33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Breast Density Characterization using Texton Distributions

Styliani Petroudi and Michael Brady

Abstract— **Breast density has been shown to be one of the most significant risks for developing breast cancer, with women with dense breasts at four to six times higher risk. The Breast Imaging Reporting and Data System (BI-RADS) has a four class classification scheme that describes the different breast densities. However, there is great inter and intra observer variability among clinicians in reporting a mammogram's density class. This work presents a novel texture classification method and its application for the development of a completely automated breast density classification system. The new method represents the mammogram using textons, which can be thought of as the building blocks of texture under the operational definition of Leung and Malik as clustered filter responses. The new proposed method characterizes the mammographic appearance of the different density patterns by evaluating the texton spatial dependence matrix (TDSM) in the breast region's corresponding texton map. The TSDM is a texture model that captures both statistical and structural texture characteristics. The normalized TSDM matrices are evaluated for mammograms from the different density classes and corresponding texture models are established. Classification is achieved using a chi-square distance measure. The fully automated TSDM breast density classification method is quantitatively evaluated on mammograms from all density classes from the Oxford Mammogram Database. The incorporation of texton spatial dependencies allows for classification accuracy reaching over 82%. The breast density classification accuracy is better using texton TSDM compared to simple texton histograms.**

I. INTRODUCTION

Breast cancer will affect between 1 to 8 women during their lifetime but the earlier the diagnosis the better the prognosis for the disease. Mammographic density which refers to the prevalence of fibroglandular tissue as it appears on a mammograms has been shown to be one of the most important risks for developing breast cancer and this has been confirmed in a number of studies [1], [2], [3]. Additionally, breast density may lower the sensitivity of mammography and obscure lesions. Thus, breast density and change thereof may be used for risk assessment, for reducing screening intervals, for the development of Computer Aided Detection (CAD) systems with higher sensitivity and specificity, but most importantly for signaling the necessity for a more thorough interpretation of certain mammograms for achieving the earliest possible diagnosis.

The American College of Radiology (ACR) proposes the following breast density classification in the Breast Imaging and Reporting Data System [4]:

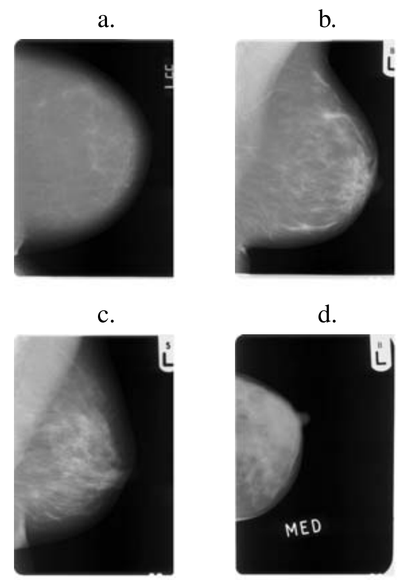

Fig. 1. Examples of mammograms from the 4 BI-RADS categories: a) BI-RADS I, b) BI-RADS II, c)BI-RADS III, d) BI-RADS IV.

- (i) the breast is almost entirely fat,
- (ii) there are scattered fibroglandular densities,
- (iii) the breast is heterogeneously dense which may lower the sensitivity of mammography, and
- (iv) the breast tissue is extremely dense, which could obscure a lesion in mammography.

Mammograms corresponding to the four BI-RADS classes can be seen in Figure 1. BI-RADS density classes *I* and *II* correspond to low density mammograms that translates to low breast cancer risk cases, whilst BI-RADS *III* and *IV* density classes correspond to high density mammograms and thus high risk cases. Recently, due to how significant a risk mammographic density is, clinicians are required to report to the woman her breast density BI-RADS classification, so that she can make more informed decisions regarding her health. Automated breast density classification methods can also be incorporated in CAD systems to achieve higher sensitivity and specificity. A number of studies have shown that CAD's sensitivity is low for mammograms with density classification BI-RADS III and IV, and almost half of interval cancers occur in mammograms classified as BI-RADS IV [5], [6]. Reporting of breast density suffers from high inter and intra observer variability [7]. Automated breast density classification algorithms can overcome this difficulty, aid the clinicians and provide objective classification of breast density.

