Classification of Breast Tissue by Texture Analysis

Peter Miller
Sue Astley

Wolfson Image Analysis Unit
Department of Medical Biophysics
University of Manchester
Stopford Building, Oxford Road
Manchester M13 9PT

Abstract

The identification of glandular tissue in breast X–rays (mammograms) is important both in assessing asymmetry between left and right breasts, and in estimating the radiation risk associated with mammographic screening. The appearance of glandular tissue in mammograms is highly variable, ranging from sparse streaks to dense blobs. Fatty regions are generally smooth and dark. Texture analysis provides a flexible approach to discriminating between glandular and fatty regions. We have performed a series of experiments investigating the use of granulometry and texture energy to classify breast tissue. Results of automatic classifications have been compared with a consensus annotation provided by two expert breast radiologists. On a set of 40 mammograms, a correct classification rate of 80% has been achieved using texture energy analysis.

1   Introduction

A national breast screening programme has recently been introduced with the aim of detecting breast cancer at an early stage in asymptomatic women. Women believed to be at risk from the disease, either because of age or family history, are subjected to regular X–ray examination of the breasts. The screening programme has resulted in a large increase in the number of mammograms requiring interpretation, and a variety of computer–based aids have been proposed to improve the performance of radiologists searching for small, subtle and infrequent abnormalities [1, 2].

One important mammographic sign of breast disease is asymmetry between the right and left breast images. Radiologists identify asymmetry by comparing the size, density and location of significant regions (including glandular tissue and any other localised opacities) in the two breasts [3]. Since there is often natural variability in the size and shape of the breast outline and regions within it, automated approaches to the detection of asymmetry which depend on direct bilateral subtraction are unlikely to be successful. A more promising approach is to detect the significant regions and use them as a basis for comparison.

The identification of fatty and glandular regions is also useful in calculating the radiation risk associated with breast screening, since it is believed that risk is related

to the X–ray dose to glandular tissue [4]. Estimates of the risk of mammographic screening for different classes of women can be obtained from data including age, compressed breast thickness, glandular composition and imaging parameters [5]. If reliable, such estimates could influence the frequency and type of mammographic examination used in screening programmes.

In mammograms, fatty regions are usually relatively smooth and dark in appearance, possibly containing bright streaks of fibrous tissue. The appearance of glandular regions is more variable, ranging from bright, fluffy blobs to sparse lines, depending on the amount and type of glandular tissue. Since mammograms are two–dimensional projections of the three–dimensional structure of the breast, overlapping structures cause a wide variation in glandular intensity. It cannot be assumed that all regions of glandular tissue will be significantly brighter than the surrounding fat, so it is appropriate to use texture, rather than intensity, as a primary classification method for breast tissue.

Little has been published on the computerised analysis of glandular patterns, other than four attempts to classify glandular tissue into the categories proposed by Wolfe [6] which are believed to be related to the natural risk of developing breast cancer. Kimme–Smith [7] described an *ad hoc* collection of statistical texture features, and claimed reasonable classification of a very limited number of images. Shadagopan [8] counted the number of ducts within the breast using a shape template, and attempted to associate the results with natural risk. Magnin [9] used conventional statistical texture features, but could not discriminate between Wolfe grades. Caldwell [10] measured the fractal dimension of the patterns, producing more accurate classifications. None of these authors attempted to automatically identify regions of glandular tissue, or to relate quantitative breast composition to mammography–induced risk.

Of those who have explored asymmetry as a cue for automated cancer detection, Hand [11] and Giger [12] employed direct subtraction of geometrically, rather than structurally, corresponding regions of the left and right breasts, and Kimme [13] measured unilateral differences in texture. A wide range of texture methods is available [14, 15]. We have investigated two approaches; the granulometric techniques proposed by Serra [16], which provide a flexible approach to the analysis of mammographic textures, and Laws' texture energy method [17], which has frequently been used for image segmentation.

