Towards Accurate Segmentation of Fibroglandular Tissue in Breast MRI Using Fuzzy C-Means and Skin-Folds Removal

Mohammad Razavi[1], Lei Wang[1(✉)], Albert Gubern-Mérida[2], Tatyana Ivanovska[3], Hendrik Laue[1], Nico Karssemeijer[2], and Horst K. Hahn[1]

[1] Fraunhofer MEVIS - Institute for Medical Image Computing, Bremen, Germany
lei.wang@mevis.fraunhofer.de
[2] Radboud University Medical Centre, Nijmegen, The Netherlands
[3] Ernst-Moritz-Arndt University Greifswald, Greifswald, Germany

Abstract. Breast density measuring the volumetric portion of fibroglandular tissue is considered as an important factor in evaluating breast cancer risk of women. Categorizing breast density into different levels by human observers is time-consuming and subjective, which may result in large inter-reader variability. In this work, we propose a fully automated fibroglandular tissue segmentation technique aiming to assist automatic breast density measurement in magnetic resonance imaging (MRI). Firstly, a bias field correction algorithm is applied. Secondly, the breast mask is segmented to exclude air background and thoracic tissues, such as liver, heart and lung. Thirdly, the segmentation is further refined by removing the skin-folds that are normally included in the breast mask and mimic the fibroglandular tissue, leading to incorrect density estimation. Finally, we apply a fuzzy c-means approach to extract the fibroglandular tissue within the breast mask. To quantitatively evaluate the proposed method, a total of 50 MR scans were collected. By comparing the volume overlap between manually annotated fibroglandular tissue with the results of our method, we achieved an average Dice Similarity Coefficient (DSC) of 0.84.

Keywords: Breast · MRI · Fibroglandular · Skin-folds

1 Introduction

Breast density is classified into four groups in the standardized report of Breast Imaging-Reporting and Data System (BI-RADS) proposed by American College of Radiologists (ACR) [5]. Dense breasts classified into groups 3 and 4 have more fibrous and glandular tissue that may obscure small masses and thus lower the sensitivity of mammography (MG). Women with dense breasts have been shown to have a four to six-fold increased risk of developing breast cancer [6]. However, the classification of an individual breast depends on the opinion of radiologists, which leads to higher inter-reader variability. Automated density quantification based on breast MG and magnetic resonance imaging (MRI)

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part I, LNCS 9279, pp. 528–536, 2015.
DOI: 10.1007/978-3-319-23231-7_47

allows for fast and reproducible assessment and thus decreases variability. As a better imaging modality to measure breast density, breast MRI scans the entire breast volumes in 3D without any tissue overlapping or projection. Therefore, density measurement in MRI tends to be more accurate than MG [4].

Several works were published aiming to segment fibroglandular tissue in breast MRI. Most algorithms used intensity-based approaches, such as fuzzy c-means (FCM) or Gaussian mixture (GM) models. Gubern-Mérida et al. developed a fibroglandular tissue segmentation using a GM model based on an atlas-based approach [2] . The overall average Dice Similarity Coefficient (DSC) reported was 0.80. Nie et al. proposed a FCM based method which was applied on a semi-automatically segmented breast mask, which requires users interactions to identify important landmarks [8]. Wu et al. adopted an atlas-aided FCM approach to segment fibroglandular tissue, which requires again large training set for atlas construction and might encounter difficulties to cope with new testing images acquired from other sites with different imaging protocols [13]. Moreover, the best average DSC achieved from their experiments was 0.69. Most recent work published by Ivanovska et al. used level-set based gradual method which simultaneously corrects bias field and segments fibroglandular tissue [3]. The method processed input images in 2D slice-by-slice and achieved DSC of 0.83 on average.

In this work, we propose a fully automated fibroglandular tissue segmentation framework based on robust breast segmentation and a skin-folds removal procedure. The entire work flow consists of four major steps which are illustrated in Fig. 1. First, a bias field correction algorithm is applied to alleviate intensity inhomogeneity of each tissue type. Second, breast region is extracted to exclude air background and irrelevant thoracic tissue, such as lung, liver and heart. Third, the breast segmentation is further refined by removing skin-folds which are typical false positive structures mimicking fibroglandular. Finally, we adopt a FCM algorithm to classify breast volume into fatty and fibroglandular classes. The performance of the proposed method is tested on 50 MR scans acquired from 50 different subjects. The volumetric overlap between manual annotations and segmented results is measured.

