Modelling Breast Tissue in Mammograms for Mammographic Risk Assessment

Zhili Chen[1,3]
zzc09@aber.ac.uk

[1]Department of Computer Science
Aberystwyth University
Aberystwyth, UK

Erika Denton[2]
Erika.denton@nnuh.nhs.uk

[2]Department of Breast Imaging
Norfolk and Norwich University Hospital
Norwich, UK

Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[3]Faculty of Information and Control Engineering
Shenyang Jianzhu University,
Shenyang, China

Abstract

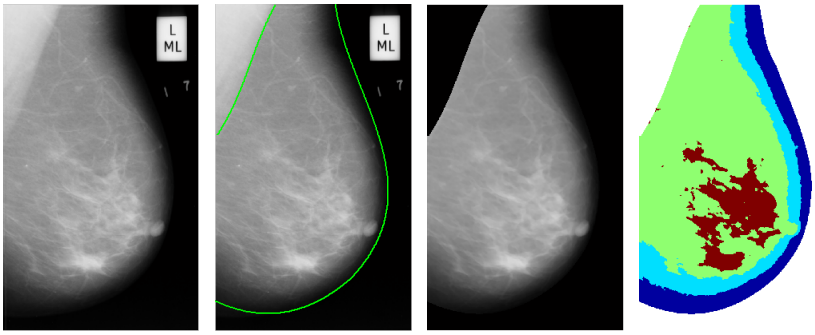
We propose to model breast tissue in mammograms covering both density and tissue patterns. The breast tissue density modelling is based on the global density distribution of the breast region. We segment the whole breast region into a number of uniform density sub-regions and construct an overall density model of the breast using the relative proportions of these sub-regions. The breast tissue pattern modelling is based on the local texture appearance, for which we use a texton based approach. The breast tissue models generated in this way can be used for mammographic risk assessment. The evaluation is based on the MIAS database. The classification results show a high agreement with the consensus of three experts according to four BIRADS categories.

1 Introduction

Mammographic density and parenchymal patterns are both strong predictive indicators of breast cancer risk, which corresponds to greylevel and texture information, respectively. A variety of methods have been developed for breast tissue characterisation [1, 3, 5, 8, 9]. A number of publications are related to breast tissue segmentation, where the breast tissue region is segmented into a number of sub-regions corresponding to different densities or appearances based on greylevel and texture information [3, 5]. The segmentation results can be used for mammographic risk assessment. On the other hand, texture representation of breast tissue has been investigated and has played a significant role in mammographic risk assessment [8, 9]. Recently, the adoption of local information to represent texture images has become a trend in texture classification. Numerous approaches to texture classification based on local texture information have been developed [4, 7, 11]. A number of publications have applied local texture information to mammographic risk assessment [1, 9].

We model breast tissue in mammograms incorporating global greylevel information and local texture information. The modelling process consists of three steps: (1) breast tissue extraction; (2) breast tissue density modelling; and (3) breast tissue pattern modelling.

© 2011. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.



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

Figure 1: (a) Original mammogram; (b) breast region segmentation; (c) breast tissue filtering; and (d) breast tissue segmentation: the breast tissue density corresponding to the four colour sub-regions increases in the order of blue, light blue, green, and red.

2 Methodology

As a preprocessing step we segment the breast region [2]. An original Medio-Lateral Oblique (MLO) mammogram and the segmentation result are shown in Figure 1(a) and 1(b), respectively. Following breast region segmentation, the breast region is filtered with an adaptive local window filter defined in [3] in order to eliminate the sensitivity to noise and small intensity inhomogeneity effects. The filtered result is shown in Figure 1(c).

To model breast tissue, a set of training images are first needed to learn the greylevel and texture information for a specific mammographic database. A greylevel histogram is generated over all the training images, which is used to describe the overall distribution of the breast tissue density. The modified Fuzzy C-Means (MFCM) algorithm [3] is used to segment the breast region into N_d density sub-regions. The segmentation result ($N_d = 4$) of Figure 1(c) is shown in Figure 1(d). The relative areas of the density sub-regions are used to represent the breast tissue density composition.

Breast tissue patterns are modelled using a texton based approach [9, 11] (as shown in Figure 2). Firstly, the original image patches (for which neither filtering nor normalisation is performed) are extracted at each breast tissue pixel from all the training images. The image patches we use are $N \times N$ square local windows. The patches entirely located within the breast area are aggregated to form a local tissue appearance space. We rearrange the pixels within the patches in row order to form N^2 dimensional vectors. The K-Means clustering algorithm is used to partition these vetors into N_t clusters. The N_t cluster centres are considered as the breast tissue textons. Subsequently, the breast tissue patterns are modelled with the frequency histogram of the N_t textons. Specifically, for a given mammographic image, image patches are first extracted, each image patch is then labelled with the closest breast tissue texton. The occurrence of each texton labelling is calculated to generate the breast tissue texton histogram, which is L_1 normalised to avoid the bias caused by the breast size.