This work was not supported by any organization
Styliani Petroudi is with the Department of Computer Science at the University of Cyprus, P.O. Box 20537 1678 Nicosia, Cyprus. styliani@ucy.ac.cy
Sir Michael Brady is with the Department of Department of Radiation Oncology and Biology, University of Oxford, Parks Road, Oxford, UK. jmb@robots.ox.ac.uk

A number of techniques have been proposed for breast density pattern classification. Boyd et al. [2] proposed a semi-automatic computer measure based on interactive thresholding and the percentage of the segmented dense tissue over the segmented breast area. Miller and Astley [8] investigated texture-based discrimination between fatty and dense breast types applying granulometric techniques and Laws texture masks. Other methods include automatic segmentation based on variance histogram discriminant analysis classification [9], and density classification using a large set of statistical and compositional features in term of BI-RADS [10]. Musta et al. [11] present an overview of the accuracy of different breast density classification methods using different features, and achieve a maximum classification accuracy of 73.3% through a selection of Haralick and Soh texture features, genetic search and wrappers. Petroudi et al. [12] proposed a scheme that uses texture models to capture the mammographic appearance within the breast area: parenchymal density patterns are modeled as a statistical distribution of clustered, rotationally invariant filter responses in a low dimensional space.

The purpose of this paper is the development of a fully automatic, and highly accurate mammographic breast density classifier based on objective and quantitative texture measures. The presented algorithm builds on the same definition of textons as clustered filter responses [13]. However, instead of using histograms for modeling the different texture classes, a texture descriptor that evaluates the spatial dependence between the textons characterizing the image is introduced and used. The descriptor follows the definition of intensity spatial dependency matrices as introduced by Haralick et al. [14], but is evaluated on texton maps.

II. METHOD

The gray-level spatial dependence or co-occurrence matrix (GLCM) measures the frequency of intensity pairs in the gray-level image of neighboring pixels at different distances and directions [14]. Haralick et al. [14] evaluated second order statistics on the corresponding GLCM for further texture description. However, the simple intensity information does not provide adequate information especially for analysis and characterization of many medical images. Thus, following the definition of GLCM presented in [14], a new texture descriptor that captures both structural and statistical texture information is defined, the texton spatial dependence matrix (TSDM), or texton co-occurrence matrix. The term texton co-occurrence matrix was first used by Liu *et al.* in [15]. However, they define textons as different shape descriptors. They define a $2x2$ grid, and if three or four of the corresponding pixel values are the same, then those pixels are set to form a texton. If a pixel belongs to a texton the pixel will keep the intensity value of the image where the five texton shapes are evaluated on. The resulting image is what Liu et al. [15] call a texton map. Thus in [15] the corresponding texton map for the intensity image is the same image with the same intensity values, except where the pixels do not match a texton shape and are set to zero. Li and Shi follow a similar approach using local binary patterns [16]. For the TSDM texture descriptor presented here, textons are defined under the operational definition of Leung and Malik [13], resulting in a very different texton map - where each texton corresponds to a vector and not to a pixel intensity value, or a gradient thereof, as in [15].

The TSDM texture descriptor captures both structural and statistical texture properties. Textons, as proposed by Julesz [17], are the primitives of texture. Structural models of texture are based on the view that texture are composed of primitives in spatial arrangements. The presented method provides structural characteristics of texture in the sense that the primitives are explicitly defined [18]. Additionally, TSDM follows the definition of Haralick [14], and can be explained as the matrix containing the frequencies or the probabilities of the textons co-occurrences, and as such provides statistical information regarding the texture.

Let the image to be analyzed defined as I and let $L_x = \{1,2,...,N_x\}$ and $L_y = \{1,2,...,N_y\}$ the spatial domains in X and Y with N_x number of columns and N_y the number of rows. Let TI be the texton map matrix where each entry identifies the texton, $T \in \{t_1,t_2,...,t_n\}$, each pixel is mapped to. There are n textons in the corresponding texton dictionary. TI can be defined as a function that assigns some texton T to each pixel: $TI : L_x \times L_y \to T$.

