The Classification of Textured Surfaces Under Varying Illuminant Direction

G. McGunnigle B.Eng.

Thesis submitted
for the
Degree of Doctor of Philosophy

Heriot-Watt University
Department of Computing and Electrical Engineering

June 1998

Table of Contents

Chapter 5 A Classification System

Chapter 6 Modelling the Classifier Tilt Response

Chapter 7 Addressing the Problem

Chapter 8: A Simulation-based Approach to Tilt Effects

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisor Dr Mike Chantler, firstly for the opportunity to conduct research, and secondly for his encouragement as well as his willingness to ask awkward questions. I would also like to express my gratitude to Dr. Laurie Linnett whose earlier work fundamentally influenced the approach of this thesis and whose software was critical to this work.

My thanks are also due to Dr. Tariq Khan and Mr Thomas Wyatt whose help in the preparation of this thesis has been invaluable. I am also indebted to the ERASMUS students Tilmann Wittig, Michel Damoiseau and Gregory Delguste for their help and stimulation. I would also like to thank the Chief Technician Alistair Houston who maintained an enthusiastic and helpful attitude to experiments which must have seemed at least slightly bizarre.

Finally, I wish to express my gratitude to my parents whose support, particularly in the last few months, has been vital to the completion of this thesis.

Principal Symbols

Signals

Spatial Quantities		Spectral Quanity	Description
Scalar Field	Vector Field		
s(x,y)		S(u,v)	Surface Height
	S(x,y)	P(u,v) Q(u,v)	Surface Derivatives
	L(x,y)		Illuminant Vector
i(x,y)		I(u,v)	Incident Image
d(x,y)		D(u,v)	Measured Data Set
$d_{f\phi}$(x,y)		$d_{f\phi}$(u,v)	Output of filter f,ϕ
	D(x,y)		Filter Outputs Vector
$f_{f\phi}$(x,y)			Feature Response derived from filter f,ϕ
	F(x,y)		Feature Vector
l(x,y)			Label field
n(x,y)		N(u,v)	Noise process
e(x,y)		E(u,v)	Residue Process

Transfer Functions

Spatial Variable	Spectral Variable	Input	Output	Function
o(p,q)		Surface derivative field	Image field	Reflectance function
	R(u,v)	Surface height spectrum	Image spectrum	Illumination
b(x,y)	B(u,v)	Image Spectrum	Data set Spectrum	Imaging
g(x,y)	$G_{\omega\phi}$(u,v)	Data Set Spectrum	Output of filter ωϕ	Gabor filter
	H(u,v)	Surface spectrum	Measure Spectrum	Combined Filter

Surface Parameters

σ_s	Rms Roughness
R_{cla}	Centre line average
m_{rms}	Rms Slope
p_{rms}	Rms slope in the x-direction
q_{rms}	Rms slope in the y-direction
m_{fg}	f^{th} and g^{th} order statistical moment.
β	Power Roll-off
k	Topothesy
λ_0	Correlation length of an isotropic surface
λ_1	Correlation length in the x-direction
λ_2	Correlation length in the y-direction
ω_0	Fundamental frequency
ω_c	Cut-off frequency

Surface Variables

t	Lag
ω	Radial frequency
θ	Polar frequency angle
u	Horizontal frequency index
v	Vertical frequency index
p	Facet slope in the x-direction
q	Facet slope in the y-direction
p_x	Second derivative of surface, in the x-direction.
q_x	Second derivative of surface, in the x-direction.
s(x)	Surface height profile.

$r_c(t)$	Autocorrelation function
$c(t)$	Autocovariance function

Illumination Variables & Parameters

σ	Slant angle
τ	Tilt Angle
R	Correlation matrix of the surface
V[a b c]	Least squares linear model of the illumination process.
i_d	Desired image
a b c	Parameters of optimal linear model.
k_1 k_2 k_3	Parameters of Kube's linear model.

Imaging Parameters

σ_b	Blur
γ_b	Exponent of camera amplification.
σ_t	Standard deviation of temporal noise.
$\sigma_{disparity}$	Standard deviation of the difference between two images
σ_n	Standard deviation of overall noise process.

Gabor Filter Parameters

ϕ	Direction of propagation
σ_x	Standard deviation of the Gabor filter envelope in the x-direction.
σ_y	Standard deviation of the Gabor filter envelope in the y-direction.
u_0	Centre frequency of filter in the x-direction.
v_0	Centre frequency of filter in the x-direction.
B_ϕ	Polar bandwidth of the filter.
B_r	Radial frequency bandwidth of the filter.
σ_p	Measured polar bandwidth

σ_p	Standard deviation of the Gabor filter spectrum in the x-direction.
σ_y	Standard deviation of the Gabor filter spectrum in the y-direction.

Feature Parameters

σ_m	Standard deviation of measure image.
μ_r	Mean of Feature image
σ_r	Standard deviation of Feature image.
a	Parameters of feature/tilt model.
b	" "
M_n	Mean vector of class *n*
Σ_n	Covariance matrix of class *n*
$p_{\mathbf{F}\|l_i}(\mathbf{F}\|l_i)$	Probability that a vector x belongs to class *n*
$k(F\|l_i)$	Probability that a vector x belongs to class *n* over the entire tilt range.

Symbols Associated with Compensation Schemes

$m(\omega)$	Parameters of Chantler' filters.
$b(\omega)$	" "
i_0,i_{90},i_{180}	images obtained from τ=°,90° and 180° respectively.
i_{NL}	Non-linear component of surface to image mapping.
i_p	Image which is a linear function of p-derivative field only.
i_q	Image which is a linear function of qderivative field only.

Abstract

This thesis sets texture analysis in a physical context. Models of the system components are obtained from the literature and integrated into a description of the process linking the rough surface to the feature set on which classification is based. The first component is the rough surface, models of the surface topography are selected from the fields of tribology and scattering. Various reflectance models are considered and a spectral model of the surface/image relationship from the literature, is evaluated and discussed. The relationship between the incident image and the captured data set is investigated and described. This model is integrated with the spectral description of the feature measures to form a model of the transition from surface to feature set.

It is clear from this model that the direction of illumination can affect the directionality of an image obtained from a given surface. Changes in the illuminant direction will result in changes in the feature outputs. If the illuminant direction is altered between training and classification, the classification rule may be inappropriate and classification poor. Several schemes are considered to combat this problem. A technique which uses a representation of the physical surface as the basis for the generation of appropriate training data is selected for further evaluation. The surface derivative fields of the training surface are estimated using photometric techniques. A rendering algorithm uses these estimates to simulate the appearance of the training surface when it is illuminated from an arbitrary direction. It is shown that where illuminant direction is varied this system is able to perform significantly better than a naive classifier, and in some cases approaches the level of accuracy obtained from training the classifier under the conditions at which classification is performed.

Texture analysis is a significant area in the field of machine vision, this is in large part due to the important role of texture in the early visual system. It follows from this that texture has been seen as being critical to general visual systems working in unconstrained environments, consequently, less emphasis has been placed on more controlled inspection tasks. In an unconstrained system it is impractical to adopt a modelling approach and most work in texture analysis takes the image as its starting point. This thesis is concerned with the inspection of rough textured surfaces. By making explicit the circumstances under which classification occurs we are able to employ modelling of the system and describe texture classification in the context of the physical system which gives rise to a textured image.

Chapter 1

Introduction

1.1 Motivation

This thesis is concerned with the application of texture analysis to discriminate between rough, textured, surfaces. In *Figure 1.1.1* we give an example where this ability is required. A fossilised ammonite is embedded in a rock matrix. Both the ammonite and the underlying rock surface have their own, distinct, surface structure which are apparent in the image. This thesis treats the problem of identifying the fossil and isolating it from the background surface as one of detecting a desired signal buried in noise.

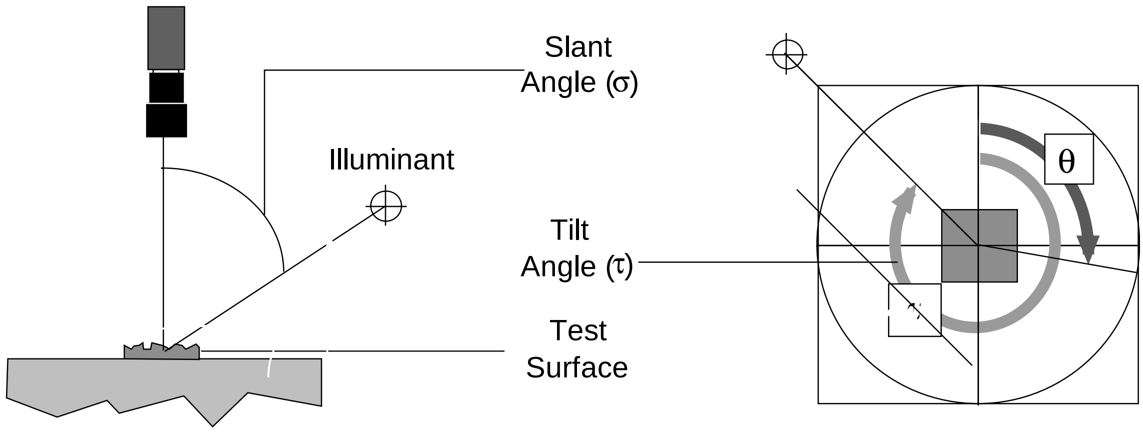
Figure 1.1.1 An ammonite (190MYA) fossilised in rock matrix.

Implicit in this approach is the assumption that the observed image of a surface is a function of the surface and is either independent of other variables, or that these variables are constant. In practice, perceived texture may be due to a projected illumination pattern, surface markings, the interaction of illumination with a rough surface or any combination of these. This work is concerned with the automated classification of textures belonging to the third group.

While this type of texture is formed by the interaction of illumination and rough surface, it is desirable to classify the surface type regardless of illumination conditions.

Consequently, work has been carried out on texture classification that is invariant to changes in the illuminant intensity [Thau90] and colour [Healey95]. However, with the exception of [Chantler94], the effect due to variations in illuminant direction is largely neglected in the published literature.

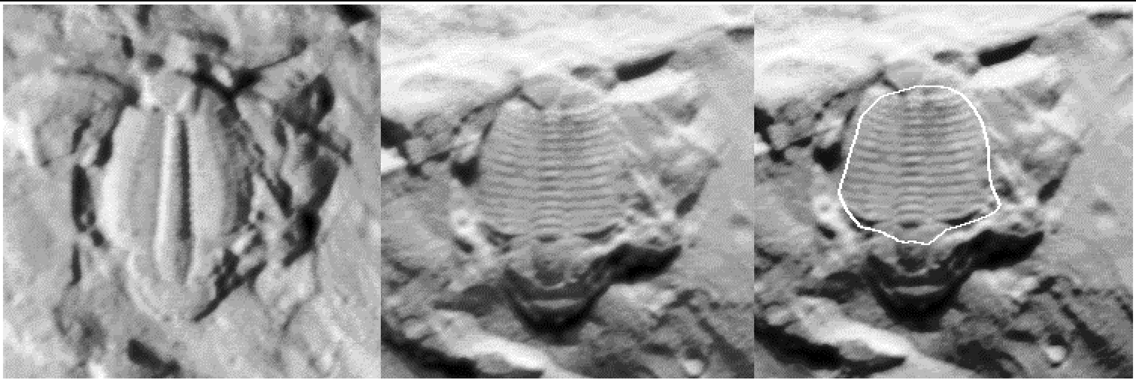
The direction of the illuminant with respect to the texture may be defined by two polar co-ordinates: slant and tilt. Slant is the angle between the camera axis and the illuminant vector; in this work it shall be held constant at 60°. Tilt refers to the polar angle of the illuminant on a plane normal to the camera axis, *Figure 1.1.2a.* The tilt angle (τ) is also shown in *Figure 1.1.2b* with the coplanar general orientation parameter θ. The aim of this work is to maintain consistently accurate classification of textures under changes in the illuminant tilt angle.



(a.) *(b.)*

Figure 1.1.2 Illuminant tilt and slant angles.

We illustrate the importance of illuminant direction for rough surface discrimination with the following example. Consider the fossilised trilobite (*Elrathia Kingii, 550 MYA, Utah*) shown in *Figure 1.1.3*a. Here the fossil is illuminated from tau=90° and there is little texture information that can be used to segment the fossil from the surrounding matrix. Illuminated from 0° (*Figure 1.1.3b*), however, the animal's pleurae become visible as vertical undulations, allowing the fossil's thorax to be identified using a vertically oriented Gabor filter (*Figure 1.1.3c).*



(a)Fossil Illuminated from Tau 90°. (b)Fossil illuminated from Tau 0°. (c) The segmented image.

Figure 1.1.3 Classification of a trilobite fossil (550 MYA).

Illuminant tilt is evidently important in attenuating or accentuating directional information that may be critical to classification. As a consequence of this, a shift in the tilt direction will alter the image texture of a given surface. It follows that a classifier which has been trained using images of surfaces obtained at a given tilt angle—which classifies well under these conditions—may not be able to accurately classify the same surfaces illuminated from different tilt angles. The goal of this thesis is to develop a classifier which will be able to classify a surface regardless of the illuminant tilt angle at which the surface was imaged.

1.2 Texture Analysis: A Brief Overview

In this thesis we will use techniques developed in texture analysis to discriminate between rough surfaces based on their visual appearance. Texture analysis, itself, has a long lineage and represents the confluence of several disciplines. In this section we briefly note the main areas which have provided either the motivation to investigate texture, or the techniques which have been used to do so.

Biological visual systems are biased towards the detection of rapid changes of intensity over either time or space, since these stimuli are most likely to be important to the organism. As a consequence of the sensitivity to spatial variation, textures are particularly useful stimuli for the investigation of visual systems. Furthermore, as Gibson stressed, biological visual systems are adept at inferring surface orientation from texture cues [Gibson50]. In fact, Gibson saw texture as being more than a visual cue but rather as having a fundamental role in his ecological view of perception[Bruce p.226]. As Bruce and Green put it "*the input for a receptor is a stream of photons, but the input for a perceiver is a pattern of light extending over space and time*" [Bruce p.376]. Gibson's work analyses the relationship between the optical pattern and the observer's environment. Texture and optical flow are considered the primary sources of information about the environment [Gibson79]. Perhaps the most convincing evidence of the biological importance of texture is its ability to affect visual systems using disruptive pattern camouflage, where texture is used to 'jam' or drown out visual cues such as outlines.

Biological systems have also contributed to texture analysis in a less predictable way. The use of a disruptive camouflage pattern requires pigmentation cells to 'switch' to one of two or more colourings. The cells are identical, and have no centralised form of communication, i.e. they have no sense of their position, yet they must switch colouring in a spatially organised manner. Turing hypothesized that the mechanism by which they do so is due to the concentration of levels of certain chemicals. The processes by which the chemicals diffuse across the surface and react with each other give rise to the global spatial pattern [Turing52]. Recently, several algorithms that simulate the reaction diffusion effect to produce textures have been developed for computer graphics [Turk90][Witkin90]. Picard includes reaction diffusion as an important class of textures in her, "society of (texture) models", and notes its ability to synthesize textures comprised of spots and stripes [Picard96].

In a typically sized image there may be approximately a quarter of a million data elements—an equivalent volume of data to a census of a small city; consequently it is hardly surprising that statistical methods should have been applied. What is special about texture images, from the statistical point of view, is the spatial interaction of both neighbouring and distant elements. Texture images have allowed the investigation of spatial interaction models in a much more visual and intuitive way. The spatial interaction may be described in terms of second order probabilities and several methods have been developed which attempt to parsimoniously extract the information held in the joint distributions. Co-occurence techniques e.g.[Haralick73] extract features from the distributions, whereas Markov models e.g.[Chellappa85] have been used to model the second order probabilities. In either case the algorithms are generally limited to observation of the interactions between immmediately neighbouring pixels.

A large class of textured images may be considered as two-dimensional stochastic signals and consequently fall into the realm of random signal processing. Analogously to the statistical approach, the interaction and interdependence of neighbouring pixels has motivated the use of techniques which are capable of predicting the likely value of a pixel's intensity from those of its neighbours. Autoregressive models have been used, both on their own and with Moving Average models, to both model and discriminate textures [Kayshap84]. Furthermore, the requirement to localise what is an essentially non-local phenomenon has motivated the use of time/frequency analysis. One of the main areas of research has been the generalisation of one dimensional signal processing algorithms to two dimensions. Wavelets [Livens97], Wigner-Ville distributions [Song92] and many other techniques have followed this path.

With the partial exception of those techniques that use texture as a cue to orientation, all the approaches mentioned above take the data set as their starting point. In this thesis we shall apply texture analysis techniques to a data set which does have a specified physical meaning—that of a rough surface interacting with light. By treating the imaged texture as a function of surface topography, which will be used to discriminate between surfaces, we must account for certain variables in the physical process that may affect the data set and our classification of it. The goal of this thesis is to develop a classifier that is immune to variation of one of these variables: *illuminant tilt.*

1.3 Scope

In defining the scope of this research it is useful to adopt a hypothetical example: consider an inspection system tasked with segmenting a rough, textured, surface into regions with uniform surface characteristics. This will be accomplished using an overhead CCD camera, directional lighting and an algorithmic classification system (*Figure 1.3.1*).

The surface is considered to be globally flat, i.e. each of its partial derivative fields sum approximately to zero. Furthermore, the surface reflectance function is considered to be diffuse, and uniform throughout the surface. The light rays incident on the surface are considered to be parallel and of equal intensity across the imaged region. This is consistent with a point source located at infinity or light originating from a parabolic reflector. The imaging device is assumed to be a monochrome CCD camera located directly overhead the texture sample. In addition, we shall assume the topography of the surface is small relative to the distance between the camera and the surface, and that the projection of the surface onto the CCD array is orthographic.

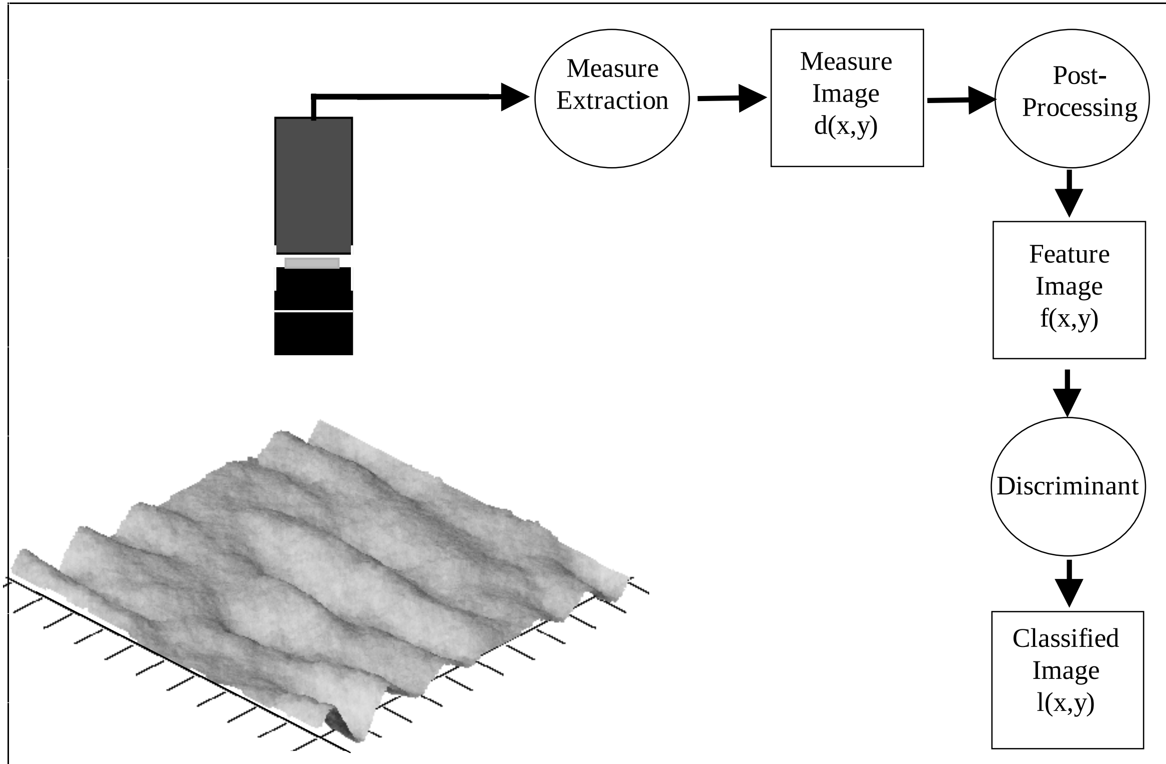


Figure 1.3.1 The classification system.

Whereas most work in texture analysis has taken the stored data set as its starting point, the work described here is concerned with modelling the transitions from surface to image to numerical and finally symbolic representation. This thesis is specifically concerned with modelling and remedying the effect of changes in the direction of the illuminant tilt for the purposes of surface classification.

1.4 Original Work

This thesis makes three main contributions to the field of texture analysis:

- The first contribution of this thesis is the combination of physically verified models, drawn from several areas within the field of computer vision and others external to it. The combination of these elements allows a much more analytical approach to texture analysis within the scope of this thesis. The analytical approach advocated in this thesis allows the systematic development and evaluation of texture analysis algorithms in the context of the actual physical system.The application of physics-based vision techniques was pioneered in the field of shape from shading, however, its application to texture analysis, as described here, is original. Pentland published a series of papers linking fractal surfaces to observed textures

[Pentland84,85,86 and Kube88]. He did not present any physical justification for the models of the surface or the reflectance function he used. Healey has been active in many of the fields considered in this thesis, such as reflectance [Healey89], imaging artefacts [Healey94] and texture analysis [Speis96]; he has not, however, attempted to relate these areas to each other.

- We have described the effect of illuminant tilt on a set of Gabor features and shown that this effect may introduce a significant degree of misclassification. The tilt induced misclassification problem was first noted by Chantler [Chantler94]. By adopting feature measures that are consistent with the surface and image models, we believe we have both verified many of Chantler's findings and also formed expressions that are more coherent and intuitively understandable.
- An important contribution of this thesis is the development of a model-based technique to reduce the effect of tilt changes on classifier accuracy. We have verified the performance of the system for tilt variation and believe it offers an effective and physically meaningful mechanism for the suppression of tilt effects.

The combined result of these contributions may be stated as follows; a physical phenomenon was accurately predicted and the effects at the algorithmic level modelled. Using our understanding of the physical system, a technique was developed which reduces the effect of the physical phenomena on the computional level. In this way, the goal of the thesis, i.e. to develop a classifier robust to illuminant tilt, is achieved.

1.5 Thesis Organisation

This thesis may be divided into two chapter groupings. The first group comprises Chapters 2 to 6 and develops a theoretical framework for the classification of rough, textured, surfaces on the basis of their visual appearance. One of the predictions of the model is that the visual texture, as well as being a function of the surface, will be affected by the illuminant tilt. The second part of this thesis, Chapters 7 and 8, focuses on techniques for the consistent classification of rough surfaces robust to illuminant direction.

Development of an analytical model is broken into the chapter aims and objectives shown in *Table 1.5.1*. Analysis begins in Chapter 2 with the rough surface. The literature associated with tribology and scattering theory is consulted to obtain firstly, a means of describing rough surfaces, and secondly, models of rough surfaces framed in terms of that form of description. In Chapter 3 a spectral model of the relationship between the surface description and the image incident on the camera lens is developed. An existing model [Kube88] is evaluated in terms of optimality and of accuracy. The conversion of this image into a usable data set is carried out by camera and framestore, and is modelled in Chapter 4. This chapter forms an analytical model of the imaged data and develops a noise engine with estimated parameters of the conversion process. The classifier developed in Chapter 5 operates on the data set and converts the two dimensional signal into symbolic form. This represents the final component in the conversion from physical surface to symbolic representation. However, in Chapter 6 it is shown that varying illuminant direction between training and classification can render the classifier useless; the second part of this thesis examines techniques designed to remove this dependency.

Thesis Goal	*To develop a tilt invariant classifier*				
Thesis Sub-goal	*To form a model of the effect of tilt*				
Chapter	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6
Aims	To model the topography of rough surfaces.	To model the process of image formation.	To model the imaging device.	To develop a texture classifier suited to rough surfaces.	To model the classifier.
Objectives	To find: • a means of description • models of rough surfaces	• To describe the local effect • To develop a spectral model.	To model: • The device transfer function • Sources of noise	Selection of appropriate algorithms for: • Measure extraction • Post-Processing • Discrimination	Modelling the effect of tilt on: • Measure spectra • Feature statistics

Table 1.5.1 Aims and objectives of first chapter grouping.

The second chapter grouping considers several schemes designed to tackle the problem of tilt dependency. Chapter 7 considers three techniques advanced by Chantler [Chantler94] as well as the feasibility of applying shape from shading techniques. In fact, whereas some of these schemes do reduce the effect of tilt on classification, none eliminates the effect. In Chapter 8 a novel technique is proposed, which models the surface and the processes analysed in the first part of this thesis during its training stage. The technique is evaluated by simulation using various model surfaces, the empirical reflectance map and the noise engine. The scheme's performance on real data is then evaluated at each process step before final evaluation on the criterion of misclassification.

Finally, Chapter 9 summarises the findings of the previous chapters and draws together the conclusions of the work carried out in this thesis.

Chapter 2

Modelling Rough Surfaces

2.1 Introduction

The goal of this thesis is to develop a classifier that is suited to the discrimination of rough surfaces, and which is robust to changes in illuminant tilt. It is the belief of the author that this goal can best be achieved by gaining an understanding of, and an ability to model, the physical, as well as the computational aspects of this problem. This chapter considers the first data structure (or signal) in the surface to symbol chain, the surface topography. In subsequent chapters we shall consider the transformations of this signal that occur at higher levels in the process of recognition.

The aim of this chapter is to develop a physically based model of the surface. The first objective is to arrive at a method for characterising textured surfaces and to define the limitations of this technique. Having found a means of description, the second objective is to adopt a group of models which are defined in terms of that method. These models will then form the basis of subsequent analysis of the recognition process.

To attain these two objectives, this chapter is structured in the following way. The first section consists of a literature survey of methods for the description of rough surfaces. The method adopted, the power spectrum, forms a complete description of a surface in conjunction with the phase spectrum. The next section develops a stochastic model for the phase spectra, which includes an *ad hoc* definition of the textures that lie within the scope of this thesis; this section also describes a simple test to determine whether a surface complies with the phase model's definition. Having adopted the power spectrum as the surface descriptor, we next consider rough surfaces models defined in terms of the power spectrum. Three models are adopted: a linear roll-off (or fractal) model; a fractal model modified to have a flat spectrum below a parameterised cut-off frequency; and a model that varies the cut-off frequency with polar angle to introduce surface directionality. These models are discussed with reference to experimental findings reported in the literature.

2.2 Rough Surface Description

2.2.1 A Brief Review of Possible Sources of Models

Several fields of research have been surveyed to consider descriptive techniques and to construct a suitable model. One source considered was visual texture analysis, where investigators have used a wide variety of techniques such as Markov modelling [Chellappa85], ARMA modelling [Kayshap84] and Gabor filtering [Jain91]. Whereas early approaches in texture analysis were based purely on discrimination, more recently, techniques which attempt to model textures have come to the fore. Furthermore, most techniques are inherently two dimensional and explicitly consider anisotropy. There are, however, two drawbacks to the use of texture analysis techniques. Firstly, most are designed to segment an image and are consequently optimised for the localisation of a texture, rather than its accurate representation. The second and more serious difficulty stems from the fact that these techniques have been developed for visual textures. Applying them to an application outside their original context has the consequence that reported parameter values cannot be used to model surfaces, so any surface model would lack experimental verification in this important respect. Texture analysis techniques will therefore not be considered as a means of either surface description or modelling.

Another potential area of interest is terrain modelling, e.g. [Austin94]; recently interest in modelling land surfaces has increased, largely motivated by the development of fractals and microcomputers. Because the subject is comparatively new, there is little literature and theoretical framework compared to some of the more established areas such as scattering. The physical basis of the application, and the overlap in descriptive methods, does however, make this a possible source of realistic modelling data.

The texture of surfaces is of interest in many engineering applications. The interaction of light with surfaces which have been “optically finished”, i.e. manufactured to an accuracy comparable with the wavelength of light, is of considerable interest in scattering theory. Surfaces manufactured to less exacting standards are also studied, usually with reference to their frictional properties in the field of tribology[1]. The combination of theoretical background and the emphasis on actual physical measurements makes these fields the most attractive areas to survey, and this chapter will be restricted to these fields.

Both scattering theory and tribology use surface description as a stepping stone towards subsequent analysis of their respective properties of interest. Despite this common interest in surface description, the fields are distinct and often use different parameters. In general, the descriptive techniques favoured in scattering theory, e.g. the Power Spectral Density (PSD) and rms roughness, are more suited to this research than their counterparts, e.g. the Autocorrelation Function (ACF) and centre line average, used in tribology. On the other hand, tribological surfaces exist on a scale that is much closer to, and which may actually overlap with, that at which a vision system is likely to operate. It follows that surfaces of this scale are more likely to resemble those which a vision system will encounter and are therefore more reliable sources of model parameters. Fortunately, there is a large degree of overlap in the subjects and we choose to treat both simultaneously.

A description of a surface may be made on several levels: a single parameter may be sufficient to characterise a surface for some purposes; in other cases a much greater degree of description is required. The organisation of this section follows an increase in the sophistication of the description and its ability to define a surface; the flow of the section is in consequence approximately chronological. We have resolved three distinct levels of description. The first level seeks to characterise some property of the surface with a single parameter such as the root mean square (rms) height or gradient. On the second level, a statistical model, in the form of a histogram, is applied to the variation of height or gradient. This is a natural extension to rms values, i.e. the standard deviation of surface height and slope, but in another respect it represents a paradigm shift since it allows the use of surface models, such as Beckmann and Spizzichino's height model [Beckmann63] or Torrance and Sparrow's slope model [Torrance67]. At the third level in our taxonomy are those techniques which incorporate spatial interaction, such as the PSD and the ACF. Given certain assumptions discussed elsewhere in this chapter, these form the basis for models of a wide range of surface textures.

[1] Tribology is the study of friction, lubrication and wear between moving surfaces.

2.2.2 Single Parameter Description

The most basic form of profile description requires the use of only one parameter. The 1930's saw new emphasis placed on profile measurement and a large number of measures were developed around this period [Parsons, p.268]. These measures vary in the degree to which they are specialised to a particular application and in their mathematical tractability. We will consider the two most common and general measures of roughness, *rms roughness* σ_s and the *Centre Line Average* (or *average roughness*) R_{cla}, in addition we will discuss the much less popular *rms slope* parameter.

σ_s (2.2.2a) and R_{cla} (2.2.2b) have, broadly speaking, been used in different fields; σ_s roughness has tended to be used to describe optical finish, whereas R_{cla} is more commonly associated with machined surfaces [Bennett89 p.39]. Despite this, even in the field of machining, most theoreticians prefer to use σ_s, since it allows the use of statistical random process techniques (these will be discussed later in this chapter). It is worth noting that for surfaces which have height distributions conforming to the normal distribution, there is an approximate relationship between the two parameters (2.2.2c). In this thesis, however, the preferred parameter will be *rms roughness* due to its compatibility with statistical analyses.

$$\sigma_s = \sqrt{\frac{1}{n}\sum_{x=o}^{n}\left[s(x)-\overline{s(x)}\right]^2} \qquad (2.2.2a)$$

where *s(x)* represents the height of the surface at a point *x* along the profile.

$$R_{cla} = \frac{1}{n}\sum_{x=0}^{n}\left|s(x)\right| \qquad (2.2.2b)$$

$$R_{cla} \approx 0.8\ \sigma_s \qquad (2.2.2c)$$

The *rms slope* parameter, m_{rms} (2.2.2d), is much less commonly used than either of the other measurements. Bennett states that the measured *rms slope* may vary by a factor of 50 depending on the profilometer and the separation of sampling points [Bennett89 p.40]. Whereas *rms roughness* is commonly used in physical scattering models, such as Beckmann and Spizzichino [Beckmann63], *rms slope* has been used in geometric optics models such as Torrance and Sparrow [Torrance67]. Despite the problems associated with this parameter, it is useful in the context of this thesis. The same texture analysis technique may be applied to images ranging from electron microscopy to satellite imagery; absolute scale is in many cases relatively unimportant. The slope of a facet is more important than its height, and the *rms slope* provides a parameter particularly suited to the purposes of this research.

$$m_{rms} = \sqrt{\frac{1}{n}\sum_{x=o}^{n}\left[s'(x) - \overline{s'(x)}\right]^2} \qquad (2.2.2d)$$

where *s'(x)* is the derivative of the function *s(x)* at point *x*.

In practice, the values of both the *rms roughness* and *slope* depend on the measuring instrument. The length of the sample, the area resolution and the sampling rate all affect the parameters—this being especially true of the *rms slope*. Strictly speaking, neither the *rms slope* nor the *rms roughness* are defined for non-bandlimited surfaces.

2.2.3 Histogram Description

In retrospect, it now seems a natural progression from considering the *rms roughness,* i.e. the standard deviation of the heights of a surface, to the adoption of a statistical model of height distributions, though this development does not seem to have followed quickly. We believe that use of histogram description represents a significant shift in the way surfaces were thought about since it not only implies a new degree of discrimination between surfaces but also allows a certain degree of modelling.

The histogram does give an insight into the nature and history of a surface. Surfaces which are the result of a large number of random events tend to have a Gaussian distribution. Wear, grinding and abrasion tend to wear down and deform summits while leaving valleys largely unaffected, which is reflected in the histogram taking on a negative skew [Bennett89 pp.42-44]. Some turning and milling operations and the presence of relatively large particulate matter on an otherwise smooth surface will result in the histogram taking on a positive skew. It is worth noting, however, that use of skew, and of higher moments of the distribution in general, is subject to artefacts in the sampling process.

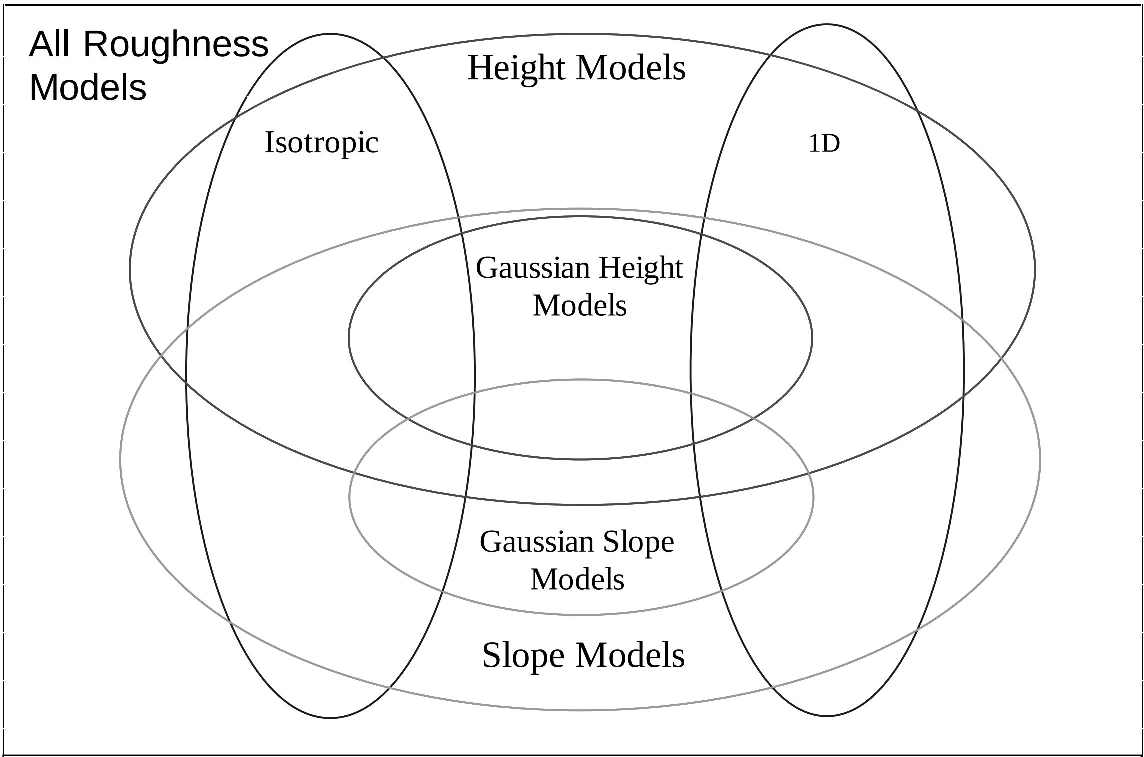


Figure 2.2.1 Stone's taxonomy of roughness models

The histogram has been used to model surfaces: Beckmann uses a normal distribution of heights in his scattering model, while Torrance uses a normal distribution of slopes for his geometric model of specular reflection from rough surfaces. Stone [Stone94] develops a taxonomy of models used for reflectance modelling in machine vision, shown in *Figure 2.2.1*. Stone defines two important classes, *isotropic* and *one dimensional* shown as the ellipses with vertical major axes. However, the major division is between models defined in terms of either slope or height. Important subclasses of each are defined by the Gaussian constraint. Since differentiation is a linear operation, Gaussian height models have Gaussian slope distributions, though unless integrability is enforced, the converse may not be true. While first order statistical surface models have proved successful in these areas, their lack of spatial information means that they are clearly inadequate for the representation of texture.

The statistical parameters of the height distribution have been further exploited in conjunction with some spatial information. The two point height probability distribution, analogous to the co-occurrence matrices used in texture analysis, allows the description of correlation and structure within the profile. The related technique of Markov chains, again with related applications in texture analysis, has also been applied to profile description

[Ogilvy p.29]. However, these techniques have not been widely used in surface description.

2.2.4 PSD and ACF

The forms of profile description discussed above do not take into account any spatial correlation within the surface. This is a profound limit to their ability to characterise surfaces to any satisfactory degree. At the next level of description the surface profile is treated as a realisation of a random process. The tools with which Wiener laid the foundations of random process theory: the autocorrelation function (ACF) and the power spectral density (PSD) are now applied to the surface profile. The autocorrelation function is the most popular approach in production engineering, while the PSD is more commonly associated with scattering theory, due to the surface spectrum's close relationship to the scattered intensity of light from the surface.

Assume the surface profile is wide sense stationary and has zero mean. The autocovariance function (ACVF) (2.2.4a) is the average product of the heights of two points on a profile separated by a distance *t*, known as lag. The autocorrelation function (2.2.4b) is the normalised (by surface variance) form of the ACVF[2].

$$c(t) = E[s(x)s(x+t)] \tag{2.2.4a}$$

$$r_c(t) = E\left[\frac{s(x)s(x+t)}{\sigma_s^{\,2}}\right] \tag{2.2.4b}$$

where *t* is lag.

In *Figure 2.2.2* we plot the ACF for various surfaces. The ACF for a white noise process approximates an impulse at zero lag; the ripple texture, shown in *Figure 2.2.3a*, has a gradually decreasing periodic ACF, whereas that of the fractal surface (*Figure 2.2.3b)* falls monotonically as lag increases. The form of the decay for real surfaces has been the subject of much controversy and several models have been proposed; the most common models are the exponential (2.2.4c) and the Gaussian (2.2.4d). The exponential form appears to be the better fit to the experimental data, however, it does suffer from a discontinuity at zero lag; Ogilvy and Foster [Ogilvy89a] show that an exponential decay is equivalent to a linear roll-off for at least part of the power spectrum. In practice, many

[2] These definitions are based on those used in scattering theory texts, [Ogilvy] and [Bennett], signal processing texts, e.g. [Therrien] define the autocorrelation and autocovariance functions as:

surfaces have ACF which do not conform to either form. Consider the ACF of the ripple surface (the image of which is shown in *Figure 2.2.3a*), this exhibits zero crossings and a large positive peak, suggesting that the surface contains a strong periodicity at that lag.

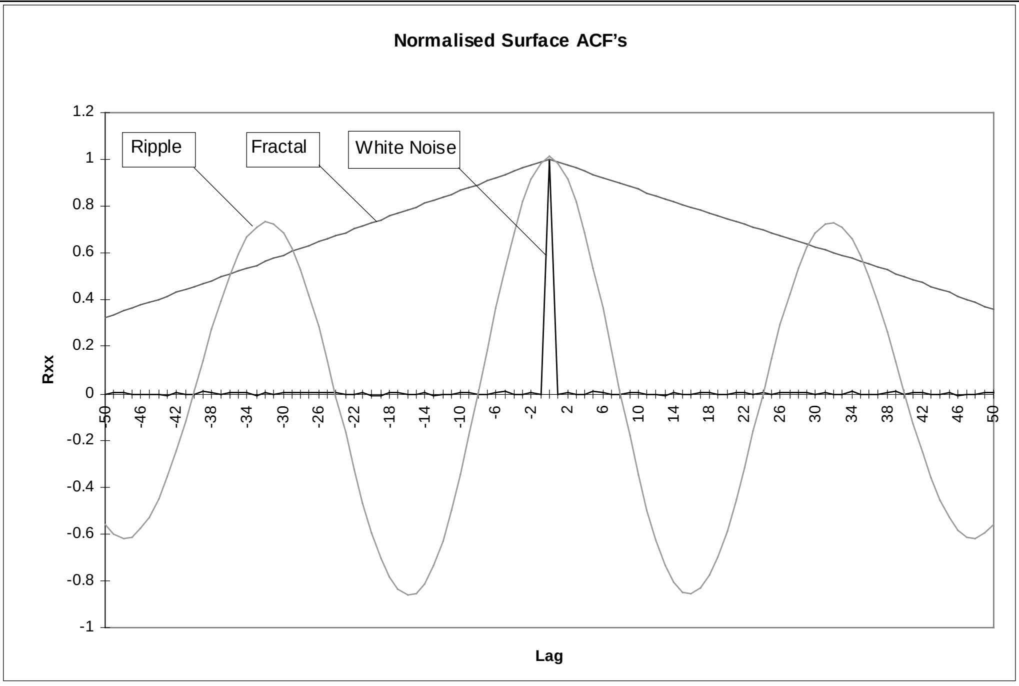
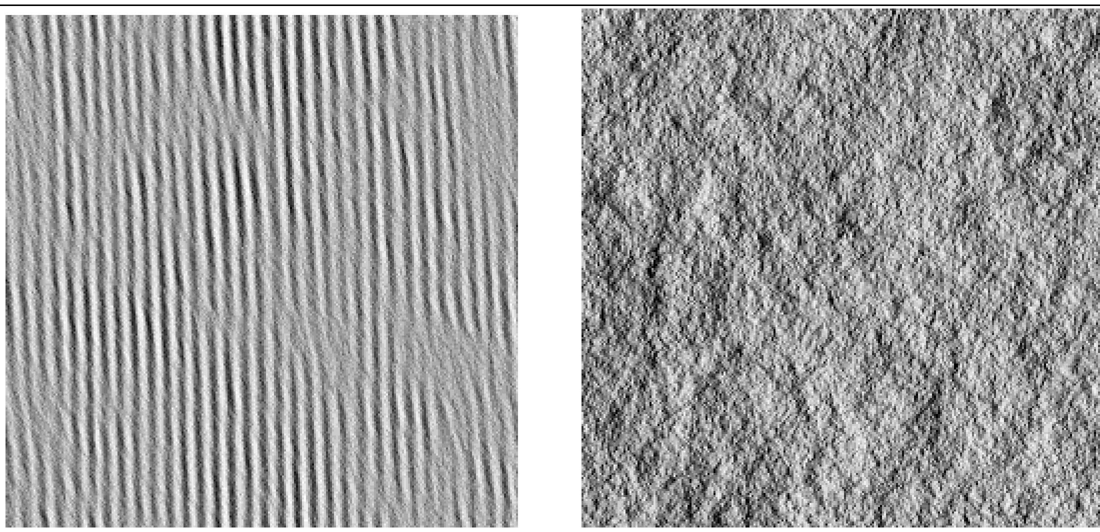


Figure 2.2.2 The 1D ACF for various surfaces



(a.) *(b.)*

Figure 2.2.3 Synthetic ripple(a) and fractal (b) textures.

$r(t) = E[s(x)s(x+t)]$ and $c(t) = E[(s(x) - \overline{s(x)})(s(x+t) - \overline{s(x)})]$ where $\overline{s(x)} = E[s(x)]$.

$$r(t) = \exp\left(\frac{-|t|}{\lambda_0}\right) \qquad (2.2.4c) \qquad\qquad r(t) = \exp\left(\frac{-t^2}{{\lambda_0}^2}\right) \qquad (2.2.4d)$$

where λ_0 is the correlation distance.

The power spectral density is equivalent to the ACF and forms a Fourier transform pair with the ACVF—it is also easily related to the *rms roughness* (2.2.4e) and slope(2.2.4f) of a bandlimited surface. Equations (2.2.4e) and (2.2.4f) also illustrate that these parameters are not defined for non-bandlimited fractal surfaces.

$$\sigma_s^2 = \int_{-\infty}^{\infty} S_{1d}(\omega)d\omega \qquad (2.2.4e)$$

$$m_{rms}{}^2 = \int_{-\infty}^{\infty} \omega^2 S_{1d}(\omega)d\omega \qquad (2.2.4f)$$

where $S_{1d}(\omega)$ is the power spectrum of the surface profile.

While the rms roughness is an isotropic parameter, the rms slope will vary with direction for a anisotropic surface. In this thesis we will assume that the dominant directionality of the surface is aligned with the image axes, and the directionality of the slope distribution may be parameterised using the rms slopes of profiles taken in the direction of the x and y axes, p_{rms} and q_{rms} respectively. The two dimensional equivalents of (2.2.4f) are shown in (2.2.4g) and (2.2.4h).

$$p_{rms} = \int_{v=-\infty}^{\infty} \int_{u=-\infty}^{\infty} u^2 S\ (u,v) du.dv \qquad (2.2.4g)$$

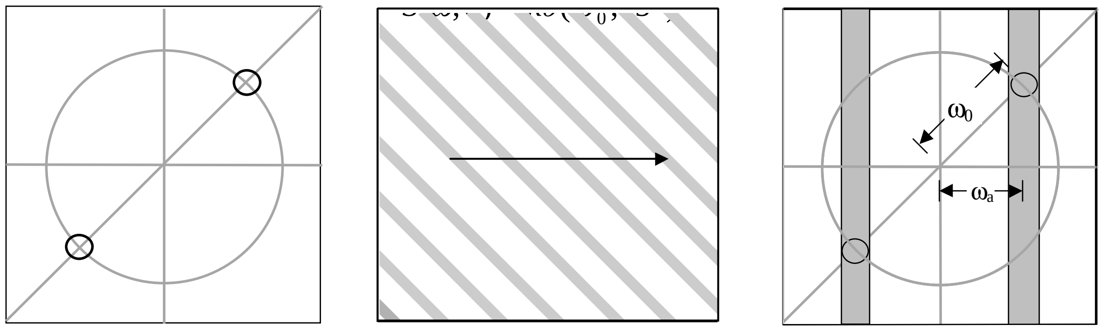
$$q_{rms} = \int_{v=-\infty}^{\infty} \int_{u=-\infty}^{\infty} v^2 S(u,v) du.dv \qquad (2.2.4h)$$

where u is the frequency in the x-direction

and v is the frequency in the y-direction

2.2.5 Relationship between Profile and Surface Spectra

Most work in tribology and scattering deals with one dimensional profiles, whereas we are primarily interested in two dimensional surfaces. The generalisation from one to two dimensions is not trivial and we give a brief description of the process based on two papers [Nayak71] and [Church83]. We begin with the calculati from a surface PSD. Consider a sinusoidal surface (shown in *Figure 2.2.4b*) with PSD:



(a.) (b.) (c.)

Figure 2.2.4 Calculation of profile spectrum of 2D spectrum.

Now, consider a profile taken in the direction of the x-axis, *Figure 2.2.4b*, this will be sinusoidal with frequency ω_a, such that $\omega_a < \omega_0$. More generally, the power spectrum of a profile at a given frequency will be the sum of the two dimensional PSD at that frequency and direction and the projections of the two dimensional PSD at all directions and frequencies that project onto this component *Figure 2.2.4c*. Again assuming the profile is in the direction of the x-axis, mathematically this can be stated as:

$$S_{1d}(u) = \int_{-\infty}^{\infty} S(u,v)\,dv \qquad (2.2.5a)$$

Consider now a two dimensional surface, such that the two dimensional power spectral density is dependent only on radial frequency ω, i.e. the surface is isotropic.

let

$$\omega = \sqrt{u^2 + v^2} \qquad (2.2.5b)$$

$$S_{1d}(\gamma) = 2\int_{\omega=\gamma}^{\infty} S_{2d}(\omega).\frac{\omega}{(\omega^2 - \gamma^2)}\,d\omega \qquad (2.2.5c)$$

where γ is the frequency parameter in the one dimensional power spectrum.

If we wish to calculate the two dimensional signal given a profile PSD, we must either assume extreme anisotropy or isotropy. Let us consider the latter case.

Given a one dimensional profile with spectrum $S_{1d}(\gamma)$, we may calculate the PSD of the two dimensional spectrum $S_{2d}(\omega)$, along one direction using:

$$S_{2d}(\omega) = -\frac{1}{\pi}\int_{\gamma=\omega}^{\infty}\frac{d}{d\gamma}S_{1d}(\gamma).\frac{1}{(\gamma^2-\omega^2)}d\gamma \qquad (2.2.5d)$$

2.2.6 Summary

In this section we have surveyed methods of surface description used in tribology and scattering theory. Three levels of description were discussed. It was argued that single parameter description is not sufficient for our purposes, though the rms slope parameter does measure a surface characteristic which is relevant. The surface height histogram is usually assumed to be unimodal Gaussian, in which case it is closely related to the rms parameters. The histogram, however, like the single parameters, lacks any description of the spatial interaction of points on the surface and therefore does not contain sufficient information for texture description. Second order statistics, i.e. the PSD and the ACF, which incorporate the spatial relationship of surface heights were then considered. The relationship of the 1D and 2D forms of the power spectra was discussed, and an analytical expression from the literature for the isotropic case was quoted.

