Abnormal Masses in Mammograms: Detection Using Scale-Orientation Signatures

Reyer Zwiggelaar[1] and Christopher J. Taylor[2]

[1] Division of Computer Science, University of Portsmouth, Portsmouth, UK
reyer@sis.port.ac.uk
[2] Wolfson Image Analysis Unit, University of Manchester, Manchester, UK
ctaylor@man.ac.uk

Abstract. We describe a method for labelling image structure based on scale-orientation signatures. These signatures provide a rich and stable description of local structure and can be used as a basis for robust pixel classification. We use a multi-scale directional recursive median filtering technique to obtain local scale-orientation signatures. Our results show that the new method of representation is robust to the presence of both random and structural noise. We demonstrate application to synthetic images containing lines and blob-like features and to mammograms containing abnormal masses. Quantitative results are presented, using both linear and non-linear classification methods.

1 Introduction

We are interested in labelling important structures in images. We assume that the position of these structures is unpredictable and that they will be embedded in a background texture. Real examples of this class of problem are ariel images - containing structures of interest such as roads, rivers and trees, and medical images containing blood vessels, ducts and focal abnormalities (e.g. tumours).

We describe an approach based on the construction of a scale-orientation signature at each pixel. This provides a very rich description of local structure which is robust and locally stationary. Given this description, standard statistical classification methods can be used - we give results for both linear and non-linear approaches for synthetic and real medical data.

2 Scale-Orientation Signatures

The Recursive Median Filter (RMF) is one of a class of filters, known as sieves, that remove image peaks or troughs of less than a chosen size [1]. They are closely related to morphological operators [8]. By applying sieves of increasing size to an image, then taking the difference between the output image from adjacent size sieves, it is possible to isolate image features of a specific size. Sieves have been shown to have desirable properties when compared to other methods [4] of constructing a scale space [2]. In particular the results at different positions on the same structure are similar (local stationarity) and the interaction between adjacent structures is minimised.

2.1 Describing Local Structure

For 2-D images, a 1-D RMF can be applied at any chosen angle, by covering the image with lines at this angle, ensuring that every pixel belongs to only one line. By performing 1-D Directional Recursive Median Filtering (DRMF) at several orientations, a scale-orientation signature can be built for each pixel. The signature is a 2-D array in which the columns represent measurements for the same orientation, the rows represent measurements for the same scale, and the values in the array represent the change in grey-level at the pixel, resulting from applying a filter at the scale and orientation corresponding to the position in the array. The grey-level changes are measured with respect to the image filtered at the next smaller scale at the same orientation.

Fig. 1 shows scale-orientation signatures for pixels located on synthetically generated structures. For instance, the response of the DRMF to a binary blob will result in a signature which has values at only one scale and is equal for all orientations (Fig. 1a). For a binary line the resulting signature is highly scale and orientation dependent, with the minimum scale related to the width of the line, and the maximum scale related to the length of the line (Fig. 1b). When the structures get more realistic, such as the Gaussian lines and blobs shown in Fig. 1 c and d, the signatures become slightly more complicated, but the overall shape remains similar. For blob-like structures the centre pixel gives a very characteristic signature, where the scales at which information is present in the signatures are related to the diameter of the structure. This is also true for pixels on the backbone of linear structures, for which the minimum scale in the signatures is related to the width of the linear structure, the maximum scale related to length, and the orientation at which this maximum occurs indicates the direction of the linear structure. Although the signatures for non-centre pixels are not identical they are usually very similar (compare the columns of signatures for each structure in Fig. 1). This local stationarity property is useful for pixel classification. The largest differences occur for the blob-like structures - there is a continuous change between the signatures from the centre of the blob-like structure to the extreme edge with small changes for pixels near the centre and relatively large changes towards the edge of the structure.

