Chapter 4

A Review of Image Processing and Analysis Techniques

This chapter surveys the range of image processing and analysis techniques that are available in the literature to study SAXS patterns. These techniques involve the disciplines of physics, statistics, mathematics and computing science. A great number of such techniques exist and entire journals (such as the *IEEE Transactions on Image Processing*) are dedicated to this task.

Only the major, more well-established models and techniques will be reviewed. Section 4.1 relates the ideal physical model of the SAXS pattern to the actual observed digital image data. Section 4.2 examines the Fourier, scale-space, fractal, random field, Bayes'ian, partial differential equation models, variational, mathematical morphology and shape analysis methods of image analysis. The final section, Section 4.3 reviews the available techniques and identifies useful concepts in the development of diagnostic models of SAXS patterns.

4.1   The Sampling Theorem & SAXS Patterns

The theory of this section demonstrates that the data is in the form of a digital image rather than a 'SAXS pattern' and that this image can reconstruct the pattern up to a finite resolution. The term 'SAXS pattern' is used in a very general sense in this thesis and in this section the nature of the data to be analysed is clarified. The data collected is in fact not a SAXS pattern but a digital image of a SAXS pattern that is projected onto the surface of a detector. A SAXS pattern is a dynamic electro-magentic field that exists in physical space and time and a detector was used to capture a representation of this pattern in a digital format. Previous authors such as Butler *et al* (2003) and Erickson (2005) have applied image analysis techniques under the assumption that the data are in the format of a digital image. The implications of this issue must be addressed in order to better understand the nature of the data that we wish to model.

Recall, in equation (2.1) that the SAXS pattern refers to a physical quantity,

$$I(\mathbf{h}) = \mathcal{F}^{\star}(\mathbf{h})\mathcal{F}(\mathbf{h}) \tag{4.1}$$

that is continuous and that exists in a physical 4-dimensional space (3 space, 1 time). We define a gray-scale digital image as a matrix $\mathbf{I}(x,y)$ indexed by location parameters (x,y), $1 \leq x \leq X_{\max}$; $1 \leq y \leq Y_{\max}$ and having an intensity $z = I(x,y)$, $z \, \epsilon \, \mathbb{R}$. Digital detectors were used in the data collection process to capture the physical SAXS pattern as a digital image (Lewis *et al* 1997a, 1997b). The pattern is recorded using the detector over a period of time, T, to reduce the impact of detector noise on the final image and to allow sufficient signal to be recorded. During this interval, the structure of the tissue and hence the form of the SAXS pattern is assumed not to change form significantly. Henceforth the data (digital image) captured by the detector in this process will be referred to as the 'SAXS image' and the actual electro-magnetic field that exists in physical space as the 'SAXS pattern'. The sampling theorem provides insight into the subtle differences between the 'SAXS image' and the 'SAXS pattern' and how these differences influence the subsequent analysis.

4.1.1 The Sampling Theorem

Shannon (1949) introduced the *sampling theorem* in the context of information theory, but as noted by Jerri (1977) significant contributions or alternative theorems were made by Nyquist (1928), ET Whittaker (1915), JM Whittaker (1929; 1935), Ferrar (1927) and Kotel'nikov (1933). The sampling theorem $\big($for a one-dimensional continuous time signal $f(t)$ $\big)$ states that:

"*If a function $f(t)$ contains no frequencies higher than W cps it is completely determined by giving its ordinates at a series of points spaced (1/2W)s apart*" where "cps" stands for counts per second (Theorem I-A-1, Jerri 1977) and the sampling rate, $2W$, is called the *Nyquist rate* (page 82, Mix & Olejniczak 2003; Nyquist 1928).

In this thesis, interest lies in a multivariate generalisation of the sampling theorem which is useful for image analysis. For a function $f(t_1, t_2)$ of two real variables, (t_1, t_2) whose two-dimensional Fourier integral $\mathcal{F}(y_1, y_2)$ exists that is identically zero outside a two-dimensional rectangle and is symmetrical about the origin, the sampling theorem can be specified as ,

$$
\begin{aligned}
\mathcal{F}(y_1, y_2) &= 0, \ \ |y_k| > |\omega_k|, \ k = (1, 2) \qquad \text{then,} \\
f(t_1, t_2) &= \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} f\big(\tfrac{\pi m_1}{\omega_1}, \tfrac{\pi m_2}{\omega_2}\big) \tfrac{\sin(\omega_1 t_1 - m_1 \pi)}{\omega_1 t_1 - m_1 \pi} \tfrac{\sin(\omega_2 t_2 - m_2 \pi)}{\omega_2 t_2 - m_2 \pi}
\end{aligned}
\tag{4.2}
$$

where $|\omega_k|$ is 2π times the frequency of the signal (adapted from Theorem IV-A-1, Jerri 1977; Rezza 1961).

The sampling theorem suggests that the digital SAXS image can be used to reconstruct a continuous function that describes the SAXS pattern, provided the detector elements are spaced sufficiently close (in physical space) to adequately sample the pattern. In practice the detector elements can be placed finitely close together, so the imaging system and hence the reconstruction has a finite resolution. The sampling theorem applies to band-limited images (a finite range of spatial frequencies), but in general, SAXS patterns are not band-limited. One solution to this problem is to interpret the function $f(t_1, t_2)$ as a perfect representation of the intensity of the SAXS pattern in the region of physical space that is captured by the detector. This function is band-limited $\big(\mathcal{F}(y_1, y_2) = 0, \ \ |y_1, y_2| > |\omega_1, \omega_2|\big)$ outside a certain range of spatial frequencies. Therefore the observed SAXS image corresponds to a sampled version of the band-limited function $f(t_1, t_2)$ that is a resolution-limited representation of the actual SAXS pattern. Throughout this thesis, this concept must be kept in the back of the readers mind when the term 'SAXS image' is used. Brillouin (1963) interpreted x-ray diffraction patterns to the *periodic* physical structure of the specimen (the electron density distribution, $\rho(\mathbf{r})$) and the observed image $I(x, y)$ using the sampling theorem. This interpretation focuses on the determination of the physical

structure, $\rho(\mathbf{r})$ of the specimen and is not as directly concerned with the statistical analysis of the data in terms of an image. For this reason the Brillouins' interpretation of x-ray diffraction patterns and hence SAXS images in terms of the sampling theorem is not considered in further depth in this thesis.

4.2   A Survey of Image Models That Are Available for SAXS Image Analysis

SAXS images of breast tissue structure have been extensively analysed in terms of parameters that describe a physical model of collagen and have also been analysed using image processing techniques (Lewis *et al* 2000; Round 2006; Sidhu *et al* 2008; Butler *et al* 2003; Erickson 2005; Falzon *et al* 2006). The advent of rapid micro-processors has allowed the regular implementation of a diverse range of these techniques and a wide variety of approaches are available that researchers may use to rigorously analyse SAXS images.

Major image analysis techniques include:

  a) Fourier analysis

  b) Wavelets & scale/space models

  c) Fractal analysis

  d) Random fields

  e) Bayes'ian statistical models

  f) Partial differential equation models

  g) Variational methods

  h) Shape analysis

  i) Mathematical morphology

  j) Fuzzy analysis

The focus of this review is to briefly introduce, examine and discuss the benefits and disadvantages of each technique as applied to SAXS images. This review is not a comprehensive examination of every single area of image analysis, nor should it be, it is a guide to those concepts in image processing/analysis (which is not restricted to SAXS images and encompassing other types of images) that the author has considered in depth in the construction of the SAXS diagnostic model. The reader must be aware that the purpose of this section is limited to surveying and selecting those modeling approaches worthy of further investigation.
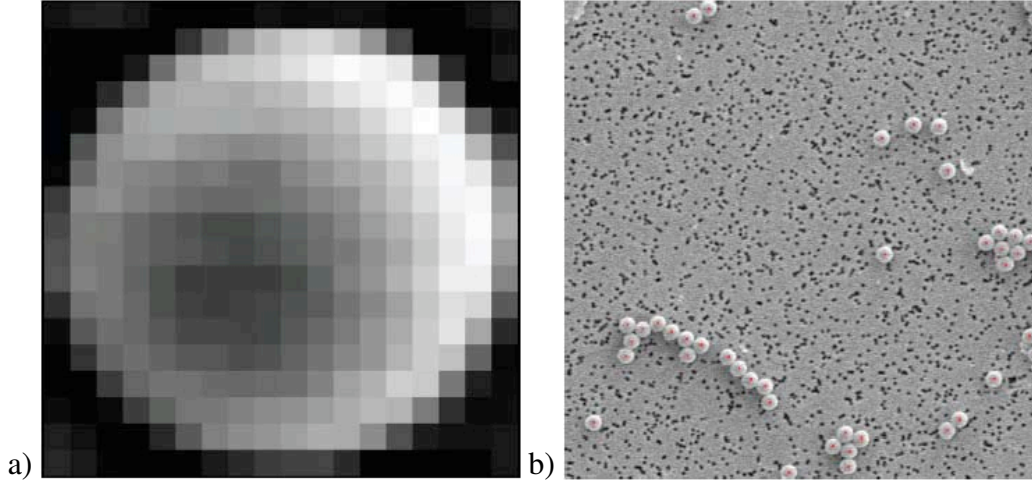
4.2.1   Fourier Analysis

Figure 4.1: Pattern recognition of latex spheres in scanning electron microscope images using the Fourier transform: (a) A reference image $R(x, y)$ that describes the latex spheres (b) successful identification of latex spheres (marked red) on the surface of a scanning electron microscope query image $Q(x, y)$ (from Figure 6.67 Russ 1992).

Fourier analysis of a digital image is performed using the discrete Fourier transform:

$$F(k, l) = \frac{1}{N_1 N_2} \sum_{y=0}^{(N_2-1)} \sum_{x=0}^{(N_1-1)} I(x, y) \exp\left[-i2\pi\left(\frac{ky}{N_1} + \frac{lx}{N_2}\right)\right] \qquad (4.3)$$

for a digital image $I(x, y)$ of size N_1 x N_2 pixels indexed by location parameters (x, y) where $x = 0, \ldots, N_1 - 1$ and $y = 0, \ldots, N_2 - 1$. The discrete Fourier transform $F(k, l)$ is indexed by parameters (k, l) where $k = 0, \ldots, N_1 - 1$ and $l = 0, \ldots, N_2 - 1$ that describe spatial frequency in cycles/pixel (equation 2.42, page 100, Rangayyan 2005).

The discrete Fourier transform has been used for a wide range of image processing and analysis tasks including:

i) Automatic alignment (registration) of pairs of images (DeCastro & Morandi 1987; Chen *et al* 1994; Stone *et al* 2001).

ii) Image filtering (Sherlock *et al* 1994; Gonzalez & Woods 2007).

iii) Deblurring (deconvolution) and restoration of images (Javidi 1990; Cunningham & Anthony 1993; Bertero & Boccaci 1998).

iv) Edge detection (Nixon & Aguado 2002; Rangayyan 2005; Engelberg 2008).

v) Fractal dimension measurement (Feder 1988; Vicsek 1989; Family & Vicsek 1991).

vi) Pattern recognition and template matching (Gardenier, McCallum & Bates 1986; Sheng & Shen 1994; Fonga 1996; Javidi & Horner 1998).

Of these many applications, pattern recognition is of particular interest for the diagnosis of breast cancer. One method that achieves pattern recognition is performed by comparing a reference image $R(x, y)$ with a query image $Q(x, y)$ in terms of image correlation:

$$\gamma_{R(x,y),Q(x,y)}(\alpha, \beta) = \sum_{y=0}^{N_2-1} \sum_{x=0}^{N_1-1} R(x, y)Q(x + \alpha, y + \beta) \tag{4.4}$$

for shift parameters $(\alpha, \beta) : \alpha \; \epsilon \; \mathbb{R}, \; \beta \; \epsilon \; \mathbb{R}$ (adapted from equation 2.72, page 118, Rangayyan 2005). Image correlation can be rapidly calculated in the Fourier domain using the convolution theorem (equation 2.70, page 118 and equation 2.73, page 118 Rangayyan 2005; equation 10.13, page 447, Boas 2006).

Russ (1992) describes the application of the cross-correlation technique on scanning electron microscope images of nucleophore filters with latex spheres. In this example, automated processing of large numbers of scanning electron microscope images was desired as was the accurate counting of the large number of latex spheres on the filter surface. An average of ten latex spheres was used to produce the reference image ($R(x, y)$ in equation 4.4) as displayed in Figure 4.2.1(a) which was correlated with the query scanning electron microscope image $Q(x, y)$ using Fourier techniques. The cross-correlation technique was able to successfully detect the presence and location of the latex particles on the surface of the nucleophore filter as seen in

the results presented in Figure 4.2.1(b). Cross-correlation could also be used to recognise SAXS images of normal, benign and malignant breast tissue. Reference images must be specified for each tissue group using a training data set with future (query) images classified according to the group with the greatest correlation.

Correlation statistic based pattern recognition is limited because it does not provide explicit probability estimates of the query image belonging to each group. A correlation statistic alone does not take into account the variability of the sample measurements about a typical value and there is a need to combine the cross-correlation approach with a statistical model. Furthermore, a high correlation measurement does not necessarily equate to a query image belonging to a particular group. Methods based upon cross-correlation consider only the second-order statistical properties of the intensity values in the image but higher-order relationships might convey important diagnostic information. Given these limitations a superior technique/model is sought in this thesis.

4.2.2 Wavelet Analysis & other scale/space methods

There are two main varieties of wavelet transform, the continuous wavelet transformation and the discrete wavelet transformation. The continuous wavelet transformation of an image, $I(x, y) \epsilon L^2(\mathbb{R}^2)$ of finite energy is:

$$\mathcal{CWT}(I(x,y)|\psi_{(b_1,b_2,\theta,a)})(x,y) = C_\psi^{-1/2} < \psi_{b_1,b_2,\theta,a}(x,y), I(x,y) > \qquad (4.5)$$

where C_ψ is a normalisation constant that is less then infinity, $\psi_{b_1,b_2,\theta,a}(x,y)$ is the analysing wavelet function, b_1 and b_2 are translation parameters, $\theta \epsilon [0, 2\pi)$ is the rotation angle and $a \epsilon \mathbb{R}^+$ is the scale parameter (adapted from equation 3, page 569, Arnéodo, Decoster & Roux 2000).

In contrast, the discrete wavelet transform is defined as,

$$
\begin{aligned}
I(x,y) &= \sum_{k_2=0}^{N_{j_0,2}} \sum_{k_1=0}^{N_{j_0,1}} c_{j_0,k_{1,j_0},k_{2,j_0}} \phi_{j_0,k_{1,j_0},k_{2,j_0}}(x,y) \\
&+ \sum_{i=1}^{3} \sum_{j \geq j_0}^{J} \sum_{k_{2,j}=0}^{N_{2,j}} \sum_{k_{1,j}=0}^{N_{1,j}} d_{j,k_{1,j},k_{2,j}}^{i} \psi_{j,k_{1,j},k_{2,j}}^{i}(x,y) \qquad (4.6)
\end{aligned}
$$

where $\phi_{j_0,k_{1,j_0},k_{2,j_0}}(x,y)$ is a low-pass filter known as the *scaling function* and the functions $\psi_{j,k_{1,j},k_{2,j}}^{i}(x,y)$ are high-pass filters known as the *wavelet functions*. The index, i indicates the particular type of wavelet function ($i = 1$: vertical, $i = 2$: horizontal and $i = 3$: diagonal), parameter j denotes scale with parameter j_0 indicating the lowest-resolution scale and parameter J the highest resolution so that $j \epsilon (j_0, \ldots, J)$ with $j \epsilon \mathbb{N}$. Shifts in the both the scaling and wavelet functions are described by translation parameters $\mathbf{k}_j = (k_{1,j}, k_{2,j})$. The *approximation* coefficients $c_{j_0,k_{1,j_0},k_{2,j_0}} = < I(x,y), \phi_{j_0,k_{1,j_0},k_{2,j_0}}(x,y) >$ describe the match between the image and the scaling function and the *detail* coefficients $d_{j,k_{1,j},k_{2,j}}^{i} = < I(x,y), \psi_{j,k_{1,j},k_{2,j}}^{i}(x,y) >$ describe the match between the wavelet functions and the image (adapted from page 157, Vidakovic 1999).

Key components in both transforms are the scale parameters (a and j for the continuous and discrete wavelet transformations respectively) which allow the examination of the image at different scales/resolutions and the location parameters ((b_1, b_2) and $\mathbf{k}_j$ respectively) that allow a *local* examination of the image. The analysis is performed with the wavelet basis function $\psi(x,y)$, which has the important property of integrating to zero:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x,y) \, dx \, dy = 0. \qquad (4.7)$$

Integration to zero implies that wavelet function $\psi(x, y)$ oscillates about a reference axis, hence the name '*wavelet*' (adapted from equation 7, page 570 Arnéodo, Decoster & Roux 2000).

Examples of wavelet functions include the isotropic Mexican hat wavelet,

$$\psi(x, y) = -\nabla \{ \exp [-\frac{1}{2}(x^2 + y^2)] \} \tag{4.8}$$

(equation 3.6, page 99, Antoine *et al* 2004; Antoine *et al* 1993) and the directional Morlet wavelet,

$$\hat{\psi}(K_x, K_y) = \sqrt{\epsilon} \exp -\frac{1}{2} [\epsilon K_x^2 + (K_y - K_0)^2] \tag{4.9}$$

where (K_x, K_y) are the frequency indices, K_0 is a reference frequency and $\epsilon \geq 1$ is an anisotropy parameter describing the shape of the wavelet function $\hat{\psi}(K_x, K_y)$ in frequency space (equation 3.9, page 16, Antoine & Murenzi 1996). Other notable wavelet functions include the Haar (Haar 1910), the Daubechies compactly supported family of wavelet basis functions (Daubechies 1992) and the biorthogonal wavelet bases (Cohen, Daubechies & Feauveau 1992) among many others.

Both the continuous and discrete wavelet transforms have had numerous applications in image processing and analysis, including:

a) compression (DeVore, Jawerth & Lucier 1992; Lewis & Knowles 1992; Antoni *et al* 1992).

b) deconvolution (Kalifa, Mallat & Rouge 1998; 2003; Stark & Bijaoui 1994; Neelamani, Choi & Baraniuk 2004).

c) edge detection (Mallat & Hwang 1992; Mallat & Zhang 1992; Zhang & Bao 2002).

d) object recognition (Oren *et al* 1997; Boles & Boashash 1998).

e) removal of noise (Kingsbury 1999; Michak *et al* 1999; Chang, Yu & Vertelli 2000; Portilla *et al* 2003).

f) watermarking: embedding information in the image (Kundur & Hatzinakos 1998; Ruanaidh, Dowling & Boland 1996; Dugad, Ratakonda & Ahuja 1998).

The wavelet transform has previously been used by Erickson (2005) and Falzon *et al* (2006) to analyse SAXS images of breast tissue. Despite the limitations of these studies (which were identified in Chapter 3), the wavelet transform allowed the examination of the SAXS images

across scales and provided accurate identification of tissue pathology. The precedent has been set by these two studies and the wavelet technique is clearly a versatile approach to the problem.

Multi-scale modeling of the SAXS images matches well with the hierarchical structure of the collagen that produced the features in these images. The research of Candes & Donoho (2003) revealed that both the continuous and discrete wavelet transforms do not provide the most efficient representation of images with geometric structures (such as scattering rings in SAXS images). A parsimonious approach to image transformation that captures the image data with relatively few coefficients of larger magnitude might be very useful in the estimation of a statistical model when there is limited data. Better model estimation combined with fewer parameters might be achieved with second-generation transforms such as curvelets (Starck, Candes & Donoho 2002; Candes & Donoho 2003; Candes *et al* 2006), contourlets (Do & Vertelii 2002; 2003; 2006) and the steerable pyramid (Simoncelli & Freeman 1995; Portilla & Simoncelli 2000) which could translate into better classification results.
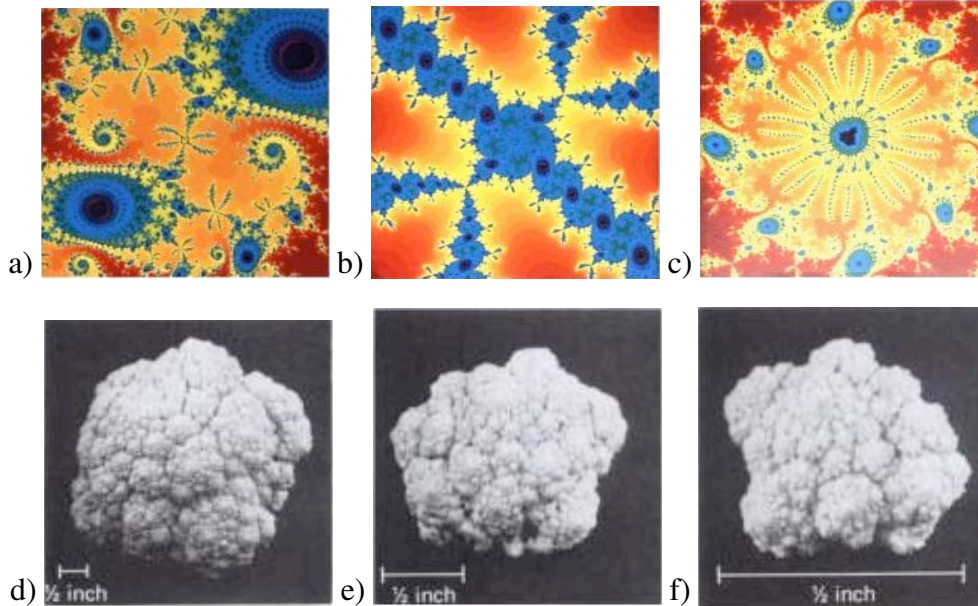
4.2.3 Fractal Analysis

Figure 4.2: Deterministic and Quasi-fractals: The Mandelbrot set, $z \rightarrow z^2 + c$, $z \epsilon \mathbb{C}$, $c \epsilon \mathbb{R}$ is an example of a deterministic fractal which characteristic shapes that are repeated when zooming into the centres of the images from (a) low to (b)-(c) increasingly finer scales. The head of a cauliflower has a quasi-fractal structure and appears self-similar across scales (d)-(f) (from colour plates 4-6, Branner 1988; Figure 2.1 Peitgen *et al* 1992).

