An Improved Method of Computing Scale-Orientation Signatures

Chris Rose* and Chris Taylor

Division of Imaging Science and Biomedical Engineering, University of Manchester, M13 9PT, UK

Abstract: Scale-Orientation Pixel Signatures provide a rich description of the neighbourhood around each image pixel. In Computer-Aided Mammography these signatures can be used as feature vectors for the classification of pixels as normal or abnormal. Previous work has focussed on the use of signatures in the classification process. In this paper we explain how signatures are generated and provide a critique of our current approach. We propose alternative methods of signature generation, along with a quantitative measure of signature 'quality'. We use this measure to compare the performance of our old and new methods and show that the new approach offers a significant advantage.

1 Introduction

As part of an on-going programme of research in Computer-Aided Mammography, we are interested in finding feature extraction methods that we can apply to digitised mammograms as a basis for classifying each pixel as representing normal healthy tissue or abnormal tissue, such as cancer, in order to aid radiologists and other mammogram readers. Previous work has established that computerised prompting can improve human detection performance, even when the prompting system is not completely accurate [1]. An individual abnormality in a mammogram can be considered as being one of the following: A cluster of microcalcifications, a mass, or a spiculated lesion (a mass with radiating abnormal tissue). Microcalcifications are small radiographically dense objects that are generally spherical in shape, clusters of which are indicative of malignancy. Masses are larger, blob-like in shape, less radiographically dense tissue that can be either benign or malignant. Spicules radiate from a central mass and can vary in length and thickness, however the mass can often be radiographically invisible and so linear structures that radiate from a common area can indicate the presence of a mass. For general utility, we are thus interested in being able to generate feature vectors that can represent blob-like and line-like structures over a range of scales.

2 Scale-Orientation Pixel Signatures

A Scale-Orientation Pixel Signature describes the local neighbourhood around a given pixel in terms of both scale and orientation. We can define a generic operator $h(s, \theta, f(x, y))$, which defines a set of new images from the scale and orientation parameters, where s is scale, θ is orientation and $f(x, y)$ represents the digitised mammogram. From these images we can construct the signatures for the original digitised mammogram. By choosing a good definition of the operator, we can construct Scale-Orientation Pixel Signatures that usefully represent the information in which we are interested. A signature is represented as a 2-D array where the rows represent scale and the columns represent orientation. The values in the array represent the change in grey level at the pixel, resulting from applying the operator at the scale and orientation corresponding to the position in the array. These changes are the grey-level differences between the result of applying the operator to the digitised mammogram at the current scale and the result of applying the operator at the next smaller scale. Pixel Signatures can be graphically represented by grey-level maps. Figure 1 shows Pixel Signatures taken from the centres of a Gaussian blob and a Gaussian line. The signatures display uniform scale over orientation (for the blob) and varying scale over orientation (for the line).

(a) (b)

Figure 1. Scale-Orientation Pixel Signatures for (a) a Gaussian blob and (b) a Gaussian line.

* E-mail: christopher.rose@stud.man.ac.uk

We have found that signatures generated for 12 orientations (from 0° to 180°) and 10 scales (arranged either linearly or logarithmically) are practically useful. Zwiggelaar *et al.* originally proposed Scale-Orientation Pixel Signatures in [2], and used a Directional Recursive Median Filter for the operator with a linear scale. M- and N-Filters [3] are a class of filters known as sieves; M-Filtering is achieved by performing morphological closing followed by morphological opening with a structuring element (readers are directed to [4] for more information). Holmes and Taylor [5] use an M-Filter with a 1-D structuring element (SE) of varying length (i.e. the scale information in the signature is found by varying the length of the SE), and orientation. The signatures shown in Figure 1 were calculated using this method, using a logarithmic series of scales. When applied to a digitised mammogram of size $M \times N$ pixels, the result is $M \times N$ signatures, each one containing 10×12 elements. The rows or columns of each signature may be concatenated to form a feature vector that can be used in a classification process and which provides a representation of each pixel of the digitised mammogram in 120-D space.

3 Improving Scale-Orientation Pixel Signatures

Generating the Scale-Orientation Pixel Signatures is the first step in our current digital mammography image analysis procedure, and so it would be desirable to have the most complete representation we can obtain. There are two main problems with the method used in [5]. Firstly the SE is one-dimensional, and so when an SE of length l has been rotated by θ (e.g. from P to Q in Figure 2 (a)) there is a sector of area $(1/2)l^2\theta$ (if we consider the non-quantised situation) that has not been sampled, as shown by the solid filled sector in Figure 2 (a). If this sector in the mammogram contained fine structure, this would result in aliasing.

