Chapter 5

A Classification System

5.1 Introduction

The goal of this thesis is the development of a rough surface classifier which operates on imaged texture, yet is robust to changes in illuminant tilt between training and classification. The aim of this chapter is to develop a classifier that operates solely on the basis of image texture and does not apply any domain knowledge of the underlying physical system.

The first part of this thesis is concerned with modelling the transition from physical surface to symbolic representation. In the previous chapters we have modelled the physical processes of image formation and data extraction. In this chapter, we will develop the classifier. The classifier forms the final link in the surface to symbol chain, extracting the symbolic representation from the textured image. This task corresponds to surface classification where the imaging conditions are held constant throughout training and classification. In the next chapter the classifier developed here will be modelled, and the effects of varying illuminant tilt considered.

The combination of algorithms, which we describe collectively as a classifier, consist of: a means of extracting the relevant signal components, a mechanism to process these components and a discriminatory mechanism to classify on the basis of this information. In this chapter we will consider all of these components, however, the emphasis will be on the first stage, that is, of signal extraction. The chapter, consists of two parts: firstly, a review of feature extraction techniques and the selection of the technique thought to be most suitable for this thesis, and secondly a description of its incorporation into a complete classifier.

5.2 Terms used in this chapter

What we mean by classification.

Ehrich and Foith [Ehrich77] define three tasks associated with texture analysis,

1. given a textured region, to which of a finite number of classes does the sample belong;
2. given a textured region, how can it be described; and
3. given a scene, how can boundaries between the major textured regions be established.

Reed and Wechsler [Reed90] quote this analysis and designate the tasks as, *classification, description* and *segmentation* respectively. We are concerned with classification and segmentation.

Unfortunately various authors give differing definitions of these terms. Tuceryan and Jain [Tuceryan92] state that "*the goal of texture classification is to produce a classification map of the input image, where each uniform region is identified with the texture class it belongs to*." This seems to imply that segmentation is taking place, if only implicitly. Chellappa et al. [Chellappa92] on the other hand, use a stricter definition: ".. *standard pattern classification techniques may be applied assuming there is only one texture in the image*." Tuceryan and Jain contrast classification with segmentation: "*The goal of texture segmentation is to separate regions in the image which have different textures and identify boundaries between them. The textures themselves need not be recognised*." [Chen, p237]. Subsequently he enlarges on this, stating that there are two general approaches; boundary-based approaches which rely on detecting differences in textures and region-based approaches which grow and merge uniform regions of texture. Clearly these definitions are incompatible, this is due to the fact that Jain bases his definitions on methodology, whereas the conflicting definitions are set in terms of objectives. We believe that all the above definitions have their own merits, but, for clarity we will adopt terms suited to the subject matter of this thesis.

We will define the purpose of our system to be *classification.* We define this term in the context of the hypothetical example given in Chapter 1; consider an inspection system applied to a rough surface, which may exhibit several different forms of roughness within the same sample. The inspection system must be able to detect and identify

different textures and the boundaries between them. The approach adopted in this thesis is to calculate the probability that each individual pixel belongs to a certain class, on the basis of a feature vector, and to label that pixel as belonging to the most likely class.

As with Jain, segmentation of the image is implicit in our definition of classification. Jain's definition of segmentation, however, is not relevant to the approach used in this thesis. We relate the term *segmentation* to that of classification in the following sense: classification is an operation applied to a pixel whereas segmentation is an image wide phenomenon. In this thesis, segmentation is considered to be the global effect of the classification process occurring at pixel level.

Terms used to describe the classifier's components.

The mechanism by which pixels are classified and the image segmented will be described as the *classifier*. The classifier represents the combination of algorithms which extract information from the image and segment the image on the basis of that information. We break this process into three stages: *measurement, feature extraction* and *discrimination.*

Discrimination occurs on the basis of evidence, generally the collection of this evidence may be broken into two stages. At the first level a *measure, d(x,y),* is obtained from the measurement of some signal component or components. This will have a tractable relationship with the original signal. While the measure is physically meaningful, it is not in a form suited to numerical discrimination and so further processing is required. The resulting *feature, f(x,y),* represents a quality of the image, it will not be tractable in general, however, it will be in a form which is suitable for the discriminant. In a multiclass problem it is usual to use more than one feature; several features associated with each pixel are treated as being orthogonal and grouped together as a *feature vector* **F(x,y)**. The discriminant function is then applied to the feature vector associated with each pixel and will allocate a label for that pixel according to the estimated class. The resulting segmentation will be described as the *label field l(x,y)*

5.3 The Choice of a Texture Measure

In this thesis, our agenda is slightly different from that of most texture analysis researchers. Whereas most researchers are concerned solely with minimising misclassification, we also aim to develop a model for the classification of rough surfaces. It follows that our main criterion for the selection of a texture measure should be

compatibility with that model, and implicit from this, that the measures should be suitable for the type of textures considered in this report. It is, however, also essential that the texture measure be illustrative of those used within the texture analysis community. The selection of a feature set for this thesis therefore rests on three criteria:

- the existence of a spectral representation of the measure,
- its popularity within the texture analysis community, and
- its suitability for random phase, broadband textures.

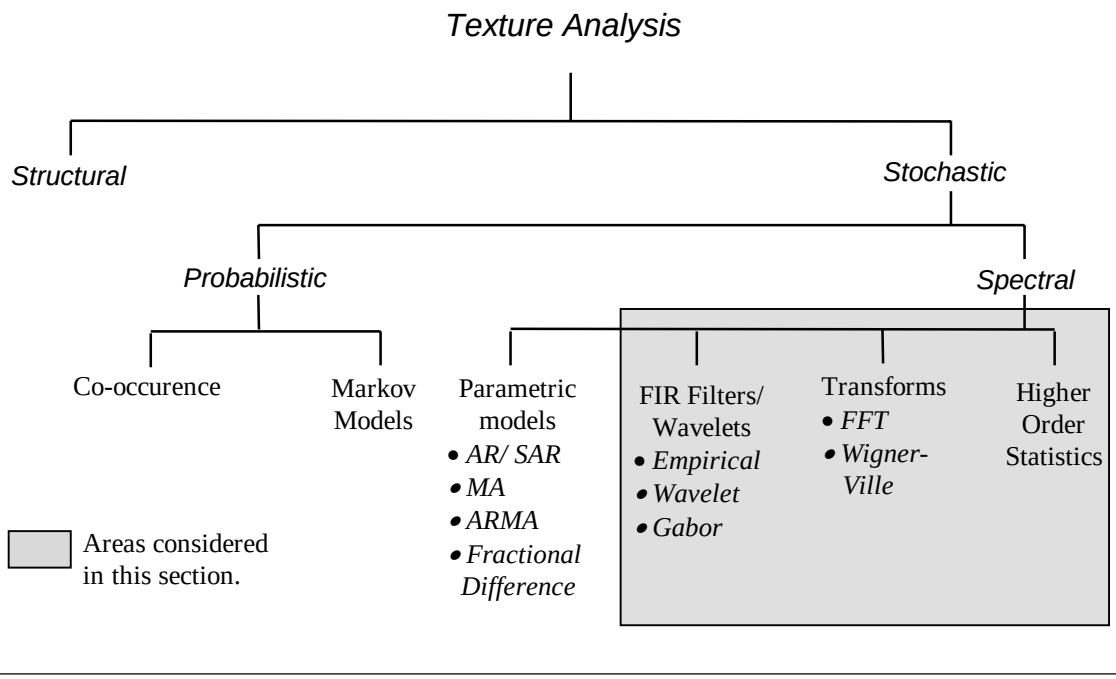


Figure 5.3.1 An (incomplete) taxonomy of texture analysis techniques.

In this section we will briefly consider a selection of texture analysis techniques which can be described in spectral terms. We show an incomplete taxonomy of texture analysis techniques in *Figure 5.3.1*. Texture analysis is usually initially divided into structural and stochastic groups. Structural techniques are generally applied to textures composed of a number of copies of a primitive placed at various locations in the image plane. Stochastic techniques are more appropriate for the type of textures—in which there is no obvious primitive—used in this thesis. We divide the stochastic techniques into those based on the texture's spectrum and those based on its probabilistic character. This distinction is not absolute; the probabilistic Gaussian Markov models are closely related to the spectral autoregressive models [Cohen91]. Furthermore, some techniques such as fractals e.g. [Linnett91] do not conform to either class.

Our first criterion is that there exists a spectral representation of the measure. This thesis is based on spectral models; in consequence we will ignore the probabilistic techniques. We do, however, note that these probabilistic techniques are often able to characterise structure which is lost by most spectral techniques. This was shown as an indirect result of the random phase experiment in section 2.3.2 where structure was detected in a signal with a whitened spectrum using second order probabilities. We justify our approach by adopting the maximum entropy phase restriction developed in Chapter 2 and only considering texture types which are described completely—for the purposes of classification—by their power spectrum.

Up until this point in the thesis, we have considered images and surfaces in which there is only one type of texture present. In any segmentation scheme this is not a valid assumption. Instead, the observed image or surface is considered as consisting of distinct areas of homogeneous texture. These regions, while being stationary, will have different spectral characteristics and the overall image may no longer be considered stationary. In consequence classical spectral analysis techniques can no longer be applied. Instead, the two dimensional equivalents of time/frequency techniques, which attempt to maximise localisation in both the spatial and the spectral domains, will be considered.

Of the spectral techniques we chose to ignore the parametric models. Although these models do have spectral representations, and indeed have been used for spectral estimation [Therrien p.596], the features on which the classification is based are difficult to relate intuitively to our theoretical models. Whereas the other techniques considered use features based on image measurements, the model based features are based on the closeness of the sample to a candidate texture. This is an effective strategy for classification, but it does not lend itself to the analytical approach of this thesis.

Historically, many texture analysis techniques have concentrated on signal magnitude or power at the expense of phase. This is not always an appropriate approach since for many textures the majority of textural information lies in the phase spectrum. More recently, techniques that utilise phase information have emerged to redress this imbalance. However, in Chapter 2 we adopted our maximum entropy phase condition and explicitly stated that no discriminatory information is held in the phase spectra of the textures used in this thesis. It would be desirable therefore for a candidate technique to identify the signal's phase component which can then be resolved from magnitude information and disregarded.

The second criterion in our selection of a texture measure is that of popularity within the texture analysis community. In the next chapter we will show that the measure adopted, and consequently the classifier, are affected by changes in illuminant tilt. For this result to have any relevance to the wider texture analysis community, the measure must be widely used within that community.

The third criterion is suitability to broadband, random phase textures. A great deal of work has been carried out on narrow band textures, e.g.[Bovik96][Weldon96], however, the development of fractal signal models and multiresolution techniques promises an equally analytical approach to broadband textures. In the context of this thesis, the only real effect of stating that the sample textures are broadband is the exclusion of narrowband techniques.

5.3.1 Wigner Ville Distribution

The application of time/frequency techniques to texture analysis represents an attempt to resolve the conflicting requirements of accurately localising a non-local phenomenon. Judged on this criterion, a candidate technique should have high spatial and spatial-frequency resolution. The Wigner Ville distribution (WVD) has twice the conjoint resolution of the STFT periodogram and superior resolution to that of the Gabor transform. It is therefore unsurprising that several authors have used the WVD as the basis of texture analysis algorithms [Reed90][Christobal91][Song92][Zhu93].

