bias field B [Fig. 2(c)] from Eq. (3) was first smoothed by a 3×3 Gaussian kernel [Fig. 2(d)] and then further smoothed by B-spline parametric surface fitting [Fig. 2(e)]. The smoothed bias field was then applied back to N to calculate the inhomogeneity corrected image NF using

$$NF_n(x)=\exp(N\hat{F}_{n-1}(x)-B(x)),\quad 1\le n\le \max\_\text{iter}$$

with

$$\widehat{NF_0}(x)=\hat{N}(x). \tag{4}$$

After correcting the bias field using B to obtain the image NF [Fig. 2(f)], an additional step was applied to make sure that the intensity of any pixel within the image field is brighter than their intensity in the previous iteration as

$$C_n(x)=\max(NF_n(x),C_{n-1}(x)),\quad 1\le n\le \max\_\text{iter}$$

with

$$C_0(x)=\max(I(x),N(x)). \tag{5}$$

This step is necessary because some fatty tissue pixels close to the fibroglandular tissue may be smoothed out and has a lowered signal intensity after the bias field correction. For these fatty tissue pixels, the lowered intensity will decrease their contrast from the nearby dense tissue pixels and lead to segmentation error. Therefore, for these pixels, their brighter intensities in the previous iteration will be kept. As will be shown later, one main purpose of the bias field correction is to increase the dynamic range, particularly to brighten the signal of the fatty tissues so they can be separated from the fibroglandular tissues, so this step is needed to ensure that the intensity of fatty tissues will increase after each iteration.

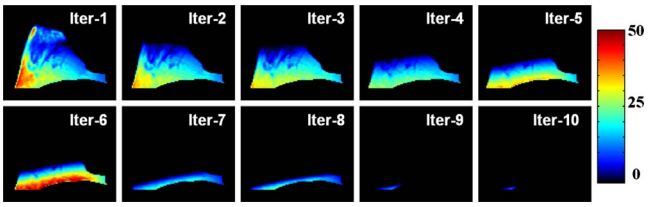

FIG. 3. The subtraction image between the pair of images before and after each iteration. Ten images from iterations 1 to 10 are shown using a normalized scale indicated by the color bar. They demonstrate that the area showing intensity difference between two iterations is shrinking. The lateral posterior breast presents the strongest field inhomogeneity, and the correction effect is clearly seen after each iteration. The iteration will stop when the number of pixels showing changes is smaller than 5% of the total number of pixels in the whole breast.

This process combining N3 and FCM will be repeated until the bias field does no longer change. Since the main purpose of the bias field correction is to brighten the intensity of fatty tissue pixels, the stopping criteria are defined based on the number of pixels that show an intensity change $>i_{thd}$. When less than 5% of the total number of pixels in the whole breast shows changes, the iteration will stop. Define SC_n as the absolute value of the difference between C_n and C_{n-1}, then the number of pixels with intensity change greater than i_{thd} can be counted by the pixels that show $SC_n(x)>i_{thd}$. In this study, the threshold was set as $i_{thd}=\max(SC_0)/256$. Using this threshold, the correction in most cases was completed between 10 and 20 iterations. The iteration process stops when one of the following conditions is met: (1) The number of pixels showing intensity change greater than i_{thd} is less than 5% of the total number of pixels in the whole breast area and (2) the number of iteration exceeds the maximum, which was set as 20 in this work. The example shown in Fig. 1 takes ten iterations to reach the stopping criteria, and the corrected images after the first, sixth, and tenth iterations are shown in Fig. 1. Figure 3 shows the subtraction image SC_n after each iteration. It clearly demonstrates that during the iteration process, the areas showing changes (coded by color) are shrinking from the nipple to the chest wall, indicating that more severe inhomogeneity occurs in the region close to the chest wall, and as such, more iterations are needed to correct pixels in that region.

II.C. Fibroglandular tissue segmentation

The skin and nipple show similar intensities compared to that of the fibroglandular tissue and need to be removed. In our previous study,[6,8] skin was identified using dynamic search based on the gradient of the signal intensities perpendicular to the skin layer. However, nipple often connects to the fibroglandular tissue and cannot be defined based on gradient search, and a fixed layer of 3 pixels (approximately 5

mm) was excluded. Generally, the breast boundary follows a smoothly varying curve and the nipple at a prone position is protruding from the breast curve. The control points of the breast boundary can be set automatically as the boundary pixels of the breast. After setting up the control points, the breast boundary was fit into a smooth curve using Bezier splines[31]

$$\begin{cases} x(u) = \sum_{k=0}^{p-1} x_k BEZ_{k,p-1}(u) \\ y(u) = \sum_{k=0}^{p-1} y_k BEZ_{k,p-1}(u) \\ BEZ_{k,p}(u) = \dfrac{(p-1)!u^k(1-u)^{p-k-1}}{k!(p-k-1)!}, \quad 0 \le u \le 1, \end{cases} \tag{6}$$

where p is the number of control points. The fitted breast boundary removed the major area of the nipple, and then within the remaining breast area, the dynamic search algorithm was applied to find and extract the skin. On the non-fat-sat T1-weighted images, the background signal is dark, the skin signal is intermediate, and the fatty tissue signal is bright, and the skin is defined as the layer in between the two gradients.