Again as in [14], the developed texture measures are angular texton nearest-neighbor spatial dependence matrices (TSDM) specified by the matrix of relative frequencies P_{t_i,t_j} with which two neighboring pixels mapped to textons t_i and t_j separated by distance d occur on the image's texton map TI.

The TSDM for displacement $d = (d_x,d_y)$ can be represented by:

$$TSDM(t_i,t_j,d_x,d_y) = \frac{1}{\sharp} \sum_{k=1}^{N_x} \sum_{l=1}^{N_y} \begin{array}{ll} 1 & if \;\; TI(k,l) = t_i \;\; and \;\; TI(k+d_x,l+d_y) = t_j \\ 0 & otherwise \end{array} \tag{1}$$

where $\sharp$ defines the total number of elements in the corresponding set. By incorporating a displacement vector in the horizontal and vertical direction the angular relationship between neighboring pixels is inherently incorporated.

For the development of the new TSDM based density classification model the following steps need to take place. Initially the texton dictionary must be derived. Following, segmentation of the breast region [19] the resulting images from the training set are filtered using the Maximum Response 8 (MR8) filter bank proposed by Varma and Zisserman [20]. After filtering using the filter bank, each pixel is associated with a vector that holds the filter response corresponding to each filter in the filter bank. The filter responses over all the pixels in the images' regions of interest are aggregated. The texton dictionary is created by clustering these aggregated filter responses over all images per BI-RADS class using the K-Means algorithm.

TABLE I

CLASSIFICATION ACCURACY RESULTS

Accuracy%	BI-RADS I	BI-RADS II	BI-RADS III	BI-RADS IV
4 Density Classes	86%	93%	80%	66%

Given the texton dictionary, each image pixel in the breast region of each mammogram in the training set is mapped to the texton closest to it in the filter response space. This step provides the TI image's texton map. TI is then used to evaluate the TSDM for different displacements as shown in equation (1). TSDM matrices for different displacements are computed for each training mammograms. The sets of the TSDMs define the breast parenchymal density models.

To achieve classification, the same steps as above are followed for a test mammogram - segmentation, filtering, evaluation of the corresponding TI and TDSMs. The resulting TSDMs are compared to the TSDMs of all learnt models and the mammogram is assigned to the BI-RADS class closest to it using χ^2 significance test in conjunction with a nearest neighbor rule. For the developed method 10 cluster centers per BI-RADS class are used for the creation of the texton dictionary, and for this paper only the TSDM with $d = 1$ and $d = 2$ are investigated.

III. RESULTS

The algorithm is evaluated on a set of 100 mammogram cases from the Oxford Database [12], 25 from each BI-RADS class for which there was independent agreement in density classification by three expert breast radiologists. 10 mammograms from each class were used for training and 15 for testing. The images correspond to 8-bit mammograms downsampled to 300μm/pixel. Despite the small size of the training set, exact agreement with the ground truth was achieved in 81% of the cases.

Table I shows the classification accuracy of the presented technique discriminating between the 4 BI-RADS categories based on the ground truth. Accuracy is calculated as the percentage of correctly classified mammograms in a breast parenchymal density category over the ground truth total number of mammograms in that category.

The classification algorithm achieves better accuracy if only two classes are used for breast density characterization. If the mammograms with breast density classification BI-RADS I and BI-RADS II are combined in one class defined as the low risk class whilst the mammograms with breast density classification BI-RADS III and BI-RADS IV are combined in another class defined as the high density class the lowest achieved classification is 90%. The results are presented in Table II. This result suggests, that breast parenchymal density and its distribution result in distinctly different texture characteristics in the lower and higher density classes. The algorithm provides the worst classification for BI-RADS IV mammograms and this may be attributed to the fact that mammograms from this category tend to have the parenchymal density more uniformly distributed in the breast region.