The term 'fractal' is used to describe irregular structure in sets that are not well described by Euclidean geometry. Mandelbrot (2006) defines a fractal as"*a set for which the Hausdorff Besicovitch dimension strictly exceeds the topological dimension*" (page 15, Mandelbrot 2006). The topological and the Hausdorff Besicovitch dimensions measure different characteristics of the sets and a detailed treatment of the topic can be found on page 198, Barnsley *et al* (1993). Key characteristics of fractals include:

  (i) an infinite length

  (ii) self-similar behaviour across scales. That is, N the number of points of the set at scale, a, obeys the power-law, $N = a^d$, $d \epsilon \mathbb{R}^+$.

(Mandelbrot 1982).

A number of fractals or fractal-like objects have been described in the literature or found in nature. Deterministic fractals are those fractals that are fully specified by a mathematical relationship and include for instance the Mandelbrot set, $z \rightarrow z^2 + c,\ z \epsilon \mathbb{C},\ c \epsilon \mathbb{R}$, whose characteristic shape repeats across scales (Mandelbrot 1982; Branner 1988). The Mandelbrot set is displayed in a series of successive magnifications in Figure 4.2(a)-(c) over the region centered on $c \approx -0.124422584 + 0.839099345\Im$. Quasi-fractals are often found in nature such as in the structure of a cauliflower head. As seen in Figure 4.2(d)-(f) the form of the cauliflower head appears similar across different scales of magnification. Finally, sets exist that are self-similar in their statistical properties across scales.

Fractal research has increased substantially since Mandelbrot's (1982) work and a range of fractal techniques and concepts exist:

a) fractal analysis of binary images or shapes extracted from the images.

b) analysis of gray-scale images.

c) multi-fractal analysis.

d) lacunarity analysis.

e) fractal analysis that considers the image as a surface.

Each of the above techniques will now be considered in further detail.

4.2.4   Fractal analysis of binary images

Shapes can be extracted as outlines from objects within images or gray-scale images made into binary images by thresholding (by setting to zero those intensities below a certain magnitude and setting those remaining to one). The set of non-zero points that remains in the binary image (or outline) can be analysed using fractal techniques. In practice the Hausdorff-Besicovitch dimension is replaced by other fractal dimension estimators such as the '*box-counting*' dimension (Mandelbrot 1982; Voss 1991). This dimension is defined as,

$$D = \lim_{m \to \infty} \left\{ \frac{\ln(\mathcal{N}_m(\mathcal{A}))}{\ln(2^m)} \right\} \qquad (4.10)$$

for a subset $\mathcal{A} \epsilon \mathbb{R}^n$ using a Euclidean metric and a covering of $\mathbb{R}^n$ of closed square boxes of side length $(1/2^n)$, where $\mathcal{N}_m(\mathcal{A})$ is the number of 'non-empty' boxes of side length $(1/2^m)$ (Theorem 1.2, page 175, Barnsley 1988; Sullivan & Hunt 1988; Hunt 1990).

The box-counting technique in action is displayed in Figures 4.3(a)-(d) where a grid of square boxes is used to calculate the box-counting dimension of an image of a histopathology slide of a tumour (Cross 1997). The slope of the regression of $\ln(\mathcal{N}_m(\mathcal{A}))$ against $\ln(2^m)$ is used to estimate D. Hall & Wood (1983) showed that the estimation of D has unacceptably high bias which is undesirable for those diagnostic applications where differences in the box-counting dimension are small. The binary fractal image analysis techniques require the thresholding of a gray-scale image, for the classification of SAXS images this operation results in a loss of information which is deemed unacceptable.
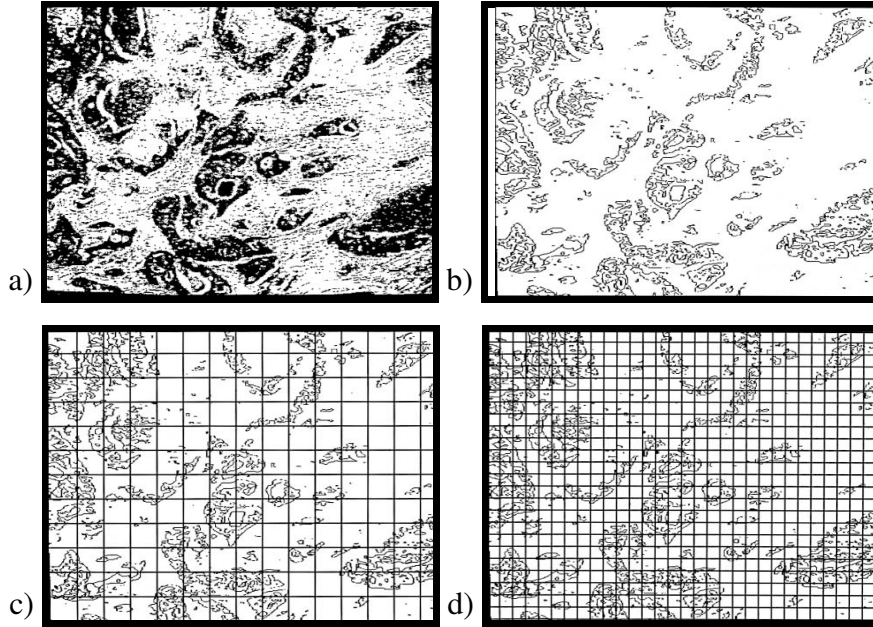

Figure 4.3: Calculating the box-counting dimension of an image of a histopathology slide: (a) the binary image of a histopathology slide of a tumour, (b) is converted to a single pixel outline using edge detection methods, (c) and (d) is partitioned using grids of different box side lengths $(1/2^m)$ and the number of squares containing the outline $\mathcal{N}_m(\mathcal{A})$ is counted in each case. The box-counting dimension is found by the slope of the regression of $\ln(\mathcal{N}_m(\mathcal{A}))$ against $\ln(2^m)$ for a large number of different box sizes m (Figure 2, Cross 1997).

4.2.5 Gray scale fractal image analysis

The fractal box-counting technique has been extended to incorporate the analysis of gray-scale images (Lundahl *et al* 1985). The intensity is treated as a third-dimension and the object is analysed as a surface. The surface area $\mathcal{A}(2^m)$ for a box of side length $(1/2^m)$ is calculated as,

$$
\begin{aligned}
\mathcal{A}(2^m) &= \sum_{y=1}^{N_2}\sum_{x=1}^{N_1}(2^m)^2 + \sum_{y=1}^{N_2}\sum_{x=1}^{N_1}(2^m)^2(|I_{2^m}(x,y) - I_{2^m}(x,y+1)| \\
&- |I_{2^m}(x,y) - I_{2^m}(x+1,y)|)
\end{aligned}
\tag{4.11}
$$

where $I_{2^m}(x,y)$ denotes the analysis of the intensity image with boxes of side length $(1/2^m)$ centered at position (x,y) (equation 8, page 390, Yaffe *et al* 2000). Regression of $\ln(\mathcal{A}(2^m))$ against $\ln(2^m)$ for a range of box sizes, m, allows the estimation of the box-counting dimension, D of the gray-scale image. This method appears to be far more suitable than the binary fractal image analysis method for the analysis of SAXS images. Nonetheless, the gray-scale fractal image analysis methods are themselves limited in that they calculate only one characteristic dimension when in fact a range of scaling relationships may exist in the image.

4.2.6 Multi-fractal image analysis

The box-counting algorithm makes no provision for the number of points located inside each box (Theiler 1990). To account for this problem, a weighting algorithm was introduced along with the generalised dimensions, D_q that describe the scaling of the of the 'bulk' of a set (the fractal) with respect to scale (box size) (Theiler 1990).

The generalised dimensions are defined as,

$$
D_q = \frac{1}{q-1}\lim_{\epsilon \to 0}\frac{\log \sum_{i=1}^{N} p_i^q}{\log \epsilon}
\tag{4.12}
$$

where p_i is the probability for a randomly chosen point on the fractal is in the ball $\mathcal{B}_i$, this probability is estimated by counting the number of points in the ith box and by dividing by the number of points in total. The radius of the ball is given by $\epsilon > 0$ and the moment index is $q \ \epsilon \ \mathbb{R}$ (equation 24, page 1060, Theiler 1990).

Plots of D_q against q reveal the nature of the fractal, with those fractals that have a response that varies with the moment index q referred to as *multi-fractals*. The generalised dimensions, D_q of multi-fractals generally decrease with increasing q. The most negative value of D_q correspond with the least-dense points on the fractal and the most positive values of D_q with the most dense
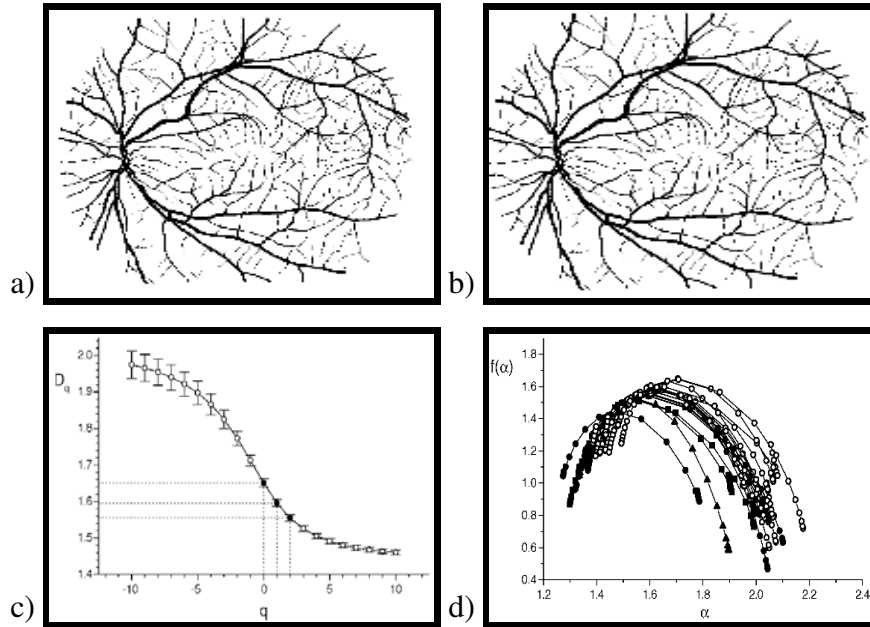
Figure 4.4: Multi-fractal analysis of binary images of the vascular structure of the human retina: (a) image of a normal healthy retina, (b) image of a retina from a patient with background diabetic retinopathy (c) the $D(q)$ curve which indicates the multi-fractal structure of these images and (d) the $f(\alpha)$ curve has a shift to lower values and a decreased range for pathological cases (marked with a full shade symbol) as compared to normal healthy cases (marked with an open symbol) (Figures 1,2,4 Stošić & Stošić 2006).

set of points (Theiler 1990). The generalised dimensions capture a range of fractal behaviour, in particular D_0 corresponds to the box-counting dimension (Theiler 1990). Multi-fractal behaviour has been found in human heart-beat dynamics, foreign stock exchange, turbulent fluids, internet traffic, cloud, rain and river flow data (Ivanov *et al* 1999; Ausloos & Ivanova 2002; Sreenivasan 1991; Muzy, Bacry & Arnéodo 1991; Meneveau & Sreenivasan 2006; Tessier, Lovejoy & Schertzer 1993; Tessier 1993; Lovejoy & Schertzer 1989; Tessier *et al* 1996).

The multi-fractal spectrum $f(\alpha)$ can also be derived (Halsey *et al* 1986). The $f(\alpha)$ curve describes the spectrum of scaling laws that hold for a fractal set and is a single point for a homogeneous fractal (Halsey *et al* 1986). Let $\tau = (q-1)D_q$, by applying the Legendre transformation to the variables (q, τ) it is possible to find $(\alpha, f(\alpha))$ as:

$$f(\alpha) = \min_q \left\{ q\alpha - \tau(q) \right\} \tag{4.13}$$

and,

$$\tau(q) = \min_\alpha \left\{ q\alpha - f(\alpha) \right\} \tag{4.14}$$

(equation 31 & equation 32, page 1061, Theiler 1990). Together the generalised dimensions D_q and the multi-fractal spectrum $f(\alpha)$ describe the characteristics of a fractal in great detail.

Multi-fractal analysis was applied to binary images of the vascular structure of the human retinal vessel by Stošić & Stošić (2006). Figures 4.4(a) and (b) display binary images of the retinal structure for normal and pathological (background diabetic retinopathy) tissue respectively. The D_q curves in Figure 4.4(c) indicate the multi-fractal nature of the images and the $f(\alpha)$ curves of Figure 4.4(d) indicate considerable differences in the fractal characteristic of those images belonging to normal healthy samples and those belonging to pathological samples. The pathological samples produce images that have a $f(\alpha)$ curve that is shifted towards lower values and have a decreased range as compared to images of normal healthy samples. Image analysis using multi-fractal techniques appears to be a richer concept than the more standard box-counting techniques of Sections 4.2.4 and 4.2.5 and might be useful in the analysis of SAXS images to detect the existence of multiple scaling relationships in the intensity values.

4.2.7   Analysis of a digital image as a rough surface

Fractal methods developed for the study of surfaces can also be applied to gray-scale images of the form $z = I(x, y)$. Of particular interest are self-affine fractal surfaces which are those fractal surfaces that are invariant under *affine* transformations (Arnéodo *et al* 2000). An affine transformation is an anisotropic rescaling of each dimension of the surface,

$$\triangle x \rightarrow \lambda_x \triangle x, \ \triangle y \rightarrow \lambda_y \triangle y, \ \triangle z \rightarrow \lambda_z \triangle z, \tag{4.15}$$

for $\lambda_x, \lambda_y, \lambda_z \ \epsilon \ \mathbb{R}$ and horizontal distances $\triangle x, \ \triangle y$ and vertical distance $\triangle z$ (equation 27, page 572, Arnéodo, Decoster & Roux 2000). Therefore any region of a self-affine surface can be made to match any other smaller or larger region by anisotropic rescaling. Deterministic and anisotropic versions of the self-affine fractal exist. In these cases the surface $z = I(x, y)$ is described by a single-valued self-affine function $f(x, y)$ has the property that for any fixed reference point $(x_0, y_0 \ \epsilon \ \mathbb{R}^2)$ all other locations $(x, y) \ \epsilon \ \mathbb{R}^2$ in the neighbourhood there exists a parameter $H \ \epsilon \ \mathbb{R}$ called the *Hurst exponent* such that for any $\lambda > 0$ and $\alpha \ \epsilon \ \mathbb{R}$,

$$f(x_0 + \lambda x, y_0 + \lambda_y^\alpha) - f(x_0, y_0) \simeq \lambda^H [f(x_0 + x, y_0 + y) - f(x_0, y_0)] \tag{4.16}$$

(equation 29, page 572, Arnéodo, Decoster & Roux 2000; Mandelbrot 1982; Feder 1988; Vicsek 1989; Family & Vicsek 1991). The magnitude of the parameter α indicates the spatial properties of the scale invariance of surface. For instance, $\alpha = 1$ indicates that the function has *isotropic* scale invariance with respect to spatial parameters (x, y), whilst $\alpha \neq 1$ indicates *anisotropic* scale invariance (Arnéodo *et al* 2000; Schertzer & Lovejoy 1985). The Hurst exponent, H, characterises the overall global 'roughness' of the surface, with higher magnitudes indicating smoother surfaces with less intensity fluctuations.

The surface fractal technique appears to be useful in the analysis of SAXS images. These images are gray-scale and can be described as a surface. Increased levels of amorphous scatter has been reported in SAXS images of malignant breast tissue as compared to normal healthy tissue (Fernández *et al* 2002; Round 2006). This increased level of scatter corresponds to an increased level of random fluctuation in the intensity magnitude in particular regions of the image (Fernández *et al* 2002). Self-affine surface fractals might well describe these regions of the SAXS image and differences in the amorphous scatter levels might be able to be detected using the Hurst exponent.
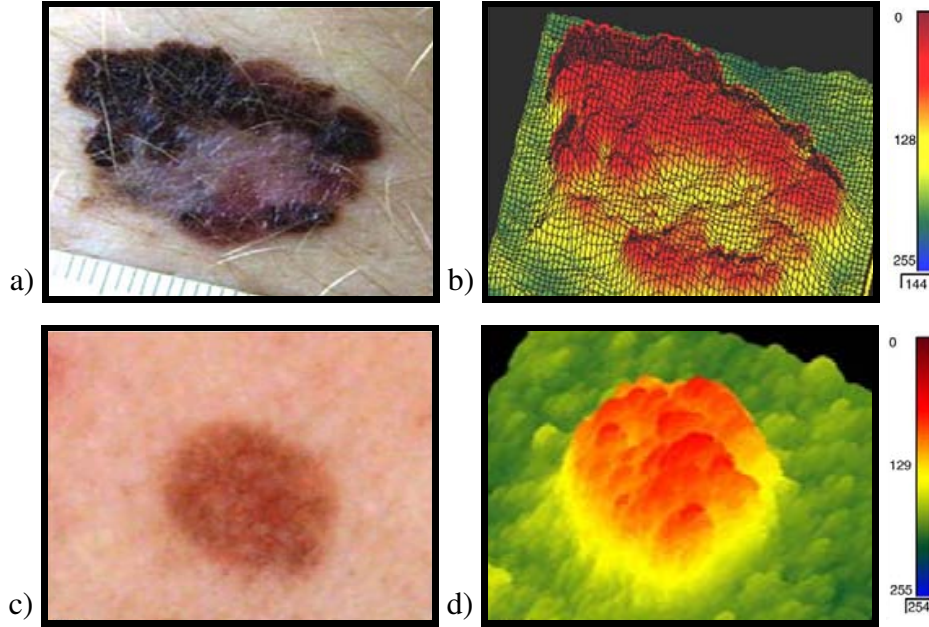
4.2.8 Lacunarity analysis of an image

Figure 4.5: Surface fractal image analysis of skin cancer: (a) image of a melanoma, (b) three-dimensional surface of the melanoma obtained from the digital image, (c) image of a dyplastic nevus (d) three-dimensional surface of the dyplastic nevus obtained from the digital image (Figures 3 and 5, Manouski *et al* 2006).

The fractal dimensions characterise only a portion of the information that may be available in an image. Fractal sets may have the same box-counting dimension, D but have a completely different visual appearance. The lacunarity measurement is a descriptor that might be capable of discerning a difference in such situations.

Denote $P(n, \epsilon)$ as the probability density function that n points from a set, ζ are contained within a z-dimensional sphere of radius ϵ which is centered on an arbitrary point in ζ (for images z = 2). The moments (about zero) of this probability density function are,

$$\mathcal{M}^q = \sum_{n=1}^{N} n^q P(n, \epsilon) \tag{4.17}$$

(equation 25, page 9 , Voss 1991) and the *lacunarity* as,

$$\Lambda(\epsilon) = \frac{< \mathcal{M}^2(\epsilon) > - < \mathcal{M}(\epsilon) >^2}{< \mathcal{M}(\epsilon) >} \tag{4.18}$$

 (equation 28, page 28, Voss 1991).

Lacunarity measures the ratio of the variance of the probability density function $P(n, \epsilon)$ to the mean at scale ϵ and hence some of the statistical properties of the spatial arrangement of the fractal. The study of lacunarity with varying scale, ϵ yields a curve $\Lambda(\epsilon)$ that conveys detailed information on the spatial arrangement of the fractal.

The lacunarity curve can be used for pattern recognition tasks such as the case of Einstein, Wu & Gil (1998) who used it to distinguish between benign and malignant breast epithelial cells. Lacunarity and surface fractal image analysis was applied to images of melanoma (malignant) and dyplastic and common melanocytic nevi (pre-cursors of melanoma) by Manouski *et al* (2006). Figures 4.5(a) and (b) display the image of a melanoma and the corresponding digital surface which are far more rugged than those of the dyplastic nevus of Figures 4.5(c) and (d). Identification of tissue pathology using the surface fractal dimension alone was difficult (melanoma $D = 2.49 \pm 0.10$ (n = 23), dyplastic nevi $D = 2.44 \pm 0.11$ (n = 44) and common melanocytic nevi $D = 2.46 \pm 0.07$ (n = 65). In contrast the measurement of lacunarity (at a fixed scale) provided a greater level of separation of the average value (melanoma $\Lambda = 0.42 \pm 0.35$, dyplastic nevi $\Lambda = 0.25 \pm 0.22$ and common melanocytic nevi $\Lambda = 0.19 \pm 0.14$) as supported by Mann-Whitney tests (Tables 1, 3 and 4, Manouski *et al* 2006).

The lacunarity analysis method can be applied to gray-scale images (such as SAXS images) and it may reveal useful diagnostic information if any major differences in the spatial arrangement of the images exist. The SAXS images of breast tissue all appear very similar when examined visually and the differences often only emerge when the intensity magnitudes are examined in detail (Falzon *et al* 2006; Sidhu *et al* 2008). Therefore, it is unlikely that the lacunarity analysis will be useful in the diagnosis of breast cancer using SAXS images and the method will not be pursued any further.

4.2.9 Random field models

Random field models provide a stochastic statistical interpretation of a digital image. In this review, we will be concerned with two major types of random field models: (i) the Gibbs random field and (ii) the Markov random field. Both models capture important information about the structure of the image.

4.2.10 Gibbs random fields

In a Gibbs random field the geometric structure and the magnitude of pixel interactions within a digital image are specified by a Gibbs probability distribution (Farag *et al* 2005). In the discrete form of the Gibbs random field, this probability is described by as a probability mass function of the form:

$$p(\mathbf{x}) = \frac{1}{Z} \exp^{-\frac{E(\mathbf{x})}{\beta}} \qquad (4.19)$$

for the Gibbs random field (vector) $\mathbf{x}$, an energy function $E(\mathbf{x})$, a normalisation constant Z known as the *partition function* and a scalar $\beta > 0$ known as the *temperature* (equation 9.1, page 482, Farag *et al* 2005). The energy function is defined in terms of local interactions those pixels in a *clique*, c (which can be a single pixel or a set of pixels that are all neighbours to each other) and a *clique potential* $V_c(\cdot)$ which is a function of the random variables (image intensity for instance) of the members of the clique. The energy function is given by,

$$E(\mathbf{x}) = \sum_{i \,\epsilon\, C} V_c(\mathbf{x}) \qquad (4.20)$$

for the set C of all possible cliques (equation 28, page 366, Fieguth & Zhang 2005) and the partition function by,

$$Z = \sum_{\mathbf{x}} \exp[-\beta E(\mathbf{x})] \qquad (4.21)$$

(equation 29, page 366, Fieguth & Zhang 2005). Evaluation of Z is impractical for many image processing tasks and parameter estimation is achieved using *pseudo-likelihood* techniques (Besag 1974).