In the skin and nipple excluded breast region, the FCM clustering algorithm was applied to segment the fibroglandular tissue. Typically, a total of six FCM clusters was used, brighter three as the fatty tissues and the darker three as the fibroglandular tissue. If the segmented dense tissues using the default setting is consistently overestimated or underestimated, the operator may change the setting of cluster numbers.[6] In this study, fixing the total cluster number as 6, 3 as fibroglandular and 3 as fatty was used for the segmentation in all 60 breasts. Therefore, the process of nipple/skin exclusion and segmentation is fully automatic without any operator interventions.

The source C++ codes of N3 algorithm and B-spline fitting algorithm was taken from this website: *packages.bic.mni.mcgill.ca/tgz*. Since the iterative process of inhomogeneity correction and the segmentation process were written in MATLAB, the C++ codes were modified to be compiled in MATLAB using GNUMEX (http://gnumex.sourceforge.net/) and MINGW (http://www.mingw.org/). The computational time of the entire segmentation process including inhomogeneity correction mainly depends on the iterations of correction. For a desktop computer (Intel Quad 2.8 GHZ, 8 G RAM), the typical computational time including inhomogeneity correction and fibroglandular tissue segmentation ranges from 20 to 90 s. Since the segmentation is fully automatic, the variation in the processing time is mainly dependent on the number of iterations for bias field correction.

II.D. Evaluation of segmentation quality by four methods

For each breast, the fibroglandular tissue segmentation was performed on four sets of images processed using four different inhomogeneity correction methods (FCM-based, N3, CLIC, and the new method using N3+FCM). First, the performance using N3+FCM correction method was compared to that using the N3 alone or the FCM alone method. Since there is no ground truth, the segmentation accuracy cannot be reliably evaluated using a quantitative criterion; therefore, we chose to compare the segmentation quality based on visual inspection of an experienced radiologist. In order to ensure a fair review, the three sets of segmented fibroglandular tissue images were presented blindly in a random order noted by (A, B, and C), and the segmentation accuracy among these three sets was ranked. Using the original nonsegmented images as references, the radiologist evaluated the segmentation quality slice-to-slice. The evaluation criteria include the number of slices in which a noticeable portion of breast tissues were wrongly assigned (fatty tissues as fibroglandular tissues or vice versa), as well as the area of wrongly segmented tissues on each slice. After completing the evaluation of all three sets of images for each breast, the radiologist determined the best, the second best, and the worst (e.g., A>B>C), or equal quality (e.g., A>B=C, A=B>C, or A=B=C). To assess the consistency of radiologist's evaluation, the rating was done twice with 1 month interval in between.

Next, the performance of N3+FCM was compared to that using the CLIC correction method. The segmentation quality was evaluated using the same criteria, and the performance between the N3+FCM and the CLIC methods were compared.

II.E. Statistical analysis

The pairwise Wilcoxon signed-rank test was applied to assess whether or not the superiority exists between the segmentation accuracy based on images generated using different correction methods. The test assigns a sign to the rankings. When comparing a pair between X and Y, a binary outcome categorization was applied: If X was ranked better than Y, then $X=1$ and $Y=0$; if Y was ranked better than X, then $Y=1$ and $X=0$; and if X and Y were ranked equivalent, then $X=1$ and $Y=1$. The performance between each pair of N3, FCM, and N3+FCM was tested using this scoring system and compared. A p-value less than 0.05 was considered significant.

III. RESULTS

III.A. Comparison of inhomogeneity correction

Since the purpose of inhomogeneity correction is to improve the accuracy of fibroglandular tissue segmentation, the comparison is made using the segmentation quality as the evaluation metrics. Figure 4 shows an example of the comparison on one image slice. The original image with heavy bias field, and the corrected images using FCM-based, N3, CLIC, and the proposed N3+FCM are presented in the top row. In the middle row, the gold standard fibroglandular tissue manually delineated by the radiologist, and the corresponding fibroglandular tissue segmentation results based on