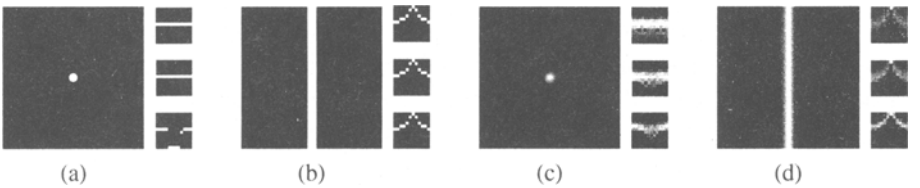

Fig. 1. Some synthetic examples of multi-scale DRMF signatures, where the larger four images show (a) a binary blob, (b) a binary linear structure, (c) a Gaussian blob and (d) a Gaussian linear structure. The twelve smaller images are the scale-orientation signatures for the centre pixel (top), for a pixel at the extreme edge of the structure (bottom) and for a pixel in between these two extremes (middle). In the smaller scale-orientation signature images, scale is on the vertical axis (with the finest scale at the bottom) and orientation on the horizontal (the background grey-level is zero, i.e. only positive values are present in the DRMF signatures).

2.2 Noise Aspects

To show the effects of noise some example structures and scale-orientation signatures are shown in Fig. 2. The signal to noise ratio (SNR) is increased from left to right from infinite (i.e. no noise present) to 25%. It is clear that the signatures remain stable for signal to noise ratios larger than 0.5. but even below this value certain features in the signatures remain stable, with a band of values across all orientations between two distinct scales. Also note that in the noisy signatures a substantial change occurs at the smaller scales, which are characteristic of the noise that is present in the images. This behaviour of the scale-orientation signatures in the presence of noise, together with the effect shown in Fig. 1, makes it possible to obtain robust classification of pixels.

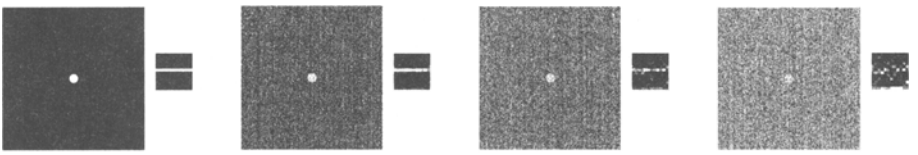

Fig. 2. Some synthetic examples of multi-scale DRMF signatures to indicate the effect of noise on the scale-orientation signatures (the same format as in Fig. 1 is used and only the signature for the centre pixel is shown) ,where the signal to noise ratio is ∞ (a), 1.0 (b) , 0.5 (c) and 0.25 (d).

3 Statistical Modelling

A brief description of the linear and non-linear modelling techniques used in our experiments are discussed in this section.

3.1 Principal Component Analysis

Principal component analysis (PCA) is a well documented statistical approach to data dimensionality reduction [3]. The principal components of a population of observation vectors are the characteristic vectors of the covariance matrix ($\mathbf{C}$) constructed from the population. Projecting the data into its principal components generally results in a compact and meaningful representation in which the first few characteristic vectors describe the major modes of data variation. The characteristic values provide the variances of the principal components. Data dimensionality reduction is achieved by ignoring those principal components which have zero or small characteristic values.

3.2 Linear Classification

The objective of the work is to classify pixels, that is to label each pixel as belonging to a certain type of image structure. Since any method is likely to be imperfect it is useful to explore a range of compromises between false negative errors (poor sensitivity) and false positive errors (poor specificity). This can be achieved conveniently by

constructing a probability for each pixel. The starting point is an observation vector, $\mathbf{x}_i$, for each pixel i, describing properties relevant to the classification task. For each class, ω_j (e.g. normal or abnormal), the mean, $\mathbf{m}_{\omega_j}$, and covariance, $\mathbf{C}_{\omega_j}$, of the observation vectors is estimated from a training set of signatures in which every pixel has been annotated with the appropriate class by an expert. The probability density of obtaining an observation vector $\mathbf{x}_i$ for a pixel of class ω_j is given by $p(\mathbf{x}_i|\omega_j)$ which can be calculated from $\mathbf{m}_{\omega_j}$ and $\mathbf{C}_{\omega_j}$ assuming a Gaussian distribution. Applying Bayes theorem, a probability image for class ω_j (e.g. abnormal) is found by calculating, for each pixel $P(\omega_j|\mathbf{x}_i)$. Detection can be performed by thresholding the resulting probability image. Different values of the threshold will result in different compromises between true positive and false positive errors. The detection performance as the threshold is varied can be summarised conveniently using Receiver Operating Characteristic (ROC) curves [5].