4.2.11   Markov random fields

Markov random fields have the property of decoupling image information into different regions of the image. The image pixels can be associated with random fields belonging to the interior $\mathbf{x}_I$, exterior $\mathbf{x}_E$ and the boundary $\mathbf{x}_B$ (Fieguth & Zhang 2005). Decoupling the image into different regions influences the posterior distributions of the random fields such that,

$$p(\mathbf{x}_I|\mathbf{x}_B, \mathbf{x}_E) = p(\mathbf{x}_I|\mathbf{x}_B), \quad p(\mathbf{x}_E|\mathbf{x}_B, \mathbf{x}_I) = p(\mathbf{x}_E|\mathbf{x}_B) \tag{4.22}$$

(equation 10, page 363, Fieguth & Zhang 2005). In practice this boundary is difficult to interpret and the posterior probability distribution for a single element $x_{i,j}$ (sampled from the field $\mathbf{x}$) is conditioned on the values of the elements of the field that are contained in a surrounding neighbourhood ($\mathcal{N}_{i,j}$). That is, $p(x_{i,j}|x_{k,l},\ (k,l)\ \epsilon\ \mathcal{N}_{i,j})$ is considered (equation 11, page 363, Fieguth & Zhang 2005). A variety of neighbourhood structures can exist that partially specify the complexity of the Markov random field model (see for instance example 6.12 and Figure 6.4 Davison 2003; or Figure 4, Fieguth & Zhang 2005). The random field can be estimated as a linear sum of weighted elements $w_{i,j,k,l}$ in the neighbourhood $\mathcal{N}_{i,j}$,

$$\hat{\mathbf{x}}_{i,j} = \sum_{(k,l)\ \epsilon\ \mathcal{N}_{i,j}} w_{i,j,k,l}\, x_{k,l} \tag{4.23}$$

(equation 14, page 364, Fieguth & Zhang 2005).

An important practical example of a Markov random field is the Gaussian Markov random field which is useful when the random field is of high dimension. A Gaussian Markov random field is a Markov random field whose elements can be described by a multivariate Gaussian distribution (Rue & Held 2005; Fieguth & Zhang 2005). The linear Markov random field model can be estimated as a linear sum of weighted elements $w_{i,j,k,l}$ in the neighbourhood $\mathcal{N}_{i,j}$,

$$\hat{\mathbf{x}}_{i,j} = \sum_{(k,l)\ \epsilon\ \mathcal{N}_{i,j}} w_{i,j,k,l}\, x_{k,l} \tag{4.24}$$

(equation 14, page 364, Fieguth & Zhang 2005).

A Markov random field model was developed and applied to digital mammograms of the breast by Katartzis *et al* (2005) in order to estimate the thickness of skin. The thickness of skin is an indicator of breast cancer, with retraction or thickening indicating abnormal changes. These changes may be evident on mammograms before they can be detected clinically. The radiographic properties of the skin region were modeled as a Gaussian Markov random field (Katartzis *et al* 2005) and incorporated as *a priori* knowledge in a binary segmentation of the image into a region of skin and no skin. Figure 4.6(a) displays the original digital image of the mammogram and Figure 4.6(b) the resulting binary image of the segmentation model where the dark region corresponds to the skin region. The results of the algorithm are assessed in Figure 4.6(c) as compared to manual measurement and indicate good overall agreement.

Random fields provide a model with stochastic structure with which to interpret an image. These models incorporate local information, maximum likelihood of model estimates can be obtained and statistical inference can be performed. Despite these attractive properties, the random field models reviewed have inherent limitations in the modeling of SAXS images. The SAXS images contain sharp intensity transitions (scattering rings) and long-range spatial dependencies (repeated orders of the scattering rings) which may not be well modeled using *linear* random field models (such as the Gaussian Markov random field model). Non-linear models based upon Gaussian mixtures might be more appropriate and a fruitful avenue of research in the modeling of SAXS images.
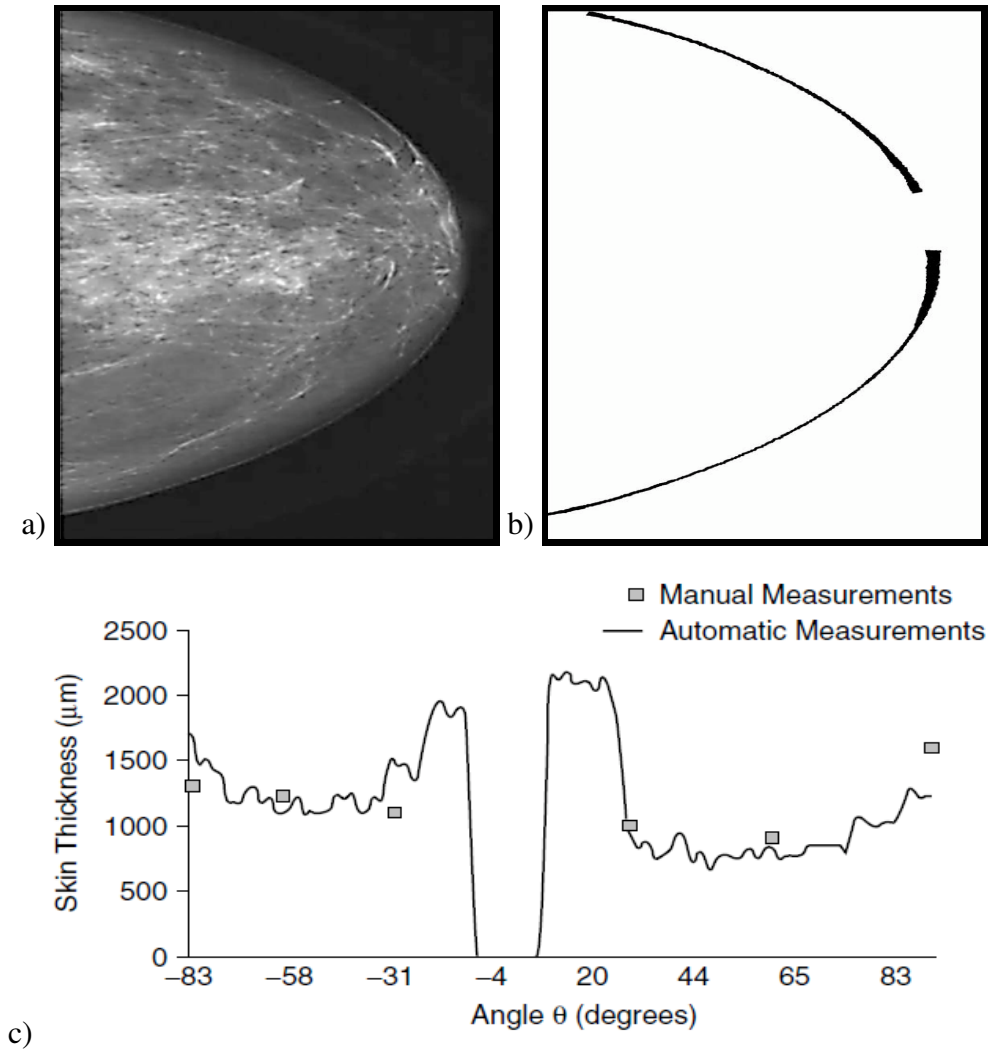
Figure 4.6: Mammogram image segmentation using a Gaussian Markov random field model: (a) original image (b) segmented result, dark regions indicates the region of the image corresponding to skin (epithelial) tissue and (c) the skin thickness (μm) as a function of angle θ, as measured from the centre of the image to the edge of the skin for the manual and automatic (random field) techniques (Figure 8.10, Katartzis *et al* 2005).
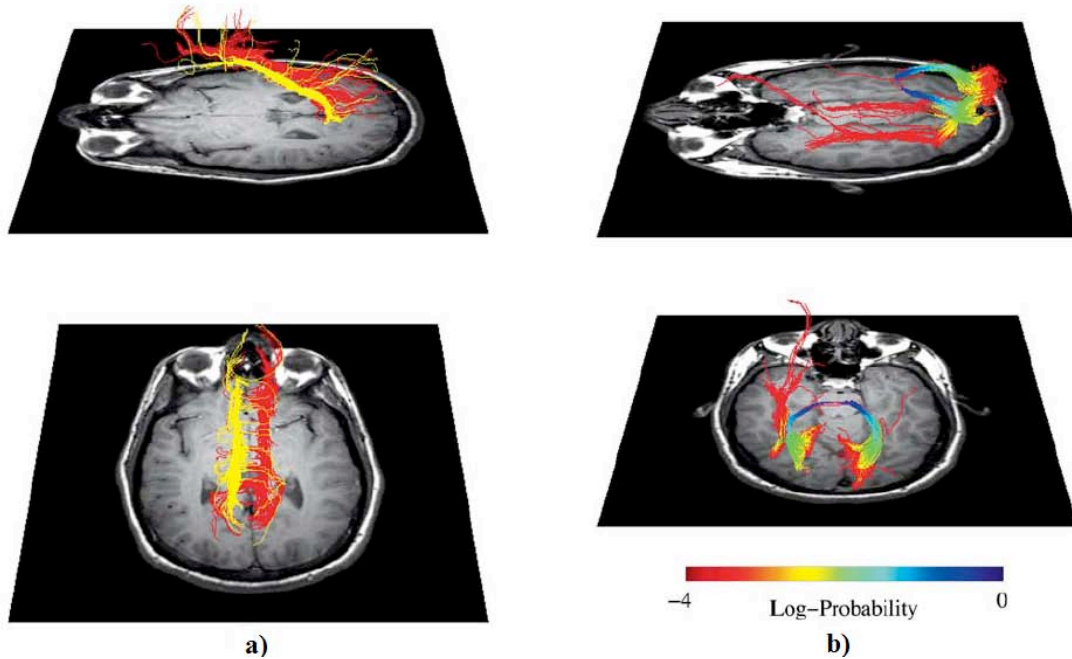
4.2.12 Bayes'ian statistical models

Figure 4.7: Analysis of white matter fibre tracts in a stack of magnetic resonance images (from two different viewpoints) using Bayes'ian methods: (a) fibre tracking in the left and right Cingulum bundles of the brain, the red and yellow fiber paths are associated with different initialisation points in the original image, (b) the model probability estimate evolving along the fiber path for 3000 samples from the splenium of the Corpus callosum of the brain (Figure 8, Frieman *et al* 2006).

Bayes'ian models permit the use of prior knowledge about a problem and as such their implementation is wide-spread throughout the image analysis literature. Bayes" theorem relates the posterior probability of an event occurring to both the likelihood of the event occurring (under a particular statistical model) and the prior probability that the event could occur.

The theorem is stated in mathematical form as,

$$p(\mathbf{x}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{x})p(\mathbf{x})}{p(\mathcal{D})} \tag{4.25}$$

where,

$p(\mathbf{x}|\mathcal{D})$ is the posterior probability of the parameter $\mathbf{x}$ being a certain value having observed the data $\mathcal{D}$,

$p(\mathcal{D}|\mathbf{x})$ is the likelihood, the probability of the data $\mathcal{D}$ having the observed magnitude given the parameter $\mathbf{x}$ under a certain statistical model,

$p(\mathbf{x})$ is the prior probability of the parameter before observing any data and,

$p(\mathcal{D})$ is the normalisation constant, the marginal probability of obtaining the data regardless of other information (equation 6, page 718, Hanson 1993).

The prior information is characterised by a user-specified probability density that is often obtained from the results of previous experiments or well-tried statistical methods. Common prior distributions used in image analysis include:

$$p(x,y) = |\nabla I(x,y)|^2 \tag{4.26}$$

$$p(x,y) = |\nabla^2 I(x,y)|^2 \tag{4.27}$$

$$p(x,y) = \exp\Big\{ -\beta \sum_{m=1}^{n} \omega_m \sum_{c=1}^{K} \phi[\nabla_c^{(m)}(x,y)]\Big\} \tag{4.28}$$

where $\beta > 0$ is a smoothing parameter, $\{\omega_n\} \; \epsilon \; \mathbb{R}^+$ are positive weights, $\phi(\cdot)$ is a potential function that is often specified as $\phi(u) = |u|/(1 + |u|)$ for a dummy variable, $u \; \epsilon \; \mathbb{R}$, ∇ is the discrete gradient operator, ∇^2 the discrete Laplacian operator and, $\nabla_c^{(m)}$ the general m-th order discrete derivative for the gray-level clique, c (equation 26, page 725, Hanson 1993 & equation 2.1, page 303, Horn *et al* 2003). The choice of the prior model is often tailored to the problem at hand and the prior models of equations (4.26-4.28) are designed with image edges in mind.

Once a statistical model has been specified, inference is often performed using the MAP (*maximum a posteriori*) estimator,

$$\hat{\mathbf{x}} = \arg\max p(\mathbf{x}|\mathcal{D}) \tag{4.29}$$

(equation 11, page 719, Hanson 1993), which can be interpreted as the maximum likelihood estimate (under the specified statistical model) weighted by the prior probability $p(\mathbf{x})$.

Bayes'ian concepts are incorporated into a wide variety of statistical models including pixel-level descriptions of an image (random fields) to higher-level structured descriptions of an image (Green 1995; Grenander & Miller 1994; Winkler 1995). Medical applications of Bayes'ian methods appear in tomography (Hanson & Wecksung 1983; Green 1990; Mumcuoylu, Leahy & Cherry 1994), radiography (Baydush *et al* 1997), echocardiography (Papademetris *et al* 2001), functional magnetic resonance (Genovese 2000), single photon emission computed tomography (Lee *et al* 1995) and ultrasound (Achim, Bezrianos & Tsakalides 2001).

The literature on Bayes'ian methods in image analysis is vast and many applications have not been mentioned but a study of particular interest is that of Frieman *et al* (2006) who used Bayes'ian models to locate and track white matter fibre bundles through stacks of $(i = 1, \ldots, n)$ magnetic resonance images of the human brain (Frieman *et al* 2006). This research is important for both surgical planning and the study of white matter degenerative diseases. The white matter tracts were described by a vector path $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, which is a collection of vectors in which each vector $\mathbf{v}_i$ describes the position of the white matter fibre in the ith image. The position of the current vector $\mathbf{v}_i$ was assumed to depend only on the position of the previous vector $\mathbf{v}_{i-1}$. The voxel intensity $I_i(x, y, z)$ was assumed to consist of a smooth μ_i and a stochastic ϵ_i component, so that $I_i(x, y, z) = \mu_i + \epsilon_i$. The smoothed component was specified using a model (Table 1, Frieman *et al* 2006) related to the underlying physical properties of the tissue,

$$\mu_i = \mu_0 \exp\left(-\alpha \mathbf{b}_i\right) \exp\left(-\beta \mathbf{b}_i (\mathbf{g}_i^T \mathbf{v})^2\right) \tag{4.30}$$

for a discrete gradient direction $\mathbf{g}_i$ and experimental parameters $\mathbf{b}_i$ (Table 1, Frieman *et al* 2006). A linear model was fit to the log-transformed intensity, $z = \ln I_i(x, y, z)$ as,

$$z_i = \ln \mu_i + \epsilon_i^\star \quad \epsilon^\star \sim \mathcal{N}(0, \sigma^2/\mu_i^2) \tag{4.31}$$

where $\theta = \{\mu_0, \sigma^2, \alpha, \beta\}$ are deemed *nuisance parameters* (page 968, Frieman *et al* 2006). The likelihood for the observed data $\mathcal{D} = [z_1, z_2, \ldots, z_n]$ in a voxel was,

$$p(\mathcal{D}|\mathbf{v}, \theta) = \Pi_{j=1}^N \frac{\mu_j}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-\mu_j^2}{2\sigma^2}\right\}(z_j - \ln \mu_j)^2 \tag{4.32}$$

(equation 11, page 968, Frieman *et al* 2006). Probability density functions were used to encode prior knowledge about the regularity of the fiber path as $p(\mathbf{v}_i|\mathbf{v}_{i-1})$ and prior knowledge of the nuisance parameters as $p(\theta)$. In this application, these prior functions were modeled as,

$$p(\mathbf{v}_i|\mathbf{v}_{i-1}) \propto \begin{cases} (\mathbf{v}_i^T \mathbf{v}_{i-1}^T)^\gamma, & \mathbf{v}_i^T \mathbf{v}_{i-1} \geq 0 \\ 0, & \mathbf{v}_i^T \mathbf{v}_{i-1} < 0 \end{cases} \tag{4.33}$$

where $\gamma \geq 0$ is a regularisation parameter influencing the smoothness of the path which was set to $\gamma = 1$ (equation 12, page 968, Frieman *et al* 2006). The nuisance parameters in the model had a prior that was a product of Dirac impulses,

$$p(\theta) = \delta(\mu - \hat{\mu}_0)\delta(\alpha - \hat{\alpha})\delta(\beta - \hat{\beta})\delta(\sigma^2 - \hat{\sigma}^2) \tag{4.34}$$

(page 968, Frieman *et al* 2001). Specialised eigenvalue factorisation and random sampling from a probability denisty function defined on a unit sphere were used to produce the model estimates (Frieman *et al* 2001). Results of the model for fibre tracking of the left and right Cingulum bundle are displayed in Figure 4.7(a), red fibers originate from the same initial point on the image as do the yellow fiber paths. Figure 4.7(b) displays the probability evolving along different fiber paths for 3000 fiber samples that were initiated in the splenium of the Corpus callosum. These results demonstrate how the fibre tracts can be estimated from a sequence of images using the model of Frieman *et al* (2006) and is one of many useful applications of Bayes'ian statistical methods.

Models that incorporate Bayes'ian methods have a wide range of useful applications, especially when relevant prior information is available. Bayes'ian methods could readily be incorporated in SAXS image models as prior experimental and computer-generated SAXS images based on the physical structure of collagen are readily available. The two main challenges that must be met with Bayes'ian models are the accurate specification of a prior distribution and model estimation. The high-dimensionality (number of pixels) of the SAXS images might make accurate and efficient model estimation difficult but the value of incorporating prior information into the diagnostic model makes the Bayes'ian approach worthy of further investigation.

4.2.13   Partial differential equation models

Partial differential equations play an important role in image processing and analysis. Dynamical, gradient, diffusion and equilibrium types of image processing problems can all be expressed in the language of partial differential equations (Chan & Shen 2005). Specific image processing tasks include denoising (Jain & Jain 1978; Chen, Vemuri & Wang 2000; You & Kaveh 2000; Lysakes & Tai 2003; Weeratunga & Kamath 2003), edge detection (Nordstrom 1990; Catte *et al* 1992; Perona & Malik 1990; Alvarez, Lions & Morel 1992), image restoration (Tschumperle & Deriche 2002; Welk *et al* 2002), segmentation (Pollak, Willsky & Krim 2000; Tsai *et al* 2001; Weickert 2001; Sofou & Maragos 2003) and motion tracking (Kornprobst, Deriche & Aubert 1999; Paragios & Deriche 2000; Sapiro 2001; Mansouri 2002).

The diffusion family of equations are prominent for analysing image sequences. The diffusion family of partial differential equations can be specified as,

$$\frac{\partial}{\partial t} I(x, y, t) = \nabla \cdot \big(D(x, y, t) \nabla I(x, y, t) \big) \tag{4.35}$$

where the t th image in the stack, $I(x, y, t) = K_\sigma * I(x, y, 0)$ is a convolution of the initial image $I(x, y, 0)$ and the Gaussian kernel K_σ with standard deviation parameter $\sigma = \sqrt{2t}$. The diffusivity function $D(x, y, t)$ is used to identify image edges, a standard form of this function is $D(|\nabla I(x, y, t)|) = 1\big((1 + |\nabla I(x, y, t)|^2)\big)$, where ∇ is the gradient operator and $\nabla \cdot$ is the divergence operator (equations 3 and 4, page 202, Weeratunga & Kamath 2003; Chan & Shen 2005a). The diffusion family of equations is similar to a multi-scale analysis of the image using Gaussian kernels of different sizes, σ. The diffusivity function $D(x, y, t)$ varies with position in the image to allow for non-linear effects and permits selective image edge enhancement. Other prominent partial differential equations that are used in image analysis include the heat and the Euler-Lagrange equations (Rudin, Osher & Fatemi 1992; Aubert & Vese 1997; Chan & Shen 2005a).

Yezzi Jr & Prince (2003) described an image analysis method based upon the solution of Eulerian partial differential equations. The method was designed to compute the thickness of a tissue boundary within a medical image or within a stack of medical images. A unique harmonic function, $u(\mathbf{x})$ was used to describe a select region, $R \subset \mathbb{R}^n$, $n = 2, 3$ of the image or the image stack.

The harmonic function $u(\mathbf{x})$ satisfied the conditions that:

$$\triangle u(\mathbf{x}) = 0 \tag{4.36}$$

and,

$$u(\partial_0 R) = 0 \quad u(\partial_1 R) = 1 \tag{4.37}$$

(equation 1, pages 1334 , Yezzi Jr & Prince 2003). The normalised gradient $\mathbf{T}(\mathbf{x})$ was introduced to perform interpolation in two regions $\partial_0 R$ and $\partial_1 R$. It was defined as,

$$\mathbf{T}(\mathbf{x}) = \frac{\nabla u(\mathbf{x})}{||\nabla u(\mathbf{x})||} \tag{4.38}$$

(equation 2, page 1334 , Yezzi Jr & Prince 2003). Two arc-length functions, $L_0(\mathbf{x})$ and $L_1(\mathbf{x})$ were also defined such that they satisfy two first-order linear partial differential equations,

$$
\begin{aligned}
\nabla L_0 \cdot \mathbf{T}(\mathbf{x}) &= 1, \quad \text{with } L_0(\partial_0 R) = 0 \\
- \quad \nabla L_1 \cdot \mathbf{T}(\mathbf{x}) &= 1, \quad \text{with } L_1(\partial_1 R) = 0
\end{aligned}
\tag{4.39}
$$

(equations 3 and 4, page 1334, Yezzi Jr & Prince 2003). The thickness, $W(\mathbf{x})$ of the tissue region of the image at point $\mathbf{x} \; \epsilon \; \mathbb{R}$ was then found using these arc-length functions,

$$W(\mathbf{x}) = L_0(\mathbf{x}) + L_1(\mathbf{x}) \tag{4.40}$$

(equation 5, page 1334, Yezzi Jr & Prince 2003).