The evidence considered in this section suggests that the power spectrum should be adopted as the principle means of description. However, the rms slope was also found to be a useful surface parameter and will be used throughout this thesis. The surface height histogram, and consequently the slope histogram, will be assumed to be Gaussian.

2.3 An Admissibility Criterion Based on Phase

In the previous section it was stated that the most sophisticated level of surface description commonly in use in tribology and scattering is the PSD (or equivalently the ACF). It follows that surfaces with identical power spectra are considered to be identical for the purposes of the analysis. This thesis is concerned with distinguishing between surfaces that have dissimilar appearance under identical conditions. The sufficiency of the PSD-only approach must therefore be considered in this context. Following from our PSD-only approach we must impose certain bounds on the phase spectra of the surfaces which we will consider in this thesis.

Our next objective is to determine for what type of surface the condition holds. Unfortunately, we do not possess a data set of surface topography. We do however have a series of textured images each of which form a two dimensional data set sharing many of the textured characteristics of the surface. By applying a test of the phase condition to this data set, we will identify those characteristics which either include or exclude a particular data set from the scope of this thesis.

We now consider some general requirements of the phase condition. Firstly, the model must be sufficiently specific to allow discrimination, yet general enough to allow the grouping of similar textures. A further requirement is that the assumptions of the model should easily be met for synthesis. Finally, the degree to which the model is appropriate for a particular texture must be measurable. With these requirements in mind, next we consider the nature of a possible model.

2.3.1 A Phase Condition

It is well established that images with the same PSD may be easily discriminable: both the power and phase spectra are required for complete reconstruction of the image. This applies equally well to the surface. If we discriminate between surfaces solely on the basis of their PSD we cannot *guarantee* a correspondence with any reasonable visual discrimination. One approach to circumvent this problem is to adopt a model of the phase spectra and to restrict our research to textures which satisfy the model's assumptions. If the model is appropriate, we will be able to extract the PSD of a texture and use it, together with a random realisation of the phase component of the model to generate a texture which can be reasonably said to belong to the same class as the original texture.

We adopt the concept of a maximum entropy phase spectrum, i.e. the phase spectrum should contain no discriminatory information and all characteristics of the texture should be encoded in the power spectrum. Clarke [Clarke92] has shown that the appearance of deterministic textures, i.e. those containing a large degree of structure such as Beans (*Figure 2.3.1)* is more sensitive to changes in phase than that of unstructured textures such as Rock. It is reasonable to conclude that unstructured textures are more effectively defined in terms of a PSD-only model than their structured counterparts. This leads us to predict that the maximum entropy requirement will be fulfilled by unstructured textures; the next step is to formalise this either directly, or indirectly, in terms of the phase spectra.

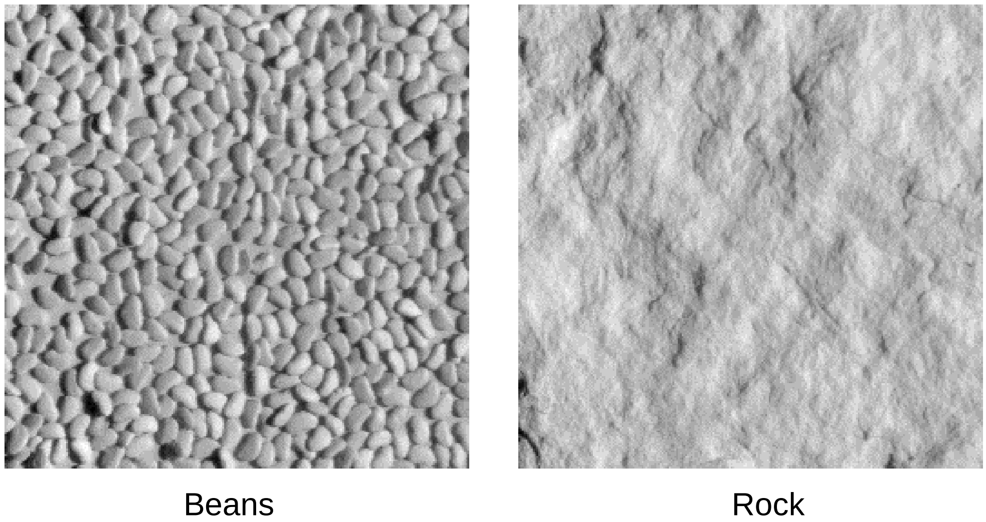


Figure 2.3.1 Two Test Textures

2.3.2 A Simple Test for the Condition

Given the difficulty in interpreting phase information, it would clearly be advantageous to express information held in the phase spectra in another form in order to assess its effect. If we inverse filter the Fourier transform of the original texture with the resultant of its real and imaginary parts, we obtain a scalar field with a uniform PSD. Any structure in the field will be due to the phase spectrum. An unstructured texture will produce an uncorrelated random field, whereas a structured texture will have a structured field. If we can differentiate between these two cases we can discover how well a texture is described by the PSD-only model.

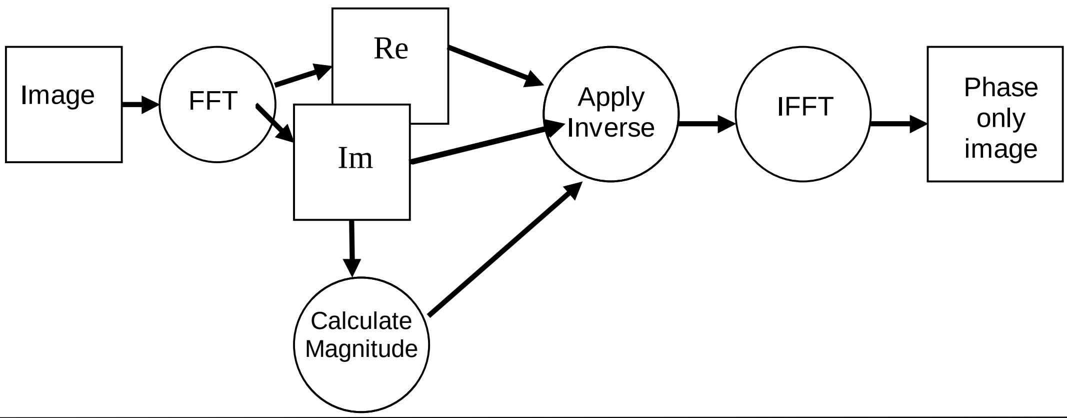


Figure 2.3.2 Recovery of Phase Information

We must now deal with the problem of detecting texture with a uniform power spectral density. Clearly frequency domain models will be inappropriate; we must therefore use probabilistic means. The second order statistics of a spatially uncorrelated random field are completely specified by the first order statistics. If the second order statistics predicted from the mean and standard deviation statistics differ significantly from the measured statistics we may infer that the field is correlated and the texture does not conform to our criteria.

We use the χ^2 goodness of fit test to ascertain whether the measured second order statistics are drawn from the same distribution as those predicted from the first order statistics. We use the following framework:

define i as the intensity of a point (x,y) and g as the intensity at point $(x+\delta x,y+\delta y)$.

For an uncorrelated field the values of i and g are independent and the joint probability function of i and g, $p(i/g)$, is equal to the product of the probabilities of each of the individual probability functions, $p(i)$ and $p(g)$, i.e :

$$p(i/g)=p(i).p(g) \qquad \text{for all } \delta x,\delta y \text{ except } \delta x=\delta y=0$$

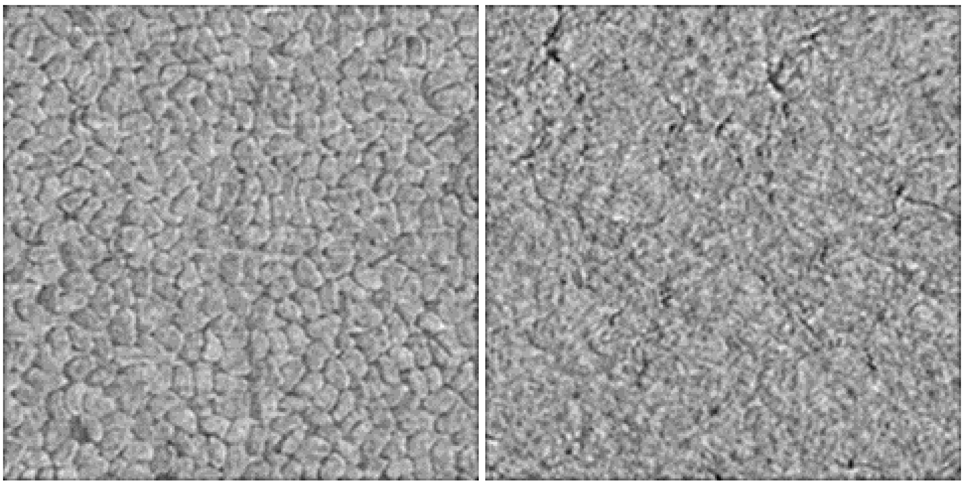


Figure 2.3.3 'Phase-only' Images of the Test Textures

since i and g share the same uniform probability function:

$$p(i/g)=p(i)^2=p(g)^2 \qquad \text{for all } \delta x,\delta y \text{ except } \delta x=\delta y=0 \qquad (2.3.2a)$$

The observed histogram will have $f(i,g)$ elements in each bin (i,g). For a random field with n elements, the expected number of elements in each bin, by (2.3.2a), will be:

$$E(i,g)= np(i/g)=n.p(i)^2$$

The form of the χ^2 test for this application is shown below:

$$\chi^2 = \sum \frac{[f(i,g) - E^2(i,j)]^2}{E^2(i,j)}$$

We adopt the following procedure: the whitened image is requantised to 32 levels and the histogram equalised; allowing all bins to be filled to a level that will avoid the problems associated with the Chi statistic when it is applied to empty or nearly empty bins. The Chi square goodness of fit criteria is then applied to assess the similarity of the function $f(i,g)$ with $E^2(i)$. For a histogram of these dimensions and a 5% level of significance, the Chi square test gives a threshold of 1098. A plot of the statistic for each displacement vector is given in a 3D histogram (*Figure 2.3.4*). Columns exceeding the 5% threshold are shaded more lightly than those that fall below the level.

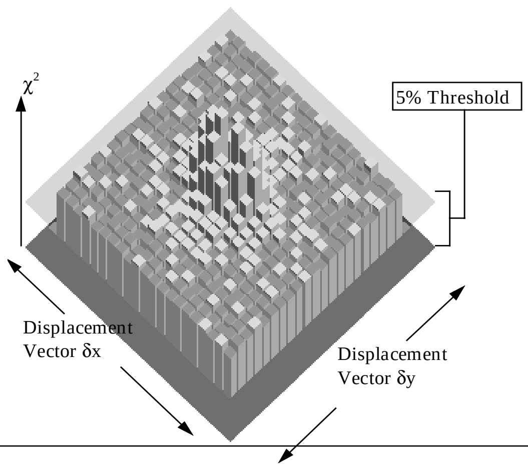
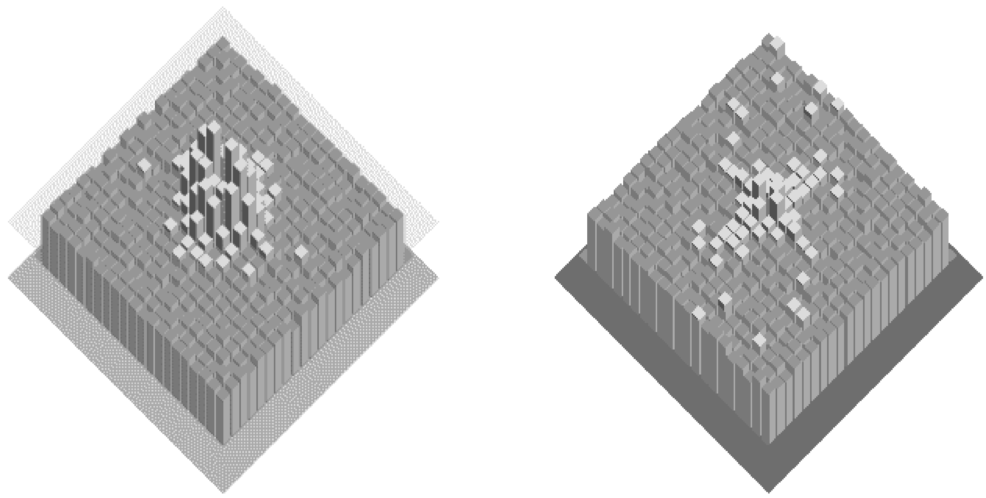


Figure 2.3.4 Chi Square Statistic/ Displacement Vector Histogram

In *Figure 2.3.5* the histograms for the Beans and Rock textures are plotted. The first point we note is that in both cases the highest values are clustered around the zero displacement vector. This indicates that the spatial interactions of pixels in the whitened textures is highly localised, suggesting that both the phase-only textures exhibit the Markov property. Beyond this, comparison of the graphs shows that while both textures do exceed the 5% threshold, the structured texture does so in a much more marked fashion.



(a) Beans *(b) Rock*

Figure 2.3.5 Histograms of Beans and Rock Test Textures

The test is now applied to a wider range of textures, these are shown in Appendix B. Since it is not practical to plot such a wide range of displacements for all the textures, we take advantage of the localisation property observed earlier and define a second order neighbourhood composed of the eight closest displacements. From this neighbourhood we extract two figures: the mean and the maximum value. The results are tabulated in *Table 2.3.1* and shown in *Figure 2.3.6.*

If we consider the statistic we note that the textures do fall into two distinct categories: those with statistics less than 1500 and those with values greater than 2000. The latter group consists of stones, chips and beans, i.e. those textures which we intuitively identified as being structured. It is also worth noting that the anaglypta textures, which consist of a regularly repeated primitive, have among the lowest Chi statistics.

Texture	Mean	Max		Texture	Mean	Max
Isoroc	1033	1055		Maze	975.5	1044
Pitted	938	965		Ripple	1025	1052
Radial	994	1023		Lip	993	1048
Rock	1081	1179		Stripl	1122	1173
Slab	1063	1109		Beans	2103	2537
Slate	983	1023		Chips	2186	2341
Striate	1057	1137		Rock (Chantler)	1304	1589
Twins	1023	1100		Stones	3380	4555

Table 2.3.1 Chi Statistics

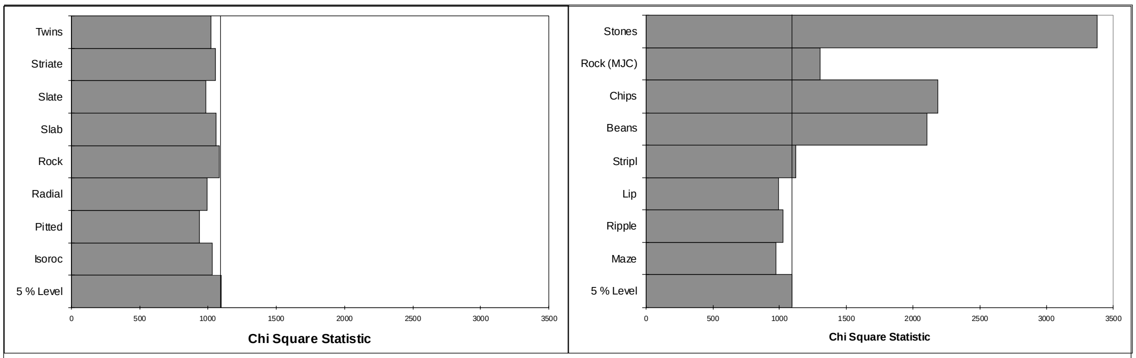


Figure 2.3.6 Histogram of Chi Square Statistic for various textures

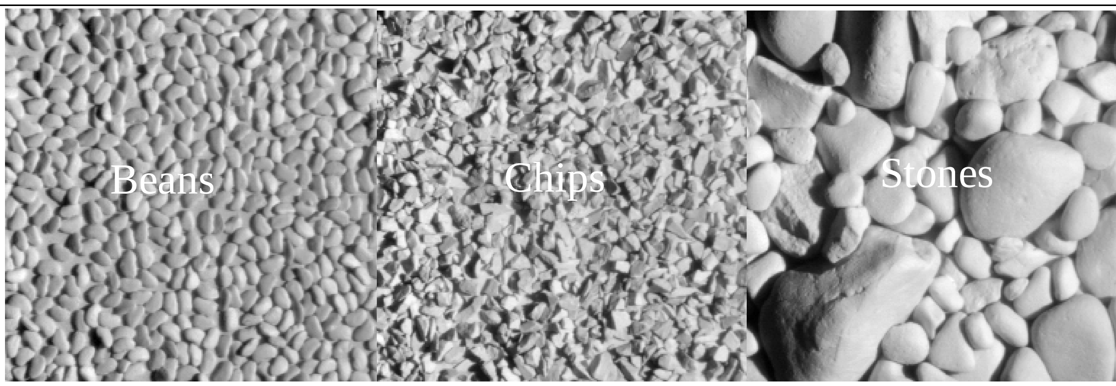


Figure 2.3.7 Phase rich textures.

It is clear that the textures Stones, Chips and Beans do contain important phase information, though the other textures also contain a certain degree of structure which is independent of PSD. The key characteristic which differentiates phase-rich textures is the presence of step changes in the value of the data set. In the image data set this occurs due to cast shadows or the clear delimiting of a texture primitive. In terms of the surface, a step change would manifest itself as a rapid change in height, or 'cliff' topography. This seems to be rare in rough surfaces which have their origin in physical, rather than biological or synthetic processes. The maximum entropy phase condition is therefore compatible with the chosen data set.

By defining random phase textures, we can restrict the scope of this thesis to a class of textures which we can confidently discriminate on the basis of the PSD alone. In the next section we consider models that are expressed in terms of the power spectrum; our interpretation of these models implicitly assumes the phase spectra to be realisations of an uncorrelated random process.

2.4 Models of Surface Roughness

In the previous sections of this chapter we considered some popular techniques for describing engineering surfaces, culminating in the power spectrum. We now look at some models which have been developed from these methods. Two aspects of the model: the *power roll-off* and the *surface directionality* are considered. In the literature these aspects are rarely treated in the same paper: most roll-off models are one dimensional, and directional models are generally stated in terms of *m parameters*, rather than explicitly in terms of the spectrum.

2.4.1 Modelling Roll-Off

The models are generally stated in one of three ways: PSD, ACF or fractal techniques. In this thesis, for the purposes of consistency we will state each model in terms of its PSD. As we would expect, the models do have characteristics in common—an inverse power law is a common feature of most for at least part of the frequency range.

The first model we consider was proposed by Sayles and Thomas in their paper in Nature [Sayles78]. The paper showed that a large number of surfaces, ranging over eight decades of surface roll-off had power spectra exhibiting an inverse square power law with increasing wavelength. The thesis of the paper was that a wide range of surfaces may be modelled with an expression of the form:

2.4 Models of Surface Roughness

In the previous sections of this chapter we considered some popular techniques for describing engineering surfaces, culminating in the power spectrum. We now look at some models which have been developed from these methods. Two aspects of the model: the *power roll-off* and the *surface directionality* are considered. In the literature these aspects are rarely treated in the same paper: most roll-off models are one dimensional, and directional models are generally stated in terms of *m parameters*, rather than explicitly in terms of the spectrum.

2.4.1 Modelling Roll-Off

The models are generally stated in one of three ways: PSD, ACF or fractal techniques. In this thesis, for the purposes of consistency we will state each model in terms of its PSD. As we would expect, the models do have characteristics in common—an inverse power law is a common feature of most for at least part of the frequency range.

The first model we consider was proposed by Sayles and Thomas in their paper in Nature [Sayles78]. The paper showed that a large number of surfaces, ranging over eight decades of surface roll-off had power spectrums exhibiting an inverse square power law with increasing wavelength. The thesis of the paper was that a wide range of surfaces may be modelled with an expression of the form:

$$S_{1d}(\omega) = \frac{2\pi k_t}{\omega^2} \qquad (2.4.1a)$$

It follows that these surfaces can be characterised by the factor k_t, known as *topothesy*. In retrospect, the similarity to fractals is obvious and while Sayles and Taylor do not mention it in their paper, Berry and Hannay do relate Sayles' work to that of Mandlebrot in their reply [Berry78]. Berry goes on to describe Sayles' treatment of data, as 'procrustean' and Sayles himself states that many machined surfaces will have their longer wavelengths suppressed by processing. This remark would seem to characterise many of the subsequent modifications to the linear roll-off model.

A more empirical approach to spectral description adopted by several authors is to split the spectrum into two regions of different fractal dimension. Both [Hasegawa93] and [Oden92] fit two lines: with low roll-off at low frequencies and high roll-off at high frequencies. Imre et al. [Imre93], split the spectrum into three regions, microscopic, fractal and macroscopic. The fractal region has dimension 2.4, while the micro and macroscopic regions have no roll-off. A similar, though more elegant, model was developed by Underwood and Banerji in their paper dealing with fractured steel samples. Instead of a linear fractal curve, they report a reversed sigmoidal curve [Underwood86].

Echoing Sayles' reservations on the power of low frequency components for machined surfaces, Mulvaney et al. have developed an alternative model from experimental data (2.4.1b) and goes onto develop an estimate for surface variance (2.4.1c) [Mulvanney89]. Interestingly the PSD of the exponential correlation function mentioned in section 2.2.4 is of a similar form (2.4.1d) [Ogilvy89a].

$$S_{1d}(\omega) = \frac{k_t}{1 + \left(\frac{\omega}{\omega_c}\right)^2} \qquad (2.4.1b) \qquad\qquad \sigma_s{}^2 = \frac{\pi k_t \omega_c}{2} \qquad (2.4.1c)$$

$$S_{1d}(\omega) = \left(\frac{2}{\lambda_0\sqrt{2\pi}}\right)\frac{1}{\left(\frac{1}{\lambda_0}\right)^2 + \omega^2} \qquad (2.4.1d)$$

where λ_0 is the correlation length, i.e. the distance from the origin it takes the ACF to fall to e^{-1} of its value at the origin. The fractal and Mulvaney power spectra is shown in *Figure 2.4.1*.

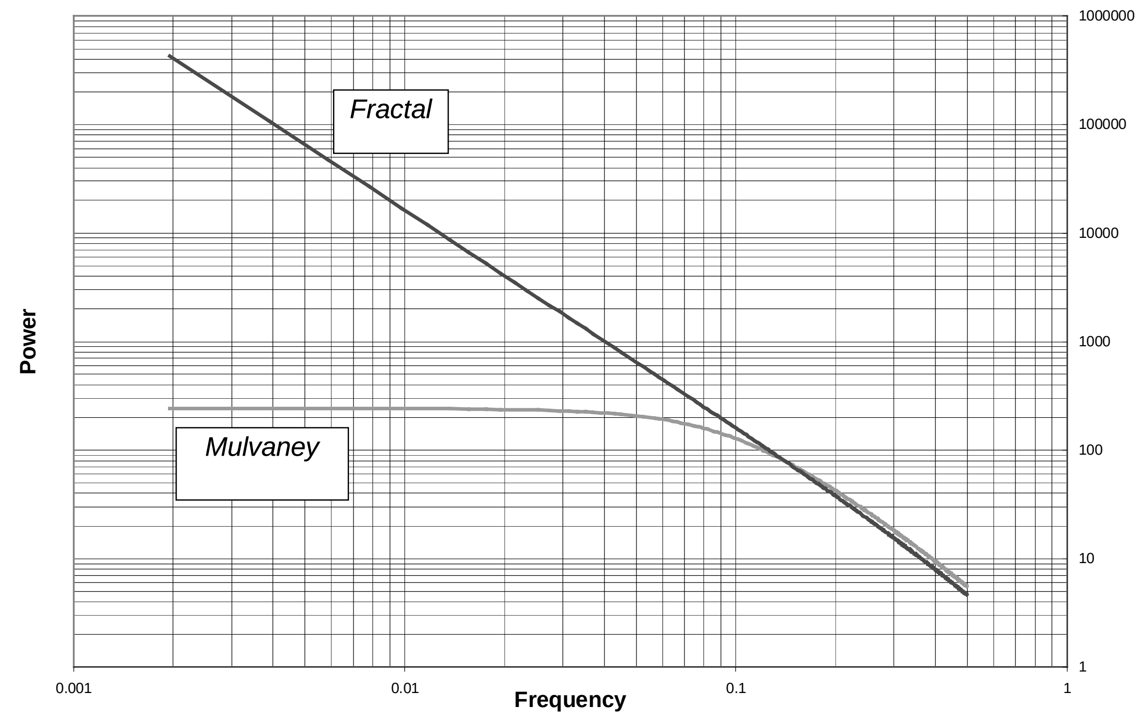


Figure 2.4.1 Power Spectra of Ogilvy and Fractal Surfaces

The illuminated fractal and Mulvaney surfaces are shown in *Figure 2.4.2*, it is interesting to note that, while both are realistic, the Mulvaney surface resembles a rough surface which has undergone a degree of physical processing.

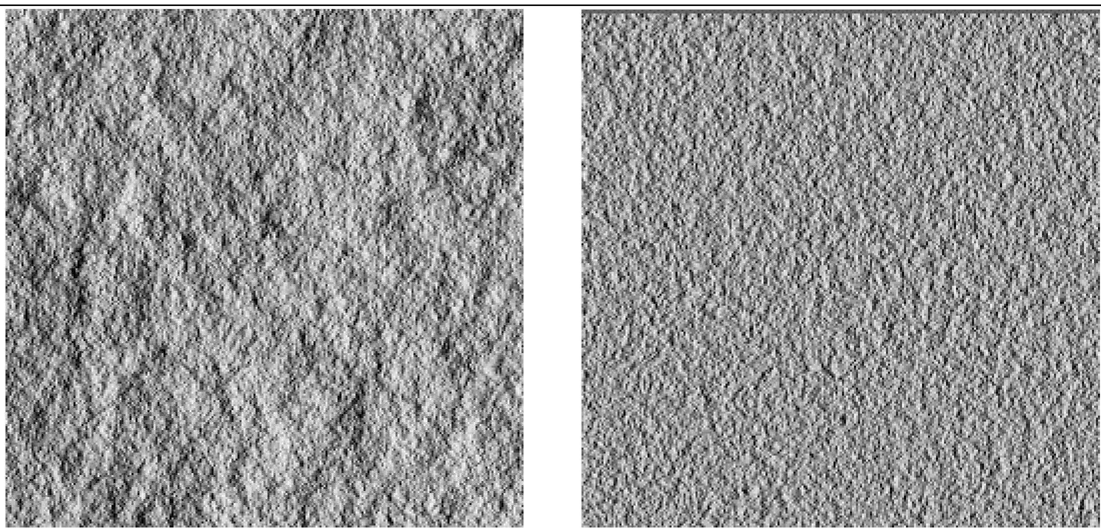


Figure 2.4.2 Isotropic fractal (left) and Mulvaney surfaces (right).

Both the fractal and Mulvaney models are stated in terms of profile spectra, using equation 2.2.5d we can predict the spectra of isotropic surfaces conforming to these models. Substituting the fractal model into equation 2.2.5d leads to:

$$s_{2d}(\omega) = \frac{-1}{\pi} \int_{\gamma=\omega}^{\gamma=\infty} \frac{d}{d\gamma}\left(\frac{k_t}{\gamma^2}\right) \frac{1}{(\gamma^2-\omega^2)^{\frac{1}{2}}} d\gamma$$

$$s_{2d}(\omega) = \frac{-1}{\pi} \int_{\gamma=\omega}^{\gamma=\infty} \frac{k_t}{\gamma^3(\gamma^2-\omega^2)^{\frac{1}{2}}} d\gamma$$

$$s_{2d}(\omega) = \frac{\pi k_t}{2\omega^3}$$

A fractal surface profile with roll-off β_{1d}=2.0 corresponds to a fractal surface with power roll-off β_{2d}=3.0. We may repeat the process for the Mulvaney profile spectrum :

let $\alpha=\omega_c^{-2}$

and $s_{1d}(\omega) = \dfrac{k_t}{\alpha\gamma^2+1}$

$$s_{2d}(\omega) = \int_{\gamma=\omega}^{\gamma=\infty} \frac{k_t}{\left(\alpha\gamma^2+1\right)^2\left(\gamma^2-\omega^2\right)^{\frac{1}{2}}} d\gamma$$

$$s_{2d}(\omega) = \frac{\pi k_t}{4\sqrt{\alpha}} \frac{\omega\sqrt{\alpha+\omega^{-2}}}{\left(\alpha\omega^2+1\right)^2}$$

$$s_{2d}(\omega) = \frac{\pi}{4\sqrt{\alpha}}\left(\alpha\omega^2+1\right)^{\frac{-3}{2}}$$

For ω small, this corresponds to a white noise spectrum, and for ω large to fractal roll-off with β=3.0.

Since this thesis is concerned with the surface image, which is related to the surface facet slopes, we differentiate the two surface models (in any direction) and the plot the power in *Figure 2.4.3*. Both models have a 1/ω characteristic at high frequencies. This does not converge to zero as frequency tends to infinity. Consequently, without bandlimiting the slope variance will tend to infinity. This serves to underline the dependency of the m_{rms} parameter on sampling frequency. Since the model is used in discrete form we assume the surface is bandlimited above the Nyquist frequency.

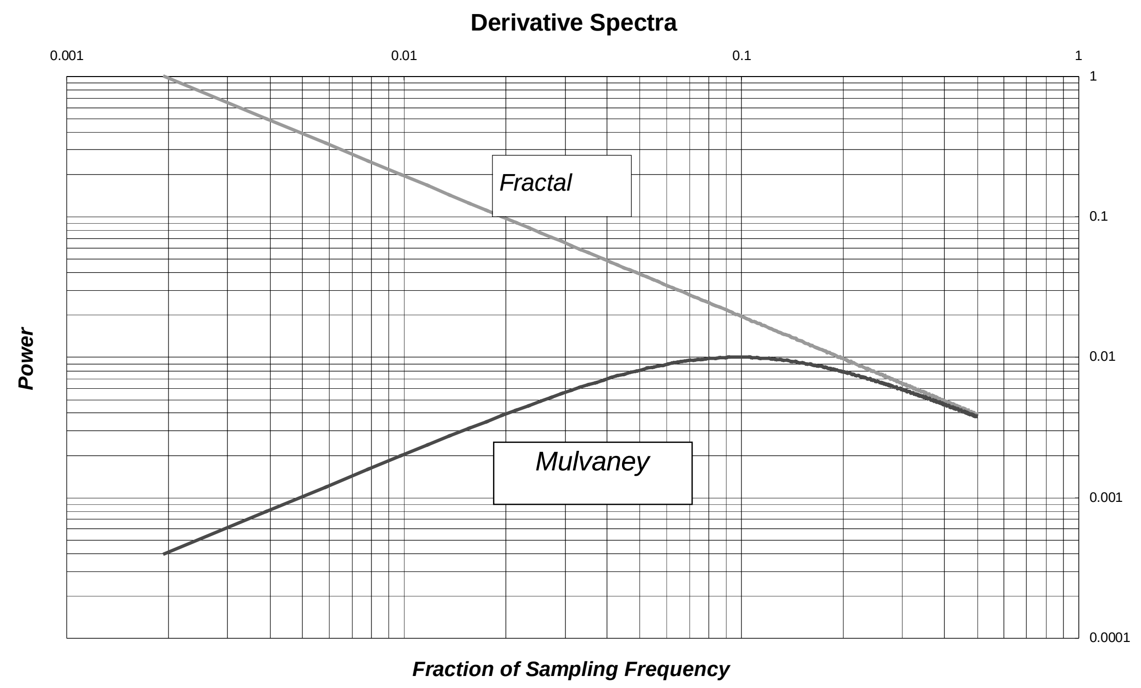


Figure 2.4.3 Comparison of Surface Derivative Spectra (not to scale)

At low frequencies, the power of the derivatives of the Mulvaney model varies quadratically with frequency and therefore tends to zero as frequency approaches zero. The amount of derivative power present at wavelengths greater than the sampling window size will depend on the frequency of the breakpoint. For the fractal model the $1/\omega$ relationship still holds at low frequencies. The derivatives are therefore non-stationary, with unstable statistical properties.

The $1/\omega$ characteristic is both very common in practice and difficult to accommodate in signal theory. However Keshner states that for a $1/\omega$ process, "*If the time over which the process is observed is short compared with the time elapsed since the process began then the exact ACF can be approximated*." He goes on to say, "*the PSD* (of a ω^{-1} process) *is stationary except for the steady state value which depends logarithmically on the time elapsed since the process was started*" [Keshner82]. For ease of analysis, we will assume that the surface spectra are bandlimited for frequencies less than the fundamental.

2.4.2 Modelling Directionality

In this section we will examine the most common model of directionality and consider some frequently made assumptions. We will also survey the literature to find what degree of directionality occurs in real surfaces. Finally, we will consider the question of whether fractal dimension is rotation invariant for anisotropic surfaces.

We begin by considering the Longuet-Higgins (LH) description [Longuet-Higgins57]. This was developed to describe the sea surface and forms the foundations of statistical geometry. Nayak [Nayak71] was the first to apply the LH approach to engineering surfaces and therefore provided the motivation for the use of random process techniques on surfaces. We have neglected the LH description, since it is first order i.e. does not incorporate any spatial information and is ultimately used to develop a number of parameters, such as summit density, which are of no direct relevance to this work. However, the first stage of the LH model does provide a useful basis for discussion of surface directionality.

If we assume that the direction of the surface grain is aligned with one of the axes, then we may use the matrix described by Bush et al. [Bush79].

$$\begin{matrix} m_{00} & 0 & 0 & -m_{20} & -m_{02} \\ 0 & m_{20} & 0 & 0 & 0 \\ 0 & 0 & m_{02} & 0 & 0 \\ -m_{20} & 0 & 0 & m_{40} & 0 \\ -m_{02} & 0 & 0 & 0 & m_{04} \end{matrix} \qquad (2.4.2a)$$

where

$$m_{fg} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u^f v^g S(u,v)du,dv \qquad (2.4.2b)$$

From equation (2.4.2b) we can see that m_{00} is independent of direction. We do however note that the one dimensional profiles in the Hasegawa study do show considerable variation in surface variance [Hasegawa93]. We presume this is due to limited data length and/or detrending removing significant low-shifted frequencies.

From (2.4.2a) we infer that surfaces are assumed to contain one dominant directionality. While this is probably the case for the majority of surfaces, we note that at least one of our real, manmade, test surfaces exhibit two well resolved directionalities.

In the previous section we noted that most surface roughness models have a power law roll-off over at least part of their spectrum. We now consider whether fractal dimension, and therefore the rate of roll-off, is rotation invariant for anisotropic surfaces. Hall et al. develop the theorem that, *"the fractal dimensions along the transects of a stationary stochastic surface are all identical, save that in one special direction."* He further states that *"the majority of processed surfaces we have examined appear to have identical fractal dimensions, even though they are markedly anisotropic in other characteristics"* [Hall95]. In his study of an anisotropic thin-film rigid disk, Majumdar and Tien found that the x spectrum consisted of two regions with β=-1.94 and -1.42, the y spectrum had β=-2.51, though the low frequency spectrum oscillated and did not follow any power-law. In the most comprehensive study of this area, Hasegawa et al. measured fractal dimension for several sample types at various orientations [Hasegawa93]. The roll-offs are shown in *Table 2.4.1*, where Hasegawa divided the spectrum into two regions; we have denoted these regions by (a) and (b). The results shown do not appear to give a definitive answer to the question of the rotation dependency of fractal dimension.

Angle	***0°***	***30°***	***60°***	***90°***	***120°***
Lapped Surface	*1.56*	*1.54*	*1.56*	*1.69*	*1.69*
Electric Discharge (a)	*1.57*	*1.87*	*1.58*	*1.56*	*1.57*
(b)	*1.66*	*1.80*	*1.86*	*1.65*	*1.74*
Grinding1 (a)	*1.54*	*1.55*	*1.55*	*1.52*	*1.59*
(b)	*1.77*	*1.79*	*1.73*	*1.78*	*1.73*
Grinding2 (a)	*1.67*	*1.72*	*1.68*	*1.73*	*1.70*
(b)	*1.68*	*1.65*	*1.64*	*1.72*	*1.70*

Table 2.4.1 Estimated Fractal dimensions in various directions, (from Hasegawa93)

The two dimensional PSD form of the exponential correlation function derived by Ogilvy (2.4.2c) offers an interesting possibility: since the only term which is direction dependent is the correlation length, directionality may be expressed in the cut-off frequency (λ_1 and λ_2) at which fractal behaviour, i.e. linear roll-off, begins to occur [Ogilvy91].

$$S(u,v) = \frac{\sigma_s^{\,2}}{\lambda_1\lambda_2\pi^2}\frac{1}{\left(1/\lambda^2{}_1 + u^2\right)}\frac{1}{\left(1/\lambda^2{}_2 + v^2\right)} \tag{2.4.2c}$$

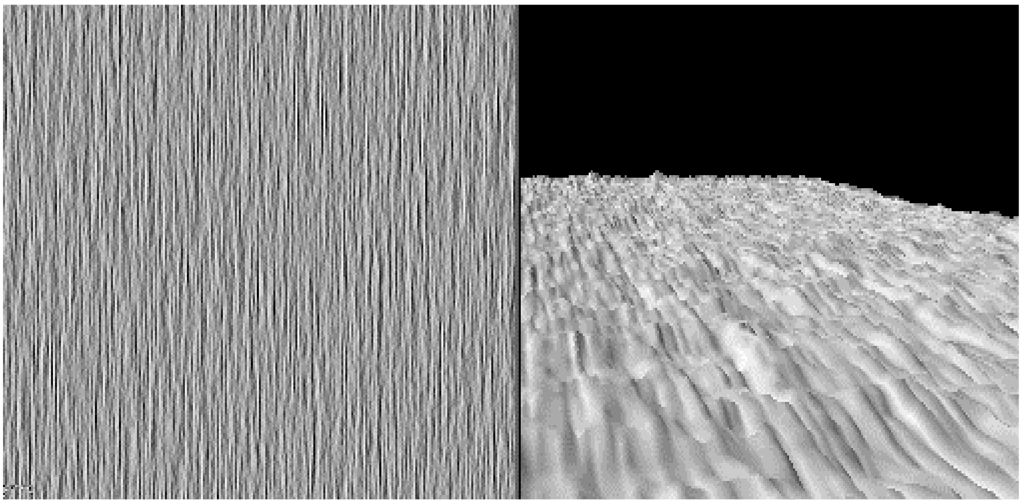
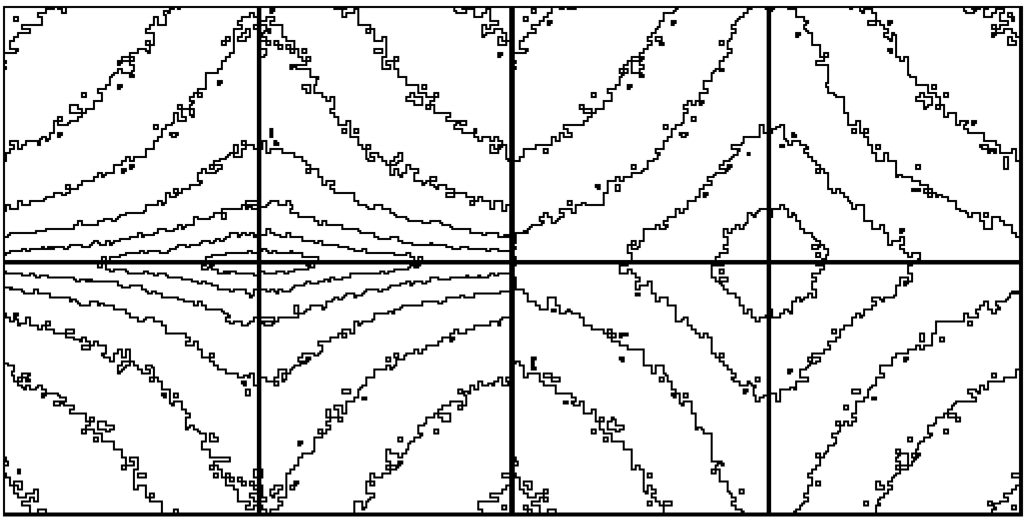


Figure 2.4.4 Overhead and isometric views of directional Ogilvy surface



(a.) *(b.)*

Figure 2.4.5 Contour Plots for Ogilvy Surface Spectra

In fact we find this model inadequate in one respect; it does not allow for a smooth transition between isotropic and highly directional textures. Setting the correlation lengths in the orthogonal directions to equal values results in a cross-like power spectrum, *Figure 2.4.5b*, rather than an isotropic spectrum. However, this model is well suited to extremely anisotropic textures and we shall use this model to produce a series of exemplar textures.

2.4.3 Assumptions of the Models

We make several assumptions in connection with the models we have adopted. We assume that a surface can be reconstructed from its gradient field i.e. the gradient field must be integrable; if a surface s is a potential field to the gradient field **S** then **S** is conservative, i.e. an integral of the gradient field around any path is equal to zero, Eq 2.4.3.

$$\mathrm{Curl}\ \boldsymbol{S} = 0 \qquad (2.4.3)$$

As a consequence of (2.4.3), the first and second partial derivatives of s are continuous and the mixed derivatives equal [Thomas p.1065].

A stricter assumption is that the surface is bandlimited. The low pass component of this assumption allows a discrete form of the surface model to be used, and in consequence the surface model can be differentiated an infinite number of times. A further consequence of this is that the *rms slope* of a surface will be finite. The high pass component of the assumption implies that the surface is stationary. It follows that the surface can be completely predicted from a realisation of the phase spectrum and the model without recourse to initial conditions. A further consequence is that the surface statistics are stable. Within the bandlimited region we assume the surface spectrum conforms to either the fractal, Mulvaney or directional Ogilvy models.

We assume the surface phase spectrum satisfies the maximum entropy condition, i.e. there is no discriminatory information held in the phase spectra and classification must be carried out purely on the power spectra. A consequence of this assumption and the central limit theorem is that the height distribution, and consequently the slope distribution will be Gaussian, allowing a more tractable analysis in subsequent chapters. We make the further assumption in this thesis that the slope distribution will have a standard deviation not exceeding 0.5.

2.4.4 Summary of the Models

We have identified three complementary models:

1. a fractal model,
2. the Mulvaney model, and
3. the Ogilvy model.

Surfaces (1) and (2) are isotropic and differ in the low frequencies. At high frequencies, the Mulvaney model exhibits behaviour identical to the fractal model, i.e. a power law roll-off, however, at low frequencies the Mulvaney spectrum is white. By varying with direction the frequency at which the transition between the two behaviours occur, Ogilvy produces our third model. The fractal and Ogilvy surfaces will be used extensively in this thesis as exemplars, though the Mulvaney model will be used to a lesser extent.

2.5 Summary of Surface Description and Modelling

The topography of a rough surface is modelled as the scalar field $s(x,y)$ which acts as a potential field to the (conservative) vector field $\boldsymbol{S}(x,y)$. The probability function associated with each of the scalar fields of $\boldsymbol{S}(x,y)$ as well as the height field itself are assumed to be Gaussian. The standard deviations of these distributions, p_{rms} and q_{rms} form useful surface parameters and will be used extensively in this thesis.

The surface power spectra $S(u,v)$ is the principle means of surface description—it is assumed to conform to one of three prototypical forms:

(1) a fractal form, this follows an inverse power law throughout its range,

(2) the Mulvaney form, exhibiting roll-off only above a cut-off frequency, and

(3) the Ogilvy form, which varies the frequency at which cut-off occurs with orientation to introduce surface directionality.

The phase spectrum of the surface is assumed to contribute no discriminatory information, i.e. surfaces with the same surface model and model parameters, but different phase spectra, will be considered to be of the same type. In fact, this is only a reasonable assumption for a limited subset of textures. Intuitively, this subset is characterised by the absence of step changes and a lack of structure in the data set.

2.5.1 Relevance of this Chapter to subsequent chapters

This chapter describes the first stage in the classification process. In consequence, many of the assumptions and decisions made in this chapter have consequences throughout the rest of the thesis. In this sub-section we show the relevance of these decisions to later chapters.

Throughout this thesis we model the topography of a rough surface with a discrete function $s(x,y)$ with an associated derivative vector field $\mathbf{S}(x,y)$. The distribution of heights and gradients is assumed to be Gaussian. In Chapters 3, 4 and 6 we model imaging and classification as a linear process, consequently the assumption of normality is maintained throughout these sections. The slope distributions are parameterised by the rms slope parameters (m_{rms},p_{rms} and q_{rms}). Unlike the rms roughness, the slope parameters are directional, and consequently provide a simple indication as to the directionality of the surface. In Chapter 3 we will see that the adequacy of the linearity assumption for the rendering process is a function of slope. Slope parameters therefore represent an important surface characteristic and are used in Chapters 3, 7 and 8. As a consequence of starting from a height, rather than a derivative, model we assume the derivative fields are conservative and integrable. This is relevant in Chapters 7 and 8 where we consider reconstruction of the surface from derivative estimates.

The scalar field $s(x,y)$ is described using the power spectrum, $S(u,v)$. Use of the PSD allows integration of the surface model with the illumination, noise and filtering models of Chapters 3,4 and 6 respectively. The power spectrum is assumed to be bandlimited. The low pass component of this assumption allows a discrete form of the surface model to be used, consequently the *rms slope* of a surface will be finite. The high pass component implies that the surface is stationary and the surface statistics are stable.

We assume the surface phase is random, i.e. there is no discriminatory information held in the phase spectra, and classification must be carried out solely on the power spectra. This will affect the classifier design in Chapter 5. This assumption in conjunction with the central limit theorem supports the assumption that the height distribution, and consequently the slope distribution will be Gaussian.

This chapter has also justified the adoption of three surface models defined in terms of the power spectrum:

(1.) an isotropic and fractal model—this obeys a power law roll-off within the observed bandwidth.

(2.) Mulvaney's model is also isotropic and can be seen as a generalisation of the fractal case, with a flat response below a cut-off frequency.

(3.) Ogilvy's model, varies the frequency at which the transition between white noise and fractal roll-off occurs with angle to produce a directional surface.

The spectral characteristics of these models has influenced the design of the classifier developed in Chapter 5. The fractal and Ogilvy models are used throughout this thesis as test cases for isotropic and anistropic surfaces respectively. The requirement for a series of isotropic surfaces in Chapter 6, is partially fulfilled using the Mulvaney model.

	3 Image Formation	4. Imaging	5. A classifier	6. The Effect of tilt on the classifier	7 Compensation Schemes	8 A Model - based scheme
s(x,y)						
normality						
m_{rms}						
curl**S**=0						
PSD						
Phase						
Fractal						
Ogilvy						
Mulvaney						

Figure 2.5.1 Assumptions developed in this chapter and their use in later chapters. (Shading indicates use of assumption in chapter)

2.6 Conclusions

The aim of this chapter was to define a system of surface modelling, whereby different realisations of the same model, with identical parameters, could be classified as belonging to the same class of surface on the basis of their visual appearance. After considering several fields it was decided to consult the related fields of tribology and scattering in order to find an established methodology of surface description and modelling. Several levels of description were discussed, before a second order technique, the PSD, was adopted for the description of rough surfaces.

The description is in terms of the power spectrum. However, surfaces may have the same PSD but radically different visual appearance due to the influence of their phase spectra. We circumvent this problem by adopting a maximum entropy condition for the textures' phase spectra: a texture's phase spectrum may have any realisation so long as it doesn't carry any information. A simple test was developed to assess whether this condition holds for a given texture. Only textures which fulfilled the criterion will be considered in this report.

The models considered are defined in terms of the description adopted: the power spectrum. Two characteristics of the spectrum were considered: the roll-off of power with increasing frequency and directionality. All the models display a power law roll-off of power with frequency for at least part of their frequency range. While the fractal model shows this characteristic throughout the frequency range, Mulvaney and Ogilvy models have a limiting process at low frequencies where the spectrum is approximately white. The Ogilvy model, models directionality by varying the frequency at which the transition from white to fractal behaviour occurs with direction.

Chapter 3

Image Formation

3.1 Introduction

The goal of this thesis is the development of a classifier which discriminates between rough surfaces on the basis of their visual appearance. This chapter links the surface description developed in the last chapter to the image incident on the camera lens. It will be shown that the characteristics of this link depend on, among other things, the illuminant tilt. This chapter develops a model of the source/surface dependency to which we aim to make the classifier robust.

The aim of this chapter is to model the transition from physical surface to incident image. This aim is attained by satisfying two objectives: firstly, the formation of a description of the mapping from surface derivatives to image intensity; and secondly, the development and evaluation of a frequency based transfer function from surface to image. Whereas the first is a purely local process, the second takes into account the spectral characteristics of the surface and forms the link between the spectral surface models described in the previous chapter, and the image, data set and frequency-based feature set considered in subsequent chapters.

The first aspect of the model is considered in section 3.2; a brief description of the underlying physical mechanisms of reflection and some of the more prominent developments in reflectance modelling, is given. The thesis then focuses on the much less researched area of diffuse reflection: three modern models are considered, two of these, as well as the classical Lambertian model, are compared with an empirical mapping. We will find the analytical function which most closely resembles the measured function under the conditions associated with rough surfaces.

The second aspect of the model is a frequency-based transfer function between surface and image—again taking an analytical model as the starting point. In this section we ask two questions:

- is the analytical model optimal in the least squares sense?
- how does the nature of the surface affect the accuracy of estimation?

This section therefore considers the optimality of the analytical model as well as the accuracy of a more empirical linear model.

Finally, the predictions of the analytical model are compared with the actual characteristics of an isotropic and an anisotropic texture.

3.2 Modelling the Reflectance Function

In this section we shall consider the purely local interaction between a surface facet and lighting geometry to give the observed intensity. This problem has been investigated by researchers in physics [Torrance67], computer graphics [Cook82] and more recently machine vision [Nayar91]. The phenomenon of reflection can be produced by one or more different physical processes. These processes give rise to various generic behaviours, as well as intermediate forms. In this section we will briefly discuss the different characteristic behaviours, as well as the physical processes behind them, before focusing on one generic type. Several models of this type will be evaluated against the reflection characteristic of the surfaces used in the experiments described in this thesis. The findings of this section will then be used in the next to produce a non-local model of illumination, which will link the surface models of the previous chapters to the feature models of Chapter 5.