The spatial/spatial-frequency representation of a two dimensional image is a four dimensional function. For simplicity of notation, and ease of visualisation, we will, in general, concentrate on the application of the WVD to a one dimensional signal *s(x)*.

The WVD is a member of the large family of time/frequency distributions known as Cohen's class. Members of this class conform to a general expression that includes a kernel function. Members differ in the form of this kernel, and in the case of the WVD the kernel is constant and equal to unity. The form of the WVD is given below:

$$W(x,f) = \int_{-\infty}^{\infty} s(x+\frac{t}{2})s^{*}(x-\frac{t}{2})\exp(-j2\pi ft)dt$$

Due to the symmetry of the 'lag', the WVD is always real, however, it does contain phase information, albeit implicitly, and is an invertible transform.

The discrete form of the WVD is known as the *Pseudo* WVD (PWVD), it is shown below:

$$W_s[x, f] = 2 \sum_{k=-\infty}^{\infty} s[x+t]s^*[x-t]w[t]w^*[-t]\exp(-2\pi ft)$$

The factor two, present before the summation and in the exponential term, is due to the fact that in order to evaluate x± t/2, we must sample at twice the Nyquist rate to avoid aliasing. Christobal et al. [Christobal91] note three methods of satisfying this requirement:

- oversampling the signal, used in [Reed90]
- using the analytical signal, used in [Zhu93] or
- low pass filtering the image, implicit in the windowing function in [Song92].

The PWVD also includes a window, $w[t]$, which reduces the effect of truncation on the estimate, and acts to suppress high frequencies and reduce the aliasing problem.

The major shortcoming of the WVD is the presence of interference, or cross terms in the distribution. Due to the intrinsic bilinearity of the distribution, frequency components interact to introduce spurious components. Zhu et al. identify two classes of interference:

(a.) interference between positive and negative frequencies, and

(b.) interference between components of different absolute frequency.

In the one dimensional case, the first problem can be avoided by eliminating the negative frequencies and using the analytical signal [Quian, p.123]. In two dimensions, Zhu notes that there is more than one possible analytical image for a given real image; furthermore, each analytical image has its own properties; which analytical image to use must be decided on the basis of the spectral properties of the original image.

While certain aspects of the second problem can be alleviated by using the analytical image, it nonetheless remains a significant problem. Quian notes that the cross-terms oscillate, especially for frequency components which are far apart [Quian p.121]. The cross-terms therefore can be reduced by low pass filtering the spectra with over time and frequency—at the cost of resolution—to give the smoothed WVD (SWVD).

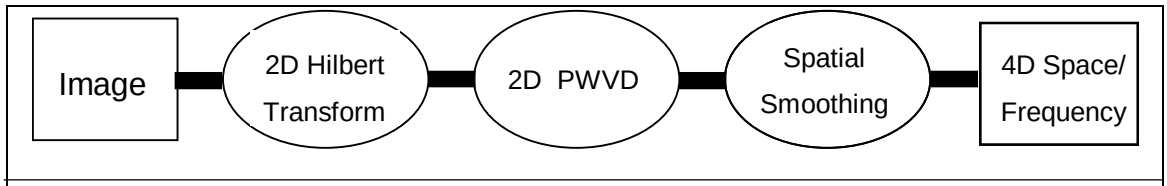


Figure 5.3.2 Zhu's scheme for local spectra estimation

Zhu adopts the scheme for the spectral estimation for spectral estimation outlined in *Figure 5.3.2*. The first step in the scheme is to suppress negative frequencies by using the 2D Hilbert transform of the image. The 2D pseudo WVD is then obtained at each spatial point in the analytic image. The local spectra obtained are then averaged over space.

Application of the WVD to images results in a 4D function (or a 2D function at each sample point). Authors adopt various schemes to extract useful information from the distribution. In the context of crack detection in a textured field, Song et al. first calculate the average distribution for the image. The average local spectra is then subtracted from the actual spectra at each image point. Pixels with residues exceeding a threshold are classified as being possible crack locations.

Other authors have extracted feature measures from the spectra. Reed and Wechsler use the location of the peak frequency component spectra as a feature, whereas Zhu used the orientation and radial frequency of the largest spectral peak as features. Christobal et al. developed five feature vectors in their scheme:

- mean of image frequency content
- variance of image frequency content
- mean of image directionality
- variance of image directionality
- variance of image intensity.

We believe these features to be largely arbitrary and, in the context of our requirement for a tractable measure, none of these techniques is entirely satisfactory.

The WVD has two problems: the most serious of which is intercomponent interference; this presents a significant problem for signals with a limited number of harmonic components, it is likely to be an even more serious drawback to application of the technique to broadband textures. Hlawatsch and Boudreaux-Bartels state that the

number of interference terms grows quadratically with the number of signal components [Hlawatsch92]. The second difficulty is that the WVD implicitly contains phase information. The fact that there is no clear method stated in the literature for resolving phase and magnitude components means that the feature images will contain information, which by our conditions, is redundant.

5.3.2 Higher Order Statistics

Higher order statistics (HOS) have recently been proposed as a means of discriminating between textures. The power spectrum gives a complete statistical description of a Gaussian process. It is however, unable to describe either the phase relations within a process or departures from normality. Hall and Giannakis developed a test of normality and applied it to nine Brodatz textures. All the test textures were found to be non-Gaussian to varying degrees—presenting a strong case for the use of HOS [Hall95b]. Higher order statistical techniques are suited to those textures which we have been at pains to exclude from the scope of this thesis. Nevertheless, we include a brief discussion of these methods partly for completeness, but also in order to gain an insight into the effect of our assumptions about the data set.

Higher order statistics form a natural extension to the second order statistics used in classical spectral analysis. If we consider the moments of a random process $s[x]$,

$M_x^{(1)}$	$E[s[x]]$
$M_x^{(2)}[t]$	$E[s^*[x]s[x+t]]$
$M_x^{(3)}[t_1,t_2]$	$E[s^*[x]s[x+t_1]s[x+t_2]]$
$M_x^{(4)}[t_1,t_2,t_3]$	$E[s^*[x]s^*[x+t_1]s[x+t_2]s[x+t_3]$

Table 5.3.1 Definition of higher order moments.

Cumulants of order 1 and 2 are identical to the mean and autocorrelation function. The third and fourth order cumulants for a non-Gaussian process *s[x]* are defined as

$$C_s^{(K)}[t_1,t_2..t_K] = M_s^{(K)}[t_1,t_2..t_k] - M_{s'}^{(K)}[t_1,t_2..t_K]$$

K=3,4

where *s'* is a Gaussian process with mean and correlation function identical to those of *s*. It follows that where *s* is a Gaussian process, the third and fourth order spectra are zero.

The higher order spectra, or polyspectra, are the Fourier transforms of the third and fourth order cumulants, and are known as the *Bispectrum* and the *Trispectrum*

respectively. The Bispectrum has been found to be useful in detecting *quadratic phase coupling*, i.e. phase relationships between harmonically related components of a random process.

Several researchers have taken advantage of the properties of higher order spectra for signal processing applications. Nikias and Mendel [Nikias93] identify four areas where high order statistics have been used.

- Suppression of additive Gaussian noise of unknown spectra.
- Identification and reconstruction of non-minimum phase systems.
- Recovery of information contained in deviations from a Gaussian process.
- Detection and characterisation of non-linear systems.

Tsatsanis and Giannakis use HOS to classify textures using a bank of filters, where each filter is matched to one of the candidate textures [Tsatsanis92]. The image is then filtered and the zeroth lag of the third order cumulant calculated for each filter output. The filter which gives the highest cumulant is assumed to be matched to the correct texture.

HOS techniques represent an attempt to utilise texture information that has been largely ignored by most spectral techniques. The proponents of HOS make a convincing case that the assumption of normality, which underlies most spectral techniques is not safe for many textures, and allows the loss of valuable discriminatory information. It is our view that this argument is, in general, justified. However, we have explicitly limited the scope of this thesis to textures which are near Gaussian and which carry little or no information in the phase spectrum. It is our belief that by adopting these restrictions we may justifiably ignore higher order statistics.

5.3.3 Empirical Techniques

The class of techniques we describe as *empirical* are FIR filters, designed purely for the purposes of discriminating between textures, either in the general case, or for a particular set of textures.

The oldest, and best known members of this class are the Laws feature measures [Laws79]. Despite their simplicity and lack of theoretical background, these form a highly effective approach, which in many cases have a classification performance comparable with modern techniques, and very modest computational expense. More recently, several authors have used Laws features as a testbed for the use of novel classification algorithms.

Greenhill and Davies used a classifier consisting of Laws filters with a neural network and mode filter [Greenhill93]. While Chen and Kundu used the Laws features in combination with a hidden Markov model [Chen95]. Moreover, the simplicity and effectiveness of Laws filters have made them a popular choice for researchers working on applications rather than investigating techniques, e.g.[Miller91][Neubauer92].

Other authors have produced schemes with analytically derived filters for a particular classification task. Ade calculated the eigenvectors of the covariance matrix of 3x3 neighbourhoods of the textures. The nine resulting eigenvectors were convolved with the image to produce principal component images. Classification is carried out on the basis of the averaged absolute values of pixels from these images [Ade83]. Randen and Husoy developed an expression for the statistics of the output of the post-processing stage of his classifier, this being a function of the texture statistics, the filter mask and an averaging filter [Randen95]. He then uses this expression as the basis of an algorithm which adjusts the filter weights as a means of maximising the distance between classes in feature space. Jain and Karu noted the similarity between the prototypical FIR filter based system and a neural network [Jain96]. In Jain's analogy, the filter coefficients correspond to the input weights of the network, while the non-linearity and averaging stages correspond to the non-linear summation junction of the neural network. He uses training data with the Back-propagation algorithm to derive filter coefficients which are optimal for a given classification task.

The use of empirically defined masks offers an effective, and computationally efficient, approach to texture classification. Recent advances in developing masks for a particular task represent a promising area of research for supervised classification. Despite these advantages, empirical filters will not be used in this thesis. Although, by definition, the modern empirical techniques will be suited to the discrimination of the test textures, due to the *ad hoc* nature of the filters, these techniques fail to meet our two remaining criteria for the selection of a feature measure:

- A given filter may not be well localised, or have a tractable expression in the frequency domain.
- The *ad hoc* sampling of the frequency domain is unique to the application -this will reduce the ability of the results to be generalised.

5.3.4 Wavelets

Wavelets have proved to be an effective and popular tool for texture analysis in recent years [Livens97]. Introductions to wavelets occur at a variety of conceptual and mathematical levels. In this section we will give a brief description of the technique at the two simplest levels to illustrate the significance of the different techniques used in texture analysis. The first explicitly deals with the wavelet—effectively treating it as an FIR filter related to the empirical schemes discussed above. The second level does not explicitly evaluate the wavelet, but uses it as a theoretical construct to describe the result of the implementation.