Applications of this method to magnetic resonance images included measurement of myocardial thickness in the heart, tibial cartilage thickness in the human knee and cortical thickness in the brain (Yezzi Jr & Prince 2003). The measurement of cortical thickness in the brain using magnetic resonance images is singled out as an important application of this model because thinning of gray matter in the brain cortex might be associated with Alzheimer's disease (Fischl & Dale 2000). Figure 4.8 displays the results of the image model produced by Yezzi Jr & Prince (2003) when applied to the analysis of magnetic resonance images of the human brain. The model appears to be able to accurately extract the complex boundaries of the cortical surface of the brain (as determined by the images) for a variety of projections. The model allows for detailed reconstruction of a complex surface and provides researchers with detailed measurements of the cortical surface without the need for invasive operation. Overall, this example serves to demonstrate the capabilities of the partial differential equation approach.

Partial differential equation image models might be very useful in the extraction of diagnostic features in SAXS images. Alternatively, they might be very useful to transform SAXS images in a manner that highlights important aspects of the images such as scattering rings. The diffusion models that can analyse the image non-linearly across scales might be a useful method to study any hierarchical structure that exists in the SAXS images. The partial differential equation image methods and models appear to be most useful in image processing tasks and may require coupling with another classification model in order to be capable of breast tissue diagnosis. Regardless of these limitations, the opportunity to incorporate partial differential equation models into the analysis of SAXS images might provide new modeling opportunities.
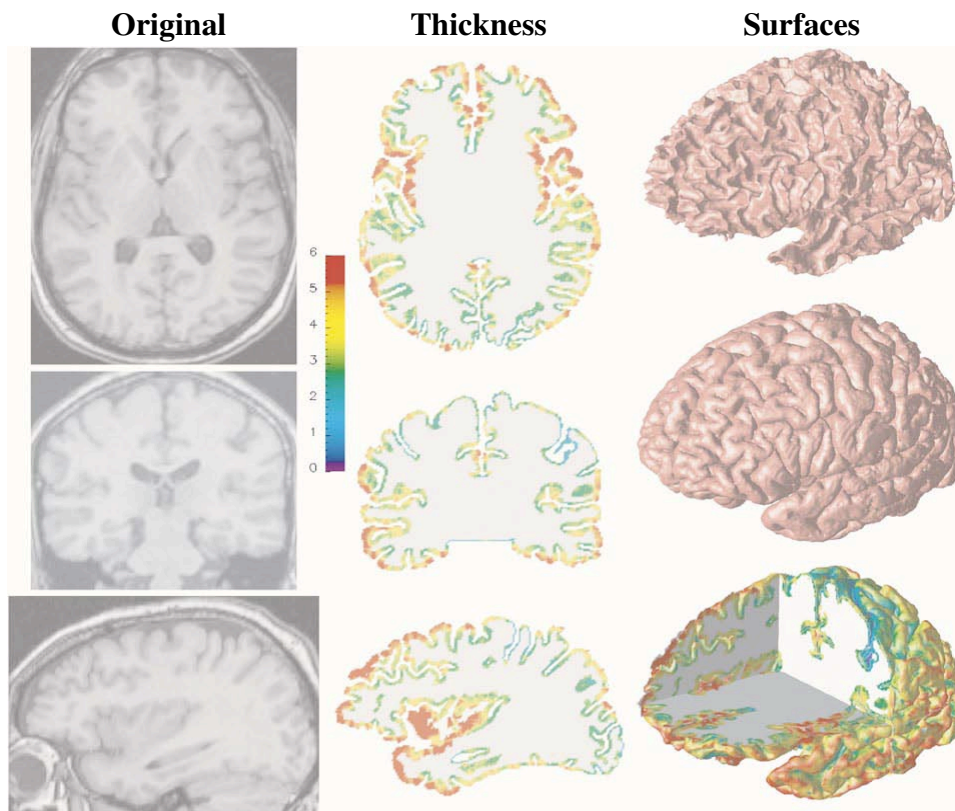

Figure 4.8: A partial differential equation image model was used to estimate the thickness of the cortical surface of the human brain using magnetic resonance images: original slices are displayed in the left column, the central column displays the gray matter thickness as calculated using the image model, the color-bar indicates the thickness of the surface in milli-metres, the upper and central objects in the right most column display the inner and outer three-dimensional surfaces (respectively) that were calculated for a 'stack' of images and the lower object in this column displays the thickness of the cortical surface as determined from the model (Figure 8, Yezzi Jr & Prince 2003).
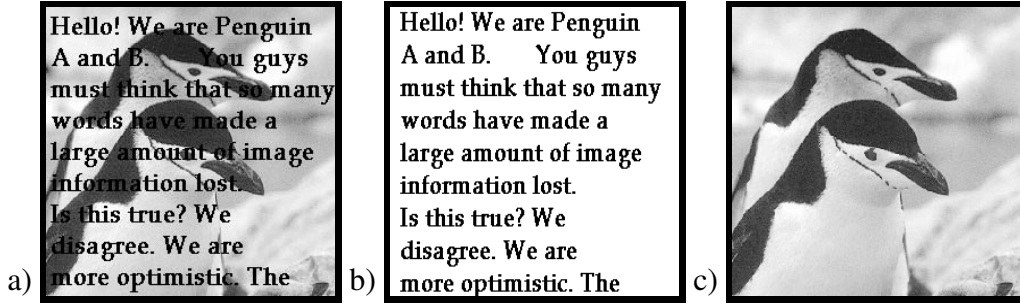
4.2.14   Variational Methods

Figure 4.9: Variational methods in image inpainting: (a) the original image to be restored consists of text overlaid on an image (b) the image mask that is to be removed (c) the result of the application of the modified Mumford-Shah model of Esedoghe & Shen (2002) that estimates the 'missing' information of the image inside the mask using variational methods (Figure 3, Esedoghe & Shen 2005).

Variational methods minimise functionals (such as a mapping from a vector space to a field) to solve image processing problems. Minimisation is based on an error functional which measures the deviations of the model from the observed data. Constraints that minimise of the residuals or a pre-specify the smoothness of the solution allow for it to be found as an optimisation problem. For instance if the error functional is of the form

$$\epsilon(\mathbf{f}) = \int_\omega L(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x}) \, d\mathbf{x}^W \tag{4.41}$$

where $\mathbf{f} = [f_1, f_2, \ldots, f_n]^T$ is a vector of features with partial derivatives $(f_p)\mathbf{x}_W$ over spatial coordinates $\mathbf{x} = (x, y)$ over the image domain $\omega \in \mathbb{R}^W$ (equation 17.8, page 468, Jähne 2005), then minimisation of the error functional $\epsilon(\mathbf{f})$ provides the solution. The function in the integrand, $L(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x})$ consists of two terms,

$$L(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x}) = S(\mathbf{f}, \mathbf{x}) + R(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x}) \tag{4.42}$$

where $S(\mathbf{f}, \mathbf{x})$ measures the similarity between the observations and the model, and $R(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x})$ is a regularisation penalty to ensure the constraints on the model are enforced. A commonly used similarity function is the L_n-norm,

$$S(\mathbf{f}, \mathbf{x}) = ||f(\mathbf{x}) - g(\mathbf{x})||_n \tag{4.43}$$

where $f(\mathbf{x})$ is the observed image and $g(\mathbf{x})$ is the model estimate (equation 17.10, page 468, Jähne 2005).

Another useful similarity function that is used in restoration problem is:

$$S(\mathbf{f}, \mathbf{x}) = ||h(\mathbf{x}) * f(\mathbf{x}) - g(\mathbf{x})||_n \tag{4.44}$$

where $h(\mathbf{x})$ is the convolution kernel or point-spread function of the image processing problem (equation 17.12, page 468, Jähne 2005). The regularisation constraint is often of the form:

$$R(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x}) = \alpha^2 |\nabla \mathbf{f}|^2 \tag{4.45}$$

where $\alpha \; \epsilon \; \mathbb{R}$ is a weight or smoothness term (adapted from equation 17.19, page 470, Jähne 2005). The variational problem can now be expressed as,

$$\min \int_\omega S(\mathbf{f}, \mathbf{x}) + R(\mathbf{f}, (\mathbf{f}_p)_{\mathbf{x}_W}, \mathbf{x}) \, d^W \mathbf{x} \tag{4.46}$$

(equation 17.24, page 471, Jähne 2005).

Variational analysis methods have been used in a large number of image processing problems. Well-known models include the total-variation and Mumford-Shah free-boundary models (Rudin, Osher & Fatemi 1992; Mumford & Shah 1989). Applications have included image segmentation (Koepfler, Lopez & Morel 1994; Chambolle 1995; Morel & Solimini 1995; March & Dozio 1997; Chan, Lom & Zhu 1998; Hewer, Kenny & Manjunath 1998; Blomgren *et al* 1999; Kimmel & Sochen 1999; Cremers 2002; Cremers *et al* 2003; Ayed 2006; Bertelli *et al* 2008), image restoration (Rudin & Osher 1994; Blomgen *et al* 1997; Teboul *et al* 1998; Chan, Marquina & Mulet 2000; Samson *et al* 2000, Chan & Shen 2001a; 2001b; Brook, Kimmel & Sochen 2003; Likas & Galatsanos 2004; Daubechies & Terchke 2005; Molina, Mateos & Katsaggelos 2006; Fundana, Overgaard & Heyden 2008), removal of noise (Aubert & Vese 1997; Nikolova 2004), inpainting (Bertalimio *et al* 2000; Chan & Shen 2001; 2002; Esedoglu & Shan 2002; Perez, Gangnet & Blake 2003; Chan & Shen 2005), registration (Droske & Rumpf 2004; Larry-Ruiz 2008), image matching (Amit 1994; Dupuis, Grenander & Miller 1998; Hermosillo & Faugeras 2001; Hermosillo, Ched d'Hotel & Faugeras 2004), medicine (Yezzi, Zollei & Kapur 2003; Faugeras *et al* 2004; Wang *et al* 2004; Ehrhardt, Schmidt-Richberg & Handels 2008; Gooya *et al* 2008) and remote sensing (Alvarez *et al* 2008).

Esedoglu & Shen (2002) applied variational methods to digital image inpainting, a restoration problem that fills in missing information in the image. The Mumford-Shah model was adapted to this task with the objective being to minimise an 'energy' functional of the form:

$$L[\widehat{I}(x,y)|I^\circ(x,y), \mathcal{D}] \tag{4.47}$$

where $\widehat{I}(x,y)$ the estimate of the restored images, $I^\circ(x,y)$ the uncorrupted part of the original image and $\mathcal{D}$ the domain over which the image information is missing (the 'mask') (page 355, Esedoghu & Shen 2005; Mumford & Shah 1989). Image intensity outside the mask is used to estimate the image intensity inside the mask. The capabilities of this image processing model are illustrated in Figures 4.9(a)-(c), the original 'corrupted' image that consists of a gray scale image overlaid with text is displayed in Figure 4.9(a), the image mask that conceals key parts of the image data is displayed in Figure 4.9(b) and the estimate $\widehat{I}(x,y)$ of the uncorrupted image obtained using the model of Esedoghu & Shen (2005) in Figure 4.9(c). Visual inspection of the restoration suggests that impressive results can be obtained using variational methods for image processing and this warrants more quantitative studies on a wider variety of images.

Variational methods are a very powerful tool for image processing tasks and they might be very useful in the extraction of features or in the pre-processing of SAXS images. Potential applications include the removal of detector noise from the images or the enhancement and detection of scattering rings. Alternatively, the images might be able to be segmented into simpler components such as the 'scattering' and 'background'. Variational methods seem better suited to feature extraction rather than classification.
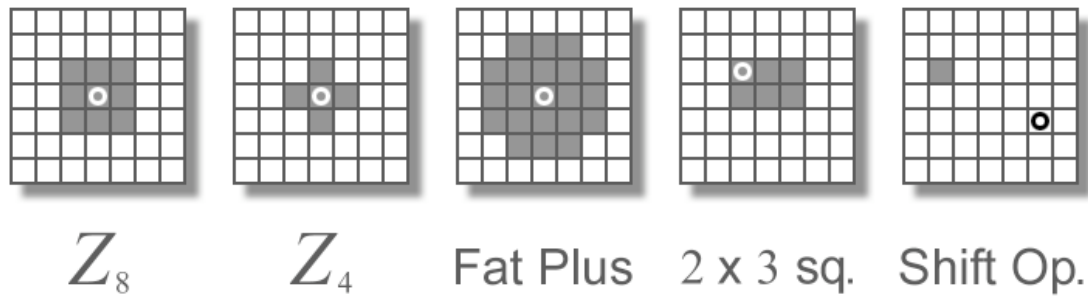
4.2.15 Mathematical Morphology

Figure 4.10: Examples of structuring elements used in morphological analysis of binary images. The foreground (in black) corresponds to $I(x, y) = 1$ and the background (in white) corresponds to $I(x, y) = 0$ with the circle denoting the origin of the structuring element (Peters II 2008).

Mathematical morphology is concerned with the geometric properties of shapes within images (Goutsias & Heijmans 2000). Matheron (1975) and Serra (1982) laid the foundations and popularised the technique which consists of analysing the shapes of objects in the image using small geometric objects (such as discs and squares) that are known as structuring elements. Examples of structuring elements used in the analysis of binary images are displayed in Figure 4.10. The specific structuring element used is often selected for the application at hand and acts as a moving window filter on the image.

Mathematical morphology was initially applied to binary images, but has since been extended to gray-scale images. In the analysis of gray scale images both the image and the structuring element are treated as if they were solids in a three-dimensional space with the third-dimension describing the intensity of the pixel values. The gray scale mathematical morphology techniques are of particular interest as they are directly applicable to SAXS images. Gray scale morphology consists of several basic operations which include erosion, dilation, opening and closing (Serra 1982), each of these operations is displayed in Figure 4.11.

For a finite number of gray levels in the image $I(\mathbf{x})$ and structuring element $Z(\mathbf{x})$, the dilation operation is defined as,

$$[I \overset{\bullet}{\oplus} Z](\mathbf{x}) = \bigvee_{\mathbf{s} \, \epsilon \, \mathrm{dom}(Z)} \left(I(\mathbf{x} - \mathbf{h}) \overset{\bullet}{+} Z(\mathbf{s}) \right) \tag{4.48}$$
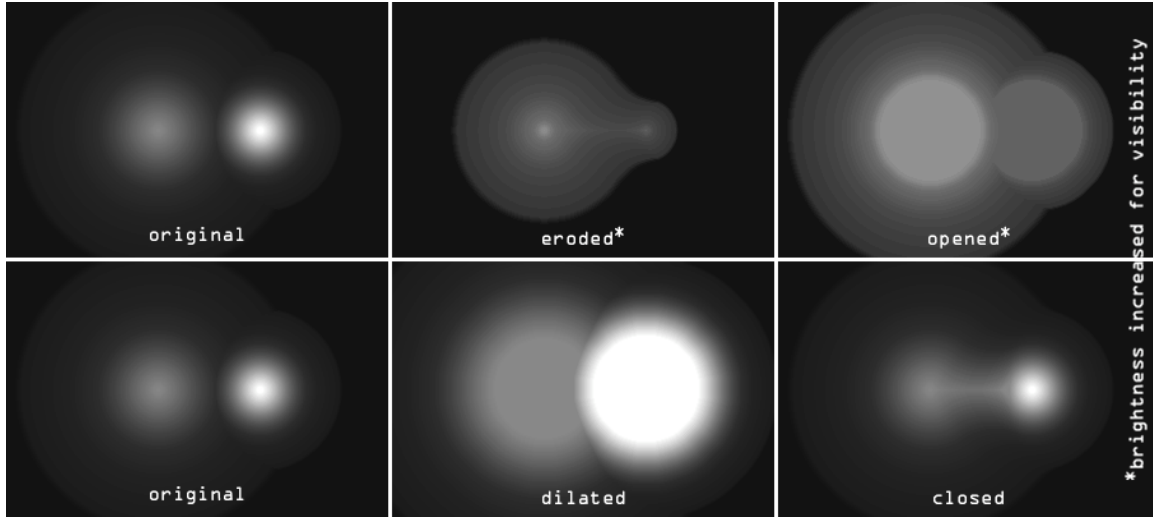
Figure 4.11: Examples of the basic operations of erosion, dilation, opening and closing on a gray scale image (Peters II 2008).

and the erosion operator as,

$$[I \stackrel{\bullet}{\ominus} Z](\mathbf{x}) = \bigwedge_{\mathbf{s} \,\epsilon\, \mathrm{dom}(Z)} \left(I(\mathbf{x} + \mathbf{s}) \stackrel{\bullet}{-} Z(\mathbf{s}) \right) \tag{4.49}$$

where the domain of the structuring element $Z(\mathbf{x} \geq 0)$ is denoted $\mathrm{dom}(z)$ with values in $\mathbb{Z}$, $\bigvee \mathcal{H}$ is the surpremum and $\bigwedge \mathcal{H}$ is the infimum of the subset $\mathcal{H}$ of the complete lattice $\mathcal{L}$ (a non-empty set furnished with a partial order relationship '$\leq$', so that every subset $\mathcal{H}$ of $\mathcal{L}$ has a least upper bound called the supremum and a greatest lower bound called the infimum and, where the operation $t \to t \stackrel{\bullet}{+} v$ is defined for $v \,\epsilon\, \mathbb{Z}$ on $\{0, 1, \ldots, N\}$ gray-levels as,

$$\begin{cases} 0 \stackrel{\bullet}{+} v = 0, \\ t \stackrel{\bullet}{+} v = 0, & \text{if } t > 0 \quad \text{and } t + v \leq 0 \\ t \stackrel{\bullet}{+} v = t + v & \text{if } t > 0 \quad \text{and } 0 \leq t + v \leq N \\ t \stackrel{\bullet}{+} v = N & \text{if } t > 0 \quad \text{and } 0 \leq t + v > N \end{cases}$$

and the operation $t \to t \stackrel{\bullet}{-} v$ on $\{0, 1, \ldots, N\}$ for $v \,\epsilon\, \mathbb{Z}$ as,

$$\begin{cases} t & \stackrel{\bullet}{-} & v & = & 0, & \text{if } t < N \quad \text{and } t - v \leq 0, \\ t & \stackrel{\bullet}{-} & v & = & t - v, & \text{if } t < N \quad 0 \leq t - v \leq N, \\ t & \stackrel{\bullet}{-} & v & = & N, & \text{if } t < N \quad t - v > N, \\ N & \stackrel{\bullet}{-} & v & = & N. \end{cases}$$

(pages 15-16 Goutsias & Heijmans 2000).

Opening is an operation that first erodes an image $I(\mathbf{x})$ by the structuring element, $\overset{\cup}{Z}(\mathbf{x})$, followed by dilation by $Z(\mathbf{x})$ on the result:

$$[I \overset{\bullet}{\circ} Z](\mathbf{x}) = [I \overset{\bullet}{\ominus} \overset{\cup}{Z}] \overset{\bullet}{\oplus} Z(\mathbf{s}) \tag{4.50}$$

where the notation $\overset{\cup}{Z}$ is the tranpose of the $Z(\mathbf{x})$ which is a reflection of $Z(\mathbf{x})$ with respect to the origin, that is the set of points such that $-Z(\mathbf{x}) \; \epsilon \; \mathbb{Z}$ (adapted from eqn II-21, page 50, Serra 1982).

Closing is an operation that first dilates and then erodes the image,

$$[I \overset{\bullet}{\bullet} Z](\mathbf{x}) = [I \overset{\bullet}{\oplus} \overset{\cup}{Z}] \ominus Z(\mathbf{s}) \tag{4.51}$$

(adapted from eqn II-21, page 50, Serra 1982). The erosion operation appears to contract and more sharply define the shape of an object in an image, whilst dilation seems to fill-out shapes in the image. Opening often removes small objects and closing often fills small holes in the foreground of the image.

Two other important gray scale morphological transforms are the 'gradient' and the 'top-hat' transform. The gradient transform is defined for digital images with a discrete number of gray levels as,

$$\text{grad}[I(\mathbf{x})] = \frac{1}{2}\big[(I \overset{\bullet}{\oplus} Z)(\mathbf{x}) - (I \overset{\bullet}{\ominus} Z)(\mathbf{x})\big] \tag{4.52}$$

(page 30, Heijmans 1995) while the open 'top-hat' transform is defined as,

$$I \,\widehat{\circ}\, Z(\mathbf{x}) = I - (I \overset{\bullet}{\circ} Z)(\mathbf{x}) \tag{4.53}$$

(equation 6.5, page 132, Dougherty & Lotufo 2003). Both transforms are useful in the detection of image features such as peaks and valleys in image intensity of particular sizes.

Mathematical morphology has a wide range of useful image processing applications including edge detection and feature generation. The research of Betal, Roberts & Whitehouse (1997) is a useful case study of the mathematical morphology technique. Digital mammograms were assessed using mathematical morphology techniques to detect micro-calcifications, an important indicator of breast cancer. A region of a digital image of a breast mammogram is displayed in Figure 4.12(a) which consists of an overall gray scale background texture and white (high intensity) speckles indicative of micro-calcifications. These micro-calcifications are in general difficult to detect visually on the mammogram. The morphological gradient (equation 4.52) was applied

to the image in Figure 4.12(a) to produce Figure 4.12(b). Regions associated with the micro-calcifications appear to have been identified and the internal structure of the breast is evident. Further processing using the open top-hat transform (equation 4.53) allowed the identification of regions of interest that might correspond to micro-calcifications. Additional image processing was then performed using the watershed transform to segment and extract features (Beucher 1994). A classification model with an accuracy of 86 % for the detection of micro-calcifications was then developed using a nearest-neighbour algorithm (Betal, Roberts & Whitehouse 1996). This example demonstrates a useful application of mathematical morphology in the medical image processing field.
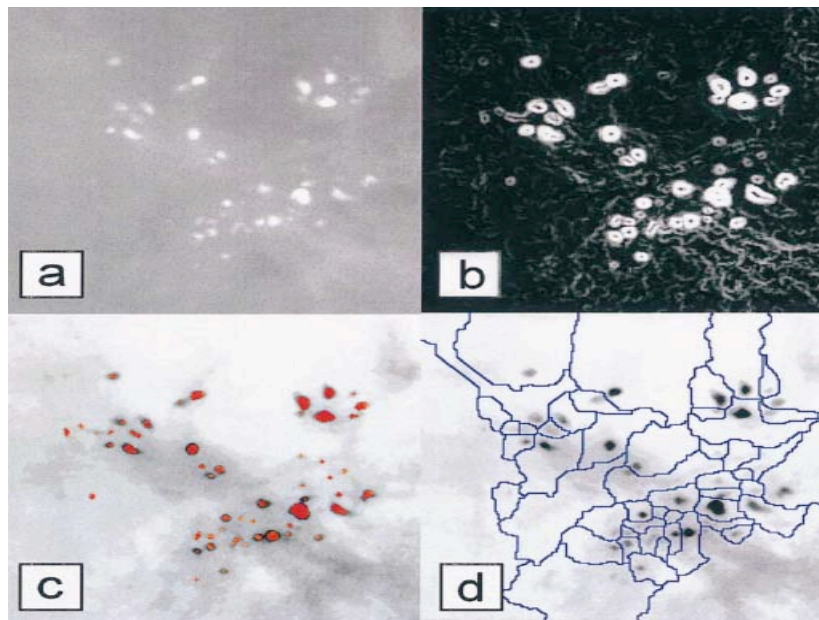

Figure 4.12: Mathematical morphology for microcalcification detection in digital mammograms: (a) original region of digital mammogram with micro-calcifications (white) (b) the result of the morphological gradient operator and (c) the identification of micro-calcifications (red) on the digital mammogram using the top-hat operator and (d) segmentation using the watershed algorithm to extract diagnostic features (Figure 1, Betal, Roberts & Whitehouse 1996).

