BMVC99

Separating Background Texture and Image Structure in Mammograms

Reyer Zwiggelaar
Division of Computer Science, University of Portsmouth,
Milton Campus, Locksway Road, Southsea PO4 8JF, UK
reyer.zwiggelaar@port.ac.uk
http://www.dcs.port.ac.uk/~reyer/

Abstract

There have been several approaches to the classification of texture in images. Most approaches will take certain local attributes or features into account and base the classification on these measures. In here we demonstrate the use of a statistical approach to separate the structure and texture background present in images. Modelling is based on *normal* images which only contain a texture background. The resulting model is applied to images which contain *abnormal* image structures as well as a *normal* texture background. Especially for mammographic (and other medical application) this can provide useful information which can be used as a pre-processing tool to obtain the structures present in the image and at the same time get a robust classification of the background texture.

1 Introduction

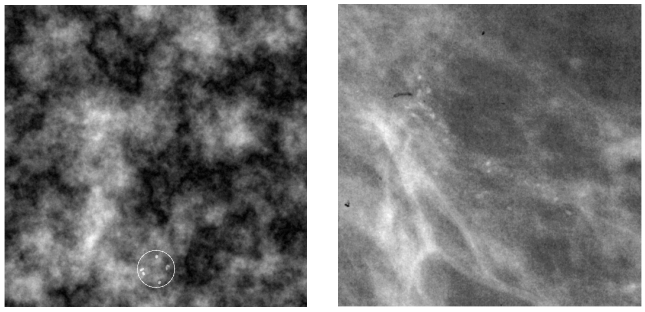
The UK Breast Screening Programme alone generates 1.5 million mammograms per annum. Potential malignancies can be detected from subtle abnormalities in radiographic appearance but it is known that radiologists fail to detect a significant proportion of these abnormalities. It has been shown that their performance would improve if they were prompted with the locations of possible abnormalities [10].

Mammographic parenchymal tissue can be divided into a number of classes. These classes (e.g. Wolf grades [9]) are used by radiologists for the assessment of risk in mammographic images. Methods to automaticly classify the tissue in the same risk classes have only had limited success [6]. However, of the used approaches the most successful have in common the assumption that fractal measures can be used for the classification of the background texture.

Besides the background texture, image structures (both linear and blob-like) can carry information about possible mammographic abnormalities. See Fig. 1b for an example of a patch of a mammographic image containing a cluster of micro-calcifications. Linear structures (spicules, vessels, ducts, etc.) can be detected by various methods [11]. This is also the case for blob-like structures (masses and micro-calcifications) [8, 10]. However, it is possible that these structures (*normal* and *abnormal*) cause confusion when trying to classify the background texture. Therefor a separation of the background texture and the image structures could make the texture classification stage more robust.

BMVC 1999 doi:10.5244/C.13.36

We propose a method, based on the information contained in Fourier space, which will capture the texture information and at the same time will separate the image structure from the background texture. Effectively the described approach is a form of image preprocessing to enhance the structures present in the images and to reduce the effects of background noise. Other methods to enhance the structures in images are well known [5].



(a) (b)

Figure 1: (a) Synthetic and (b) mammographic examples of images that contain *abnormal* structures and *normal* texture background. For the synthetic image the abnormalities are within the circle and for the mammographic example the abnormalities can be found as brighter blobs along the linear structure which extends approximately as $x = -y$.

2 Fourier Space Revisited

There is a direct link between the information contained in Fourier space and between the fractal model which describes the texture in the image. A short description of the Fourier space theory [1] and the relation with fractal descriptors [7] is given. The discrete Fourier transform $F(\mu, \nu)$ of the function $f(x, y)$ is given by

$$F(\mu,\nu) = \sum\sum f(x,y) e^{-\imath 2\pi (x\mu + y\nu)/N} \tag{1}$$

Here the coordinates (x, y) refer to the spatial frame (i.e. the real world image) and (μ, ν) to the frequency frame. It should be clear that the resulting function F is a complex function (indicated by the $\imath$ in Eq. 1). Because the function f is a real function (no imaginary part) the real part of F is even and the imaginary part of F is odd (F is an Hermitian function).