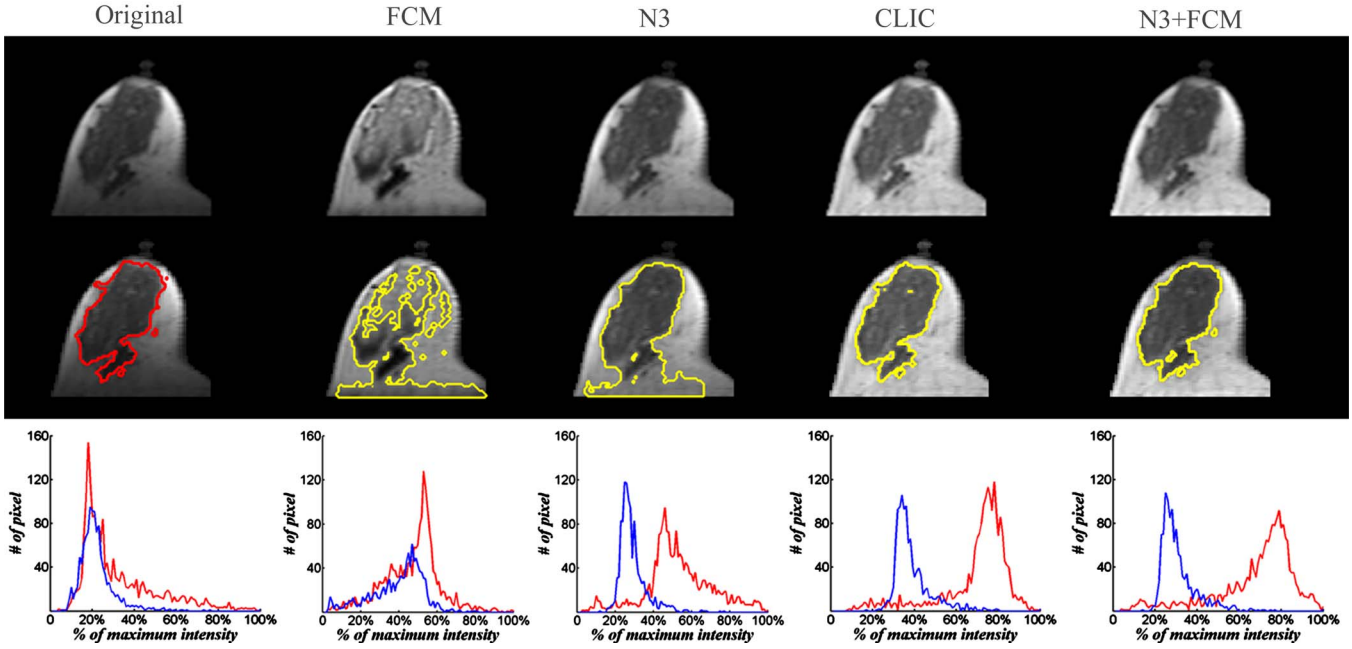

FIG. 4. Comparison of the fibroglandular tissue segmentation quality based on images corrected using four methods. The top row from left to right shows original image, FCM corrected image, N3 corrected image, CLIC corrected image, and N3+FCM corrected image. The middle row from left to right shows the truth fibroglandular tissue delineated by a radiologist and the segmentation results based on FCM, N3, CLIC, and N3+FCM corrected images. It can be seen that both CLIC and N3+FCM produce the most accurate results close to the truth outlined by the radiologist, and FCM has the worst segmentation quality. The bottom row shows the corresponding histograms from pixels in the radiologist outline fibroglandular and fatty tissues on each image. The two curves denote the histograms of fibroglandular tissue and fatty tissue, respectively. It clearly shows that both CLIC and the proposed N3+FCM algorithms increase the dynamic range and widen the separation between the histogram peaks of fatty tissue and fibroglandular tissue.

FCM, N3, CLIC, and the new N3+FCM corrected images are shown. On this particular imaging slice where the fibroglandular tissues are clustered inside surrounded by fatty tissue outside, manual drawing to obtain the ground truth is feasible. The segmentation done based on both CLIC and N3+FCM corrected images is close to the gold standard and would be considered as correct. Both FCM-based and N3 corrected images lead to misclassification of posterior fatty tissues as fibroglandular tissues. For FCM-based correction, some fibroglandular tissues in the anterior region of the breast close to the nipple are brightened so much, resulting in misclassification as fatty tissues. Therefore, the FCM-based correction is the worst among all four methods. The corresponding histograms of the pixels in the truth fibroglandular tissue and fatty tissue regions (based on radiologist's manual segmentation) analyzed from the original image, FCM-based, N3, CLIC, and the N3+FCM correction are shown in the bottom row. For FCM-based correction, the histograms of the fibroglandular tissue and fatty tissue still have a substantial overlap, only slightly better compared to the overlap in the original histograms without correction. The peaks of FCM corrected histograms shift to the right (toward higher intensity range), indicating that both fibroglandular and fatty tissues are brightened. This result clearly shows that the FCM-based correction is not sufficient to allow an accurate segmentation. The N3 algorithm widens the separation between the two histogram peaks. The proposed N3+FCM and CLIC both show a clean separation between these two histogram peaks, thus allow a clean differentiation between fibroglandular tissues and fatty tissues to achieve an accurate segmentation result. For this imaging slice, since the ground truth, algorithm-based segmentation results, and histograms are available, it is possible to define quantitative evaluation metrics (e.g., percentage of accurately segmented pixels or the variation within the fibroglandular and fatty tissue peaks with respect to their separation on the histogram); however, this comparison is only done on one imaging slice, and it is not feasible to obtain ground truth by manually delineating the fibroglandular tissues on each of 20–30 imaging slices contained in one breast.