The gray scale mathematical morphology method of image analysis appears to useful for a number of tasks, especially in the detection of features and edges. The main problem with the mathematical morphology approach for our applications is that these methods are focused on the geometric shape of objects within images. SAXS images do have geometric features (such as the scattering rings) but they may also contain important diagnostic information in the amorphous scattering intensity (Round 2006). The mathematical morphological operations might ignore this useful diagnostic information and they also do not directly consider the hierarchical structure of features in the image.

4.2.16   Shape Analysis

A wide variety of approaches have been adopted to analyse the shape of objects extracted from images. In this section, several of the key concepts and techniques in shape analysis are discussed. The notion of shape refers to that information that remains about an object after all other information about scale, location and orientation has been removed (Small 1996; Dryden & Mardia 1998). Methods for analysis include landmark/procrustes analysis (Bookstein 1991; Small 1996; Dryden & Mardia 1998; Kendall *et al* 1999; Lele & Richtsmeir 2000), methods involving mathematical morphology (Jang 1990; Iwanowski & Serra 2000; Tek & Kimia 1998; Pizer *et al* 1998), deformable templates (Kass, Witkin & Terzopoulos 1988; Cootes *et al* 1995; Jain, Zhong & Lakshmanan 1996; McInerneyy & Terzopoulos 1996; Jain *et al* 1998), contours & Fourier descriptors (Morse 1968; Zahn & Roskies 1972; Freeman 1974; Persoon & Fu 1977; Kuhl & Giardina 1982), and other general descriptors (da Fontoura Costa & Cesar Jr 2001). The scale-space representation that is provided by wavelets in another interesting approach to the problem (Antoine *et al* 1997; Shen & Ip 1999; Li & Li 2000; da Fontoura Costa & Cesar Jr 2001; Davatzikos, Tao & Shen 2003). Two of the most promising techniques for the analysis of SAXS images (deformable templates and elliptical Fourier descriptors) will be reviewed in greater depth.

4.2.16.1   Deformable Templates

Deformable templates are active contour models that are useful for segmenting, extracting and tracking geometrical objects (shapes) in images. These models combine 'low-level' information such as jumps in image pixel intensity with 'high-level' prior knowledge of the type of shapes that one wishes to extract from the images. A popular, special type of deformable template is that of a *snake* which is a parametric contour embedded in the image plane. The contour is represented parametrically as $\mathbf{v}(s) = \big[x(s), y(s)\big]^T$ where $s \; \epsilon \; [0, 1]$ and $(x, y) \; \epsilon \; \mathbb{R}^2$. The shape of the contour is determined by the functional,

$$\mathcal{E}[\mathbf{v}(s)] = \mathcal{S}[\mathbf{v}(s)] + \mathcal{P}[\mathbf{v}(s)] \tag{4.54}$$

(equation 1, page 3, McInerney & Terzopoulos 1988). The term $\mathcal{S}[\mathbf{v}(s)]$ describes the 'internal deformation energy',

$$\mathcal{S}[\mathbf{v}(s)] = \int_0^1 \{w_1(s)|\frac{\partial \mathbf{v}(s)}{\partial s}|^2 + w_2(s)|\frac{\partial^2 \mathbf{v}(s)}{\partial s^2}|\}ds \tag{4.55}$$

where $w_1(s)$ and $w_2(s)$ are weights controlling the 'smoothness' of the contour (equation 2, page 3, McInerney & Terzopoulos 1988) and,

$$\mathcal{P}[\mathbf{v}(s)] = \int_0^1 P[\mathbf{v}(s)]ds \qquad (4.56)$$

is a *external potential energy* function driving the fit of the contour to the image data (equation 3, page 3, McInerney & Terzopoulos 1988). The integrand, $P(v[\mathbf{s}])$ is selected such that the snake seeks out specific image features. An integrand of the form,

$$P(x, y) = -C|\nabla[G_\sigma * I(x, y)]| \qquad (4.57)$$

is often used to design a snake that seeks out edges in the image where G_σ is a Gaussian kernel that smooths the image to control the influence of minor edges and $C \in \mathbb{R}$ is a constant that controls the magnitude of the potential (page 200, Yezzi Jr *et al* 1997). Once a contour is selected, the potential function $P(x, y)$ is put in parametric form, $P[v(\mathbf{s})]$ for use in equation (4.56). Minimisation of the energy functional $\mathcal{E}[\mathbf{v}(s)]$ via repeated iterations to solutions of the Euler-Lagrange equation (McInerney & Terzopoulous 1996).

Yezzi Jr *et al* (1997) describe the application of a modified snake algorithm to medical image data. The modified method allowed automated handling of splitting and merging contours. Figure 4.13(a) displays the results of this algorithm when applied to computed tomography data of the bone. Two disconnected objects (that correspond to bone) are visible embedded in a tissue matrix. An initial contour is set (solid line) by the user over the region of interest (as visible in Figure 4.13(a)). This contour rapidly splits and converges onto the outlines of the bones (Figures 4.13(b)-(d)), finally settling on the outlines displayed in Figure 4.13(e). The example displayed in Figure 4.13 demonstrates the utility of the snake algorithm in the extraction of the shapes of objects of interest in medical images.
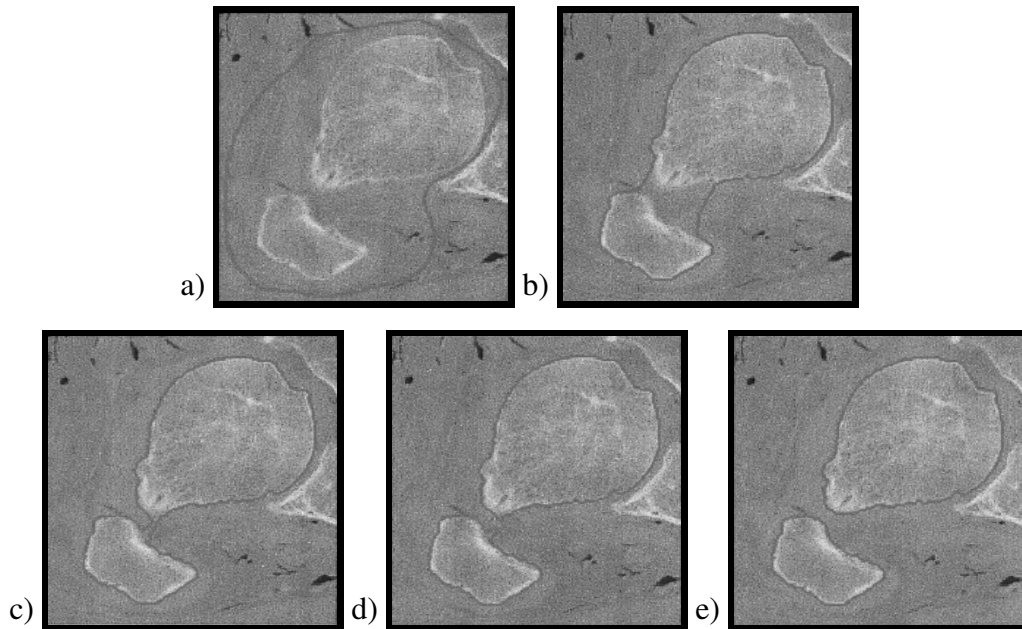

Figure 4.13: Extraction of key contours related to bone structure in a computed tomography image using a snake algorithm: (a) original image and the initial contour (solid line), (b)-(d) intermediate iterations of the algorithm, note how the contour is beginning to merge and (e) the final contour, at the 67 iteration of the algorithm (Figure 6, Yezzi Jr *et al* 1997).

4.2.16.2   Shape outline analysis

The analysis of shape outlines consists of detecting the outer (defining) edge of an object in an image and converting it to a binary image. A parametric representation of the outline $[x(t), y(t) : t \in \mathbb{R}]$ is used to describe the shape of the object.

Kuhl & Giardina (1982) described an elliptical Fourier transform that is useful for the analysis of such parametric curves. The elliptical Fourier transform basis functions are a set of circles and ellipses. A set of transform coefficients are used to provide a mathematical description of the shape of the object outline. The parametric functions $x(t)$ and $y(t)$ are represented using the elliptical Fourier transform as,

$$x(t) = A_0 + \sum_{n=1}^{N} a_n \cos(nt) + \sum_{n=1}^{N} b_n \sin(nt) \tag{4.58}$$

(equation 2.42, page 39 , Lestrel 1997) and,

$$y(t) = C_0 + \sum_{n=1}^{N} c_n \cos(nt) + \sum_{n=1}^{N} d_n \sin(nt) \tag{4.59}$$

(equation 2.43, page 39, Lestrel 1997), for each harmonic number, n, out of a total of N harmonics over the interval $[0, 2\pi)$. The elliptical Fourier transform coefficients that encode the description of the outline are given by,

$$a_n = \frac{1}{n^2 \pi} \sum_{p=1}^{q} \frac{\triangle x_p}{\triangle t_p} [\cos(nt_p) - \cos(nt_{p-1})] \tag{4.60}$$

$$b_n = \frac{1}{n^2 \pi} \sum_{p=1}^{q} \frac{\triangle x_p}{\triangle t_p} [\sin(nt_p) - \sin(nt_{p-1})] \tag{4.61}$$

$$c_n = \frac{1}{n^2 \pi} \sum_{p=1}^{q} \frac{\triangle y_p}{\triangle t_p} [\cos(nt_p) - \cos(nt_{p-1})] \tag{4.62}$$

$$d_n = \frac{1}{n^2 \pi} \sum_{p=1}^{q} \frac{\triangle y_p}{\triangle t_p} [\sin(nt_p) - \sin(nt_{p-1})] \tag{4.63}$$

where the shape is represented by a series of straight-line segments from points t_p to t_{p+1} over the period 2π such that $\triangle t_p = [\triangle x_p^2 + \triangle y_p^2]^{\frac{1}{2}}$ and q is the total number of points on the polygon (equations 2.44-2.47, page 39, Lestrel 1997).

The constant terms are given by,

$$A_0 = \frac{1}{2\pi} \sum_{p=1}^{q} \frac{\triangle x_p}{2\triangle t_p}[t_p^2 - t_{p-1}^2] + \alpha_p[t_p - t_{p-1}]$$ (4.64)

$$C_0 = \frac{1}{2\pi} \sum_{p=1}^{q} \frac{\triangle y_p}{2\triangle t_p}[t_p^2 - t_{p-1}^2] + \beta_p[t_p - t_{p-1}]$$ (4.65)

(equations 2.50 and 2.51, page 40, Lestrel 1997) with,

$$\alpha_p = \sum_{j=1}^{p-1} \triangle x_n - \Big[\frac{\triangle x_p}{\triangle t_p} \sum_{j=1}^{p-1} \triangle t_j\Big]$$ (4.66)

and,

$$\beta_p = \sum_{j=1}^{p-1} \triangle y_j - \Big[\frac{\triangle y_p}{\triangle t_p} \sum_{j=1}^{p-1} \triangle t_j\Big]$$ (4.67)

with $\alpha_1 = \beta_1 = 0$ (equations 2.50 and 2.51, page 40, Lestrel 1997).

Bailey & Lynch (2005) applied the elliptical Fourier transform to discriminate between images of teeth from Neanderthals and anatomically modern humans. The elliptical Fourier transform with $N = 32$ harmonics, was used to quantify observed shape differences (extracted from outlines of images of the teeth) in the mandibular P4 tooth for 136 samples. The mandibular P4 tooth is particularly useful for fossil identification of hominids and may in some cases be the only feature available for classification. Figure 4.14 identifies the regions of the P4 tooth that appear to be different between Neanderthals and anatomically modern humans.

Key characteristics that might distinguish a Neanderthal tooth (from anatomically modern humans) are:

a) a strong transverse crest

b) a well-developed and mesially placed metaconid
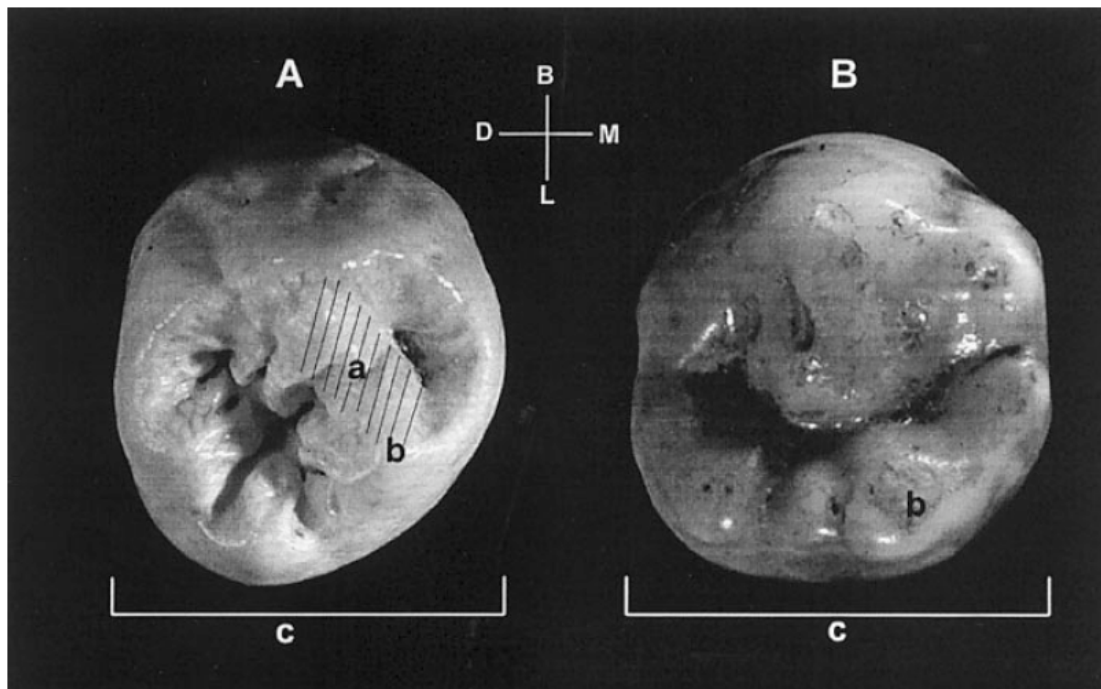
c) an asymmetrical lingual contour.

Figure 4.14: Differences between images of the (A: left) Neanderthal P4 and (B: right) the P4 of the anatomically modern human. Characteristics indicative of a Neanderthal specimen include a. a strong transverse crest (hatched area on the Neanderthal P4), b. a well-developed metaconid, c. an asymmetrical lingual contour. Symbols B indicates buccal, L - lingual, M- mesial and D - distal directions (Figure 1, Bailey & Lynch 2005).

Principal component analysis was applied to the data set of elliptical Fourier coefficients that described the P4 molar shape. The first two principal components described 64.1 % of the variation in the data set and a plot of the eigenvalues of the first two principal components (as in Figure 4.15) resulted in some separation between Neanderthal and other groups of anatomically

modern humans. The negative region of the first principal component corresponded to highly irregular P4 outlines and the positive region with smooth P4 outlines. The anatomically modern humans had a much greater variability in shape. Discriminant analysis on the 32 harmonics of the elliptical Fourier coefficients was also used to produce a discriminant vector that approximately separated the samples from the Neanderthal and anatomically modern human groups. Analysis of variance indicated that the Neanderthals had P4 teeth different from anatomically modern humans in:

i) the symmetry of the crown of the tooth and,

ii) the relative size of the lingual crown.

The model based upon the shape of the P4 molar provided correct identification for 65.0 % of the Neanderthal samples and 98.1 % of the anatomically modern human samples (Bailey & Lynch 2005). This study suggests that the shape of the P4 molar might be one of several useful features that can be used to identify hominids and provides an example of how outline shape features can be successfully incorporated into a classification model using the elliptical Fourier transform.

4.2.16.3 Conclusion of Shape Analysis Techniques

Both the deformable template and outline shape analysis models have the potential to be very useful in the analysis of SAXS images. Deformable templates could be used to extract the shape objects (such as scattering rings) and the elliptical Fourier techniques could be used to analyse these extracts. The scattering rings in the SAXS images convey shape information which is related to the underlying collagen distribution and structure of the sample. Inclusion of this shape information might be an important diagnostic feature in a classification model of tissue group. A limitation of shape analysis as applied to SAXS images is that this method extracts only a certain type of information from the images.
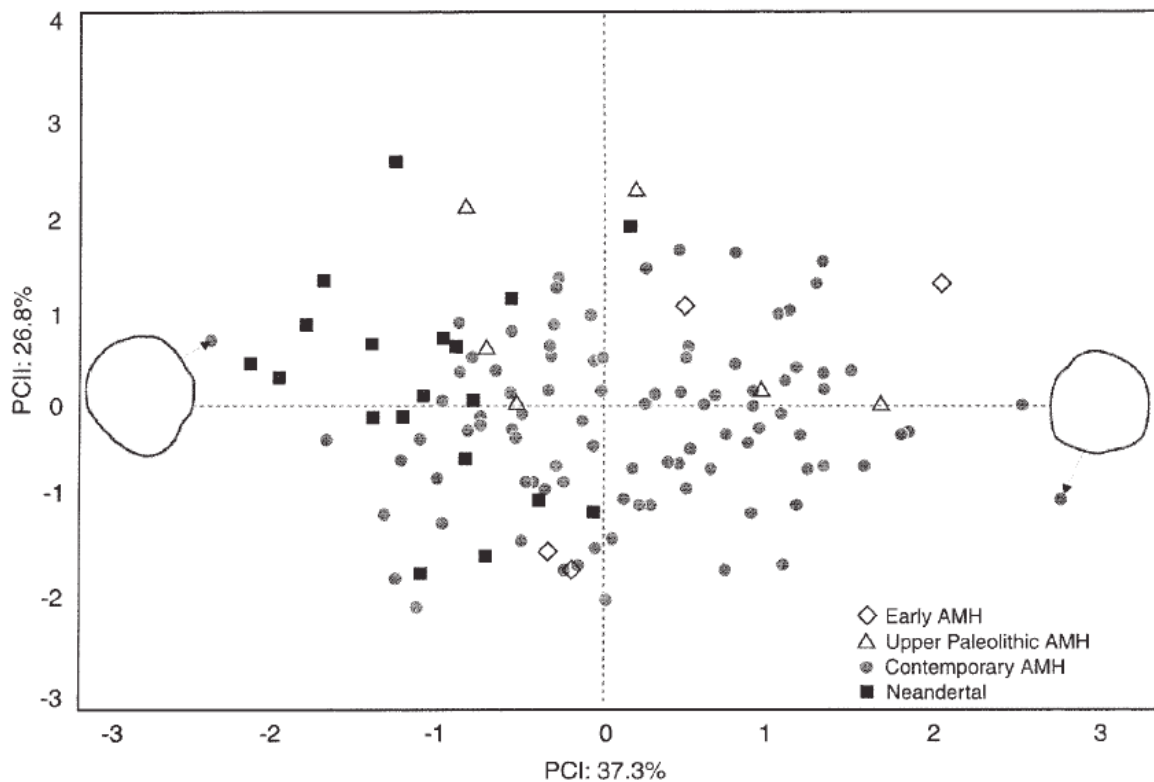
Figure 4.15: Principal component analysis of the elliptical Fourier coefficients that describe the shape of the P4 molar of Neanderthals and anatomically modern humans. The data is plotted on the first two components (PC1) and (PC2) for the Neanderthal and a variety of anatomically modern human subgroups. The first component appears to describe outline irregularity and is able to separate the Neanderthal from the anatomically modern human (AMH) samples (Figure 2, Bailey & Lynch 2005).

4.2.17 Fuzzy Analysis

Fuzzy image analysis can be loosely described as those image processing and analysis techniques that rely on fuzzy set theory as a key component. Developed by Zadeh (1965), fuzzy set theory is a mapping of data (for instance wavelet coefficient magnitude) from a set X onto the unit interval. This mapping is expressed mathematically as $\gamma : X \rightarrow [0, 1]$ (from definition 1.7, page 17, Höppner *et al* 1999). The value that the mapping γ indicates how much an observation x can be considered to be in the set X. A mapping value of $\gamma_x \approx 0$ indicates that the observation has a low amount of membership to the set and a value of $\gamma_x \approx 1$ indicates that the observation has a high amount of membership to the set.

