Classification of Mammogram Patterns using area measurements and the Standard Mammogram Form (SMF)

Styliani Petroudi[a*], Kostas Marias[a], Ruth English[b] and Michael Brady[a]

[a] Medical Vision Laboratory, Engineering Science, Oxford University, Parks Road, Oxford OX1 3PJ, [b] Breast Care Unit, Churchill Hospital, Oxford OX3 7LJ, UK

Abstract. Our aim is to propose a new approach to breast pattern classification that will aid in the development of an automated mammographic density analysis procedure. Breast patterns broadly classify the mammographic density and density distribution of each mammogram in order to provide a framework for assessing the risk of breast cancer according to density. However, basic weaknesses of classifying mammogram films are the inter and intra-subject variability in interpretation and coding of the patterns by different radiologists as well as the fact that the intensities of the mammogram do not indicate the actual volume of the breast regions. Our endeavour is the development of *objective descriptors of breast density* based on the digital mammogram. Unlike previous efforts to classify breast patterns based on image texture, we propose the use of density measures based on a normalised representation of the breast (h_{int} representation or SMF) and demonstrate that they can provide a more intuitive quantitative description of mammogram patterns.

1 Introduction

Breast cancer is one of the leading causes of death for women. Objective quantitative measures of breast density are crucial tools for assessing the association between the risk of breast cancer and mammographic density as well as for quantification of density changes to the breast. However, there is a need to validate any suggested measure of density and the most logical choice that can allow comparison between clinical descriptions and computer measures is the classification of *mammographic parenchymal patterns*. In 1976 Wolfe [1] suggested a way of classifying breast parenchymal tissue, in order to associate mammographic density and appearance, to the risk of breast cancer. They are briefly described in the next section.

Previous efforts to classify mammogram patterns are mostly based on image texture measures (e.g. in [2]). However, we denote two basic weaknesses of textural approaches:

1. It is intrinsically difficult to correlate any grey-level texture measures with the actual image properties and in extension with the consistent characteristics of each mammogram pattern class in an intuitive manner.
2. Because of the relatively weak control over the image acquisition process, it is difficult to eliminate variability in image characteristics, such as contrast and brightness. This means that differences in imaging conditions lead to a non-rigid variation in the mammographic intensity distribution a fact that diminishes the possible use of texture for mammographic pattern recognition.

For these reasons we base our measures on the normalised h_{int} representation of interesting tissue [3] which allows a standardised representation of mammograms (SMF) factoring out variations due to imaging conditions.

2 Definition of Breast patterns

Breast pattern are defined on the basis of the appearance of the 2D projection of the breast parenchyma under projection. The breast parenchyma is comprised by the lobules, ducts and the interlobular fibrous tissue and it's diffusely distributed within the fat tissue [4]. Ducts, lobular elements and fibrous connective tissue comprise the visible density that makes up the mammogram.

Mammographic patterns characterise the appearance of the breast density (glandular and stromal tissue) and assess the relative risk of breast cancer accordingly. In general, the higher the density the higher the risk of breast cancer and for this reason it is important to assess parenchymal density and patterns in a quantitative image-based manner. A measure that is objective and quantitative will therefore be valuable in recommending alternative breast screening paradigms and preventive measures [5].

*Author for correspondence: styliani@robots.ox.ac.uk

According to Wolfe [1], the four suggested patterns are:

N1: This category represents radiologically lucent mammograms. The breast consists mostly of fatty tissue with no ducts visible. This essentially represents an essentially normal breast (varies somewhat with the age of the patient) and it is considered to be a *lower risk* pattern

P1: This category represents a fatty breast with predominant ducts in anterior portion up to $1/4^{th}$ of the breast. The ducts are visible as cord-like structures in the sub-areolar area or as a thin band of duct-like structures extending into one of the quadrants (typically the UOQ). It is considered as the *low risk* pattern.