III.B. Segmentation quality based on N3+FCM compared to using N3 and FCM alone

The visual evaluation results of the segmentation quality made by the radiologist in two separate reading sessions are summarized in Table I. All together, there are a total of 60 breasts. In the first reading session, the (N3+FCM>N3>FCM) ranking is found in 17 breasts, i.e., the performance of N3+FCM is better than using N3 alone, and also the performance using N3 is better than using FCM. The (N3+FCM>N3=FCM) ranking is found in seven breasts, i.e., the performance of N3+FCM is better than using N3 alone, and the performance of N3 and FCM is equivalent. The (N3+FCM=N3>FCM) ranking is found in 32 breasts, i.e., the performance using N3+FCM and using N3 alone is equivalent, and they are better than using the FCM. The (N3+FCM=N3=FCM) ranking is found in two breasts, i.e.,

TABLE I. Radiologist's visual ranking of the fibroglandular tissue segmentation quality based on images corrected by these three methods. ">" means superior quality and "=" means equal quality. The two reading sessions are 1 month apart, performed independently.

	Reading 1	Reading 2
N3+FCM>N3>FCM	17	19
N3+FCM>N3=FCM	7	8
N3+FCM=N3>FCM	32	30
N3+FCM=N3=FCM	2	2
N3>N3+FCM>FCM	2	1

the performance using all three methods is equivalent. The (N3>N3+FCM>FCM) ranking is found in two breasts, i.e., the performance using the N3 alone is better than using N3+FCM. The results of the second reading session are slightly different, but overall it also shows that in most cases the proposed N3+FCM is better or equal to using N3 alone, and that the FCM method is the worst. Figures 5–7 show examples of three different comparison results. For the case shown in Fig. 7, all three correction methods can lead to an accurate segmentation. Compared to the other cases shown in Figs. 5 and 6, the breast shown in Fig. 7 is much smaller (especially in terms of the protruding depth of the breast into the coil), and the signal intensity is homogeneous on the entire image. For this case without a strong bias field, the correction is probably not needed, and all three methods yield satisfactory results.

The significance level between each pair of methods was evaluated using the Wilcoxon signed-rank test based on the assigned outcome categorization score of 1 or 0. In the first reading session, the N3+FCM is better than N3 in 24 cases, with equal performance in 34 cases, and worse than N3 in 2 cases. The Wilcoxon test shows that N3+FCM is significantly better than N3, with $z=4.315$ and $p<0.001$. The N3+FCM is better than FCM in 58 cases and with equal performance in 2 cases with $z=7.616$ and $p<0.001$. The N3 is better than FCM in 51 cases and with equal performance in 9 cases with $z=7.141$ and $p<0.001$. The results of the second reading session were more in favor of N3+FCM. The N3+FCM is better than N3 in 27 cases, with equal performance in 32 cases, and worse than N3 in only 1 case. Therefore, overall, the proposed N3+FCM method is superior compared to the FCM or the N3 method, and that the N3 is significantly better than the FCM method.

III.C. Comparison of segmentation quality based on N3+FCM and CLIC

The radiologist compared the segmentation quality made based on N3+FCM and CLIC corrected images and found that N3+FCM and CLIC have comparable performance in the majority, 57 out of 60, cases. In the remaining three cases, N3+FCM is better than CLIC in two cases and worse than CLIC in one case, but the differences were very subtle. Overall, N3+FCM and CLIC did not show a significant difference, and they could allow a satisfactory segmentation quality for all cases. The segmentation performance based on the CLIC corrected images is also demonstrated in the three case examples shown in Figs. 5–7. Although the segmented areas based on N3+FCM and CLIC show slightly different results, they were considered satisfactory without apparent errors.

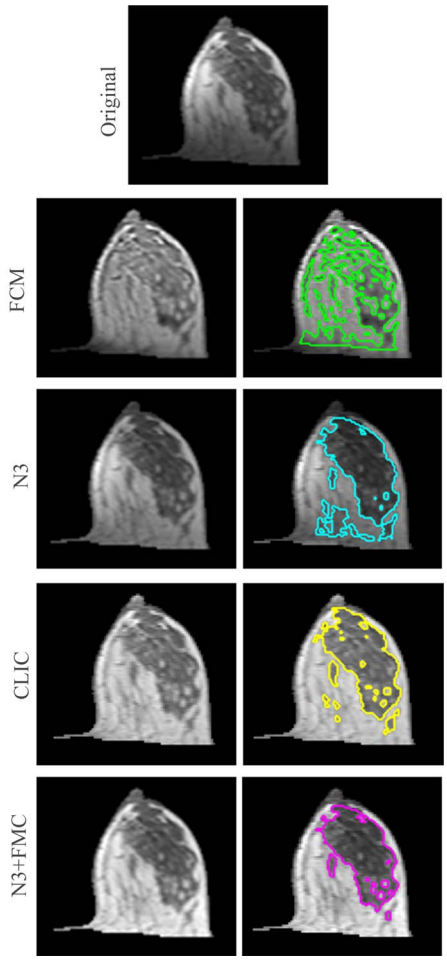

FIG. 5. A case example of "N3+FCM=CLIC>N3>FCM." It can be seen that both the CLIC and N3+FCM correction brighten the signal of fatty tissues in the medial posterior breast and allow the correct classification of pixels in that area as fatty tissues. N3 did not completely correct the bias field, and some tissues in that area are misclassified as dense tissues. The FCM gives the worst performance. Not only that some fatty tissues in the medial posterior breast are misclassified as dense tissues, but also some dense tissues in the anterior breast close to the nipple are misclassified as fatty tissues.