3.2.1 Terms used in this chapter

We begin by stating the standard radiometric definitions. These are given in more complete terms in most textbooks [Watt p.91] and many papers [Nayar91], we reiterate them here for completeness. *Flux* is the rate of emission or reflection of light energy, and is measured in Watts. *Irradiance* is the incident flux per unit surface area (Wm^{-2}), *radiant intensity* is the flux radiated per solid angle in a particular direction ($Wm^{-2}\ St^{-1}$) . *Radiance* is the radiant intensity, per unit, of *foreshortened* area ($Wm^{-2}\ St^{-1}$), where the foreshortened area is the surface area times the cosine of the angle between the radiated light and the surface normal. For a Lambertian surface the radiance is independent of the viewing angle. Horn has shown that image irradiance, i.e. the pixel intensity value is proportional to the scene radiance [Horn86].

The ratio of surface radiance to incident irradiance is defined as the bi-directional reflection distribution function (BRDF) (equation 3.2.1a), which Horn attributes to Nicodemus. For a facet with isotropic microstructure, the BRDF is invariant to rotation of the surface about the surface normal and may be expressed in the form shown in (3.2.1b) [Horn p.210].

$$f(\theta_r,\theta_i,\eta_r,\eta_i) \tag{3.2.1a}$$

$$f(\theta_r,\theta_i,\eta_r\text{-}\eta_i) \tag{3.2.1b}$$

where θ is the tilt angle and η is the slant angle. The subscript denotes whether the angle refers to the incident (i) or reflected (r) ray.

In this thesis, we will borrow the concept of the reflectance map, developed by Horn [Horn77] and widely used in *shape from shading* research. The reflectance map assumes a specific lighting orientation and describes the radiance/irradiance ratio in terms of the surface normal. It is therefore a two dimensional function with two independent variables, the directional derivatives. The measured intensity is the dependent variable.

While the reflectance map is stated in terms suited to our purposes, the fact that it has two independent variables makes it difficult to show graphically and we seek to describe the effect with a single variable. The Lambertian model has the assumption that reflectance is a function purely of the angle of incidence and is independent of viewing direction. This assumption has been challenged in both [Oren94] and [Wolff94], though Healey's model does retain the assumption [Healey89]. We will show the reflectance function as a function of incident angle averaged over the relevant range of viewing angle. That is, the reflectance function is calculated in terms of surface gradient and expressed in terms of incident angle.

3.2.2 Underlying Physical Processes

Light incident on a flat surface will be reflected and transmitted. The degree to which either occurs depends on the electrical characteristics of the material. In a conductor electrons are loosely bound to atoms and will oscillate at the same frequency as the incident field. The field itself will be quickly attenuated as it penetrates the material, however, the surface currents induced by the field will in turn reradiate an electromagnetic field in the mirror direction. If on the other hand, the electrons are tightly bound to atoms, as they are in a dielectric, they will interact with the field to a much lesser degree and the light will, in the main, be transmitted. We would therefore expect reflection to be the dominant effect in conductors and transmission in dielectrics.

In fact, dielectrics are rarely transparent; light *is* reflected from the material, though by a different mechanism and in a much less directional manner than with the surface reflection associated with conductors. The actual mechanism is unknown, though the most popular hypothesis is that of 'bulk (or body) scattering'. This model assumes the material to be inhomogeneous; light incident on the surface is refracted at the boundary and is transmitted through the material until it encounters an inhomogenity boundary from which it is scattered. Light is repeatedly reflected and refracted from these internal boundaries and some is scattered in the direction of the air boundary and will emerge in a randomised direction. In the limiting 'Lambertian' case, the surface will appear equally bright from all viewer orientations.

The previous discussion attributed optical effects solely to the material characteristics and yielded the two extreme cases: specular reflection for conductors and Lambertian reflection for dielectrics. In fact, both conductors and dielectrics exhibit transitional behaviour. We consider surface reflection from dielectrics first. The ratio of surface reflected light to incident light is given for both dielectrics and conductors in terms of the Fresnel equations. These are defined in terms of the complex refractive index, $n+iK$, and the angle of incident light with respect to the surface normal. If we plot the Fresnel expression for unpolarised light incident on a smooth dielectric, we obtain the curve shown in *Figure 3.2.1*. The surface reflection rises sharply as the incident light approaches the grazing angle and, assuming a conservative system, the total body reflection will decrease accordingly.

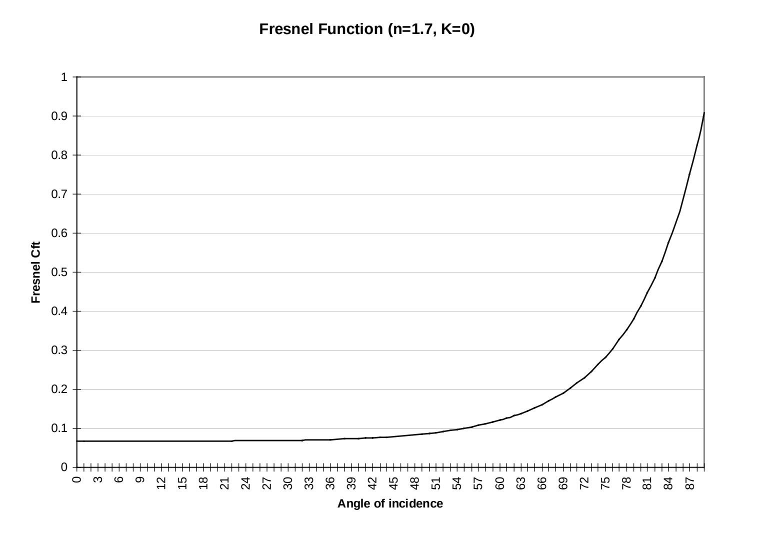


Figure 3.2.1 Fresnel Function

Just as dielectric reflection shows some characteristics associated with conductors, conductor reflection may also fall into the transitional region. Although metals attenuate transmitted light far too quickly for body reflection to be significant, surface roughness will cause the reflected light to be more diffuse. A rough surface no longer has a single mirror direction, but instead a statistical distribution of directions, that manifests itself as a variation of intensity. This effect will be most pronounced in the conductor case, though it will also have an effect for dielectric reflection.[a1]

3.2.3 Specular Models

The most general methodology for dealing with surface reflection, both in terms of the degree of roughness and surface conductivity, is scattering theory. A less general and rigorous technique, which models the surface as a collection of perfectly mirror-like microfacets whose slopes conform to a specified distribution was developed by Torrance and Sparrow [Torrance66]. The Torrance-Sparrow (TS) model is only relevant where the level of roughness is much greater than the wavelength of light. Although this restriction excludes surfaces with optical finish, the TS model forms the basis for the majority of specular models recently developed in both computer graphics and machine vision. We argue that the substitution of the scattering model by a TS type model is particularly valid in texture analysis given the large body of experimental evidence that the roughness spectrum is continuous to almost atomic scales.

The TS model consists of three sub-components: a slope distribution (analytically modelled as a distribution of V-grooves), a geometric factor used to model facet shading and a Fresnel coefficient. Blinn is the first to have applied the model to computer graphics, appending a Lambertian component [Blinn77]. Cook further developed the model by introducing a new slope distribution, defined in terms of real measurable parameters and introducing a spectral dependency [Cook82].

3.2.4 Diffuse Models

While specular reflection has been thoroughly investigated in the field of computer graphics, researchers in this area seem to be content to assume diffuse reflection to be adequately modelled by the Lambertian model. However, diffuse reflection is also widely encountered in machine vision and recently the need for accurate reflectance maps has motivated the investigation of diffuse reflection. We consider three analytical models developed in recent years, the model developed in [Oren94] is the least physically-based and, like the TS model it models the surface as a group of microfacets, although in Oren's model the facets are perfectly Lambertian, rather than specular, reflectors. Healey [Healey89] adapts existing theory of diffuse reflection and integrates it closely with the TS model of specular reflection. Wolff [Wolff94a] develops a more complex model, which includes a viewer orientation dependency. This is valid for smooth surfaces only, though in a later paper the model is extended to include rough surfaces. We will briefly discuss these models with reference to each other and the Lambertian model in the context of textured surfaces.

Lambertian Model

In intuitive terms, the Lambertian model states that the perceived intensity of a surface facet is dependent only on the relative geometry of that facet and an illuminant. Mathematically, this is represented as the dot product of the surface derivative vector with the illuminant vector.

While the radiance of the facet is constant with respect to the viewer's position, we note that the radiant intensity, that is the flux per unit solid angle, varies with the foreshortened area of the facet.

The radiance is plotted against the angle of incidence in *Figure 3.2.2*; as with all the other reflectance functions in this thesis the function is scaled to give unity radiance for a flat surface, i.e. at an angle of incidence of 60° where the slant angle is held constant at 60°. This is compatible with our experimental practice of using the image of a flat surface with a uniform albedo equal to that of the textures to remove variations in the irradiance incident on the surface.

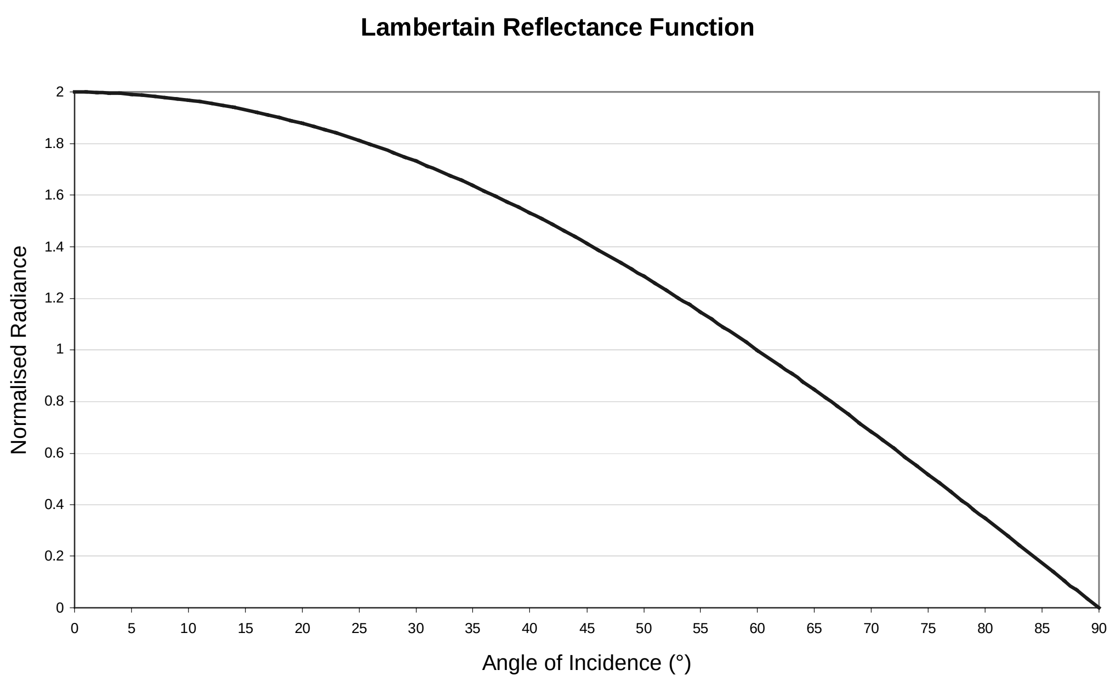


Figure 3.2.2 The Lambertian Reflectance Function

Oren's Reflectance Function

Oren states that the Lambertian model is inadequate for many applications and that this failure is most apparent when viewer direction (relative to the surface normal) is varied. It is Oren's thesis that the discrepancies are primarily due to subpixel surface roughness and he does not pursue the physics of diffuse reflection. Like the TS model, Oren models the surface as a set of long V-grooves consisting of two facets, although in Oren's model these are perfect Lambertian, rather than specular reflectors. Another difference is in the models' facet distributions: the TS facets are of equal area, and the distribution represents the number of facets with a specific normal that lie within a unit area. Oren's distribution, in contrast, represents the fraction of the area occupied by facets of a given normal; this, he argues, makes the model less sensitive to variations in the actual roughness distribution. Oren uses these facets to model shadowing, masking as well as interreflection, though this last term is subsequently neglected. Finally Oren develops a functional approximation for the model.

We plot Oren's reflectance function of radiance against angle of incidence for three degrees of sub-facet roughness *Figure 3.2.3*. This function, as with all the functions reported in this section, has been normalised at $\theta_i = 60°$, i.e. it assumes a value of unity for the radiance of light reflected from a flat surface

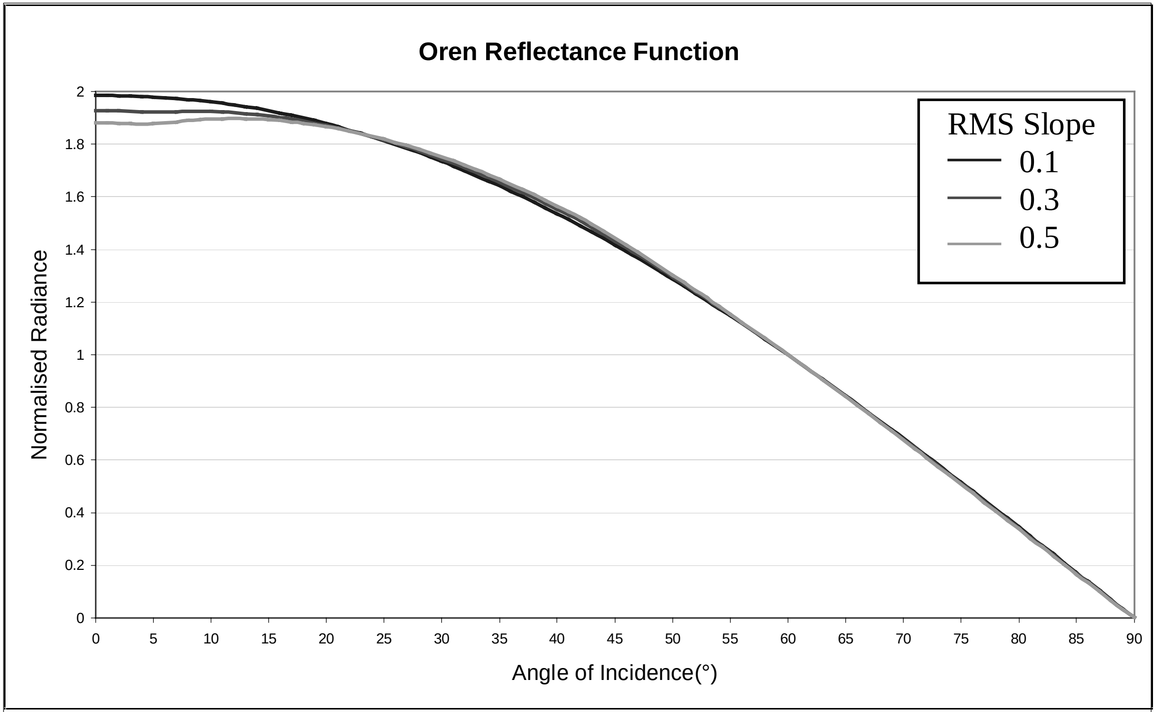


Figure 3.2.3 Oren's reflectance function for various degrees of surface roughness

Oren then experimentally verifies his model on a set of 2X2 inch samples imaged with a CCD camera at 6ft, with the average pixel value as the sample irradiance. He finds the model to be in strong agreement with his experiments for samples of wallplaster, painted sandpaper and white sand (though a small specular component was detected in this case). Other notable results are that as roughness increased, the reflectance map varied from the Lambertian to a more linear form. Another interesting result is that backscatter is greater than foreground irradiance—not only illustrating an inadequacy in the Lambertian model but also contrasting with the case of specular reflection from a rough surface. For samples of cloth, foam and woodshaving, Oren appends a TS model for specular reflection, with varying degrees of success. Several parameters of the model are estimated by fitting the data to the model using non-linear optimisation. He assumes the Fresnel coefficient to be unity due to difficulties in measuring *n*.

Healey's Reflection Model

Healey's dielectric model must be set in the context of the rest of his paper [Healey89], which develops a dichromatic reflection model with the aim of discriminating between materials based on their reflectance properties. He therefore considers metals as well as dielectrics and wavelength forms an important part of his model. This thesis is concerned only with dielectric surfaces and only considers monochrome images, wavelength dependencies are therefore neglected in the interest of clarity.

Healey models surface reflection using the TS model, with the Cook and Torrance slope distribution [Cook82]. He notes that for perfect dielectrics the refractive index is purely real, simplifying the Fresnel equations. Internally, scattering is based on a modified form of the Kubelka-Munk (KM) theory. The original KM theory treats the dielectric as consisting of many elementary layers, in which colorant particles are embedded. These layers are characterised by two parameters: α_h the fraction of light absorbed per unit path length and σ_h the fraction of light scattered per unit path length. The reflectance function predicted by the KM model is:

$$R_\infty = \frac{2 - w_h - 2[1 - w_h]^{0.5}}{w_h}$$

where

$$w_h = \frac{\sigma_h}{\sigma_h + \alpha_h}$$

In the case of Lambertian scattering, *w* is equal to unity.

The KM theory was extended by Reichmann [Healey89] to allow non-diffuse light, and vehicles with refractive indices which differ from those of air to be modelled:

$$R_b = (1 - R_s)\frac{C(\theta)(1 - r_i)[R_\infty - D(\theta)]}{2[1 - r_i R_\infty]\cos\theta}$$

where r_i is the internal diffuse surface reflectance, described by:

$$r_i = \frac{1 - 1.439 - 0.7099n + 0.3319n^2}{n^2}$$

where *n* is the real part of the refractive index,

and

$$C(\theta) = \frac{w_h \cos\theta(2\cos\theta + 1)}{1 - 4[1 - w_h]\cos^2\theta}$$

$$D(\theta) = \frac{2\cos\theta - 1}{2\cos\theta + 1}$$

It is worth reiterating that, in contrast to the models of Oren and Wolff, Healey's model does not incorporate any viewer dependency.

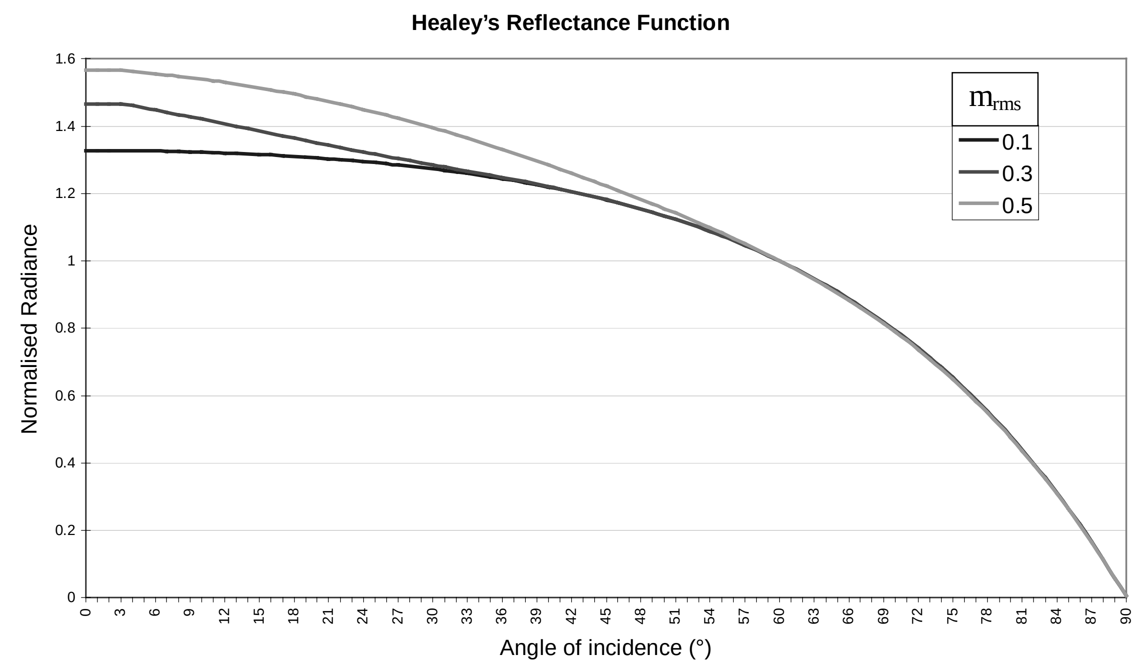


Figure 3.2.4 Healey's reflectance function for various surface roughnesses

Wolff's Reflectance Model

Wolff [Wolff94] also develops a model for sub-surface scattering, and in [Wolff94a] he extends this model to include rough surfaces and a surface reflection term by linearly adding the TS model. Wolff bases his model on the theory of Chandrasekhar; like the KM model this assumes a vehicle with refractive index equal to that of air. Wolff couples this model to the Fresnel boundary conditions, noting changes in the direction and subtended solid angle of the incident light.

Unlike the other models considered in this section, the Wolff model was not implemented due to its complexity. We are therefore unable to evaluate its effectiveness.

Comparison of Models

The relative characteristics of the illumination models, can be examined by rendering three spheres with Lambert's, Oren's and Healey's reflectance functions (see *Figure 3.2.5*). The most obvious difference between the renderings is in the regions of maximum reflectance. Lambert's function gives a distinct peak in intensity while Healey's model does not produce any discernible bright spot. Oren's model is intermediate between these forms.

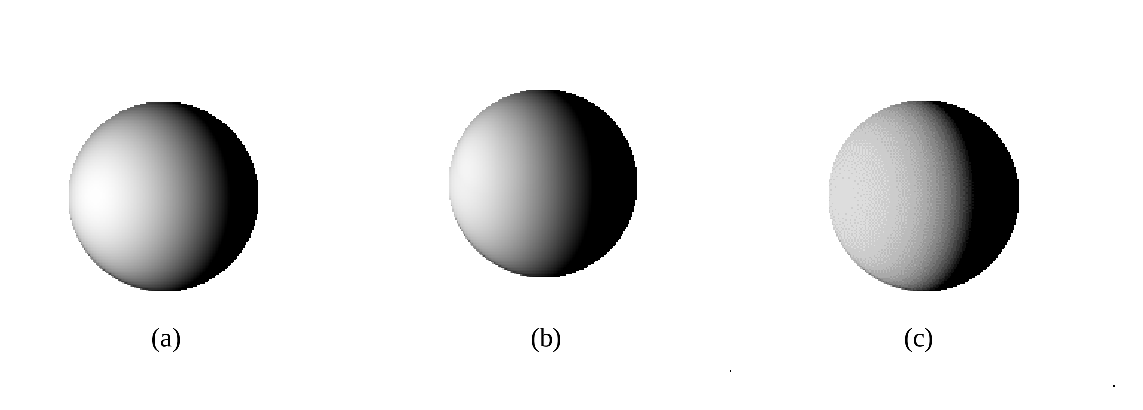


Figure 3.2.5 Spheres rendered with Lambertian (a), Oren(b) and Healey Models (c)

The highlight effect is also visible in the reflectance maps of the models (*Figure 3.2.6*). The lessening of the effect suggests that the relationship between surface derivative and radiance is more linear for Oren's model than Lambert's, as was noted by Oren himself.

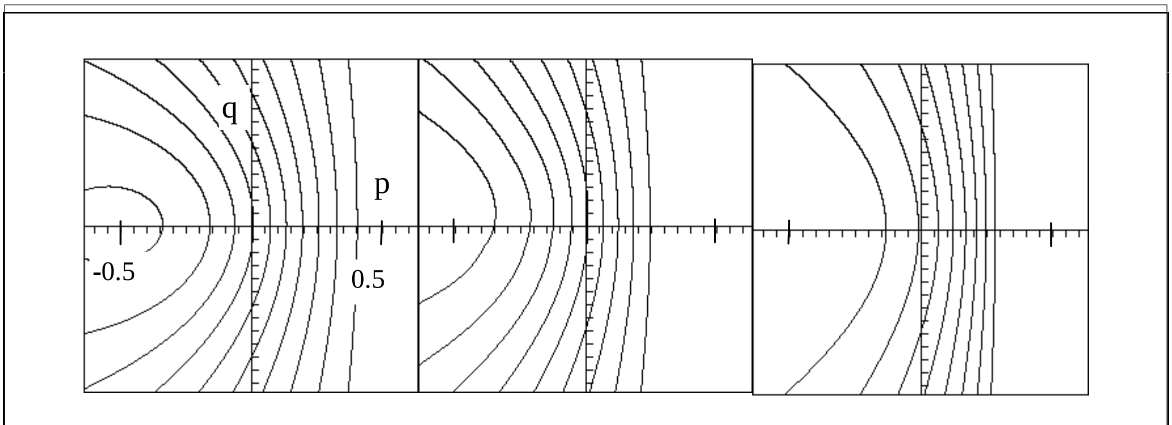


Figure 3.2.6 Reflectance Maps for (a)Lambert's, (b)Orens's and (c)Healey's Models

Comparison of the models with the empirically measured function is of more relevance to this thesis. All the surfaces considered in this thesis have been sprayed matt white. The reflectance map associated with this surface characteristic was derived by measuring the intensity of an object of known geometry, i.e. a sphere, and recording the mapping between surface derivatives and radiance. It is shown together with the analytical reflectance models in *Figure 3.2.7*.

It can be seen that Healey's model performs relatively poorly for the surface type used in this report. The Lambertian model is a much better approximation, however, it is Oren's model that is closest to the measured reflectance function.

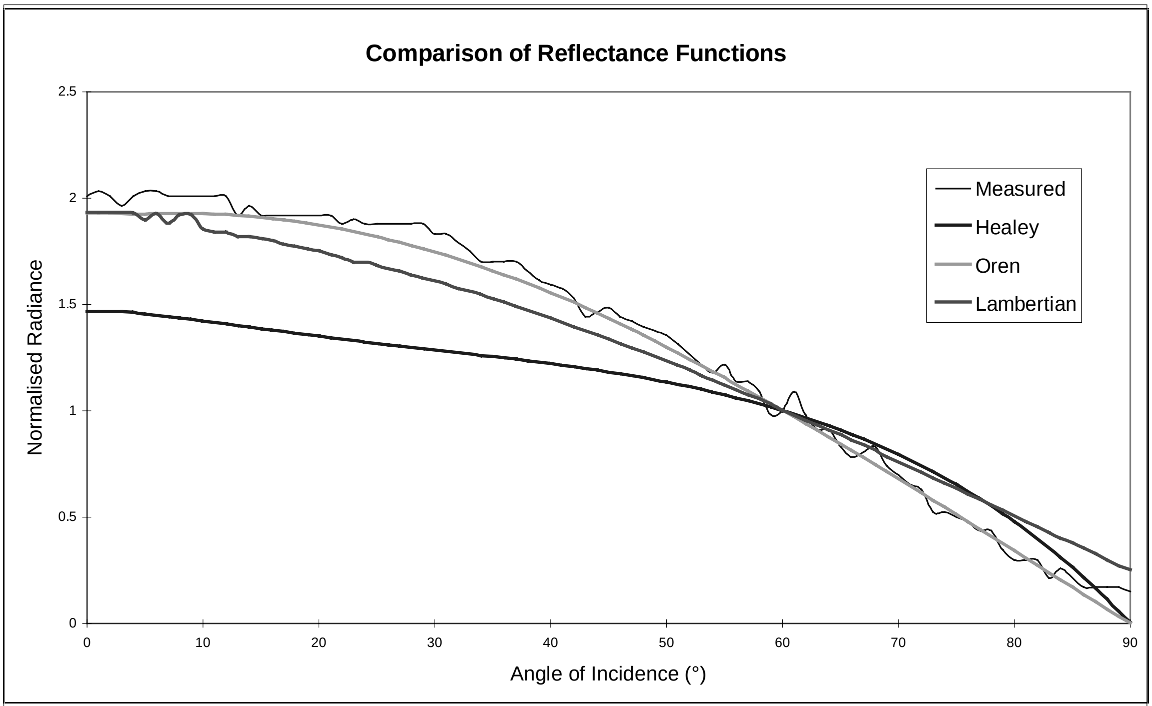


Figure 3.2.7 Comparison of Reflectance Function

While the comparison of reflectance functions is useful from an analytical point of view, it is also important to assess how well the reflectance functions perform on rough surfaces of the type considered in this thesis. The accuracy of the reflectance function will vary with incident angle. Light incident on a rough surface will have a distribution of incident angles, which will be partially dependent on the slope distribution. In order to assess the accuracy of the models for textured surfaces we must take the angular distribution into account. We do so by applying the models to synthetic surfaces with different slope distributions. This approach will therefore give a better indication of the actual accuracy of the function for our application. We evaluate the accuracy by rendering a series of synthetic height maps using an experimentally measured reflectance map. These results are then compared with analytical estimates derived using the three models. A signal to residue ratio (S/R) is used to perform this comparison. It is defined as

$$S/R = 10 Log_{10}\left(\frac{Var[i(x,y)]}{Var[e(x,y)]}\right)$$

where Var[*i(x,y)*] is the variance of the image predicted using the empirical function and Var[$e(x,y)$] is the variance of the difference between this, and the image predicted by the relevant theoretical model.

The first group of surfaces we consider is based on an isotropic fractal surface which is scaled to give varying degrees of roughness. The results show the Lambertian function to be the most accurate model for the reflectance function of our experimental surfaces, followed closely by Oren's model, with Healey's model performing comparatively poorly on our surfaces (see *Table 3.2.1)*

	S/R (dB)		
RMS Slope	**Lambertian**	**Oren**	**Healey**
0.125	18.96	17.05	12.12
0.250	18.77	17.95	10.27
0.500	16.50	15.81	12.10

Table 3.2.1 The accuracy of image prediction from isotropic surfaces of various roughnesses.

We now consider the highly directional Ogilvy surface defined in the previous chapter. In the first case, the grain of the surface runs at right angles to the direction of illumination, i.e. the illuminant is aligned with the direction of maximum slope variance *Table 3.2.2*. We note an increase in accuracy over the isotropic case, however the fall in accuracy associated with increasing roughness is still apparent.

RMS Slope		**S/R (dB)**		
$\mathbf{p_{rms}}$	$\mathbf{q_{rms}}$	**Lambertian**	**Oren**	**Healey**
0.125	0.034	20.43	18.76	12.42
0.250	0.068	20.62	21.30	10.01
0.500	0.136	16.26	16.26	12.39

Table 3.2.2 The accuracy of image prediction for directional surfaces, illuminant perpendicular to surface grain.

If we illuminate the same series of directional surfaces in the same direction as the surface grain, we find a significant fall in the accuracy of the Lambertian, and even more so in the Oren functions, whereas Healey's model is relatively unaffected (*Table 3.2.3*) . The poor performance of the Lambertian and Oren models is not solely an effect of the low slope in the direction of illuminant. We infer this from the fact that the maximum *rms*. slope in the direction of the illuminant for this category of surface (p_{rms} =0.136) exceeds the minimum slope parallel to the illuminant for two previous surface types (p_{rms} =0.125), yet the accuracy is still substantially below that of the first two cases.

RMS Slope		**S/R (dB)**		
$\mathbf{p_{rms}}$	$\mathbf{q_{rms}}$	**Lambertian**	**Oren**	**Healey**
0.034	0.125	11.13	2.18	11.12
0.068	0.250	13.66	9.87	12.48
0.136	0.500	14.22	13.47	10.82

Table 3.2.3 Accuracy of prediction for directional surface illuminated in the direction of surface grain.

We conclude that of the models tested here, the reflectance function of the surfaces used in this thesis is best approximated using the Lambertian model. The model performs well for isotropic surfaces and directional surfaces of low slope, which have their axis of maximum slope variance aligned with the illuminant. The Lambertian model performs most poorly on directional surfaces of low slope when the illuminant is parallel to the grain of the surface. For the directional surfaces considered in this thesis, the accuracy of the Lambertian function never falls below 11dB while for isotropic surfaces it remains above 16.5 dB.

3.2.5 Summary

In this section we have reviewed three analytical models of diffuse reflection, and evaluated these against the empirically measured reflectance map for our test surfaces. We found that, as a function of incident angle, Oren's was the closest to our measured function. If the test is weighted by the frequency with which facet orientations are predicted by our surface models, however, it was found that the Lambertian function modelled the empirical mapping more accurately. We found that the Lambertian model performed best on isotropic surfaces with low slope angles and on directional surfaces of low slope when the illuminant is perpendicular to the surface grain. The model performed most poorly on directional surfaces of low slope when the illuminant is parallel to the surface grain. However, even in the worst case, the S/R ratio of the Lambertian model never fell below 11 dB for the range of surfaces considered in this report.

3.3 An Optimal Linear Model of Image Formation

In the previous section we treated the reflectance function as a purely local phenomenon. However, any work concerned with texture is inherently concerned with the interaction of a point with its neighbours. In the previous chapter we modelled this interaction in the frequency domain. The radiance of a facet is a function of its surface derivatives—the spectra of which are easily related to the surface height spectrum. It is therefore attractive to model the transition from surface to image in the frequency domain.

For a tractable frequency domain analysis we require a linear model of the rendering process. If we assume a single reflectance function is valid over the entire surface, and that the illumination conditions are constrained, we can state that the only factor that will affect the degree to which the transfer function can be described as being linear will be the topographical characteristics of the surface. These will affect:

- the area and location of the region of the reflectance map which the model must approximate, and
- the degree of cast and self-shadowing[1].

From this it follows that a linear approximation is more valid for some surfaces than others.

The first objective of this section is to identify the degree of accuracy with which a linear model can represent the rendering of various surface types. This will be carried out in section 3.3.1 by using a least squares filter to obtain the optimal linear reflectance function for a given surface and illumination conditions. The surfaces will be scaled versions of two prototypes: the isotropic fractal, and the directional Ogilvy surface. The synthetic surfaces will be rendered using the empirical reflectance map and the least squares derived linear model. The criterion of accuracy will be stated in terms of the signal to residue ratio. This section will conclude with the statement of a number of constraints on the characteristics of surfaces which can be rendered, using an optimal linear model, to a specified S/R ratio.

[1] Shadowing may take two forms. A *cast* shadow occurs where one part of the surface prevents another part from being illuminated, by blocking the direct path between the light source and the shadowed area. *Self-shadowing* occurs when a facet is oriented such that it does not present an area on which light is incident, i.e. $\underline{\mathbf{I}}.\underline{\mathbf{N}}<0$. While the latter is a function of facet orientation and is a purely local phenomenon, the former is a function of surface height and is non-local in effect.

3.3.1 A Linear Filter

In this section we treat the problem as one of developing an optimum filter for a given illumination condition and also for a given surface. The optimum filter represents the transfer function from the surface derivatives to the image. Applying the filter to the surface derivative fields will give an approximation to the image. The filter will have the general form shown below:

$$i(x,y) = \mathbf{S}(x,y)^T \mathbf{V}$$

where:

$$\mathbf{S} = \begin{bmatrix} p(x,y) \\ q(x,y) \\ 1 \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

i(x,y) is the image intensity at point *(x,y)*,

p(x,y) and *q(x,y)* are the surface derivatives at *(x,y)*,

and *a,b* and *c* are the filter weights.

Let

$$\mathbf{R} = E[\mathbf{S}\mathbf{S}^T] \qquad \mathbf{U} = E[i_d \mathbf{S}]$$

where i_d is the desired signal.

To find the optimum parameter values we must minimise the quantity *e*, where

$$e = E[i_d^2] + \mathbf{V}^T\mathbf{R}\mathbf{V} - 2\mathbf{U}^T\mathbf{V}$$

The transfer function is approximated by a least squares linear filter in *p* and *q* calculated for the surface and image to which it will be applied.

3.3.2 The Signal to Residue Ratio

The analytical work of Chapter 5 is based on Kube and Pentland's frequency domain model [Kube88]. In order to link the surface and image spectra, Kube was obliged to use a linear approximation to the reflectance map. In this section we aim to assess the conditions under which the assumption of linearity is reasonable by comparing renderings using the optimal linear model *for a given surface and illuminant* with those of the empirically derived reflectance map. We also seek to discover the nature of the errors in spectral terms.

We will consider the transition from surface *s(x,y)* to images *i(x,y)* and *i'(x,y)* as being performed by process *o(p,q)*(derived experimentally using the matte-white sphere) and its linearised version *o'(p,q)* respectively. The difference between the desired signal *i(x,y)* and the linear estimate *i'(x,y)* is denoted by *e(x,y)*, the residual signal (*Figure 3.3.1*).

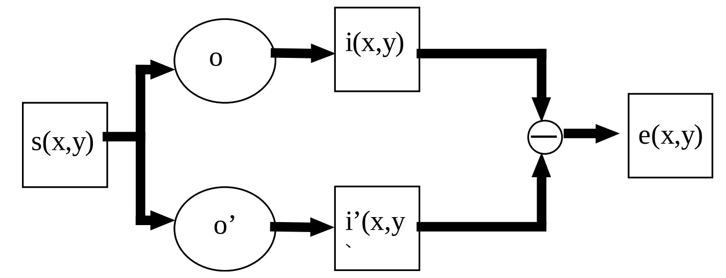


Figure 3.3.1 Calculation of residue image.

The signal to residue ratio is defined as:

$$S / R = 10 Log_{10}\left(\frac{Var[i(x,y)]}{Var[e(x,y)]}\right)$$

where Var represents the variance of a signal.

The surface vector field, ***S(x,y)***, is estimated from the potential field, *s(x,y)*, using two simple two-point estimators. The two point estimator is the simplest form of differentiator, and we note that it underestimates the magnitude of high frequencies, nevertheless we use the two-point estimator due to its simplicity, see [Bentum96] for a readable treatment of the subject.

3.3.3 The Effect of Surface Characteristics

Surface Roughness

We now consider the effect of surface roughness, parameterised by *rms slope*, on the accuracy of the linear model. An isotropic fractal surface was scaled and rendered using the empirical reflectance map. The accuracy of prediction is plotted against *rms slope* in *Figure 3.3.2*. As we would expect, the accuracy of the approximation falls with increasing slope—falling below the 10 dB mark for surfaces with an rms slope greater than 0.25.

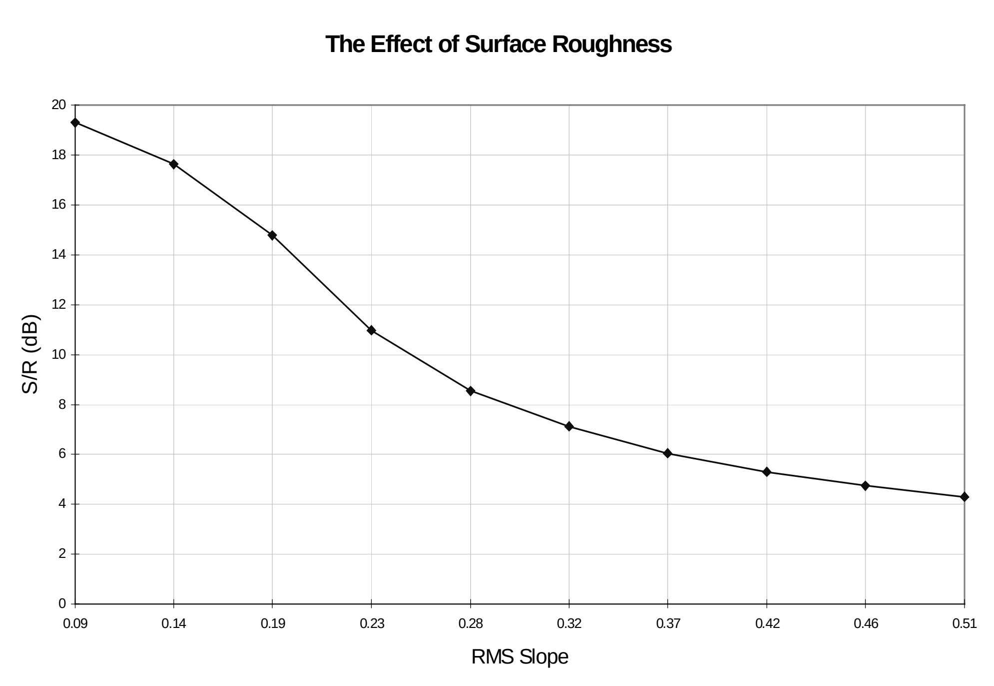


Figure 3.3.2 The effect of surface slope on the accuracy of an optimal filter.

Surface Directionality

If we consider Ogilvy's directional surface and alter the tilt of the illuminant vector we see that the S/R of the prediction is highly dependent on the tilt of the illuminant relative to the directionality (*Figure 3.3.3*). The prediction is least accurate when the illuminant vector is parallel with the grain of the surface, falling to a minimum of 5 dB. We also note that the increasing rms roughness in the principal direction causes a significant decrease in the prediction accuracy.

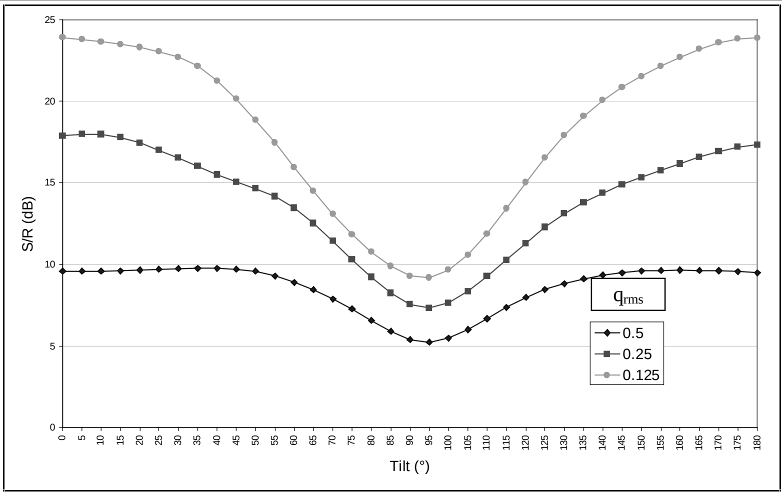


Figure 3.3.3 The accuracy of image prediction for a directional surface.

If we again set an arbitrary signal to noise limit of 10 dB we may form the following nominal bounds for the application of a linear approximation: surfaces should have an rms gradient not exceeding 0.25, whereas strongly directional surfaces should not be illuminated from tilt angles at less than 15 ° to the perpendicular of their primary axis.

3.3.4 Summary

In this section we found that the accuracy of the approximation falls as *rms slope* increases. Setting a nominal lower limit of 10dB for the S/R ratio we showed that isotropic surfaces could be modelled as long as the *rms slope* did not exceed 0.25. The accuracy with which directional surfaces could be modelled is highly dependent on the angle between the grain of the material and the illuminant. Again taking the 10dB limit, we found that surfaces with *rms slopes* of less than 0.25 could be accurately modelled except in the region ±15° of the material grain direction.

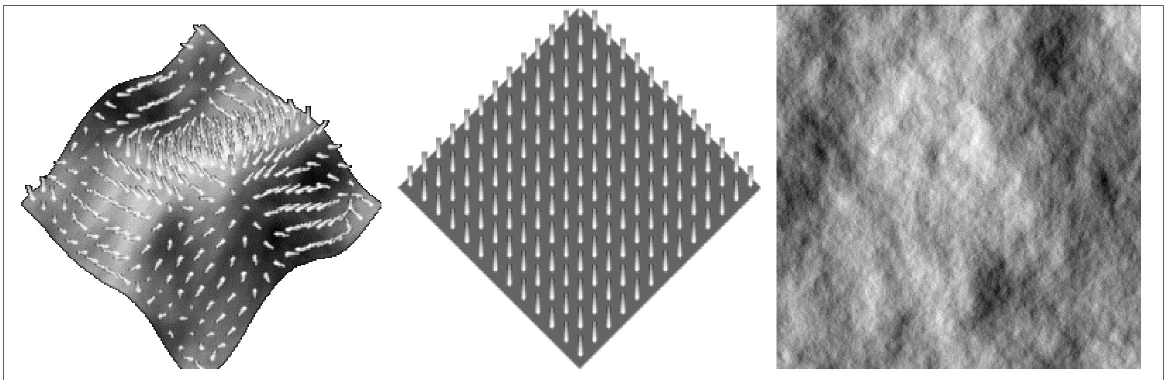
3.4 Kube's Model

Having considered the accuracy and scope of an optimal linear filter, in this section we shall consider an analytical linear model reported in the literature [Kube88]. Since the characteristics of the function, which the optimal model approximates vary with surface characteristics, a single analytical filter cannot be optimal for all possible surfaces. The *form* of the model may, however, be common to the optimal filters of a range of surfaces. The next question we ask is under what circumstances does the form of the theoretical model coincide with the behaviour of the least squares filter? Kube's model predicts that the intensity will be a linear combination of the surface derivatives scaled by a trigonometric function of tilt as well as a mean component. If the coefficients of the optimal linear model, recalculated for each tilt condition, vary in the same manner as the trigonometric function, we may conclude that the form of Kube's model is optimal in the least squares sense.

3.4.1 An Analytical Expression

We restate Kube's model in our own notation, dropping the assumption that the surface is fractal.

Consider the bandlimited surface with height map defined by the scalar function $s(x,y)$.



(a.) **S(**x,y**)** *(b.)* **L(**x,y**)** *(c.) i(x,y)*

Figure 3.4.1 The components of image formation.

(a) The Lambertian Image

let **S(x,y)** (*Figure 3.4.1*a) be the derivative field of the bandlimited scalar field *s(x,y)* such that

$$\mathbf{S(x,y)} = \quad \text{grad } s(x,y) = \begin{bmatrix} \frac{\delta z}{\delta x} & \frac{\delta z}{\delta y} & \frac{\delta z}{\delta z} \end{bmatrix}$$

Also define the illumination vector field **<u>L</u>(x,y)**, assuming illumination is produced by a point source located an infinite distance from the scene, the magnitude and direction of the vectors will be uniform throughout the field (*Figure 3.4.1*b).

$$\mathbf{L(x,y)} = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}$$

where

$k_1 = Cos\tau Sin\sigma$ $\quad$ $k_2 = Sin\tau Sin\sigma$ $\quad$ $k_3 = Cos\sigma$

Assume that a surface has a reflectance function which is:

(i) homogeneous over the surface, and

(ii) Lambertian.

Furthermore, assume that the surface is constrained such that there is no significant:

(i) cast or self shadows, or

(ii) degree of interreflection.

Adopting these assumptions we may state that the image field *i(x,y)* (*Figure 3.4.1*c)is the normalised scalar product of the surface derivative field and the illumination vector field.

$$i(x,y) = \frac{\mathbf{L}(x,y)^T \mathbf{S}(x,y)}{\left|\mathbf{L}(x,y)^T \mathbf{S}(x,y)\right|}$$

This results in a non-linear operation at each facet,

$$i = \frac{-k_1 p - k_2 q + k_3}{\sqrt{p^2 + q^2 + 1}} \qquad (3.4.1a)$$

where i, *p* and *q* are functions of *x* and *y*.

(b) Kube's Linearisation

Kube uses the Taylor Expansion to form a linear approximation to equation 3.4.1, shown below, truncated beyond the quadratic term.

$$\text{i}= k_1 p + k_2 q + k_3 - \frac{k_3}{2}\left(p^2 + q^2\right)$$

Using the first three terms he forms a linear estimate

$$\hat{i} = k_1 p + k_2 q + k_3 \qquad (3.4.1b)$$

or $$\hat{i} = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix}$$

This approximation is reasonable where:

$p>>p^2$ and $q>>q^2$

i.e. where

$p<<1$ and $q<<1$

(c) Frequency Domain Dual

Since equation 3.4.1b is linear, we can easily form a frequency domain dual. For simplicity we will work in terms of the magnitude (denoted by the *m* subscript rather than the power spectrum.

$$I_m(u,v)=P_m(u,v).k_1+Q_m(u,v)k_2+k_3\delta(0,0)$$

where $P_m(u,v)$ and $Q_m(u,v)$ are the spectra of the directional derivative fields.

However, the directional derivatives are related to $S(u,v)$

$$P_m(u,v) = iuS_m(u,v)$$
$$Q_m(u,v) = ivS_m(u,v)$$

where *i* represents a 90° phase shift.

If we restate these in polar co-ordinates:

$$P_m(\omega,\theta) = i\omega\cos\theta.S_m(\omega,\theta)$$
$$Q_m(\omega,\theta) = i\omega\sin\theta.S_m(\omega,\theta)$$

Ignoring the mean and using polar co-ordinates:

$$I_m(\omega,\theta) = k_1.i.\omega.\cos\theta.S_m(\omega,\theta) + k_2 \ i.\omega.\sin\theta.S_m(\omega,\theta)$$

Resolving the k terms into their trigonometric components and gathering like terms:

$$I_m(\omega,\theta) = i.\omega.\sin\sigma.S_m(\omega,\theta).[\cos\theta.\cos\tau + \sin\theta\sin\tau] \tag{3.4.1c}$$

Which can be simplified to

$$I_m(\omega,\theta) = i.\omega.\sin\sigma.\cos(\theta-\tau).S_m(\omega,\theta)$$

From the point of view of this thesis, the most important effect of this expression is that the image spectrum is a function of both the surface and the illuminant tilt.

3.4.2 A Comparison of Kube's Model with the Optimum Linear Model

Kube's model gives an analytical expression as to the relationship between surface and image, however, we can also empirically define a least squares mapping from *p* and *q* to intensity as we did in section 3.3.1.

We can therefore define a least squares model, with parameters *a* and *b*, to map the surface derivative fields to the image intensity fields for a given tilt. If we then vary tilt we can compare the behaviour of the parameters (*a* and *b*)with that predicted by Kube, i.e. $k_1=\cos\tau.\sin.\sigma$ and $k_2=\sin\tau.\sin.\sigma$ (σ is held constant at 60° in this work).