An image analysis example that motivates fuzzy techniques is displayed in Figure 4.16. Some images analysis problems desire the partitioning of the regions into distinct objects in a process that is known as segmentation. Figure 4.16 displays a pair of overlapping circles superimposed on a uniform background. The 'dark' region belongs to the background, the 'light' region to the intersection of the two circles and the 'gray' region to those regions outside the intersection of the circles but still within their union. A challenge that must be addressed in the segmentation problem is the assignment of each pixel in the image to a particular object or class label. In this case unique assignment of each pixel to a particular object is not desirable. Those pixels within the intersection of belong to both circles with equal weight. The fuzzy interpretation of this problem is that the pixels within the intersection can be described by a membership vector of $\mathcal{M} = [\gamma_b, \gamma_{c1}, \gamma_{c2}] = [0, 0.5, 0.5]$ where γ_b denotes the membership of an image pixel to the set 'background', γ_{c1} to the set 'circle 1' and γ_{c2} to set 'circle2'. This example motivates the use of fuzzy set theory as a possible solution in image segmentation problems when the boundaries of objects within the image are not clearly defined.
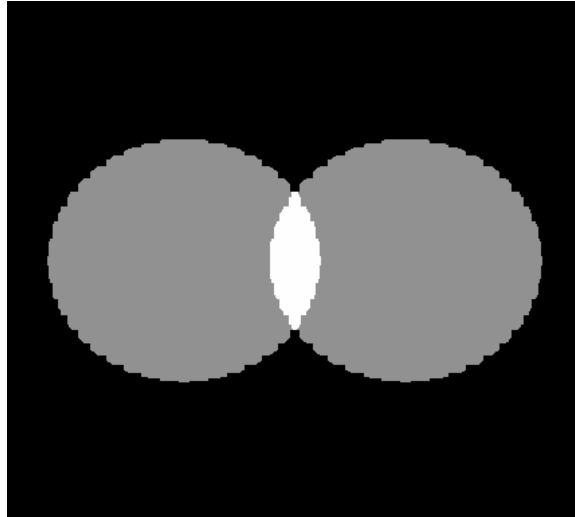
Figure 4.16: An image analysis problem where fuzzy sets might be useful. Two circles are super-imposed on a dark background, it is desired to segment the image and thereby detect those pixels belonging to each circle. Those pixels within the intersection of the two circles can be thought of as equally belonging to both circles, making the assignment of each pixel exclusively to one circle or the other or the background a troublesome task.

As suggested by the above example, fuzzy image analysis has been widely applied to image segmentation problems including the work of Huntsherger, Jacobs & Cannon (1985), Trivedi & Bezdek (1986), Lim & Lee (1989), Nguyen & Cohen (1993), Lin, Cheng & Mao (1996), Udupa & Samarsekara (1996), Moghaddamzadeh & Bourbakis (1997), Liew, Leung & Lau (2000), Rezaee *et al* (2000), Ahmed *et al* (2002), Hall *et al* (1992), Leung, Wang & Lau (2004) and Petrosino & Salvi (2006). Other prominent image analysis/processing applications include pattern recognition (Pedrycz 1990; Pal 1992; Jahns, Nielsen & Paul 2001; Bezdek *et al* 2005; Binaghi *et al* 2007), removal of image noise (Huang & Wang 1995; Mansfield *et al* 1998; DeVille *et al* 2003; Rosales-Silva *et al* 2007; Morillas *et al* 2007; Schulte *et al* 2007; Tang, Wang & Qi 2007; Zhang & Zhang 2007; Nair & Wilscy 2008), image enhancement and event detection (Goetcherian 1980; Pal & King 1980; Gong 1993; Tao, Thompson & Tour 1993; Kuo, Lee & Liu 1997; Russo 1998; Tizhoosh 2002; Liang & Looney 2003; Wu, Yin & Xiang 2007), math-ematical morphology (Sinha *et al* 1997; Strauss & Comby 2005; Bloch, Martino & Petrosino 2006; Mansoor *et al* 2007). A major application has been in the analysis of medical images (Bezdek *et al* 1997; Udupa *et al* 1997; Park, Hoffman & Sonka 1998; John, Innocent & Barnes 2000; Mansfield *et al* 1998; Pavlopoulos *et al* 2000; Suri 2001; Guliato *et al* 2003; Hassanien,
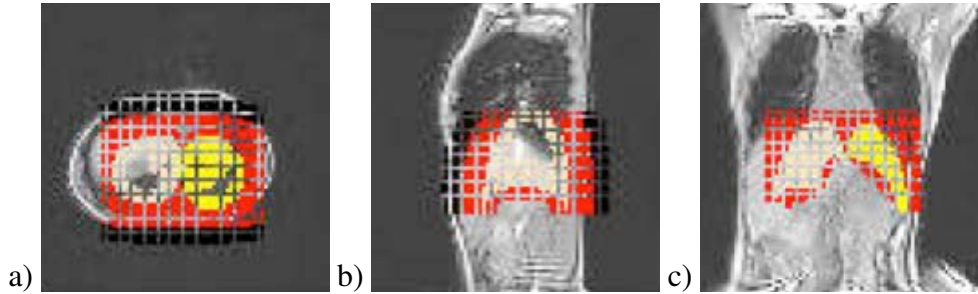
Figure 4.17: Segmentation of thoracic magnetic resonance images using fuzzy surface templates: (a) axial (b) sagittal and (c) coronal projections. Red dots indicate tissue regions between the lungs and thoracic outer wall whilst yellow dots regions associated with the left lung as identified by the algorithm (adapted from Figure 5.6, Lelieveldt *et al* 2000).

Ali & Nobuhara 2004). Reference works that discuss image analysis using fuzzy techniques include Kandel (1982), Pal & Dutta-Majumder (1986), Bezdek & Pal (1992), Ghosh & Pal (1992), Klir & Yuan (1995), Chi, Yan & Pham (1996), Höpner *et al* (1999); Pal & Mitra (1999), Kerre & Nachtegael (2000) and, Nachtegael *et al* (2003).

Lelieveldt *et al* (2000) describe the segmentation of thoracic magnetic resonance images using fuzzy implicit surfaces. Three-dimensional templates of relevant anatomical structures (such as the heart and lungs) were modeled as surfaces of the form, $f(x, y, z)$, which were used to describe the appearance of the structure as it would appear when projected into a magnetic resonance image. A 'crisp' surface would have a boundary of the form of $f(x, y, z) = c$ for a constant $c \; \epsilon \; \mathbb{R}$ with values less than this constant being *inside* the boundary and those greater than this constant being *outside* the boundary. In contrast, a surface with a 'fuzzy' boundary is less well defined with all points in $\mathbb{R}^3$ having membership to each of the three sets, 'inside boundary', 'on boundary', 'outside boundary'. The result is a surface that is modeled as a gradual (rather than sharp) transition from object to background. Multiple organs were modeled as three-dimensional fuzzy surfaces with respect to their expected position in the image and an algorithm developed that allowed automated, accurate (within 6mm on average) identification of the boundary surfaces of the heart and lungs in the images. Figures 4.17(a)-(c) display the results of the automated detection of the boundary surfaces of the lungs in a variety of magnetic resonance images. Yellow dots indicate regions associated with the left lung and red dots indicate those tissues between the lungs and the thoracic outer wall. Identification of the vessel boundaries within acceptable accuracy (approximately 6mm) was reported for an average of 90 % of the contour length of the boundary.

The fuzzy image analysis techniques might have useful applications in the analysis of SAXS images. One possible application is to transform the intensity values in the image into a fuzzy set. Membership could be encoded in terms of 'not edge' and 'edge' or in terms of 'high intensity', 'medium intensity' and 'low intensity' that might provide a different insight into the data but it is necessary to verify that important diagnostic information is not lost in the transformation process. Another approach could include a fuzzy classification with each image being assigned membership to each of the three groups 'normal', 'benign' and 'malignant'. The observed image can then be said to be more like 'normal' or more like 'malignant' depending on these membership values. Instead of distinct classifications into tissue groups, the images are assigned to a continuous 'spectrum' of states.

The concept of a sample belonging to multiple states or groups simultaneously could be very useful in SAXS image diagnosis. A sample may well contain both normal, benign and malignant tissue. Classification models that would account for this characteristic would be beneficial. Diagnostic models could be built that incorporate the concept that an observation can belong to multiple groups but that still allow for probability estimates to be developed that allow accurate detection of an abnormality. For instance a tissue sample could contain a 50-50 mixture of normal and malignant tissue, traditional fuzzy analysis would provide high membership values of the sample for both the normal and malignant groups. In contrast, the diagnostic model could be fined tuned for such cases such that it produces a high probability of malignancy even when other tissue is present and a high probability of being normal only when normal breast tissue is present in the sample.

4.3 Summary & Conclusions to the Image Analysis Review

The interpretation of the data as a digital SAXS image was discussed in Section 4.1 and a wide range of image processing and analysis techniques were reviewed in Section 4.2. This review serves as a guide to some of the major image processing and analysis techniques that are available to study SAXS images. It is by no means comprehensive in comparison to the vast literature that exists on the topic but it is sufficient enough to obtain ideas on how to model SAXS images.

Selection of a 'best' technique is a difficult and subjective task that inherently depends on the application at hand. The majority of image processing and analysis techniques reviewed in this chapter could certainly be incorporated as part of a diagnostic model of SAXS images. Fourier analysis techniques might be useful because they allow back-transformation to recover the original image, match well with the theory of SAXS image formation and capture spatial frequency information. Wavelets and the scale-space approach matches very well with the hierarchical structure of the collagen that generated the images, allows directional filtering to separate the meridional and equatorial scattering rings and have already produced good results when used by Erickson (2005) and Falzon *et al* (2006). The concept of a fractal surface might be useful in the description of the amorphous scatter in the SAXS images that was of interest to both Fernández *et al* (2002) and Round (2006). Random fields might also provide an alternative model to describe this background scatter. Prior information (from previous SAXS experiments) is available and could be incorporated using Bayes'ian techniques. Both partial-differential equations and variational methods could be used to pre-process SAXS images to provide one of several components of a SAXS diagnostic model. This would have to be done in a way that ensures that key diagnostic information is not lost. Mathematical morphology operations such as the top-hat transform (equation 4.53) could also be used in such a manner to detect scattering rings in the images. The shapes of the scattering rings could also be extracted as one of several components of a diagnostic model but again they may have to be included as additional model component to the intensity values of the SAXS image. The concepts of fuzzy set theory could be useful for future demands of the classification model such as the diagnosis of malignancy grade or when a sample belongs to multiple categories simultaneously. For instance, a future role of the SAXS diagnostic model might be to provide the exact pathological type of the tumour, not just classification of whether it is benign or malignant. This could be very challenging as some breast tumour pathologies such as Phyllodes tumor, can be both benign, malignant or somewhere in between (borderline) (Scarff 1981). Models using fuzzy set theory might be well-suited to such problems.

Given the inherent hierarchical nature of the structure (collagen) that produced the images, it seems sensible to analyse the SAXS image using image filters such as wavelets. A flexible custom-designed image transform that incorporates a range of analysis functions seems to be the key to extract and separate different information on tissue structure. As such, the transform approach will be given emphasis in this thesis and serve as a initial starting point for model development. The other desirable qualities possessed by the other image processing and analysis techniques will also be incorporated into the model where possible.

Chapter 5

Frequency and Time-Frequency Image Transforms

Chapter 4 reviewed the wide range of image processing and analysis tools that are available to the researcher in the study of SAXS images. The Fourier and wavelet transforms were identified as being of particular interest because of their ability to study the image in terms of either spatial frequency or scale-space features. A wide range of frequency and time-frequency image transforms have been reported in the literature that merit further consideration. This chapter will briefly survey and describe a range of Fourier, wavelet and related transforms. Section 5.1 will review the continuous, discrete and windowed Fourier transforms in greater detail, Section 5.2 examines the integral, discrete and stationary wavelet transforms. More advanced, custom-designed image transforms such as the wavelet packet, steerable pyramid, curvelet and contourlet transforms are described in Section 5.3. Adaptive multi-scale image transformations that describe the image using a library of waveforms are reviewed in Section 5.5. The best orthogonal basis, matching pursuit and basis pursuit methods are all discussed in this section. The chapter ends with a brief summary and a review which emphasises the broad, diverse and extensive range of transforms available to the researcher.

5.1   The Fourier Transform Family

The Fourier transform is useful to analyse the *spatial frequency* content of the image. The discrete Fourier transform (discussed in Section 4.2.1) is one of several transforms in the Fourier transform family that are available for image analysis. Other alternatives include the windowed Fourier and the discrete cosine transforms. All three transforms are useful tools in different settings. Section 5.1.1 reviews the discrete Fourier transform, Section 5.1.2 the windowed Fourier transform and Section 5.1.3 the discrete Cosine transform.

5.1.1   The Discrete Fourier Transform

Recall equation (4.3) defines the discrete Fourier transform of a digital image $I(x, y)$ of size N_1 x N_2 pixels that is indexed by location parameters (x, y) as,

$$F(k, l) = \frac{1}{N_1 N_2} \sum_{y=0}^{(N_2-1)} \sum_{x=0}^{(N_1-1)} I(x, y) \exp\left[- i2\pi\left(\frac{ky}{N_1} + \frac{lx}{N_2}\right)\right] \tag{5.1}$$

for frequency parameters (k, l) where $k = 0, \ldots, N_1 - 1$ and $l = 0, \ldots, N_2 - 1$ (equation 2.42, page 100, Rangayyan 2005). The discrete Fourier transform $F(k, l)$ is in fact a sampled version of the continuous Fourier transform $\mathcal{F}(\omega_1, \omega_2)$ which is defined as,

$$\mathcal{F}(\omega_1, \omega_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp[-i2\pi(\omega_1 x + \omega_2 y)] dx\, dy \tag{5.2}$$

when $f(x, y) \in L_2(\mathbb{R})$ is a continuous function (adapted from equation 2.32, page 47, Jähne 2005). This continuous function is the 'ideal' representation of a digital image that has infinite resolution. The frequency index (k, l) used in the discrete Fourier transform is a subset of those frequencies $\boldsymbol{\omega} = (\omega_1, \omega_2)$ described by the continuous Fourier transform. The discrete Fourier transform is nonetheless sufficient to describe the discrete sampled intensity values of a gray-scale image in terms of spatial frequency (cycles/pixel). As the rate of sampling increases, the discrete Fourier transform of an image more closely resembles the continuous Fourier transform of that same image. Both versions of the Fourier transform analyse the image in a global context, that is the support (range of non zero values) of the sinusoidal basis functions $\{\exp\left[- i2\pi(\frac{ky}{N_1} + \frac{lx}{N_2})\right]\}$ is over the entire image. Therefore, the value of the discrete Fourier transform coefficients, $F(k, l)$ depends on the values of the image $f(x, y)$ over the entire pixel grid N_1 x N_2. This characterises the *global* nature of Fourier analysis, in that it is possible to measure the spatial frequency content of the image but it is difficult to determine the locations of where these frequencies occur in the image.

5.1.2   The Windowed Fourier Transform

The windowed Fourier transform analyses an image in terms of both spatial location and spatial frequency. It provides insight into not only the spatial frequency content of the image but also the locations where these frequencies occur. The transform uses Gabor functions of the form,

$$g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x,y) = g(x - \mu_1, y - \mu_2) \exp\left[i(\epsilon_1 x + \epsilon_2 y)\right] \tag{5.3}$$

and,

$$G_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(\boldsymbol{\omega}) = G(\omega_1 - \epsilon_1, \omega_2 - \epsilon_2) \exp\left\{ -i[\mu_1(\omega_1 - \epsilon_1) + \mu_2(\omega_2 - \epsilon_2)]\right\} \tag{5.4}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{R}^2$, $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2) \in \mathbb{R}^2$, $\boldsymbol{\omega} = (\omega_1, \omega_2) \in \mathbb{R}^2$ (adapted from equation 1.5, page3, Mallat 1998). The windowed Fourier transform correlates the continuous image function $f(x, y)$ with the Gabor function $g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x, y)$. The transform is defined as,

$$\mathcal{SF}(\boldsymbol{\mu},\boldsymbol{\epsilon}) = < f(x,y), g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x,y) > = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x,y) \, dx \, dy \tag{5.5}$$

(adapted from equation 1.7, page 4, Mallat 1999). Approximations of this transform can be readily obtained for discrete data (such as digital images) by convolution with discrete versions of the filter $g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x, y)$. The energy of $g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x, y)$ is concentrated over the neighbourhood of $\boldsymbol{\mu}$ over an interval of size $\boldsymbol{\sigma_\mu} = (\sigma_x, \sigma_y)$. In the spatial frequency domain the function $G_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(\boldsymbol{\omega})$ is localised over the frequency region $\boldsymbol{\epsilon}$ of size $\boldsymbol{\sigma_\omega} = (\sigma_{\omega_1}, \sigma_{\omega_2})$. The Heisenberg uncertainty principal limits the precision of simultaneous measurements in space and frequency which means that the product of $\boldsymbol{\sigma_\mu}$ and $\boldsymbol{\sigma_\omega}$ has a finite value greater than zero (page 104, Prasad & Iyengar 1997). The Gabor function corresponds to the case when $g_{\boldsymbol{\mu},\boldsymbol{\epsilon}}(x, y)$ is a *Gaussian*, it allows maximum resolution in both space and frequency. The windowed Fourier transform provides a powerful alternative to the discrete Fourier transform when the spatial and frequency aspects of the image need to be examined simultaneously.

5.1.3   The Discrete Cosine Transform

The discrete cosine transform is an alternative to the discrete Fourier transform. A key difference between these two transforms is that the coefficients of the discrete cosine transform are all real-valued. There are at least eight versions of the discrete cosine transform reported in the literature, all of which project the image onto various forms of cosine basis functions. Strang (1999) discusses these various forms in detail. In this section the version of the transform that is often presented in image processing textbooks is examined. The type II discrete cosine transform (as labeled by Strang (1999)) is defined as,

$$F_C(k,l) = \frac{2}{\sqrt{N_1 N_2}} C(k,l) \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} I(x,y) \cos\left[\frac{\pi k(2x+1)}{2N_1}\right] \cos\left[\frac{\pi l(2y+1)}{2N_2}\right] \qquad (5.6)$$

for a digital image $I(x,y)$ of size N_1 x N_2 and discrete frequency indices $k = 0, 1, \ldots, N_1 - 1$ and $l = 0, 1, \ldots, N_2 - 1$ where the filter coefficients $C(k,l)$ are defined as,

$$C(k,l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \quad \text{or } l = 0 \\ 1 & \text{otherwise} \end{cases} \qquad (5.7)$$

(adapted from page 71, Acharya & Ray 2005). The inverse transform allows recovery of the original image,

$$I(x,y) = \frac{2}{\sqrt{N_1 N_2} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} C(k,l) F(k,l) \cos\left[\frac{\pi k(2x+1)}{2N_1}\right] \cos\left[\frac{\pi l(2y+1)}{2N_2}\right]} \qquad (5.8)$$

for $x = 0, 1, \ldots, N_1 - 1$ and $y = 0, 1, \ldots, N_2 - 1$ (adapted from page 71, Acharya & Ray 2005). The cosine basis functions are orthogonal and are separable, that is the basis functions can be formed as a product of one-dimensional functions. Therefore, the type II discrete cosine transform can be performed by applying a one-dimensional transform along the rows of the image followed by another one-dimensional transform along the columns. The discrete cosine transform is well-known for its energy compaction properties and corresponding sparse representation, that is the transform has few large magnitude coefficients. This property has lead to applications in image compression such as the joint photographics experts group (JPEG) image format that is widely used in digital media (Nixon & Aguado 2002). The discrete cosine transform is *shift-variant*, which means that the transform of the translated image is not equivalent to a translation of the same size of transform of the original image (Acharya & Ray 2005).

5.2 The Wavelet Transform Family

The wavelet family of transforms describe a variety of approaches to the analysis of an image over space and scale. The integral wavelet transform, which is introduced in Section 5.2.1 is a redundant scale-space representation of an image that can be calculated using convolution of a wavelet basis function and the image. In contrast, the discrete wavelet transform that is described in Section 5.2.2 uses a combination of scaling and wavelet functions to describe the image. Section 5.2.3 describes a translation-invariant but redundant version of the discrete wavelet transform that is known as the stationary wavelet transform. It is useful in those pattern recognition tasks where shifts of the patterns can occur. The wavelet packet transform is discussed in Section 5.2.4. It is a generalisation of the discrete wavelet transform that uses a greater combination of image filters which makes it more versatile in some image analysis applications.

5.2.1 The Integral Wavelet Transform

The integral wavelet transform is defined as,

$$
\begin{aligned}
\mathcal{W}_\psi[f(x,y)](a,\mathbf{b},\theta) &= C_\psi^{-\frac{1}{2}} < \psi_{a,\mathbf{b},\theta}(x,y), f(x,y) > \\
&= C_\psi^{-\frac{1}{2}} a^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\psi^\star\big(a^{-1}r_{-\theta}(x-b_1,y-b_2)\big)\,dx\,dy
\end{aligned}
\tag{5.9}
$$

where $C_\psi^{-\frac{1}{2}}$ is a constant, $\psi_{a,\mathbf{b},\theta}(x,y) \,\epsilon\, L^2(\mathbb{R}^2, dx\,dy)$ is a wavelet function such that $\psi_{a,\mathbf{b},\theta}(x,y) = a^{-1}\psi[a^{-1}r_{-\theta}(x-b_1,y-b_2)]$, $a \,\epsilon\, \mathbb{R}^+$ is a scale parameter, $\mathbf{b} = (b_1, b_2)$, $b_1, b_2 \,\epsilon\, \mathbb{R}$ are translation parameters and $\theta \,\epsilon\, [0, 2\pi)$ is a rotation parameter (adapted from equations 1 and 13, page 569, Arnéodo, Decoster & Roux 2000). The symbol $r_\theta(x,y)$ denotes the rotation operation,

$$
r_\theta(x,y) = \big[x\cos(\theta) - y\sin(\theta), x\sin(\theta) + y\cos(\theta)\big]
\tag{5.10}
$$

(equation 2, page 569, Arnéodo, Decoster & Roux 2000) and $f(x,y) \,\epsilon\, L^2(\mathbb{R}^2, dx\,dy)$ if $||f(x,y)||^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x,y)|^2 dx\,dy < \infty$ (equation 2.1, page 32, Antoine *et al* 2004). The Fourier domain representation of the transform is,

$$
\begin{aligned}
\mathcal{W}_\psi[f(x,y)](a,\mathbf{b},\theta) &= C_\psi^{-\frac{1}{2}} < \Psi_{a,\mathbf{b},\theta}(\omega_1,\omega_2), f(\omega_1,\omega_2) > \\
&= C_\psi^{-\frac{1}{2}} a^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\big[i(\omega_1 b_1 + \omega_2 b_2)\big] \Psi^\star(ar_{-\theta}\,\omega_1, ar_{-\theta}\,\omega_2) f(\omega_1,\omega_2)\,d\omega_1\,d\omega_2
\end{aligned}
$$

$$
\tag{5.11}
$$

where $f(\omega_1, \omega_2)$ is the Fourier transform of the image $f(x, y)$, $\Psi(\omega_1, \omega_2)$ is the Fourier transform of the wavelet function $\psi(x, y)$ and $\boldsymbol{\omega} = (\omega_1, \omega_2)$ are the spatial frequency indices for the x and y directions of the image respectively (equation 5, page 569, Arnéodo *et al* (2000)). The original image $f(x, y)$ can be recovered from the transform using the following relation,

$$f(x, y) = C_\psi^{-\frac{1}{2}} \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \int_{-\infty}^\infty a^{-3} \mathcal{W}_\psi[f(x, y)](a, \mathbf{b}, \theta) \psi_{a,\mathbf{b},\theta}(x, y)\, db_1\, db_2\, d\theta\, da$$

$$(5.12)$$

(equation 8, page 570, Arnéodo, Decoster & Roux 2000). Which holds as long as,

$$C_\psi = (2\pi)^2 \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{|\Psi(\omega_1, \omega_2)|^2}{(\omega_1)^2 + (\omega_2)^2}\, d\omega_1\, d\omega_2 \; < \; \infty \qquad (5.13)$$

(equation 6, page 569, Arnéodo, Decoster & Roux 2000). Equation (5.13) is known as the *admissibility condition*, this condition puts a requirement on the type of wavelet functions that are acceptable for recovery of the original image from the integral wavelet transform. The admissibility condition implies that the wavelet function oscillates around zero,

$$\int_{-\infty}^\infty \int_{-\infty}^\infty \psi(x, y)\, dx\, dy = 0 \qquad (5.14)$$

(equation 7, page 570, Arnéodo, Decoster & Roux 2000). The integral wavelet transform is a redundant representation of an image as it uses four parameters (a, b_1, b_2, θ) to describe an image that is a function of only two parameters (x, y). Nonetheless it provides insight into the structure of the image over a continuum of scales. The choice of analysis function $\psi_{a,\mathbf{b},\theta}(x, y)$ in the integral wavelet transform is very flexible and options include the Mexican Hat, Morlet and Poisson wavelet functions (Vidakovic 1999). The ability to reconstruct the original image from the transform means that the integral wavelet transform preserves the image information.