3.3 Non-Linear Classification

To assess the non-linear aspects of the scale-orientation signatures a basic back-propagation artificial neural network (ANN) was used [6]. The same network was used in all experiments. The architecture comprised an input layer of the 132 components (12 orientations $\times$ 11 scales) of the the scale-orientation signatures, two hidden layers and the output layer comprising three units for the synthetic data and two units for the mammographic data (these numbers were used to obtain a direct comparison with the linear classification which provides three and two class probabilities, respectively). The network was fully connected between the second hidden layer (10 units) and both the output layer and the first hidden layer (23 units). The connections between the input layer and the first hidden layer were, however, more restricted. The information from one scale (summed across all orientations) or one orientation (summed across all scales) was connected to each of the units of the first hidden layer.

4 Test Data

To assess the potential of the scale-orientation signatures for the classification of structures in images two different datasets were used: one synthetic, the other derived from real mammographic data. In both cases the datasets were divided into three equal sized subsets - facilitating a training, validation (only used for the non-linear classifier) and test set.

The synthetic data consisted of three separate datasets of 12288 scale-orientation signatures each (see Fig. 2d for a typical example). Each dataset contained equal numbers of signatures obtained from Gaussian linear structures, Gaussian blob-like structures and texture background (which was based on a combination of Gaussian and shot noise). The signatures were extracted from the top 5% of brightest pixels (before noise was added) from 768 images. All the samples were given class expectation values which could be used to determine the mean square error for a certain network.

The second dataset used in our experiments was drawn from mammograms. The signatures were extracted from 54 mammograms of which half contained an abnormality (a spiculated lesion). The available data was separated into three datasets, each comprising

signatures from 18 mammograms. The resulting mammographic data comprised three separate datasets of 2700 scale-orientation signatures, each containing equal numbers of "normal" and "abnormal" signatures. The "abnormal" signatures were taken from the annotated central mass while the "normal" signatures were randomly selected from the normal mammograms.

5 Classification of Synthetic Data

We present results for classifying the synthetic data into three classes, using both linear and non-linear classifiers.

A principal component model was trained for every synthetic training data set. The first five principal components cumulatively explained approximately 49%, 60%, 64%, 69% and 73% of the training set variance respectively. A linear classifier was used to classify each signature in the datasets as belonging to the class of linear structure, blob-like structure or the texture background as described in Sec. 3.2. This information was used to obtain class probabilities.

For this data the ANN had three output neurons for the three classes; linear structures, blob-like structures and background texture. The class expectation values used were 0.90 for class samples and 0.05 for non-class samples. The three expectation values were thus directly comparable to the probabilities obtained using a linear classifier with their sum equal to 1.0. The optimal network was found (as described in Sec. 3.3) and used to obtain the expectation values which were used for classification of the data.

Classification accuracy was assessed by producing ROC curves as shown in Fig. 3. There is no difference between the use of the full scale-orientation signature and the reduced dimensionality representation. A sensitivity of 80% is obtained at a false positive rate of 0.15.

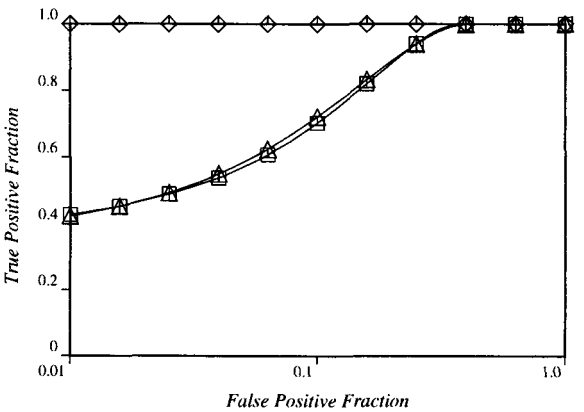

Fig. 3. Synthetic signature classification results for linear classification (△), PCA model based linear classification (□) and non-linear classification (◇).