Consequently if Kube's model is near optimal in the least squares sense we would expect that the *a* and *b* parameters of the linear model to be equivalent to the parameters k_1 and k_2 and would follow $\cos\tau$ and $\sin\tau$ curves respectively as tilt is varied and the slant angle is held constant. We therefore plot the estimates of *a* and *b* for an isotropic surface against these expected functions in *Figure 3.4.2*.

If we test this prediction with an isotropic surface we see a linear relationship between the parameters estimated by the least squares criterion and those predicted by Kube's equation *3.4.1c*. However, it is immediately obvious that the slope of the line varies with the rms slope of the surface; this was not predicted by Kube's model. If we tabulate these results we can see that both the estimated parameters are highly correlated with Kube's predictions (see *Table 3.4.1*). It is also clear that the scaling factor falls as *rms slope* increases. Finally we note the estimated a and b parameters at a given tilt are scaled by almost equal amounts. We measure the degree of correlation between the parameters *a* and *b* with Kube's predictions.

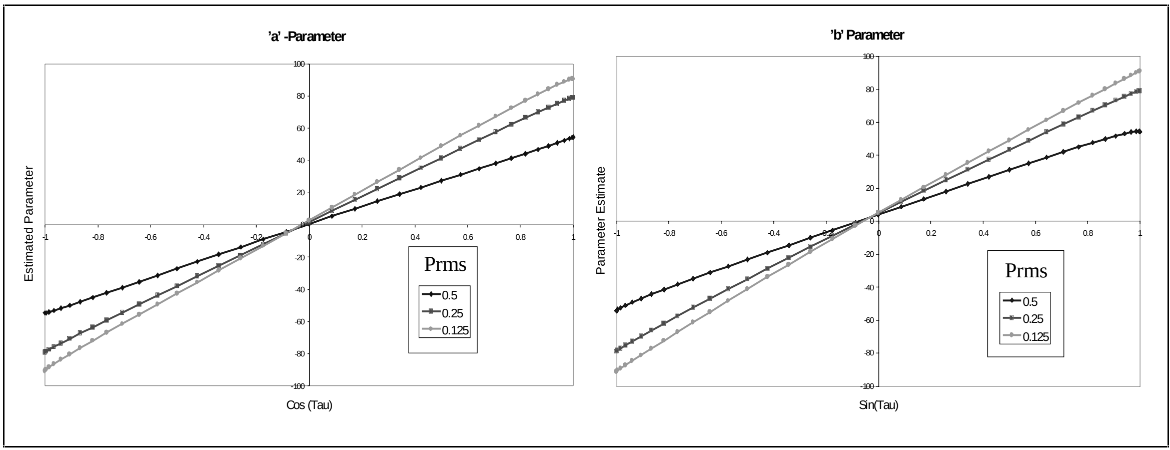


Figure 3.4.2 The variation of the least squares parameters plotted against those predicted by Kube (isotropic fractal surface).

	Scaling		**Correlation**	
RMS Slope	**a**	**b**	**a**	**b**
0.125	90.83	90.89	0.99988	0.99963
0.250	79.21	78.89	0.99995	0.99956
0.500	54.60	54.37	0.99995	0.99945

Table 3.4.1 The relationship of LS parameters to Kube's predictions.

If we now consider Ogilvy's directional surface we note the relationship is less linear (*Figure 3.4.3*), though the correlation coefficients are still high (*Table 3.4.2*). The trend of decreasing line slope with rising rms surface slope is still evident, though there is a growing disparity between the parameter scaling values as slope increases. This would seem to be in agreement with the inverse slope relationship found in both the isotropic and directional cases.

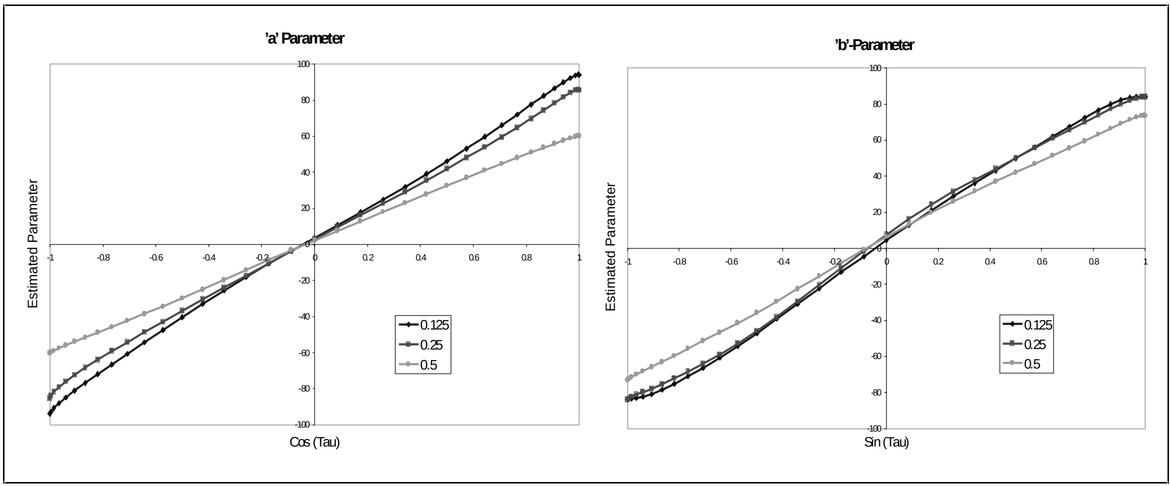


Figure 3.4.3 The variation of LS parameters for directional Ogilvy surface to Kube's predictions.

Slope		**Scale**		**Correlation**	
$\mathbf{p_{rms}}$	$\mathbf{q_{rms}}$	**a**	**b**	**a**	**b**
0.125	0.34	91.72	89.46	0.999537	0.9985
0.250	0.067	82.90	87.77	0.999554	0.9983
0.500	0.134	60.75	74.64	0.999833	0.999217

Table 3.4.2 The relationship of LS parameters for directional surface to Kube's predictions.

3.4.3 Summary

We conclude that Kube's model is near optimal in the least squares sense for isotropic surfaces, albeit with the proviso of a scaling factor dependent on the *rms slope* of the surface. Furthermore, while some distortion is apparent in the parameter estimates, Kube's model agrees to a satisfactory degree for Ogilvy's directional surfaces as long as the slope is moderate. The scaling effect is also apparent for the directional surface.

3.5 Real Textures

3.5.1 Test textures and their Spectra

In this section we shall use two exemplar textures and will consider whether the predictions of Kube's model and our simulations are borne out by real data. In this chapter we will concentrate on two textures: the Rock texture, *Figure 3.5.1*, and the Striate texture, *Figure 3.5.2*. The Rock texture is isotropic and of the textures used in this thesis, is the most similar in appearance to the fractal simulations. The Striate texture is highly anisotropic, and is the closest experimental texture to the synthetic Ogilvy surface.

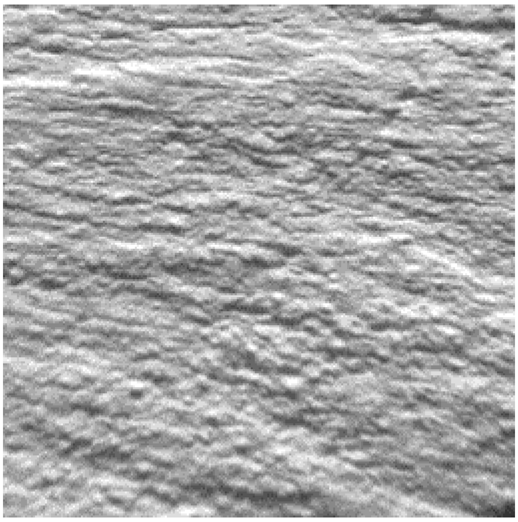


Figure 3.5.1 Rock Texture *Figure 3.5.2 Striate Texture*

If we consider the log magnitude spectra of the above images (*Figure 3.5.3*), we note a valley running at right angles to the direction of illumination in both spectra. This is consistent with Kube's prediction of attenuation of frequency components at right angles to the illuminant direction. In addition, the anistropic texture has an obvious ridge running in the perpendicular direction which is not apparent in the isotropic case.

Figure 3.5.3 Magnitude Spectra of Rock (left) and Striate (right) Textures τ=0°

If the light source revolves through 90° the power spectrum of the isotropic surface is effectively rotated by the same amount, however, the change in tilt causes the ridge of the striate texture to be lost. While the effect on the isotropic texture is equivalent to one of rotation, the image spectrum of the directional surface is of a fundamentally different image texture.

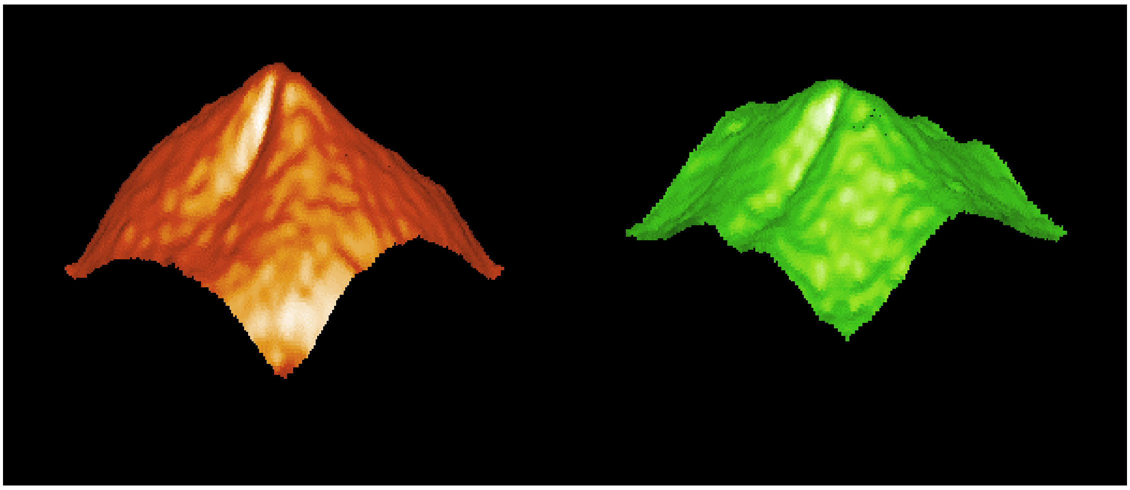


Figure 3.5.4 Magnitude Spectra of Rock (left) and Striate (right) Textures τ=90°

3.5.2 The Polar Power Distribution

In order to verify Kube's model, however, it is necessary to consider the effect more analytically. We focus our attention on the variation of spectrum power with polar angle. This is carried out by integrating the power along radii of the polar spectrum and plotting the quantity as a function of polar angle. Due to the finite quantisation of frequency at the low end of the spectrum, and the effects of noise at the high end, this integration is performed within certain frequency bounds. The bounds used for our first experiment are $\omega = 0.156f_s$ to $0.469f_s$ shown in *Figure 3.5.5*.

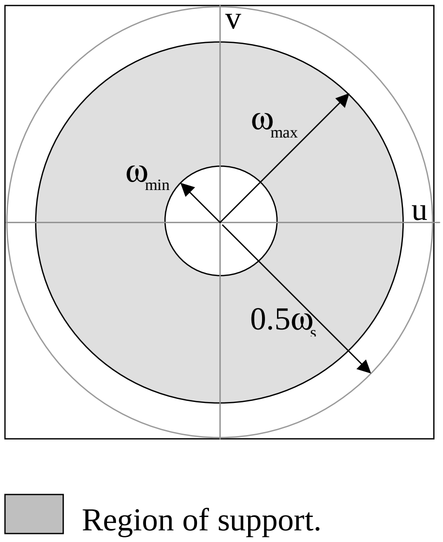


Figure 3.5.5 Region of support for polarogram.

Plotting this quantity for the rock texture *Figure 3.5.6* with the least squares fit of both the Tau 0 and 90° images, we find that it is in agreement with the $|\mathrm{Cos}(\theta-\tau)|$ (or $\mathrm{Cos}^2(\theta-\tau)$ for the power spectrum) relationship predicted from Kube's work with the addition of a small bias term. This result is also in agreement with that reported by Chantler [Chantler94] for a similar isotropic rock surface.

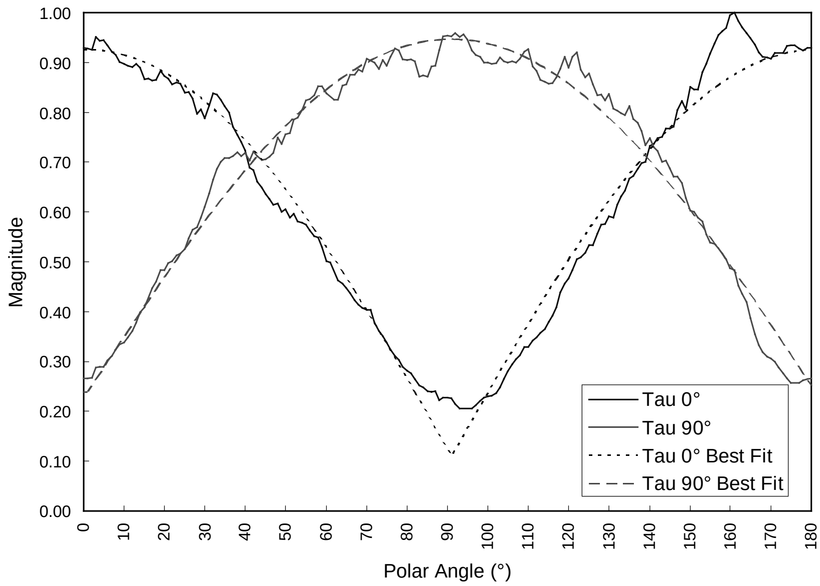


Figure 3.5.6 The variation of image magnitude with polar angle.

In the case of an isotropic surface, the dominant source of image directionality is due to directional illumination. However, if the surface is anisotropic, the surface directionality will also contribute to the polar distribution of power, *Figure 3.5.7*.

Figure 3.5.7 Polar distribution of image power for Striate texture.

3.5.3 Summary

In this section we took two real textures and showed that their image spectra are directionally attenuated as predicted by Kube's model. Using the isotropic test texture it was possible to verify the $|\cos(\theta-\tau)|$ relationship predicted by Kube. For the isotropic surface, varying illuminant tilt is equivalent to rotating the texture, however, for the directional surface the attenuation causes a fundamental change in the visual texture which cannot be modelled as rotation.

3.6 Conclusions

This chapter has considered the underlying mechanisms of reflection and discussed four analytical descriptions of the phenomenon. Three of the models were implemented and compared with the function measured for the surface type used in the experimental work. Both Oren's and Lambert's models were found to be effective methods of modelling, and, whereas Oren's model was the more accurate over a range of incident angles, Lambert's Law was found to be more accurate when the slope distribution of a rough surface was taken into account. We then considered the circumstances under which a linear filter could maintain adequate accuracy. Taking a nominal accuracy threshold of 10dB we concluded the linear form could accurately predict images of surfaces with an rms slope of less than 0.25. The highly directional test surface had the additional restriction that the illuminant should not make an angle of less than 15° from the grain direction of the material.

The third part of this chapter considered an analytical frequency based description of the link between surface and image. A model reported in the literature, [Kube88] was investigated to see whether it is optimal in the least squares sense. We concluded that the form was optimal for most surfaces, provided that a scaling factor dependent on the surface roughness was used.

The observation of the real isotropic surface *Rock* confirmed Kube's predictions for the angular distribution of power. The behaviour of the directional surface *Striate* illustrated that the tilt effect can either attenuate or accentuate useful directional information. This implies that tilt variation will have serious consequences for the discrimination of rough surfaces.

Chapter 4

The Imaging Process

4.1 Introduction

In the previous chapters we have considered the modelling of rough surfaces and their interaction with light to form a textured image. However, strictly, machine vision algorithms do not operate on images. An algorithm is applied to a data set. The aim of this chapter is to describe the link from the image, whose formation was described in the previous chapters to the data set, on which the classifier discussed in the next chapter will operate.

The motivation for the development of this model is two-fold. First, a recurring theme in this thesis is that texture analysis cannot be properly investigated in isolation, but must be observed in the context of the physical process of which it is part. The need for an analytical description of the imaging process follows from this premise. Since CCD cameras are the most common imaging devices in machine vision, it is logical that we concentrate on these devices. There is however a second, more practical, motivation: in the latter part of this thesis several algorithms will be proposed. In order that their performance may be assessed under realistic conditions, *yet in a controlled manner,* we must develop a 'simulation engine' which characterises noise and distortion in a realistic way. This engine will be based on the work reported in this chapter.

4.1.1 Noise and Texture

We note that noise has largely been ignored in the texture analysis field. Those papers which do consider the effect of noise, such as [Liu90], [Szirani96] and [Porter97], do not cite experimental evidence to support their models. Liu and Jernigen state that their noise is random, additive and, the author presumes, white. They use the model to select a subset of measures from a feature set defined in terms of the power spectrum, and then evaluate their performance for various S/N ratios. Szirani investigates the use of neural networks for the tasks of deblurring and texture segmentation. In the case of texture segmentation he investigates the effect of additive noise and the number of quantisation bits upon the classification accuracy. Porter and Canagarajah use additive white noise, with a signal to noise ratio in the range 0-25dB, to compare the robustness of Markov models, Gabor filters and wavelet approaches, though again with no physical justification of either the form or the magnitude of the noise model. Outwith texture analysis, Terrillon, [Terrillon96] adopts a more analytical approach, relating specific noise mechanisms, i.e. speckle and film grain noise to rotation-invariant classification. We note that unlike many of the effects observed in this chapter, these effects are either isotropic, or rotate with the image.

4.1.2 Organisation of this Chapter

The aim of this chapter is the development of a spectrally defined model of the imaging process. We noted in the previous section that where texture analysis does account for noise, it is usually considered to be white and additive, without any theoretical or experimental justification. In this thesis, we have taken a physically-based approach to the effects considered in the belief that this approach is more rigorous and leads to a more accurate and more flexible understanding of how surfaces can be classified. In this context, we do not believe that it is appropriate to treat noise as being additive and white without obtaining supporting evidence.

This support may come from two direct sources, the literature, or experiment. Unfortunately the literature is fragmented and deals mostly with mechanisms, rather than the spectral effects which are of relevance to this thesis. Similarly, the experimental approach is also limited, since many of the effects are difficult to resolve from the correct image. Where the effect is not easily separable from the correct image, we may use a third approach—simulation. This allows us to take a mechanism from the literature and observe the nature and magnitude of its effect for a textured image of the type considered in this thesis.

The process of developing an imaging model consists of four stages (*Figure 4.1.1)*. In the first, the literature is surveyed in order to identify noise mechanisms and, where possible, their spectral effects. At the end of this section, the noise mechanisms will be categorised as belonging to four functional classes. These classes are organised on the basis of the experimental approach which will be required to characterise and measure them in sections 4.4 and 4.5. In the experimental section, we set about measuring those categories of imaging artefact which can be directly observed. Those which cannot be directly measured are further divided into categories we are unable to analyse, and those which can be investigated using simulation. In the former case we are limited to a qualitative description of the effect, while in the latter, a model of the effect *specific to an exemplar texture* will be developed. The descriptions obtained will be integrated with those from the simulation section to form a complete noise model.

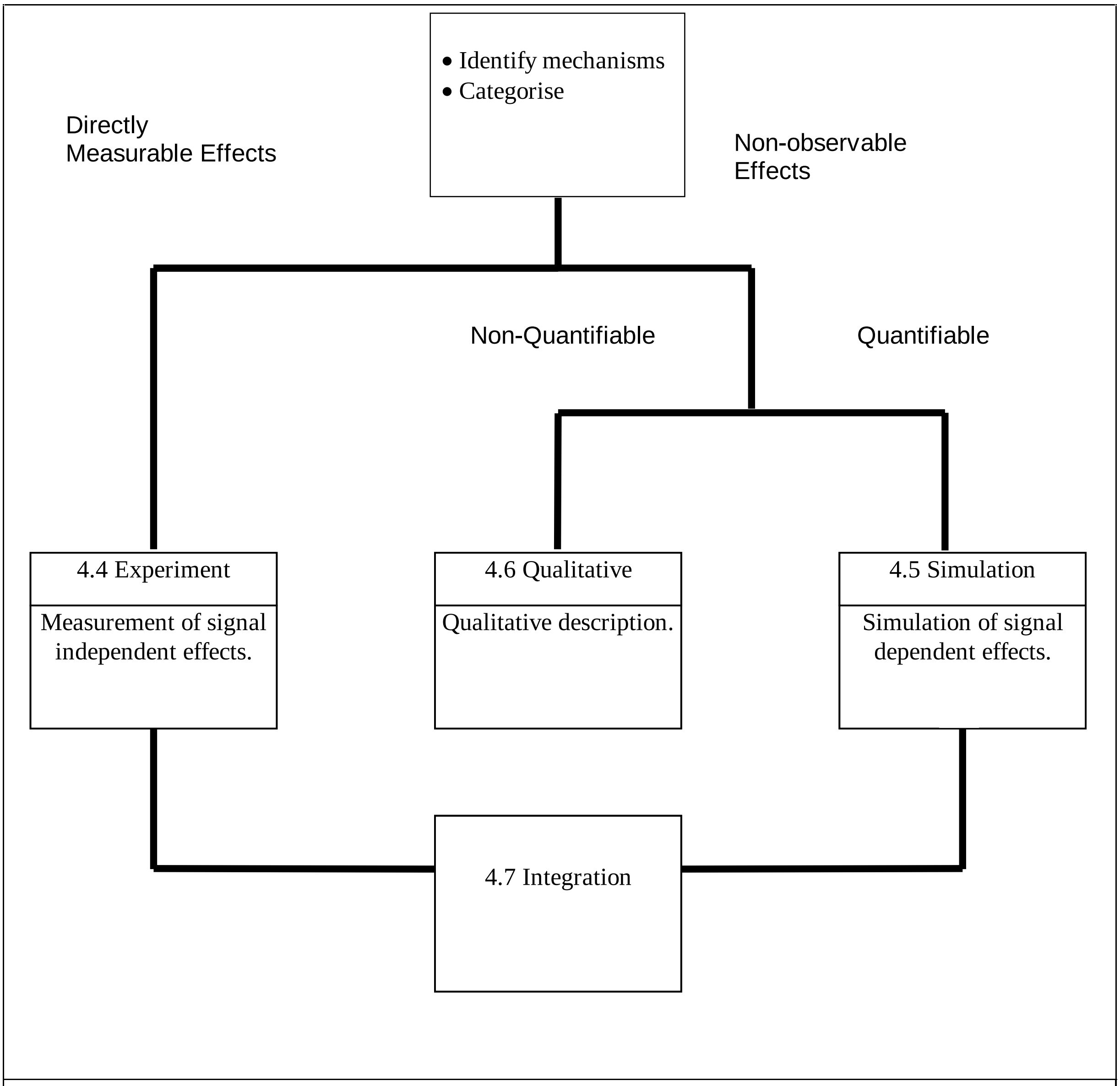


Figure 4.1.1 The organisation of the noise investigation.

Since this chapter is the first point at which we can directly observe the data set, in the final section we compare both the first order statistics and the measured spectra with those predicted by our models. This represents our first opportunity to use our surface, rendering and imaging simulations in order to verify our models.

4.2 Literature

The aim of this section is firstly to identify noise mechanisms from the literature and then to categorise them into classes according to the experimental approach required to quantify them.

4.2.1 Overview

Though several authors do give descriptions of the entire process, e.g. [Healey94] and [Lenz90], most of the literature relevant to the imaging process is usually limited to a single effect. This specialisation is problematic in our application; we are interested in the overall effect, rather than the individual components. This section surveys the literature to obtain a description of the mechanisms which form components in the transition from image to data set. The organisation of this section is therefore arranged by mechanism, rather than perceived effect, and follows the order in which the signal encounters these mechanisms. We use three categories; system optics, the CCD array, and the framestore (*Figure 4.2.1)*

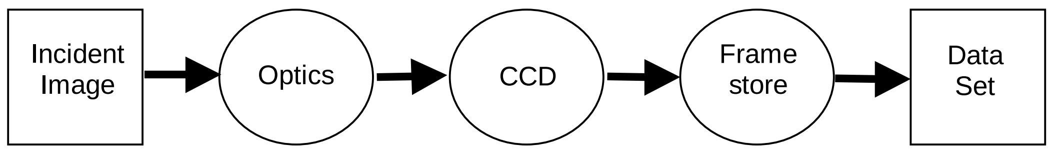


Figure 4.2.1 Section approach.

While this organisation is imposed by the nature of the literature, it is of limited use in our application. This section therefore concludes by categorising noise mechanisms by the experimental approach which will be used to quantify them.

4.2.2 System Optics

We retain the assumption of orthographic projection, but now consider sub-optimal focusing. Drawing on the work carried out in the field of shape from focus, we adopt the term Point Spread Function (PSF) which represents the image *i(x,y)* resulting from a single point source imaged by the camera. This is equivalent to the two dimensional impulse function of systems theory. Although usually associated with optics, the PSF may be used to characterise the transfer function of the entire imaging system.

Subbarao and Lu use paraxial geometric optics and the assumption of spatial invariancy to develop an cylindrical PSF [Subbarao94]:

$$b(x,y) = \frac{4}{\pi d_m^2} \qquad \text{where } x^2 + y^2 \le \frac{d_m}{4}$$

and (4.2.2a)

$$b(x,y) = 0 \qquad \text{where } x^2 + y^2 > \frac{d_m}{4}$$

where d_m is the diameter of the blur circle in millimetres. As the diameter of the blur circle decreases, the image becomes sharper; as the diameter approaches zero, the impulse response of the optical system approximates the ideal impulse function. Pentland argues that the PSF is best modelled as a two dimensional Gaussian function $b(r,\sigma_b)$ where σ_b is a spatial constant proportional to the diameter of the blur circle [Pentland87], (in fact Subbarao also provides the option for a Gaussian PSF in his simulation engine). This model is also used by Nayar and Nakagawa [Nayar94] who state it both in the spatial and frequency domains:

$$b(x,y) = \frac{1}{2\pi\sigma_b^{\,2}} \exp\left[-\left(\frac{x^2+y^2}{2\sigma_b^{\,2}}\right)\right] \tag{4.2.2b}$$

$$B(u,v) = \exp\left[-\left(\frac{u^2+v^2}{2}\right)\sigma_b^{\,2}\right] \tag{4.2.2c}$$

where *B(u,v)* is the Fourier transform of *b(x,y)*.

An example of the effect of the blurring function is shown in *Figure 4.2.2*.

Figure 4.2.2 Focused and defocused (σ_b=0.04) images.

We briefly note one other effect of the optical system: vignetting is an artefact of the compound lens system which causes a roll-off in sensitivity with distance from the centre of the image. The effect is most significant when using wide angle lenses or when using a lens at full aperture. Kolb et al. [Kolb95] give an example where intensity at the edges of the sensor plane is one third that of the centre. However, like Subbarao we explicitly ignore this effect, and assume it to be corrected for in the set up of the system.

4.2.3 CCD Array

We now consider the CCD array itself, a clear description of the noise mechanisms and observed effects is difficult since each mechanism may contribute to one or more observed effect. Each effect in turn may be due to a combination of noise sources which are not necessarily confined to the array. Since researchers generally investigate only part of the process, or to varying degrees of depth, taken as a whole, the literature often appears to be inconsistent. The aim of this section is to present a coherent model of effects occurring on the array in a form that is relevant to texture analysis.

We begin with dark current noise: as with any silicon device at ambient temperature, thermally generated electrons are produced within the body of the sensor. If these electrons wander to the sensing surface they may be captured by a photoelectron site, giving rise to noise. This effect is stochastic: the generation of an electron is a random event. However, due to the inhomogeneous nature of the lattice, electrons are more likely to be generated at some locations than others. While the subsequent movements of the electrons will also be random, some sites will be more prone to electron capture than others [Talmi80]. In this sense dark current noise may be argued to contribute to both fixed pattern and spatially independent noise (as shot noise). It should also be noted that this effect is highly dependent on temperature; each 6.7° rise in temperature will cause a doubling in the dark current, so changes in temperature that imperceptible to the human observer may cause significant changes in both the fixed pattern and spatially independent noise.

A fundamental limitation of all imaging devices is the quantum nature of light: the number of photons which arrive at a site in a given period is a Poisson process and may be described as shot noise [Healey94]. In most practical applications the large number of incident photons will cause the distribution to tend towards a Gaussian distribution with variance equal to its mean. The larger the number of incident electrons, the less significant the standard deviation will be relative to the mean.

Two other effects occurring at this stage are blooming and non-uniform sensitivity. Blooming occurs when some sites are intensely illuminated causing a large number of electrons to be generated at a site, some of which then overflow to adjacent sites; we will neglect this effect. Non-uniform sensitivity is due to variations in the quantum efficiency of sites caused by fabrication effects [Healey94], we shall consider this to be fixed pattern noise.

4.2.4 Framestore

The previous sections described the process of measuring image intensity over a discrete grid, this section is concerned with the transfer to numerical representation. The process may be described in two steps: first a serial readout of the rows, with low pass filtering of the pulse train and the addition of the synchronisation signal. In the second stage, the signal is stripped of synchronisation information, resampled and quantised before being stored in memory. We describe three problems associated with this process identified in the literature:

1. serial filtering effects,
2. errors in synchronisation, and
3. quantisation errors.

Serial Filtering

Having collected the image, it is necessary to convert it to a continuous, interlaced, serial signal. In their paper Boie and Cox develop a camera model in electronic terms [Boie92].

The circuit model consists of both serial and parallel components, the parallel noise mechanisms i.e. leakage currents and Johnson noise are modelled by the i_n term, the serial component models noise associated with amplification and is also modelled as a current source, $e_n 2\pi\omega C_t$. These components are combined in (4.2.4a) to model the noise PSD due to the camera:

$$2G_n(\omega) = \overline{i_n^2} + \overline{e_n^2}(2\pi\omega)^2 C_t^2 \qquad (4.2.4a)$$

where

$G_n(\omega)$ is the total noise current power density,

i_n is the total noise current,

e_n is the equivalent input noise voltage,

C_t is the total camera capacitance,

The model is high pass in character, and predicts that noise power will increase with the square of frequency. However, since the camera is being readout in a serial manner this characteristic only applies to the rows, the noise at different points on the same column will be uncorrelated and the noise signal along columns is therefore predicted to be white. In fact Boie found both spectra to be essentially white and goes onto argue that the bandlimiting in the preamplifier acts as a prewhitening filter.

Synchronisation

The next stage we consider covers the removal of synchronisation information. The CCD array is read out in rows, each of which corresponds to a low pass filtered signal bounded by a blanking interval. If the rows are not perfectly synchronised, sampling points will be out of alignment with those of adjacent rows. Wang and Chen report this error is within the bounds ±0.4 pixels [Wang96].

Quantisation

The final stage considered is the quantisation of the signal. Most video analogue to digital converters (ADC) have a resolution of 8 bits, giving a random error uniformly distributed between $\pm\frac{1}{512}$ of the full scale deflection. The quantisation noise may be treated as a white noise process with variance $\frac{0.5^2}{12}$ [Healey94] . In practice quantisation noise is more significant than this figure would suggest since the full dynamic range of the ADC is rarely utilised and the presence of both random and fixed pattern noise reduces the proportion of the ADC range which is available to the signal.

4.2.5 Conclusions

In this section, we have discussed the noise mechanisms described in the literature. In order to incorporate these mechanisms into a spectral model of the observed effects, we must:

- characterise the spectral shape, and
- measure its magnitude.

The approach we take will depend on the nature of the mechanism. We define four functional categories of artefact; transfer function, fixed pattern noise, temporal noise—which we define to be variations in the image of the same scene over consecutive frames, and signal dependent noise. A given mechanism may contribute to one or more of these categories.

The predominant effect of the optical system considered in the literature is the blur function, we adopt Pentland's model and treat it as a Gaussian transfer function.

Dark current noise was considered to be stochastic, contributing to both temporal and fixed pattern noise. Shot noise was described as being temporal, having a white spectrum and conforming to a Poission distribution, which tends to the Gaussian case as irradiance increases.

We have been unable to find any work on the spectral effects of line jitter, and hence we are forced to resort to simulation in order to estimate the effects line jitter will introduce into our model.

Serial filtering will affect the transfer function, however, the effect may not be noticeable due to the dominance of optical blur. The filtering effect may manifest itself in the filtering of noise generated subsequent to the optical system, most obviously in temporal noise.

Quantisation noise has on the one hand been described as being signal dependent [Lim, p.592] and on the other, as being white [Healey94]. While these descriptions are not necessarily incompatible, we believe it is necessary to verify that the white noise assumption is safe for the type of images used in this thesis. We again use simulation to assess the character and degree of the effect.

	Transfer function	Fixed Pattern Noise	Temporal Noise	Signal Dependent Noise
Blur	√			
Dark noise		√	√	
Shot noise			√	
Filtering	√		√	
Jitter			√	√
Quantisation				√

Table 4.2.1 Noise categories.

If we assume a linear transfer function, the system transfer function is easily measured using calibration images. Temporal noise is also relatively easy to measure, using disparities in multiple images of the same scene. Fixed pattern noise can be observed by preventing light from entering the lens and observing the resulting image. However, if the camera does not have its autogain disabled, it is not possible to measure the effect relative to a textured image. Unfortunately, we found that the autogain was required for stable operation of the camera/framestore, and we are unable to measure the effect, though we do characterise the effect in section 4.4. Signal dependent noise is also difficult to measure and model. We adopt the following approach. The effect is simulated with a synthetic exemplar texture, the difference between the correct image and the noisy image is treated as being additive noise, (we shall describe it as pseudo-additive). In this way we treat imaging artefacts as being either a transfer function or as (pseudo) additive.

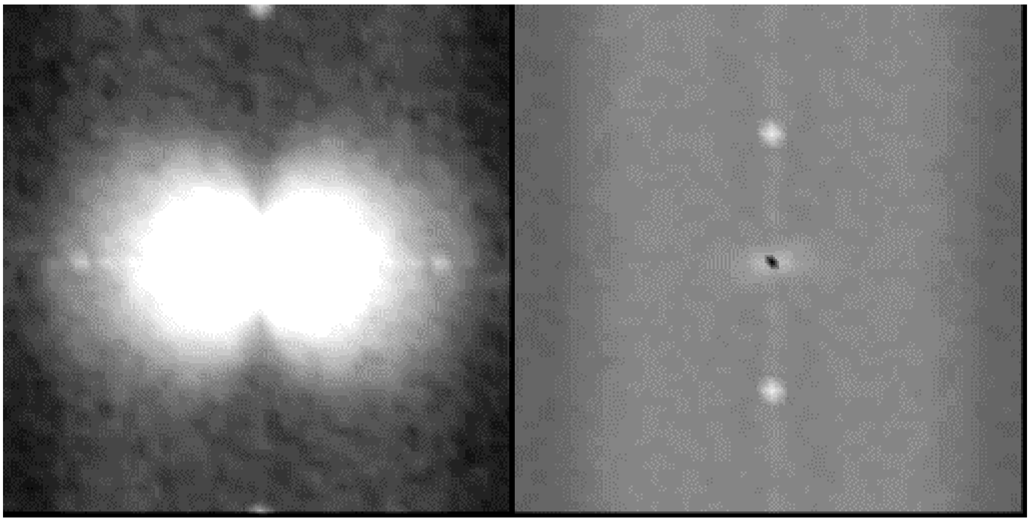
4.3 The Case for Sub-sampling

4.3.1 Introduction

The easiest form of noise to identify is that disclosed by temporal variation of the image. Using multiple images of the same scene we observe the spectral characteristics of temporal noise relative to that of the static part of the image. The spectral characteristics of the S/N ratio lead us to implement sub-sampling in the test images. In the next section we estimate the overall transfer function of the system and the remaining temporal noise, subsequent to pre-processing.

4.3.2 Justification of Sub-sampling

We average 10 frames to produce a 'clean' image and produce residual error images from the original images, we then average the PSD of these images, *Figure 4.3.1*. We note that the temporal noise is clearly coloured. This immediately calls into question the white noise assumption made in most papers which deal with the effects of noise on texture.



(a.) (b.)

Figure 4.3.1 Spectra of Time averaged image (a) and averaged residual spectra (b)

We define two nominal, possibly overlapping, classes of noise: spectral artefacts and temporal variation. Taking the PSD of a captured rock image, *Figure* ***4.3.4***, we identify three suspected artefacts of the imaging process in *Figure* ***4.3.3***: a wide vertical-running noise plateau (1), and 2 groups of spikes (2) and (3). We confirm that these are artefacts of the imaging process by rotating both the light source and the surface by 90°, *Figure* ***4.3.2***. While the characteristic illumination effects rotate, the artefacts (1)

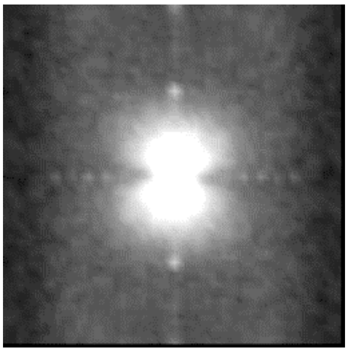


Figure 4.3.2 Rotated Rock PSD

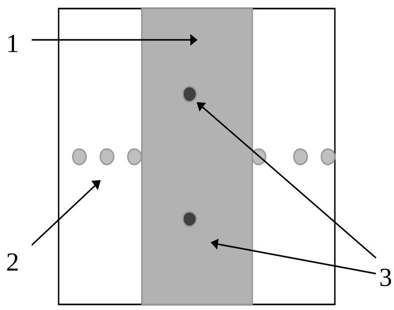
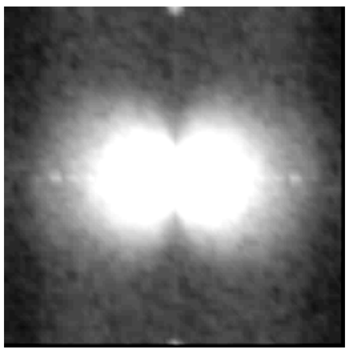


Figure 4.3.3 Key Figure



4.3.4 Rock Image Spectrum

and (2) remain static, showing that these effects are associated with the imaging process.

We observed from *Figure 4.3.1* that the temporal noise spectra is not white. While the spectra is uniform in the vertical direction, there is a pronounced "plateau" running in the horizontal direction. In *Figure 4.3.5* we plot the profile of the 1D spectrum in the horizontal averaged over the spectrum columns. There is a clear trend which is remarkably uniform for all the textures.

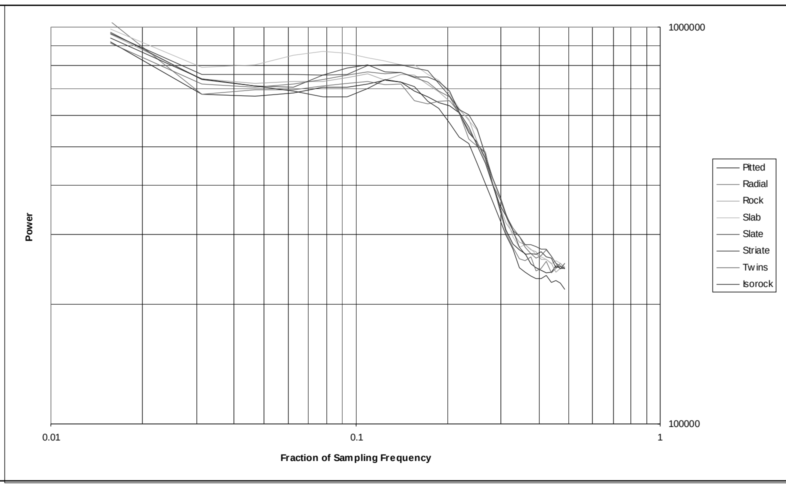


Figure 4.3.5 Horizontal Profiles of Residual Power Spectra

We now consider noise in the context of imaged test surfaces. We measure the signal to temporal noise ratios of several textures along both the vertical and horizontal axes of the frequency plane. The textures are illuminated from the vertical direction, and the signal is at its maximum in this direction. According to our models, the signal along the horizontal axis is composed of the non-linear components and noise. The frequency dependent signal to temporal noise ratio is defined below:

$$E_i(\omega) = 10\log_{10}\left(\frac{\sqrt{I_i(\omega)^2 - E_r(\omega)^2}}{E_r(\omega)}\right)$$

where

$I_i(\omega)$ = the power of the average measured image at frequency ω.

$E_r(\omega)$ = the power of the residue image.

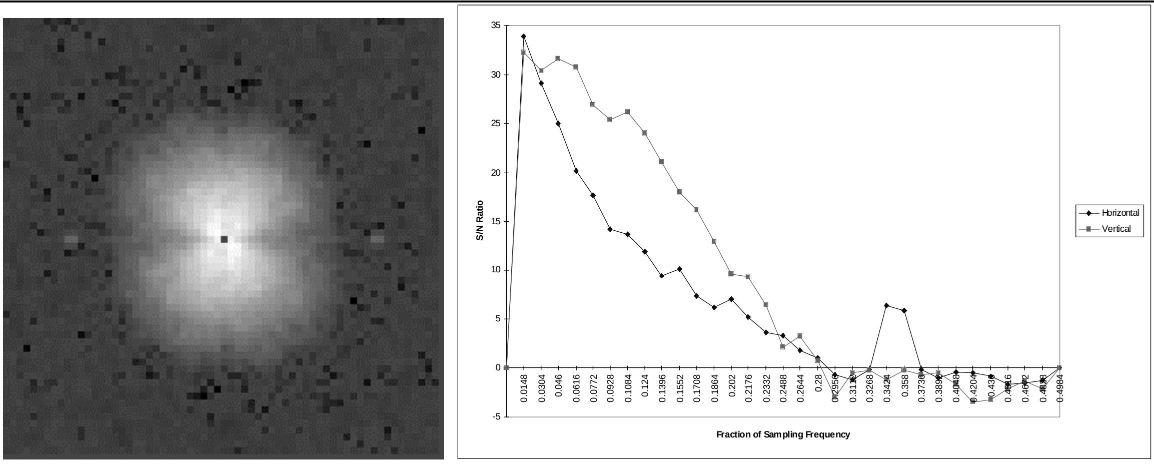


Figure 4.3.6 Signal to Temporal Noise Spectrum for Rock Texture

The signal to temporal noise ratio, *Figure 4.3.6*, shows a clear frequency dependency as we would expect from the fractal nature of the image. High frequencies are characterised by low signal to noise ratios of 0dB or less. This leads us to believe that little information, useful for classification, will be found at these frequencies. This immediately commends low pass filtering and decimation for two reasons.

1. Low pass filtering will remove the majority of noise spikes and the associated attenuation of high frequency, vertical running components will mean that the temporal noise process can be effectively described as being white.
2. The suppression of high frequencies will allow decimation of the images with the associated reduction in processing time for the classification algorithm.

This procedure was therefore adopted, and all the images were low pass filtered with a 7th order Butterworth filter implemented in a 25X25 mask prior to decimation. A linear image rescaling was also carried out at this stage.

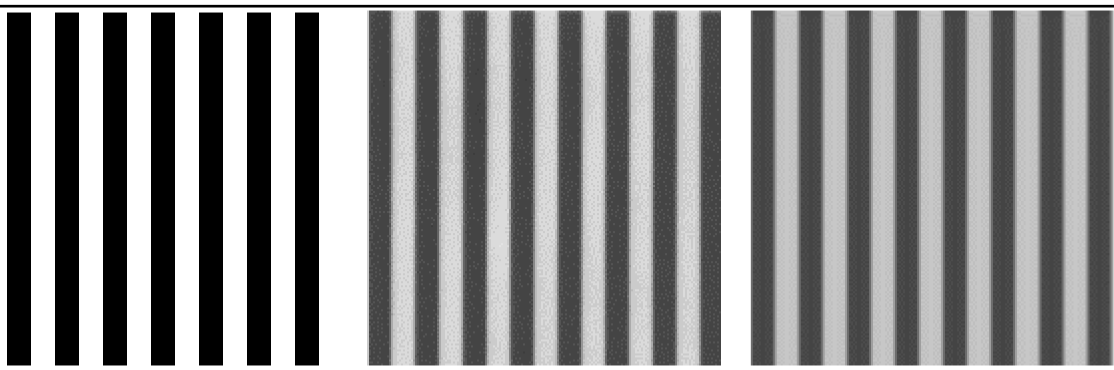
4.4 Experimental Investigation

4.4.1 Introduction

This section is concerned with the experimental measurement of image artefacts. We estimate the overall transfer function of the system, including pre-processing, using a calibration image, and fitting an experimental model to the observed data. As with the previous section we use multiple images of the same scene to identify temporal variation before quantifying the effect in terms of its first and second order statistics.

4.4.2 Transfer Function

Having adopted the down-sampling procedure, we now investigate the overall effect of the blur function. The blur function is assumed to be isotropic and Gaussian in accordance with Pentland [Pentland87]. We compare a calibration image with a synthetic image and a blurred synthetic image *Figure 4.4.1*, the spectral results are shown in *Figure 4.4.2*.



(a) Synthetic *(b) Actual* *(c) Blurred synthetic*

Figure 4.4.1 Synthetic, actual and blurred calibration images.

The blur function with σ_b=0.02 accurately predicts the magnitude of the peaks in the calibration spectra, with the exception of the very low frequencies, i.e. below 0.025 f_s. We therefore adopt this as our optics model.

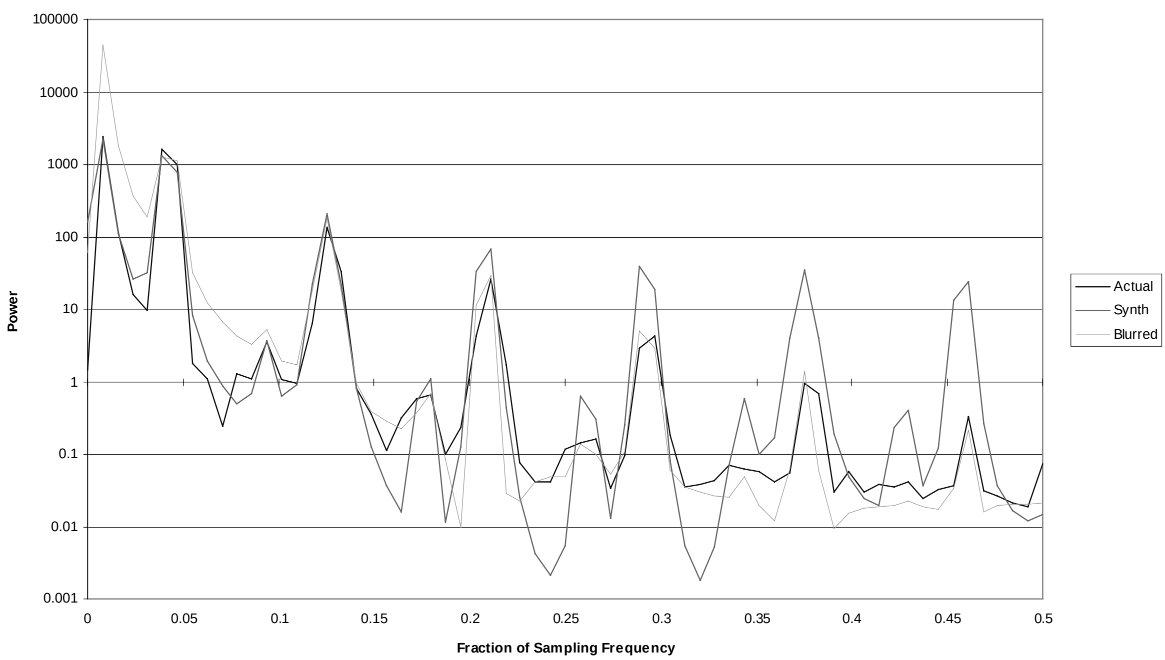


Figure 4.4.2 Comparison of Actual, Synthetic and Blurred Synthetic Calibration Spectra

4.4.3 First Order Statistics of Temporal Noise

Now, let us consider temporal noise. We treat any variation over consecutive frames of the image of a given scene as being noise. We call this temporal noise and model it as being additive. We begin by estimating the first order statistical properties of this noise. Averaging a series of temporal images will reduce temporal noise. The additive noise will be of the same order of magnitude as the quantisation noise and a disparity histogram calculated from a mean image will have a low standard deviation, making it difficult to form a good estimate of its properties. Instead we mitigate this problem by subtracting two images of the same scene; this is equivalent to the subtraction of two random noise processes.

Texture	Mean Disparity	Standard Deviation of Disparity	Texture	Mean Disparity	Standard Deviation of Disparity
Isorock	0.45	1.78	Slab	0.03	1.69
Pitted	0.08	1.58	Slate	0.40	1.68
Radial	-0.02	1.59	Striate	-0.02	1.67
Rock	-0.35	1.64	Twins	-0.07	1.62

Table 4.4.1 Table of image disparities.

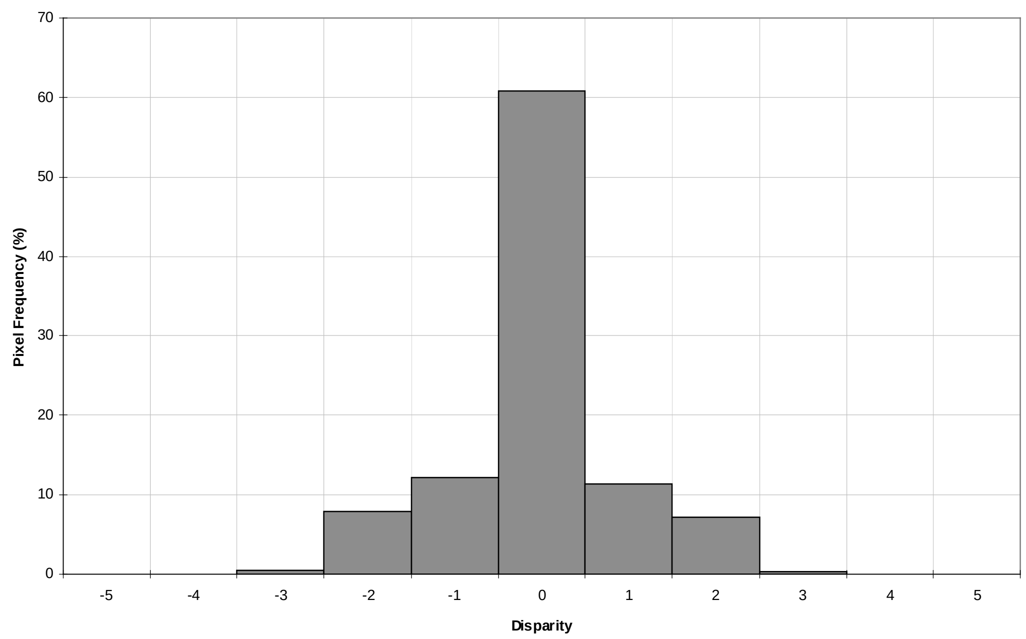


Figure 4.4.3 Disparity histogram for rock texture.