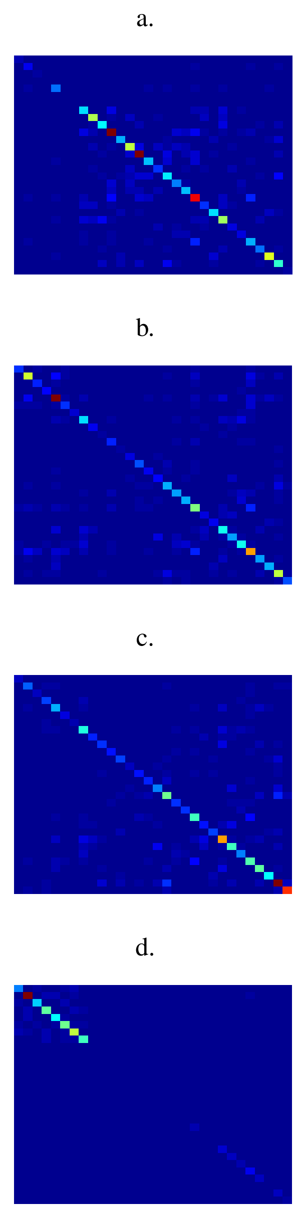

Fig. 2. Examples of the TSDM matrices with distance 1 for the mammograms from the 4 BI-RADS categories shown in Figure 1: a) BI-RADS I, b) BI-RADS II, c)BI-RADS III, d) BI-RADS IV.

IV. DISCUSSION

This paper proposes a new effective texture descriptor that captures both structural and statistical properties. The paper introduces TSDM, which evaluates the relative frequencies with which neighboring pixels are mapped to textons in the texton dictionary. Evaluation of different distance TSDM matrices and calculation of different texture features from the corresponding matrices will allow for even better texture characterization and improved performance and consistency.

TABLE II

CLASSIFICATION ACCURACY RESULTS FOR A TWO CLASS CHARACTERIZATION OF BREAST DENSITY

Accuracy%	BI-RADS I and II	BI-RADS III and IV
Two Risk Classes	93.3%	90%

The images in the BI-RADS I and BI-RADS IV appear as much more homogeneous than do the other two classes, thus evaluation of homogeneity should allow for better characterization of the two classes. Moreover, BI-RADS II and and BI-RADS III have textures at different scales which may be a reason why the algorithm can achieve good classification accuracy for mammograms from these two categories.

The texture descriptor is incorporated in a method for breast density classification. The results are very good and the TSDMs for the different density classes show good separation between them. Figure 2 shows the TSDM matrices for the four mammograms corresponding to the four BI-RADS classes in figure 1.

The algorithm achieves good classification accuracy compared to other methods in the literature. Using simple texton histograms achieves a classification accuracy of about 76%. Oliver et al. [21] also achieve comparable classification accuracy of 76% by using a combination of morphological and texture features from segmented breast and fatty density regions. However, the presented algorithm needs to be evaluated on a much larger dataset

V. CONCLUSION

A breast density classification approach is presented based on the development of a new texture model TSDM that captures both structural and statistical texture information through the use of textons which are texture primitives and texton co-occurrence which provides frequency information regarding the distribution of textons at certain distances. The method builds on the definition of textons in [13] and spatial and angular dependence matrices in [14]. The presented method defines texture classes as TSDM matrices over "texton" dictionaries developed from a training set. Classification is simply a matter of comparing the texton spatial dependence relative frequency matrices using an appropriate distance measure. The results compare favorably to other methods in the literature [11]. However, further evaluation using different filter banks for the texton dictionary and combination of the TSDMs evaluated at different distances can result in significant improvement.

In the future, TSDMs corresponding to different distances will be evaluated on a larger dataset. Evaluation of texture features on TSDMs will also be investigated.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of Dr. Ruth English, Dr. Rosie Adams and Dr. Basil Shepstone for providing the breast density classifications.

REFERENCES

[1] K. Kerlikowske, "The mammogram that cried wolfe," *New England Journal of Medicine*, vol. 356, pp. 297–300, 2007.