IV. DISCUSSION

We propose a new bias field correction method combining N3 and FCM-based algorithm and have demonstrated that this new correction method can be used to improve the accuracy in segmentation of the fibroglandular tissue. The bias field correction and segmentation are fully automatic and require no manual operations. The most noticeable benefit of the N3+FCM correction method is in its ability to correct the strong bias field near the posterior breast. Tissues in this area fall in the low sensitivity region of the coil, which makes the fatty tissues appear dark and misclassified as the fibroglandular tissue. The proposed N3+FCM algorithm can

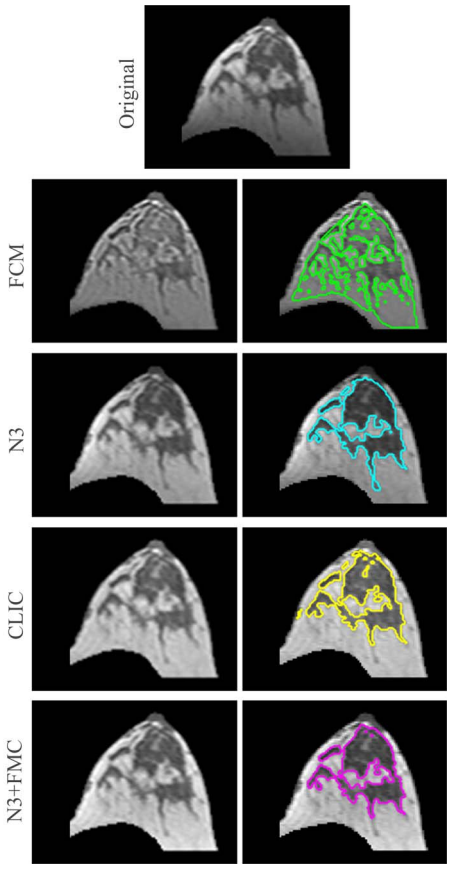

FIG. 6. A case example of "N3+FCM=CLIC=N3>FCM." Although the N3, CLIC, and N3+FCM show slightly different results, all are acceptable, and their performances are rated equal. The same problems indicated in Fig. 5 for FCM corrected images (fatty tissues misclassified as dense tissues and dense tissues misclassified as fatty tissues) are seen, and that makes the performance of FCM inferior to the other three methods.

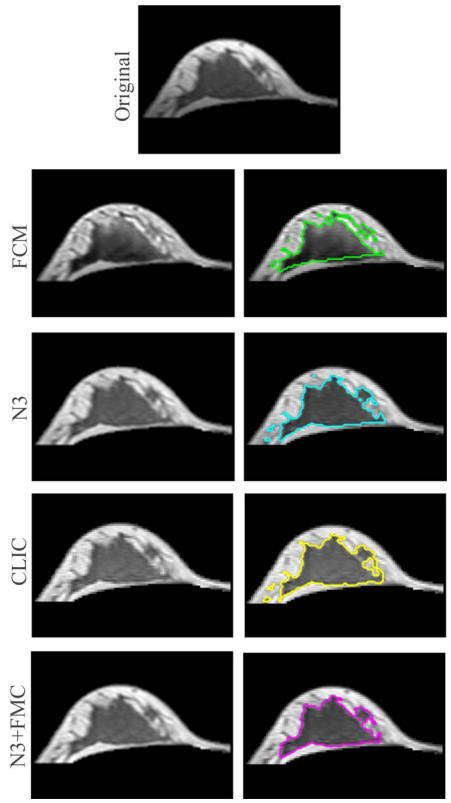

FIG. 7. A case example of "N3+FCM=CLIC=N3=FCM." The segmentation based on all four methods yields similar results, and their performances are rated equal. Note that this breast is fatter (with a smaller protruding depth into the coil), and there is no visually discernable strong bias field. For this case, the correction is probably not needed, and all four methods perform equally well.

be used to brighten the fatty tissue in this area without affecting the intensity of the fibroglandular tissue elsewhere.

The motivation of this work came from the poor performance of the FCM-based correction method in the presence of a strong bias field. Although the FCM method worked well in our previous datasets acquired using a closed-form bra-shaped breast coil,[6–9] this method could not correct the strong bias field on images acquired using the flat-bed breast coil and often led to wrong segmentation results. Because the optimization function of FCM is designed to detect local valleys instead of global minimum, the FCM-based correction is very sensitive to noise. Therefore, in our previous approach, an iterative low-pass filter was added to smooth the neighborhood in the standard objective function of the FCM algorithm.[6] However, this smoothing filter may cause problem in a large breast, where fatty tissues along the boundary of the chest wall may be smoothed out to be close to the outside background. In addition, this correction method assumes that the bias field is of low spatial frequency and other components in the residual image have higher frequencies, which is usually not true for cases with dense breasts, and the correction would lead to erroneous contrast between fibroglandular tissue and fat, as shown in Fig. 4.