If we assume that the temporal noise distributions, which underlie the disparity distribution, are identical and Gaussian, we may infer their standard deviations from that of the disparity histogram.

$$\sigma_1 = \sigma_2 = \sqrt{\frac{\sigma^2_{disparity}}{2}}$$

This results in an estimated standard deviation (σ_t) of 1.17 averaged over all the textures. We may also usefully state the temporal noise level in terms of the signal to noise residue.

Texture	**S/R (dB)**	**Texture**	**S/R (dB)**
Rock	25.90	Slate	15.90
Striate	24.08	Pitted	19.63
Isorock	22.67	Twins	20.76
Slab	26.23	Radial	17.20

Table 4.4.2 Signal to temporal noise levels for the test textures.

4.4.4 Spectral Characteristics of Temporal Noise

Having observed the statistics of the disparity images, and inferred those of the temporal noise process, we now consider the spectral characteristics of the temporal noise. We do so by observing the spectra of the same disparity images which will be the scaled image of the noise spectra.

In section 4.3 we advocated the use of sub-sampling for the images used in this thesis, we now make a polar plot of the temporal noise spectra of the subsampled *Rock* image (*Figure 4.4.4*), for comparison we also show the polar plot for the original Rock image prior to sub-sampling. We can see that the noise power in the subsampled image is far more evenly distributed throughout the angular range than that of the original image. With the exception of the peak at θ=90°, the plot shows the noise power distribution to be essentially isotropic in the case of the sub-sampled image.

If we now plot power against radial frequency for the *Rock* image (*Figure 4.4.5*), we note a gentle fall in power with increasing frequency. A least squares fit over the entire frequency range gives an estimated β=0.70, if the beta parameter is measured over the midband range f_s=0.14 to 0.38, a much lower estimate of β=0.39 is made.

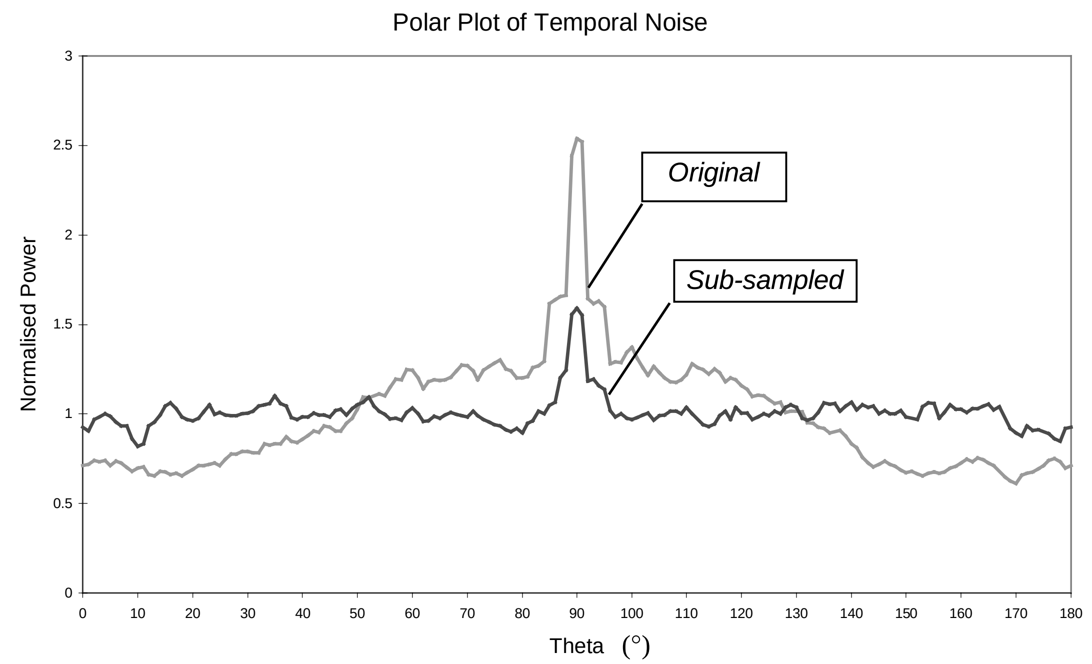


Figure 4.4.4 Polar plot of temporal noise spectra of original and sub-sampled images.

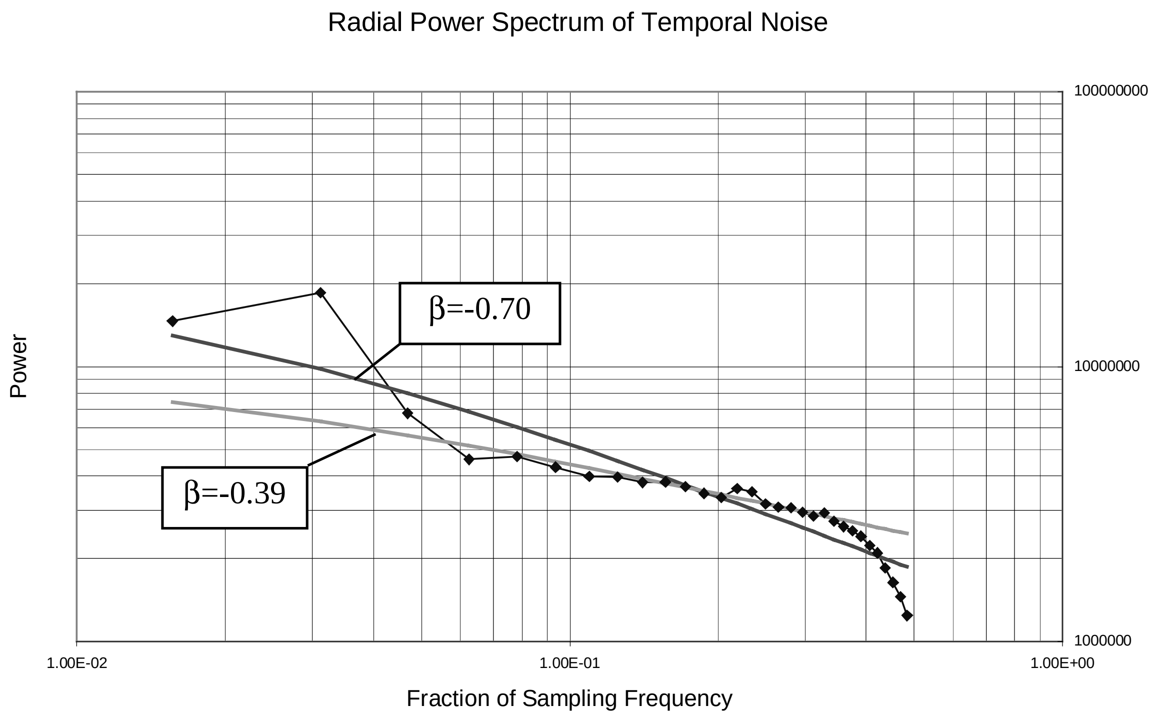


Figure 4.4.5 Radial power spectra of temporal noise associated with (sub-sampled)Rock image.

The result of the polar and angular plots is the description of the noise process as being of isotropic and mildly low pass in character. Given the relatively low estimates of the beta parameter, and the convenience of the assumption, we will assume in this thesis that the temporal noise is white.

Comparison of the radial noise spectra for different textures (*Figure 4.4.6)* show them to be similar. Some differences are apparent at low frequencies, though we note that these correspond to only a few spectral components. Overall, we believe the similarity of the spectra vindicate our assumption that temporal noise can be treated as being signal independent, at least within the class of textures considered in this thesis.

We therefore conclude that the spectra of temporal noise associated with the textures of this thesis may be modelled as being realisations of the same noise process, independent of the texture characteristics, and, furthermore, that the process may be well approximated by a white noise process.

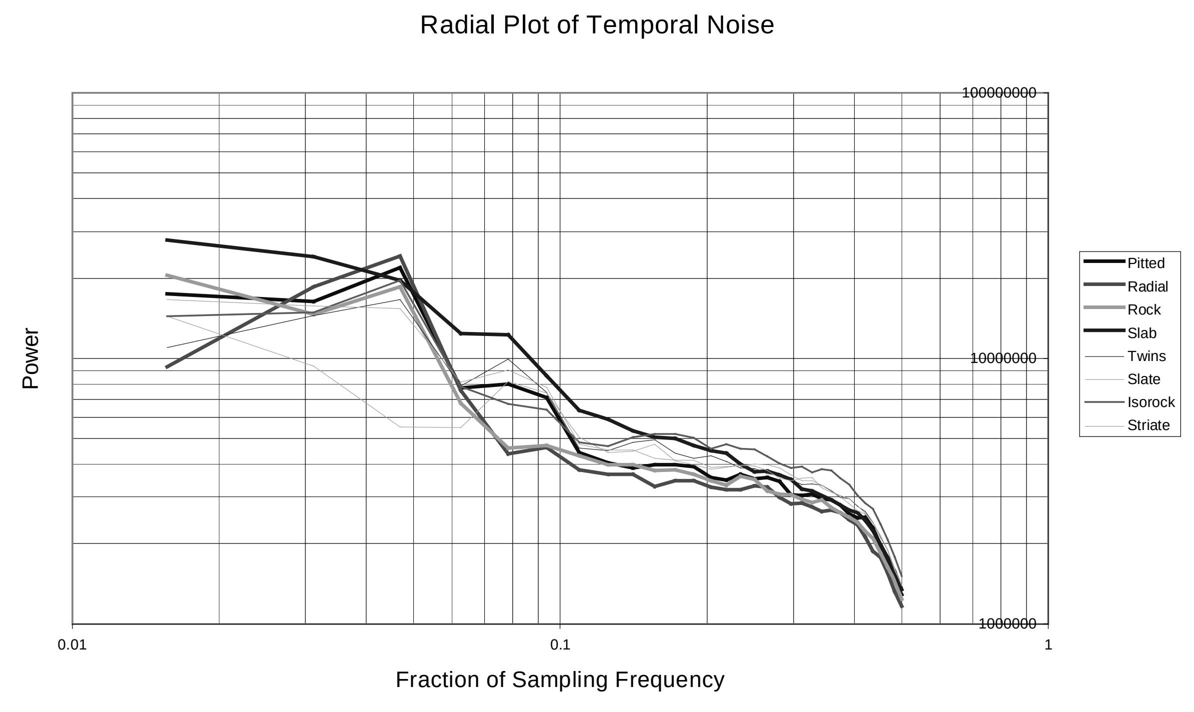


Figure 4.4.6 Radial temporal noise spectra.

4.4.5 Summary

In this section we have found the transfer function to be well approximated by a Gaussian with parameter σ_b = 0.02. It was also concluded that the temporal noise associated with the subsampled image is reasonably approximated as white noise of variance 1.17.

4.5 Use of Simulation to Investigate imaging phenomena

In the literature section, several noise mechanisms were described which cannot be easily resolved from the uncorrupted image. In the cases considered here, it was not possible to either observe the mechanism or measure the effect directly. We therefore use simulation to estimate the magnitude and nature of the spectral effects caused by the noise sources. It is, however, important to note the limitations of simulation as an investigative tool. We cannot prove causal links between a mechanism and an observed effect. Instead, we are limited to stating that the results predicted by the model are compatible with the observed results.

In order to make the experiment as relevant to texture analysis as possible, we will use a synthetic test surface with realistic parameters. A fractal surface will be illuminated with the empirical reflectance map. Noise simulations will then be carried out on this image and the measured and simulated spectra compared.

4.5.1 Quantization Noise

While most cameras/framestores are 8-bit, in practice, due to lighting variations and the field of view not consisting entirely of texture, it is likely only a fraction of this range will be utilised to convey useful texture information. To illustrate this point we show the standard deviations of the pixel intensities for the test surfaces illuminated from $\tau = 0°$ (*Table 4.5.1)*. In the third column, the number of bits required to quantise the range from -2σ to $+2\sigma$ is shown.

Texture	σ	Bits	Texture	σ	Bits
Isorock	9.48	6	Slate	6.19	5
Radial	6.61	5	Striated	12.20	6
Rock	11.37	6	Twins	8.45	6
Slab	13.10	6	Pitted	6.68	5

Table 4.5.1 Effective quantisations of data sets.

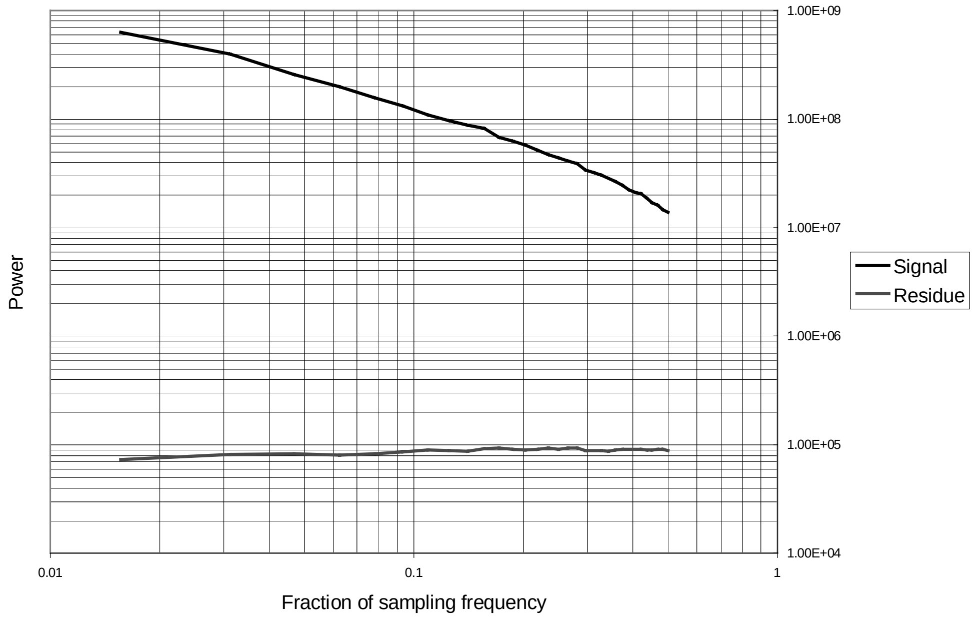


Figure 4.5.1 Power spectra of original signal and residue signal.

If we assume that the active region ranges from -2σ to +2σ, all of our images are effectively quantised in 5 or 6 bits. This was clearly an area of concern and the effect was simulated by comparing the original rendering with that of the image quantised to 6 bits. The spectrum of the residual image was found to be small and effectively white, *Figure 4.5.1*. This is consistent with Szirani who found that at least four bits were required for reliable classification.

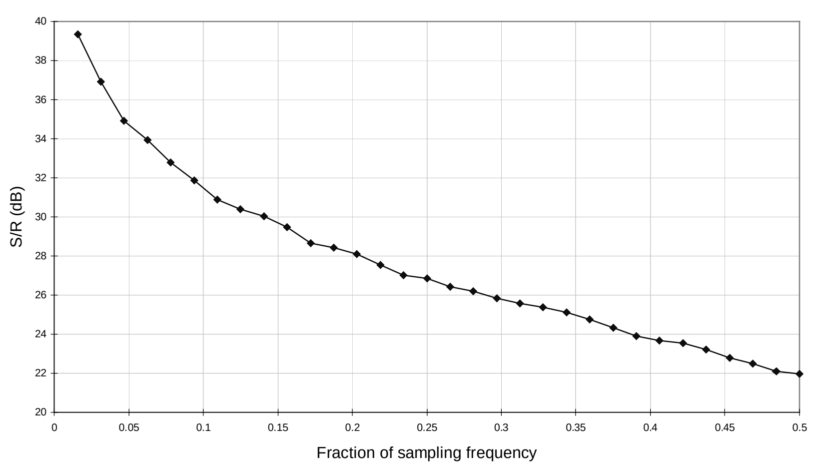
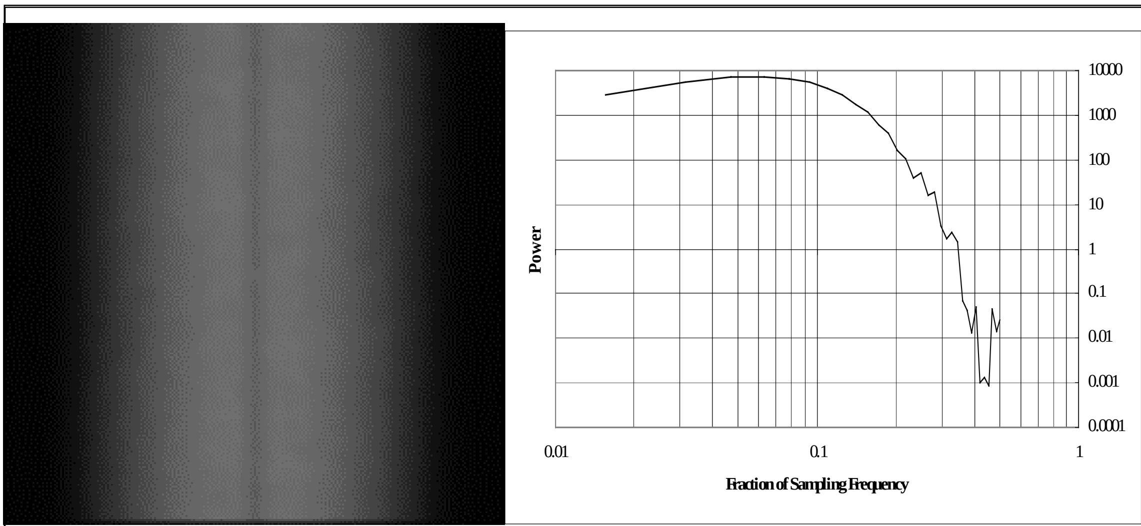


Figure 4.5.2 Signal to residue ratio vs. frequency

We therefore conclude that it is safe to assume that the residual spectrum due to quantisation is white. Furthermore in our simulations the S/R ratio varies from approximately 40dB at low frequencies to 22dB at high frequencies for a 6 bit quantisation.

4.5.2 The Effect of Jitter

We again use our experimentally based synthetic surface to generate a blurred image which is then supersampled (X10) in a raster scan with a random jitter drawn from a uniform distribution of width -0.2 to 0.2 pixels being added at the beginning of each line. The serial signal is then subsampled to its original rate and the power spectrum examined for artefacts. The resulting residual spectrum is shown in *Figure 4.5.3*.



(a) Linear frequency scale, log power scale. (b) Log frequency;log power scale.

Figure 4.5.3 Jitter induced 2D residual spectrum (a) and mean row spectrum (b).

The simulated effect is similar to the plateau effect observed for temporal noise prior to subsampling.

4.5.3 Serial Filtering

The image is read from the CCD array in discrete, serial form. This pulse train is then low pass filtered to yield a continuous waveform. A one-dimensional filter applied to this signal will produce an attenuation with frequency of the signal in the rows, while leaving the columns unaffected. We hypothesise that a white noise process filtered in this way will produce a spectrum with a characteristic vertically running plateau.

We test this hypothesis by simulation, using a Butterworth filter (arbitrarily designed to be 10th order) with breakpoint at 0.25 f_s, we filter a two dimensional white noise process in serial form. The resulting spectrum *(Figure 4.5.4)* is of similar form to that observed for the jitter simulation.

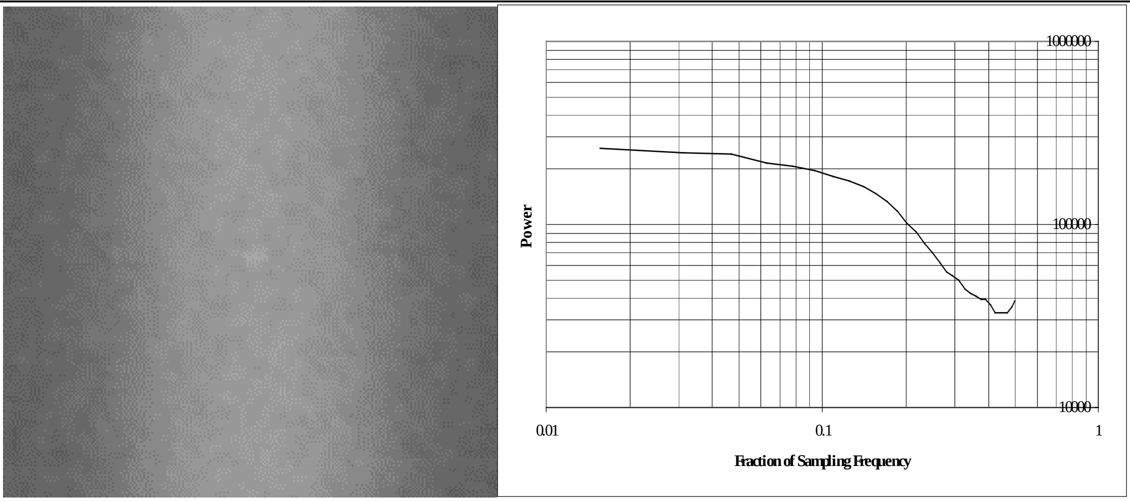


Figure 4.5.4 The spectral effects of an anti-aliasing filter.

Like the jitter simulation this produces a residue spectrum of the “plateau” form. The fact that two mechanisms may form rival hypotheses for a particular observed phenomenon serves to underline the inability of simulation to establish causal relationships. In any case, the results of the simulation indicate that the effect can be reduced to a white noise source if subsampling is employed.

4.5.4 Summary

In this section we found that the members of the data set were effectively quantised in 5 to 6 bits. The residual signal associated with quantisation was found to be white, with the result that quantisation noise is most significant at high frequencies. This section also disclosed that serial filtering of a white noise source and line jitter both give rise to the 'plateau' effect, and are possible sources of this plateau. However, as with the real data set after subsampling, these are effectively white noise sources.

4.6 Non-Quantifiable Noise Mechanisms

The aim of this chapter is to develop and parameterise a noise model. There are some noise mechanisms which are not readily measured either by direct experimentation, nor by simulation. Since we cannot measure these mechanisms in a useful way, we are unable to include them in a noise model. We do, however, briefly discuss two effects for completeness.

4.6.1 Camera Non-linearity

We conclude this section by briefly noting the effect of camera non-linearity. While many cameras may be set to give a linear response, the large range of intensity values, and the limited number of bits, mean that it is often more convenient to use a camera with γ=0.5[1]. *Figure 4.6.1* shows the spectra of a "Chessboard" calibration image. While the measured and synthetic spectra share the same spectral peaks, particularly at low frequencies, both the measured spectra have additional peaks due to non-linearity-induced harmonics.

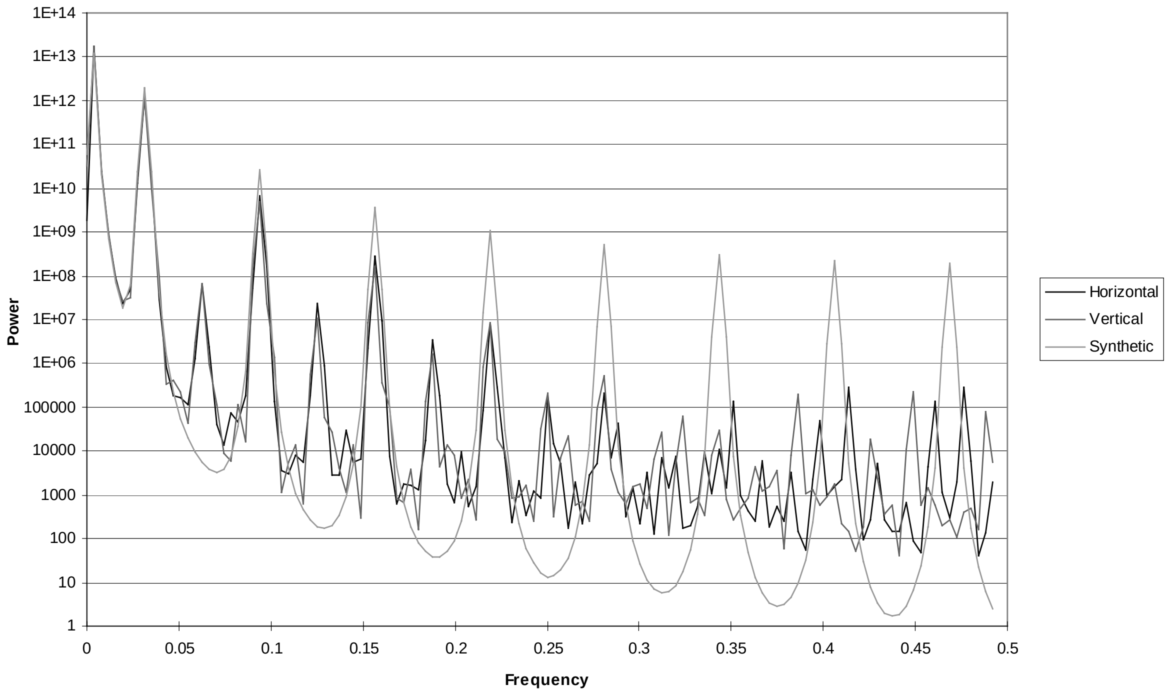


Figure 4.6.1 The generation of harmonics due to camera non-linearity.

In the previous chapter we measured the reflectance function in terms of camera response, effectively lumping together camera and reflectance non-linearity. We note that unlike reflectance, the effects of camera non-linearity are dependent on the incident intensity. As a consequence of this, the amplitude of the harmonics *relative to the fundamental* will change if the incident intensity is varied. A change in the intensity of illumination may therefore have serious effects on the classification of textures which cannot be corrected for using standard normalisation techniques.

[1] The camera amplifier can be modelled using the relationship $i_{out} = k.i_{in}^{\gamma}$,, where *k* is a constant. In linear operation γ=1.0, more commonly however γ=0.5 to allow quantisation of a large dynamic range in

We also note that a non-linear camera means that even the additive noise measured earlier will be signal dependent in effect. We justify our use of an additive model by measuring the effect of noise on the data set itself, we are in effect forming a linear approximation around the actual operating point. The uniformity of the temporal noise spectrum measured in *Figure 4.3.5* supports the validity of this assumption.

4.6.2 Fixed Pattern Noise

Associated with several of the noise sources considered in section 4.2 is the phenomenon of fixed pattern noise, that is an imaging effect which is present regardless of the image content. Given a camera without autogain and with a linear gain function, it is possible to measure fixed pattern noise by preventing light from entering the lens. However, we found both the autogain and the non-linear gamma function to be essential when imaging textures. Consequently, while it is possible to observe the effects of fixed pattern noise, it is not possible to predict its effect when imaging textures.

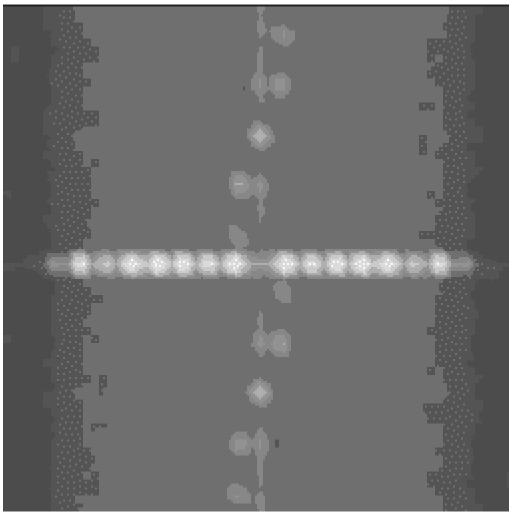


Figure 4.6.2 Fixed Pattern Noise Spectrum

We are therefore unable to quantify or model the effect of fixed pattern noise and while we do illustrate the spectral properties in *Figure 4.6.2* we do not include it in our noise model. Comparison with *Figure* ***4.3.4*** shows that several of the peaks observed in the FPN spectrum are also apparent in the imaged texture spectrum. We note however, that these are only obvious at high frequencies, and the effect will be reduced by subsampling.

---

a limited number of bits.

4.6.3 Summary

This section has briefly discussed those noise effects which we are unable to integrate into our model. Ideally it would be possible to use a linear camera without autogain. However, in practice the required dynamic range make this difficult to use with only eight quantisation bits. Due to non-linearity it is important that a constant level of incident intensity be maintained.

Fixed pattern noise is not in itself a problem to classification, though its non-linear interaction with the correct image may generate spurious spectral components.

4.7 An Imaging Model

Up until this point we have concentrated on describing and quantifying the noise and distortion associated with the imaging process. We now draw together the strands of this investigation to form a model of the imaging process. This model will serve as both an analytical tool and as the basis of simulations of the imaging process.

4.7.1 Integrating noise components

Two of the simulated effects, serial filtering and line jitter, are temporal in nature and their effect will be apparent in the temporal noise observed in section 4.5. The effect of quantisation, conversely, is static and is not accounted for in the temporal measurements. We now combine the temporal noise with quantisation noise to form an estimate of the error between the measured and actual values.

The pdf of a random process which is the sum of two random variables may be obtained by convolving the distributions of the two variables [Peebles, p.102]. We form expression (4.4.2) from the convolution of the uniform distribution with the postulated Gaussian distribution. Assuming a zero mean Gaussian process:

$$p_x(e) = \frac{1}{\sqrt{2\pi\sigma_t^{\,2}}} \int_{t=-0.5}^{0.5} e^{\left(\frac{(e-t)^2}{2\sigma_t^{\,2}}\right)} dt \qquad (4.4.2)$$

The parameter σ_t, controls the shape of the distribution, *Figure 4.7.1*. For temporal noise with a relatively large variance ($\sigma_t = 1$), the resulting distribution resembles a Gaussian curve, smaller variances ($\sigma_t = 0.1$) give rise to distributions closer to the uniform case.

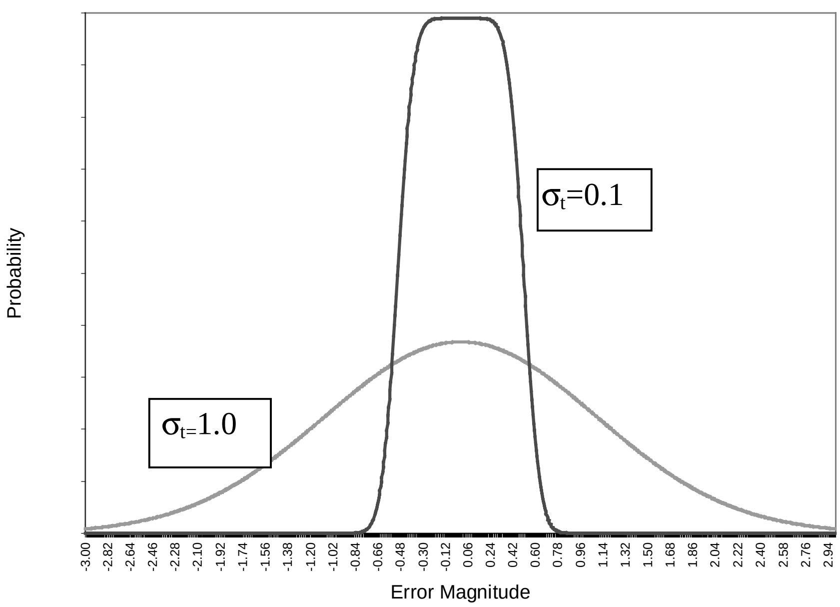


Figure 4.7.1 The effect of the sigma parameter on the PDF of a linear combination of Gaussian and uniform processes.

If we combine the estimated temporal noise distribution with the quantisation distribution (which exists in the interval ±1.2 due to rescaling due to pre-processing) we can form the distribution shown in *Figure 4.7.2.* We find that the resulting distribution can be modelled as being Gaussian of standard deviation (σ_n) 1.36.

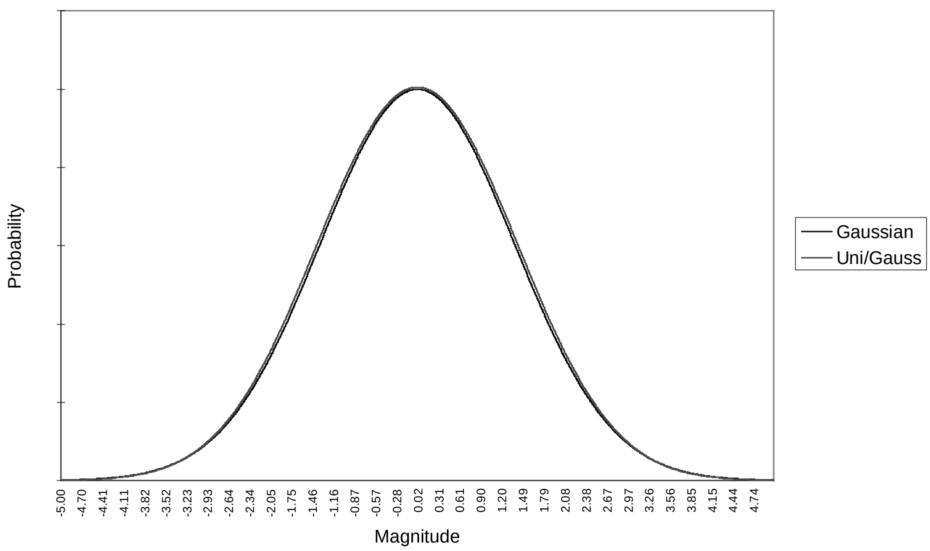


Figure 4.7.2 Estimated distribution of combined temporal and quantisation noise.

Since both the subsampled temporal and the quantisation residual spectra are considered to be white, the linear combination of the two will also be white.

4.7.2 Development of a Noise Engine

We now use the work carried out in this chapter to develop a noise engine which will be used later in this report. This engine forms a computational component in our system model.

The decision to use downsampled images allows several simplifications to be made to our model, such as the elimination of noise spikes at $0.25f_s$, the reduction of the relative effect of the blur function and the suppression of those directional noise mechanisms that only become apparent at high frequencies.

The model itself, *Figure 4.7.3*, consists of only two components, a blur function and an additive white noise process, both of which are considered to be global in character. The estimated blur function was found to fit well with the Gaussian model presented by Pentland. While localised distortions were experienced in the experimental work, these occurred in regions of the image which were not utilised. Their absence in our working data and the relative difficulty of their description lead us to omit them from our model.

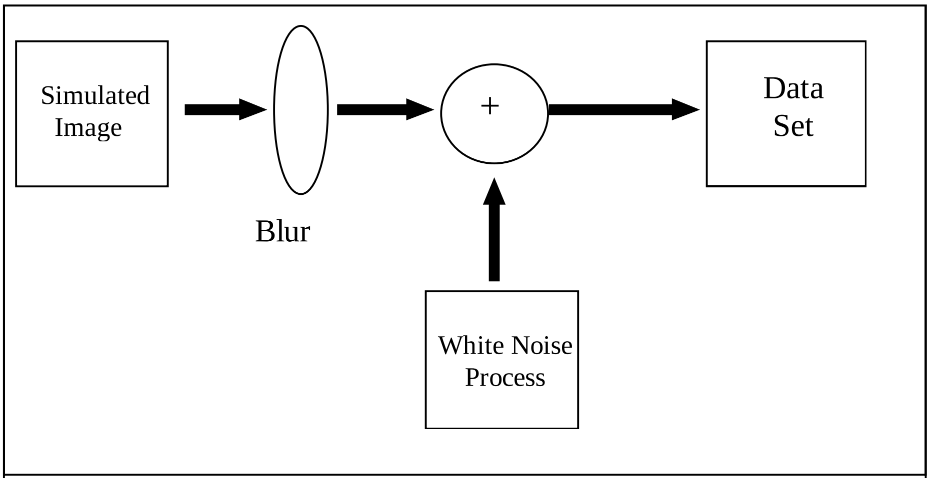


Figure 4.7.3 Noise model for sub-sampled images.

The second component of our noise model, additive white noise, lumps together several noise mechanisms. Dark noise, shot noise and quantisation noise are considered as being white in accordance with the literature, while jitter noise and filtered noise are also considered to be white as a consequence of subsampling.

Since we only consider texture which is caused by topography with no albedo variation, we will assume that camera non-linearity is lumped together with the reflectance function and modelled implicitly with the reflectance map.

Noise Mechanism	Description	Notes	Modelled
Dark noise	Temporal	Subsampled temporal noise found to be white.	Gaussian White
Shot Noise	Temporal	Subsampled temporal noise found to be white.	Gaussian White
Jitter	Temporal	Approximately white when subsampled	Gaussian White
1,2 Noise	Narrow band noise at 0.25 fs	Suppressed by lpf prior to sub-sampling.	Ignored
Quantisation		Simulation confirms accepted white noise assumption for broadband textures. While quantisation error is itself drawn from a uniform distribution, when lumped together with temporal noise of the magnitude observed here, the resulting distribution is Gaussian in character.	White noise.

Table 4.7.1 Summary of noise mechanisms and their effects.

The engine will have default parameters as follows: a Gaussian blur function with σ_b=0.02 and a noise function with standard deviation 1.36, quantisation will be carried out to 6 bits. The effect on a synthetic fractal image is shown in *Figure 4.7.4*.

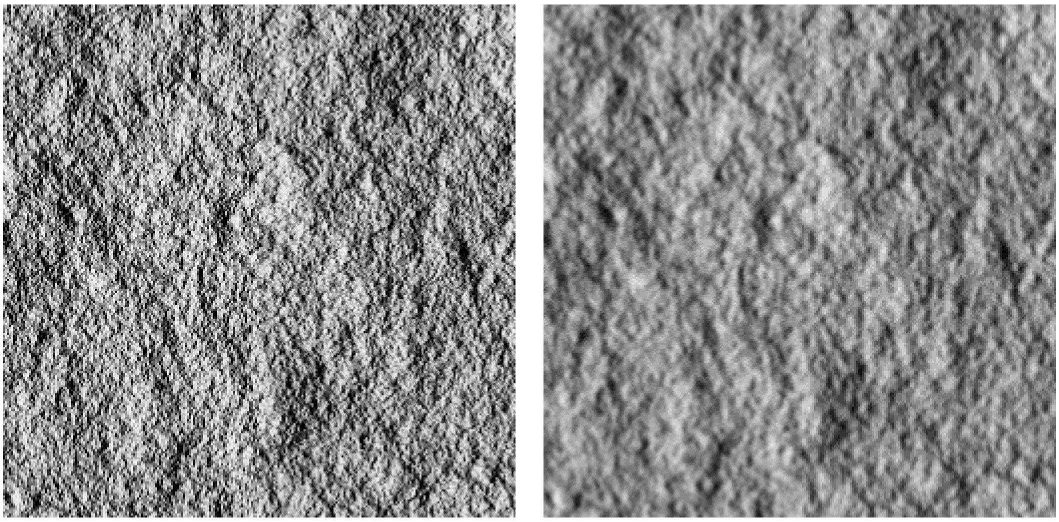


Figure 4.7.4 Original and processed images.

4.8 Assessment of the Physically Based Model of Texture Images

The work carried out in this thesis has so far consisted of modelling the formation of textured images from rough surfaces. We have now reached a stage in the imaging process at which we are able to compare the predictions made by our models with actual measurements. Comparison of measurement and prediction will give an indication of the accuracy of our models.

4.8.1 First Order Statistics

The theoretical work carried out in the previous chapters has been performed with the (non-essential) assumption that the data had, at all stages, a Gaussian distribution. This was justified on the basis of the height distribution of our surface models, and the linearity of the subsequent stages of the process. We have now reached the point where we experimentally evaluate this assumption in a direct way.

The standard test for testing whether a distribution conforms to a model is the Chi square test (used in the phase section). This will give the probability that the observed measurements are drawn from a particular distribution. If the sampled distribution really is drawn from the model distribution we would expect the measured pdf to converge to the models as the number of samples increase. Accordingly, the probability of a fit for *a given pdf* will decrease as the number of samples increases. Since in this application we use a very large number of samples the measured pdf must be almost exactly Gaussian.

In fact, the measured distributions are not Gaussian, but may be described as near-Gaussian. The large sample size means that the Chi square test will return a vanishingly small probability for distributions which are near-Gaussian. In this case we must rephrase the question, instead of asking *is this distribution Gaussian*? we must ask *how close is the observed distribution to the Gaussian form?* We therefore seek a figure of merit which operates on the pdf. and gives a smoother transition as the actual pdf. moves away from the model.

The requirement for a smooth transition is shared by criteria used in parameter estimation. With this in mind we define a modified form of the L_1-norm [Siva, p.67]. We define the figure in equation 4.9.

$$L_1 = \sum_{i=0}^{N} \left| p_o(i) - p_g(i) \right|$$

where p_o is the observed probability density function, and p_g is the pdf of a Gaussian distribution with the same mean and standard deviation.

In order to relate the figure to a meaningful form, we show the pdf of the textures with the lowest and highest measures, (*Figure 4.8.1*) before tabulating the criterion in *Table 4.8.1*.

In all cases the Gaussian model is parameterised by mean and standard deviation measured in the usual way, with no parameter optimisation.

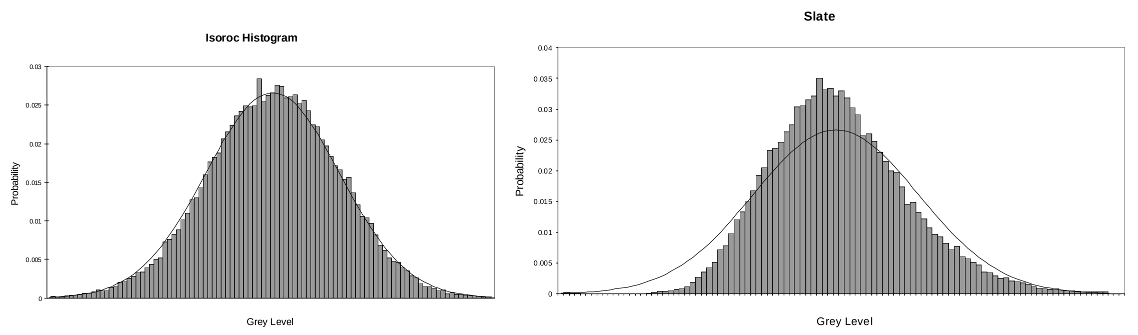


Figure 4.8.1 Histogram distributions of grey levels for most and least Gaussian distributions.

Texture	**L_1 Statistic**	**Texture**	**L_1 Statistic**
Isoroc	0.039	Pitted	0.095
Rock	0.077	Radial	0.173
Slab	0.053	Slate	0.173
Striate	0.061	Twins	0.099

Table 4.8.1 Degree of 'normality' in grey level distribution.

The distributions range from the very Gaussian-like *Isoroc* to the positively skewed, *Slate* and *Radial* textures. However, even the *Slate* and *Radial* textures are broadly Gaussian and we shall retain the assumption of normality in the theoretical models used in this work.

4.8.2 Second Order Statistics

We have described the data set as being the output of a linear system with input conforming to one of the surface models described in Chapter 2.

$$I(\omega,\theta) = S(\omega,\theta).\,\omega.\,k.\cos^2(\theta-\tau).\,B(\omega,\theta)+W$$

We plot the power spectra of the images of the test surfaces illuminated at τ 0° in *Figure 4.8.2*.

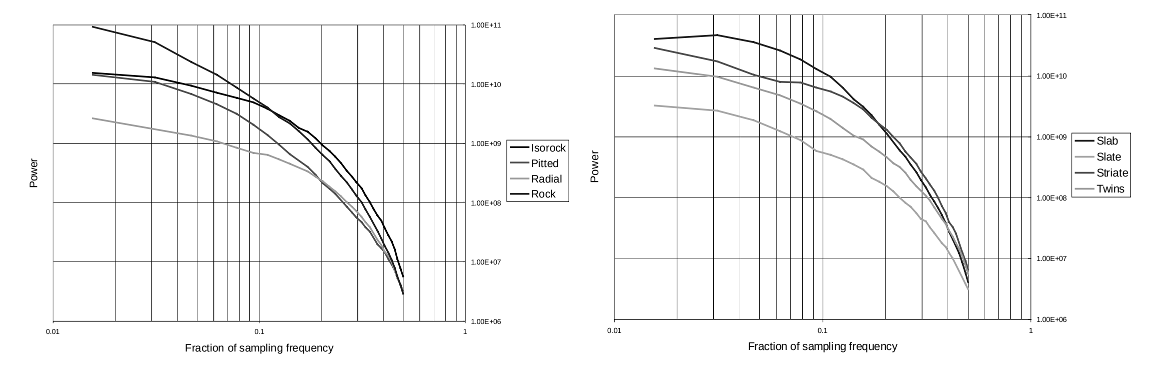


Figure 4.8.2 Power spectra of test textures illuminated at τ=0°.

The first point we note from *Figure 4.8.2* is that the power spectra are no longer fractal, i.e. they do not have a linear roll-off when plotted on a log-log graph. We believe this to be primarily due to the Gaussian transfer function, associated with blur acting as a low pass filter. In *Figure 4.8.3* we plot the image power spectra against that of the fractal surface rendered and 'imaged' using our models. The spectra are not identical, however, they are in reasonable agreement and this result does give some verification of the models used in this thesis.

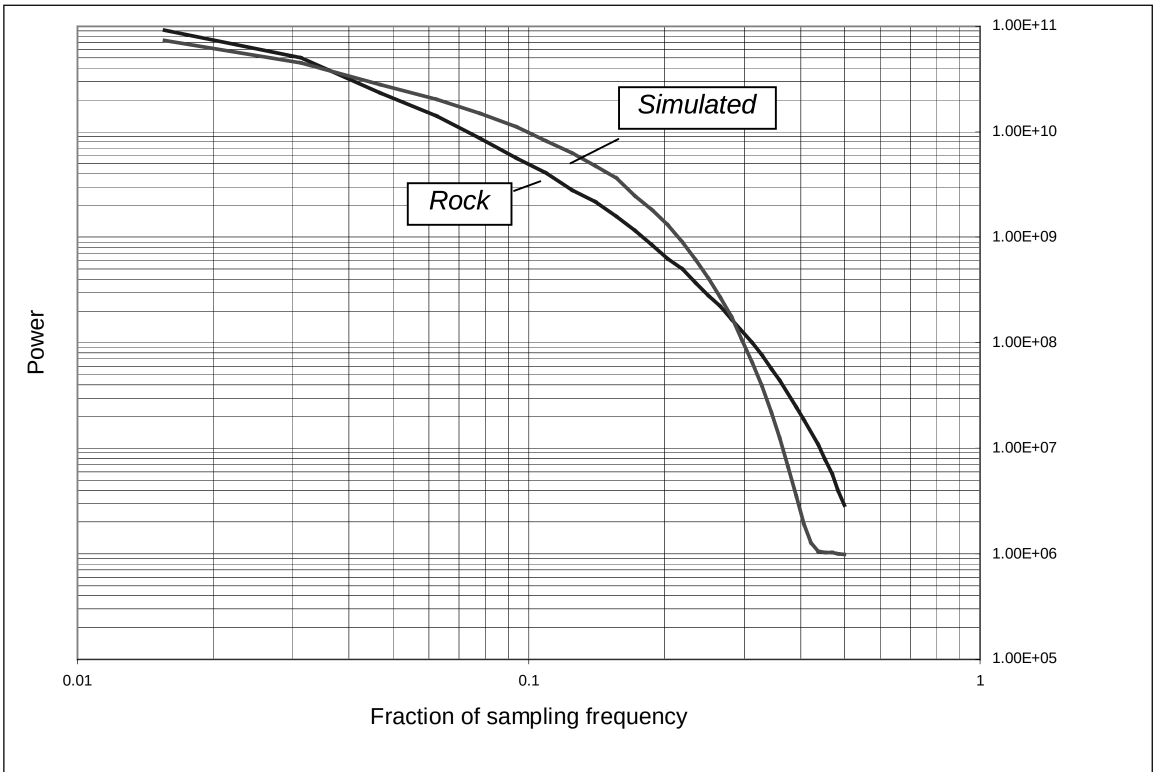


Figure 4.8.3 Comparison of Rock spectra with the blurred fractal spectra.

The result discussed above while analytical, does not give a very intuitive idea of the accuracy of our models. In *Figure 4.8.4* we show the original *Rock* image as well as a second semi-synthetic image. The second image was obtained by imposing the predicted power spectrum on the *Rock* image while retaining the original's phase spectrum.

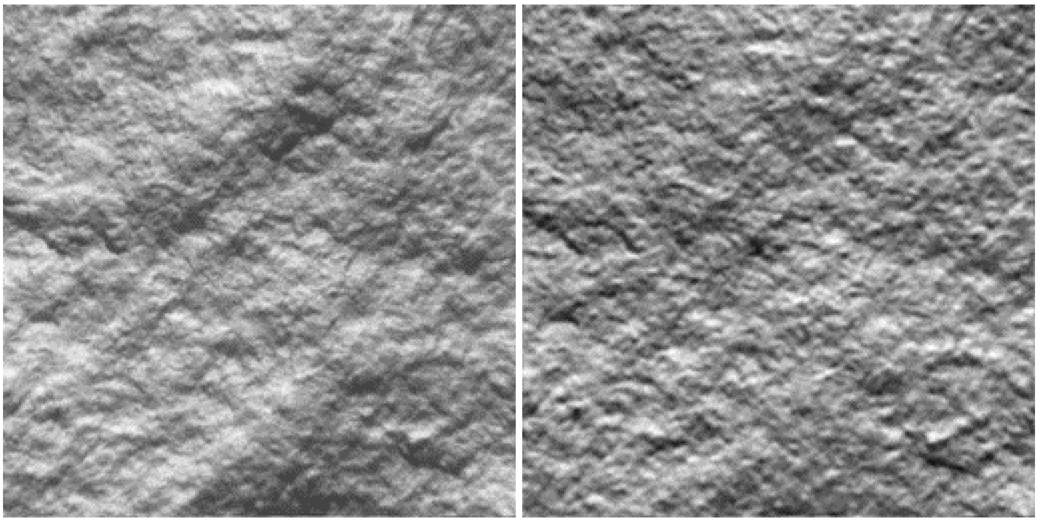


Figure 4.8.4 Comparison of original rock surface with modified PSD reconstruction.