2 Material and Method

2.1 Material

For this study, we used a set of 50 coronal T1-weighted MR breast volumes from 50 different patients collected within the years 2003 and 2009 from Radboud University Medical Center. Patients were scanned in prone position. The age of screened women ranged from 23 to 76 years (45.84 ± 11.97 on average). The breast MRI examinations were performed on either a 1.5 or 3 Tesla Siemens scanner (Magnetom Vision, Magnetom Avanto and Magnetom Trio), with a dedicated breast coil (CP Breast Array, Siemens, Erlangen). The clinical imaging parameters varied; matrix size: 256×128 or 256×96; slice thickness: 1.3 mm; slice spacing: 0.625 - 1.25 mm; flip angle: 8, 20 or 25 degrees; repetition time: 7.5

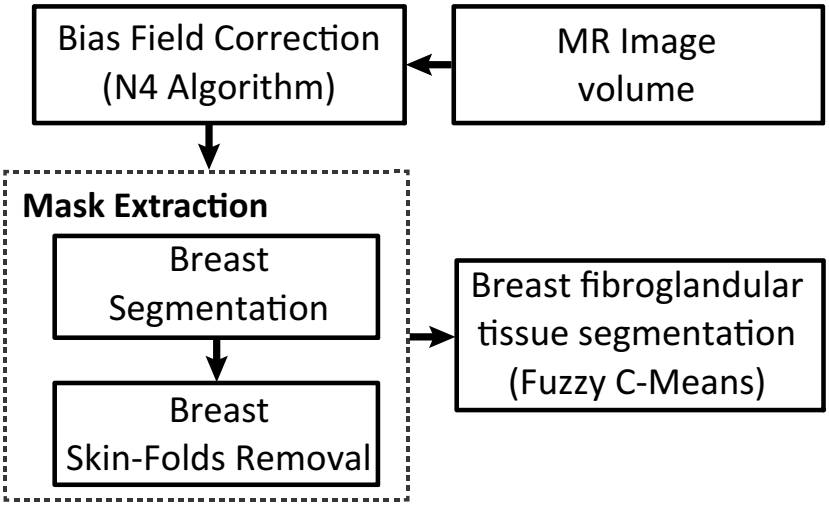

Fig. 1. General overview of the process for fibroglandular tissue segmentation in breast MRI.

- 9.8 ms; echo time: 1.7 - 4.76 ms. Each MRI exam in the test set was manually segmented into 7 tissue classes by an experienced expert which are used as the reference image to compare the quality of segmentation. A detailed description of the annotation is given in [2].

2.2 Method

Bias field Correction. Intensity inhomogeneity in MRI can attribute to the imperfection in radio-frequency coils or to the problems associated with acquisition protocols. It results in a slowly varying intensity change over the image that can produce errors with conventional intensity-based methods to distinguish different tissue types. The N4 bias field correction algorithm which is used in this work is a variant of the popular non-parametric nonuniform intensity normalization (N3) algorithm introduced by Sled et al. [9]. By assuming that the corruption of low frequency bias field can be modeled as a convolution over the intensity histogram by a Gaussian kernel, the algorithm iterates the following steps: de-convolving the intensity histogram by a Gaussian; re-mapping the intensities; spatially smoothing the result by a B-spline model[10]. By removing such artifacts, segmenting different tissue types in MR images can be done more accurately (see Fig. 2).

Breast Segmentation. Breast mask extraction separates the breast area from the other body parts such as lung, heart, pectoral muscle as well as air background presented in MRI scans. We previously implemented a fully automatic segmentation method specially designed for processing non-fat suppressed breast

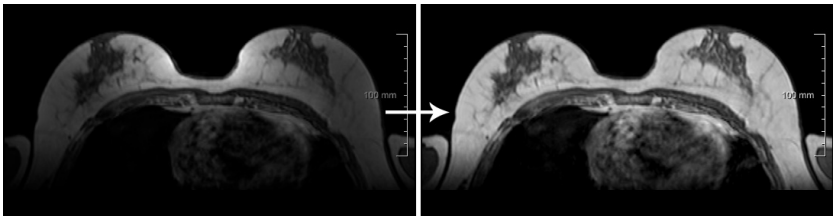

Fig. 2. (**Left**) The original breast MR image with the bias field artifacts from magnetic coil. (**Right**) The bias field corrected MR image using N4 algorithm.

MRI [12]. The key observation of this method is that the pectoral muscle and the breast-air boundaries exhibit as smooth sheet-like surfaces in 3D, which can be simultaneously enhanced by a Hessian-based sheetness filter [11]. The method consists of four major steps: enhancing sheet-like structures, segmenting the pectoral muscle boundary which defines the lower border of breast region, segmenting the breast-air boundary which delimits the upper border of the breast region, and extracting the region between the upper and lower borders that finally captures the area of breast tissue.

