A Novel Breast Image Preprocessing For Full Field Digital Mammographic Segmentation and Risk Classification

Wenda He[1]
weh@aber.ac.uk
Minnie Kibiro[2]
minnie.kibiro@nnuh.nhs.uk
Arne Juette[2]
arne.juette@nnuh.nhs.uk
Peter Hogg[3]
p.hogg@salford.ac.uk
Erika R. E. Denton[2]
erika.denton@nnuh.nhs.uk
Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1] Department of Computer Science, Aberystwyth University, Aberystwyth, SY23 3DB, UK
[2] Department of Radiology, Norfolk & Norwich University Hospital, Norwich, NR4 7UY, UK
[3] School of Health Sciences, University of Salford, Salford, M6 6PU, UK

Abstract

To obtain optimal breast image quality during the image acquisition, a compression paddle is used to even the breast thickness. Clinical observation has indicated that breast peripheral areas may not be fully compressed, and may cause unexpected intensity and texture variation within these areas. Such breast parenchymal appearance discrepancies may not be desirable for tissue modelling within computer aided mammography. This paper describes a novel mammographic image preprocessing method to improve the image quality before analysis. Mammographic segmentation and risk classification were performed to facilitate a quantitative and qualitative evaluation, using digital mammographic images. Visual assessment indicated significant improvement on segmented anatomical structures and tissue specific areas when using the processed images. The achieved risk classification accuracies are 80% and 79% for Birads and Tabár risk scheme, respectively. The developed method has demonstrated an ability to improve the quality of mammographic segmentation, leading to more accurate risk classification. This in turn can be found useful in early breast cancer detection, risk-stratified screening, and aiding radiologists in the process of decision making prior to surgery and/or treatment.

1 Introduction

Within screening mammography, a compression paddle is used to even out the breast tissue in order to obtain high quality mammographic images. However, when the breast is subjected to compression, the peripheral areas may not be compressed due to a reduction

of breast thickness. This results in air gaps above and beneath the uncompressed areas, leading to X-ray scattering, degradation in contrast, and limitation of the quantitative usefulness of radiographic images [2]. Compression paddle related image quality issues vary due to difference in breast size and composition; *e.g*. in extreme cases, the visibility of breast parenchyma is too low in the peripheral areas to be examined. Therefore, it is necessary to develop an image processing method to improve the visibility of peripheral uncompressed area of the projected breast, which facilities presentation and can be beneficial to mammographic analysis [1]. Image processing (enhancement) methods developed to address the aforementioned problem can be categorised into two groups; parametric [1, 8] and non-parametric [9] approaches. Methods [8] and [9] are critically dependent on the accuracy of interactive segmentation of the dense breast tissue and fatty tissue interpolation, whilst [9] can only be applied to unprocessed digital mammograms before logarithmic transformation. The reader is referred to [1, 8, 9] for the details of the developed methodologies.

Regarding mammoraphic risk assessment, Tabár *et al*. [10] proposed a model based on a mixture of four mammographic building blocks representing normal breast anatomy, mammographic risk is categorised into five classes along these building blocks (*i.e*. [nodular%, linear%, homogeneous%, radiolucent%]). Pattern I [25%, 15%, 35%, 25%]; pattern II/III [2%, 14%, 2%, 82%]; pattern IV [49%, 19%, 15%, 17%]; and pattern V [2%, 2%, 89%, 7%] [10]. Alternatively, American College of Radiology's Breast Imaging Reporting and Data System (Birads) [6] was developed, mammographic risk is categorised based on the percentage of dense breast tissue in four risk classes: Birads 1, the breast is almost entirely fat ($< 25\%$ glandular); Birads 2, the breast has scattered fibroglandular densities ($25\% - 50\%$); Birads 3, the breast consists of heterogeneously dense breast tissue ($51\% - 75\%$); and Birads 4, the breast is extremely dense ($> 75\%$ glandular).

We present a novel mammographic image preprocessing technique, which models a breast by estimating relative breast thickness ratios using both Mediolateral Oblique (MLO) and Cranio-Caudal (CC) views, the correction process is marginally similar to [5, 11]. An additional automatic selection method was developed to better target images requiring enhancement in a systematic way. A quantitative and qualitative evaluation was conducted to assess the robustness of the method; all processed images were subjected to mammographic segmentation and subsequent risk classification using both Tabár and Birads risk scheme.