Since the N3 algorithm was the most well-established inhomogeneity correction method, it was implemented to test its correction effect on breast images. N3 can detect the smooth, slowly varying, multiplicative field that maximizes the high frequency content of the distribution of the tissue intensities. Although it is originally designed for brain MRI, since the *priori* knowledge regarding segmentation is not required, this algorithm is applicable to breast MRI as well. In a recent review paper about inhomogeneity correction, N3 is still considered as an optimal and widely applicable method.[17] However, the test results using the N3 algorithm for inhomogeneity correction still showed problems (shown in Fig. 4) and could not allow an accurate segmentation. This was probably due to the much larger size of the breast compared to the brain, as well as the design of the breast coil as a surface coil not a volume coil.

Inspired by the advantages and drawbacks of the two correction methods, we proposed to combine them. How these two methods are combined to improve the correction is illustrated in Fig. 2. Unlike the FCM-based correction, the optimization of the N3 algorithm does not depend on the local minima,[25] so it will not change the overall contrast between fibroglandular tissue and fatty tissue on the image. Therefore, the N3 algorithm was applied first to make the initial correction [Fig. 2(a)], and then the FCM-based correction is ap-

plied to further correct the inhomogeneity [Fig. 2(b)]. However, as shown in Fig. 2(b), although the fatty tissues in the posterior breast are brightened, some fibroglandular tissues in the anterior breast are also brightened. In order to suppress this erroneous change in contrast due to the local minima, B-spline surface fitting is applied to smooth the estimated bias field. This is an important process to fully utilize the advantage of FCM and minimize the impact of erroneous contrast. As shown in Fig. 3, this process can be repeated iteratively, and the area that needs bias field correction is shrinking after each iteration.

The new algorithm combing N3+FCM yields a significant improvement in the segmentation quality. As shown in the radiologist's blind evaluation results, during the first reading session, the combined approach is superior to N3 in 24 of the 60 breasts, with equality quality in 34 cases, and inferior in only 2 cases. A more favorable result was found in the second reading session. The combined approach is superior to N3 in 27 cases, with equality quality in 32 cases, and inferior in only 1 case. Figures 5–7 show three examples. It can be seen that for cases with a very strong bias field (e.g., Fig. 5), N3+FCM is better than N3, and both are better than FCM. For cases without a discernable bias field (e.g., Fig. 7), correction is probably not needed, and all three methods produce good segmentation quality and are rated equally. Therefore, the choice of the correction method should be dependent on the quality of the image. The results based on these three case examples suggest that for a large breast that protrudes deep into the flat-bed breast coil, the coil cannot produce a large homogeneous field to cover the whole breast and, consequently, the acquired images show a strong field inhomogeneity. On the other hand, for a small breast that does not protrude deep, the coil can produce a homogeneous field to cover the whole breast, and the acquired images do not show a strong field inhomogeneity. For cases with small breasts, the FCM-based method works just as well as the other two. The images reported in our previous publication were acquired using a closed-form breast coil,[6–9] with the breast tissue fitted into a confined bra-shaped space. The produced images are more homogeneous, and FCM-based correction method is sufficient.

We also compared the segmentation results based on the N3+FCM correction method to a recent method of bias field estimation reported by Li *et al.*[30] based on a CLIC criterion function. This segmentation-based bias field correction method modifies the standard FCM clustering energy function by assuming that the variation in the neighbor intensity can be modeled by a truncated Gaussian kernel. Based on this assumption, the inhomogeneity of a pixel is affected by its neighborhood, which follows the spatially coherent nature of bias field. Therefore, CLIC ensures the smoothness of the estimated bias field naturally during iteration, while the standard FCM requires an additional low-pass filtering, which may suffer from noise and local minima. Our results clearly show that CLIC is more robust and accurate compared to the standard FCM. Unlike CLIC, which is segmentation-based, the proposed N3+FCM uses the combination of segmentation, surface fitting and histogram-based approaches. The algorithm uses N3 to maximize the high frequency content of intensity distribution and remove the major bias field based on the global maxima, and then uses FCM to detect and remove the residual bias field. As shown in Sec. III, both N3+FCM and CLIC algorithms can remove strong bias field and provide satisfactory segmentation results for all cases analyzed in this study. Since they use different approaches, they may provide alternative options for other segmentation studies involving a strong bias field that needs to be corrected.