4.9 Conclusions

Summary

In this chapter the case for down-sampling the test image has been demonstrated on grounds of noise reduction and the advantages of subsampling to classification, rotation invariant classification and image modelling have been highlighted.

The chapter has consulted both the literature and experimental data to model the transition of the signal from incident image to data set. The form and parameter values of the model have been estimated for the rock surfaces used in the work described in this thesis.

Blurring was found to be the dominant artefact of imaging, and a Gaussian model of the imaging system transfer function was adopted. We have shown that the common assumption of a white noise model is not justified, in our case at least, without down-sampling.

We have established signal to noise ratios for the textures considered in this thesis, giving a measure of the quality of these images. Furthermore, we have measured the S/N ratio as a function of sampling frequency, allowing the effects on spectrally-definable classification feature measures to be considered.

We have also compared the statistics of the imaged textures with those predicted by our models. We found that the first order statistics while not *exactly* Gaussian could *reasonably* be described as such. We also found that the spectra were not fractal, though the measured spectra were reasonably approximated with our surface, rendering and imaging models.

Implications for Texture Analysis

The fractal nature of the signal and the observed effects of blur and white noise discourage the use of high frequencies as consistent sources of discriminatory information. While the blur function reduces the feature mean and variance, particularly at high frequencies, the additive white noise increases the texture's variance in feature space. The magnitude of the degradation in classification will depend on the proximity of the textures in feature space as well as the level of the noise itself.

While the absolute power of noise seems to be similar in all images, the power relative to the texture varies depending on which texture is used. Consequently, we cannot apply Wiener filtering to the image prior to classification. Use of the related technique of spectral subtraction is possible, albeit with the introduction of serious distortions. In line with the philosophy of this thesis, processing beyond sub-sampling will be avoided.

The directionality of noise and imaging effects for the original image makes the unprocessed image unsuitable for rotation invariant algorithms. Fortunately, these directional effects can largely be suppressed by subsampling the image and we recommend the operation prior to the use of a rotaion-invariant algorithm.

The findings of this chapter are relevant to the remainder of this thesis. The results affect the design of the classifier in chapter 5 and its modelling in chapter 6. The most significant contribution of this chapter is, however, to chapter 8, where an experimental algorithm is evaluated using the simulation engine developed in this chapter.

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* $\mathbf{F(x,y)}$. 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.

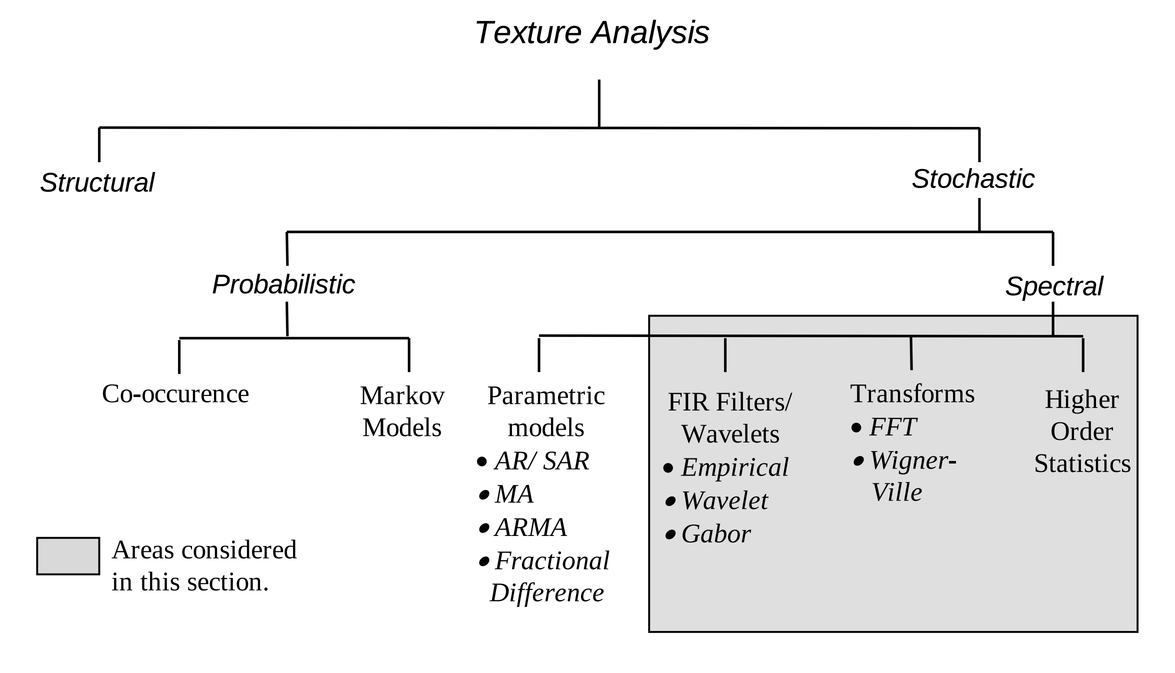


Figure 5.3.1 An (incomplete) taxonomy of texture analysis techniques.

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:

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

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).

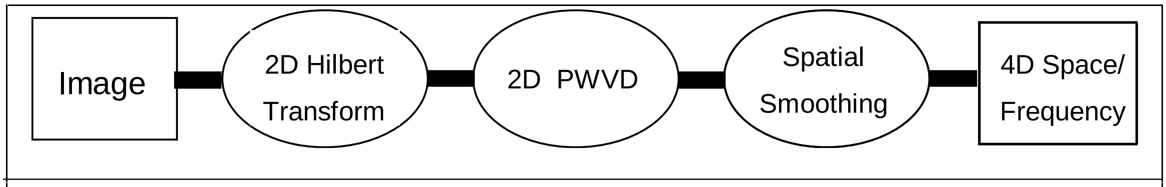


Figure 5.3.2 Zhu's scheme for local spectra estimation

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.

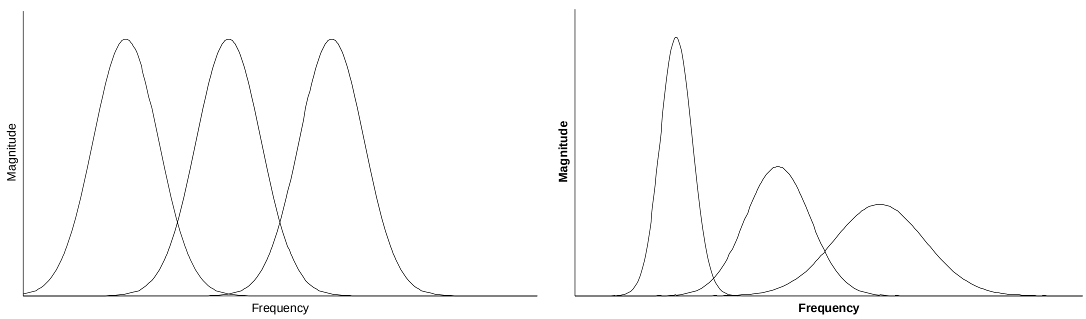
Explicitly evaluated wavelets

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

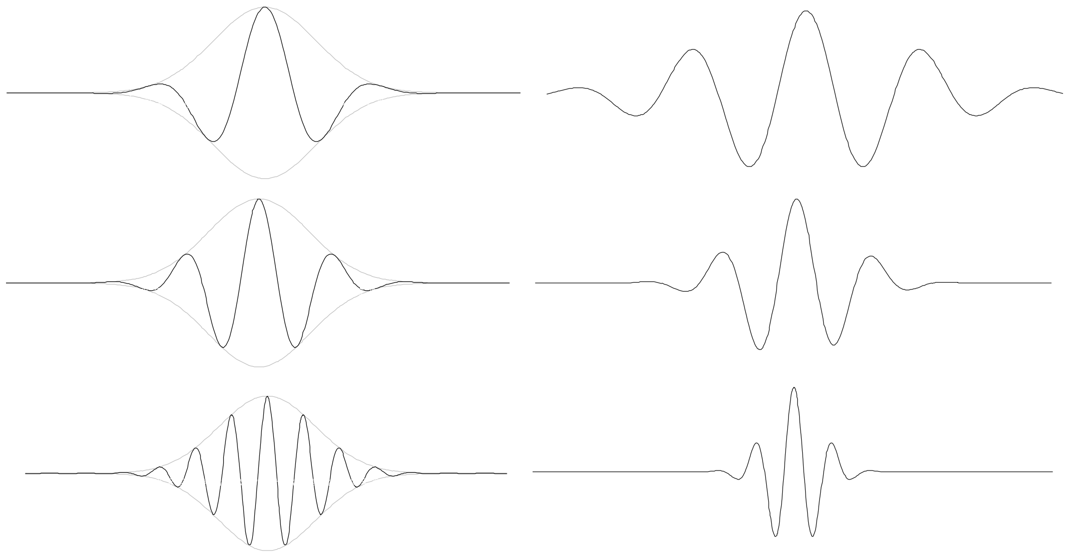
(2.) sampling with elementary functions each possessing a bandwidth which is constant *relative* to the function's central frequency, *Figure 5.3.3b.*



(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.*

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

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.

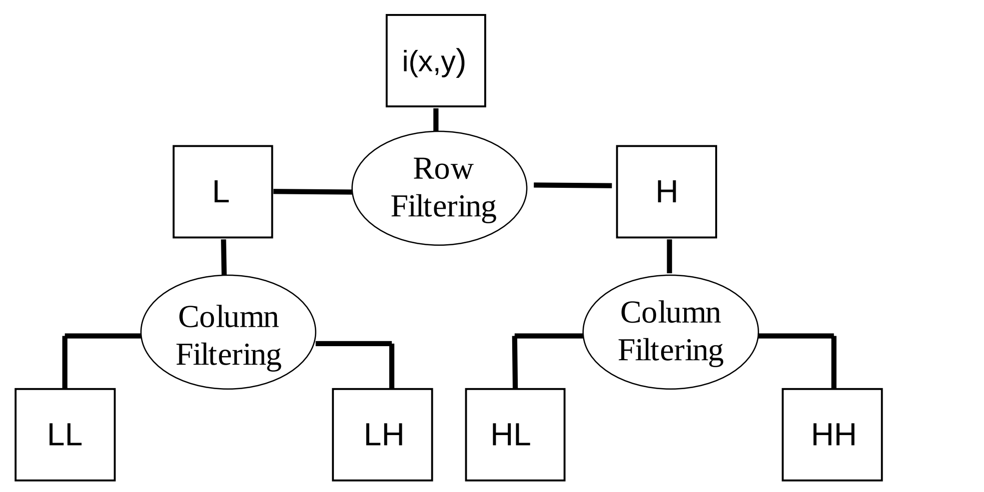


Figure 5.3.5 Two dimensional wavelet implementation.

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*.

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

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

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.

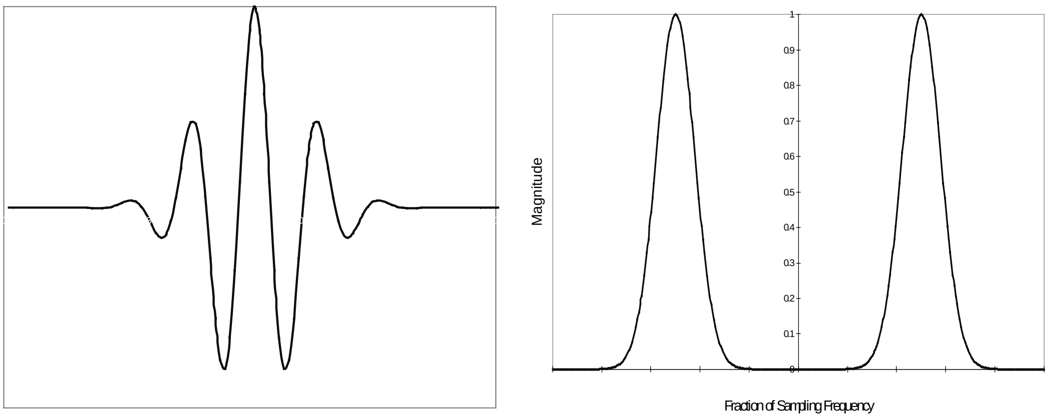


Figure 5.3.6 Real spatial part and frequency domain representation of Gabor function.

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

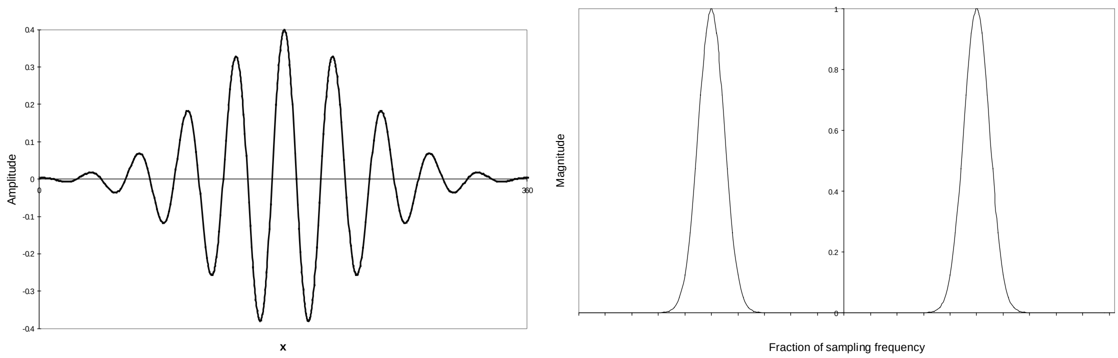
By linking envelope extent to the sinusoid's wavelength, we may define a series of functions which are scaled versions of each other. A family of Gabor functions may be described as a series of dilated forms of a single prototype, and a transform may be defined in terms of equation 5.3.5c [Lee96].

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

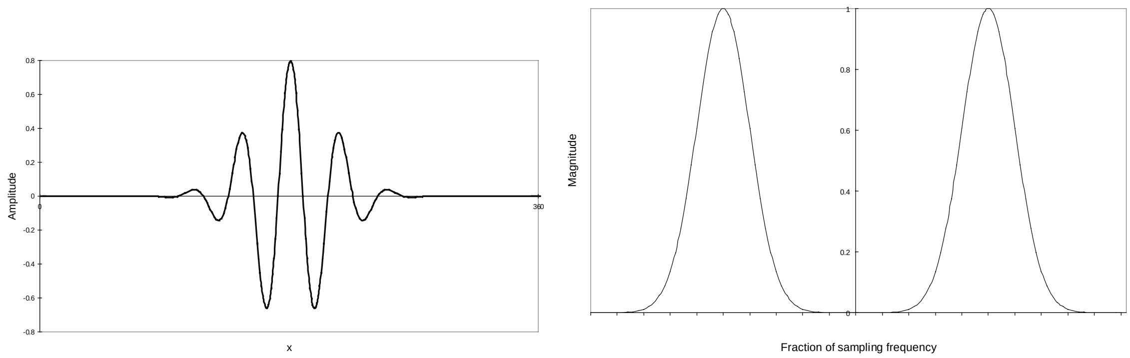
where x_0 and y_0 are the spatial co-ordinates of the filter at a given time,

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.



(a.) *(b.)*



(c.) *(d.)*

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.

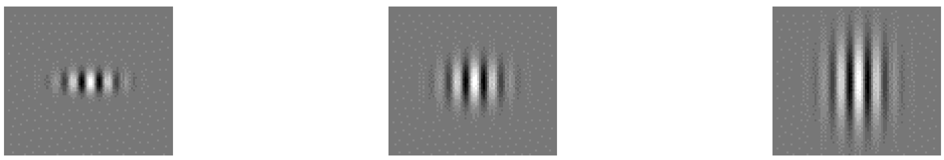
σ_x is the standard deviation of the Gaussian envelope in the direction of the x axis in the spatial domain.

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*.



(a.) σ_y=3.54 *(b)σ_y=7.07* *(c)σ_y=14.14*

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*.

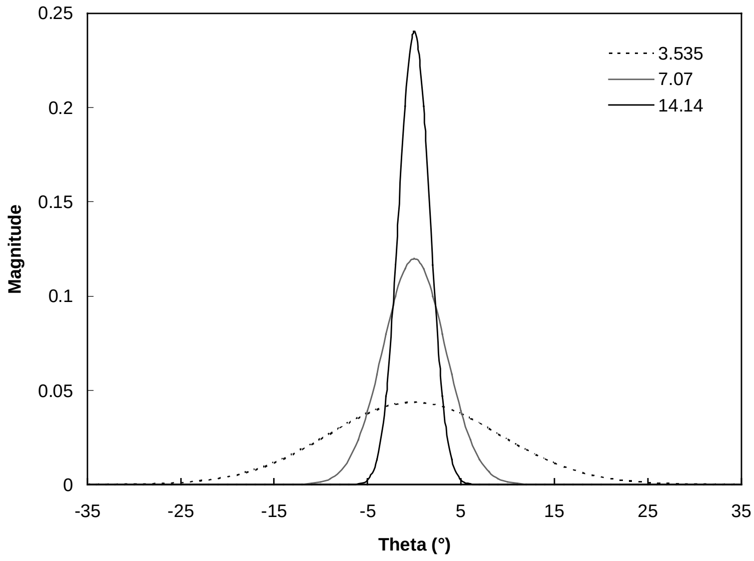


Figure 5.3.9 Polar plot of Gabor filters angular magnitude responses for various σ_y.

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.

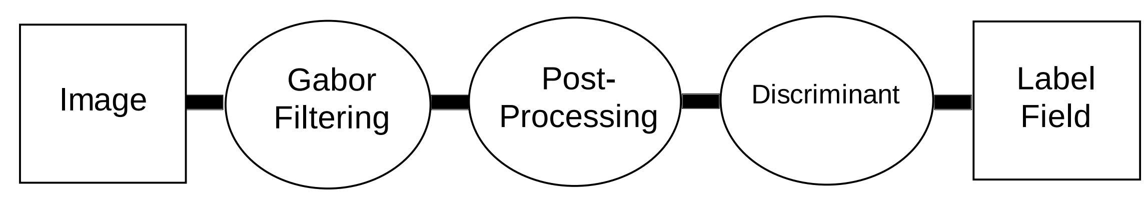


Figure 5.4.1 The Gabor-based Classifier

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

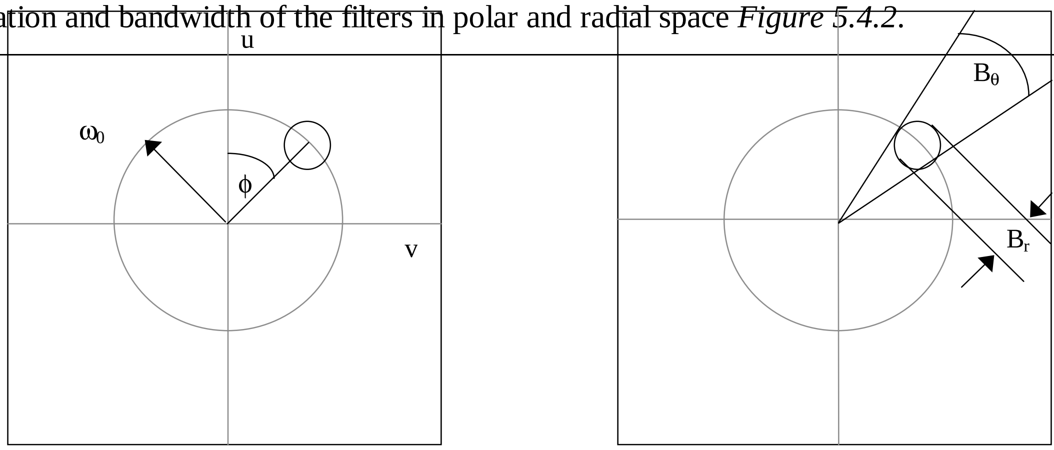
Design vs. Systematic Structure

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:

$$\mathbf{F}\omega_0\,\mathbf{d}\phi$$

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

and ϕ is the filter orientation in degrees.

Polar Properties

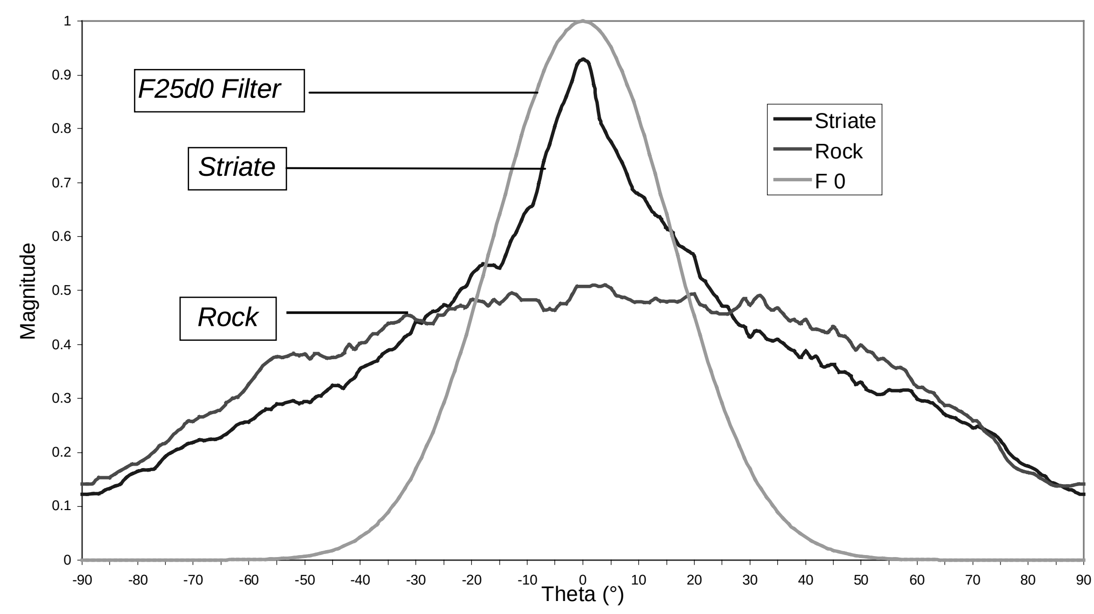


Figure 5.4.3 Polar Spectra of Rock and Striate Surfaces illuminated from Tau 0° and a Gabor filter orientated at Tau 0°.

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.

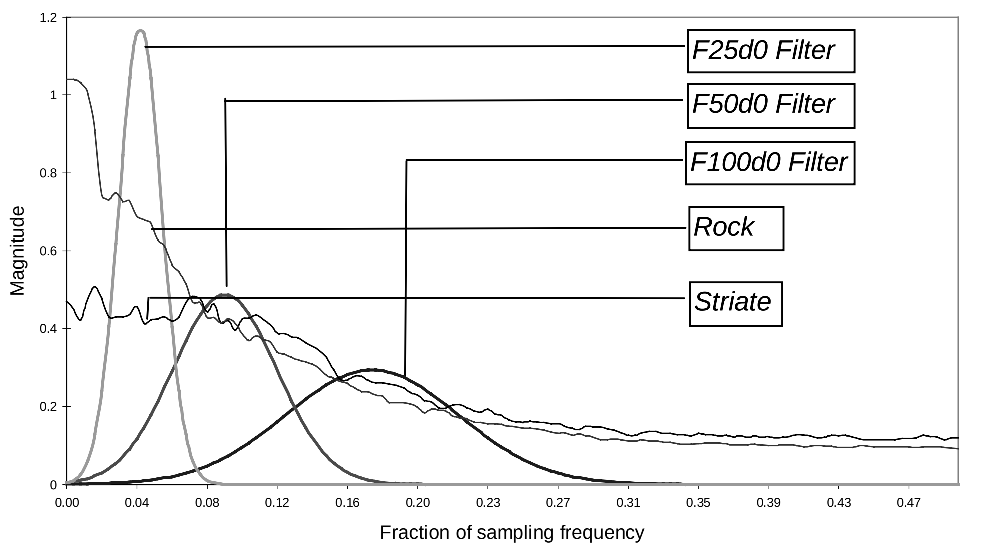


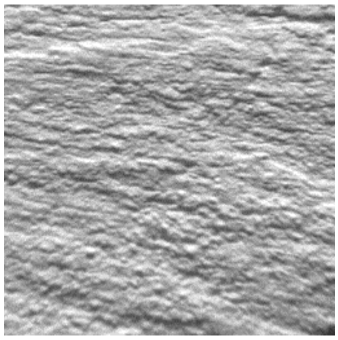
Figure 5.4.4 Radial Responses of Filters with Texture Spectra

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*.

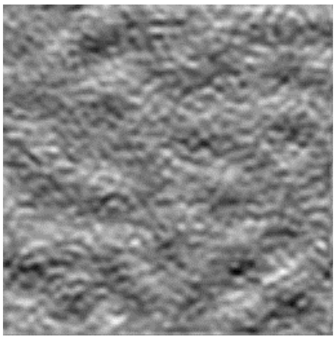
Spectral Support

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.

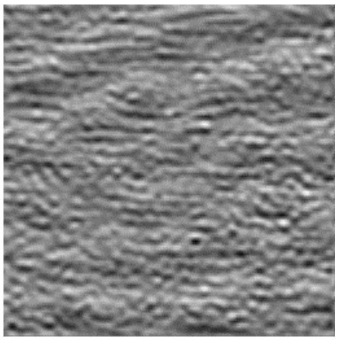
(a.) Original Rock image.



(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

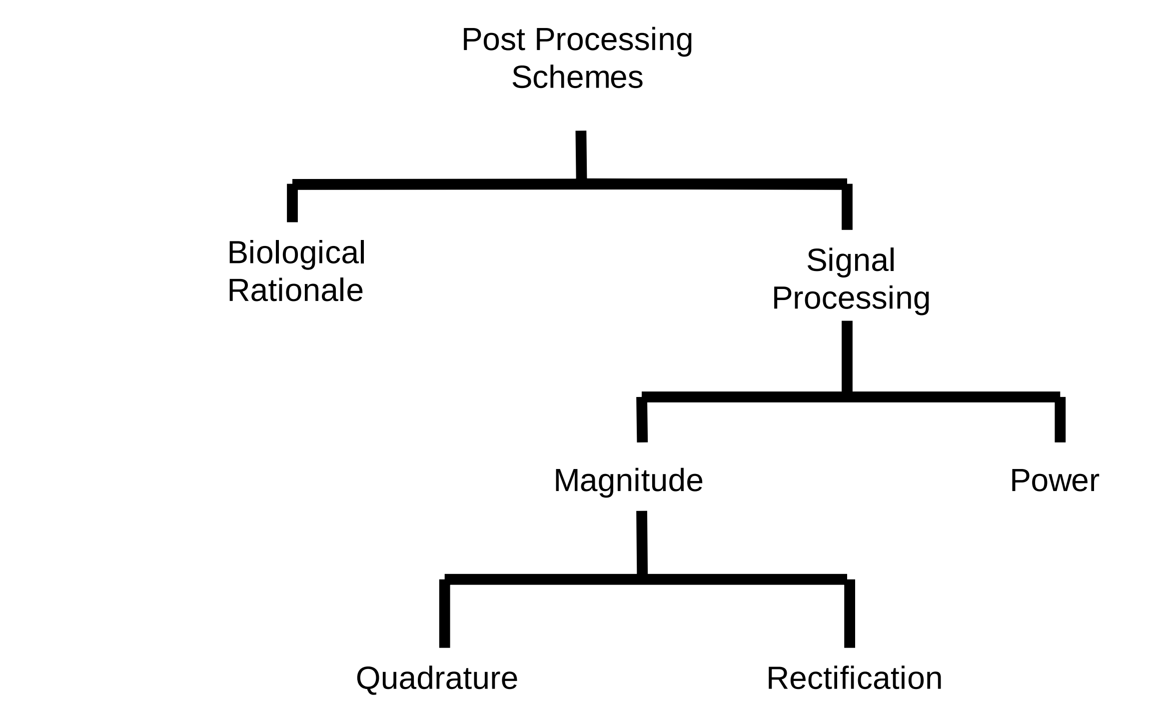


Figure 5.4.6 Taxonomy of post-processing techniques.

(*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*.

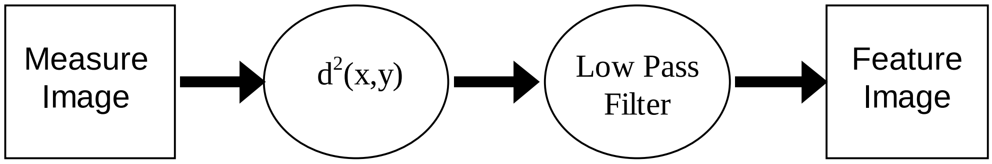


Figure 5.4.7 General form of feature based on channel power measurement.

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.

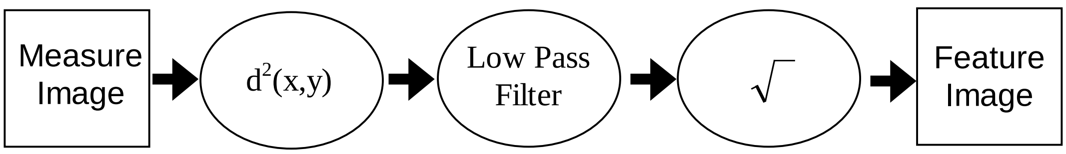


Figure 5.4.8 Unser's scheme.

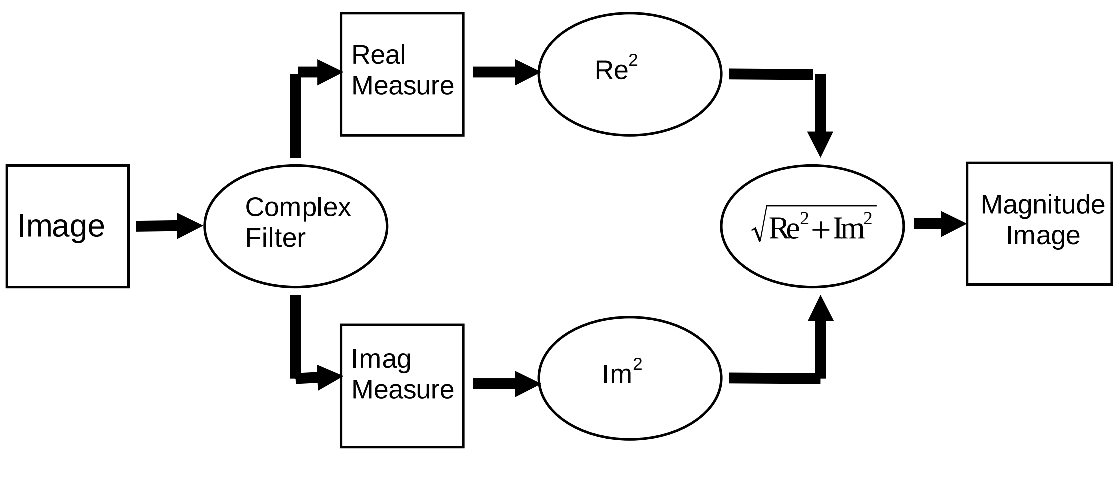


Figure 5.4.9 Magnitude estimation using quadrature filtering.

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].

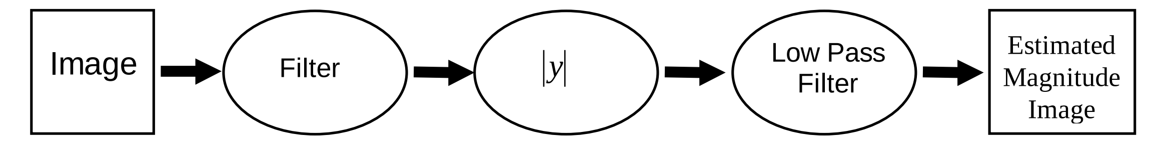


Figure 5.4.10 Magnitude estimation using ABS approximation.

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.

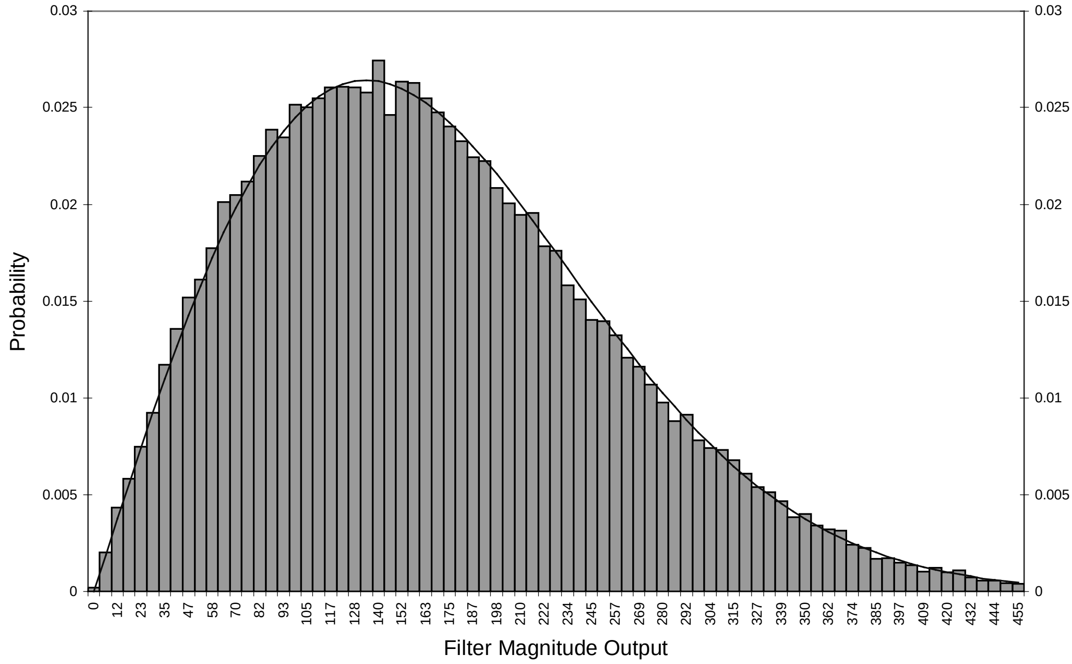


Figure 5.4.11 Pdf of modulus of quadrature filter outputs.

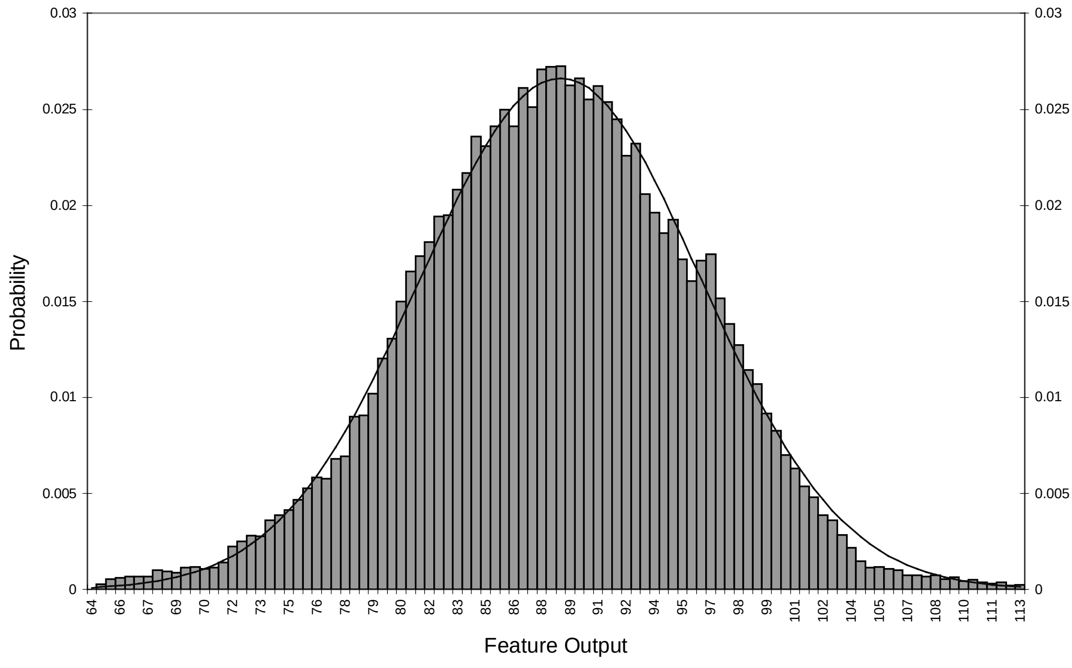


Figure 5.4.12 PDF of Feature Image

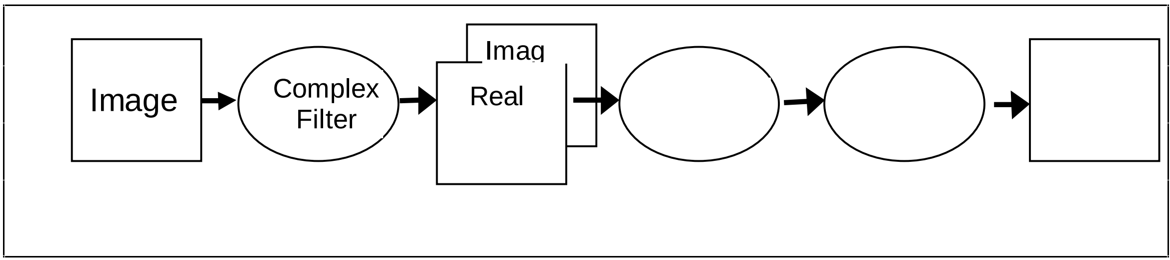


Figure 5.4.13 Post Processing

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

Chapter 6

Modelling The Classifier Tilt Response

6.1 Introduction

This thesis can be broken into two parts: the first describes the development of a model for rough surface classification, the second part develops techniques to classify surfaces invariant to illuminant tilt. This chapter represents the transition between the analytical and the problem solving phases of this thesis. It has two aims: to model the effect of tilt on the classifier, and to observe the effect of tilt variation on classification accuracy. We also model the classifier developed in Chapter 5; and are therefore concerned with the final stage of the model that describes the process of classification from imaged data set to symbolic representation. We will also measure the effect of tilt variation on classification accuracy. This can therefore be seen as defining the problem that the second part of this thesis is designed to confront.

This chapter consists of four sections. In the first, the tilt response of the texture features will be predicted using theory. The predicted relationships will then be tested using both simulation and experimentation in the second section. The discriminant part of the classifier cannot be modelled in general terms since it depends on the second order interactions between members of the sample set. The next section is designed to extend our analysis and understanding of the tilt effect at the discriminant stage in the classification. The third section of this chapter is designed to give an intuitive understanding of the effect on classification. Using synthetic textures it is possible to obtain an adequate classification using only two features. By using such a small feature set, the discrimination process can be observed on a two dimensional scatter plot—allowing the reader to observe the tilt induced movement of feature clusters across discriminant boundaries with the associated rise in misclassification. Finally, the full classifier is applied to three montages of real textures and the accuracy of the classifier observed as tilt is varied. This chapter will therefore use theory, simulation and experimentation to model the relationship of features with tilt; and use experiments to observe the effect on classification accuracy.

6.2 Modelling The Feature/Tilt Response

In Chapter 5, a classifier was developed. We now set about the formation of an analytical model which encompasses both the classifier and the illumination model of Chapter 3. This will enable us to make predictions as to the effect of illuminant tilt on surface classification.

Chantler made the analogy between the tilt effect and a linear filter [Chantler94b]. We believe this to be a useful model, and combine it with the Gabor filters to form a single linear stage (highlighted in *Figure 6.2.1*), bridging the gap between surface and measure image. We will model the tilt response of the combined filter (sub-section 6.2.1), the power spectrum of the resulting measure images (sub-section 6.2.2), and the first order statistics of the feature images (sub-section 6.2.3).

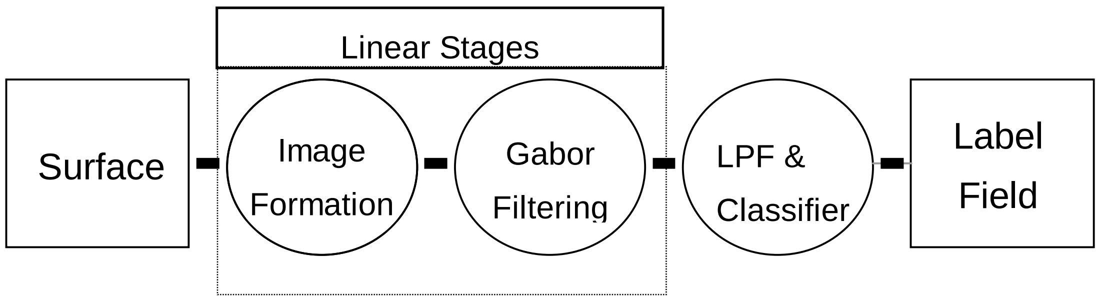


Figure 6.2.1 The Signal to Symbol Chain

6.2.1 Combined Filter Tilt Response

The transfer function relating surface spectra to image spectra may be described as a filter with magnitude response:

$$R(\omega,\theta \mid \tau) = \omega|\cos(\theta-\tau)|k|B(\omega)| \qquad (6.2.1a)$$

where $B(\omega)$ is the transfer function of the imaging device, The magnitude response of a Gabor filter oriented in direction ϕ is

$$G\left(u,v\middle|u_o,v_o,\sigma_x,\sigma_y\right) = \exp\left(-0.5\left[\frac{(u-u_o)^2}{\sigma_x{}^2}+\frac{(v-v_0)^2}{\sigma_y{}^2}\right]\right)$$

$$+\exp\left(-0.5\left[\frac{(u+u_o)^2}{\sigma_x{}^2}+\frac{(v+v_0)^2}{\sigma_y{}^2}\right]\right)$$

$$\text{u} = \omega\cos\theta \qquad u_0 = \omega_0\cos\phi$$

$$v = \omega\sin\theta \qquad v_0 = \omega_0\sin\phi$$

In polar form,

$$G(\omega,\theta|\omega_0,\phi,\sigma_u,\sigma_v) = \exp\left(-0.5\left[\frac{(\omega\cos\theta-\omega_0\sin\phi)^2}{\sigma_u{}^2}+\frac{(\omega\sin\theta-\omega_0\cos\phi)^2}{\sigma_v{}^2}\right]\right)$$

$$+\exp\left(-0.5\left[\frac{(\omega\cos\theta+\omega_0\cos\phi)^2}{\sigma_u{}^2}+\frac{(\omega_0\sin\theta+\omega_0\sin\phi)^2}{\sigma\,{}_v^2}\right]\right) \quad (6.2.1b)$$

If we combine the rendering response Eq. 6.2.1a with the response of the Gabor filter Eq. 6.2.1b, we obtain the magnitude response of the combined Gabor and rendering filter:

$$Z(\omega,\theta,\tau|\omega_0,\phi,\sigma_u,\sigma_v) = \omega k|\cos(\theta-\tau)|.B(\omega).G(\omega,\theta)+W.G(\omega,\theta) \qquad (6.2.1c)$$

We now numerically integrate equation 6.2.1c over radial frequency, ω, to plot a series of angular response curves in theta for various values of τ, with ω_0=0.125ω_s and ϕ=0°.

$$\int_{\omega=0}^{0.5\omega_s} Z(\omega,\theta,\tau)d\omega \qquad (6.2.1d)$$

where ω_s is the sampling frequency.

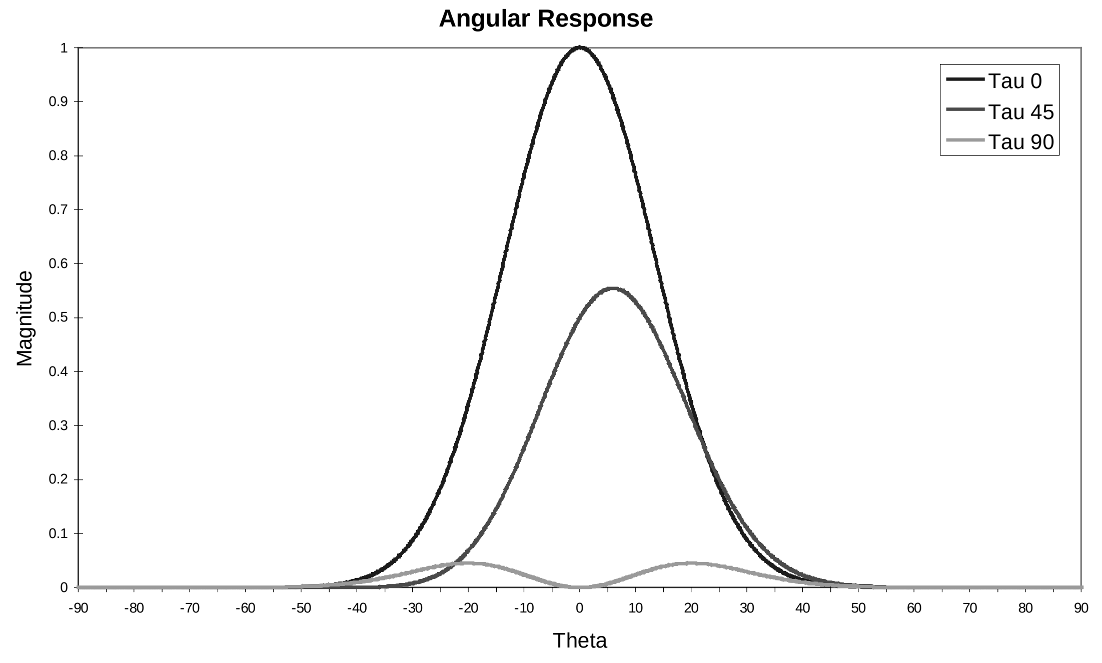


Figure 6.2.2 Angular Transfer Function of Combined Illuminant/Gabor Model

We note three points from this experiment:

1. Firstly, the response of the combined filter is attenuated as the tilt angle τ increases from 0° to 90°. Consequently, the variance of the measure image associated with the filter will fall as the tilt angle is varied through this range.

2. Secondly, the direction of peak sensitivity is shifted in the τ =45° and 90° cases. For an isotropic surface, the directionality of the measure image may no longer be aligned with the directionality of the filter. In the case of a directional surface, the directionality of the measure image will be a function of the surface, illuminant and filter directional properties.

3. Thirdly, even at τ=90°, the filter still gives a significant response due to the bandwidth of the filter orientation. Since we are using a linear model, there is no response at θ=0°, however, there is still a response in other directions due to the two sidelobes present. One consequence of this is that the variance of the measure image for an isotropic image will always be greater than zero. Even with a linear rendering model and in the absence of noise, the feature/tilt response will always be greater than zero.

6.2.2 Modelling The Measure Images

In Chapters 3 and 4, we developed an expression for the power spectral density (PSD) of the data set,

$$I(\omega,\theta \mid \tau) = \omega^2 \cos^2(\theta - \tau)\sin^2 \sigma . S(\omega,\theta).|B(\omega)|^2 + W \qquad (6.2.2a)$$

where $S(\omega,\theta)$ is the surface power spectrum.

Each Gabor filter produces a real and an imaginary output image. The PSD of the real and imaginary measure images will be identical, and equal to:

$$D(\omega,\theta | \tau,\omega_0,\phi,\sigma_u,\sigma_v) = I(\omega,\theta).|G(\omega,\theta)|^2$$

Expanding the image term:

$$D(\omega,\theta,\tau | \omega_0,\phi,\sigma_u,\sigma_v) = \omega^2 k^2 \cos^2(\theta - \tau) S(\omega,\theta).|B(\omega)|^2.|G(\omega,\theta)|^2 + N.|G(\omega,\theta)|^2 \qquad (6.2.2b)$$

Since the operation is linear, and the original image is assumed to be Gaussian, we may assume the measure images will also have a Gaussian distribution. Integrating (6.2.2b) over both frequency and angle, we will obtain a quantity, which we will call *signal energy*, σ_m^2 , equivalent to the variance of the measure images.

$$\sigma_m^2(\tau) = \int_{\theta=0}^{2\pi} \int_{\omega=0}^{fs} D(\omega,\theta,\tau) d\omega d\theta \qquad (6.2.2c)$$

If we integrate a two dimensional Gabor function, with the parameters used in this thesis, over radial frequency, we obtain a polar plot that is approximately Gaussian with a peak at $\theta=\phi$, *Figure 6.2.3*. We can use this approximation to develop an analytical expression for σ_m^2 by integrating the product of the Gaussian polar response of the Gabor function and the sinusoidal tilt response Eq. 6.2.2d. This is equivalent to the variance of a measure image calculated from the image of an isotropic 1/f surface in the absence of imaging artefacts.

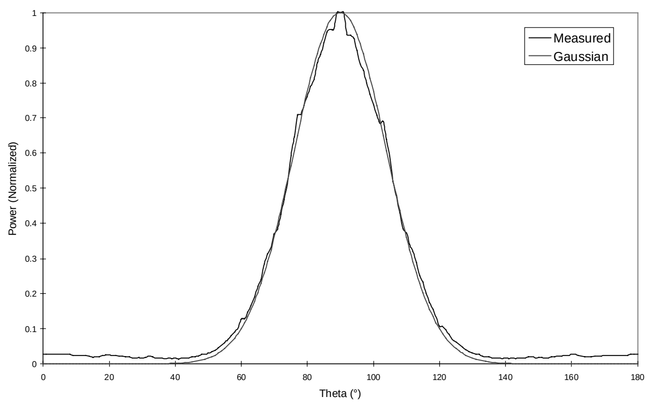


Figure 6.2.3 Comparison of polar response (power) of F25d90 filter with Gaussian function.

$$\sigma_m^2(\theta,\tau,\phi) = \int_{\theta=\phi-\pi}^{\phi+\pi} \cos^2(\theta-\tau)\exp(-0.5\left(\frac{\theta-\phi}{\sigma_p}\right)^2)d\theta \qquad (6.2.2d)$$

For the filters used in this thesis σ_p is less than 20° and the value of the integral is negligable beyond the limits used in 6.2.2d. It is therefore permissable to redefine the range of integration as being from θ=-∞ to ∞.