The ANN performs almost perfect classification which compares very favourable with the linear classifier. 80% sensitivity is achieved at a false positive rate of 0.000013 and 98% sensitivity at 0.00005.

6 Classification of Mammographic Data

The classification of the mammographic data is into two classes, either abnormal mass or non-mass. Again both classification based on linear and non-linear methods are investigated and compared. In addition, application of the approach to images is discussed, with a typical example illustrating the differences in results leading to free response operating characteristic (FROC) curves and the number of false positives as a function of probability threshold.

A principal component model was trained on the mammographic data (individual results for the three datasets were similar). The first five principal components cumulatively explained approximately 31%, 49%, 57%, 62% and 66% of the training set variance respectively. A linear classifier was used to classify pixels as mass/non-mass. The resulting ROC curves are shown in Fig. 4, indicating the effect of using the full signatures or the first 17 principal components (using the principal component model to explain 85% of the variation in the data) in the dataset. Using the full signatures, slightly worse results are obtained. For the results based on the first 17 principal components a sensitivity of 80% is obtained at a false positive fraction of 0.25.

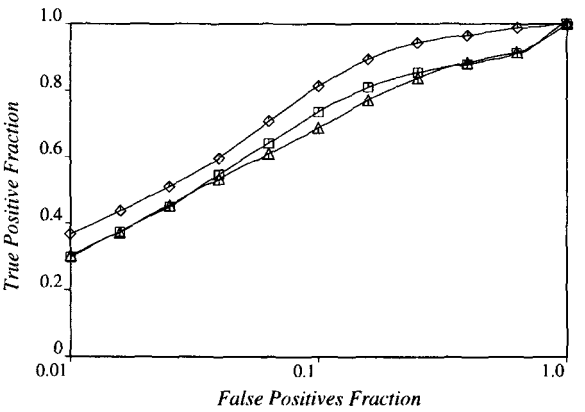

Fig. 4. Mammographic signature classification results for linear classification (△), PCA model based linear classification (□) and non-linear classification (◇).

In this case the ANN had two output neurons for the two classes, mass and non-mass. Again, every sample was given class expectation values. In this case 0.9 for mass samples and 0.1 for non-mass samples. The same approach as in Sec. 5 was followed to obtain ROC curves. The results are shown in Fig. 4 which show that a sensitivity of 80% was achieved at a false positive rate of 0.10.

6.1 Region Classification

An example of a spiculated lesion with central mass is shown in Fig. 5a. A class probability image resulting from the full scale-orientation signatures and a linear classifier is shown in Fig. 5b, results from the first 17 principal components and a linear model are shown in Fig. 5c, and from he non-linear classifier in Fig.5d. This shows that the mass is detected by all three approaches. It should be noted that the area (i.e. the region of pixels) detected has an improved overlap with the annotated region when going from full signature linear model (Fig. 5b), principal component linear model (Fig. 5c) to full signature non-linear model Fig. (5d).

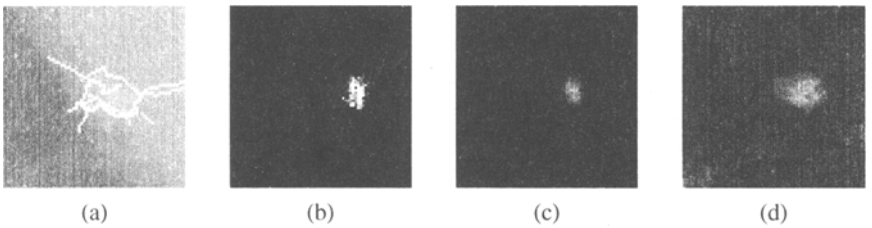

Fig. 5. Example of applying the linear and non-linear classification approach to a section of a mammogram. (a) original mammogram (b) original mammogram with central mass and spicules annotated (c) full signature linear classification probability image (d) first 17 principal component based linear classification probability image (e) non-linear classification probability image.