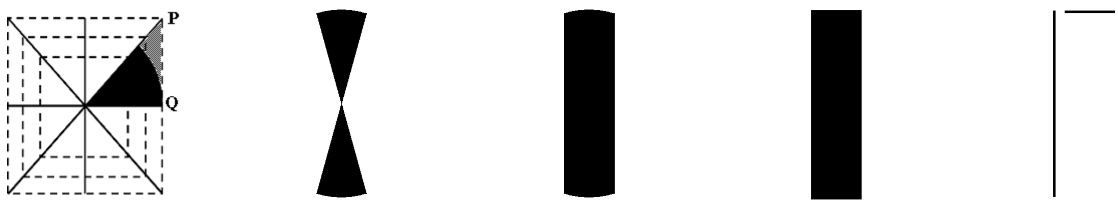


Figure 2. (a) A simplified representation of the SE used in [5], with 3 scales and 4 orientations, showing where under-sampling occurs; (b) The bow tie SE; (c) Simplified bow tie SE; (d) Rectangular SE; (e) 1-D approximation to the rectangular SE using two perpendicular SEs.

The second problem is that as θ increases, the SE traces out a square, rather than a circle. So at certain angles, the SE is longer than it should be (and it could be argued that the area that is not sampled extends to the grey shaded area in Figure 2 (a)). So we have aliasing, and at some orientations the scale information we compute is incorrect. The latter explains the two troughs in the signature in Figure 1 (a). Figure 2 (b) to (e) shows several potential solutions to these problems that do not leave any gaps when rotated by θ, and which trace out filled circles ((d) would produce an approximately circular filled shape) as they are rotated through all 180° at the discrete steps that form the columns of the signature. Using any of the SEs in Figure 2 (b) to (d) would be computationally expensive compared to the 1-D SE previously used. Experimental work showed that using two 1-D SEs (Figure 2 (e)) of length l and $w = 2l\sin(\theta/2)$ (see Figure 3 (a)) perpendicular to each other, and sieving with the SE of length l first is a close approximation to the rectangular SE, where around 60% to 90% of the pixel values are exactly the same as for the rectangular method, the remaining pixel values are close to those of the rectangular method, and the resultant images are visually almost identical. Using 1-D SEs, Soille's algorithm can be used to efficiently perform the sieving [6]. Figure 3 (b) illustrates how the new method solves the problems illustrated in Figure 2 (a).

Figure 3. (a) As the SE of length *l* is rotated by θ, *w* should be such that the edges of this SE touch when rotated. (b) A Scale-Orientation Pixel Signature generated with the new method, taken from the centre of a Gaussian blob (note the improved structure over that in Figure 1 (a)).

4 Entropy as a Measure of Signature 'Quality'

Now that we have a new operator with which to construct Scale-Orientation Pixel Signatures, how can we compare signatures generated by the two methods to determine which method is better? An obvious approach would be to conduct an experiment to classify Regions of Interest (ROIs) from digitised mammograms using each of the methods, and then determine which method gave the best classification result. However this is not really what we want to achieve – this approach would tell us which type of signature was best for a particular classification technique, and we would need to use many different types of classifiers to determine which type of signature was the best overall. An alternative measure of signature 'quality' is required. In producing signatures, we are hoping to encapsulate information about the structures in the digitised mammogram. We propose that a sensible approach is to measure the information content of a set of signatures, using the concept of Entropy [4] from Information Theory. If we have a set of symbols $A = \{a_1,...,a_n\}$ and $P(a_i)$ is the probability of the *i*-th symbol, then Entropy can be defined as:

$$H = -\sum_i P(a_i)\log_2 P(a_i)$$

If the base of the logarithm is 2, the Entropy is expressed in bits. We can regard the quantised value of each element of a signature as a symbol and so, by constructing a frequency table of the values contained in a signature, we can compute the probability of the symbols occurring. Note that we are not constructing a histogram where the bins each have an associated range; an element in the signature that has a unique value is placed in its own bin. As the number of unique values within a signature increases, so does the Entropy, and so the more information the signature contains.