Skin-folds Removal. Skin-folds artifact could appear due to either large breasts do not entirely fit in the coils [2] or MRI technician's fault as breasts are not pulled perfectly into the coils while taking images [7]. Since the voxels of skin-folds have similar intensity levels with the fibroglandular tissue, normally they are included in the breast mask obtained in segmentation step (see Fig. 3), which will be erroneously recognized as the fibroglandular tissue in subsequent steps. Therefore, skin-folds need to be removed from the breast mask.

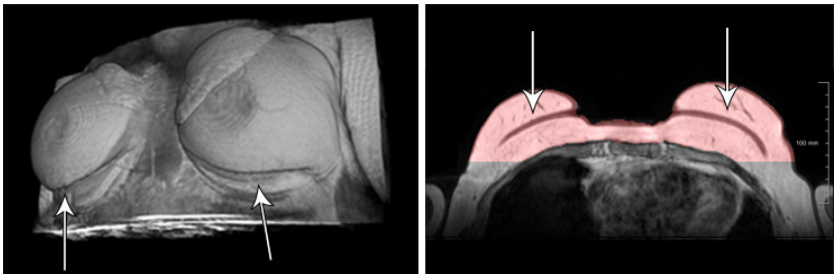

Fig. 3. On the **left**, the skin-folds artifact in 3D view. On the **right**, skin-folds which are usually included in the breast segmentation mask (red overlay area).

Based on the fact that skin-folds have lower intensity level compared to fatty tissue, we applied a first stage of FCM on entire MR volume to classify all the voxels into two classes i.e. dark and bright structures, which yields a binary

image. As shown in Fig. 4(c) skin-folds, air background, fibroglandular tissue and other thoracic tissues are classified into one class. Considering the 3D spatial connectivity of the skin-folds and the background, we then carry out a 3D region growing procedure on the binary images starting from the seeds in the background and propagating to the skin-folds. Although in 2D transversal view, the skin-folds might be partially surrounded by fatty tissue and not corrected to the background in several slices, the 3D region growing process guarantees these parts will be reached. For the cases where the fibroglandular tissue connects to the background near the nipples, region growing might leak into the fibroglandular tissue. To prevent the leakage near the nipple, we cut a small patch on the top of the segmented breast mask (see Fig. 4(b)) and paste it to the corresponding location in the binary images resulted from the first stage of FCM (see Fig. 4(c)). More specifically, the patch is obtained by cutting the breast mask through a cutting-line placed in the middle distance of the peak point (i.e., the closest point of the breast mask to the top border of image in transversal view) and concavity point (i.e., the first intersection of the breast mask with the center line) automatically detected on the breast mask (see Fig. 4(b)). After blocking the leakage, the region growing procedure results in a union of connected dark structures, such as background, skin-folds, lungs etc, but except fibroglandular tissue (see Fig. 4(d)). Hence, a subtraction between the breast mask and the region growing results will produce a refined breast mask excluding the skin-folds (see Fig. 4(e)).

Fibroglandular Tissue Segmentation. In the previous pre-processing steps, intensity inhomogeneity is corrected, and a breast mask excluding air background and skin-folds artifact is obtained. Within the breast mask, a second stage of FCM is carried out. The class number is set to three, to capture the fatty tissue, the fibroglandular tissue and any transition structures resulting from either partial volume effect or imperfect inhomogeneity correction. The class with the lowest mean intensity level is recognized as the fibroglandular tissue.

3 Evaluation

To evaluate our approach, automatic segmentation of 50 MRI scans were compared to the reference masks created manually. The agreement between the reference annotations and the segmented results was measured by calculating Dice Similarity Coefficient (DSC) and Jaccard Coefficient (JC) and Absolute Volume Error (AVE), which measure the volumetric overlap between segmented and reference volumes. The definitions of the volumetric metrics are given in the following equations:

$$DSC(V_{Ref}, V_{Seg}) = \frac{2 \times |V_{Ref} \bigcap V_{Seg}|}{|V_{Ref}| + |V_{Seg}|} \tag{1}$$

$$JC(V_{Ref}, V_{Seg}) = \frac{|V_{Ref} \bigcap V_{Seg}|}{|V_{Ref} \bigcup V_{Seg}|} \tag{2}$$