2.1 Fractal Modelling

If the information contained in the function $f(x, y)$, which is the image, can be described as being fractal the power spectrum is given by [7]

$$\begin{aligned} P(\mu,\nu) &= F(\mu,\nu)F^*(\mu,\nu) \\ &= (reF + imF\,\imath)(reF - imF\,\imath) \\ &\propto (\mu^2+\nu^2)^{-\beta}, \end{aligned} \tag{2}$$

where β is related to the Hausdorff dimension for fractal structures and reF and imF represent the real and imaginary part of the Fourier transform F. For most images Eq. 2 is an approximation (which might only be correct over a limited spectral range, or a series of ranges all with their own local Hausdorff dimension).

Our hypothesis is that this approach to image (texture) modelling can be used to separate the (non-fractal) structures in the image from the (fractal) background. To achieve this an approximation of the original power spectrum $P(\mu, \nu)$, which is indicated by $P'(\mu, \nu)$, is used to reconstruct the original image by transforming reF taking the original imF and Eq. 2 into account. The transformed reF is indicated by reF' and is given by

$$|reF'| = \sqrt{P'(\mu, \nu) - imF^2} \tag{3}$$

After this reconstruction the information in reF' and imF are used in an inverse Fourier transform operation to obtain a reconstruction of the original image. Even though it was suggested previously that the imaginary part of F does not carry any significant information for texture classification [2], the above described approach does not results in a robust and satisfactory separation of the image structures and the background texture. As imF was not changed the original image was almost perfectly reconstructed [2].

2.2 Statistical Modelling

In the previous section it was suggested that the fractal texture information can be separated from the overall image through reconstruction of the real and imaginary part in Fourier space. However, the need to change both reF and imF (else reconstruction is almost perfect) makes a direct approach (as described above) not possible. To obtain an approximation of both reF and imF we have investigated the use of principal component analysis with respect to the information in Fourier space.

Principal component analysis (PCA) is a well documented statistical approach to data dimensionality reduction [4]. The principal components of a population of observation vectors are the characteristic vectors of the covariance matrix ($\mathbf{C}_{\omega_j}$) 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. Observation vectors (in this case the reF and imF information in the Fourier domain) can be approximated from a PCA model using (4).

$$\mathbf{x}_i \approx \mathbf{P}_{\omega_j}\ \mathbf{b}_i + \mathbf{m}_{\omega_j} \tag{4}$$

where $\mathbf{x}_i$ is the i^{th} observation vector, $\mathbf{m}_{\omega_j}$ is the mean observation over the population, $\mathbf{P}_{\omega_j}$ is the matrix of the most significant characteristic vectors and $\mathbf{b}_i$ is a vector of lower dimensionality than $\mathbf{x}_i$ (hence the approximation sign). The weights of the principal components ($\mathbf{b}_i$) for an observation ($\mathbf{x}_i$) can be estimated (since $\mathbf{P}_{\omega_j}^{-1} = \mathbf{P}_{\omega_j}^T$) by

$$\mathbf{b}_i = \mathbf{P}_{\omega_j}^T\ (\mathbf{x}_i - \mathbf{m}_{\omega_j}) \tag{5}$$

Observation vectors can be reconstructed by substituting (5) into (4).

3 Artificial Images

To investigate the proposed approach to the separation of the structure and texture in images synthetic images were generated which either contained a (*normal*) fractal texture or a fractal texture in combination with a cluster of (*abnormal*) blob-like structures. A typical example of a *abnormal* synthetic image is shown in Fig. 1a.

The PCA model was based on a large number ($N \gg 1000$) of grey-level normalised sections (the regions of interest which were 32×32 pixels in size) extracted from *normal* synthetic images. For each region of interest the ${}^{10}log$ of reF and imF in Fourier space were determined. To obtain orientation independent results the extracted sections were rotated through an angle of $\pi/2$ radians four times and all four samples were used.

With the available information there are to possible approaches. Firstly it is possible to combine the information provide in reF and imF and build one observation vector for the PCA model. However, it is also possible to build two individual PCA models, one based on reF and the second based on imF.

3.1 Combined reF and imF PCA Model

The combination of the reF and imF information from the Fourier domain results in an observation vector which has 2048 dimensions. The resulting PCA model needs 1021 components to describe this data (the reduction by a factor of two is expected as the Fourier domain is rotational symmetric). To describe 25% of the data variation 5 principal components were needed and to describe 75% of the data variation close to 300 principal components were necessary. The effect of the first three principal components and the mean of reF and imF are shown in Figs 2 and 3, respectively.