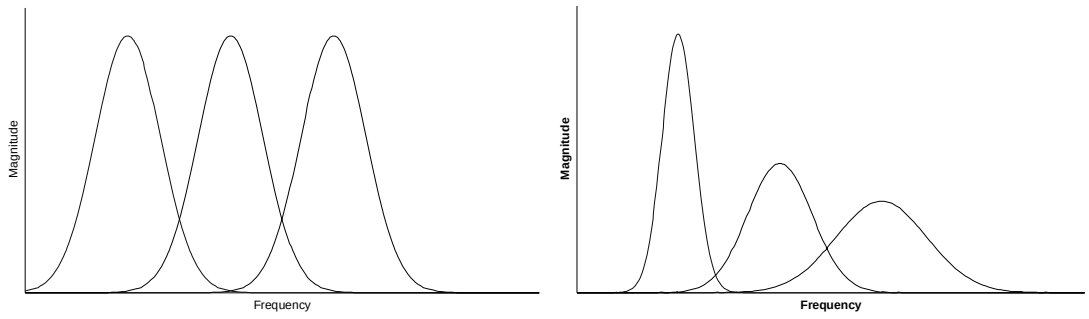
Explicitly evaluated wavelets

The trade-off inherent in time/frequency techniques is (in our case) spatial localisation against spectral localisation. It is desirable that the bandwidth and the spatial extent of the elementary functions should both be minimised. Unfortunately they are inversely related and a compromise must be sought.

Given the relationship between spatial and spectral localisation, we may pursue two distinct strategies for sampling the signal spectrum:

(1.) using elementary functions with bandwidths which are constant throughout the signal spectrum, *Figure 5.3.3a*, and

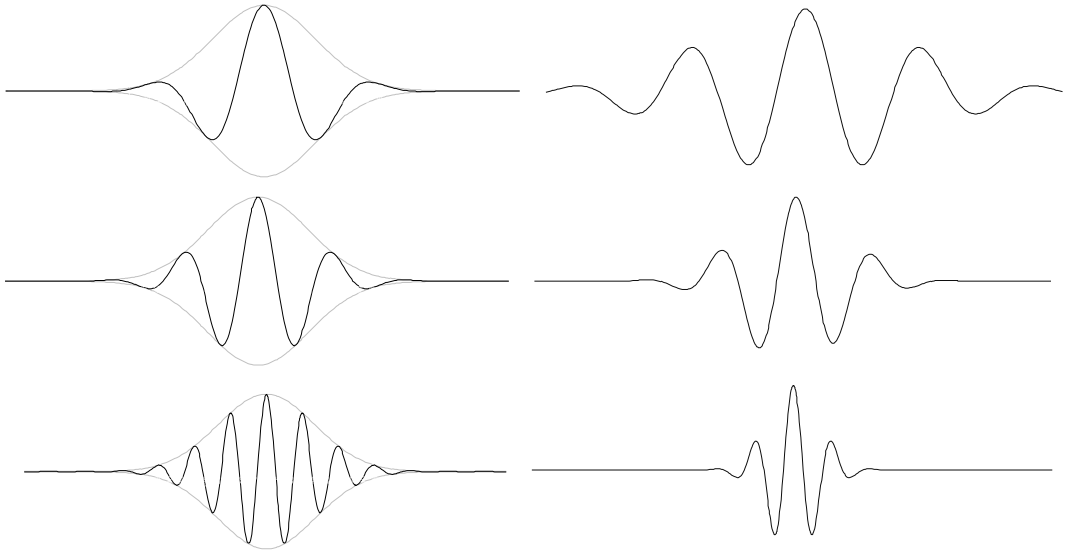
(2.) sampling with elementary functions each possessing a bandwidth which is constant *relative* to the function's central frequency, *Figure 5.3.3b*.



(a.) *(b.)*

Figure 5.3.3 Spectral sampling schemes, (a) constant absolute bandwidth, (b) constant relative bandwidth.

The first is the approach undertaken by the Gabor Transform and represents the spectral dual of a series of analysis functions with envelopes of the same spatial extent. Spatial localisation is uniform throughout the frequency range, and is therefore limited by the bandwidth requirements of the analysis function of lowest frequency, *Figure 5.3.4a*.



(a) Gabor Transform *(b.) Wavelet*

Figure 5.3.4 Analysis functions of (a) the Gabor Transform and (b) the wavelet approach, (after Quian p.77).

The second approach, used in wavelet analysis, varies bandwidth with frequency. In the spatial domain this results in the extent of the envelope function varying with the wavelength modulating function. As a consequence of this, the number of wavelengths in an envelope will be constant and the functions will be scaled versions of a single prototype, *Figure 5.3.4b*.

The spatial effect of variable bandwidth is that low frequency filters will give good frequency resolution, but poor spatial resolution, high frequency filters will give poor frequency resolution and good spatial resolution. Low signal frequencies will be easily discerned from neighbouring frequencies but will be difficult to localise spatially whereas the converse will apply to high frequency signals.

Implementation-based explanation

In practice, wavelets are not generally implemented as convolution filters, instead most implementations are based on Mallat's multiresolution approach. A brief description of the technique is given here.

Define an elementary function: $\phi(x-k)$, known as a *scaling function*, together with translated versions of itself, this forms a basis for the space V_0. Scaled versions of this function, i.e. $\phi_l(2^l) = \phi(2^l x - k)$ each form a basis for a corresponding space space V_l. Intuitively, as l increases the ability of the function to detect detail also increases. Furthermore, V_l is a subspace of V_{l+1}

and define W_l to be the complement of V_l on the space V_{l+1} such that,

$$\mathrm{V}_{l+1} = V_l \oplus W_l$$

That is, a signal defined on V_{l+1} can be represented in terms of a low detail signal defined on V_l and a detail signal defined on W_l

Since V_l is a subspace of V_{l+1}, the elementary function ϕ_l can be represented as a linear combination of shifted version of the function ϕ_{l+1}, *i.e.*

$$\phi_l(\frac{x}{2}) = 2\sum_t h_t \phi_{l+1}(x-t) \qquad \textit{(5.3.4a)}$$

Expression 5.3.4a is known as the dilation equation. The coefficients h_t may be thought of as the weights of a FIR filter, furthermore it can be shown that this is low pass in character (Qian p.88). Since Eq. 5.3.4a is the convolution of the coefficients h_t and $\phi_{l+1}(x-t)$, we may express it in the frequency domain as the product of the Fourier transforms of these terms:

$$\phi_l(2\omega) = H(\omega)\phi_{l+1}(\omega)\,.$$

or equivalently as:

$$\phi_l(\omega) = H\left(\frac{\omega}{2}\right)\phi_{l+1}\left(\frac{\omega}{2}\right).$$

or more generally as

$$\phi_l(\omega) = \prod_{K=1}^{\infty} H\left(\frac{\omega}{2^k}\right)\phi(0)$$

$\phi(0)$ is a constant and $\phi_l(\omega)$ is a function of $H(\omega)$ only.

The low pass filter represents a mechanism to move from a space V_{n+1} defined in terms of $\phi_{l+1}(\omega)$ to a less detailed space V_n defined in terms of $\phi_l(\omega)$.

The detail which is residual to this transition is defined in the space W_l. Let elementary function ψ_l form the basis for W_l. The detail signal at level *n* may be observed by high pass filtering the image defined in subspace V_{n+1}.

The high pass filter should satisfy (5.3.4b)

$$H(\omega)G^*(\omega) + H(\omega+\pi)G^*(\omega+\pi) = 0 \qquad (5.3.4b)$$

where $H(\omega)$ and $G(\omega)$ are quadrature filters. One solution of (5.3.4b) is shown below:

$$G(\omega) = -\exp(-j\omega)H^*(\omega+\pi)$$

The detail image at level *l* represents the result of the low pass filterings required to reach level *l* followed by a high pass filtering. The net effect of which is of a bandpass filter

$$\psi(\omega) = G\left(\frac{\omega}{2}\right)\phi\left(\frac{\omega}{2}\right)$$

$$\psi(\omega) = G\left(\frac{\omega}{2}\right)\prod_{K=2}^{\infty} H\left(\frac{\omega}{2^K}\right)$$

$$\psi(x) = 2\sum_K g_K \phi(2x-K)$$

Consequently, the wavelet transform may be evaluated over a range of scales by repeated low pass filtering followed by high pass filtering, and neither the wavelet, nor the scaling function need ever be explicitly evaluated.

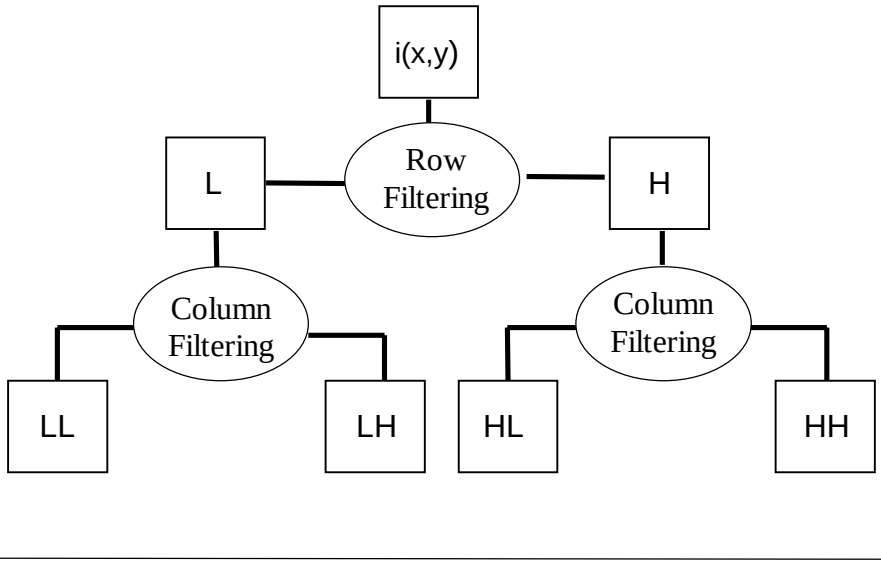


Figure 5.3.5 Two dimensional wavelet implementation.

Directionality

The two dimensional wavelet transform is applied consecutively along the rows and columns of the data set, in a similar fashion to that of the FFT. The separability of the two dimensional transform makes its implementation highly efficient. The approach is shown in *Figure 5.3.5*, giving four data sets for each level.

While the above approach is used in almost all schemes, it gives poor polar resolution—which is particularly relevant to this thesis. Antoine [Antoine93] has applied a directional Morlet wavelet, though this is implemented in the two dimensional frequency domain and will lack the computational efficiency of the separable scheme. Freeman and Adelson propose the use of a steerable pyramid [Freeman91]. This approach uses basis filters oriented at 0°, 45°,90° and 135° which can be combined to allow the image to be filtered in any arbitrary direction. The filters are themselves separable, allowing an efficient implementation.

Summary

Wavelets have become highly popular in the literature over recent years, they offer a computationally efficient method of localising spectral components in the frequency domain. Unfortunately, most implementations are separable, being applied to rows and columns in turn. The result of this is a rather crude sampling in the polar frequency domain. Given the importance of directionality in this thesis, this is a rather more serious drawback than in most applications.