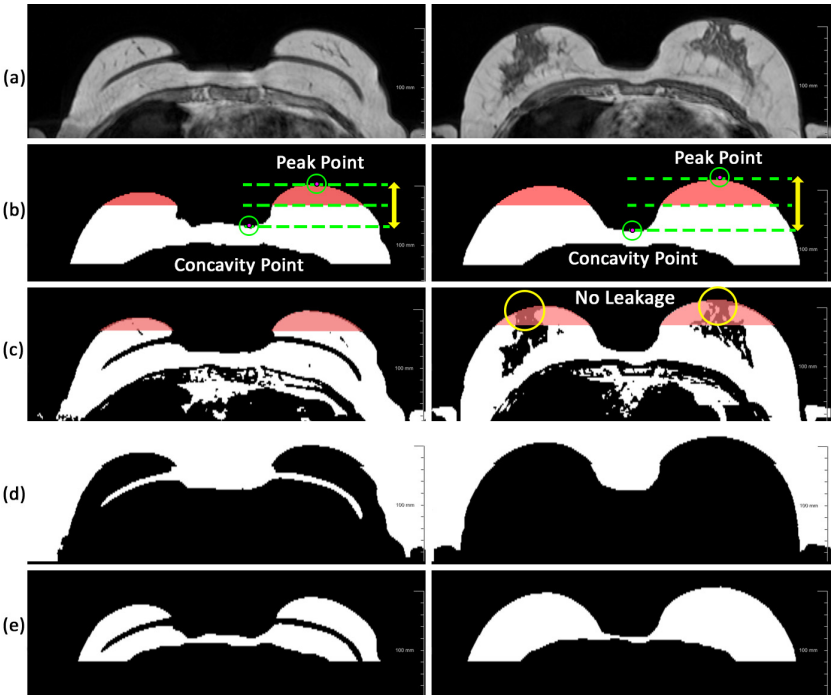

Fig. 4. Skin-folds removal and leakage prevention procedures applied on two representative slices (left: the skin-folds are present; right: the fibroglandular tissue near the nipples are connecting to the background). **(a)** results of bias field correction; **(b)** the segmented breast masks and the obtained patches (red overlay) by identifying the cutting-lines between the peak and concavity points. **(c)** the binary masks achieved from the first stage FCM with overlay of the patches that block the leakage near the nipples (see right figure). **(d)** the results of region growing. **(e)** the refined breast masks without the skin-folds (labeled as background) by subtracting **(b)** and **(d)**.

$$AVE(V_{Ref}, V_{Seg}) = |V_{Seg}/V_{Ref} - 1| * 100 \tag{3}$$

where V_{Seg} and V_{Ref} represent the segmentation and reference volumes. Additionally, average symmetric Root Mean Square Distance (asRMSD) between the boundary surfaces of the two volumes is computed [1]. First, the boundary voxels of segmentation and reference are determined. For each voxel in one set, the closest voxel in the other set is determined (using RMSD). All these distances are stored for boundary voxels from both reference and segmentation. The average of all these distances gives the averages symmetric distance.

The evaluation results are summarized in Table 1, which shows how the skin-folds removal and the N4 bias field correction influence the accuracy of the results. As seen in the first row, the best results were achieved by using a combination of both steps, since typical false positive segmentation in terms of skin-folds and inhomogeneous fatty tissue are removed. By combining both,

an average DSC of 0.84 with a standard deviation of 0.08 was achieved, which resulted in the best volumetric overlap and lowest boundary surface distance. Although taking advantage of each steps adds up more time to the whole segmentation process, we finally managed to optimize the computing process and ended up with an average computation time of 39.78 seconds per case.

In Fig. 5, a comparison is made between the references and the automated segmentation among the cases with the highest and lowest DSC values. By observing the results, it turns out that the best outcomes are mostly among the cases with dense breasts and the worst are among the ones with major fatty tissues. The remaining errors in the results are either due to inaccurate segmented breast masks, or due to incomplete manual annotations in reference images.

Table 1. The statistical results (mean ± standard deviation) of the metrics and time comparison using the skin-folds removal (SFR) and N4 bias field correction methods and without using them.