Generally, bias field correction is a preprocessing step in order to improve the quality of the segmentation results. If the gold standard (ground truth segmentation, e.g., BrainWeb: http://www.bic.mni.mcgill.ca/brainweb/) is available, the quality of the resulted segmentation based on different correction methods can be evaluated using quantitative metrics, e.g., Jaccard similarity.[30] Unfortunately, there is no gold standard method for segmentation of fibroglandular tissue in the breast. Although manual segmentation may work well for some images (e.g., Fig. 4), for majority of cases showing scattered fibroglandular tissue distributions (e.g., Fig. 6), manual segmentation will not yield a consistent ground truth. Some studies[25,26] used a manually trained minimum distance classifier[32] together with manual edition for evaluation, and the variation in intensity within a tissue class is assessed by the coefficients of joint variation (CJV) based on the assumption that a lower CJV corresponds to a better inhomogeneity correction. While this assumption is true for brain images that are normally acquired using a volumetric coil, it is not applicable for breast images that are acquired using a surface coil, which is known to have strong location-dependent intrinsic inhomogeneities. Within each class of fibroglandular or fatty tissues, the intensities would vary between different image slices, as well as between anterior and posterior region on the same image. Therefore, in this study the evaluation of the segmentation quality was made based on the visual reading of an experienced radiologist. Although it is not based on a quantitative criterion, it is a feasible way to yield reliable evaluation results from many imaging slices contained within a 3D volume of a breast.

In summary, in this work, we described a new bias field correction method by combining the N3 and FCM-based inhomogeneity correction algorithm. This algorithm first utilizes the advantage of N3 for a global correction and then by iteration of FCM and B-spline fitting to gradually correct the bias field presented on the original images without erroneously changing the tissue contrast. It is shown that the N3+FCM method can lead to an improved segmentation quality compared to using either the N3 or the FCM method alone, and that the performance was equivalent to a new method, CLIC, proposed by Li *et al.*[30] The N3+FCM and CLIC methods are both useful for correcting the MR images with a severe regional bias field, which is commonly presented in the MR images of large breasts acquired using a flat-bed breast coil. Choosing an appropriate bias field correction method is a very important preprocessing step to allow an accurate segmentation of the fibroglandular tissues based on breast MRI for measurement of breast density.

ACKNOWLEDGMENTS

This work was supported, in part, by the NIH (Grant Nos. R01 CA127927 and R03 CA136071), by the California BCRP (Grant Nos. 14GB-0148 and 16GB-0056), and by the National Science Council of Taiwan (Grant No. NSC-98-2221-E-039-009).

[a]This work was conducted at Tu and Yuen Center for Functional Onco-Imaging at University of California, Irvine.
[b]Author to whom correspondence should be addressed. Electronic mail: msu@uci.edu; Telephone: (949) 824-4925; Fax: (949) 824-3481.