5.3.5 Gabor Functions

Definition of 1D Gabor function

The Gabor filter is defined as consisting of a Gaussian envelope modulated by a complex exponential. The one dimensional spatial and spectral forms are expressed in equations (5.3.5a) and (5.3.5b) respectively. The function's real part and its frequency domain description are shown in *Figure 5.3.6*.

$$g(x) = \exp\left[\frac{-x^2}{2\sigma_f{}^2}\right]\exp\left[j(2\pi\omega_0 x + \varphi)\right] \qquad (5.3.5a)$$

$$G(\omega) = \exp\left[-2\pi\sigma_f{}^2(\omega-\omega_0)^2\right] + \exp\left[-2\pi\sigma_f{}^2(\omega-\omega_0)^2\right] \qquad (5.3.5b)$$

where x is the spatial variable

σ_f parameterises the extent of the Gaussian envelope

ω_0 is the centre frequency of the filter

φ is the phase displacement.

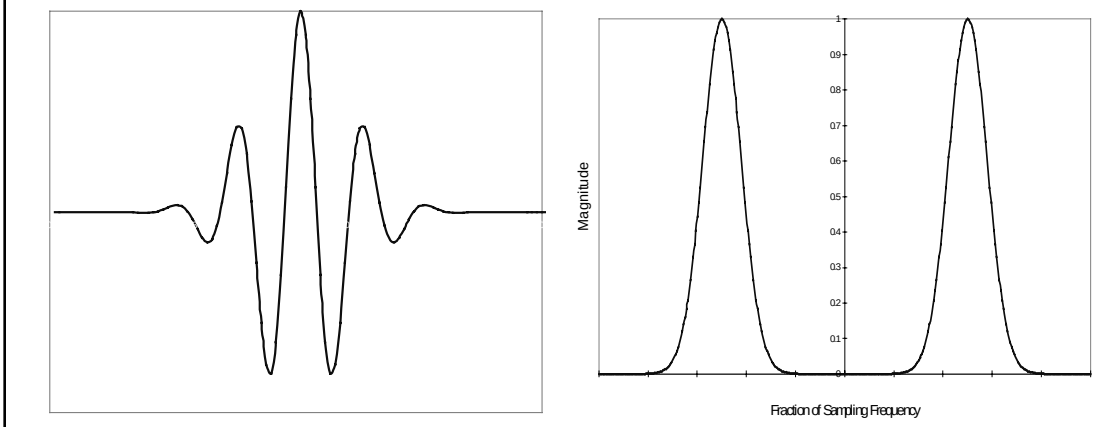


Figure 5.3.6 Real spatial part and frequency domain representation of Gabor function.

Bandwidth Characteristics of the Gabor Function

The spectral location of the bandpass region of a Gabor filter is governed by the modulating function. The bandwidth, however, is governed by the standard deviation of the Gaussian envelope, regardless of the frequency modulating function. A filter with a large spatial variance (*Figure 5.3.7a*) will have a relatively localised spectral representation (*Figure 5.3.7b*). Decreasing the extent of the spatial envelope, i.e. increasing its spatial localisation(*Figure 5.3.7c*), will decrease the spectral selectivity of the filter(*Figure 5.3.7d*). A large variance will increase the spectral resolution at the cost of decreased spatial resolution.

Gabor as a wavelet

By linking envelope extent to the sinusoid's wavelength, we may define a series of functions which are scaled versions of each other. A family of Gabor functions may be described as a series of dilated forms of a single prototype, and a transform may be defined in terms of equation 5.3.5c [Lee96].

$$g(a_d, x, y, x_0, y_0) = \|a_d\|^{-1} g\left(\frac{x - x_0}{a_d}, \frac{y - y_0}{a_d}\right) \qquad (5.3.5c)$$

where x_0 and y_0 are the spatial co-ordinates of the filter at a given time,

and a_d is the dilation parameter (typically a power of two).

In this sense, a Gabor function may be treated as a wavelet, albeit lacking in the usual requirements for admissibility and orthogonality.

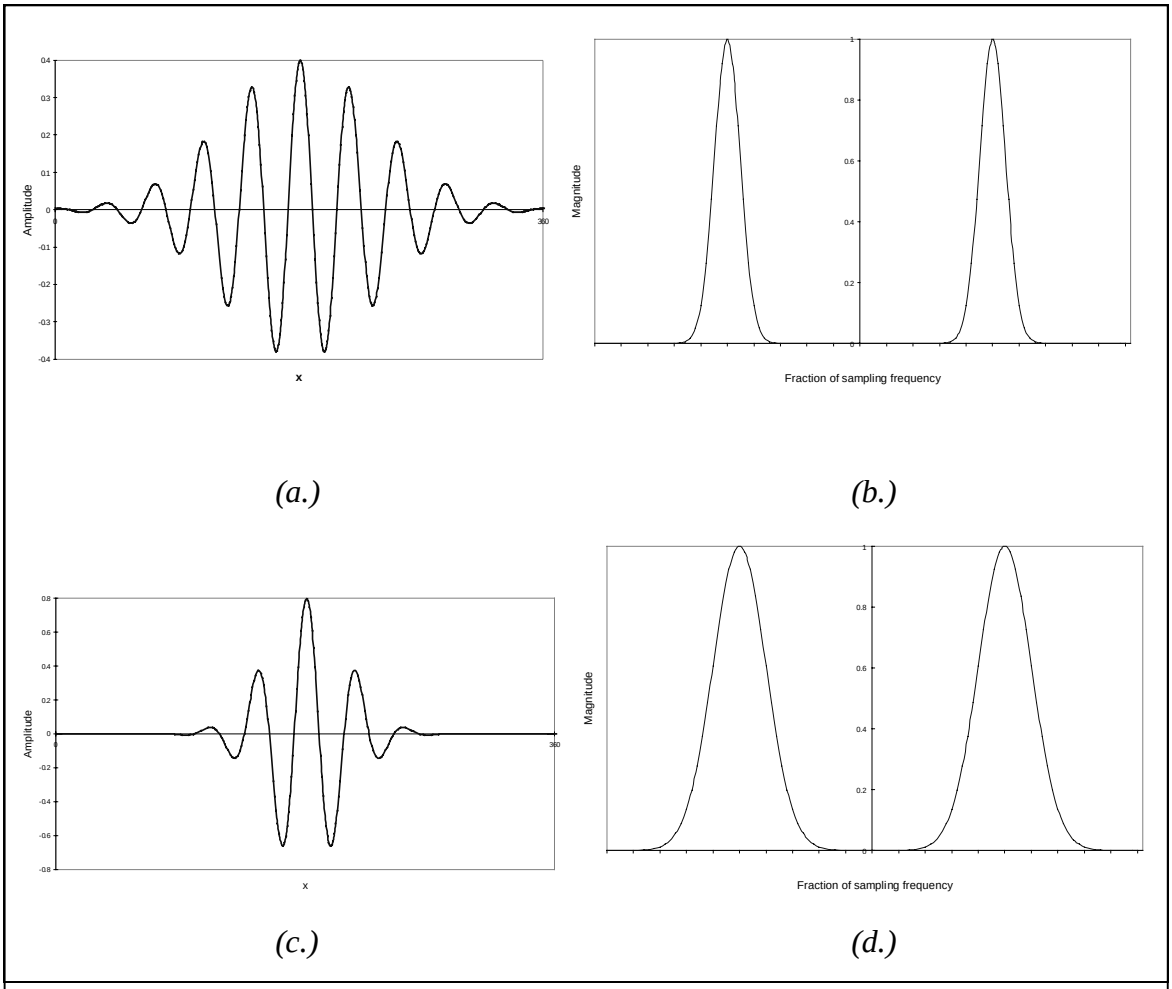


Figure 5.3.7 The effect of decreasing the extent of the Gaussian envelope on the frequency domain representation.

The admissibility condition requires that a wavelet function has zero mean. While the imaginary part of the function is admissible, the real part has a non zero mean. Lee develops an additional term to the carrier function to eliminate the mean response and satisfy the admissibility criterion, the modified expression is shown below.

$$\exp[j\omega_0(xCos\phi + ySin\phi)] - \exp\left[\frac{-\kappa^2}{2}\right]$$

where ϕ is the orientation of the filter.

and $\kappa=\pi$ for a 1 octave bandwidth filter.

Navarro et al. report that the mean component is very small [Navarro95].

The orthogonality condition is clearly of relevance to coding schemes, but how relevant is it to pattern recognition? Non-orthogonal filters will certainly exhibit a degree of redundancy, but Navarro et al. argues that with their use, there will be an increase in the robustness of classification: "*Biological vision...lacks orthogonality, producing a*

redundancy that is highly expensive, this being the price of robustness." This implies that a degree of redundancy is desirable to our application, and that orthogonality is not a prerequisite for a classifier.

Although classifying the Gabor function as a wavelet is debatable, it is certainly wavelet-like. As the number of terms in the dilation function of a wavelet increases (Eq. 5.3.4a) the smoothness of the wavelet and its spectral compactness increase. For *t* large, many families of wavelets resemble windowed sinusoids similar in appearance to Gabor functions. In fact, the relationship of several wavelets is even more closely established; Unser et al. have shown that wavelets based on a B-spline scaling function converge to the Gabor function as the power of the spline increases [Unser92]. Antoine et al. describe the Morlet wavelet as a Gabor function with an additional term to ensure admissibility, furthermore, this term tends to zero for high frequencies [Antoine93].

Definition of the two-dimensional Gabor Function

The two dimensional form of the Gabor filter was first defined by Daugman [Daugman85]. The filter can be described as an elliptical Gaussian modulated by a complex sinusoid with direction of propagation ϕ. The spatial and spectral forms are shown in equations 5.3.5d and 5.3.5e respectively. For consistency, we use the notation described in [Jain91].

$$g(x,y) = \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi u_0 x) \qquad (5.3.5d)$$

$$G(u,v) = A\left(\exp\left\{-\frac{1}{2}\left[\frac{(u-u_0)}{\sigma_u{}^2}+\frac{v^2}{\sigma_v^2}\right]\right\}+\exp\left\{-\frac{1}{2}\left[\frac{(u+u_0)}{\sigma_u{}^2}+\frac{v^2}{\sigma_v^2}\right]\right\}\right) \qquad (5.3.5e)$$

where

$$\sigma_u = \frac{1}{2\pi\sigma_x} \qquad \sigma_v = \frac{1}{2\pi\sigma_y}$$

g(x,y)	is the filter impulse response.
G(u,v)	is the spectral transfer function.
u	is the component of frequency in the direction of the x-axis.
v	is the component of frequency in the direction of the y-axis.
v_0	is the centre frequency of the filter along the y-axis.

u_0 is the centre frequency of the filter along the x-axis.

σ_x is the standard deviation of the Gaussian envelope in the direction of the x axis in the spatial domain.

σ_y is the standard deviation of the Gaussian envelope in the direction of the y axis in the spatial domain.

σ_u represents the extent of the Gaussian envelope in the spectral domain in the direction of the x-axis.

σ_v represents the extent of the Gaussian envelope in the spectral domain in the direction of the y-axis.