If we make the substitution $\lambda = \frac{\theta-\phi}{\sqrt{2}\sigma_p}$ then:

$$\theta = \sqrt{2}\sigma_p\lambda + \phi \qquad \frac{d\lambda}{d\theta} = \frac{1}{\sqrt{2}\sigma_p}$$

when θ=-∞ then λ=-∞ and when θ=-∞ then λ=∞

Therefore when integrating between θ=-∞ to ∞, the equivalent limits for integration are λ=-∞ to ∞.

$$\sigma_m^2(\theta,\tau,\phi) = \int_{\lambda=-\infty}^{\infty} \sqrt{2}\sigma_p \cos^2\left(\sqrt{2}\sigma_p\lambda + \phi - \tau\right)\exp\left(-\lambda^2\right)d\lambda$$

$$\sigma_m^2(\theta,\tau,\phi) = \frac{\sigma_p}{\sqrt{2}} \int_{\lambda=-\infty}^{\infty} \left(1 + \cos\left(\sqrt{8}\lambda\sigma_p + 2(\phi-\tau)\right)\right)\exp\left(-\lambda^2\right)d\lambda$$

Let $\alpha = \sqrt{8} \qquad \beta = 2(\phi-\tau)$

$$\sigma_m^2(\theta,\tau,\phi) = \frac{\sigma_p}{\sqrt{2}} \int_{\lambda=-\infty}^{\infty} \left(1+\cos(\alpha\lambda+\beta)\right)\exp\left(-\lambda^2\right)d\lambda$$

Using the general results:
$\int_{\lambda=-\infty}^{\infty} \exp\left(-\lambda^2\right)d\lambda = \sqrt{\pi}$
$\int_{\lambda=-\infty}^{\infty} \cos(\alpha\lambda-\beta)\exp\left(-\lambda^2\right) = \sqrt{\pi}\cos(\beta)\exp\left(\frac{-\alpha}{4}\right)$

$$\sigma_m^2 = \frac{\sigma_p\sqrt{\pi}}{\sqrt{2}}\left(1+\cos(\beta)\exp\left(-\frac{\alpha}{4}\right)\right)$$

$$\sigma_m^2(\tau,\phi) = \frac{\sigma_p\sqrt{\pi}}{\sqrt{2}}\left(1+\left(2\cos^2(\phi-\tau)-1\right)\exp\left(-\frac{\alpha}{4}\right)\right)$$

If we gather the angular variables, we obtain an expression for the variance of the measure image, which is a function of the filter orientation and the illuminant tilt:

$$\sigma_m^2(\phi,\tau) = a\cos^2(\phi-\tau)+b \qquad (6.2.2e)$$

where a and b are constants.

Consequently, we predict that the variance of the measure image will vary with the square of the cosine of the angle between the illuminant vector and the direction of maximum sensitivity of the Gabor filter. We can clearly observe this result if we integrate equation 6.2.2c numerically. The resulting function can then be sampled at values of ϕ corresponding to our filter set and plotted as a fuction of tilt, with ω_0=0.125ω_s.

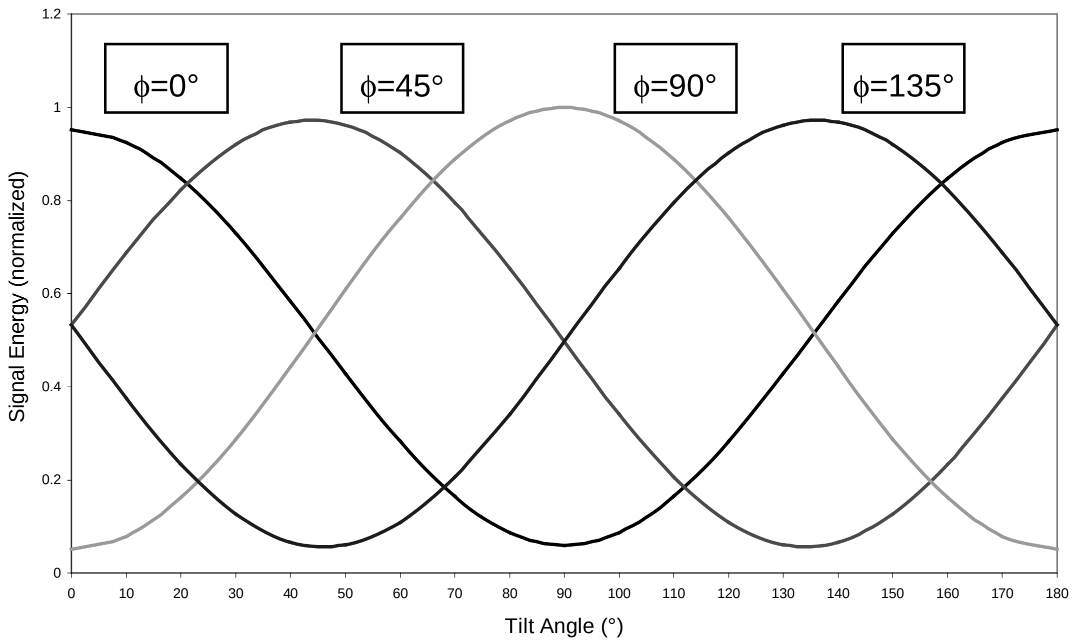


Figure 6.2.4 Numerical estimate of measure variance for an isotropic fractal surface.

The result of the numerical integration and evaluation are plotted against $cos^2(\phi-\tau)$ in *Figure 6.2.5*. In all cases a near linear relationship exists. The function $a\cos^2(\phi-\tau)+b$ is also plotted, where the parameters a and b are the fitted using least squares to give the best fit to the integration results. In the case of the ϕ=0° and ϕ=90° curves, the best fit is visually indistinguishable from the evaluated curves and is not plotted. In the case of the ϕ=45° and ϕ=135° the numerically evaluated curves exhibit a small amount of hysteresis—presumably due to numerical approximations. The hysteresis is symmetrical about the best fit curve. The numerical results agree well with the analytical predictions, (6.2.2e).

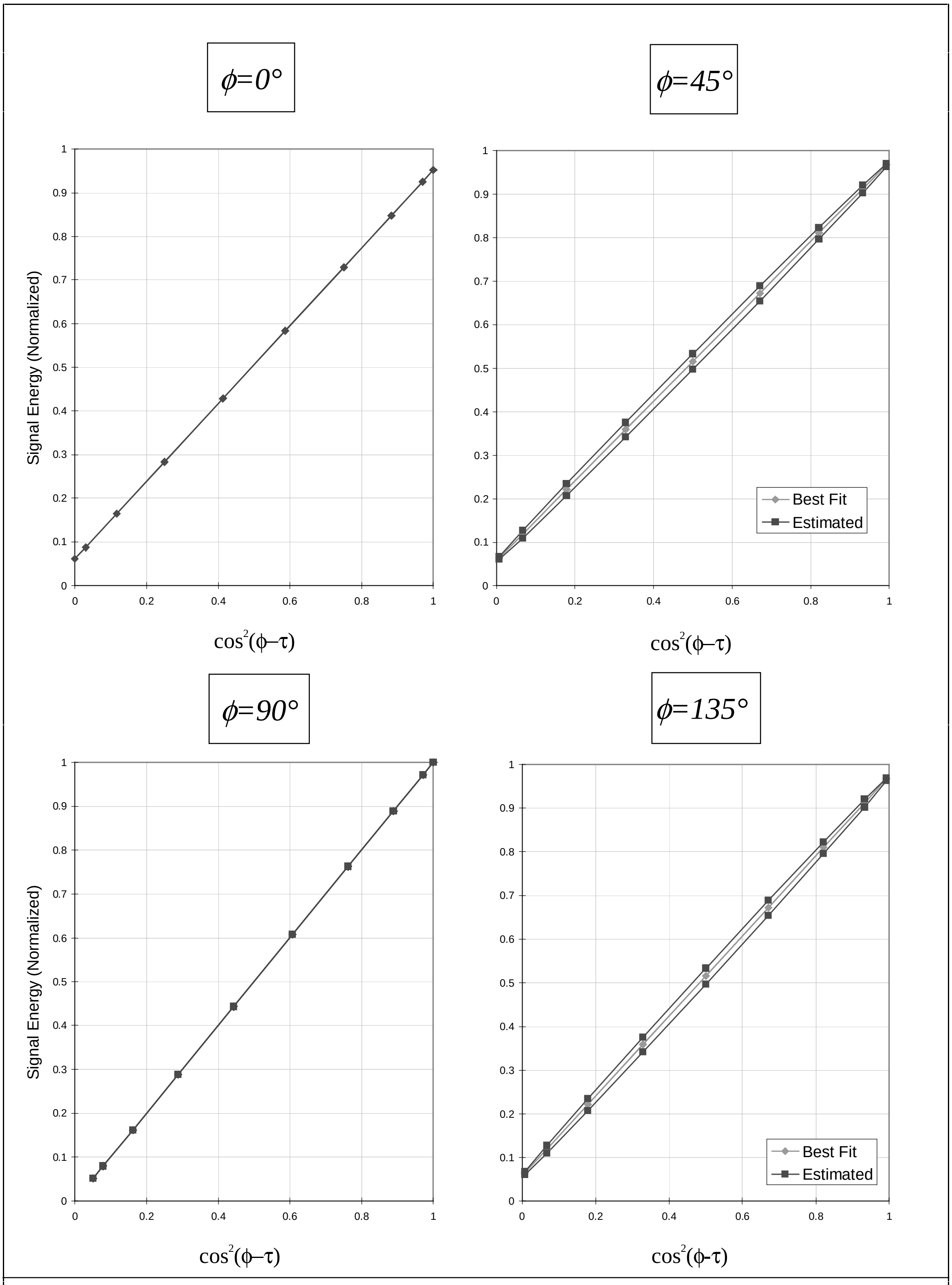


Figure 6.2.5 The numerical integration of Eq. 6.2.2c for various values of ϕ.

6.2.3 Feature Image Statistics

Assuming both the measure images to be zero mean Gaussian with standard deviation σ_m, the magnitude image will have a Rayleigh distribution, Eq. 6.2.3a, parameterised by σ_m:

Rayleigh distribution: $$p(f) = \frac{f}{\sigma_m^{\ 2}} \exp\left(\frac{-f^2}{2\sigma_m^{\ 2}} \right) \quad (6.2.3a)$$

with mean, $$\mu_f = \sqrt{\frac{\pi}{2}} \sigma_m \quad (6.2.3b)$$

and standard deviation, $$\sigma_f = \sqrt{2 - \frac{\pi}{2}} \sigma_m \quad (6.2.3c)$$

As was noted in the previous chapter, it is necessary to low pass filter the magnitude image. Due to the earlier non-linear operation of calculating the magnitude image from the quadrature images, it is not possible to model this in the frequency domain in the context of the earlier results. Nevertheless, the mean response of the magnitude will be largely unaffected by the low pass filtering, merely scaled by the sum of the filter weights. Since the mean of the distribution is a linear function of σ_m , (Eq. 6.2.3b), we predict from Eq. 6.2.2e that the feature mean will vary with a relationship of the form:

$$\mu_f(\tau) = a\left|\cos(\phi - \tau)\right| + b$$

Although we cannot predict the value of the standard deviation, we can hypothesise its relationship with tilt angle. The low pass filter is isotropic, and our (linear) model of rendering induced directionality is independent of radial frequency (if we ignore imaging artefacts). Consequently, it would be reasonable to suggest that the standard deviation of the filter shares the tilt characteristics of the magnitude image, since the standard deviation of the magnitude image is also a linear function of σ_m (Eq. 6.2.1k). We therefore hypothesise that both the mean and standard deviation of the feature follow a relationship of the form:

$$\sigma_f = a\left|\cos(\phi - \tau)\right| + b$$

We shall now compare this prediction with the filter output obtained from both simulation and experiment.

6.3 Testing the theoretical predictions

6.3.1 Verification by Simulation

The figure below shows the measured mean and standard deviation of a feature derived from a single ϕ=0° filter applied to an synthetic, isotropic, fractal surface. Both the mean and the standard deviation are scaled to give a maximum of unity. We also plot two additional waveforms for comparison, the absolute value of the cosine and the square of the cosine. The parameters *a* and *b* have, in both cases, been estimated using a least squares fit to the average of the mean and the standard deviation values.

$$O(\tau) = a\left|\cos(\tau)\right| + b$$

$$O(\tau) = a\cos^2(\tau) + b$$

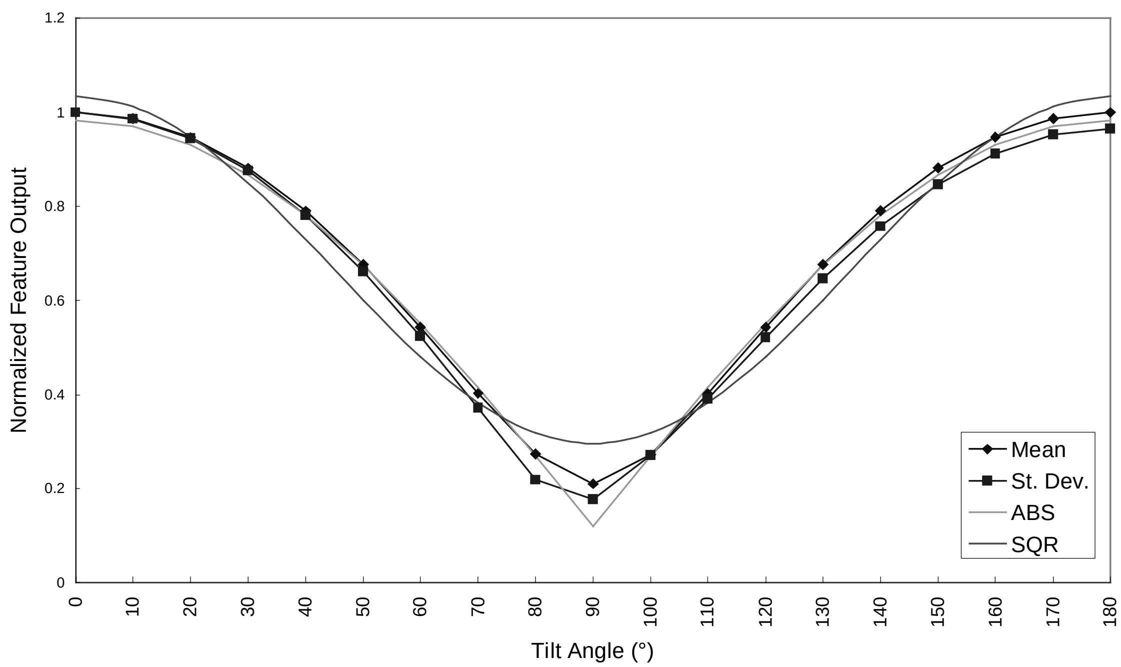


Figure 6.3.1 The effect of tilt angle variation on feature output

We note several points from the above graph:

(i) the mean and standard deviation have very similar tilt responses,

(ii) both have a similar a/b ratio,

(iii) both the mean and the standard deviation appear to be most successfully modelled using the absolute function. This is in agreement with the theoretical predictions.

In this thesis we will adopt the $a\left|\cos(\phi-\tau)\right|+b$ relationship as an approximation for the dependency of both the mean and standard deviation of the feature images on the illuminant tilt.

In the previous graph we only considered the case of a single filter, in any application it would be normal to use a larger feature set. In *Figure 6.3.2* we apply all the filters of the set within a frequency band, each is scaled by the same amount. As was predicted in Eq. 6.2.2e, the waveforms are effectively phase shifted versions of that shown in *Figure 6.3.1* each of which reaches its maximum when the tilt angle coincides with the filter orientation, i.e. $\tau = \phi$.

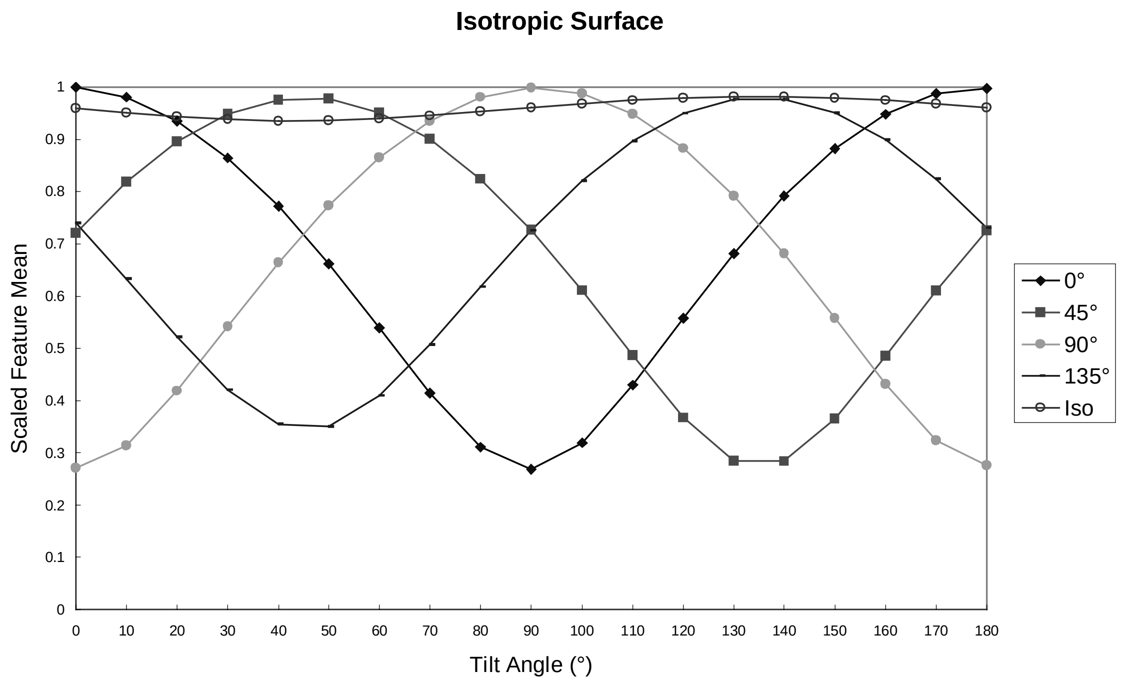


Figure 6.3.2 Tilt effects on the means of features obtained over a range of orientations in the F25 band.

Since the waveforms are phase shifted versions of each other, varying the illuminant tilt may be considered as being equivalent to rotation of the image in the case of an isotropic texture. To demonstrate this point, we plot the mean output of a fifth feature. This feature is derived from an isotropic, Gaussian bandpass filter[1]. As expected, this is almost unaffected by tilt variation.

We now consider the effect on feature output of tilt variation on the directional *Ogilvy* surface, again the feature means have been scaled by the maximum value of the largest waveform.

[1] Developed by T. Wittig

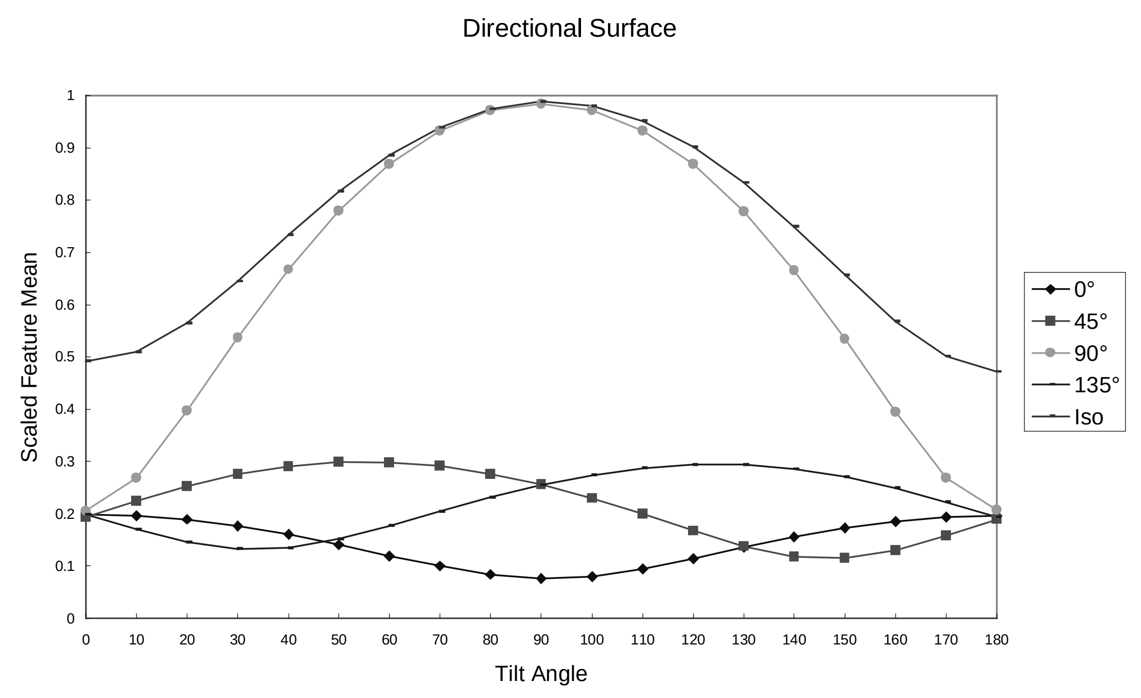


Figure 6.3.3 F25 Filter responses to directional texture with varying tilt.

We note the following points:

(i) on visual inspection the $|\cos(\phi-\tau)|$ relationship appears to hold for both textures and all filters.

(ii) the output of the filter (ϕ=0°) orientated at right angles to the grain of the surface is much larger than that of the other filters in the band.

(iii) the feature mean of the isotropic filter is heavily dependent on the tilt direction.

6.3.2 Verification by Experiment

We now consider the two real test samples *Rock* and *Striate*. As with the isotropic synthetic surface, the feature means for the *Rock* texture all vary in a cosine manner of approximately equal amplitude, though with a slightly lower a/b ratio.

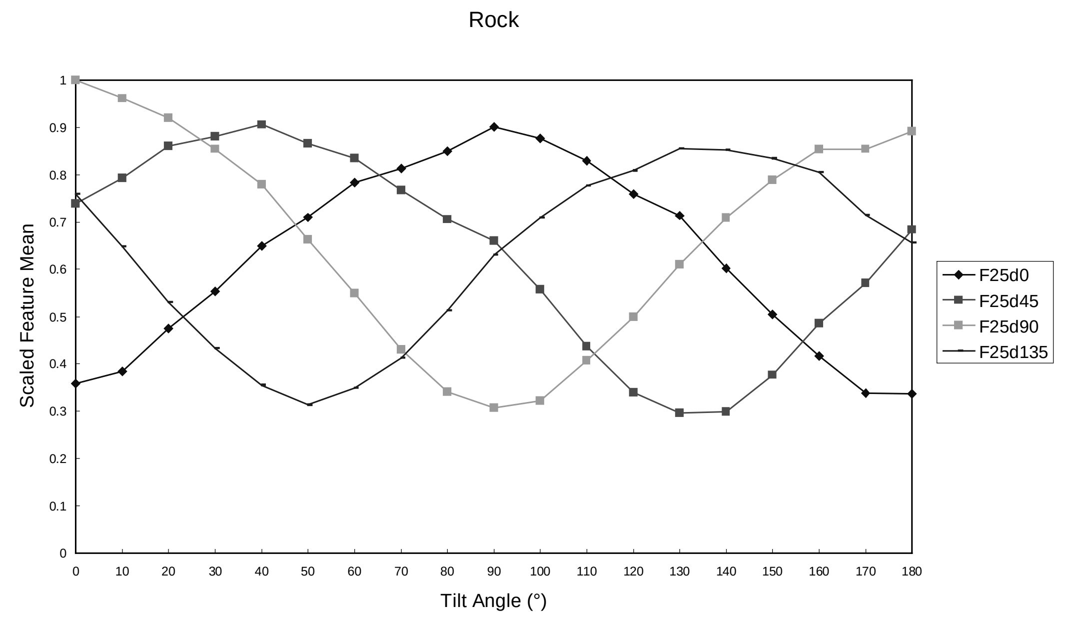


Figure 6.3.4 Filter responses to 'Rock' surface illuminated from various tilt angles.

The 'Striate' texture is analogous to the Ogilvy surface. As with the synthetic case all features vary in an approximately sinusoidal manner. The filter means can all be approximated by sinusoids of the same amplitude save that feature measured perpendicularly to the grain, which has a much larger amplitude.

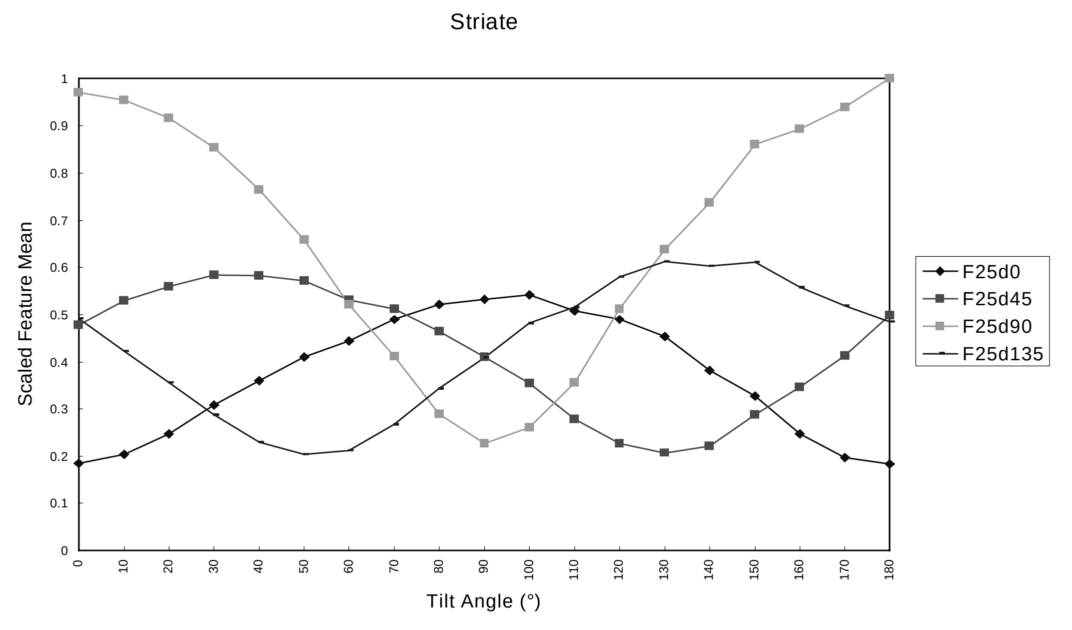
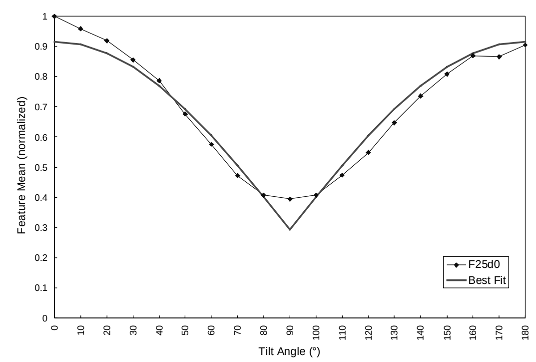
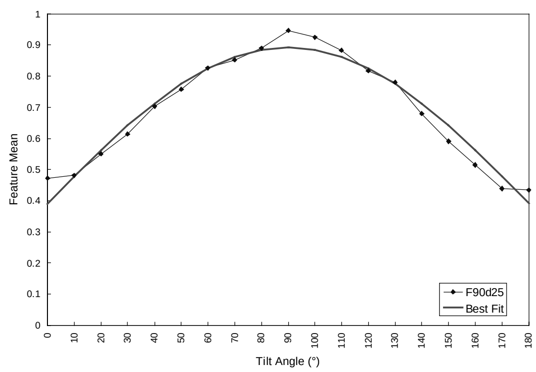


Figure 6.3.5 Filter responses to 'Striate' texture.

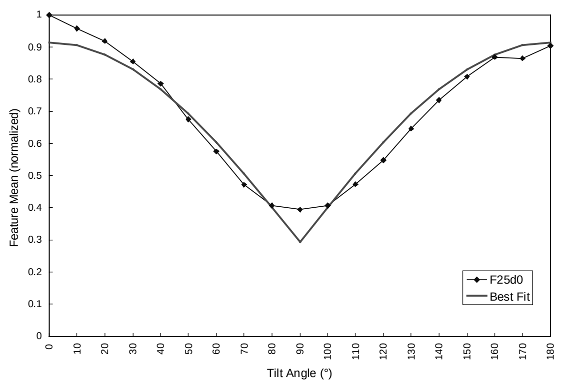
Although the feature means are important, in the context of discrimination the standard deviation of the feature must also be considered. In *Figure 6.3.6* and *Figure 6.3.7* the means and standard deviations of features derived from filters oriented at ϕ=0° and 90° are plotted for the *Rock* and *Striate* surfaces respectively. An $a\left|\cos(\phi - \tau)\right| + b$ curve is also shown in each graph where the a and b parameters have been estimated using least squares.



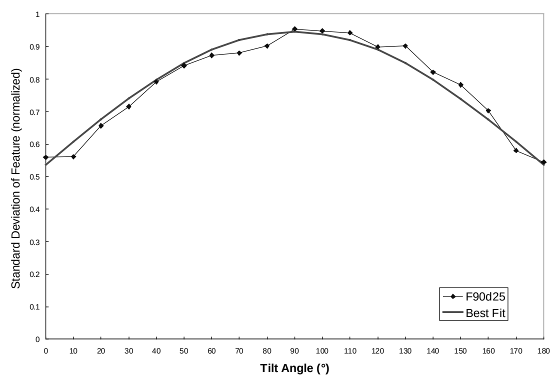
F25d0 Mean



F25d90 Mean



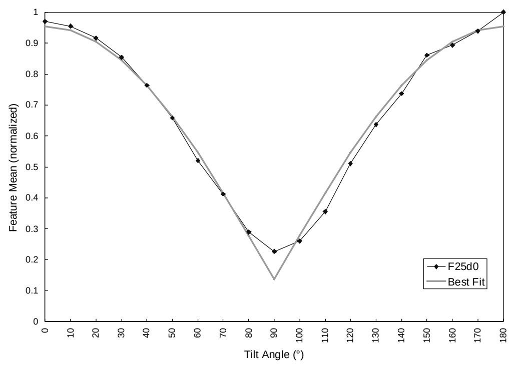
F25d0 Standard deviation



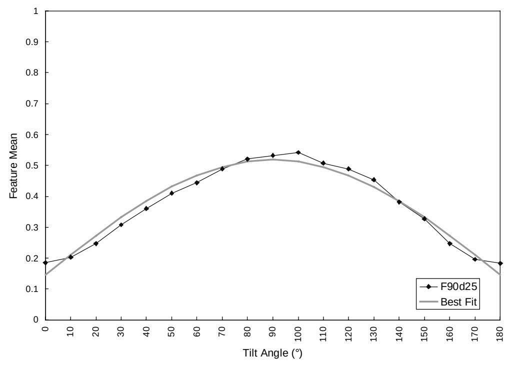
F25d90 Standard deviation

Figure 6.3.6 Variation of Rock feature statistic with tilt angle.

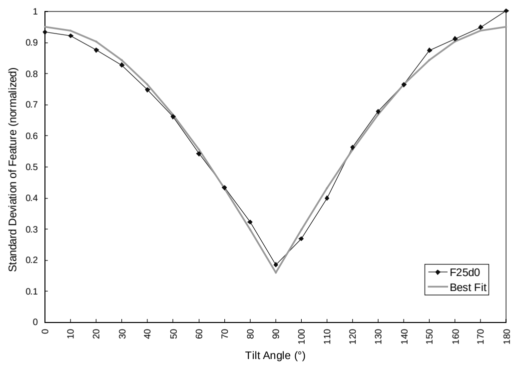
In the case of *Rock* the feature mean fits well to the parameterised curve, though the minima of the F25d0 graphs are much smoother than predicted. Both the mean and standard deviation of the F25d0 features derived from *Striate* appear to be a better fit in this respect, though there is a small phase disparity between the predicted and the observed curves in the F25d90 case. Nevertheless, the proposed relationship does appear to be a reasonable approximation to the experimental findings.



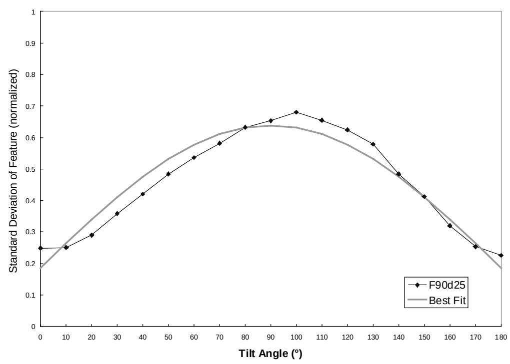
(a.) F25d0 Mean



F25d90 Mean



F25d0 Standard deviation



F25d90 Standard deviation

Figure 6.3.7 Variation of Striate feature statistics with tilt angle.

In *Table 6.3.1* and *Table 6.3.2* the estimated parameters for the feature mean/tilt relationship are tabulated for the *Rock* and *Striate* surfaces respectively. It was noted earlier that the mean curve of the F25d0 feature for the *Striate* texture is much larger than the means of the other features in the set. The *a* (tilt-dependent) parameter reflects this; the F25d0 feature parameter is almost twice that of the other features. The *b* parameter estimate for F25d0 is noticeably larger than for the features, however, the difference is much less significant than was observed with the *a* parameter.

Striate	Mean			Standard Deviation		
Filter	a	b	a/b	a	b	a/b
F25d0	14.077	5.987	2.35	2.440	1.181	1.348
F25d45	7.702	4.740	1.62	1.608	0.969	1.659
F25d90	6.870	4.125	1.67	1.534	0.938	1.636
F25d135	7.086	4.914	1.44	1.839	1.083	1.698

Table 6.3.1 Feature statistics for Striate surface.

The parameter estimates for the *Rock* show a significant amount of variation in the *b*-parameter, although the variation is small relative to that observed in the *Striate* parameters. The *a* parameter is more consistent for the *Rock* than for the *Striate* surface, and the variation which does exist does not show the same degree of linkage to the *b*-parameter as in the *Striate* estimates.

Rock	Mean			Standard Deviation		
Filter	a	b	a/b	a	b	a/b
F25d0	9.245	5.642	1.64	1.522	1.043	1.46
F25d45	7.606	5.687	1.34	1.321	1.147	1.15
F25d90	7.950	5.642	1.41	1.498	1.235	1.21
F25d135	8.736	5.205	1.68	1.101	1.162	0.947

Table 6.3.2 Feature statistics for Rock surface.

The results obtained are to some degree ambiguous, however, we make the following conclusions, albeit based on a very limited data set: the *a* parameter is largely dependent on the nature of the surface in the direction of the filter—it will therefore vary widely from surface to surface. The *b*-parameter also contains an element of dependency on the surface characteristics within the filter's bandpass region, however, it is largely invariant to the surface characteristics.

6.3.3 Summary of Classifier Modelling

In this section we have shown that measures derived from directional Gabor filters vary with tilt. Both mean and standard deviation vary with a relationship approximated by Eq. 6.3.3.

$$a\left|\cos(\phi - \tau)\right| + b \qquad (6.3.3a)$$

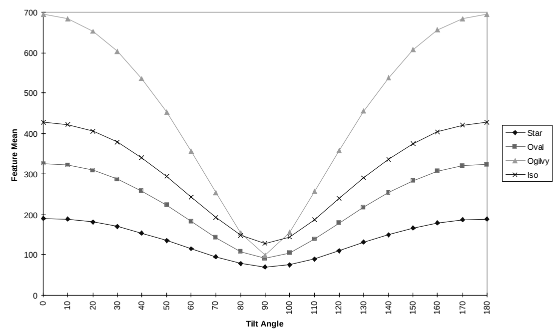
We have shown this using theory, simulation and experiment.

We have also shown that an isotropic filter, applied to the image of an isotropic surface is unaffected by tilt variations. However, an isotropic filter applied to the image of a directional texture *will* be affected by tilt variation. In addition, an isotropic feature will not, of course be able to capture important directional information.

By demonstrating that tilt affects the output of features, we have shown that there exists a potential problem for classification. In the following sections we show tilt induced classification failure occurring for several texture classification tasks.

6.4 The Effect on a Classifier

We shall now consider the effect of tilt variation in terms of classification accuracy. Due to interdependency of classes associated with classification, any treatment of this subject is necessarily empirical. We use the controllability of the synthetic textures to define a relatively simple classification task. This enables us to reduce the feature set to just two features (Gabor filters orientated to 0° and 90°) and still maintain a good level of classification. With only two features we may characterise the classification process with a two dimensional scatter plot. We also plot the feature measures for each texture in *Figure 6.4.1*. This section is designed to use this window into classification to gain an intuitive understanding of how clusters behave during tilt variation and how this will affect classification.



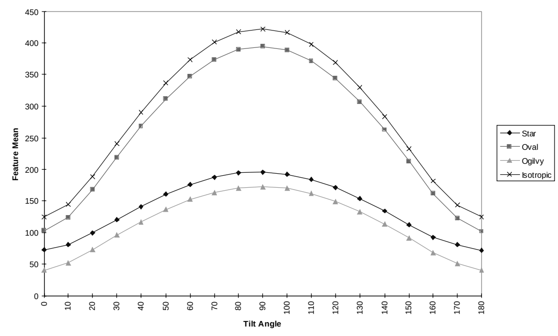


Figure 6.4.1 Feature Means derived from F64d0 and F64d90 Filters

The classifier is trained on surfaces illuminated from τ=0°, classification at this stage being generally good, the most prominent misclassification being mutual confusion between the isotropic and mildly directional surfaces, (see *Figure 6.4.2)*. The relatively compact nature of the clusters in the vertical direction reflects the attenuation of the vertical frequencies.

As the illumination is rotated towards the vertical, the vertical frequencies are accentuated and the horizontal frequencies attenuated. The cluster centres now begin to trace out an approximation to simple harmonic motion in feature space whereas the clusters themselves contract and expand along their feature axis. By the time the illuminant has reached the vertical, all the clusters lie in the area of feature space assigned to the surface which displayed most vertical energy at τ=0°, which in this case is the isotropic surface. Misclassification of the other textures is almost complete.

Figure 6.4.8

6.5 Classification Experiments For Real Textures

The thesis of this chapter is that variation in illuminant tilt between training and classification can cause the classifier to fail catastrophically. In this section, this effect will be demonstrated experimentally using real data.

6.5.1 Test Criteria

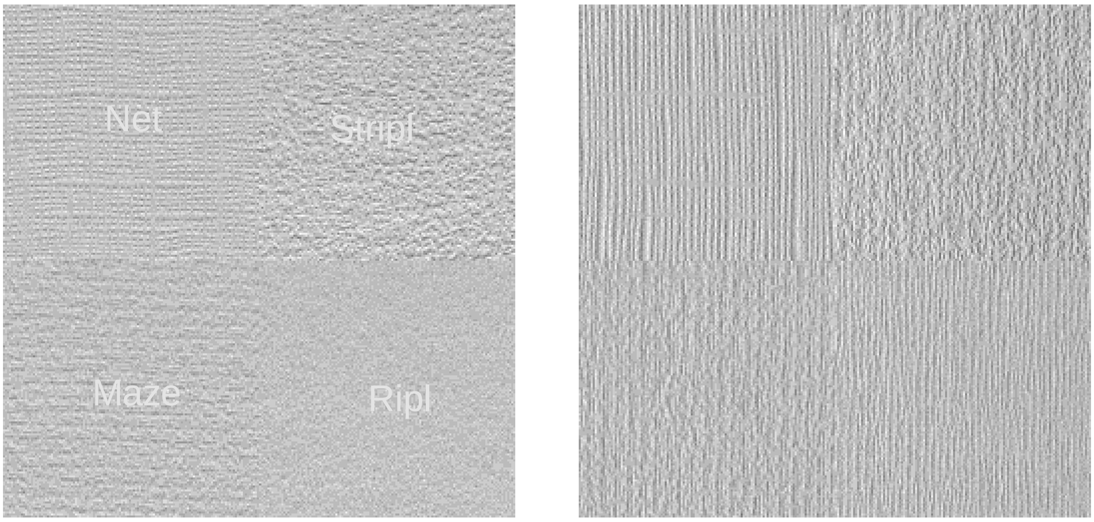
In order to demonstrate this effect, we must show three things:

1. The classifier performs well when the tilt angle is identical for training and classification.
2. The misclassification rate increases progressively, though not necessarily linearly, with the cosine of the angle between the illuminant vectors of the training and classification,
3. Implied from (2) is that the misclassification rate at τ=180° should be approximately equal to that at τ=0°.

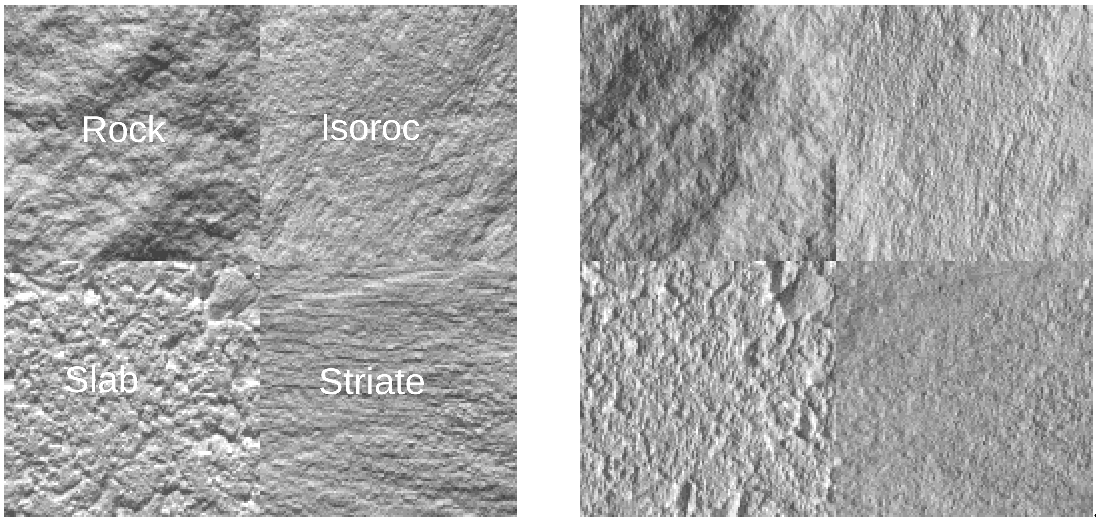
6.5.2 The Data Set

In this section we will use three texture montages: *Anaglypta, Stones1* and *Stones2,* shown in *Figure 6.5.1*. The *Anaglypta* montage consists of highly directional, and highly uniform textured surfaces. It consequently represents the easiest classification task. The *Stones1* montage consists of three approximately isotropic rock surfaces and one highly directional surface, *Striate,* in which the directionality is aligned approximately with the Y-axis, (θ=0°). The textures comprising the *Stone2* montage are all directional to some degree. The directionality of the *Slate* and *Pitted* surfaces is aligned with the θ=0° direction. The *Twins* direction has its most prominent directionality in the direction θ=90°, while the radial texture is directional at θ=45°.

Anaglypta Montage



Stone 1 Montage



Stone 2 Montage

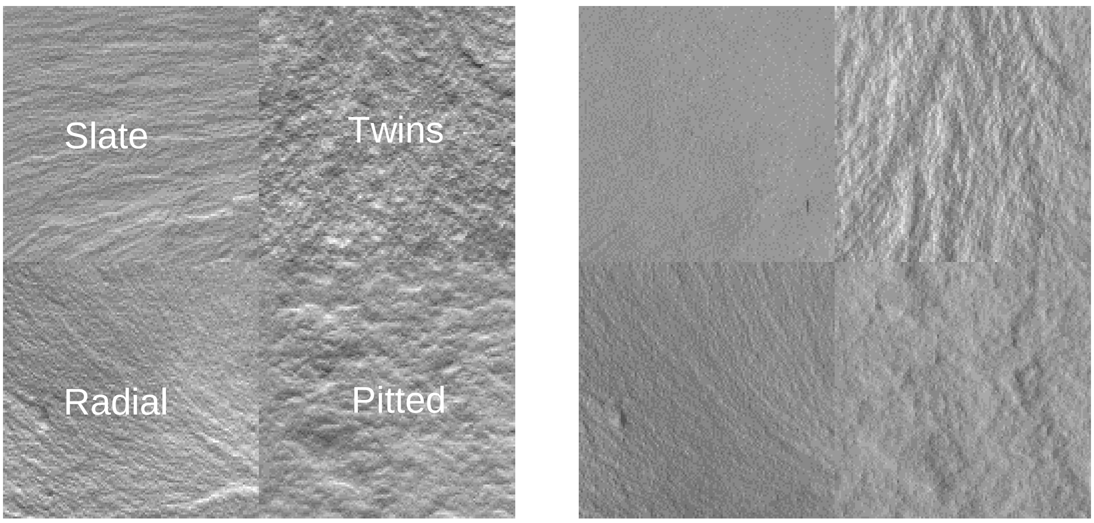
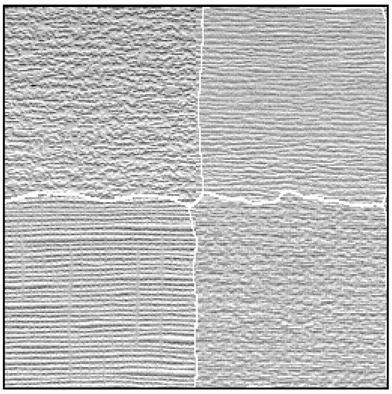


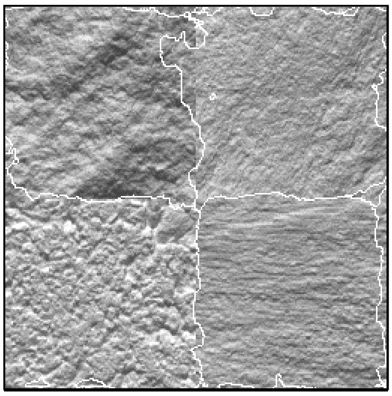
Figure 6.5.1 Test montages of surfaces, illuminated at Tau 0° (left) and 90° (right).

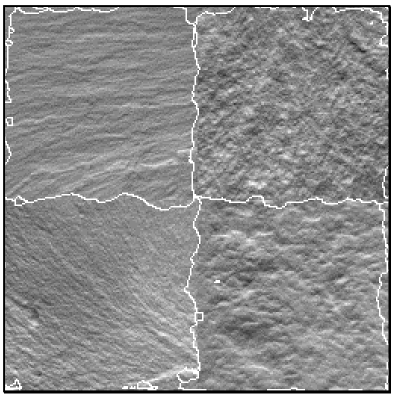
6.5.3 Experimental Work

First Criterion

Our first criterion is that the classifier should be perform classification accurately for the τ=0° image. Real textures were classified using the same system as the synthetics in the previous section, though a much larger feature set with twelve members and a 12*12 mode postprocessing filter were used to achieve an acceptable classification accuracy. Unfortunately the ability of the classifier to cope with the heterogeneous nature of surface roughness on the samples is poor, and to obtain an acceptable level of accuracy, it was necessary to use the entire montage for training, this means that we are only testing the ability of the features to describe the textures and are not testing the classifiers ability to generalise from the training data. The classified images and the misclassification rates are shown in *Figure 6.4.2*.







Montage	Misclassification (%)
Anaglypta	2.88
Stones1	5.09
Stones2	3.29

Figure 6.5.2 Montages trained and classified at Tau 0°.

As predicted, the anaglypta montage is the easiest classification. Whereas the Stones montages are classified to a much lower degree of accuracy.

Second Criterion

The second criterion for our thesis is that the classifier should be progressively less accurate as tilt angle increasingly differs from the training angle. The classifier was trained at τ=0° and then tested on images of the surfaces illuminated from τ=0 to 180° in 10° degree increments. The degree of misclassification for this classifier was recorded for each illumination condition in *Figure 6.5.3-5* In addition to the classifier trained at 0°, a classifier retrained for each tilt angle, labeled "*Best*", was also employed. This acts as a control in displaying the level of difficulty of classification inherent in a particular classification, which enables us to resolve the misclassification rate which is due to inappropriate training data.

In only one of the three cases does the misclassification rate increase in an approximately monotonic fashion as the cosine of the tilt angle increases—fulfilling our second criterion. However, deviations from the expected behaviour occur in regions of high misclassification, and are still of a magintude such that classfication at these tilt angles, with this classifier, is pointless. In this way we argue that the second criterion has been fulfilled to a satisfactory degree.