2 Granulometry

Matheron first proposed a morphological sieving operation to filter particles in binary images according to their sizes [18]. Texture information is given by the rate at which the image is altered by a sieving process over a range of sizes. Serra described granulometric analysis using a variety of morphological operations [16].
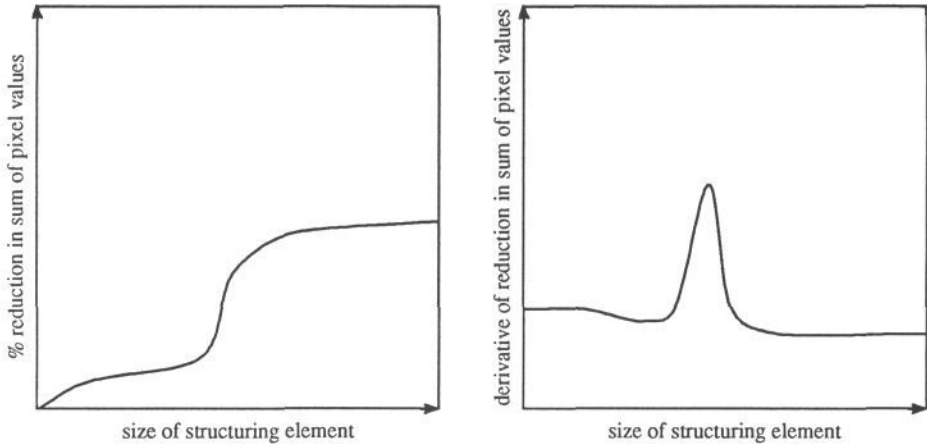
Figure 1. Granulometry graphs for opening

The morphological opening operation (erosion followed by dilation) is the most commonly used for granulometry. In the binary case, the effect of opening is to remove any objects or parts of objects which cannot completely enclose the structuring element (usually a small circle). Any change in the image after opening can be quantified as a reduction in the number of pixels. Grey level opening has a similar effect, but keeps brighter pixels where the structuring element can fit inside a bright object, and darker pixels if not; hence the corresponding change is in the sum of *pixel values* in the image.

Texture information is provided in the form of size distributions (figure 1). The first graph shows the effect on an image of a sequence of openings using structuring elements of increasing size. It illustrates a typical distribution for an image in which most features are of similar size; there is a sudden reduction in the sum of pixel values at the point where the structuring element can no longer fit inside those features. The second graph shows this effect as a peak in the *change* in sum of pixel values. A useful granulometric texture feature is the size of structuring element producing the greatest change in pixel value sum (the highest peak in the derivative graph). This corresponds to the most prominent size of structure in the image; it gives a low response for smooth images and a higher response for images containing mainly thick lines or blobs.

3   Texture Energy

A successful methodology for image segmentation using texture analysis has been designed by Laws [17]; it comprises a sequence of relatively simple operations (figure 2). Firstly, the image is convolved with a number of small masks, each of which extracts quantitative information about a different textural feature, such as edges, lines, spots or ripples. The variance for each feature image is then computed over a moving window, measuring the local texture energy. This step is also performed on a
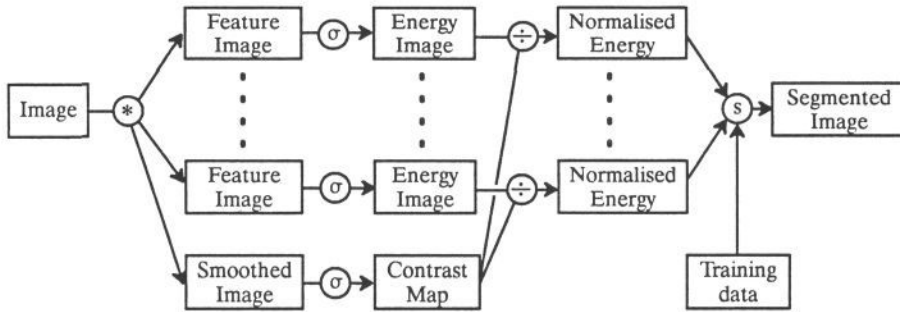
Figure 2. Schematic diagram of Laws' texture energy method

smoothed version of the original image, to produce a map of local contrast values. Each feature is normalised so that it is invariant to the effects of variable illumination and contrast; this is achieved by dividing the texture energy values by the corresponding contrast values. These normalised feature images are used to segment the original image on the basis of textural similarity.