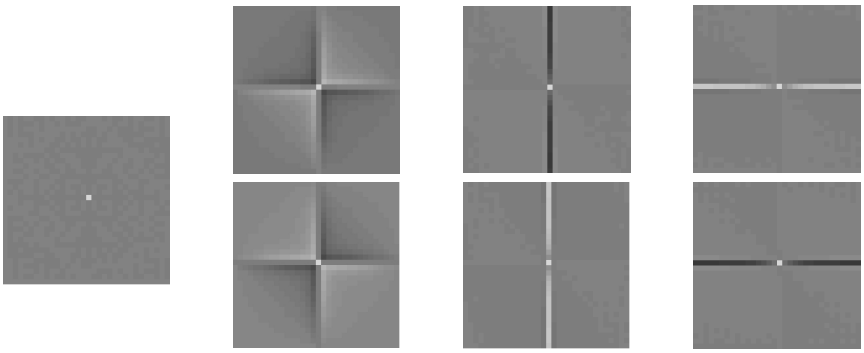


Figure 2: The mean (left image) and first three (from left to right) principal components (± 2 standard deviations) of reF for the PCA model based on the combined information from reF and imF.

The results in Fig. 2 indicate several effects. The first principal component indicates rotational aspects whilst the second and third principal components indicate the presence of horizontal and vertical structures in the image, respectively. No variations in the intensity were expected as the regions of interest were normalised with respect to the local grey-level distribution. The mean of reF is almost a delta function. These effects are

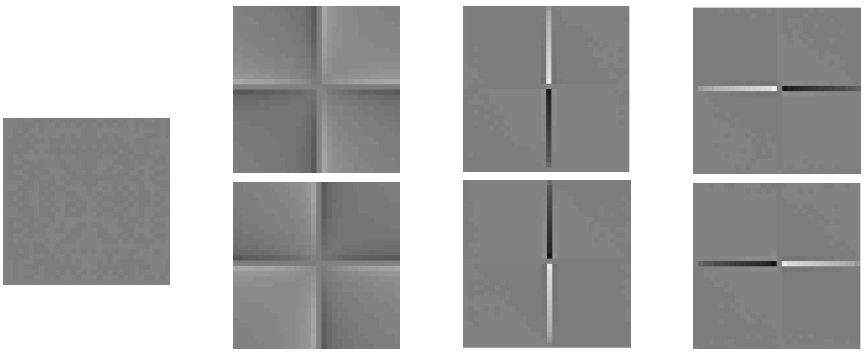


Figure 3: The mean (left image) and first three (from left to right) principal components (± 2 standard deviations) of imF for the PCA model based on the combined information from reF and imF.

supported by the results shown in Fig. 3, but in this case the mean of imF is almost a flat response.

3.2 Separate reF and imF PCA Model

The individual use of the reF and imF information from the Fourier domain results in two PCA models. each based on observation vectors which have 1024 dimensions (and 512 components to describe this data). To describe 25% of the data variation 5 principal components were needed and to describe 75% of the data variation close to 180 principal components were necessary. The effect of the first three principal components and the mean of reF and imF are shown in Figs 4 and 5, respectively.

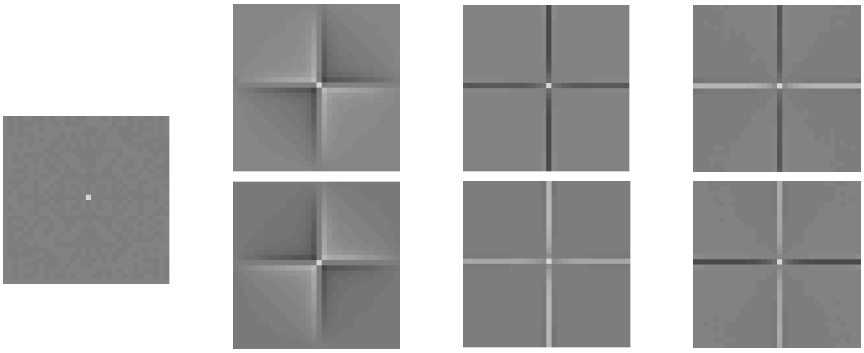


Figure 4: The mean (left image) and first three (from left to right) principal components (± 2 standard deviations) of reF for the PCA model based on the reF information.