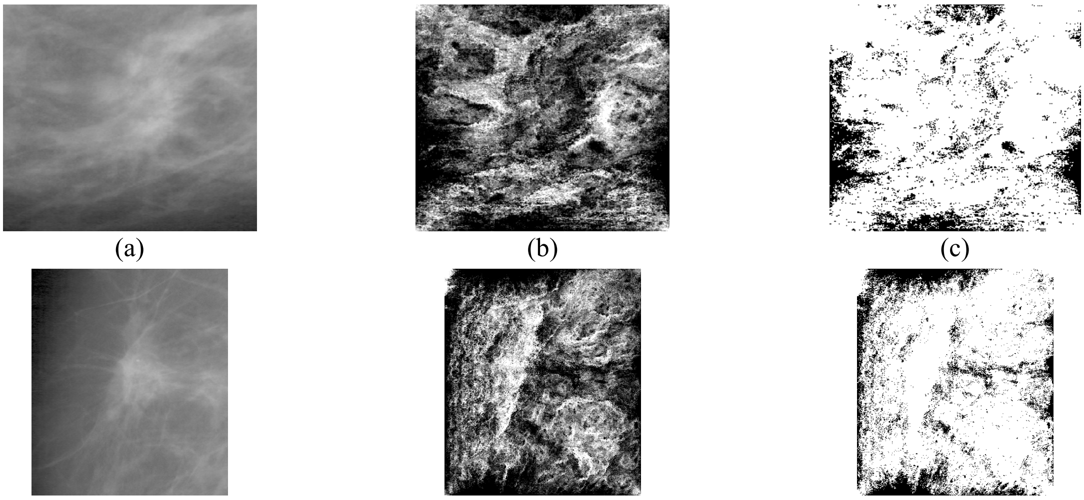
5 Method and Results

To compare the operator shown in Figure 2 (e) for constructing Scale-Orientation Pixel Signatures with the previous operator we selected a set of 20 ROIs from digitised mammograms from the Digital Database for Screening Mammography (DDSM) [7], 10 of which were from normal mammograms and 10 contained spiculated lesions (with variable mass visibility). Due to time constraints on the experiment microcalcifications were not considered though we may explore this research in the near future. The ROIs were selected to represent a range of subtlety, from obvious spiculated lesions to more subtle ones; the normals were selected to represent a range of normal tissue, from homogenous to more highly structured areas. As a result, the sizes of the images varied. The images were sub-sampled from 43.5µm to 100µm to reduce computational cost. Note that although it would appear that we only used a small number of images, we are actually comparing the signatures produced, of which there is one per pixel: Therefore the sample we looked at was actually quite large, at nearly 2 million. A summary of our results is shown in Table 1.

No. Pixel Signatures	Total Entropy Using Orig. Operator	Total Entropy Using New Operator	Average Difference Per Pixel (Orig. Operator - New Operator)	Average Percentage Increase
1,997,921	4,631,821 bits	5,132,849 bits	0.251 bits	10.82%

Table 1. Results of comparing the previous and new operators on 10 normal and 10 abnormal ROIs.

The previous operator typically achieved between 1.9 bits/pixel and 3.1 bits/pixel. There was no significant difference between the performance increase offered by the new operator for the normal vs. abnormal ROIs, each giving approximately 11% more information. For each ROI we produced images that represent the difference in signature Entropy between the previous operator and the new operator; these images tell us about where in the ROI the information is distributed. Figure 4 shows two examples of these images.



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

Figure 4. (a) and (d) show ROIs containing spiculated lesions; (b) and (e) show the difference between the information contained in the signatures generated using the previous operator and those generated with the new operator (light areas are where the new operator performed better); (c) and (f) are the images (b) and (e) thresholded at zero – white is where the new operator performs better.

6 Discussion

Scale-Orientation Pixel Signatures can be used to detect features such as masses and linear structures in mammograms. We have described elsewhere [2] how these methods can be combined to detect lesions as part of a system for computer-aided mammography. The aim of this work was to investigate how the previous operator for generating Scale-Orientation Pixel Signatures could be improved, and to devise a quantitative measure of how good signatures generated by a particular operator are. Using our proposed operator on 20 ROIs we increased the information contained in the resultant signatures by an average of 11%. From inspecting the signature Entropy images it is interesting to note that the proposed operator gave a higher performance on tissue surrounding mass-like areas (see Figure 4 (e)), and often on linear structures such as spicules (see Figure 4 (b)). In none of our experiments did the new operator produce a set of signatures for an image that contained less information than the previous operator. We will shortly be replicating an experiment described in [8] using data from 36 mammograms, to test for evidence that our new method for computing signatures has improved diagnostic accuracy. Future work may include determining the maximum entropy that a signature that usefully represents mammographic structure can have, experimenting with other operators and studying the performance of such operators on synthetic models of digitised mammograms.

Acknowledgements

Chris Rose is funded by the EPSRC, and would like to thank Anthony Holmes for his encouragement and eager help.

References

1. S. M. Astley, R. Zwiggelaar, T. C. Parr and C. J. Taylor, "Prompting in mammography: how good must prompt generators be?". *International Workshop on Digital Mammography*, 1998.
2. R. Zwiggelaar, T. C. Parr, J. E. Schumm et al. "Model-based detection of spiculated lesions in mammograms". *Medical Image Analysis* **3(1)**, pp. 39-62, 1999.
3. R. Harvey, A. Bosson & J. A. Bangham. "The robustness of some scale-spaces". In *British Machine Vision Conference*, pp. 11-20, 1997.
4. M. Sonka, V. Hlavac, R. Boyle. *Image Processing, Analysis, and Machine Vision*. Brooks/Cole Publishing Co., 1999.
5. A. S. Holmes & C. J. Taylor. "A Run-time Method of Estimating Pixel Signature Similarity". In *Medical Image Understanding and Analysis*, pp 85-88, 2000.
6. P. Soille, E. J. Breen & R. Jones. "Recursive implementation of erosion and dilations along discrete lines at arbitrary angles". *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18(5)**, pp. 562-567, 1996.
7. M. Heath, K. W Bowyer, D. Kopans et al. "Current status of the Digital Database for Screening Mammography". In *Digital Mammography*, pp. 457-460, Kluwer Academic Publishers, 1998.
8. A. S. Holmes. *Computer-aided Detection of Abnormalities in Mammograms*. PhD thesis, 2001 (submitted).