Segmentation may either be unsupervised, using a multivariate clustering algorithm to determine the number and type of regions present, or supervised, in which case intra–class texture distributions are made available to a suitable classifier.

4 Experimental Approach

Experiments were performed using a Sun 3 workstation and a commercially available computer vision system (AVS IPB–3000), with software written in Pascal. A consultant radiologist selected a representative data set of 40 normal screening mammograms, which were digitised with a Pulnix CCD video camera, using a spatial resolution of 2.5 pixels/mm (512 x 512 pixels), and a grey resolution of 6 bits/pixel.

In order to establish a standard against which the automatic methods could be tested, two consultant radiologists drew round the glandular regions on transparent overlays. A consensus annotation was produced by combining the individual annotations; regions in which the two opinions differed were declared to be undefined, and excluded from subsequent analysis. The accuracy of computer–based methods was calculated as the percentage of correctly classified pixels within the breast area, as defined by an automated breast–finding algorithm [19].

Intensity thresholding was also applied to the digitised mammograms, for comparison with the texture methods. However, thresholding is known to be unreliable, since the image intensities for fat and glandular tissue depend on tissue thickness, film processing and illumination. Better results from thresholding could be obtained by normalising the images prior to analysis, but this is impractical, as it requires the identification of sample patches of known tissue type and thickness in each image.

4.1 Granulometric Analysis

Granulometry parameters were selected by examining the results of a series of experiments performed on a limited data set of four representative mammograms. Classification element size defines the resolution of the grid of texture measurements covering the breast area; it was found that a size of 4 x 4 pixels gives a satisfactory trade–off between segmentation accuracy and execution time. Neighbourhood size defines the area used for gathering local texture information for each element; a neighbourhood of 30 x 30 pixels is large enough to contain a representative sample of texture primitives, but not so large that the information can no longer be considered local. Granulometry involves applying a sequence of morphological operations using structuring elements of increasing size to cover the expected range of structure sizes in the image. Structuring element shape determines the texture characteristic to be measured. We used a set of eight circular structuring elements, of diameter 1, 3, 5, . . . , 15 pixels, to produce images in which intensity represents texture coarseness.

4.2 Texture Energy

Of the 16 feature masks suggested by Laws, we found the 5 x 5 versions of the edge and spot filters RR, LL and SR [17] to be the most powerful breast texture classifiers, based on an analysis of 40 mammograms which were selected to be representative of the normal range of appearance found in breast screening. Local variance was measured using 31 x 31 macro–window statistics. Segmentation was attempted using both single features and combinations of features. A Bayesian classifier [20] was trained on a leave one out basis, to produce glandular probabilities for an unseen mammogram based on combinations of texture energy features.

5 Results

Receiver operating characteristic (ROC) graphs were used throughout these experiments to compare the performance of one classifier with another, and to optimise the operation of each classifier. Intrinsic images in which intensity represents image brightness, granulometric coarseness, texture energy, or glandular probability were thresholded to classify the pixels as fatty or glandular. The ROC curves shows the true and false glandular classification rates, plotted over the complete range of possible thresholds. Hence the optimum threshold corresponds to the point on the curve closest to 100% true glandular, 0% false glandular (the top left corner); and the curves of two classifiers can be compared by noting which passes closer to that ideal point.