2 Data and Method

A total of 360 digital mammographic images were used in the experiment. A consensus ground truth was obtained based on three radiologists. Modelling Tabár's mammographic building blocks, a collection of mammographic patches consisting of a total of 344, 89 and 457 examples of nodular, homogeneous and radiolucent (similar to Birads dense, semi-dense and non-dense) tissue was cropped to various sizes from randomly selected images from the dataset. Note that the linear structure was not specifically included in the experiment, instead this tissue type was considered part of the other tissue classes as it appears in combination with all three classes. Therefore, only nodular, homogeneous and radiolucent tissues were modelled and used in the segmentation process.

The developed methodology starts off with automatic image selection, the image preprocessing consists of four steps; 1) X-ray penetration probability weighting map generation, 2) intensity balancing, 3) intensity ratio propagation, and 4) boundary stitching. All the processed images were subjected to geometric moments based mammographic segmentation [3]

and risk classification.

2.1 Automatic Image Selection

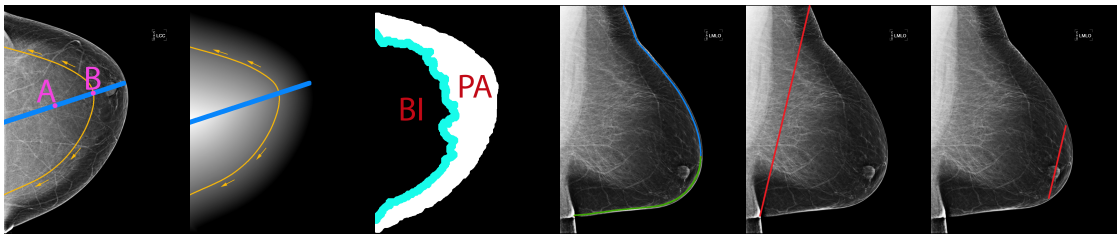
Thorough empirical observations of the collected data indicated that a mammographic image may require processing if its automatically calibrated parameters *acp* (*i.e.* compression force (*CF*), breast thickness (*BT*) and peak kilovotage (*KVP*)) are less than the mean values. It was less robust to use the *acp* alone as a constraint due to large breast tissue density, composition and size variations. Additional constraints based on prior knowledge were incorporated to strengthen the selection robustness. Specifically, Otsu [7] automatic binary image segmentation technique was used to separate the peripheral area (*PA*) and breast interior (*BI*). Three properties were calculated: total peripheral area (*TPA*), vertical peripheral area coverage (*VPAC*) in image rows and pectoral coverage (*PC*) in image rows. A mammographic image (*img*) was selected to be processed if $CF_{img} <= \overline{CF}$ & $BT_{img} <= \overline{BT}$ & $KVP_{img} <= \overline{KVP}$ & $15\% < TPA_{img} < 50\%$ & $VPAC_{img} > 75\%$ & $PC_{img} < 30\%$. The threshold values were empirically defined through trial and error, to achieve the best separation overall.

2.2 Mammographic Image Preprocessing

The X-ray penetration probability has a direct correlation with breast thickness. Due to physical complexity (*e.g.* unknown combination factors in X-ray beam spectrum and breast tissue composition), the X-ray penetration probability was modelled in a simplified way by encompassing all other elements (*e.g.* dosage, filter, anode) in a "black box" only considered as inversely proportional to the breast thickness. An X-ray penetration probability weighting map was generated for each image by calculating and propagating the relative breast thickness ratios based on its pair. For example, to a CC view, the relative breast thickness ratio (r) can be estimated based on the projected physical contour of the compressed breast as seen on MLO view. In particular, the skinline was firstly extracted from the MLO view and split in two at the furthest pixel to the chest wall to form upper and lower skinlines (*e.g.* blue and green lines in Figure 1 (c)). For each pixel in the top skinline, a corresponding pixel was sought in the lower skinline, to form a parallel line (*p-line*) (*e.g.* red line in Figure 1 (d)) to the chest wall by linking the two pixels. This process was repeated for all the pixels in the top skinline, and resulting in a series of parallel lines (*e.g.* Figure 1 (d)(e)) . To the CC view, the r at a given point (p) (*e.g.* 'A' in Figure 1 (a)) is calculated based on the boundary pixel (p_{base}) (*e.g.* 'B' in Figure 1 (a)) which separates *PA* and *BI* as $r = \frac{p-line(p)}{p-line(p_{base})}$); both pixels are on the thickest projected section (*e.g.* blue lines in Figure 1 (a)(b)) on the CC view. To complete the X-ray penetration probability weighting map for the CC view, the remaining pixels on the thickest projected section with the estimated breast thickness ratios were assigned in the same way, and the calculated ratios were propagated to the pixels that have the same distance to the skinline (S) (*e.g.* pixels on the yellow lines in Figure 1 (a)(b)) .