The results in Fig. 4 indicate several effects. The first principal component indicates rotational aspects whilst the second and third principal components indicate the presence of horizontal and vertical structures in the image. The mean of reF is almost a delta

function. Again, these effects are supported by the results shown in Fig. 5 and also in this case the mean of imF is almost a flat response.

It must be pointed out that there are subtle differences between the modelling based on the combined and separated reF and imF parts of the Fourier space information.

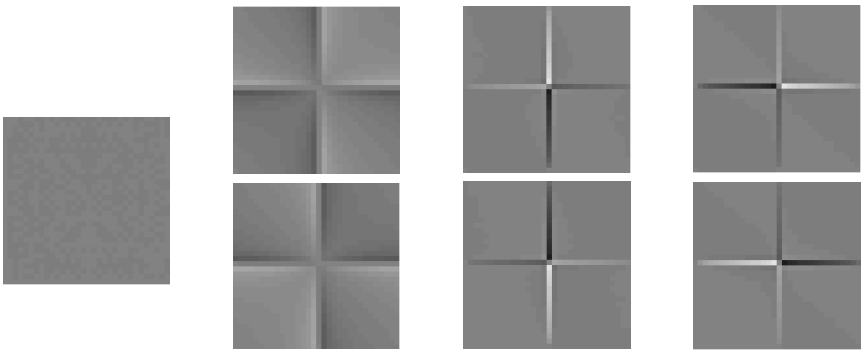
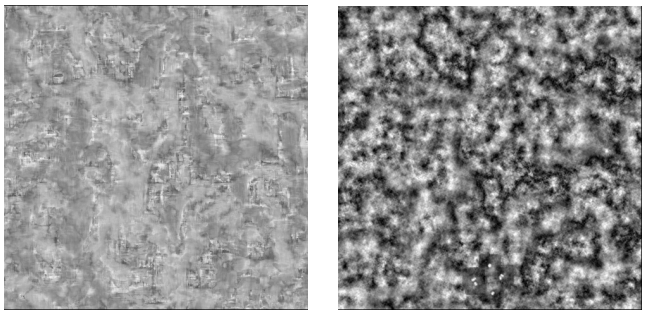


Figure 5: The mean (left image) and first three (from left to right) principal components (± 2 standard deviations) of imF for the PCA model based on the imF information.

3.3 Pixel Classification

The models described in the previous sections were used to obtain a pixel classification for images containing *abnormal* structures (see Fig. 1a). Reconstruction of reF and imF was based on 25% or 75% of the data variation and reconstruction of the image shown in Fig. 1a are shown in Fig. 6.



(a) (b)

Figure 6: Reconstructed images (see Fig. 1a) based on (a) 25% and (b) 75% of the data variation with the PCA modelling based on the individual aspects of reF and imF.

The reconstructed image, which should only contain the *normal* texture background (this is the case for Fig. 6a, but some *abnormal* structure is left in Fig. 6b) can be subtracted from the original image. The resulting image should only contain the *abnormal* structures. If some of the structure is still present in the reconstructed image the resulting

difference image can have concave areas were the abnormal structures were present. Receiver operating characteristic curves can be used to obtain a direct comparison between the original and the resulting *normal* images with the assumption that the reconstructed *normal* image should contain less (i.e. harder to detect) of the *abnormal* structures. This is done by a simple threshold applied at various intensities. The results for individual PCA models based on reF and imF of such an approach are shown in Fig. 7.

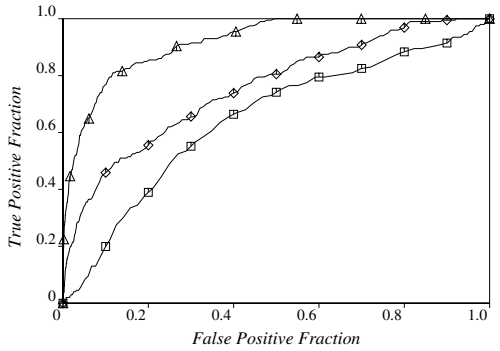


Figure 7: ROC results for the detection of the abnormal structures where: $\triangle$: original image, $\diamond$: reconstruction based on 75% PCA modelling, and $\square$: reconstruction based on 25% PCA modelling.