To generate a breast tissue model incorporating both breast tissue density and patterns, we combine the relative proportions of the density sub-regions and the texton histogram to form an n ($n = N_d + N_t$) dimensional feature space. Consequently, breast tissue is represented based on the joint distribution of breast tissue density and local breast tissue appearance.

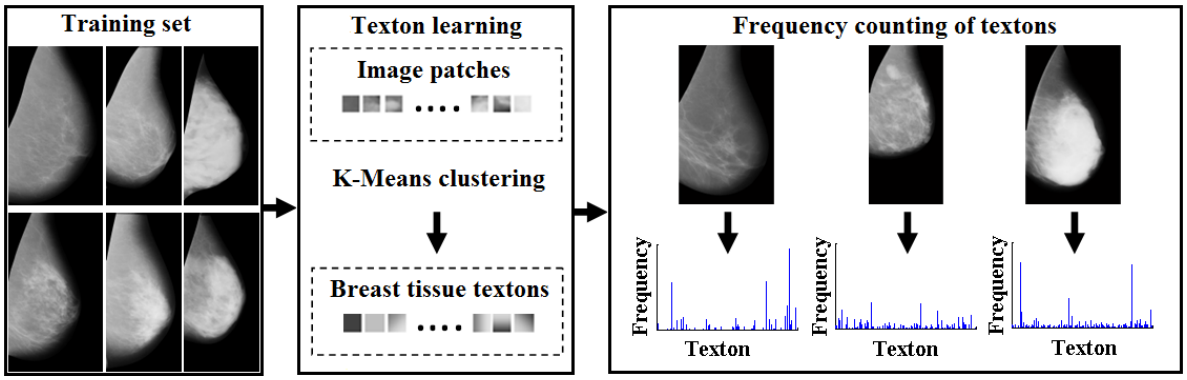


Figure 2: Breast tissue pattern modelling: the three texton histograms correspond to the top mammograms with the breast density increasing from left to right.

3 Results and Discussion

To test the capability of the developed breast tissue models for mammographic risk assessment, the Mammographic Images Analysis Society (MIAS) database [10] has been used. This database contains 322 MLO mammograms from 161 women. Three experts classified 321 available mammograms (mdb295ll was excluded for historical reasons) according to the Breast Imaging Reporting and Data System (BIRADS) standard [6]. The consensus was considered as the classification ground truth.

We used 5 different training sets, for each of which we randomly selected 40 mammograms from the MIAS database (we did not make an implicit assumption that mammograms in each class have the same appearance). The breast region was segmented into 8 density sub-regions. The original spatial resolution of the images is $50\mu m \times 50\mu m$ per pixel. For memory and efficiency reasons, we downsampled the full resolution to $800\mu m \times 800\mu m$ per pixel at the stage of breast tissue texton learning. The image patch size was 3×3. 160 textons were learned from the training set. The number of features ($N_d + N_t = 8 + 160$) in the breast tissue model was 168 in total.

A leave-one-woman-out evaluation methodology was used for the classification. When classifying one mammogram, the other mammogram from the same woman was excluded from the training samples to avoid bias (left and right mammograms from the same woman might have similar tissue features). We used a k-Nearest Neighbours (kNN) based classifier. A value of $k = 8$ was selected, but small variations in k produced similar results. The similarity between two models was measured using the χ^2 distribution comparison, where $\chi^2(x,y) = 0.5\sum_{i=1}^{n}(x_i - y_i)^2/(x_i + y_i)$. To prevent the similarity measurement being dominated by the features scaled in a range of larger values, all features in the models were normalised between 0 and 1 (other normalisation methods have been tried giving similar results). The performance was improved after normalisation. Initially, a classic kNN was used, but when equal classes were indicated, we used a weighted kNN approach. In addition, the traditional kNN classifier weights all the features equally, taking no account of their discriminating capability. To solve this problem, we used the sequential forward selection (SFS) algorithm to select a set of discriminative features. For different training sets, on average 40 features were retained. Among these features, a subset of top-ranking features played an important role in the classification (which contained most of the density ralated features), while the remaining features slightly increased the performance.

Table 1 shows the classification results based on 321 mammograms in the MIAS database

		Automatic				
	BIRADS	I	II	III	IV	CA
Truth	I	78	9	0	0	90%
	II	7	87	9	0	84%
	III	0	16	73	5	78%
	IV	1	1	8	27	73%

Table 1: Confusion matrix of breast density classification, with overall classification accuracy (CA) equal to 83%.

according to the four BIRADS classes. The classification accuracy (CA) for the four-class classification is 83%, and increases to 92% when considering the two-class (i.e. low and high density) classification. When analysing each BIRADS class, the correct classification percentages are 90% for BIRADS I, 84% for BIRADS II, 78% for BIRADS III, and 73% for BIRADS IV, respectively. The relatively poor performance for BIRADS IV might be explained by the lower number of samples for that class. Moreover, it should be noted that only two mammograms were mis-classified by more than one BIRADS class. Similar classfication results were obtained when using different training sets. The average results were $81 \pm 2\%$ for four classes and $90 \pm 2\%$ for two classes.