To reduce the intensity distribution variation, a base weight (w_{base}) was firstly calculated as $w_{base} = \frac{\sum_{i=0}^{N} W_i(x,y)}{N}, \forall W_i(x,y) \in S$, where W and S denote the weighting map and the boundary between *PA* (uncompressed breast peripheral area) and *BI* (breast area not part of *PA*). For each pixel within the *BI*, the intensity value *P(x, y)* was altered as $P'(x,y) = \frac{w_{base}}{W(x,y)} P(x,y)$.

After the intensity balancing, the local intensity ratio was propagated as a means of improving tissue appearance in *PA* (similar to the process described in [11]). From pixels at the boundary S to the skinline within the *PA*, each intensity value was altered by calculating



(a) (b) (c) (d) (e) (f)

Figure 1: Image illustration, from left to right: (a) CC view, (b) its distance map (pixel to skinline), (c) areas of *PA*, *BI*, and *S* (cyan line), (d) pair MLO view, (e) parallel lines proximal to pectoral muscle and (f) proximal to nipple.

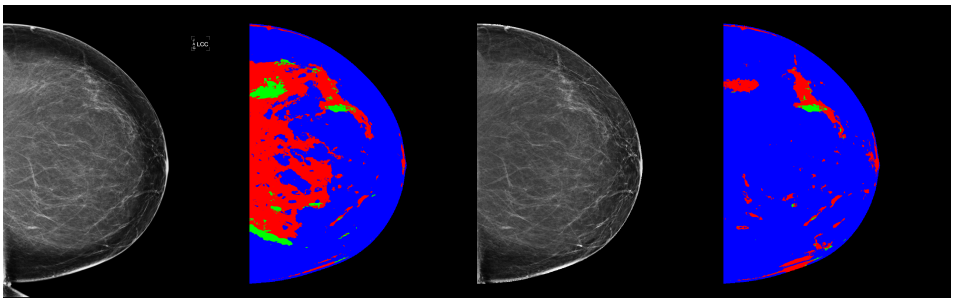
the propagation ratio (*pr*) for each pixel *P(x, y)* with distance to the skinline *D(x, y)* and within an empirically defined 17 $\times$ 17 neighbourhood for the efficiency and robustness, as $I_{avgP_1} = \frac{\sum_{j=0}^{M} P_j(x,y)}{M} \forall P_j(x,y) = D(x,y)+1$, $I_{avgP_2} = \frac{\sum_{i=0}^{N} P_i(x,y)}{N} \forall P_i(x,y) = D(x,y)+2$, $pr = \frac{I_{avgP_2}}{I_{avgP_1}}$, and $P'(x,y) = pr \times P(x,y)$; where $D(x,y)+1$ and $D(x,y)+2$ are pixel distances to skinline 1 and 2 steps further away from the observed pixel.

Boundary stitching was applied to seamlessly normalise pixel intensity within boundary *S*, thickened to 5 pixels band, in order to gradually smooth the transition from *BI* to *PA*; the maximum and minimum intensity values were determined within an empirically defined 7 $\times$ 7 neighbourhood (see Section 3 for the effects of using different neighbourhood sizes).

2.3 Mammographic Segmentation and Risk Classification

A geometric moments based mammographic segmentation [3] was applied to all the images. To be concise, this method is used to extract texture features from a set of mammographic patches using geometric moments; the derived feature vectors are expected to contain not only texture primitives but also geometric information. The methodology was modified by incorporating a feature and classifier selection process using a collection of attribute selection algorithms and classifiers available in Weka [4] through four stages: 1) A set of neighbourhoods (*i.e.* {7, 17, 27, 37, 47, 57, 67}) covering small to large anatomical structures were predefined and used in the feature extraction process over mammographic patches containing different breast tissue examples (*e.g.* nodular, homogeneous and radiolucent). 2) The raw feature vectors for all the patch pixels were filtered utilising all available filtering methods in Weka for reduction of attributes. Empirical testing indicated that the most frequent and prominent features are within the top half of the output attributes list after filtering. 3) The filtered feature vectors were then subjected to all available attribute selection (wrapper) methods in Weka to select the most discriminative subset of the filtered features. The most frequent attributes in the output subsets were used as the final selected optimal features. 4) All the available classifiers in Weka were used to perform (10-fold) cross-validation based evaluation over the selected optimal features. The classifier achieving the highest classification results (random committee in this study) was used in conjunction with the selected features in the mammographic segmentation using Tabár tissue modelling.

An unseen pixel's tissue class is determined by the trained classifier. The relative proportions of the mammograpic building blocks (tissue composition) were calculated from the resultant mammographic segmentation as features, and used in a leave one-image out (10-fold) cross-validation as a means of risk classification.