Daugman also introduced the concept of orientation bandwidth. While the direction of maximum sensitivity is defined by ϕ, the degree to which adjacent directions are attenuated is governed by the extent of the Gaussian in the direction perpendicular to propagation. In *Figure 5.3.8* three Gabor filters, which have identical spatial parameters save σ_y, are shown. If we apply these filters to an isotropic field we obtain the polar plot shown in *Figure 5.3.9*.

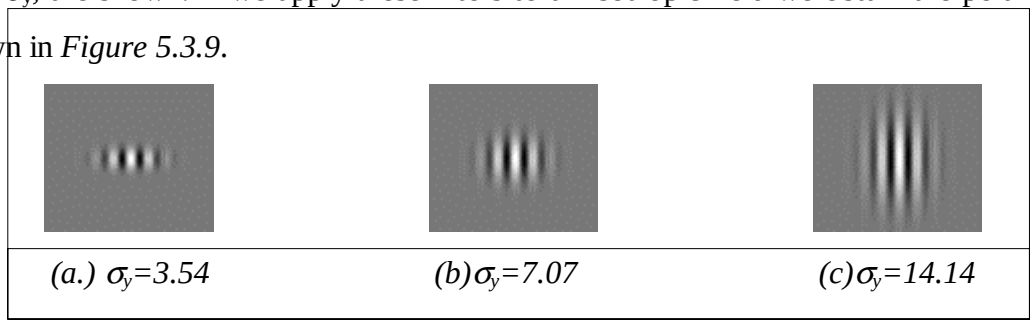


Figure 5.3.8 Gabor filters with various σ_y parameters.

As with radial frequency where there exists a trade-off between spectral and spatial resolution, there is an analogous trade-off between directionality and spatial resolution: a highly directional filter will not be well spatially localised perpendicular to propagation, *Figure 5.3.8* & *Figure 5.3.9*.

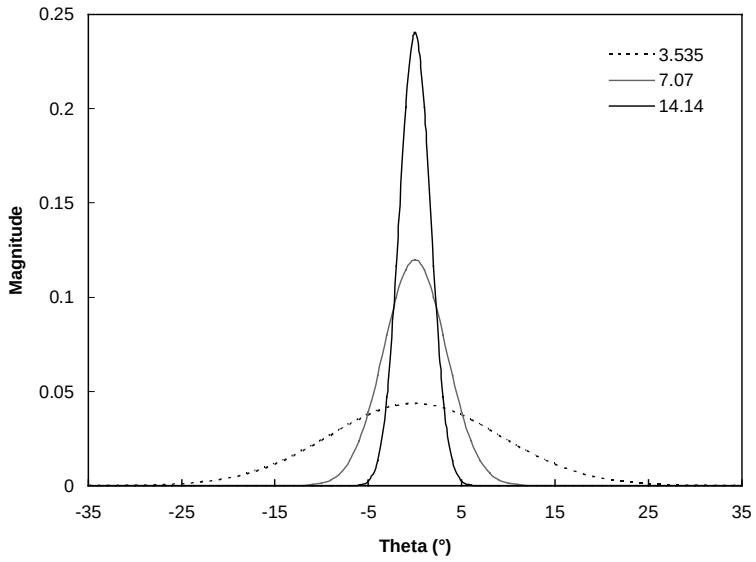


Figure 5.3.9 Polar plot of Gabor filters angular magnitude responses for various σ_y.

From Jain and Farrokhnia, the radial and orientation bandwidths of the filter, B_r and B_θ, are defined by equations 5.3.5f and 5.3.5g respectively. Kiernan has criticised the accuracy of these expressions on the basis of geometric arguments [Kiernan95]. However, in the context of this thesis, these inaccuracies are not critical and we will continue to use these expressions due to their accessibility and popularity in the literature.

$$B_r = \log_2\left(\frac{u_o + (2\ln 2)^{1/2}\sigma_u}{u_o - (2\ln 2)^{1/2}\sigma_u}\right) \tag{5.3.5f}$$

$$B_\phi = 2\tan^{-1}\left(\frac{(2\ln 2)^{1/2}\sigma_v}{u_0}\right) \tag{5.3.5g}$$

5.3.6 The Selection of a Measure Set

At the beginning of this section we stated three criteria for the selection of a texture measure:

- the existence of a spectral representation of the feature,
- its popularity in the texture analysis community, and
- its suitability for random phase, broadband textures.

The scope of this review was largely dictated by the first two criteria, we must therefore choose an algorithm on the basis of the third. The rationale behind HOS techniques is to utilise phase information. The random phase condition developed in Chapter 2 removes this justification for the type of textures used in this thesis. We

concluded that the Wigner Ville distribution was not suitable for broadband textures due to the interference of components, and the distribution's implicit treatment of phase information made it difficult to decorrelate useful power information from the random phase. Empirical techniques were considered, however, their *ad hoc* nature was considered to be incompatible with the analytical approach of this thesis. Wavelets represent a promising feature measure, however, their lack of polar resolution is a serious handicap to their application in the context of this thesis.

The related, though non-orthogonal and non-separable, two dimensional Gabor wavelet is much more versatile in the polar domain. Though this versatility is gained at the expense of the computational efficiency associated with a separable transform. In our assessment, of all the measures considered in this thesis, Gabor functions are closest to fulfilling our three criteria.

Gabor filters are defined explicitly both in the spatial domain and in the polar frequency domain and therefore can be easily integrated into our model. They may be used to extract and resolve magnitude and phase components of the image [Bovik90]. The magnitude information allows us to relate the output of the Gabor filters to the surface and image descriptions developed in the earlier chapters.

Besides being suited to the model used in this thesis, Gabor filters also represent an important and popular area of research within texture analysis. This is due to two properties of the Gabor filters: the space/frequency characteristics of the filter, and the similarity to the operation of the early human visual system. The complex form of the Gabor filter represents the optimum space-frequency resolution and is therefore of interest to analysts attempting to localise textures in the spatial domain using their spectral characteristics. Furthermore, the early stages of the human visual system can be accurately modelled using Gabor filters. These properties have popularised the Gabor function, and by establishing the tilt dependency of the Gabor based classifier, we will demonstrate the relevance of this work to a wide area of texture research.

In Chapters 2 and 3 we concluded that the surface types considered in this thesis (and their images) could be described as being broadband and having random phase. Gabor filters have been used to measure narrowband signals, however, their use for broadband textures is equally valid. Reduction in the spectral resolution corresponds to an increase in the spatial resolution. Therefore Gabor filters can be applied to accurately locate broadband textures. From our random phase condition we regard the phase

information as being redundant. Gabor filters provide a set of features in which this redundant information can be decorrelated from the relevant information and discarded. These filters are therefore suited to both the nature of the classification task, and to the characteristics of the test textures.

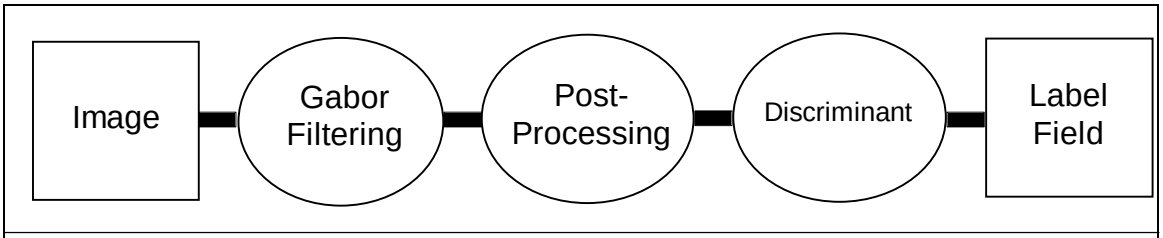


Figure 5.4.1 The Gabor-based Classifier

5.4 A Classification System

5.4.1 Overview

In this section we develop a classifier that is tailored to the task of classifying rough surfaces. We do not, however, make any claim of optimality. We base our classifier on the generic form shown in *Figure 5.4.1*. A textured image is processed using Gabor filters; the resulting measures are passed through a non-linear post-processing stage. The features extracted are passed to a statistical classifier which labels each pixel as belonging to a certain class on the basis of its feature vector and the *a priori* probability of that vector being a member of each class. In the following sub-sections we shall discuss each of these stages.

5.4.2 The Implementation of Gabor Filters

Design vs. Systematic Structure

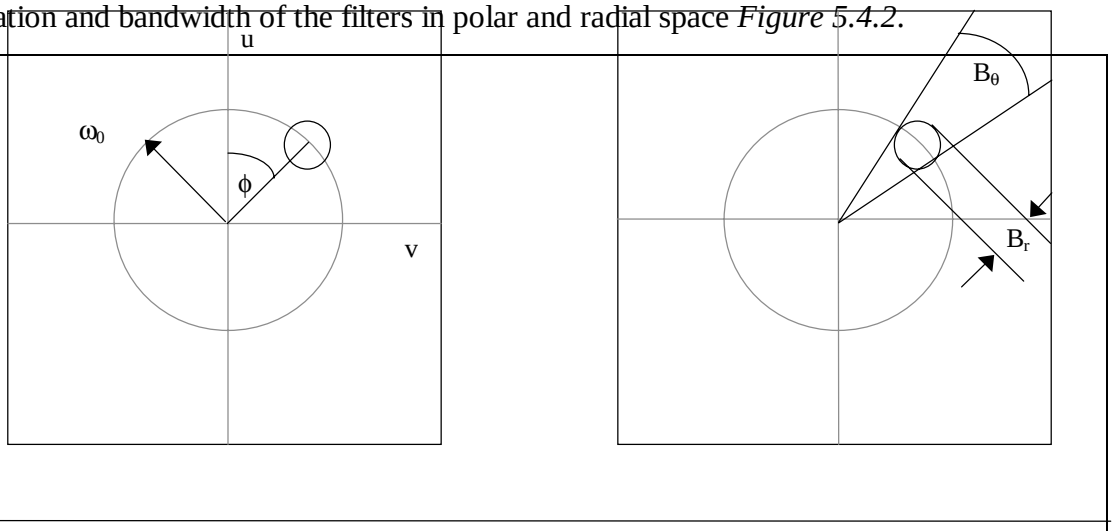
The generally accepted definition of an n-dimensional Gabor function is that of a Gaussian function modulated by a complex exponential. The generality of this statement is striking. The two dimensional expression of this statement, *Eq. 5.3.5a*, has four independent parameters. This lack of structure has, on balance, been beneficial: the most important characteristic of the Gabor filter, the space-frequency optimality, is encapsulated in its definition. Furthermore, the looseness of the definition allows the filter design algorithms the necessary latitude to be worthwhile.

While a great deal of work has been devoted to the algorithmic design of optimal Gabor filters for a particular task (e.g. [Dunn95],[Kiernan95], [Weldon96]), in this thesis

we shall use a more standard, less *ad hoc*, filter implementation. This is less computationally efficient than that produced by the design techniques, and will almost certainly give a poorer classification. Nevertheless we adopt this approach for two reasons:

- A uniform sampling of a particular frequency range is more compatible with our analytical approach. It allows us to draw conclusions not only about which features are most affected but also to analyse the effect on features using the same analytical framework used to model the earlier stages in the process considered in previous chapters.
- Generality; illustrating the effect with a small number of irregularly placed filters is less convincing proof that any irregularity is due to our predicted effect than a much wider polar frequency sampling.