The class probability images where obtained for all the mammograms in dataset. Region classification (region size equal to 10 mm) was based on these images and results are shown in Fig. 6. At a sensitivity of 80% the number of false positives per image reduces from 7.5 for linear classification to 3.2 for non-linear classification.

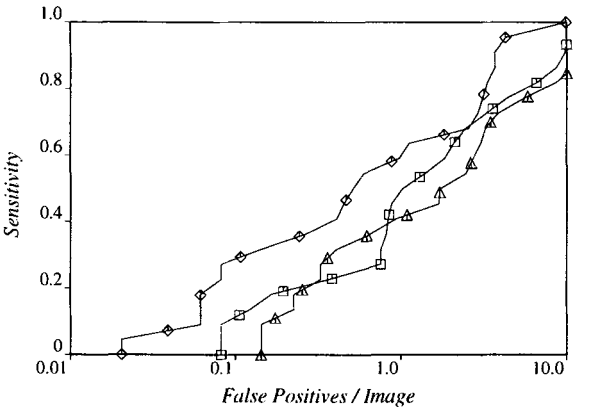

Fig. 6. FROC mammographic region classification results for linear classification (△), PCA model based linear classification (□) and non-linear classification (◇).

7 Discussion and Conclusions

We have described a principled method for the classification of pixels based on scale-orientation signatures which provide a compact description of local structure. Linear and non-linear aspects of scale-orientation signatures can be used for the labelling of structures in images. It is possible to classify pixels into several classes such as linear structures, blob-like structures or background texture.

For the synthetic data sets the results obtained using the ANN are considerably better for unseen data than the classification using a linear approach implying better generalisation. At a sensitivity of 80% the false positive rate reduces from 0.15 for the linear method to 0.000013 for the ANN.

When comparing the linear and non-linear classification results for the mammographic data it is clear that allowing for non-linear behaviour of the scale-orientation signatures provides as overall improvement. At a sensitivity of 80% the false positive rate improved from 0.25 to 0.10. When applied to images the results show an overall better performance for the non-linear modelling approach. For region classification at a sensitivity of 80% the number of false positives per image reduces from 7.5 for linear classification to 3.2 for non-linear classification. These results are comparable with methods presented in the literature [7, 9].

Results for both synthetic and mammographic data indicate that the non-linear aspects of the scale-orientation signatures provide additional information which improves the classification results. Using the non-linear technique described the detection of spiculated lesions is improved.

The approaches described for the classification of structures in images could be improved by the normalisation of the scale- orientation signatures to provide a better classification of the various structures, although some specificity may be lost.

References

1. J.A. Bangham, T.G. Campbell, and R.V. Aldridge. Multiscale median and morphological filters for 2d pattern recognition. *Signal Processing*, 38:387–415, 1994.
2. R. Harvey, A. Bosson, and J.A. Bangham. The robustness of some scale-spaces. In *Proceedings of the* 8^{th} *British Machine Vision Conference*, pages 11–20, Colchester, UK, 1997.
3. I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.
4. S.N. Kalitzin, B.M. ter Haar Romeny, and M.A. Viergever. Invertible orientation bundles on 2d scalar images. In *Lecture Notes in Computer Science*, 1252:77–88, 1997.
5. C.E. Metz. Evaluation of digital mammography by roc analysis. *Excerpta Medica*, 1119:61–68, 1996.
6. University of Stuttgart. *Suttgart Neural Network Simulator*. User Manual, Version 4.1, 1995.
7. N. Petrick, H.-P. Chan, B. Sahiner, M.A. Helvie, M.M. Goodsitt, and D.D. Adler. Computer-aided breast mass detection: false positive reduction using breast tissue composition. *Excerpta Medica*, 1119:373–378, 1996.
8. P. Soille, E.J. Breen, and R. Jones. Recursive implementation of erosions and dilations along discrete lines at arbitrary angles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):562–567, 1996.
9. W.K. Zouras, M.L. Giger, P. Lu, D.E. Wolverton, C.J. Vyborny, and K. Doi. Investigation of a temporal subtraction scheme for computerized detection of breast masses in mammograms. *Excerpta Medica*, 1119:411–415, 1996.