[2] N. Boyd, J. Byng, R. Jong, E. Fishell, L. Little, A. Miller, G. Lockwood, D. Tritchler, and M. Yaffe, "Quantitative classification of mammographic densities and breast cancer risk: Results from the canadian national breast screening study," *Journal of the National Cancer Institute*, vol. 87, no. 9, pp. 670–675, 1995.

[3] J. Byng, N. Boyd, E. Fishell, R. Jong, and M. Yaffe, "Automated analysis of mammographic densities," *Physics in Medicine and Biology*, vol. 41, pp. 909–923, 1996.

[4] *Illustrated Breast Imaging Reporting and Data System (BI-RADS)*, 3rd ed. American College of Radiology, 1998.

[5] C. Romero Castellano, C. Varela Nuez, R. Cuena Boy, A. Almenar Gil, J. Pinto Varela, and M. Botella Lopez, "Impact of mammographic breast density on computer-assisted detection in a breast imaging department," *Radiologia*, 2011.

[6] L. Li, Z. Wu, L. Chen, F. George, Z. Chen, A. Salem, M. Kallergi, and C. Berman, "Breast tissue density and cad cancer detection in digital mammography," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, jan. 2005, pp. 3253 –3256.

[7] B. Yankaskas, M. Schell, R. Bird, and D. Desrochers, "Reassessment of breast cancers missed during routine screening mammography: a community based study," *American Journal of Roentgenology*, vol. 177, pp. 535–541, 2001.

[8] P. Miller and S. Astley, "Classification of breast tissue by texture and analysis," *Image and Vision Computing*, vol. 10, pp. 277–282, 1992.

[9] T. Matsubara, D. Yamazaki, M. Kato, T. Hara, H. Fujita, T. Iwase, and T. Endo, "An automated classification scheme for mammograms based on amount and distribution of fibroglandular breast tissue density," *Excepta Medica International Congress Series*, vol. 1230, pp. 545–552, 2001.

[10] Y. Chang, X. Wang, L. Hardesty, T. Chang, W. Poller, W. Good, and D. Gur, "Computerized assesment of tissue composition of digitized mammograms," *Academic Radiology*, vol. 9, pp. 899–905, 2002.

[11] M. Mustra, M. Grgic, and K. Delac, "Feature selection for automatic breast density classification," in *ELMAR, 2010 PROCEEDINGS*, 2010, pp. 9 –16.

[12] S. Petroudi, T. Kadir, and M. Brady, "Automatic classification of mammographic parenchymal patterns: A statistical approach," in *Proceedings of EMBC, International Conference on Engineering in Medicine and Biology*. IEEE, 2003, pp. 798–801.

[13] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials seing three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, June 2001.

[14] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 3, pp. 610 –621, 1973.

[15] G.-H. Liu and J.-Y. Yang, "Image retrieval based on the texton co-occurrence matrix," *Pattern Recognition*, vol. 41, no. 12, pp. 3521 – 3527, 2008.

[16] Q. Li and Z. Shi, "Texture image retrieval using compact texton co-occurrence matrix descriptor," in *Multimedia Information Retrieval*, 2010, pp. 83–90.

[17] B. Julesz, "Textons, the elements of texture, perception, and their interactions." *Nature*, vol. 290, no. 5802, pp. 91–97, 1981.

[18] R. M. Haralick, "Statistical and structural approaches to texture." *Proc IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[19] S. Petroudi and M. Brady, "An automated algorithm for breast background segmentation," in *Medical Imaging Understanding and Analysis*, 2003, pp. 181–184.

[20] M. Varma and A. Zisserman, "Classifying images of materials: Achieving viewpoint and illumination independence," in *Proceedings of the European Conference on Computer Vision, Copenhagen, Denmark*, 2002, pp. 255–271.

[21] A. Oliver, X. Llado, R. Marti, J. Freixenet, and R. Zwiggelaar, "Classifying mammograms using texture information," in *Medical Imaging Understanding and Analysis*, Wales, UK, 2007, pp. 223–227.