If we opt for a systematic sampling of the image spectrum we must decide on the location and bandwidth of the filters in polar and radial space *Figure 5.4.2*.



(a.) Filter location (b.) Filter bandwidth

Figure 5.4.2 Filter parameters.

Several parameter schemes have emerged [Jain91][Lee96]. These are mostly based on biological justifications that constrain the parameters and which have been adopted by most non-design researchers. The most commonly used constraints are those adopted by Jain and Farrokhnia [Jain91]. Jain uses real filters with frequency and angular (half peak magnitude) bandwidth of 1 octave and 45° respectively. The polar frequency domain is sampled at intervals of 1 octave and 45°, giving a total of 32 filters for a 512x512 image. With this, and the polar distribution, in mind we adopt the following filter designation:

$$F\omega_0 d\phi$$

where ω_0 is the centre frequency of the filter in cycles per image

and ϕ is the filter orientation in degrees.

Polar Properties

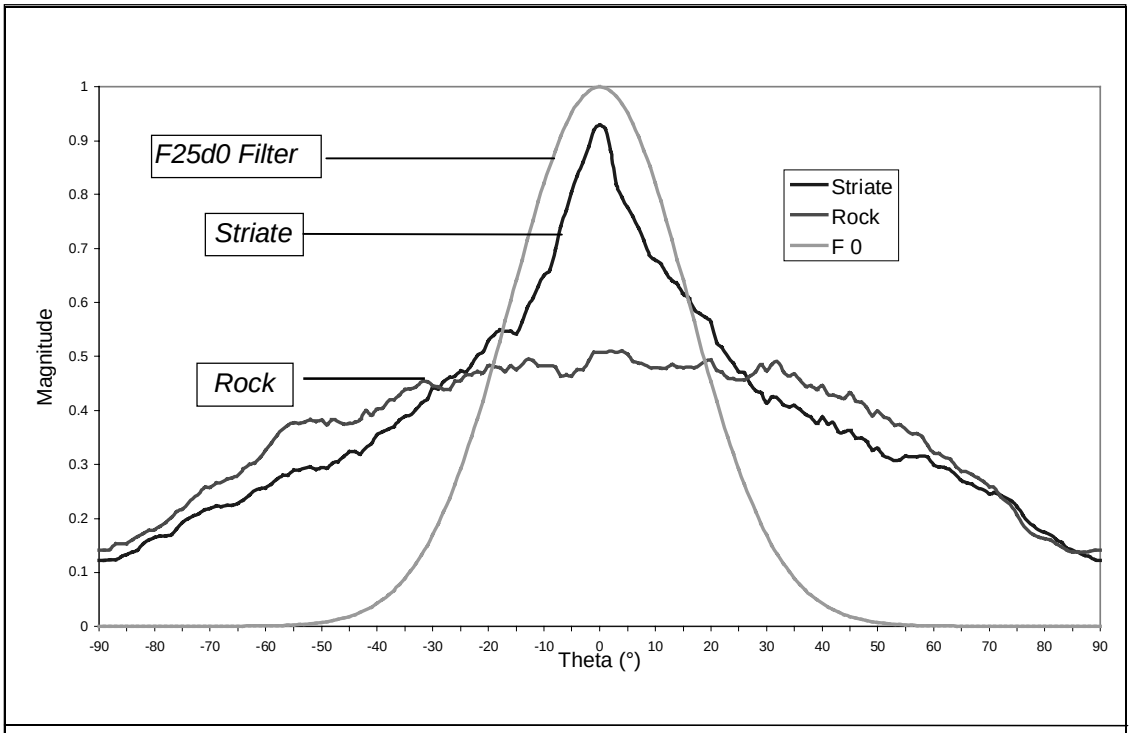


Figure 5.4.3 Polar Spectra of Rock and Striate Surfaces illuminated from Tau 0° and a Gabor filter orientated at Tau 0°.

Given the directional nature of the tilt effect we believe it is important for the purposes of analysis that the polar spectrum is uniformly sampled. Therefore, each frequency band will be sampled by filters each with a bandwidth of 45°, oriented at 0°,45°,90° and 135°, after Jain. In *Figure 5.4.3* we show the polar response of a filter oriented at ϕ=0° superimposed on the polar spectra of images of the isotropic rock surface and directional striate surface illuminated from τ=0°. The figure suggests that the 2D Gabor is able to detect directional information important for discrimination.

Radial Sampling

Several papers such as [Jain91][Nestares96][Namuduri94] adopt a wavelet approach to the implementation of Gabor filters. Kube and Pentland [Kube88] state that assuming a linear reflectance function, a fractal surface will give rise to a fractal image with roll-off $\beta_i=\beta_s+1$. In Chapter 2, we concluded that most surfaces have roll-off of

β=3.0 in their fractal region. A multi-resolution approach would seem to be redundant for textures which are entirely fractal. However, both the Mulvanney and Ogilvy type surfaces exhibit fractal behaviour over only a range of frequencies. The spectral characteristics below this range, and the point at which the transition to fractal behaviour occurs, are rich sources of discriminatory information. Moreover, the radial plots in the previous chapter show that the data sets are not fractal. A multi-resolution approach is therefore compatible with our surface models and our measurements, as well as being an efficient method of sampling the radial spectrum. The image is low pass filtered (lpf) and decimated before being passed to a set of Gabor filters. The lower the measured frequency, the more decimations its input data set has undergone. The wavelet implementation means that the central frequencies of filters will rise in one octave steps. As with Jain's implementation, the filters will have a bandwidth of an octave.

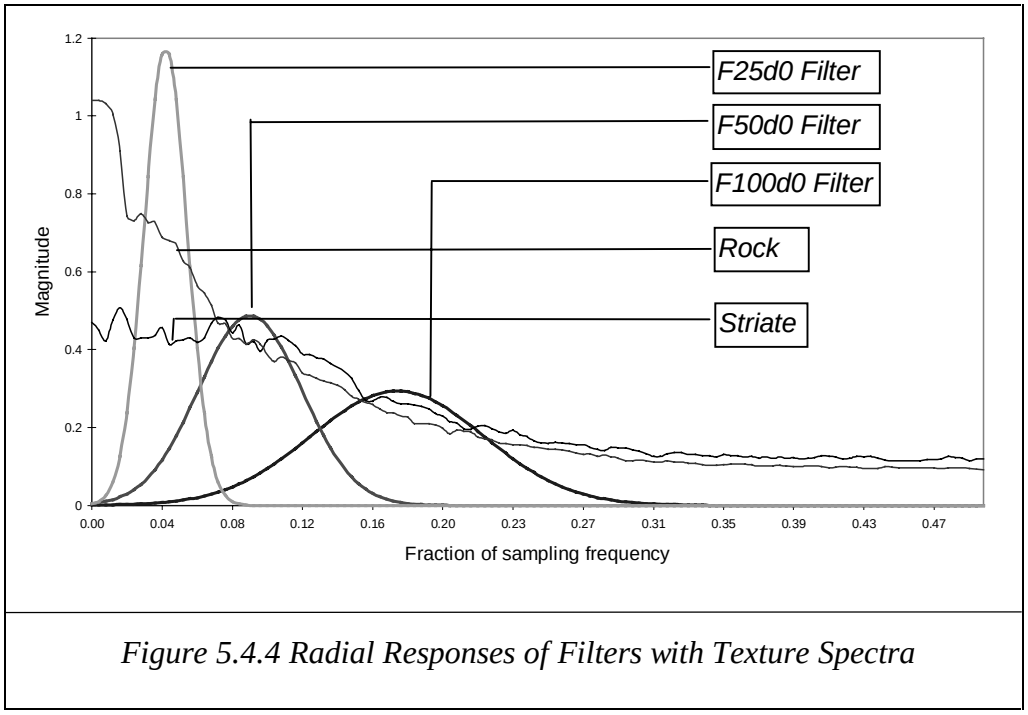


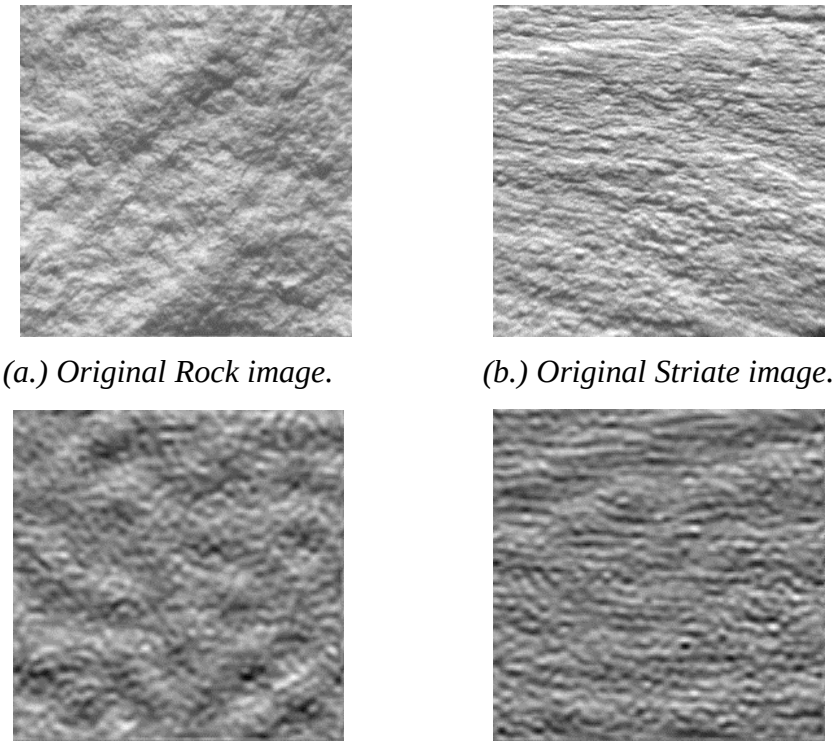
Figure 5.4.4 Radial Responses of Filters with Texture Spectra

This leads to the question of which range of frequencies the filter should be applied to. In Chapter 2 and 3, we saw that the radial spectra of different surface models (and consequently of their images) differ most markedly at low frequencies. Furthermore, in Chapter 4 we noted the attenuation of high frequencies due to blurring, and the decrease in the S/N ratio with increasing frequency. This would suggest that it is the low frequencies that should be most closely scrutinised by a classifier. On the other hand, in this chapter we have noted that low frequency filters have poor spatial resolution. As an empirical compromise, we will use three sets of filters, which range from the low

frequencies to the midband, beginning at 25 cycles per image. The responses of the F25d0, F50d0 and F100d0 filters are shown in *Figure 5.4.4*, plotted against one dimensional spectra of the columns of our two exemplar textures *Rock* and *Striate*.

Spectral Support

Given the emphasis on low frequencies, and the linear nature of the filters, it is interesting to discover how much image information is captured by the feature set. to code the image. The two exemplar textures, *Rock* and *Striate*, were filtered by the Gabor filters on which the classifier is based (the F25,F50 and F100 sets). The sum of the measure images are shown in *Figure 5.4.5c* and *d*. While the combinations of the measure images do not bear an obvious resemblance to the exemplars, it is noticeable that the directionality of both textures is captured, and the more pronounced directionality of the *Striate* surface is particularly well recorded.