P2: Describes a breast involuted with prominent duct pattern of a moderate to severe degree, occupying more than $1/4^{th}$ of the breast volume. The visible duct pattern can occupy the entire breast. It is considered as the *high-risk* pattern.

DY: Diffuse or nodular densities. DY describes a dense parenchyma, which usually denotes connective tissue hyperplasia (or as described by Wolfe "dysplasia", referring to the epithelial and connective tissue densities). It is considered as the *highest risk* pattern and can appear homogenous due to the overall increased density. The prominent duct pattern cannot be seen.

Patterns P2 and DY can lower mammographic sensitivity and/or obscure a lesion on mammography.

3 Methods Used

As mentioned before the proposed quantitative measures are based on the h_{int} representation of interesting tissue [3]. In addition, we use the density area measures proposed by Boyd et al. [6]. Figure 1.a illustrates the "conventional" method for isolating the area of glandular tissue, called interactive thresholding. In this method, the user has to define a global threshold that includes as much "density", non-fatty tissue, as possible. The percentage of density in the whole mammogram is used to describe the density content of the breast. As a mammographic measure, the area of the projected glandular tissue has two basic flaws:

- It does not account for the thickness of the projected tissue since the resulting intensity is not directly proportional to the thickness and can differ significantly due to the imaging conditions.

- The interactive segmentation of the "dense tissue" is a subjective procedure and therefore the measure obtained is not consistent. In Wolfe's classification the inter-observer variation has ranged from 52% to 97% for exact agreement and the intra-observer variation has ranged from 69% to 87 % [7], but this can be minimised by classifying the mammograms by consensus opinion [4] as it is done in this paper.

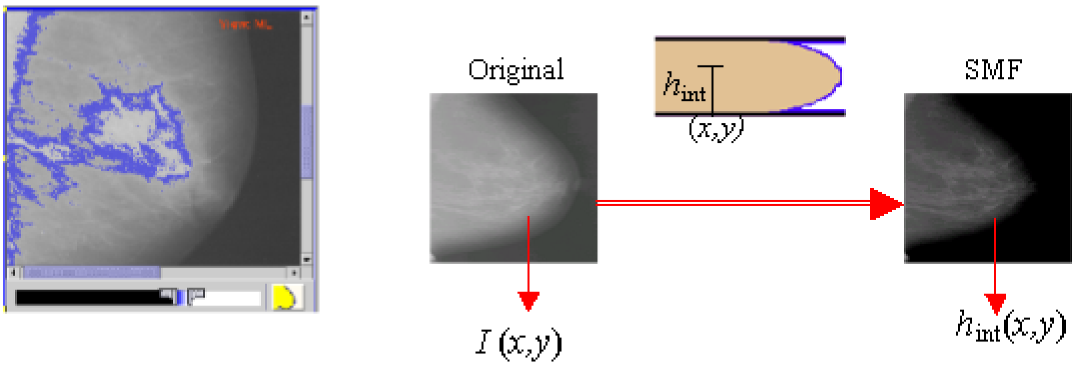


Figure 1. (a): Interactive thresholding is used to define the area of dense tissue in a mammogram, (b): The h_{int} representation allows the estimation of the height of non-fatty tissue under compression per pixel.

On the other hand the volume of "interesting" tissue is based on the normalised h_{int} representation of the breast, a physics-model based approach [3]. As is illustrated in Figure 1.b the Standard Mammogram Form (SMF) utilises image formation information (e.g. height of compression and exposure time) to calculate the height of non-fatty tissue under compression that corresponds to each pixel. The h_{int} can most usefully be regarded as a surface that conveys information about the anatomy of the breast. This enables us to have a truly quantitative density measure since the values at each entry (x,y) are in millimetres; that is they are metric.

The volume of "interesting" tissue is defined as:

$$V(h_{\text{int}}) = \frac{\sum_x \sum_y h_{\text{int}}(x, y)}{H \times A}$$

Where A is the area of the images (number of pixels and H is the height of the compressed breast in each mammogram. This h_{int} based measure is chosen because it is based on a normalised representation as well as because it is related with the actual parenchymal density of the breast (unlike the area measure that calculates the projected area of the parenchyma under compression). However, we have to note that possible errors in the calculation of H, can influence the accuracy of the V_{int} measure.