Method	DSC	asRMSD	JC	AVE	Time/Case
with N4, with SFR	0.84±0.08	4.22±2.82	0.73±0.12	18.61±14.48	39.78 s
with N4, no SFR	0.72±0.18	8.44±3.26	0.60±0.20	18.83±16.43	28.38 s
no N4, with SFR	0.61±0.25	7.04±2.54	0.48±0.25	44.84±23.66	24.58 s
no N4, no SFR	0.57±0.24	8.38±2.12	0.44±0.23	50.09±26.83	13.78 s

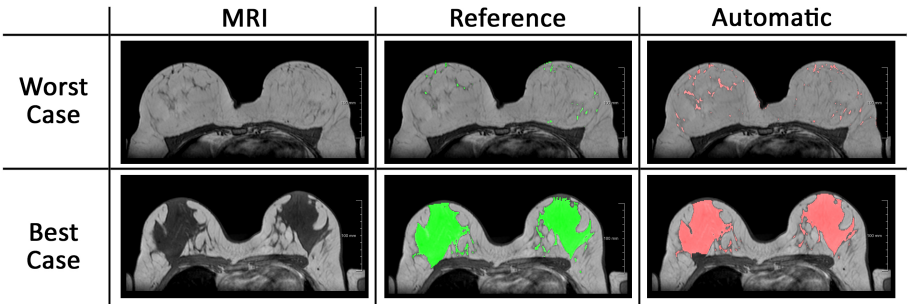

Fig. 5. Comparison of reference and automated segmentation of one axial slice from the worst case (**top**) with DSC of 0.4 and best case (**bottom**) with DSC of 0.95.

4 Conclusion and Future Works

The presented framework for automatic fibroglandular tissue segmentation shows high quality results in most of the testing cases. By incorporating skin-folds removal and bias filed correction steps, the method is more robust against intensity inhomogeneity and skin-folds artifacts. The experiments proved the importance of these pre-processing steps. Compared to previous works, we achieved

slightly higher overall average DSC (0.84) than the work reported by Gubern-Mérida et al. [2] (0.80) by using the same datasets. In addition, our results showed higher accuracy than the work by Wu et al. [13] with DSC of 0.69. Our method is comparable to the work reported by Ivanovska et al. [3], who achieved an average DSC of 0.83 using different datasets comprising 37 MRI scans. Nevertheless, we find that FCM is quite sensitive to the remained inhomogeneity that is not fully recovered by the bias field correction. Therefore, an advanced multi-dimensional FCM, which classifies the voxels not only based on their intensity similarities but also spatial connectivity, will be considered to further improve the segmentation quality in the future works.

Acknowledgments. This research leading to these results has received funding from the European Unions Seventh Framework Programme FP7 under grant agreement No. 306088.

References

1. Gerig, G., Jomier, M., Chakos, M.: Valmet: a new validation tool for assessing and improving 3d object segmentation. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 516–523. Springer, Heidelberg (2001)
2. Gubern-Merida, A., Kallenberg, M., Mann, R., Marti, R., Karssemeijer, N.: Breast segmentation and density estimation in breast MRI: a fully automatic framework. IEEE Journal of Biomedical and Health Informatics **19**(1), 349–57 (2015)
3. Ivanovska, T., Laqua, R., Wang, L., Liebscher, V., Völzke, H., Hegenscheid, K.: A Level set based framework for quantitative evaluation of breast tissue density from MRI data. PloS One **9**(11), November 2014
4. Lee, N.A., Rusinek, H., Weinreb, J., Chandra, R., Toth, H., Singer, C., Newstead, G.: Fatty and fibroglandular tissue volumes in the breasts of women 20–83 years old: comparison of X-ray mammography and computer-assisted MR imaging. AJR. American journal of roentgenology **168**(2), 501–506 (1997)
5. Liberman, L., Menell, J.H.: Breast imaging reporting and data system (BI-RADS). Radiologic Clinics of North America **40**(3), 409–430 (2002)
6. McCormack, V.A., dos Santos Silva, I.: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer epidemiology, biomarkers, prevention **15**(6), 1159–1169 (2006)
7. Morris, E.A., Liberman, L.: Pitfalls in analysis of carcinomas. Breast MRI: Diagnosis and Intervention, pp. 488–501 (2005)
8. Nie, K., Chen, J.H., Chan, S., Chau, M.K.I., Yu, H.J., Bahri, S., Tseng, T., Nalcioglu, O., Su, M.Y.: Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. Medical Physics **35**(12), 5253–5262 (2008)
9. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging **17**(1), 87–97 (1998)
10. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4itk: improved N3 bias correction. IEEE Transactions on Medical Imaging **29**(6), 1310–1320 (2010)

11. Wang, L., Filippatos, K., Friman, O., Hahn, H.: Fully automated segmentation of the pectoralis muscle boundary in breast MR images. In: SPIE Medical Imaging, pp. 796309–796309 (2011)
12. Wang, L., Platel, B., Ivanovskaya, T., Harz, M., Hahn, H.K.: Fully automatic breast segmentation in 3D breast MRI. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1024–1027 (2012)
13. Wu, S., Weinstein, S.P., Conant, E.F., Kontos, D.: Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method. Medical Physics **40**(12), 122302 (2013)