(a.) Original Rock image. *(b.) Original Striate image.*

(c.) Coded Rock image. *(d.) Coded Striate image.*

Figure 5.4.5 The Rock and Striate exemplar textures and their feature set coding.

5.4.3 Post-Processing

Although the form of the Gabor filter itself is relatively standard, there are several approaches to the subsequent processing of the signal prior to classification, although these generally take the form of a non-linearity followed by low pass filtering. Although

the non-linearity limits our ability to analyse the process, it is clearly central to the scheme. If it were not (and textures could be classified on the basis of linear operations) then classification could be carried out by a statistical classifier purely on the basis of grey levels and their displacements with no need to filter.

In selecting a post-processing approach we must attempt to satisfy certain criteria. Firstly, the technique must perform well and reliably in conjunction with the discriminant. Since many classifiers are optimal for Gaussian data only, it is also necessary that the output of the post-processing stage should have at least an approximately Gaussian distribution. Specific to the approach of this thesis, it is also desirable that the algorithm should permit some degree of analysis and have a physically meaningful output. The schemes may be broadly split into two groups: those based on a biological rationale and those grounded in signal processing theory. A taxonomy is shown in *Figure 5.4.6*.

Biologically Based Schemes

Jain and Farrokhnia were the first to suggest the use of a hyperbolic tan function

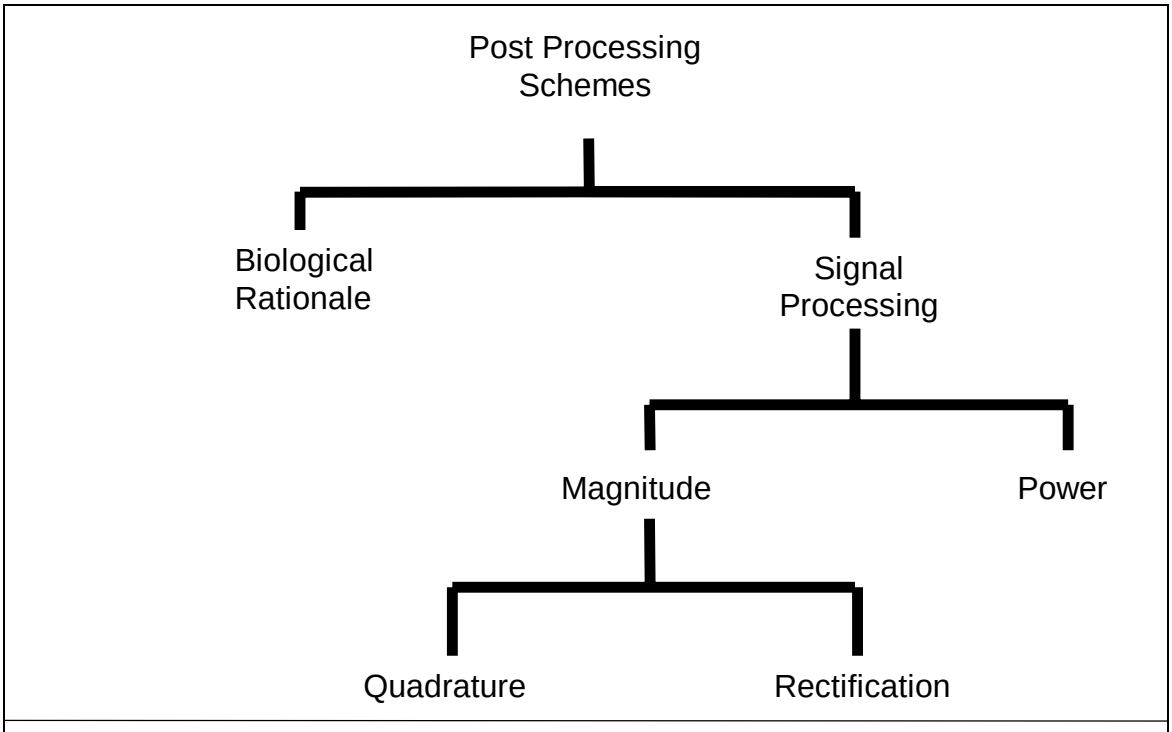


Figure 5.4.6 Taxonomy of post-processing techniques.

(*tanh*) as the non-linearity. Their argument was based on the Julesz texton theory [Jain91]. This proposed that textures are discriminated on the attributes of elongated

blobs or textons. The preattentive visual system cannot determine the location of terminations but can count their numbers or their first order statistics. Most of Julesz's work was conducted with binary images and critics have pointed out that the theory is not directly applicable to most realistic textures. Jain proposes Gabor filters as the first element of a mechanism to extend Julesz theory to naturalistic textures. The rapidly saturating *tanh* function will force the filtered output into an almost binary form—effectively composed of blobs. A low pass filter will then be used to measure the density of the textons. Randen and Husoy [Randen94] as well as Tang et al. [Tang95] also use a *tanh* non-linearity, though Tang pushes the biological analogy further by replacing the averaging filter with non-linear interaction both within and between feature images. He forms an analogue between feature maps and neurone-fields: pixels in a feature image are excited by other pixels in a surrounding annulus, and inhibited by pixels from a feature image of different orientation. This process is stepped through an unspecified number of generations before finally a dominant feature is declared. The biologically based techniques are not mathematically tractable. Consequently, we opt for the more tractable techniques employed by the signal processing based approaches.

Signal Processing Based Schemes

In this section we will consider post-processing techniques that are, or can be, stated in terms of signal processing terminology. In order to make this discussion as thorough as possible, we will not confine ourselves to the post-processing of Gabor filters, but rather to the output of any type of bandpass filter used in texture analysis. The post-processing algorithms we classify as being signal processing based measure, or approximate, one of two signal quantities: the signal magnitude or the quantity known as texture energy.

We do not believe the use of the term energy to be physically meaningful in the sense in which it has been applied and its use in the literature is not consistent. Instead we prefer to use the term *signal power*. Schemes which use this technique have the general form shown in *Figure 5.4.7*.

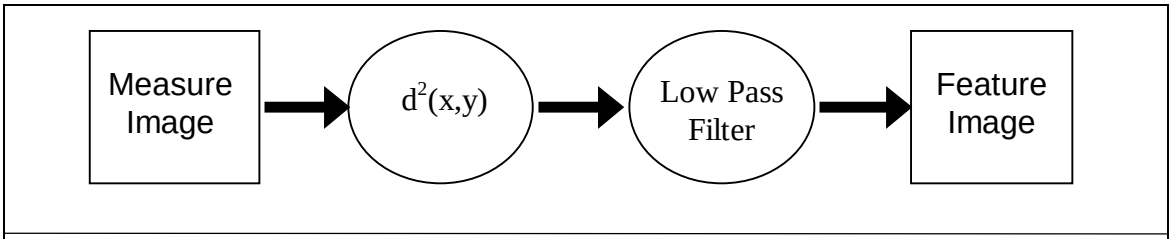


Figure 5.4.7 General form of feature based on channel power measurement.

Laws and Ade both use power measurement (as well as magnitude approximations) in [Laws79] and [Ade83]. Randen et al. also use power-based features in [Randen95] and [Randen96]; in the latter paper this allows the derivation of an expression for signal mean and variance given a description of the bandpass and low pass filters. Livens et al states that the majority of wavelet schemes also use this approach, [Livens97].

Unser and Eden [Unser90] advocated the use of a second non-linearity, taking either the log or the square root of the low passed feature, *Figure 5.4.8*. Their motivation for doing so was to make features, which have been passed through different non-linearities more comparable, simplifying subsequent feature reduction and clustering work, although they do note the tendency towards a Gaussian distribution. We note that this is equivalent to a power transform where the random variable is raised to a power to produce a more Gaussian distribution, [Fukunaga p.76]. We have noted a modest improvement in classification accuracy with a quadratic classifier whose features which have been subjected to a power transform. While the signal power is an attractive quantity on which to base a feature, in our experience, the large variance associated with the approach can introduce problems of stability with some, numerically sensitive, discriminants. Unser's approach does avoid these problems, as well as giving a more Gaussian output. However, it does not relate to any physical concept such as power.

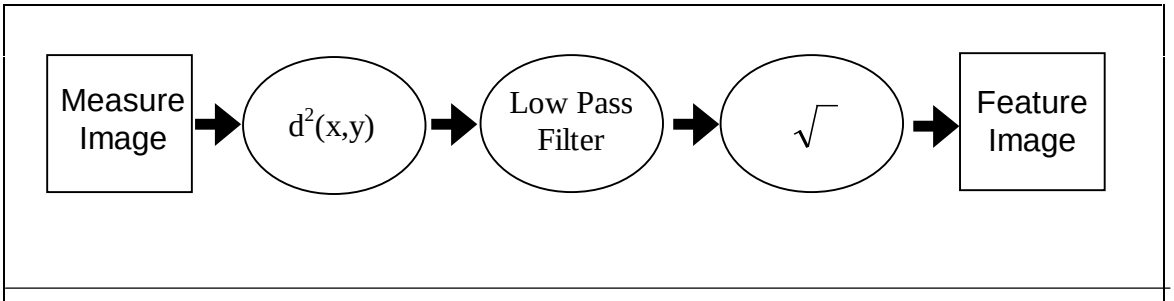


Figure 5.4.8 Unser's scheme.

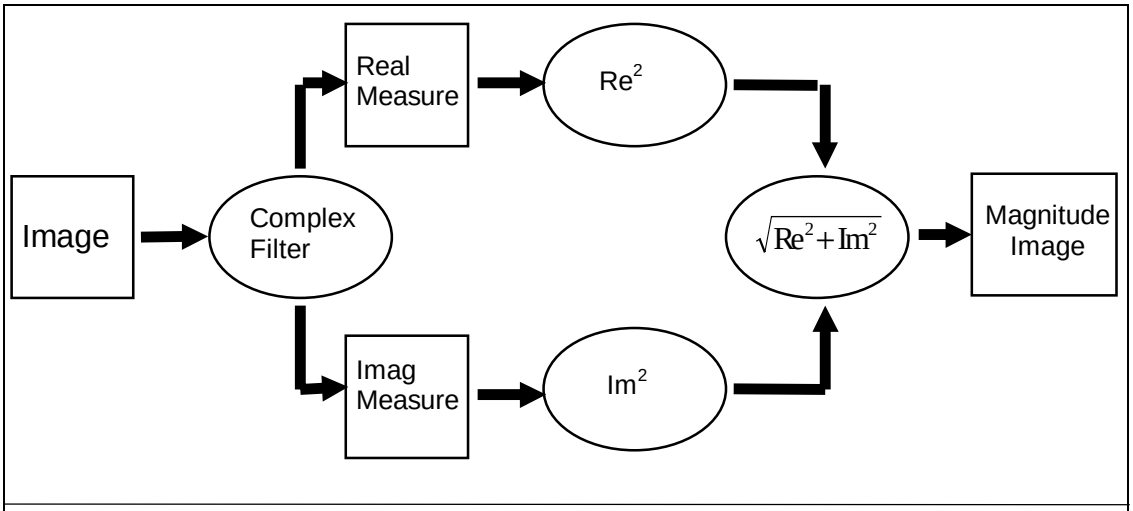


Figure 5.4.9 Magnitude estimation using quadrature filtering.