The results shown in Fig. 7 indicate that the reconstruction might be beneficial for texture classification purposes as some of the *abnormal* structure has been removed and is more difficult to detect. Decreasing the number of principal components has the effect of making it more difficult to detect the *abnormal* structures.

When comparing the two approaches to building PCA models it is clear that the overall description of the resulting models are very similar, with the first few principal components describing orientation effects and the presence of horizontal and vertical structures in the images. The ROC results are as well virtually independent of the modelling approach taken. From this point of view it might be beneficial to use the modelling based on the individual information contained in reF and imF, as this results in a significant reduction in the size of the model ($(2N)^2 \Rightarrow 2(N)^2$).

4 Mammographic Images

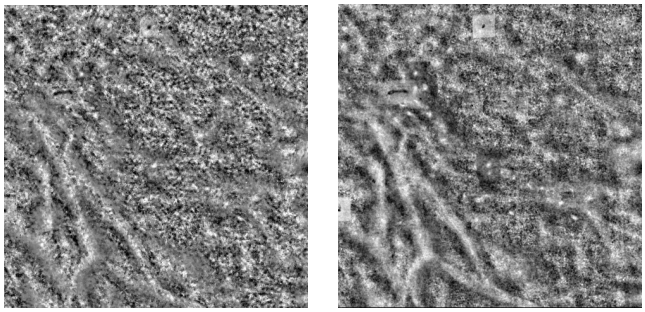
One of the application areas within mammographic screening would be the pre-processing of mammographic images with the purpose of separating out the abnormal structures from the normal texture background. In Figure 1b a section of a mammographic image is shown which contains a cluster of micro-calcifications. In the remainder of this section we will endeavour to separate out the micro-calcifications from the *normal* background (although care has to be taken as the surrounding tissue might be affected by the presence of abnormalities and no clear separation might be possible). Based on the results presented in Sec. 3 we will only use PCA modelling based on the individual information contained in reF and imF.

4.1 PCA Modelling

Two separate PCA models were build for the reF and imF information based on mammographic images containing no *abnormal* structures (although it must be emphasized that some *normal* structures, which can be very similar in appearance to abnormal structures, were present). The resulting PCA models showed similar effects as were shown in Sec. 3 for the synthetic images, with the first few principal components covering rotational and horizontal/vertical predominance aspects. One of the differences between the synthetic and mammographic modelling was the number of principal components needed to describe a certain percentage of the data variation. For the mammographic data to describe 25% of the data variation 50 principal components were needed and about 300 principal components were needed to describe 75% of the data variation. This indicates a less even distribution in Fourier space as was the case with the synthetic examples.

4.2 Pixel Classification

The PCA models can be applied to mammographic images containing abnormal structures. (as shown in Fig. 1b). The reconstructed images resulting from such an approach based on 25% and 75% of the data variation are shown in Fig. 8.



(a) (b)

Figure 8: Reconstructed images (see Fig. 1b) based on (a) 25% and (b) 75% of the data variation with the PCA modelling based on the individual aspects of reF and imF.

These images show that some of the *abnormal* structures present have not been removed. This is a large difference with the synthetic experiments described in Sec. 3. However, it can be explained by the fact that the *normal* images in the synthetic experiments did not contain any *abnormal* structures, whereas for the mammographic experiments there are structures in the *normal* images used for building the PCA model that are very similar to the *abnormal* structures we are trying to remove. The failure to remove the *abnormal* structures is confirmed with an ROC analysis which is shown in Fig. 9. This even shows that the described approach can be used as a pre-processing step to enhance the *abnormal* structures, exactly the opposite of what was expected (although it must be mentioned that the effects are not very large).

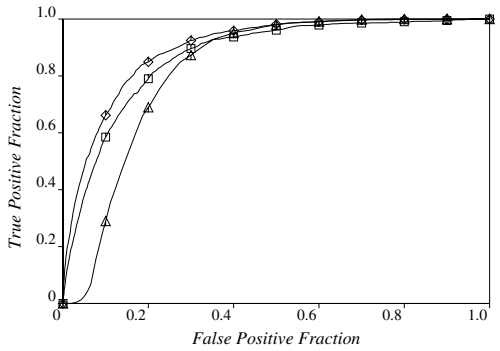


Figure 9: ROC results for the detection of the micro-calcifications where: △: original image, ◇: reconstruction based on 75% PCA modelling, and □: reconstruction based on 25% PCA modelling.