4 Results

To evaluate the suggested measures, 54 mammograms (with the background and pectoral muscle removed) were classified in 4 categories according to the Wolfe criteria (N_1, P_1, P_2, DY). Two engineers trained in understanding breast patterns and a radiologist performed the classification by consensus opinion in order to minimise possible differences in the interpretation [4]. The calculated measures for all the mammograms are shown in Figure 2.a (%Area) and Figure 2.b (V_{int}). Three thresholds are defined from each graph in order to evaluate the accuracy of each measure in classifying the mammogram data set. The agreed classification served as the ground truth for comparison and the results for both measures are presented in Table 1. For each zone defined by the three thresholds we calculate the (%) of correctly classified mammograms. Table 2 presents the accuracy of both measures in discriminating between low density (N_1+P_1) and high density (P_2+DY) classes.

	N_1 zone	**P_1 zone**	**P_2 zone**	**DY zone**
% Area	0-10%	10-17%	17-27%	>27%
Accuracy	63%	73%	75%	73%
V_{int}	0-16%	16-20%	20-25%	>25%
Accuracy	69%	80%	67%	70%

Table 1. Accuracy in classification for each suggested density measure for the four classes

	Low density (N_1 +P_1) zone	**High density (P_2+DY) zone**
% Area	<17%	>17%
Accuracy	85%	91%
V_{int}	>20%	>20%
Accuracy	93%	92%

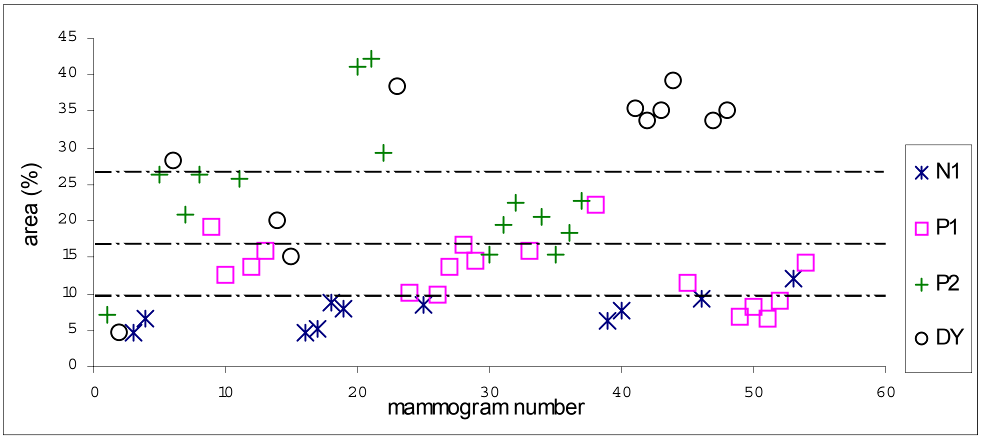
Table 2. Accuracy in classification in just low and high density classes for each suggested density measure