An ROC comparison shows that Laws single and combined texture energy features are the most accurate of the classifiers evaluated in this study (figure 3). The accuracy given for each method is the average classification accuracy obtained over 40 mammograms, using the optimum working point, and a routine to remove small isolated patches in the segmentation. The texture energy feature RR is the simpler
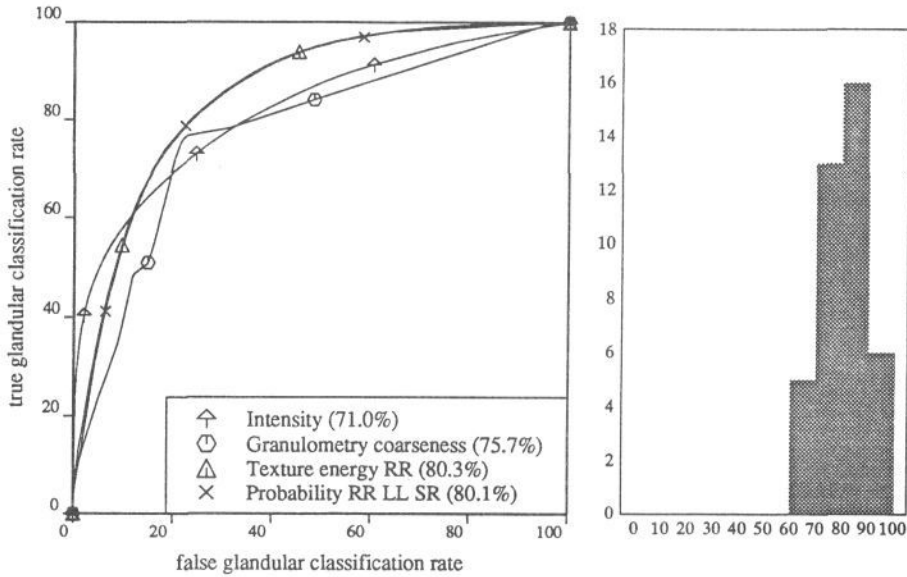
Figure 3. ROC curves for breast tissue classifiers

Legend:
- Intensity (71.0%)
- Granulometry coarseness (75.7%)
- Texture energy RR (80.3%)
- Probability RR LL SR (80.1%)

Figure 4. Histogram of RR texture energy classification accuracy for test data
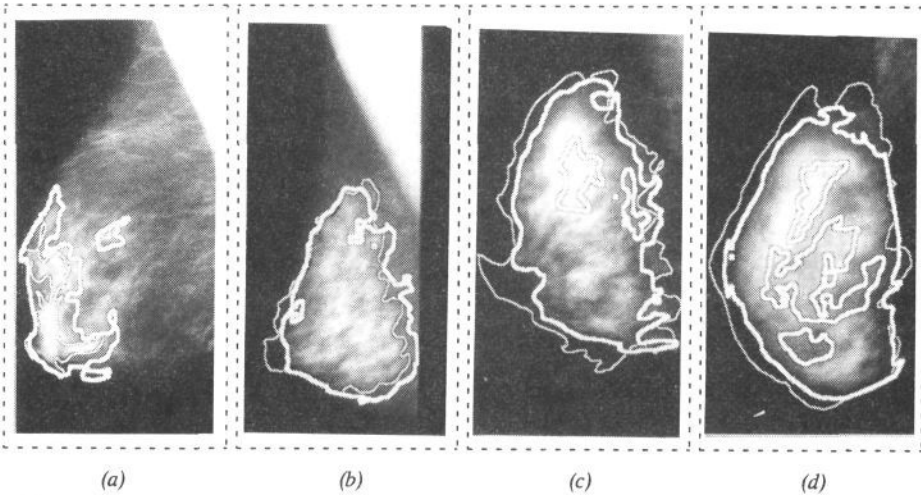

(a) (b) (c) (d)

Figure 5. Comparison of regions identified by radiologists with automatic classifications (thicker lines)

of the two Laws methods but is slightly more accurate, giving a narrow distribution of results over the data set (figure 4).

Examples of mammograms segmented using texture energy are presented in figure 5. Figure 5a shows a predominantly fatty breast, figure 5b a fatty–glandular breast, and figure 5c a glandular breast. In all of these cases a satisfactory segmentation has been obtained. The example shown in figure 5d is a dense glandular breast, representative of those in which the classification accuracy is lower.