Like signal power, signal magnitude is a physically meaningful concept, though, unlike power, it is in a range which is comparable to that of the original image. This means that numerical stability is less of an issue than for power, and outlying samples are less significant. Two approaches to the measurement of the signal magnitude have been used in the literature. The first calculates the magnitude response directly from the quadrature response of real and imaginary filters, *Figure 5.4.9*, [Bovik91]. The second method approximates the magnitude by low pass filtering the absolute value of the output of a real filter, *Figure 5.4.10*, [Randen94].

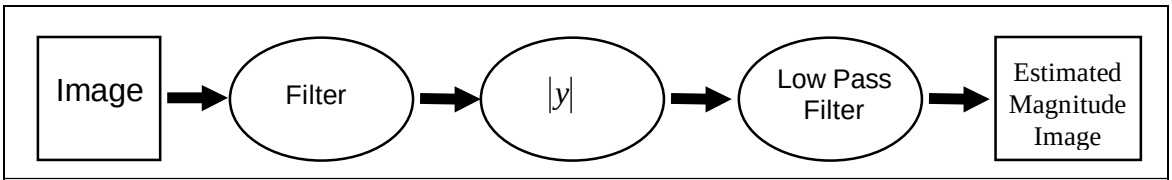


Figure 5.4.10 Magnitude estimation using ABS approximation.

Aach et al. consider an analogue of the rectifier detector with the absolute value of the feature output being low pass filtered to produce the feature image [Aach95]. They then compare this approach with the quadrature filtering approach where textures are classified by magnitude and in some cases also phase from complex feature images. Aach considered the approximation, common in communications theory, by which the magnitude of real and imaginary components is approximated by averaging the absolute output of a real filter only. He notes that low pass filtering the quadrature images produces a result almost identical to the texture energy technique. He concludes that:

1. for edge based techniques where localisation is less of a requirement and smoothing is normally carried out the estimation method is sufficient, and
2. for pixel-based classification, which may be adversely affected by smeared transitions between areas of different textures caused by low pass filtering, quadrature filters are more appropriate.

We adopt the quadrature filtering approach for the following reasons.

1. Our scheme is pixel-based and the optimal space/frequency characteristic of the Gabor filter only holds for the complex form of the function.
2. The calculation of magnitude from quadrature filters makes explicit our suppression of phase in favour of magnitude information as the basis for classification.
3. This thesis is primarily concerned with an analytical approach. Implementation efficiency is less important than tractability.

In fact, we have found that even where the texture is filtered in quadrature, the class feature distributions have unacceptably large variances, presumably due to the filtered images not being sufficiently narrowband. It is therefore necessary to filter the magnitude image before classification; this conclusion is, in effect, supported by Kiernan [Kieran95]. She calculated magnitudes from complex features but used an averaging filter before classification. In this work we shall adopt a feature set based on the signal magnitude calculated from quadrature filters. Let us model the output from the Gabor filters as a zero mean Gaussian process. While the second assumption follows from our random surface models and the linearity of the transforms up until this stage, the non-admissibility of the Gabor filters, discussed in the wavelet section (*Section 5.3.4*), means that the signal will have a non-zero mean. We justify the assumption on the grounds that the mean is small relative to the signal [Novarro95].

The pdf of the resultant of two uncorrelated Gaussian processes will follow a Rayleigh distribution (Couch p.546). *Figure 5.4.11* shows the pdf of the modulus of the response of a complex Gabor filter to an illuminated fractal surface. The theoretical Rayleigh distribution is superimposed onto the histogram. After low pass filtering, however, the distribution more closely approximates the Gaussian case (*Figure 5.4.12*) — satisfying one of the optimality conditions of the discriminant.

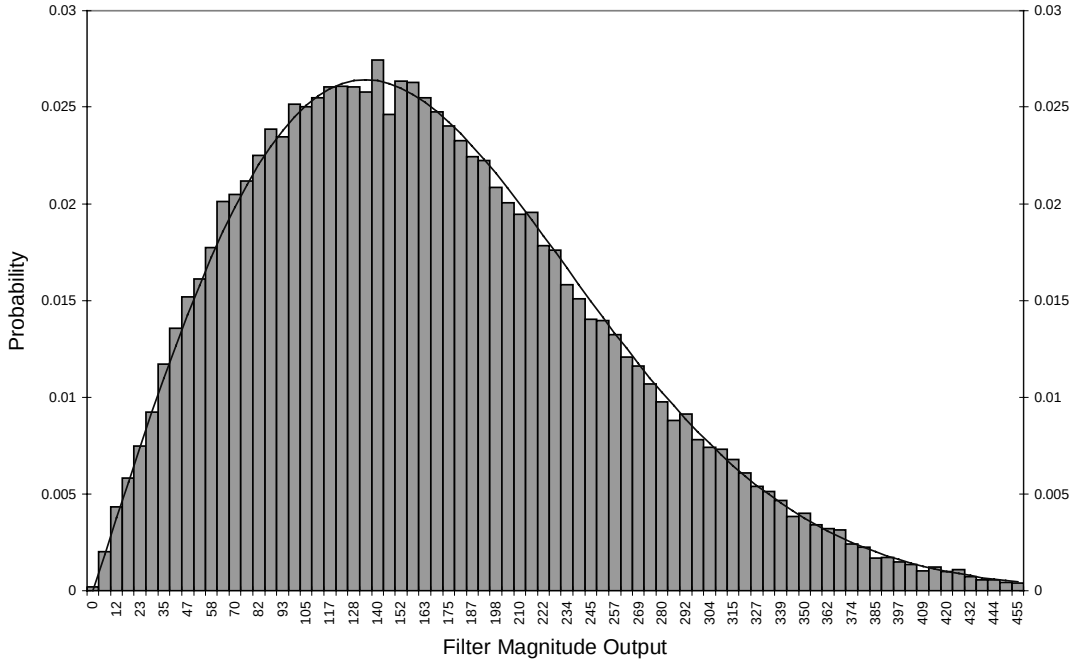


Figure 5.4.11 Pdf of modulus of quadrature filter outputs.

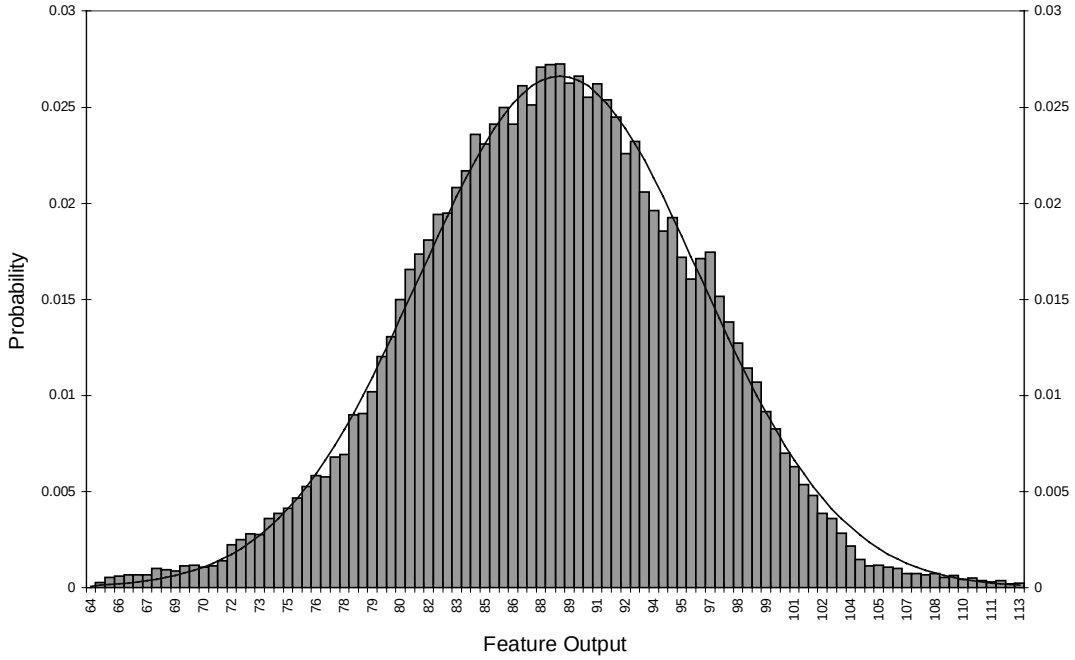


Figure 5.4.12 PDF of Feature Image

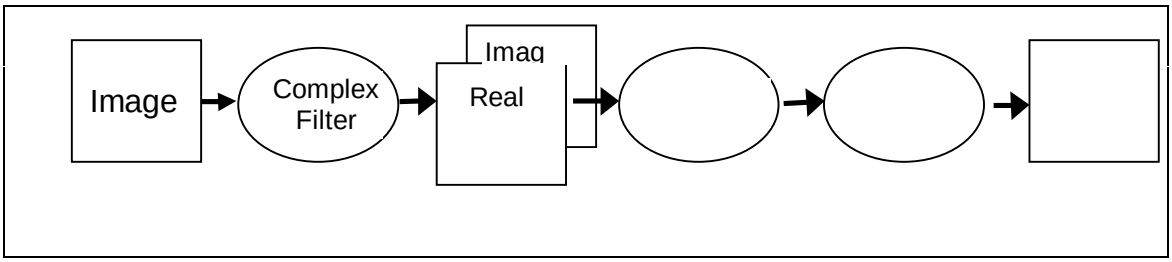


Figure 5.4.13 Post Processing

To summarise, we shall adopt the post-processing scheme shown in *Figure 5.4.13.*

5.4.4 The Discriminant

Having developed the evidence upon which the classification is carried out, a mechanism to perform the actual decision is required.

Choosing The Decision Algorithm

The nature of the problem which this thesis addresses immediately restricts us to supervised classification, i.e. there exist *a priori* classified training examples. We therefore do not consider the important area of unsupervised classification.

Several candidate techniques have been used for supervised classification in the texture analysis literature. Neural networks have been applied to texture analysis: Greenspan et al. evaluated a back-propagation network [Greenspan94], while other investigators have proposed novel networks that incorporate spatial interactions e.g. [Tang95]. Greenspan et al. also evaluated the K-nearest neighbours algorithm, as did Ohanian and Dubes [Ohanian92]. Unser applied traditional Bayesian classification [Unser95] whereas Weldon et al. used a Bayesian classifier modified to model texture classes as mixture distributions [Weldon96]. Given the application, it is clear that spatial interaction of pixel labels is an important cue to segmentation. Aach et al. used a region growing technique [Aach89] while a growing number of authors explicitly incorporate into their classifiers the effect the labelling of its neighbours has on the label probability of that pixel [Song92][Raghu96].

Both neural networks and the K-nearest neighbour techniques are widely used, however, both are largely intractable to analysis. Statistical techniques are much more responsive to analysis. We therefore adopt the statistical approach to classification due to its compatibility with the overall approach of this thesis. As mentioned earlier, decision algorithms incorporating spatial information into the labelling process are becoming