5 Discussion

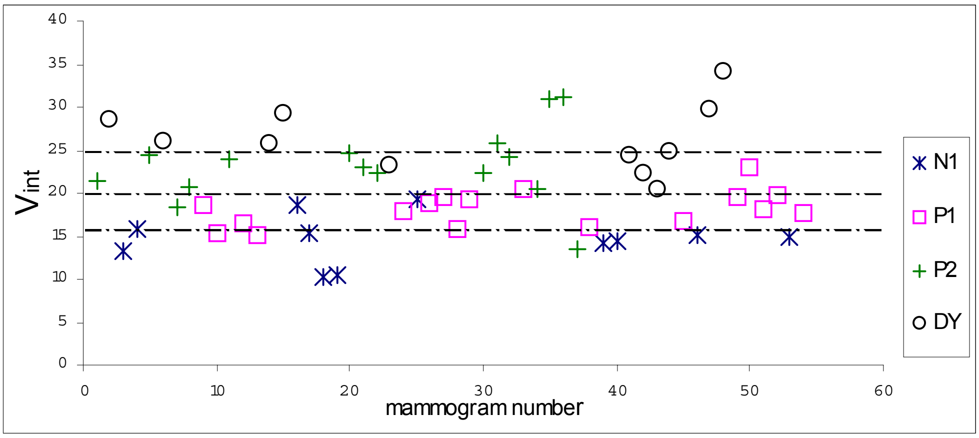
In this paper we presented results on mammogram pattern classification based on the normalised mammmogram representation (SMF). By using the SMF we overcome the difficulty in using intensity related parameters which do not indicate the actual volume of dense regions. The values of the SMF are independent of image parameters, e.g. energy, plate thickness, film characteristics, and relate to the physical volume of the breast. From the results presented in the previous section we can also conclude that:

1. The volume of interesting tissue performed better that the area (%) of dense tissue. This indicates that the projected area, although a good indicator of density, it's inssuficient to describe breast density on its own.
2. The (%) values of the volume of interesting tissue have a good anatomical correspondence with the expected tissue density of each mammogram class. For example around 20% V_{int} separates low density (N_1, P_1) from high density (P_2, DY) patterns while according to Wolfe [1] the separation level is 25% density. This clearly verifies that the h_{int} is a representation of the actual anatomy of the breast.

In addition the high accuracy in classification between low density (N_1+P_1) and high density (P_2+DY) classes indicates that purely density measures could help clearly separate lower risk from higher risk patterns automatically. This could provide important information for breast cancer screening. The results for the four class classification problem show that the classification of breast density according to Wolfe is challenging; due to the vast variation of the mammographic appearance of the breast. Our future work aims to combine density measures with texture analysis in normalised mammograms; which should also improve the four class classification problem, experimenting on a larger database of pre-characterised (according to pattern) mammograms as well as utilising pattern recognition to develop an automated mammographic density classifier.



(a)



(b)

Figure 2. (a): A plot of the calculated values of the (%) Area of dense tissue calculated for the 54 mammograms that had been previously classified into the 4 Wolfe patterns (N_1, P_1, P_2, DY). By introducing 3 thresholds (at 10%, 17%, and 27% we calculated the accuracy of the measure in classifying the mammograms in the 4 pattern classes (results presented in Table 1), (b): The same for the volume of interesting tissue (V_{int}).

Acknowledgements

The authors would like to acknowledge the Cancer Research Campaign (CRC) for supporting this work. In addition the authors thank Mirada Solutions (http://www.mirada-solutions.com/) for the SMF generation software.

References

1. J. N. Wolfe. "Breast pattern as an index for developing breast cancer", *American Journal of Roentgenology*, vol.126, pp 1130-1139, 1976.
2. P.Miller and S.Astley. "Classification of breast tissue by texture analysis", *Image and Vision Computing*, 10(5), pp 277-282, June 1992.
3. R. Highnam and M. Brady. *Mammographic Image Analysis*. Kluwer academic, 1999.
4. T. Salminen. *Causes and Consequences of Mammographic Parenchymal Patterns*. Univ. of Tampere, Finland 2000.
5. P. K. Saha, J. K., E. F. Conant, D. P. Chakreborty, and D. Sullivan. "Breast Tissue Density Quantification Via Digitized Mammograms", *IEEE Trans. Medical Imaging*, vol. 20, no. 8, pp 792-803, August 2001.
6. N.F. Boyd, G.A. Lockwood, J.W. Byng et al. "Mammographic densities and breast cancer risk", *Cancer Epidimiol Biomarkers*, vol. 7, pp 1133-1144, 1998.
7. A. M. Oza and N.F. Boyd, "Mammographic parenchymal patterns: a marker of breast cancer risk", *Epidemiol Rev*, vol. 15, pp196-208, 1993.