5 Discussion and Conclusions

We have shown that in principle the described approach can be used to separate *abnormal* structures and *normal* background texture in images. However, these positive results were based on synthetic data with far less convincing results based on real mammographic data. The main reason for this would be the presence of *normal* structure in the training images which were very similar to the *abnormal* structures we were trying to remove. This will need further investigation. For the mammographic data reducing the number of principal components to less than 50 (covering less than 25% of the data variation) might be an option.

For mammographic risk assessment the resulting reconstructed images can be used and a comparison with texture classification based on the original images will be investigated.

Care has to be taken because the statistical modelling applied here uses a one dimensional ($1D$) vector to describe two dimensional ($2D$) information. As soon as the $2D$ information is transformed to a $1D$ vector the spatial relations which are clear and closely connected in $2D$ space might be less well-defined and no longer connected in the $1D$ case. These non-linear aspects are currently under investigation for local pixel signature descriptions [3].

One of the other drawbacks of the proposed method is that to obtain a robust local description in the Fourier domain the size of the local regions of interest have to be sufficiently large. However, an increase in the size of the region of interest will result in an increase in the size of the PCA model (effectively the size of the covariance matrix $\mathbf{C}_{\omega_j}$). This means that it becomes problematic (with respect to memory aspects) when a region of interest size of larger than 32×32 is needed. Some enlargement of this region might be possible depending on how the reF and imF are used in the PCA model. We are also investigating novel methods to obtain a more compact description to start of with so larger regions of interest can be used for texture classification purposes. The reduction of the data-volume is one of the main priorities and we are investigating methods which will reduce the amount of data used for modelling without a major degradation of the

information that is described by the data.

A more compact model could have been obtained when the rotational symmetry of the Fourier domain had been taken into account. However, this would only reduce the data-volume by a factor of four (or increase the size of the regions of interest by a factor of two).

An alternative to the modelling of the data which has not been covered in the current document is the option to extract the amplitude and phase information from the Fourier transform information. This might be a more robust route and does fit in more appropriate with the theory suggested in Sec. 2.1.

In summary, we have described initial results for a method which shows the potential of separating out the structures and texture background present in images. The novel approach showed promising results based on synthetic images, but results based on mammographic data was less convincing. Additional work will be needed to fully investigate these aspects. A method like this could be used to improve the robustness of texture classification for risk-assessment in mammographic images.

References

[1] R.N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1986.

[2] J.O. Eklundh. On the use of fourier phase features for texture discrimination. *Computer Graphics and Image Processing*, 9:199–201, 1979.

[3] A.S. Holmes and C.J. Taylor. Improving euclidean space metric properties of scale-orientation signatures for use in computer-aided mammography. 3^{rd} *Conference on Medical Image Understanding and Analysis*, Oxford, UK:accepted, 1999.

[4] I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.

[5] N. Karssemeijer. Adaptive noise equalization and recognition of microcalcification clusters in mammograms. In K.W. Bowyer and S.M. Astley, editors, *Sate of the Art in Digital Mammographic Image Analysis*, number 9 in Machine Perception and Artificial Intelligence, pages 148–166, Singapore, 1994. World Scientific.

[6] N. Karssemeijer. Automated classification of parenchymal patterns in mammograms. *Phys. Med. Biol.*, 43:365–378, 1998.

[7] H.O. Peitgen and D. Saupe. *The Scienec of Fractal Images*. Springer Verlag, New York, 1988.

[8] G.M. te Brake and N. Karssemeijer. Comparison of three mass detection methods. 4^{th} *International Workshop on Digital Mammography*, Nijmegen, The Netherlands:119–126, 1998.

[9] J.N Wolfe. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, 37(5):2486–2492, 1976.

[10] R. Zwiggelaar, T.C. Parr, J.E. Schumm, I.W. Hutt, S.M. Astley, C.J. Taylor, and C.R.M. Boggis. Model-based detection of spiculated lesions in mammograms. *Medical Image Analysis*, 3(1):39–62, 1999.

[11] R. Zwiggelaar, T.C. Parr, and C.J. Taylor. Finding orientated line patterns in digital mammographic images. In *Proceedings of the* 7^{th} *British Machine Vision Conference*, pages 715–724, Edinburgh, UK, 1996.