5.2.2 The Separable Discrete Wavelet Transform

The integral wavelet transform can be made into a discrete version by *critical sampling* of the scale and translation parameters. The sampling is such that any coarser sampling would not allow the original image to be recovered from the coefficients (page 50, Vidakovic 1999). Critical sampling of the scale (a) and translation parameters (b_1, b_2) is specified by,

$$a = 2^j, \ b_1 = k\,2^j, \ b_2 = k\,2^j, \ \ j, k \in \mathbb{Z} \tag{5.15}$$

(adapted from equation 3.6, page 50, Vidakovic 1999). The discrete wavelet transform was used by both Erickson (2005) and Falzon *et al* (2006) to analyse SAXS images of breast tissue. Recall from equation (3.2) the two-dimensional separable discrete wavelet transform is defined for a square digital image, $I(x, y) \in \ell_2(\mathbb{R})$, for a total of J resolution bands as,

$$I(x,y) = \sum_{k_1=1}^{n_{j_0}} \sum_{k_2=1}^{n_{j_0}} c_{j_0,k_1,k_2}\phi_{j_0,k_1,k_2}(x,y) + \sum_{l=1}^{3} \sum_{j \geq j_0}^{J} \sum_{k_1=1}^{n_j} \sum_{k_2=1}^{n_j} d_{j,k_1,k_2}^l \psi_{j,k_1,k_2}^l(x,y) \tag{5.16}$$

which is indexed by scale parameter $j = 1, 2, \ldots, J$, a direction index $l = 1, 2, 3$ which corresponds to the vertical, horizontal and diagonal directions and the location parameters $\mathbf{k} = (k_1, k_2)$ (page 157 Vidakovic 1999). The notation $\ell_2(\mathbb{R})$ denotes the space of square-summable sequences which is the discrete counter-part to the space $L_2(\mathbb{R})$. Data $\mathbf{x}$ belongs to the space of square-summable sequences when the following holds, $\mathbf{x} = \{x_n\} \in \ell_2(\mathbb{R})$ if $\sum_{n \in \mathbb{Z}} |x_n|^2 < \infty$ (page 25 Vidakovic 1999).The separable discrete wavelet transform produces a set of coefficient matrices $\left\{ \mathbf{C}_{J,k_1,k_2}, \mathbf{D}_{j,k_1,k_2}^1, \mathbf{D}_{j,k_1,k_2}^2, \mathbf{D}_{j,k_1,k_2}^3 \right\}$. The coefficients $\mathbf{C}_{J,k_1,k_2}$ describe an approximation to the image found by projecting the image onto the scaling basis functions $\phi_{j_0,k_1,k_2}(x,y)$ and the coefficients $\mathbf{D}_{j,k_1,k_2}^l$ that describe progressive image detail in a certain direction (as indexed by parameter l) with increasing scale, j. These coefficients are found by projecting the image onto the wavelet basis functions $\psi_{j,k_1,k_2}^l(x,y)$. According to the critical sampling requirements the size of the coefficient arrays increase by a power of two for each increase in scale. A wide variety of wavelet bases may be used in the separable discrete wavelet transform including the Daubechies compactly supported, Coiflet and biorthogonal bases (Daubechies 1992; Daubechies 1993; Cohen, Daubechies & Feauveau 1992). Selection of the most appropriate wavelet basis can be tailored to the task at hand. Orthogonal wavelet basis functions are readily available and are useful in many statistical applications because they remove repetition of image information in the transform coefficients.

The discrete wavelet transform and its inverse (which is used to obtain the original image from the transform coefficients) are found in practice by applying a series of discrete filter banks that represent the scaling and wavelet functions. The decomposition filter bank that is used to convert the image data into coefficients is given by,

$$\mathbf{C}_{j+1,k_1,k_2} = \sum_{m=0}^{M}\sum_{n=0}^{N} h_{m-2k_1,n-2k_2} \cdot \mathbf{C}_{j,k_1,k_2} \tag{5.17}$$

$$\mathbf{D}_{j+1,k_1,k_2}^{l} = \sum_{m=0}^{M}\sum_{n=0}^{N} g_{m-2k_1,n-2k_2}^{(l)} \cdot \mathbf{C}_{j,k_1,k_2}^{l} \tag{5.18}$$

and the reconstruction filter bank that is used to recover the image from the coefficients is given by,

$$\mathbf{C}_{j,k_1,k_2} = \sum_{k_1=1}^{n_j}\sum_{k_2=1}^{n_j} h_{m-2k_1,n-2k_2} \cdot \mathbf{C}_{j+1,k_1,k_2}$$

$$+ \sum_{l\,\epsilon\,1,2,3}\sum_{k_1=1}^{n_j}\sum_{k_2=1}^{n_j} g_{m-2k_1,n-2k_2}^{(l)} \cdot \mathbf{D}_{j,k_1,k_2}^{l} \tag{5.19}$$

for image bank filters $h_{m-2k_1,n-2k_2}$ and $g_{m-2k_1,n-2k_2}^{(l)}$ corresponding to the discrete digital representations of the scaling and wavelet functions respectively (adapted from page 158 Vidakovic 1999). The use of filter banks allows rapid calculation of the transform and the ability to recover the original image from the transform coefficients preserves the image information in the same sense as the integral wavelet transform.

5.2.3 The Stationary Wavelet Transform

The separable discrete wavelet transform lacks translation invariance, that is the wavelet transform of a translated image is not equal to a translated version of the transform of the original image $\mathcal{DWT}[I(x-t_1, y-t_2)](a, b_1, b_2) \neq \mathcal{DWT}[I(x,y)](a, b_1-t_1, b_2-t_2)$ for $t_1, t_2 \, \epsilon \, \mathbb{R}$, where the notation $\mathcal{DWT}$ represents the two-dimensional separable discrete wavelet transform. Lack of translation invariance of a transform can be a problem in image recognition tasks. Shifts in the position of a feature within an image can result in substantial differences in magnitude between the coefficients of the original and the translated images. The stationary wavelet transformation overcomes the lack of translation invariance in the separable discrete wavelet transform by omitting the sampling operations of equation (5.15) (Dutilleux 1989). This improvement comes at the cost of increasing the redundancy of the transform. Each scale j of the stationary wavelet transform of an image will contain as many coefficients as there are pixels in the original image.

5.2.4 The Wavelet Packet Transform

The wavelet packet transform produces an orthogonal basis from linear combinations of the wavelet functions that are used in the separable discrete wavelet transform. This produces a more diverse representation of the image over scale and space which might assist in the identification of particular image features such as rings or lines. The wavelet packet transform is defined by a collection of packet functions, $\mathcal{W}_{j,n,k_1,k_2}(x,y)$ that obey the following relationship,

$$\mathcal{W}_{j,n,k_1,k_2}(x,y) = 2^{\frac{j}{2}}\mathcal{W}_n(2^j x - k_1, 2^j y - k_2) \tag{5.20}$$

for scale parameter $j = 0, 1, \ldots, J-1$, location parameters k_1 and $k_2 \, \epsilon \, \mathbb{Z}$ and index parameter $n \, \epsilon \, \mathbb{N}$ (equation 5.13, page 135, Vidakovic 1999). The packet functions $\mathcal{W}_{j,n,k_1,k_2}(x,y)$ are defined recursively according to the following equations,

$$\mathcal{W}_{2n}(x,y) = \sum_{i=0}^{N_h} h_k \sqrt{2}\mathcal{W}_n(2x - k_1, 2y - k_2) \tag{5.21}$$

$$\mathcal{W}_{2n+1}(x,y) = \sum_{i=0}^{N_g} g_k \sqrt{2}\mathcal{W}_n(2x - k_1, 2y - k_2) \tag{5.22}$$

where $\mathcal{W}_0(x,y) = \phi_0(x,y)$, $\mathcal{W}_1(x,y) = \psi_0(x,y)$, the weights h_i and g_i are coefficients of the quadrature mirror filters (**h** and **g**) that are used in the two-dimensional separable discrete wavelet transform. The number of coefficients in these filters is determined by using $|\mathbf{h}| = N_h$

and $|\mathbf{g}| = N_g$ (adapted from pages 56 and 133 and equation 5.12, page 135 Vidakovic 1999). An orthonormal basis of $L_2(\mathbb{R}^2)$ can be found from the family of wavelet packet functions and a variety of different combinations of parameters are available for the user to select that produces different image representations. Image analysis in a non-orthonormal basis is also possible by projecting the image $I(x, y)$ onto each of a collection of wavelet packet functions $\mathcal{W}_{j,n,k_1,k_2}(x, y)$ by using the inner product in order to obtain a set of transform coefficients $\mathbf{D}_{j,n,k_1,k_2} = < I(x, y), \mathcal{W}_{j,n,k_1,k_2} >$. The wavelet packet transform offers a flexible alternative to the separable discrete wavelet transform. It is used as a component in the best orthogonal basis algorithm which is an adaptive image transform which is discussed in Section 5.4.1.

5.3   Second generation Wavelet-like Transforms

A range of second-generation scale-space image transformations also exist that provide alternatives to the integral, discrete and stationary wavelet transforms. Prominent examples include the steerable pyramid, curvelet and contourlet transforms. All of these transformations are characterised by analysis functions that vary across position, scale and direction/orientation. These image transforms will be the subject of this section. The steerable pyramid transform is examined in Section 5.3.1, the curvelet transform in Section 5.3.2 and finally the contourlet transform in Section 5.3.3.

5.3.1   The Steerable Pyramid Transform

The steerable pyramid transform is a multi-scale, multi-directional representation of an image. It is formed by a set of filters that are rotated copies of each other and have the property that a filter at any orientation can be computed as a linear combination of filters at other orientations. These filters are defined in Fourier space for the ith orientation out of $\mathcal{I}$ direction bands as,

$$B_i(\boldsymbol{\omega}, \theta) = A(\theta - \theta_i)B(\omega^\star) \tag{5.23}$$

for the frequency indices $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $\omega^\star = |\boldsymbol{\omega}|$ which are defined for parameters $-\pi < \omega_1, \omega_2 < \pi$ and with the orientation specified by using the parameter $\theta = \tan^{-1}(\frac{\omega_2}{\omega_1})$ relative the reference orientation $\theta_i = 2\pi/i$ (page 445, Simoncelli & Freeman 1995). Because the filter $B_i(\boldsymbol{\omega}, \theta)$ is separable into a radial $B(\omega^\star)$ and angular $A(\theta - \theta_i)$ component, it is sufficient to describe this two-dimensional function in terms of one-dimensional functions. The angular component, $A_i(\theta) = A(\theta - \theta_i)$ of the filter is given by,

$$A_i(\theta) = \begin{cases} \alpha_\mathcal{I} \left[\cos(\theta - \frac{\pi i}{\mathcal{I}}) \right]^{\mathcal{I}-1} & |\theta - \frac{\pi i}{\mathcal{I}}| < \frac{\pi}{2} \\ 0 & \text{, otherwise} \end{cases} \tag{5.24}$$

where,

$$\alpha_\mathcal{I} = 2^{i-1} \frac{(\mathcal{I}-1)!}{\sqrt{\mathcal{I}[2(\mathcal{I}-1)]!}}. \tag{5.25}$$

Whilst the radial component of the filter, $B(\omega^\star)$ is given by,

$$
B(\omega^\star) = \begin{cases} \cos\left[\frac{\pi}{2}\log_2(\frac{2\omega^\star}{\pi})\right] & \frac{\pi}{4} < \omega^\star < \frac{\pi}{2} \\ 0 & \omega^\star \leq \frac{\pi}{4} \\ 1 & \omega^\star \geq \frac{\pi}{2}. \end{cases}
\tag{5.26}
$$

(page 55, Portilla & Simoncelli 2000).

The transform is defined using a recursive algorithm that involves three types of filters, a low-pass filter $L_{j-1}(\boldsymbol{\omega}, \theta) = L_{j-1}(\omega_1)L_{j-1}(\omega_2)$, a high pass filter $H_0(\boldsymbol{\omega}) = H_0(\omega_1)H_0(\omega_2)$ and bandpass filters $B(\omega^\star)$. The lowpass filter, $L_{j-1}(\omega_i)$ at scale j is defined as,

$$
L_{j-1}(\omega_i) = \begin{cases} 2\cos\left[\frac{\pi}{2}\log_2(\frac{2\omega_i}{\pi})\right] & \frac{\pi}{4} < \omega_i < \frac{\pi}{2} \\ 2 & \omega_i < \frac{\pi}{4} \\ 0 & \omega_i \geq \frac{\pi}{2}. \end{cases}
\tag{5.27}
$$

for $(i = 1, 2)$, $(j = 2, \ldots, J + 1)$ (page 55, Portilla & Simoncelli 2000). At scale $j = 1$, we specifically set $L_0(\boldsymbol{\omega}) = L_1(\boldsymbol{\omega}/2)$ in order to define the initial low-pass filter of the transform. Similarly we define the high-pass filters $H_0(\omega_1)$ and $H_0(\omega_2)$ as $H_0(\omega_1) = H_0(\omega_2) = B(\omega^\star/2)$.

The low-pass filter, $L_0(\boldsymbol{\omega})$ produces a smooth approximation to the original image, whilst subsequent applications of the bandpass filter, $B_i(\boldsymbol{\omega}, \theta)$ extract increasingly finer details from the image. The response of filters $L_0(\omega_1)$, $H_0(\omega_1)$ and $B(\omega^\star)$ are displayed as a function of ω_1 in Figures 5.3.1(a)-(c) and the response of filter $A_2(\theta)$ is displayed in Figure 5.3.1(d) as a function of θ.

a) **Low-pass Filter** $L_0(|\omega_1|)$ b) **High-pass Filter** $H_0(|\omega_1|)$

c) **Band-pass Filter** $B(\omega^\star)$ d) **Direction Filter** $A_2(\theta)$
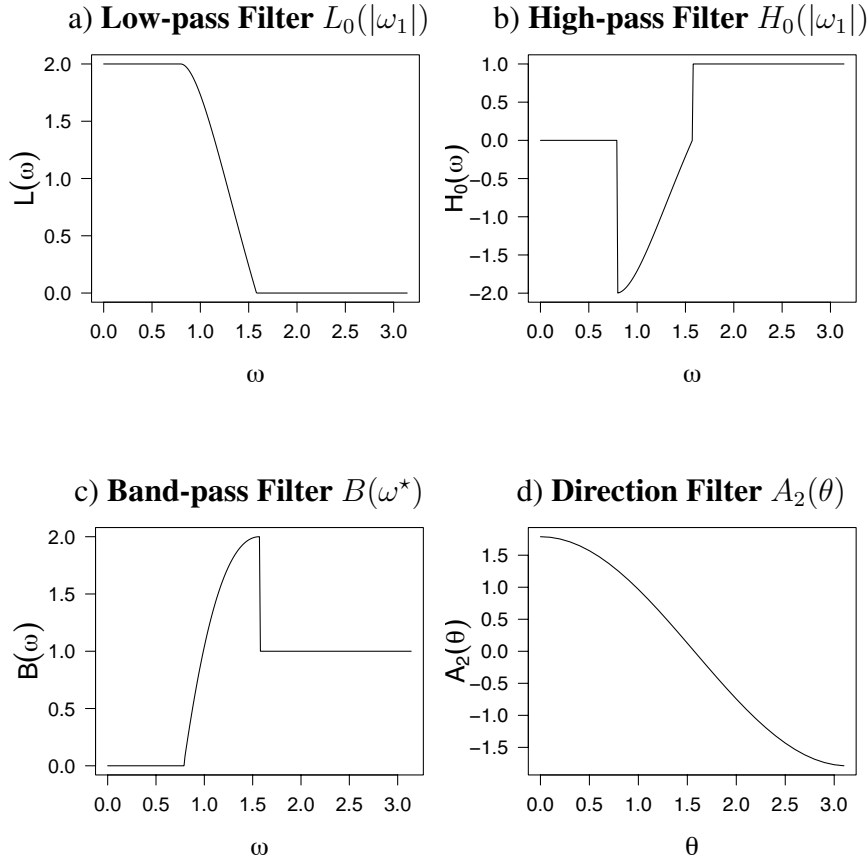

Figure 5.1: Filters used in the Steerable Pyramid Transform: (a) the low-pass filter $L_0(|\omega_1|)$ attenuates all of those frequencies greater than $|\omega_1| \approx 1$ and omits those frequencies $|\omega_1| > 1.6$, (b) the high-pass filter $H_0(|\omega_1|)$ removes the frequencies $|\omega_1| < 0.75$, applies negative weights to those frequencies $0.75 \leq |\omega_1| \leq 1.6$ and leaves unchanged those frequencies $|\omega_1| > 1.6$, (c) the band-pass filter $B(\omega^\star)$ removes those frequencies $\omega^\star < 0.75$, enhances those frequencies in the range $0.75 \leq \omega^\star \leq 1.6$ and leaves unchanged those frequencies $\omega^\star > 1.6$, (d) the direction filter $A_2(\theta)$ enhances data in the directions $0 \leq \theta < 1.0$ radians, attenuates data in the directions in the directions $1 \leq \theta \leq 2$ radians and down-weights data in the directions $2 < \theta < 3$.

The algorithm to implement the steerable pyramid transform can be described as follows:

a) Apply the filters $L_0(-\boldsymbol{\omega})$ and $H_0(-\boldsymbol{\omega})$ to the image $I(x, y)$.

b) Retain the coefficient matrix $\mathbf{D}_0$ which is obtained by filtering the image $I(x, y)$ with the high-pass filter $H_0(-\boldsymbol{\omega})$, this matrix corresponds to coefficients at the first scale of detail.

c) Apply the filters $B_i(-\boldsymbol{\omega}, \theta)$ $(i = 1, \ldots, \mathcal{I})$ and $L_1(-\boldsymbol{\omega})$ to the coefficient matrix $\mathbf{C}_0$ which is obtained by filtering the image $I(x, y)$ with the filter $L_0(-\boldsymbol{\omega})$.

d) Retain the coefficient matrix $\mathbf{D}_{1,i}$ which is produced by the application of the filter $B_i(-\boldsymbol{\omega}, \theta)$ to the matrix $\mathbf{C}_0$. This matrix corresponds to coefficients at the second scale of detail in the ith direction.

e) Down-sample by two the coefficient matrix $\mathbf{C}_1^+$ which is produced by applying the filter $L_1(-\boldsymbol{\omega})$ to the coefficient matrix $\mathbf{C}_0$. Down-sampling is achieved by removing the entries along every second row and second column. Denote this down-sampled matrix as $\mathbf{C}_1$.

f) Apply steps (iii)-(v) above recursively for the filter functions and coefficient matrices corresponding to scales $j = 2, \ldots, J$.

The low-pass filter $L_0(\boldsymbol{\omega})$ retains only the low-frequency information in the image, with the high-frequency information being modulated towards zero. The high-pass filter $H_0(\boldsymbol{\omega})$ is used in the initial step of the transform to retain a high-frequency residual band from the original image. This filter is not applied to any other scales. The bandpass filter, $B_i(\boldsymbol{\omega}, \theta)$ retains information only in a select frequency-direction band, multiple applications of this filter partition the image into different frequency-direction bands. The direction filter $A_2(\theta)$ detects image features at a particular orientation it is the part of the bandpass filter $B_2(\boldsymbol{\omega}, \theta)$ that provides the transform with direction selectivity. Like the two-dimensional separable discrete wavelet transform, application of the filter banks of the steerable pyramid transform produces a set of coefficients that describe the image over multiple scales and directions.

The original image can be recovered from the set of steerable pyramid transform coefficients by applying the steps of the algorithm specified above in reverse. The transform preserves the image information but expresses it redundantly, with the number of coefficients being greater than the number of pixels in the original image by a factor of $4\mathcal{I}/3$ (Simoncelli & Freeman 1995). Any number of directional bands can be accommodated (as specified by the user), so

the level of redundancy can be quite high for a large number of directions, $\mathcal{I}$. Nonetheless, these directional subbands (that correspond to different $B_i(\boldsymbol{\omega}, \theta)$ filters) are essentially alias free unlike the two-dimensional separable discrete wavelet transform whose directional bands within a scale are aliased (Simoncelli & Freeman 1995). Spatial frequencies in certain image directions can therefore be more reliably associated with different coefficient matrices in the steerable pyramid transform as compared to the separable discrete wavelet transform. This makes the steerable pyramid transform a useful addition to the image analysis tool kit.