We have compared the results with some closely related work which also used the BIRADS criteria for the classification. Petroudi *et al*. [9] modelled mammographic parenchymal patterns with a statistical distribution of clustered filter responses of the Maximum Response 8 (MR8) filter bank, and obtained a classification accuracy of 76% for the Oxford Database. Oliver *et al*. [8] used a set of morphological and texture features. They obtained a correct percentage of 77% using the combination of the kNN classifier and the SFS algorithm, which increased to 86% when the Bayesian combination of the kNN classifier and the C4.5 decision tree was used. He *et al*. [5] developed a number of mammographic image segmentation methods for mammographic risk assessment, and the classification accuracy they recently obtained was 75%. A number of these publications (e.g. [5, 8]) used the same database and the same ground truth, so a direct comparison is possible.

At the breast tissue density modelling stage, the breast tissue density was described by the relative areas of the segmented density sub-regions to the whole breast area. A range of other properties of the dense tissue sub-regions related to shape measurement (e.g. Euler number and solidity) can be investigated to link the BIRADS criteria and morphological information of dense tissue. At the breast tissue pattern modelling stage, we used reduced-resolution mammographic images, where useful small texture information might be lost. The multi-resolution effects on the modelling will be investigated. In addition, there are two important parameters in the developed breast tissue models: the number of segmented density sub-regions, and the number of breast tissue textons, respectively. They were set to 8 and 160 in our experiments. We have experimented with different parameter settings. From a thorough evalution, the selected parameters performed better, and small variations in this setting provided similar results. We used a leave-one-woman-out methodology in our experiments, where the distribution of training samples in the four classes was unbalanced. BIRADS IV was not well represented due to a small proportion in the training set, which might explain the lower performance for this class. The unbalanced training sample problem will be further investigated to optimise the classification.

4 Conclusions

We have presented a breast tissue modelling method for mammographic risk assessment, which combined density and tissue patterns based on global greylevel and local texture information. The breast tissue density based segmentation showed realistic results corresponding to different tissue densities. The breast tissue patterns were represented by frequency histograms of breast tissue textons. The developed breast tissue models have been used to discriminate mammographic images in the MIAS database according to the BIRADS standard. A high agreement has been achieved with the consensus classification from three experts. The classification results are comparable with the state of the art. This work provides an overall framework for modelling breast tissue in mammographic images. As future work, we will incorporate different types of breast tissue features to generate more integrated breast tissue models. Different methods for breast tissue representation will be investigated on the basis of this framework. In addition, feature selection and multiple classifier combination will be used to improve the classification.

References

[1] A. Bosch, X. Muñoz, A. Oliver, and J. Martí. Modeling and classifying breast tissue density in mammograms. In *Proc. CVPR*, pages 1552–1558, 2006.

[2] Z. Chen and R. Zwiggelaar. Segmentation of the breast region with pectoral muscle removal in mammograms. In *Proc. MIUA*, 2010.

[3] Z. Chen and R. Zwiggelaar. A modified fuzzy c-means algorithm for breast tissue density segmentation in mammograms. In *Proc. ITAB*, 2010.

[4] M. Crosier and L. D. Griffin. Using basic image features for texture classification. *IJCV*, 88(3):447–460, 2010.

[5] W. He, E. R. E. Denton, and R. Zwiggelaar. Mammographic image segmentation and risk classification using a novel texture signature based methodology. In *LNCS*, volume 6136, pages 526–533, 2010.

[6] American College of Radiology. *Illustrated Breast Imaging Reporting and Data System BIRADS*. Amer. College of Radiol, 1998.

[7] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution grey-scale and rotation invariant texture classification with local binary patterns. *PAMI*, 24(7):971–987, 2002.

[8] A. Oliver, J. Freixenet, R. Martí, J. Pont, E. Perez, E.R.E. Denton, and R. Zwiggelaar. A novel breast tissue density classification methodology. *TITB*, 12(1):55–65, 2008.

[9] S. Petroudi, T. Kadir, and M. Brady. Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Proc. EMBC*, pages 798–801, 2003.

[10] J. Suckling, *et al*. The mammographic image analysis society digital mammogram database. In *International Congress Series 1069*, pages 375–378, 1994.

[11] M. Varma and A. Zisserman. A statistical approach to material classification using image patch exemplars. *PAMI*, 31(11):2032–2047, 2009.