6 Discussion

There appear to be two main causes of misclassifications. Inaccuracies in the automatic breast feature location routines, such as incorrect delineation of the pectoral muscle, are problematic. The most significant error occurs in breasts with dense glandular patterns, where areas of the glandular disc appear to be relatively smooth and are falsely classified as fat. This is likely to be resolved by improved digitisation; higher spatial and grey resolution would reveal more subtle textural details.

We have so far been unable to improve segmentation results using combined evidence from a set of Laws features, but it is believed that this goal will be achieved using principal component analysis of the feature distributions to reveal the optimum combination of features. We are also currently investigating alternative segmentation strategies [21] using the texture features, to ascertain whether such methods offer any improvement over threshold segmentation. Granulometry could be explored further to extract features which are more representative of mammographic texture, and which could be used in combination with the Laws features.

In summary, we believe that texture analysis forms a good basis for automatically classifying breast tissue. Further work should establish a viable automatic method for detecting and quantifying glandular tissue, and supplement the available tools for computer–assisted diagnosis of breast cancer.

Acknowledgements

References

1. Astley SM, Taylor CJ. Combining cues for mammographic abnormalities. Proceedings of the BMVC Conference, 1990, pp. 253–258.

2. Chan HP, Doi K, Vyborny CJ et al. Improvements in radiologists' detection of clustered microcalcifications on mammograms: the potential of computer–aided diagnosis. Investigative Radiology, 1990; 25:1102–1110.

3. Tabar L, Dean PB. Teaching atlas of mammography. Thieme, New York, 1985.

4. Law J. Patient dose and risk in mammography. British Journal of Radiology, 1991; 64:360–365.

5. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. American Journal of Roentgenology, 1976; 126:1130–1139.

6. Hufton A, Boggis CRM et al. The combined effects of breast composition and size in determining radiation risks from mammography. Annual Meeting of the British Institute of Radiology, 1990.

7. Kimme–Smith C, Frankl G, Wassel G, Sklansky J. Toward reliable measurements of breast parenchymal patterns. IEEE Computer Applications in Radiology and Analysis of Radiological Images VI, 1979, pp 118–121.

8. Shadagopan A, Alcorn FS, Semlow JL, Ackerman LV. Computerized quantification of breast duct patterns. Radiology, 1982; 143:675–678.

9. Magnin IE, Cluzeau F, Odet CL, Bremond A. Mammographic texture analysis: an evaluation of risk for developing breast cancer. Optical Engineering, 1986; 25:780–784.

10. Caldwell CB, Stapleton SJ, Holdsworth DW et al. Characterisation of mammographic parenchymal pattern by fractal dimension. Physics in Medicine and Biology, 1990; 35:235–247.

11. Hand W, Semmlow JL, Ackerman LV, Alcorn FS. Computer screening of Xeromammograms: a technique for defining suspicious areas of the breast. Computers and Biomedical Research, 1979; 12:445–460.

12. Giger ML, Yin F, Doi K et al. Investigation of methods for the computerised detection and analysis of mammographic masses. SPIE Medical Imaging and Image Processing IV, 1990; 1233:183–184.

13. Kimme C, O'Loughlin BJ, Sklansky J. Automatic detection of suspicious abnormalities in breast radiographs. Data Structures, Computer Graphics and Pattern Recognition, Academic Press, New York, 1975.

14. Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE, 1979; 67:786–804.

15. Du Buf JMH, Kardan M, Spann M. Texture feature performance for image segmentation. Pattern Recognition, 1990; 23:291–309.

16. Serra J. Image analysis and mathematical morphology. Academic Press, New York, 1982.

17. Laws KI. Textured image segmentation. Report 940, Image Processing Institute, University of Southern California, Los Angeles, 1980.

18. Matheron G. Random sets and integral geometry. John Wiley, New York, 1975.

19. Miller PI. Implementation of a breast finding algorithm. Internal Report, Wolfson Image Analysis Unit, University of Manchester, 1990.

20. Pearl J. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, San Mateo, 1988.

21. Spann M, Wilson R. A quad–tree approach to image segmentation which combines statistical and spatial information. Pattern Recognition, 1985; 18:257–269.