(a) (b) (c) (d)

Figure 2: (a) original image (Tabár II/Birads 1), (b) segmentation of (a) with over segmented glandular (*i.e.* red nodular and green homogeneous) tissue, (c) processed image, (d) segmentation of (c) showing more segmented blood vessels in the peripheral area, segmented glandular tissue is more realistic and fatty tissue (blue) is more correctly identified than (b).

Birads	1	2	3	4
1	149	21	0	0
2	17	100	6	0
3	0	12	23	8
4	0	0	7	17

Tabár	I	II/III	IV	V
I	122	21	4	0
II/III	19	128	1	0
IV	13	0	24	8
V	1	0	8	11

Table 1: Risk classification confusion matrices; Birads left, Tabár right.

3 Results and Discussion

Visual assessment was conducted to assess the quality of enhanced images. The majority of cases showed processed images to have improvements in textural appearance and contrast in the peripheral areas as expected; see Figure 2 (a)(c) for example. However, over and under enhancement may occur in some cases if *PA* and *BI* are not separated correctly, or the breast thickness ratios are wrongly estimated. During boundary stitching, undesirable artefacts may be created when a larger neighbourhood is used. Incorrect local intensity alternation affects textural appearance which can be perceived as an artefact. However, the processed images seem to have minimum textural distortion and were suitable for the follow up image analysis. The segmentation results have shown significant improvements in terms of correctness of the segmented anatomical structures over the breast parenchyma and tissue specific areas.

The risk classification accuracies increased on average 7% when compared with the results obtained before the image preprocessing was applied. The total classification accuracies were 80% and 79% for Birads and Tabár risk scheme, respectively. Table 1 shows the confusion matrices with respect to the two risk schemes when the developed methodology was applied. In both cases, mammographic images in high risk classes seem to have more misclassification (percentage wise), which may relate to the intensity over balancing for homogeneous tissue with structureless densities. The achieved risk classification results are in line with results achieved by the state-of-art method [11], but different segmentation principle (*e.g.* two class fatty/dense segmentation) and datasets were used. Further segmentation improvement can be made by widening the tissue patch variation, leading to a more robust classifier which in turn could also improve risk classification accuracies. However, this is outside the scope of this study and is considered focus for future investigation on mammographic segmentation and risk classification methodology.

4 Conclusions

The developed mammographic preprocessing can be used to reduce intensity and textural appearance discrepancies. Results have shown more anatomically accurate and consistent segmentation over the breast parenchyma when using the processed images in conjunction with the selected feature and classifier. This in turn improved subsequent risk classification accuracies. Incorporating the novel image preprocessing into mammographic segmentation methodology could prove useful in quantifying change in relative proportion of breast tissue, aiding radiologists' estimation in mammographic risk assessment, and providing risk-stratified screening to patients.

References

[1] J. W. Byng, J. P. Critten, and M. J. Yaffe. Thickness-equalization processing for mammographic images. *Radiology*, 203(2):564–568, 1997.

[2] J. L. Ducote and S. Molloi. Scatter correction in digital mammography based on image deconvolution. *Physics in Medicine and Biology*, 55:1295–1309, 2010.

[3] W. He, E. R. E. Denton, K. Stafford, and R. Zwiggelaar. Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments. *Biomedical Signal Processing and Control*, 6(3):321–329, 2011.

[4] E. Frank I. H. Witten and M. A. Hall. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 3 edition, 2011.

[5] N. Karssemeijer and G. Te Brake. Combining single view features and asymmetry for detection of mass lesions. *Computational Imaging and Vision*, 13:95–102, 1998.

[6] American College of Radiology. *Breast Imaging Reporting and Data System BI-RADS*. Reston, VA: American College of Radiology, 4th edition, 2004.

[7] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13:146–165, 2004.

[8] P. R. Snoeren and N. Karssemeijer. Thickness correction of mammographic images by means of a global parameter model of the compressed breast. *IEEE Transactions on Medical Imaging*, 23(7):799–806, 2004.

[9] P. R. Snoeren and N. Karssemeijer. Thickness correction of mammographic images by anisotropic filtering and interpolation of dense tissue. *SPIE*, 5747:1521–1527, 2005.

[10] L. Tabár, T. Tot, and P. B. Dean. *Breast Cancer: The Art And Science Of Early Detection With Mamography: Perception, Interpretation, Histopatholigic Correlation*. Georg Thieme Verlag, 1 edition, December 2004.

[11] M. Tortajada, A. Oliver, R. Martí, M. Vilagran, S. Ganau, L. Tortajada, M. Sentís, and J. Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. *Lecture Notes in Computer Science*, 7361:561–568, 2012.