[1]N. F. Boyd, G. S. Dite, J. Stone, A. Gunasekara, D. R. English, M. R. McCredie, G. G. Glies, D. Tritchler, A. Chiarelli, M. J. Yaffe, and J. L. Hopper, "Heritability of mammographic density, a risk factor for breast cancer," N. Engl. J. Med. **347**, 886–894 (2002).
[2]M. J. Yaffe, N. F. Boyd, J. W. Byng, R. A. Jong, E. Fishell, G. A. Lockwood, L. E. Little, and D. K. Tritchler, "Breast cancer risk and measured mammographic density," Eur. J. Cancer Prev. **7**, S47–S55 (1998).
[3]V. A. McCormack and I. dos Santos Silva, "Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis," Cancer Epidemiol. Biomarkers Prev. **15**, 1159–1169 (2006).
[4]C. M. Vachon, V. S. Pankratz, C. G. Scott, S. D. Maloney, K. Ghosh, K. R. Brandt, T. Milanese, M. J. Carston, and T. A. Sellers, "Longitudinal trends in mammographic percent density and breast cancer risk," Cancer Epidemiol. Biomarkers Prev. **16**, 921–928 (2007).
[5]R. J. Santen, N. F. Boyd, R. T. Chlebowski, S. Cummings, J. Cuzick, M. Dowsett, D. Easton, J. F. Forbes, T. Key, S. E. Hankinson, A. Howell, and J. Ingle, "Critical assessment of new risk factors for breast cancer: Considerations for development of an improved risk prediction model," Endocr. Relat. Cancer **14**, 169–187 (2007).
[6]K. 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**, 5253–5262 (2008).
[7]K. Nie, D. Chang, J. H. Chen, C. C. Hsu, O. Nalcioglu, and M. Y. Su, "Quantitative analysis of breast parenchymal patterns using 3D fibroglandular tissues segmented based on MRI," Med. Phys. **37**, 217–226 (2010).
[8]K. Nie, D. Chang, J. H. Chen, T. C. Shih, C. C. Hsu, O. Nalcioglu, and M. Y. Su, "Impact of skin removal on quantitative measurement of breast density using MRI," Med. Phys. **37**, 227–233 (2010).
[9]J. H. Chen, K. Nie, S. Bahri, C. C. Hsu, F. T. Hsu, H. N. Shih, M. Lin, O. Nalcioglu, and M. Y. Su, "Decrease in breast density in the contralateral normal breast of patients receiving neoadjuvant chemotherapy: MR imaging evaluation," Radiology **255**, 44–52 (2010).
[10]J. Yao, J. A. Zujewski, J. Orzano, S. Prindiville, and C. Chow, "Classification and calculation of breast fibroglandular tissue volume on SPGR fat suppressed MRI," Proc. SPIE **5747**, 1942–1949 (2005).
[11]H. Li, Y. Chu, A. F. Salem, and R. A. Clark, "Image segmentation and 3D visualization for MR mammography," Proc. SPIE **4684**, 1780–1789 (2002).
[12]J. Wei, H. P. Chan, M. A. Helvie, M. A. Roubidoux, B. Sahiner, C. Zhou, S. Paquerault, and M. M. Goodsitt, "Correlation between mammographic density and volumetric fibroglandular tissue estimated on breast MR images," Med. Phys. **31**, 933–942 (2004).
[13]C. Klifa, J. Carballido-Gamio, L. Wilmes, A. Laprie, C. Lobo, J. Gibbs, and N. Hylton, "Quantification of breast tissue index from MR data using fuzzy cluster," IEEE Eng. Med. Biol. Mag. **3**, 1667–1670 (2004).
[14]M. 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 [Collaborators in the United Kingdom Medical Research Council Magnetic Resonance Imaging in Breast Screening (MARIBS) Study], "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**, 2268–2274 (2008).
[15]D. J. Thompson, M. O. Leach, G. Kwan-Lim, S. A. Gayther, S. J. Ramus, I. Warsi, F. Lennard, M. Khazen, E. Bryant, S. Reed, C. R. Boggis, D. G. Evans, R. A. Eeles, D. F. Easton, and R. M. Warren [The UK study of MRI screening for breast cancer in women at high risk (MARIBS)], "Assessing the usefulness of a novel MRI-based breast density estimation algorithm in a cohort of women at high genetic risk of breast cancer: The UK MARIBS study," Breast Cancer Res. **11**, R80 (2009).
[16]C. 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**, 8–15 (2010).
[17]U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," IEEE Trans. Med. Imaging **26**, 405–421 (2007).
[18]A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge, "Phase and sensitivity of receiver coils in magnetic resonance imaging," Med. Phys. **13**, 121–128 (1994).
[19]W. W. Brey and P. A. Narayana, "Correction for intensity falloff in surface coil magnetic resonance imaging," Med. Phys. **15**, 241–245 (1988).
[20]J. Y. Chiou, C. B. Ahn, L. T. Muftuler, and O. Nalcioglu, "A simple simultaneous geometric and intensity correction method for echo-planar imaging by EPI-based phase modulation," IEEE Trans. Med. Imaging **22**, 200–205 (2003).
[21]M. S. Cohen, R. M. DuBois, and M. M. Zeineh, "Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging," Hum. Brain Mapp **10**, 204–211 (2000).
[22]L. Q. Zhou, Y. M. U. Zhu, C. Bergot, A. M. Laval-Jeantet, V. Bousson, J. D. Laredo, and M. Laval-Jeantet, "A method of radio-frequency inhomogeneity correction for brain tissue segmentation in MRI," Comput. Med. Imaging Graph. **25**, 379–389 (2001).
[23]P. Vemuri, E. G. Kholmovski, D. L. Parker, and B. E. Chapman, "Coil sensitivity estimation for optimal SNR reconstruction and intensity inhomogeneity correction in phased array MR imaging," presented at the Proceedings of the 19th International Conference of IPMI, Glenwood Springs, CO, 2005 (unpublished).
[24]E. A. Vokurka, N. A. Watson, Y. Watson, N. A. Thacker, and A. Jackson, "Improved high resolution MR imaging for surface coils using automated intensity non-uniformity correction: Feasibility study in the orbit," J. Magn. Reson Imaging **14**, 540–546 (2001).
[25]J. 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**, 87–97 (1998).
[26]B. Likar, M. A. Viergever, and F. Pernus, "Retrospective correction of MR intensity inhomogeneity by information minimization," IEEE Trans. Med. Imaging **20**, 1398–1410 (2001).
[27]J. D. Gispert, S. Reig, J. Pascau, J. J. Vaquero, P. Garcia-Barreno, and M. Desco, "Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error," Hum. Brain Mapp **22**, 133–144 (2004).
[28]A. H. Andersen, Z. Zhang, M. J. Avison, and D. M. Gash, "Automated segmentation of multispectral brain MR images," J. Neurosci. Methods **122**, 13–23 (2002).
[29]W. Chen and M. L. Giger, "A fuzzy C-means (FCM) based algorithm for intensity inhomogeneity correction and segmentation of MR images," International Symposium on Biomedical Imaging (ISBI), pp. 1307–1310, 2004 (unpublished).
[30]C. Li, C. Xu, A. W. Anderson, and J. C. Gore, "MRI tissue classification and bias field estimation based on coherent local intensity clustering: A unified energy minimization framework," Proceedings of Information Processing in Medical Imaging (IPMI), pp. 288–299, 2009 (unpublished).
[31]D. Hearn and M. P. Baker, *Computer Graphics C Version*, 2nd ed. (Prentice-Hall, Upper Saddle River, NJ, 1997).
[32]J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms* (Springer, New York, 1981).