5.3.2   The Curvelet Transform

The curvelet transform is a multi-scale, multi-directional image decomposition designed to provide optimal sparse representations of certain types of images. The specific types of the images that the curvelet transform is designed to provide optimal representations are those that are C^2 smooth (functions that are bounded in magnitude and have both a first and second-derivative) except at discontinuities that are themselves described as C^2 edges. In other words, the specific digital images of interest can be segmented into regions that correspond to 'smooth' and 'edge' components which are at least twice differentiable and bounded in magnitude. The basis functions otherwise known as curvelets ($\gamma_{j,l_j,k_1,k_2}(x,y)$) used in the transform are defined as,

$$\gamma_{j,l_j,k_1,k_2}(x,y) = 2^{\frac{3j}{2}}\gamma\left(\mathbf{D}_j r_{\theta_{j,l_j}}(x,y) - \mathbf{k}_\delta\right) \tag{5.28}$$

for the scale parameter $j = 0,\ldots,J-1$, the orientation parameter $l_j = 0,1,\ldots,2^j$, the translation parameters $k_1, k_2, \ \epsilon\ \mathbb{Z}$ and the *parabolic scaling matrix* $\mathbf{D}_j$ which is defined as,

$$\mathbf{D}_j = \begin{pmatrix} 2^{2j} & 0 \\ 0 & 2^j \end{pmatrix} \tag{5.29}$$

the rotation operator $r_{\theta_{j,l_j}}(x,y)$ (refer to equation 5.10) along with the *rotation angle* $\theta_{j,l_j} = 2\pi 2^{-j} l_j$ and the *translation vector* $\mathbf{k}_\delta = (k_1 \cdot \delta_1, k_2 \cdot \delta_2)$ for $\delta_1, \delta_2 > 0$ (adapted from equations 1.4 and 1.5, page 223, Candés & Donoho 2003). The curvelet basis functions vary across scale, position and orientation. In fact, the curvelet function $\gamma_{j,l_j,k_1,k_2}(x,y)$ has a length $\approx 2^{-j}$ and a width $\approx 2^{-2j}$. Hence the curvelet functions obey a parabolic scaling relation whereby at each scale $j\ \epsilon\ (0,\ldots,J-1)$, the width of the curvelet function $\gamma_{j,l_j,k_1,k_2}(x,y)$ approximately equals the square of its length.

The curvelet transform is performed by projecting the image $I(x,y)$ onto each of the curvelet basis functions $\gamma_{j,l_j,\mathbf{k}}(x,y)$ by taking the inner product $< I(x,y), \gamma_{j,l_j,\mathbf{k}}(x,y) >$ across the range of the scale, orientation and location parameters. A low-pass band is also found by projecting the image $I(x,y)$ onto a Lemarié scaling function $\phi_{k_1,k_2}(x,y)$ from the Meyer wavelet basis (page 12, Candès & Donoho 2003). These projections produce a set of coefficients, $\left\{\mathbf{C}_0, \mathbf{D}_{1,l_1}, \mathbf{D}_{1,l_2}, \ldots, \mathbf{D}_{J-1,2^{J-1}}\right\}$ that describe the image $I(x,y)$ in terms of scale, direction and location.

The original image $I(x, y)$ (of dimensions (N_1, N_2)) can be recovered from the curvelet coefficients using,

$$I(x,y) = \sum_{j=0}^{J-1} \sum_{l_j=0}^{2^j} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} < f(x,y), \gamma_{j,l_j,k_1,k_2}(x,y) > \gamma_{j,l_j,k_1,k_2}(x,y) \qquad (5.30)$$

(page 13, Candés & Donoho 2000).

The squared error $||I(x,y) - I_n(x,y)||^2$ of the n-term partial reconstruction, $I_n(x,y)$ is found by taking the curvelet transform of $I(x,y)$, selecting the n largest magnitude curvelet coefficients and applying the inverse curvelet transform. For certain types of images (those that can be modeled as being as C^2 smooth except at discontinuities which are also C^2 smooth), the curvelet transform captures the key image features with relatively few coefficients. The squared error $||I(x,y) - I_n(x,y)||^2$ of the n-term partial reconstruction $I_n(x,y)$ converges as n^{-2}, as compared to the wavelet transform which converges as n^{-1} and the Fourier transform which converges as $n^{-\frac{1}{2}}$ as $n \rightarrow \infty$ (page 221, Candés & Donoho 2003). This gives the curvelet transform a distinct advantage over both the wavelet and Fourier transforms in representing such images.

The curvelet transform is a very redundant representation of a digital image with the ratio of the number of coefficients to image pixels being $16J+1$, for a total of J scales in the transform (page 88, Do & Vertelli 2003). Nonetheless, the curvelet transform offers very sparse, close to optimal representations of those images that can be modeled as being as C^2 smooth except at discontinuities which are also C^2 smooth.

5.3.3   The Contourlet Transform

The contourlet transform is a discrete multi-scale, multi-direction and local image transform. It applies a Laplacian pyramid (multi-scale) filter to separate the image into different resolution sub-band matrices, followed by a directional filter on each of these sub-bandmatrices (Burt & Adelson 1983; Bamberger & Smith 1992; Do & Vertelli 2006). The Laplacian pyramid detects changes in image intensity (edges) and the directional filter detects the orientation and smoothness along such edges.

The first step in the Laplacian pyramid is to apply a low-pass filter to the original image $I(x, y)$, this filtered image is then down sampled by removing every second pixel along the rows and columns. This low-pass filter is again applied to the down-sampled filtered image, followed again by further down-sampling. This filtering process continues in an iterative manner for all J scales to produce a set of low-pass filtered images which are denoted $[I_1(x, y), \ldots, I_J(x, y)]$. This set of low-pass filtered images is called the *Gaussian pyramid*. Different low-pass filters can be used to generate the Gaussian pyramid but a filter that resembles a bivariate Gaussian function is most often used.

i) Low-pass filter original image $I(x, y)$ to obtained smoothed image $G_1(x, y)$. The number of elements in the image $I(x, y)$ is reduced by a quarter in this smoothing operation.

ii) Low-pass filter image $G_1(x, y)$ to get smoother image $G_2(x, y)$. This filtering is achieved according to the following formula,
$G_2(x, y) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m, n) G_1(2x + m, 2y + n)$ (equation 1, page 533, Burt & Adelson 1983), $w(m, n)$ is a weighting function, otherwise known as a generating kernel. A range of generating kernels can be selected but often a generating kernel resembling a Gaussian function is used.

iii) Continue iterative smoothing process for a total of J times, to get sequence of smoothed images, $G_1(x, y), G_2(x, y), \ldots, G_J(x, y)$, which is called the Gaussian pyramid.

iv) Generate the expanded Gaussian pyramid, $G_1^{\mathrm{e}}(x, y), G_2^{\mathrm{e}}(x, y), \ldots, G_J^{\mathrm{e}}(x, y)$ by interpolating between each entry along the rows and columns of the Gaussian pyramid $G_1(x, y), G_2(x, y), \ldots, G_J(x, y)$. The expansion is done according to the formula,
$G_{J+1}(x, y) = 4 \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m, n) G_J(\frac{x-m}{2}, \frac{y-n}{2})$ where only terms for which $\frac{x-m}{2}$ and $\frac{y-n}{2}$ are included in this sum (equation 2, page 534, Burt & Adelson 1983).

(v) Each matrix in the Laplacian pyramid is generated by the difference between the smoothed images in the expanded Gaussian pyramid at a particular scale and the smoothed images in the Gaussian pyramid at the previous scale.

That is,

$$
\begin{aligned}
L_0(x,y) &= G_1^{\mathrm{e}}(x,y) - I(x,y), \\
L_1(x,y) &= G_2^{\mathrm{e}}(x,y) - G_1(x,y), \\
&\vdots \qquad\qquad\qquad \vdots \\
L_J(x,y) &= G_{J+1}^{\mathrm{e}}(x,y) - G_J(x,y).
\end{aligned}
$$

The second step in the contourlet transform is to apply a directional filter to each matrix produced by the Laplacian pyramid. A wide range of directional filters can be used the key requirement being that the Laplacian pyramid matrix can be recovered from the directional filter coefficients. The 'pkva' directional filters are particularly useful as they high spatial frequency selectivity whilst still allowing perfect reconstruction (Phoong *et al* 1995).

The contourlet transform uses analysis functions of the form,

$$
\left\{ \phi_{j_0,k_1,k_2}(x,y), \rho_{j,l_j,k_1,k_2}(x,y) \right\} \tag{5.31}
$$

for scale parameter $j = 0, \ldots, J-1$, orientation parameter $0 \leq l_j \leq 2^{l_j} - 1$, and location parameters $0 \leq k_1 \leq N_1 - 1$ and $0 \leq k_2 \leq N_2 - 1$ for an image $I(x,y)$ of dimensions N_1 x N_2 (equation 4.3.1, page 18, Do & Vertelli 2003). The scaling filter $\phi_{j_0,k_1,k_2}(x,y)$ provides a low resolution analysis of the image $I(x,y)$, whilst the directional filters $\rho_{j,l_j,k_1,k_2}(x,y)$ analyse the detail of the image $I(x,y)$ in a certain direction.

The transform produces a low-pass coefficient matrix $\mathbf{C}_0 = < I(x,y), \phi_{j_0,k_1,k_2}(x,y) >$ by taking the inner product of the image and the scaling filter $\phi_{j_0,k_1,k_2}(x,y)$. Similarly, a set of detail coefficients indexed by scale and direction $\mathbf{D}_{j,l_j} = < I(x,y), \rho_{j,l_j,k_1,k_2}^{l_j}(x,y) >$ are found by taking the inner product of the image and the detail filters $\rho_{j,l_j,k_1,k_2}^{l_j}(x,y)$. Like the wavelet and curvelet transforms, the original image can be recovered perfectly from these contourlet transform coefficients.

The contourlet transform offers a sparse representation of images with edges and a flexible user choice of the number of directions per scale. This transform is specifically designed for discrete grids of bivariate data. This can be a significant advantage over the curvelet transform when analysing digital images or two dimensional histograms. The curvelet transform is designed specifically for the continuous space bivariate data.

The contourlet transform is far less redundant than the curvelet transform. There will be no more than 4/3 times more contourlet transform coefficients than there are entries (or pixels) in the original bivariate data (image) (page 102, Do & Vertelli 2003). The contourlet filter-bank structure is also very flexible and has the potential to use a range of filters (analysis functions) in both the multi-scale and directional steps.

5.4 Adaptive Image Transformations

A variety of adaptive image transforms exist that use a library of functions taken from a range of image filters. Prominent functions in these libraries are those from the Fourier and wavelet families of transformations. Adaptive image transforms seeks to match the most suitable function to the image, with the most appropriate function being measured by a cost functional. The function that minimises the cost functional is the one selected for use in the adaptive image transformation. The use of a range of different functions in the transform offers alternatives to the more standard fixed basis transforms such as the wavelet and curvelet transforms. Selection of the functions from the library can be computationally demanding and a number of algorithms have been developed to meet this challenge. Section 5.4.1 examines the best orthogonal basis algorithm which is often used in combination with wavelet packets. The matching pursuit algorithm is discussed in Section 5.4.2 and its alternative, the basis pursuit algorithm in Section 5.4.3. Both the matching and basis pursuit algorithms can utilize a large number of library functions but can be time consuming to compute for images.

5.4.1 The Best Orthogonal Basis Algorithm

The best orthogonal basis algorithm uses orthogonal functions from a library of such functions to represent an image. The functions used in the library are often derived from wavelet packet functions (Section 5.2.4). A library of orthogonal functions is very useful for image compression applications as the energy of noise in the transform cannot exceed the energy of the noise in the original image (Coifman & Wickerhauser 1992). This property assists with the reliable transmission of image information. The use of a library of orthogonal functions with which to represent the image further assists with this task. Those functions that best compress the image information into a few coefficients can be selected from the library using a cost functional, $\mathfrak{C}$. This cost functional is a criterion with which to compare functions in the library, $\mathfrak{L}$. A functional based upon Shannon's entropy is often used, which for a discrete probability distribution $\{p_1, p_2, \ldots, p_n\}$ is given by

$$\mathfrak{S}(\mathbf{x}) = -\sum_{i=1}^{n} p_i \log p_i \tag{5.32}$$

for a coefficient vector $\mathbf{x} = (x_1, \ldots, x_n)$ where probability $p_i = \frac{|x_i|^2}{||x||^2}$ and $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$ (adapted from equation 5.16, Vidakovic 1999).

The cost functional based upon Shannon's entropy $\mathfrak{S}(\mathbf{x})$ is defined as,

$$\mathfrak{C}_\mathfrak{S}(\mathbf{x}) = -\sum_{i=1}^{n} |x_i|^2 \log |x_i|^2 \qquad (5.33)$$

(equation 5.17, page 142, Vidakovic 1999). Minimisation of the cost functional $\mathfrak{C}_\mathfrak{S}(\mathbf{x})$ in equation (5.33) is equivalent to minimisation of Shannon's entropy. The notion of Shannon's entropy is deeply rooted in information theory and concerns (among other things) the minimum amount of information needed to describe a signal (image). Selecting the function from the library $\mathfrak{L}$ minimises Shannon's entropy is equivalent to selecting the function that best compresses the signal. The best basis algorithm selects the best function from the library for all scales j and locations $\mathbf{k} = (k_1, k_2)$ in line with the wavelet packet indexing. The algorithm can be organised into a binary tree structure that allows for fast ($\mathcal{O}(N \log N)$ for N image pixels) (page 713 Coifman & Wickerhauser 1992). The best orthogonal basis algorithm does not always perform well for highly non-stationary images (images whose statistical properties are dependent upon location), the selection of the most appropriate functions can be driven by those transient features of highest energy (such as high intensity image edges). Other transient features might be of equal importance but that have small energy can be inadvertently suppressed using the best orthogonal basis algorithm. Mallat & Zhang (1993) claim that the best orthogonal basis algorithm is better suited to those images with stationary statistical properties and use this as one reason to advocate the use of their matching pursuit algorithm which is the subject of the next section.

5.4.2   The Matching Pursuit Algorithm

The matching pursuit algorithm is also used to decompose the image into a linear combination of functions that are selected from a library. It offers a flexible image decomposition and might be useful in those cases where a single fixed basis, filter or transform does not provide sufficient insight into the structure of the image. The matching pursuit algorithm is based upon successive approximations of an image $f(x, y) \in L_2(\mathbb{R}^2)$ via orthogonal projections onto the functions $\mathfrak{G}_{j,k_1,k_2}(x, y)$ from the library $\mathfrak{L}$. For a function $\mathfrak{G}_{j,k_1,k_2}(x, y) \in \mathfrak{L}$ the image is decomposed into,

$$f(x, y) = <f(x, y), \mathfrak{G}_{j,k_1,k_2}(x, y) > \mathfrak{G}_{j,k_1,k_2}(x, y) + R(x, y) \qquad (5.34)$$

where $R(x, y)$ is the residual image remaining after approximating the image $f(x, y)$ using the function $\mathfrak{G}_{j,k_1,k_2}(x, y)$ (equation 7, page 3340, Mallat & Zhang 1993). Since $\mathfrak{G}_{j,k_1,k_2}(x, y)$ is

orthogonal to $R(x, y)$ the energy of this projection can be split into two components,

$$||f(x,y)||^2 = | < f(x,y), \mathfrak{G}_{j,k_1,k_2}(x,y) > |^2 + ||R(x,y)||^2 \qquad (5.35)$$

(equation 8, page 33401, Mallat & Zhang 1993). Because $||f(x,y)||^2$ is fixed the residual $||R(x,y)||$ can be minimised by selecting the function $\mathfrak{G}_{j,k_1,k_2}(x,y) \; \epsilon \; \mathfrak{L}$ such that the inner product $| < f(x,y), \mathfrak{G}_{j,k_1,k_2}(x,y) > |$ is maximum. Finding the optimum function that minimises $||R(x,y)||$ can be computationally demanding for libraries with a large number of functions and a compromise that is close to an optimal choice can be found using,

$$| < f(x,y), \mathfrak{G}_{j,k_1,k_2}(x,y) > | \; > \; \alpha \sup_{\mathfrak{G}_{j,k_1,k_2}(x,y) \; \epsilon \; \mathfrak{L}} | < f(x,y), \mathfrak{G}_{j,k_1,k_2}(x,y) > | \qquad (5.36)$$

where $\alpha \; \epsilon \; (0,1]$ is an optimality factor that controls the rate of acceptance of the function $\mathfrak{G}_{j,k_1,k_2}(x,y)$ in the library. Relaxing the criterion (reducing α) for function selection significantly reduces the computational complexity of the problem (page 326, Vidakovic 1999). The matching pursuit algorithm involves recursive applications of the selection of the 'best' function from the library to represent the image. It decompose an image $f(x,y)$ into the form,

$$f(x,y) = \sum_{m=0}^{M-1} < R^m(x,y), \mathfrak{G}_{j,k_1,k_2}^m(x,y) > \mathfrak{G}_{j,k_1,k_2}^m(x,y) + R^M(x,y) \qquad (5.37)$$

for a total of M iterations (orthogonal projections) where the initial residual is $R^0(x,y) = f(x,y)$ (page 326, Vidakovic 1999). The iterations are halted when $||R^{m+1}(x,y)||^2 \; < \; \epsilon^2 ||f(x,y)||^2$, for a pre-defined tolerance criterion $\epsilon < 1$.

The matching pursuit algorithm selects at each iteration the function (from the library) best adapted to approximate a region of the image. As such the matching pursuit algorithm only looks one step ahead in each iteration and may miss the optimal decomposition. Bergeaud & Mallat (1995) applied the matching pursuit algorithm to images using Gabor functions and developed a fast matching pursuit algorithm developed specifically for images that has a complexity of $\mathcal{O}(N \log^2 N)$ per iteration for an image of $N \times N$ pixels. This suboptimal approach to the function selection is not desirable and the basis pursuit algorithm has been proposed in an attempt to produce an even sparser image representation (Chen, Donoho & Saunders 2001).

5.4.3　Basis Pursuit

Basis pursuit considers a linear image representation of the form,

$$f(x,y) = \sum_{\boldsymbol{\gamma}} \upsilon_{\boldsymbol{\gamma}} \mathfrak{G}_{\boldsymbol{\gamma}}(x,y) \tag{5.38}$$

where the vector $\boldsymbol{\gamma}$ describes a collection of indices that are appropriate to describe the function $\mathfrak{G}_{\boldsymbol{\gamma}}(x,y)$. For instance, $\boldsymbol{\gamma}$ may index scale and location or alternatively scale, direction and location. The coefficients $\upsilon_{\boldsymbol{\gamma}}$ describe the contribution of each function $\mathfrak{G}_{\gamma}(x,y)$ from a library of such functions. A large range of functions can be used in the library and the main requirement is that they can be specified in discrete form. This allows a finite dictionary that consists of K functions to be described using a three-dimensional array $\boldsymbol{\Psi}$. The two-dimensional matrices that describe each function are stacked along the third dimension of this array. The image decomposition problem can then be represented as,

$$\mathbf{f} = \boldsymbol{\Psi}\boldsymbol{\upsilon} \tag{5.39}$$

where $\boldsymbol{\upsilon} = (\upsilon_{\boldsymbol{\gamma}})$ is the vector of coefficients in equation (5.38) and $\mathbf{f}$ is the discrete matrix representation of the image $f(x,y)$ (adapted from equation 2.2, page 134, Chen, Donoho & Saunders 2001). A cost functional is selected that minimises the ℓ_1 norm of the coefficient vector $\boldsymbol{v}$, it is of the form $\mathcal{C}_{\ell_1} = |\boldsymbol{v}|_1 = \sum_{i=1}^{K} |\boldsymbol{v_i}|$. A cost functional based upon a minimal ℓ_1 norm of the coefficient vector $\boldsymbol{v}$ is driven by the desire to achieve a sparse image representation.

Basis pursuit is principle of global optimisation rather than a specific algorithm, the key concept is to find a linear representation such that the coefficients have minimal ℓ_1 norm. This optimisation problem is computationally demanding that has been tackled using linear programming techniques such as those described in Gill, Murray & Wright (1991). Chen, Donoho & Saunders (page 140, 2000) describe two such algorithms, the BP-Simplex and the BP-Interior that use linear programming to solve the basis pursuit problem for one-dimensional signals. A range of alternative algorithms could also be developed using linear programming for use in basis pursuit. Despite the success of linear programming techniques in the application of basis pursuit to one-dimensional signals, the routine application of the basis pursuit methods to image analysis problems is quite limited. The global optimisation problem is on larger scale for images and is difficult to compute in practice. Advances in global optimisation algorithms may overcome this obstacle in the future and if so the basis pursuit method might become a practical option for image analysis.

5.5 Summary and Conclusions to the Survey of Frequency and Time-Frequency Image Transforms

This chapter has surveyed the a range of spectral and scale-space image transforms. The discrete, windowed and discrete cosine transforms are designed to analyse the spatial frequency content of the image. The windowed Fourier transform is of particular interest as it analyses the spatial frequency structure in a local manner. The integral, discrete and stationary wavelet transforms adds an additional layer to the analysis allowing the image to be examined over a range of resolutions. These transforms are very useful in the detection and description of transient events such as image edges. Second generation wavelet-like transforms such as the steerable pyramid, curvelet and contourlet transforms are designed to represent an image (especially those with edges) with as few large magnitude coefficients as possible. The sparse image representation (one that has relatively few large magnitude coefficients) offered by these transforms might be useful in image compression and analysis applications. Associating image features with a few coefficients might reduce the complexity of any statistical model used to analyse the image. Adaptive image transforms that develop a description of the image using a library of functions. These transforms offer flexible representation of the image and can be designed to a specific task at hand. The computational complexity of selecting the optimal set of functions from the library to represent the image is a challenge that needs to be met in order for these methods to reach their full potential. The variety of transforms available to the researcher is quite large and allows for a number of different approaches to the image analysis problem. Selection of the most appropriate problem depends on the task at hand and the focus of the analysis. The nature and structure of the image that is under study is important, the Fourier transform seems more appropriate for images consisting of smooth periodic components whilst images with sharp edges seem better suited to analysis with the curvelet transform. This challenge of selecting the most appropriate image transform must be met for the analysis of SAXS images and will be the subject of later chapters.