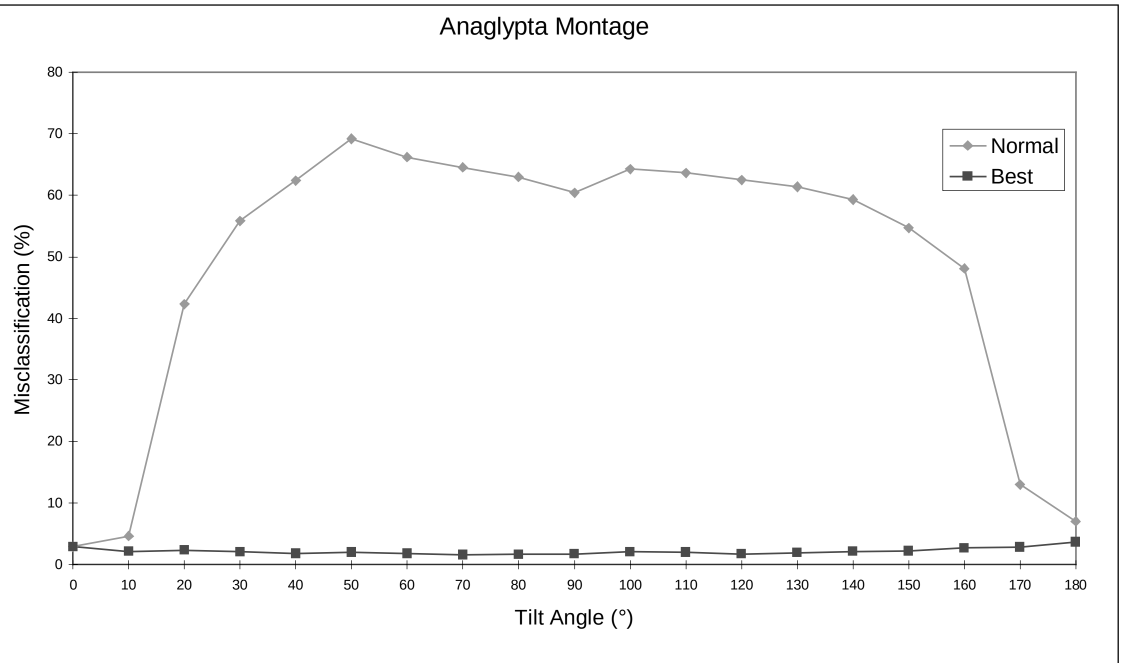


Figure 6.5.3 Misclassification rates for Anaglypta montage.

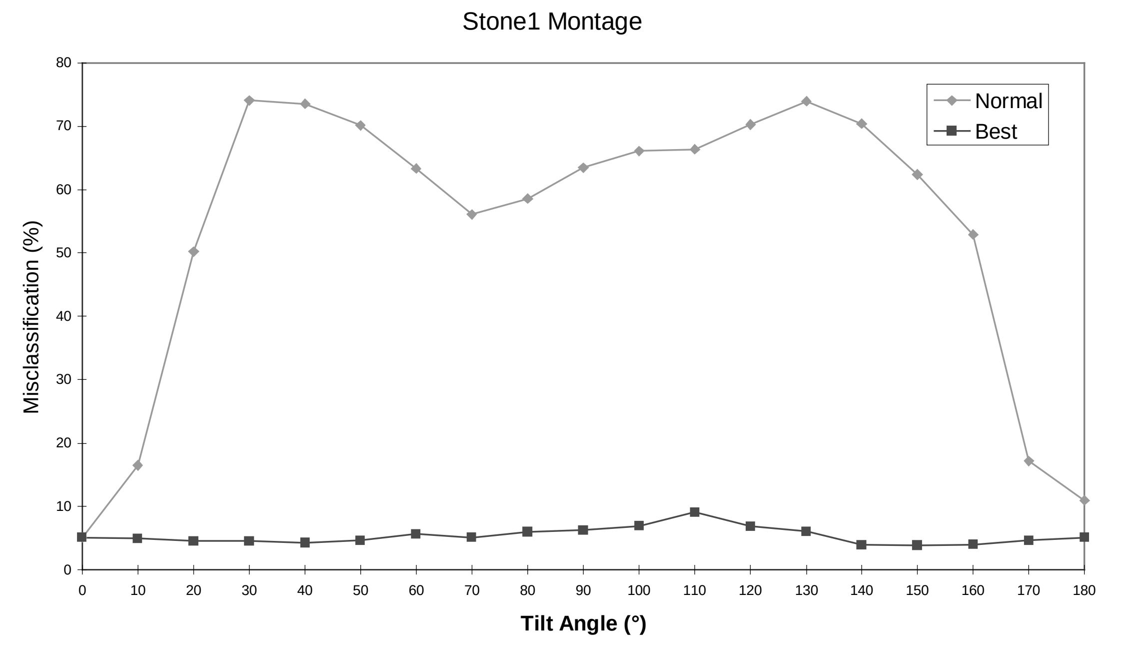


Figure 6.5.4 Misclassification rates for Stones 1 montage.

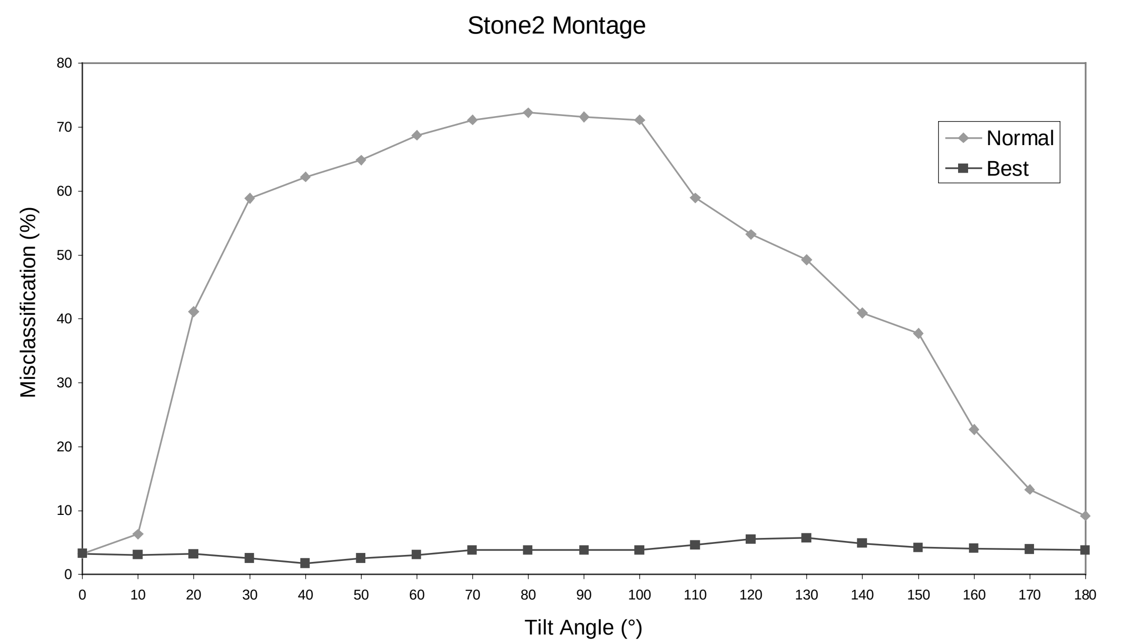


Figure 6.5.5 Misclassification rates for Stone2 montage.

Third Criterion

In order to show that the increased rate of misclassification is due to the effect of illuminant tilt our third criterion, i.e. an accurate classification at τ=180°, must be met. In all cases the classification at τ=180° is significantly poorer than that at 0°, however, in the context of the classifications at intermediate values of τ, we believe the third criterion has been fulfilled .

	Misclassification at τ=0°	Misclassification at τ=180°
Anaglypta	2.88	4.30
Stone1	5.09	10.95
Stone2	3.29	9.19

Table 6.5.1 Comparison of misclassification rates.

6.5.4 Summary of results

In order to show that classification is dependent on the illuminant tilt angle we set three criteria:

1. the classifier must be able to classify surfaces imaged under the same illuminantion conditions as those at which the training data was obtained,
2. the level of misclassification should increase with the cosine of the difference between the tilt angles at training and classification,
3. the classifier should be able to classify the τ=180° image accurately,

Using experiments on real textures, it was shown that these criteria were fulfilled to a level that provides strong evidence for the tilt dependency. We therefore argue that classification is tilt dependent, and, where illuminant tilt cannot be held constant between training and classification, the naive classifier developed in the previous chapter is not adequate.

6.6 Summary

At the beginning of this chapter we stated two aims: to model the effect of illuminant tilt on the classifier, and to observe the effect of tilt on the accuracy of classification.

Sections 6.2 and 6.3 were concerned with modelling the tilt response of the features. In the first section, the analytical model of the imaging process was extended to include the linear stage of the classifier, i.e. Gabor filtering. The analysis allowed the prediction that the first order statistics of the feature images would vary with an $a|\cos(\phi-\tau)| + b$ relationship. This was verified approximately by the simulations and experiments performed in section 6.3.

While the calculation of features is amenable to analysis, the process of discrimination is inherently non-linear and cannot be integrated into our model. In order to extend our analysis we used a synthetic classification task to observe the effect of tilt on the movement of feature clusters across discriminant boundaries. Finally the degradation in classification accuracy due to tilt variation was shown on three montages of real textures.

In this chapter, we have shown, using theory, simulation and experiment, that varying the tilt angle of the illuminant induces the movement of feature clusters. Where training and classification images are obtained at different tilt angles, this movement may cause clusters to move across discriminant boundaries. Using simulation and experiment, it was shown that this may occur to such a degree as to seriously degrade the performance of the classifier. In the next chapter we will consider several schemes to mitigate this effect.

Chapter 7

Addressing The Problem

7.1 Introduction

The goal of this thesis is to develop a classifier which discriminates between surfaces on the basis of their visual appearance. The development of such a classifier is described in Chapter 5. However, using the models developed in the earlier chapters it is shown in Chapter 6 that the direction of illumination is a critical factor in the performance of the classifier. This chapter will review a range of candidate techniques on which a classifier that is robust to tilt variations may be based. One technique will be selected for further investigation in the next chapter.

Having observed the effect, we now consider some techniques aimed at reducing tilt induced misclassification error. Chantler proposed four schemes [Chantler94], we will begin by discussing all four and evaluating the three which we believe most appropriate to the terms of this thesis. Since we wish to classify textures on the basis of their surface properties, we then proceed to review the field of shape from shading. Finally, we propose a novel scheme which uses a model based technique to overcome the problem of tilt induced classifier failure.

7.2 Review of Chantler’s Feature Space Proposals

In [Chantler94] Chantler proposes four techniques for conferring tilt-robustness on a classifier:

(i) a single classification rule obtained by training over the tilt range,

(ii) a series of rules indexed by tilt,

(iii) a segmentation based technique, and

(iv) a filter based technique designed to reverse the directional effects of illumination.

The first three techniques operate in feature space, whereas the last pre-processes the image before feature extraction. In his thesis, Chantler chose to investigate only the last option. In this work we will discuss each and investigate those which are relevant to this thesis.

7.2.1 Multiple Training Samples

Chantler's first proposal is to train the classifier over the range of illuminant conditions it is likely to encounter. He notes that the increased variance of each class may reduce classification accuracy and pursues the idea no further. We reconsider the scheme in the following terms:

In Chapter 5 we introduced a discriminant which classified a pixel, with feature vector ***F***, as having label $\hat{l}$ on the criterion shown below:

$$\hat{l}(\mathbf{F}) = \underset{l_i}{\arg\max}\,[(p_{\mathbf{F}|l_i}(\mathbf{F}|l_i))\ ]$$

where $p_{\mathbf{F}|l_i}(\mathbf{F}|l_i)$ is the probability density function of feature ***F***, given that a pixel belongs to class *i*.

In Chapter 6 we showed that the probability density function of vector ***F***, $p(\mathbf{F}|l_i)$, is a function of the illuminant tilt. We treat tilt angle as a random variable, randomly distributed between 0 and 2π, and describe using the joint pdf:

$$p(\mathbf{F}, \tau\ |l_i)$$

Chantler's proposal attempts to classify pixels on the basis of an approximation to the marginal density of *p*. The marginal distribution will be the integral of the joint density over tilt (7.2.1a), assuming the illuminant tilt angle is drawn from a uniform distribution, and the feature distributions vary continuously. This can be approximated by the summation of a finite number of sample distributions (7.2.1b).

$$K(\mathbf{F}\ |l_i) = \int_0^{\pi} p(\mathbf{F}, \tau\ |l_i)d\tau \qquad (7.2.1a)$$

$$K'(p) = \sum_{\tau=0}^{\pi} f(\mathbf{F}, \tau \mid l_i) \tag{7.2.1b}$$

Where classification is made on the basis of:

$$\hat{l}(\mathbf{F}) = \underset{l_i}{\arg\max} \left[(K_{\mathbf{F}|C_i}(\mathbf{F}|l_i)) \right]$$

The probability function $K_i(p)$ will occupy at least an equal, or more probably, a larger volume of feature space than any of its component (tilt conditional) density functions. There will consequently be a greater degree of overlap between clusters of different classes. The success of the classifier is dependent on the degree of tilt induced cluster movement being small relative to the distribution separation. In *Figure 7.2.1* we plot the feature means derived from filters oriented at 0° and 90° which have been applied to the exemplar textures: *Rock* and *Striate*.

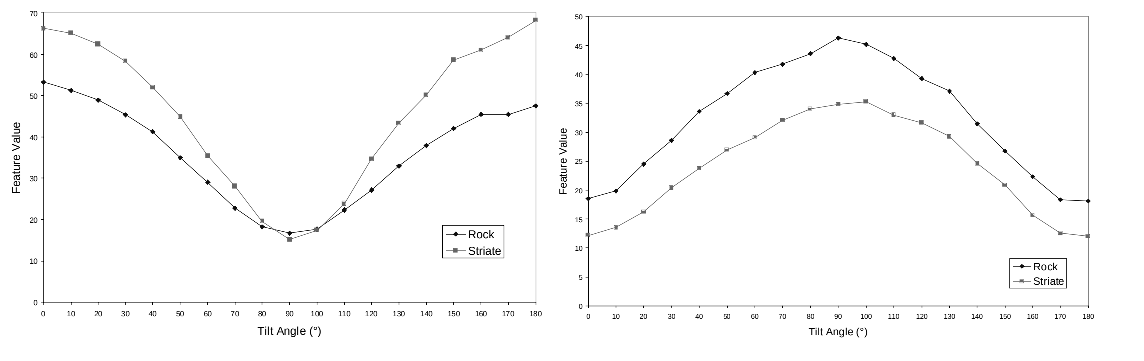


Figure 7.2.1 Feature means for Rock and Striate textures calulated from F25 filters oriented at 0° and 90°.

In *Figure 7.2.1*, it is shown that the assumption that cluster separation is large relative to cluster movement is not reasonable for the texture features measured on the *Rock* and *Striate* textures. While the clusters may or may not overlap at a given tilt, the movement of the means with tilt, in both cases, clearly show that it is not possible to set a single threshold to discriminate between the textures throughout the tilt range.

Although this technique is not appropriate in the above case, we have not ruled out the possibility that it may be effective for data sets which are well separated in feature space. While the technique is not universally applicable to the tilt problem, its simplicity of implementation makes it worthy of investigation for a given data set. We also conclude from the inadequacy of a single threshold that *a priori* knowledge of the illuminant tilt angle is a necessary condition for the illuminant invariant classification of rough surfaces. This leads us to Chantler's second proposal.

7.2.2 Multiple Discriminants

Chantler's second proposal uses multiple training samples captured under various illumination conditions to build a library of discriminant functions which can then be indexed by tilt angle. Any practical system will have a finite number of training samples. In the context of the schemes considered later in this chapter, we limit ourselves to three training images.

We conducted experiments on the *Anaglypta* and both the *Stone* montages, in each case using three discriminants, developed at 0°,90° and 180°, switching between discriminants at 50° and 140° respectively.

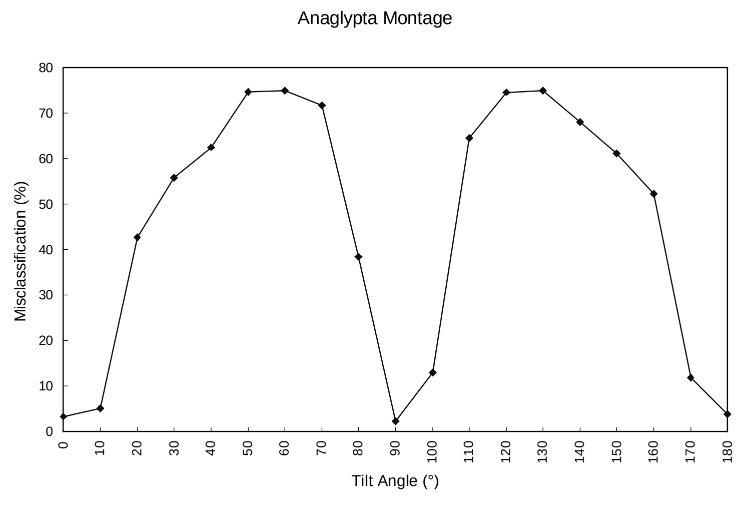


Figure 7.2.2 The use of three discriminants with the anaglypta montage.

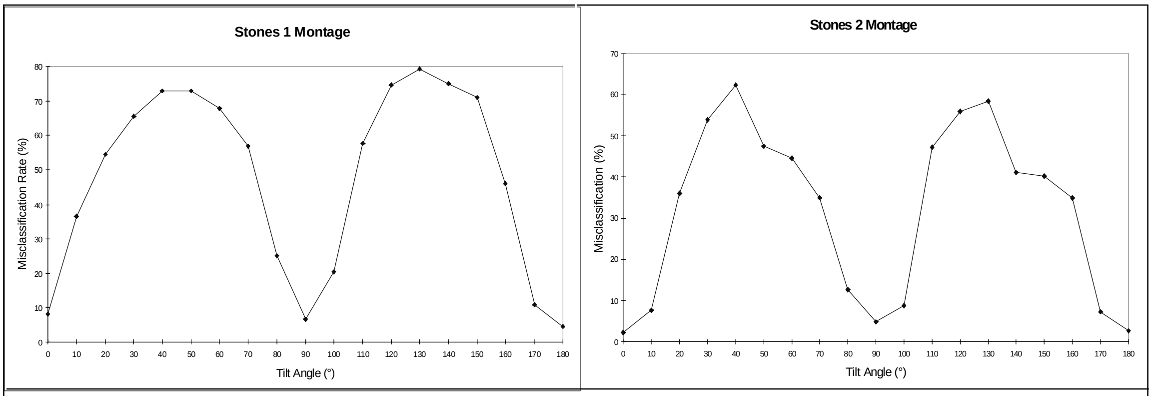


Figure 7.2.3 The use of three discriminants with Stone montages.

As we might expect, the misclassification graph has minima at 0°,90° and 180°. However, the majority of intermediate points show an unacceptably high misclassification rate. The speed with which classification errors increase as we move away from the training angle also suggests that the necessary interval between training samples is so small, certainly less than 20°, as to be uneconomical for a practical classification system. This immediately suggests an interpolative scheme, however non-linearities in both the imaging and the classification processes make this problematic. Later in this chapter we will propose a technique that is a derivative of this scheme, but which operates in image, rather than feature space, circumventing the problems of non-linearity.

7.2.3 Segmentation/Classification

The rationale for this scheme is that since the problem is caused by movement of clusters across discriminant boundaries an appropriate strategy is to track the clusters using cluster analysis techniques. The postulated clusters then may be identified as belonging to an *a priori* defined class at a higher level in the image understanding hierarchy. Identification at this level would be carried out by some feature measure (Chantler suggests features based on the power spectra) which for reasons of computational cost, or the requirement for homogenous regions of data, could not be effectively integrated into the lower levels. We note that cluster analysis is being used as a segmentation tool, and is therefore interchangeable with other segmentation techniques, such as edge-based or region growing approaches. This scheme is effectively an unsupervised technique, and consequently outwith the terms of this thesis. We do not investigate it further.

7.3 Chantler's Filters

While Chantler proposed four techniques he selected only one for further investigation. Unlike the earlier techniques which seek to deal with the feature space effects of tilt variation, this technique is proactive, and seeks to remove, or at least mitigate the effects of tilt before features are extracted.

7.3.1 Review

Chantler proposed and evaluated a system of filters to reduce the effect of tilt effects on the classification of texture [Chantler94]. After verifying Kube's work on real textures, Chantler used this model as the basis on which he could develop a frequency domain technique to remove the directional effect introduced by illumination—treating the problem essentially as one of inverse filtering. In this section we consider four issues that arise from this scheme.

- *Linearity*: the filters are subject to the restrictions on surface type and lighting conditions considered in Chapter 3. Chantler attempted to model the effects of non-linearity by adding an empirical term, *b*, to form his F1 filter class (7.3.1a). In effect treating the unwanted signal components as additive white noise.

$$H_{F1}(\omega,\theta) = \frac{1}{m\left|Cos(\theta-\tau)\right| + b} \qquad (7.3.1a)$$

where

m and b are empirically derived coefficients.

- *Frequency dependency*: if we attempt to fit the F1 model to radial plots taken at different frequencies we find a significant radial frequency dependency in both terms. This led Chantler to develop the modified F2 model, where the second order parameters are estimated in a two stage process: first, radial plots are taken for overlapping frequency ranges, the *m* and *b* parameters are fitted to each using least squares; second a linear least squares model is fitted to each parameter as a function of frequency.

$$H(\omega,\theta) = \frac{1}{m(\omega)\left|Cos(\theta-\tau)\right| + b(\omega)} \qquad (7.3.1b)$$

- *Directionality*: as the estimation process can only be applied to isotropic or near isotropic textures, it assumes that directional textures, for which parameters cannot be estimated, will be similarly affected by changes in tilt. In Chapter 3 it was shown that rough directional surfaces do exhibit behaviour incompatible with Kube's model. In combination with the linearity restrictions this forms a significant limit to the utility of the scheme. The degree of restriction is illustrated by the fact that the isotropy condition, strictly applied, would rule out all the test montages.

• *Optimality*: The work reported in Chapter 3 was based on the performance of a filter which was optimal for a particular texture. In a classification problem, we must accept the fact that any general filter will be sub-optimal for each texture. If we reconsider the frequency variation of model parameters with frequency for different textures we see that there is a wide variation from texture to texture. Chantler tackled this problem by averaging the model parameters for each texture. Whether the resulting sub-optimal filter will be sufficiently effective will depend on the similarity of the textures. This interdependency limits the generality of any experimental results obtained.

In order to address these issues, prior to an evaluation of the technique, a new set of synthetic textures *Figure 7.3.1* is introduced.

• *Linearity:* in chapter 3 we concluded that the major factor affecting how well a linear model describes rendering is the rms slope of the surface. The rms slopes of the synthetic surfaces are shown in *Table 7.3.1*.

• *Directionality*: the requirement for isotropic surfaces is satisfied by using Malvaney and fractal surface models.

• *Frequency dependency:* The synthetic textures can be used to gauge the effect of frequency dependency in the optimal parameter values. The distinct radial frequency characteristics of the fractal and Mulvaney models provides one point of comparison. The fractal surfaces (1 and 2) differ in their roll-off rates, β=3.0 and 4.5 respectively. The Mulvaney surfaces differ in the cut-off frequency at which the transition between white noise and fractal roll-off occurs.

• *Optimality:* The issue of optimality is accomodated into the evaluation in two steps. Firstly, experiments are carried out on a texture by texture basis. In each case the filter is designed purely for that texture. In the second stage the filter parameters of the texture specific filters are averaged to produce a general filter which is applied to all the textures in the montage.

The spectra and rms slope of the surfaces are shown in *Figure 7.3.2* and respectively.

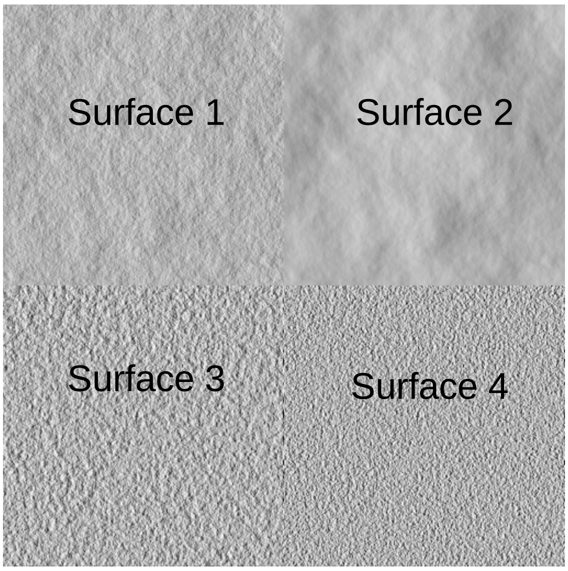


Figure 7.3.1 Isotropic texture montage.

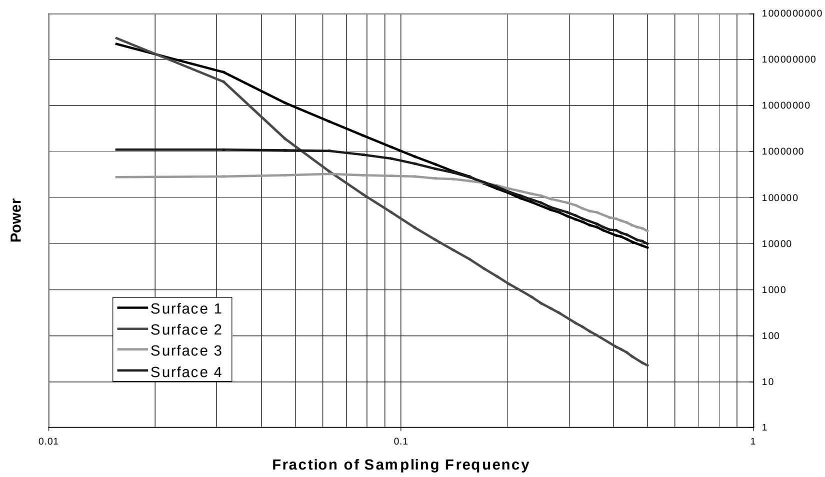


Figure 7.3.2 Spectra of isotropic surfaces.

Surface	m_{rms}
1	0.0924
2	0.0653
3	0.249
4	0.223

Table 7.3.1 RMS Slopes of test surfaces.

7.3.2 A Texture Specific Filter

In the last section it was concluded that a significant limitation of the technique would be the requirement for it to generalise, i.e. to operate effectively for a range of textures. Our assessment of the algorithm proceeds in two stages. In this section we ignore the question of generality and test the *form* of the filter—each texture is processed by a filter designed specially for that surface. In the next section the impact on performance of the general filter is assessed. This experiment represents a more realistic test of the algorithm. By resolving our assessment into two distinct stages we believe we will gain a better understanding of the technique's performance.

We begin by estimating the model parameters for each texture in the test set. This is done by splitting the power spectrum of each texture into fourteen radial frequency bands.The polar distribution of signal magnitude for each band is measured and a curve of the form $m\cos(\theta-\tau)+b$ is then fitted to the polar plot using least squares. The m and b parameters plotted against frequency in *Figure 7.3.3*. For both parameters the family of parameter curves can immediately be split into two groups which correspond to whether the surface is fractal (surfaces 1&2) or of the Mulvanney type (surfaces 3&4). The fractal textures exhibit a gradual decline in the value of the m-parameter with frequency, whereas for the Mulvanney surfaces the parameter actually increases before saturating. In the case of the b-parameter the Mulvanney surfaces are largely independent of frequency while the fractals show a significant drop with frequency. We therefore conclude that the form of the filter is a function of the surface type.

In our first application of the filters we postpone the issue of optimality and treat each texture independently. The F2 model is fitted to and applied to each texture. This involves fitting linear functions of frequency to both the parameter curves. We then examine the output of the 0 and 90° features of the f64 filter set, which is concentrated in the model's linear region ($\omega<0.25\omega_s$). In all the compensated cases the variation of feature output with tilt angle is much flatter, i.e. more stable than the uncompensated outputs (*Figure 7.3.4* and *Figure 7.3.5)*. However, we note the decrease in the effectiveness of the filters with the rougher surfaces. Even with F2 filters optimised to a particular texture the effect has not been eliminated.

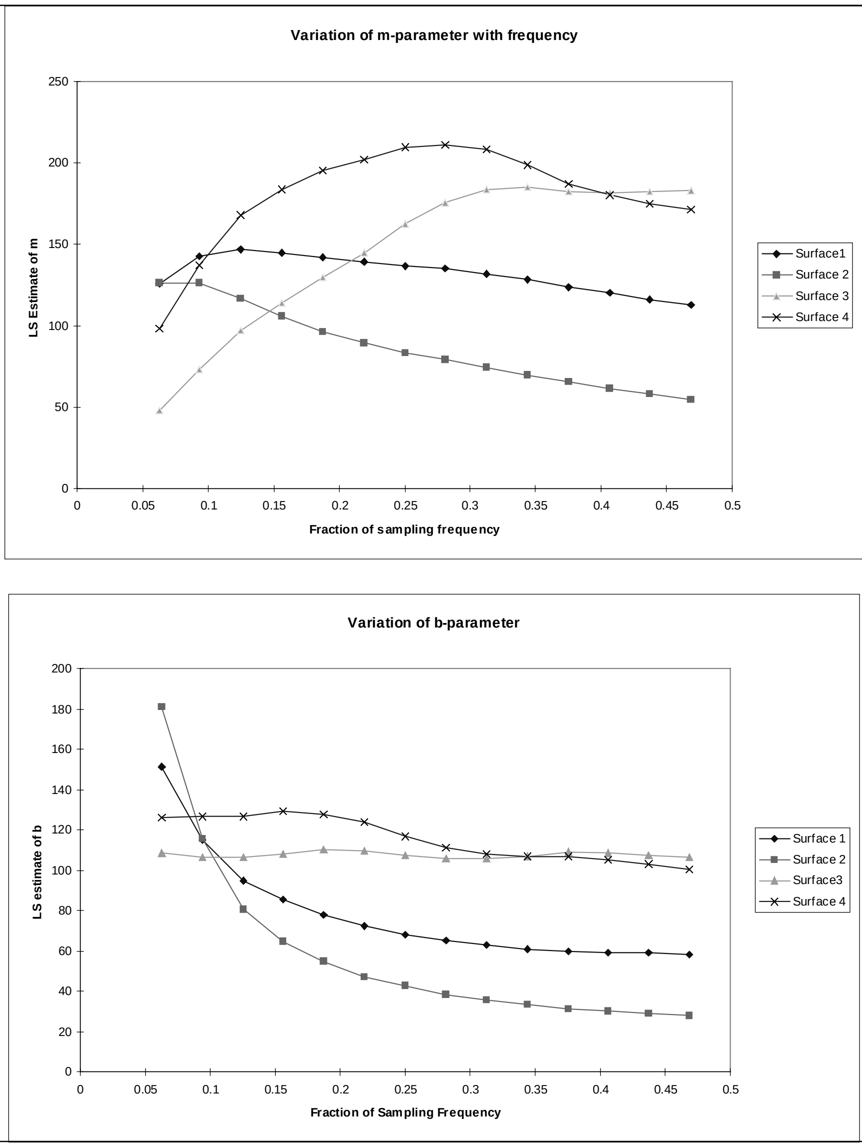


Figure 7.3.3 Frequency dependency of filter parameters for test textures.

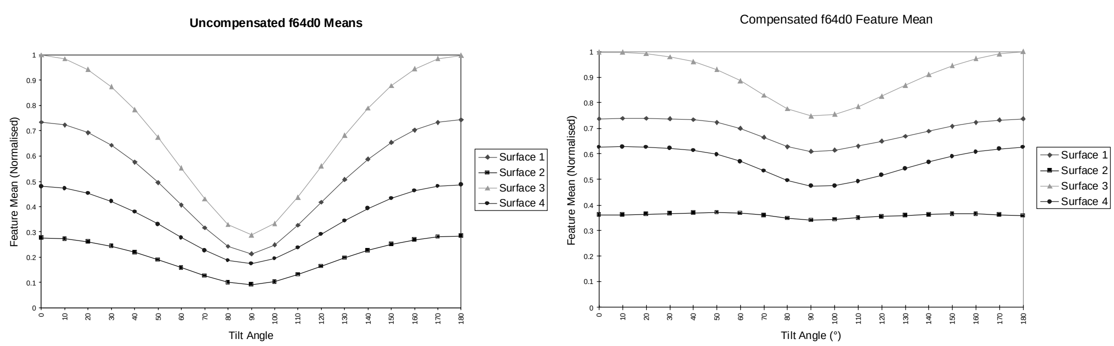


Figure 7.3.4 Uncompensated and compensated F64d0 feature means.

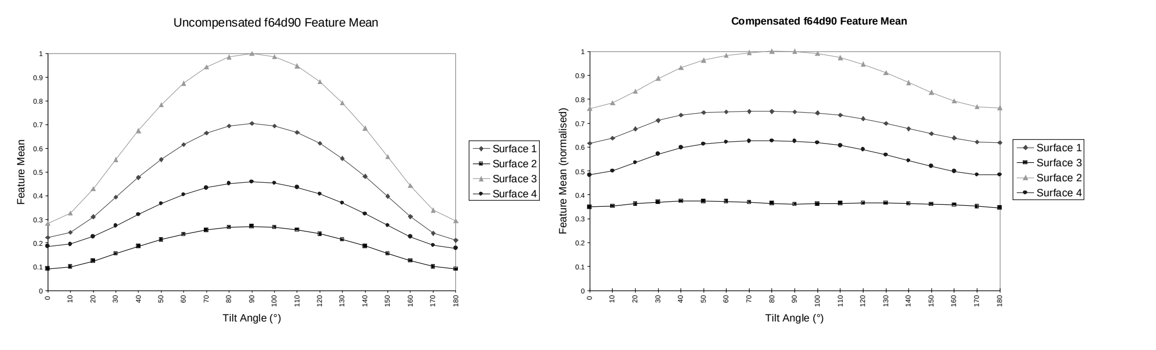


Figure 7.3.5 Uncompensated and compensated F64d90 feature means.

7.3.3 A General Filter

As stated earlier, for segmentation purposes we must adopt filter parameters which are a compromise for the textures in the test set. The compromise parameters are are estimated by averaging the *m* and *b* parameters associated with each texture and are plotted against frequency in *Figure 7.3.6,* linear functions of frequency are then fitted to both of these parameter curves. Application of the compromise filters shows a marked decrease in the stabilising properties of the filter (*Figure 7.3.7)*. If we classify on the basis of the two compensated filters we find that misclassification rates quickly rise to an unacceptable level (*Figure 7.3.8*).

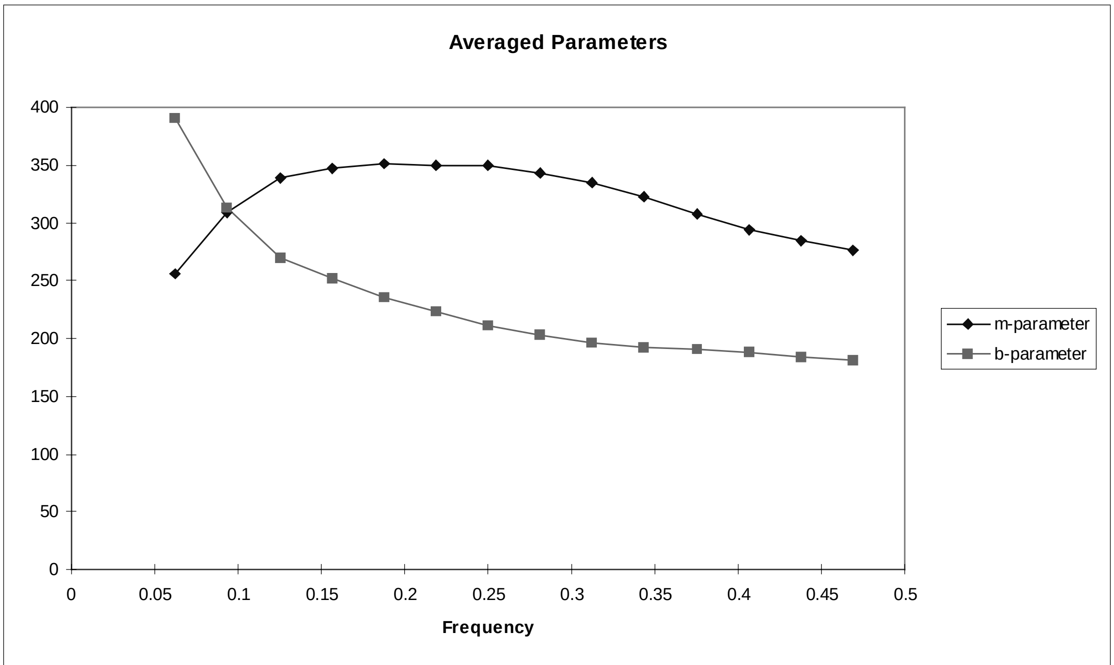


Figure 7.3.6 'General' filter parameters.

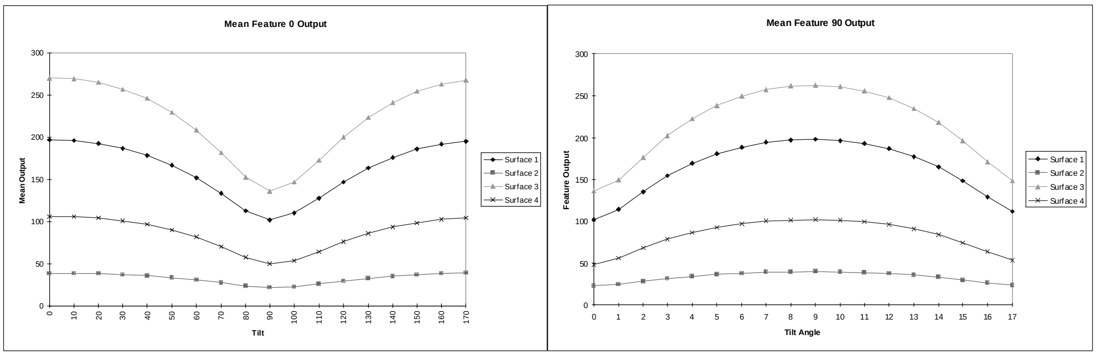


Figure 7.3.7 Feature means compensated using 'general' filter.

The filters do seem to work effectively on certain textures, specifically those with low slope angles and which are fractal within the corrected frequency range. In combination with the isotropy condition we believe the range of textures to which the technique is applicable is limited. We therefore conclude that this approach does not form an effective general approach to the problem of tilt induced failure.

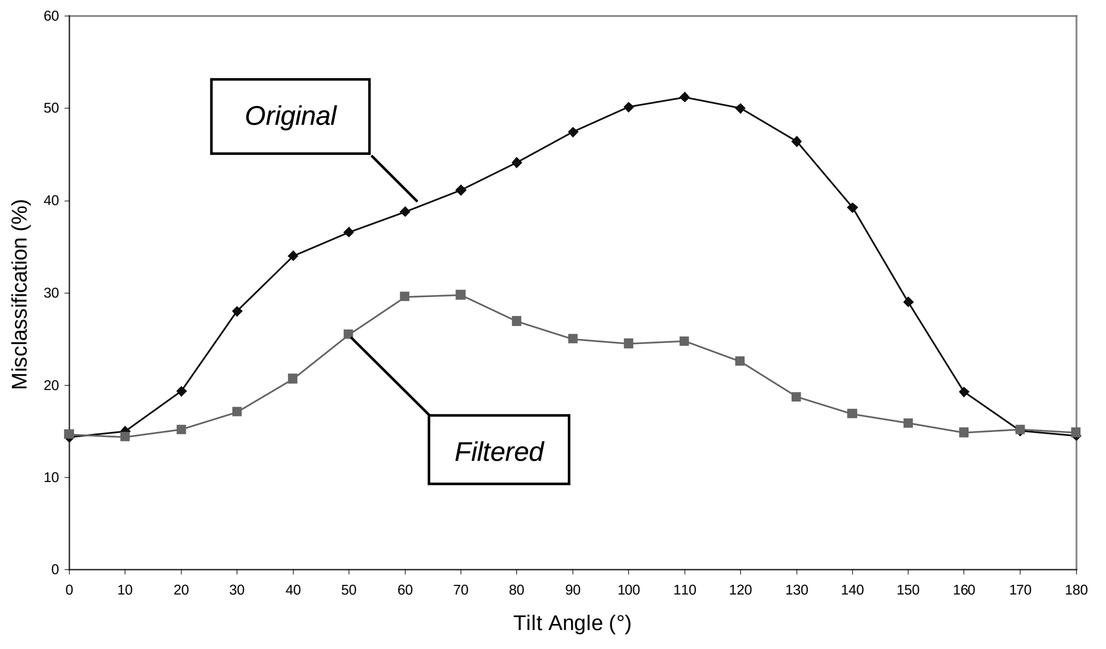


Figure 7.3.8 Misclassification rate of original and compensated montage.

7.4 *Single Image Shape From Shading Techniques*

7.4.1 Motivation

The goal of this thesis is to develop a scheme which can classify objects on the basis of their surface texture. There are a wide variety of texture analysis techniques which may be applied to the classification of a textured image. However, classification of surfaces on the basis of the appearance of textures is only valid if appearance is invariant under all conditions likely to be encountered. Chantler has shown that, for a large class of textures, changes in lighting orientation can radically alter the appearance of the texture [Chantler94]. This would suggest that where illumination direction, *relative to the texture*, cannot be strictly controlled, classification based purely on the appearance of a texture may not be an appropriate approach.

Classification of rough surfaces should ideally be carried out on the basis of the texture's surface, rather than its image structure, i.e. either on $s(x,y)$ or $\mathbf{S}(x,y)$ instead of $i(x,y)$. Several approaches exist for the recovery of surface topography, however shape from shading (SFS) is the most suited to our requirements since it does not require hardware additional to that already used for the classification task. Use of a single image SFS technique is particularly attractive since it would operate on the same data set as a naive classifier. The classifer would need only relatively minor modifications, applying the same texture analysis techniques to the surface representation rather than the image.

Shape from shading techniques have been used in computer vision for a quarter of a century, for most of this time they have been restricted to smooth surfaces. More recently, [Horn90] has applied the technique to complex wrinkled surfaces, while Pentland has applied his own novel technique to fractal surfaces [Pentland90]. This section will briefly describe the main approaches to shape from shading before examining the difficulties peculiar to their application to textures; finally the various techniques will be assessed in light of these difficulties.

Shape from shading techniques tend to capture high frequency surface variations, though they appear less adept at recovering low frequencies [Frankot88]. This seems to be true for human image interpretation [Knill90a]. Pentland [Pentland88] suggested the combination of low frequency data from a binocular stereo system and high frequency information from SFS techniques in the frequency domain. This was recently successfully implemented by Cryer et al. [Cryer95] by filtering stereo and shading derived depth maps with low pass and high pass filters respectively. Since we are concerned with textures, we are principally interested in the higher frequencies; shape from shading therefore seems to be a promising approach.

7.4.2 Shape From Shading

Introduction

Recovery of shape from shading is an ill-posed problem; a particular intensity value may be caused by any one of an infinite number of surface orientations. The intensity/orientation relationship for a given surface type is described by the reflectance map. This has the form shown in *Figure 7.4.1* where each contour corresponds to a particular image intensity. Therefore a surface facet with a particular intensity may have any orientation *p,q* which lies on the appropriate contour. There are some intensities which do correspond to a single orientation, e.g. maxima, these are known as singular points. However most intensities are not in this category and cannot be mapped uniquely to a single orientation. This ambiguity is the central problem of shape from shading, and the various SFS algorithms can be categorised by the approach they adopt to solve this problem.

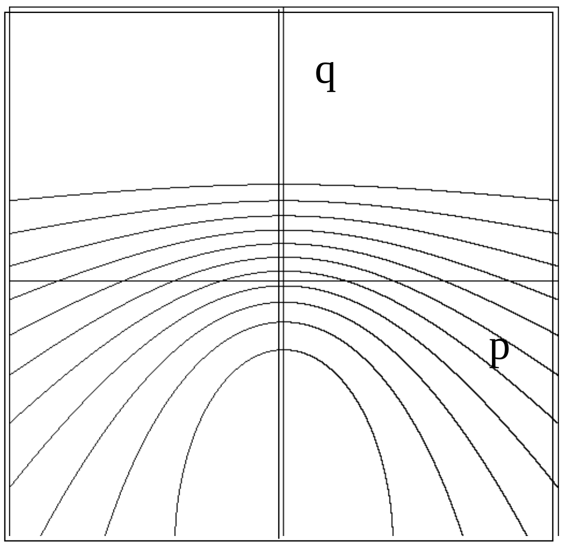


Figure 7.4.1 General form of the reflectance map

Classical Methods

The first attempt to solve the shape from shading problem was proposed by Horn [Horn70]. This treated the problem as that of solving a first order non-linear partial differential equation. Proceeding from a singular point the equation was solved to give characteristic curves of known orientation which were then grown to give the orientations of the entire surface. This technique has several difficulties: it has not been amenable to computer implementation, it is sensitive to measurement noise, and the areas grown from characteristic strips do not always merge well. This technique has largely been superceded by iterative schemes.

Iterative Techniques

Many iterative schemes have been proposed since [Strat79], most iterate on two criteria: the closeness of the simulated image of the recovered surface to the original image, and the smoothness of the resulting surface. Several algorithms also include integrabilty as a criterion—equation 7.4.2a is cited in [Zheng91] as a typical cost function:

$$\iint \underbrace{\left[i(x,y) - o(p,q)\right]}_{\text{Closeness of intensity}} + \lambda_s \underbrace{(p_x^2 + p_y^2 + q_x^2 + q_y^2)}_{\text{Smoothness}} + \mu_s \underbrace{\left[(z_x - p)^2 + (z_y - q)^2\right]}_{\text{Integrability}} dxdy \quad (7.4.2a)$$

where λ_s, μ_s are constants,

o(p,q) is the reflectance map

p_x, q_x, p_y, q_y, are the second derivatives

z_x and z_y are the derivatives of the estimated surface reconstruction.

The smoothness term restricts the applicability of the algorithm to smooth surfaces, and even for smooth surfaces may prevent convergence to the optimum. Recently however, there has been a more critical approach to the use of the smoothness criterion. Horn presents several refinements to the system, these include representation of both height and gradient to enforce integrability, a local linearisation of the reflectance map around the current gradient estimate and the ability to suppress the smoothing term as the optimum is approached [Horn90]. Using these techniques he is able to recover complex wrinkled surfaces.

Zheng and Chellapa have pointed out that most smoothness terms take no account of abrupt changes in the original image, and among other results, he presents a modified smoothing term [Zheng91]:

$$\left[R_p(p,q)p_x + R_q(p,q)q_x - I_x(x,y)\right]^2 + \left[R_p(p,q)p_y + R_q(p,q)q_y - I_y(x,y)\right]^2$$

Although Zheng conducted his experiments on locally smooth surfaces, in several images it is possible to observe discontinuities, albedo was also recovered.

A Frequency-Based Approach

In a highly original paper [Pentland90] operates in the frequency domain to recover complex surfaces including a synthetic fractal surface. Pentland uses a linearised form of the Lambertian equation as an invertable transform. This can be applied to the frequency domain representation of the original image and the resulting image can be returned to the spatial domain to yield the surface. The inverse transform is shown in Eq. 7.4.2b. Pentland has developed a modified version of this, incorporating a Wiener filter to suppress noise and non-linearities in the image, Eq. 7.4.2c.

$$H(\omega,\theta) = \frac{e^{-i\pi/2}}{2\pi\omega\left[k_1\cos\theta + k_2\sin\theta\right]} \tag{7.4.2b}$$

where $k_1 = \cos\tau \sin\sigma$

and $k_2 = \sin\tau \sin\sigma$

$$H(\omega,\theta) = \frac{e^{-i\pi/2}}{2\pi k_o\omega\left[sd + k_1\cos\theta + k_2\sin\theta\right]} \tag{7.4.2c}$$

where s= Signum[cos(τ-θ)]

and d is in the range 0.5 to 0.7

Since the Lambertian equation has been linearised, surface components perpendicular to the illuminant direction are not illuminated. The Fourier components of these patches must either be set to a default value, or estimated from other sources such as singular points. Pentland reports that using default values produces good approximations to complex and irregular surfaces, though it is less effective when dealing with regular geometric shapes. In a later paper [Pentland91] concerned with photometric effects in optical flow, Pentland uses a sequence of three images to first linearise the images before recovering both albedo and shape.

A Probabilistic System

Knill and Kersten [Knill90b] also use a linear approximation to the reflectance map as the basis for their Bayesian scheme. Rather than tackle the problem of under-determination by deterministic means, Knill seeks to use *a priori* knowledge of the surface type to estimate the most likely surface for a given image. Using 800 synthetic, illuminated, fractal surfaces as training data, Knill used the Widrow-Hoff algorithm to calculate the coefficients for two, 2 dimensional, FIR filters. These filters were used to model the mapping between the surface normals and the intensities of the corresponding pixel and those of its neighbours. Using the assumption of surface isotropy, variation in tilt angle is modelled by simple rotation of the masks.

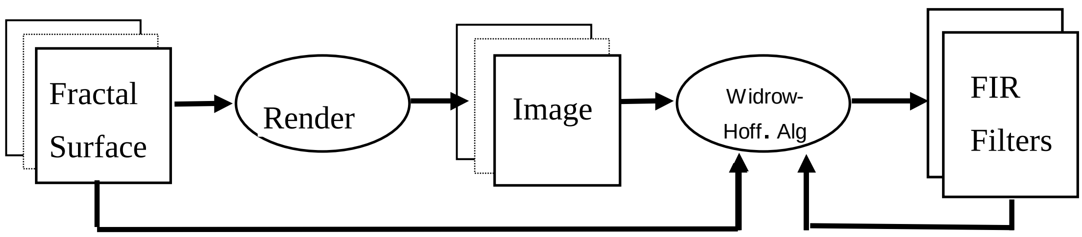


Figure 7.4.2 Deriving surface recovery filters.

The difficulty with the application of this scheme to the tilt-invariant problem lies in the requirement for *a priori* knowledge of the surface. It therefore follows that identification of the surface type is a prerequisite for the application of the optimum filter for surface recovery. Since, for our purposes, surface recovery is a means towards the end of identification this is clearly problematic, though if surfaces are sufficiently similar, it may be possible to apply a sub-optimal filter to recover the surfaces to a level of accuracy which will allow reliable classification.

7.4.3 Summary

Until recently, SFS techniques have been applied exclusively to predominantly smooth surfaces, and their application to textures raises several issues. If the texture surfaces are assumed to be fractal, then due to the self-similarity property, any smoothness in the image will be due to camera effects. Iterative techniques, unfortunately, use smoothness as a cost function. Horn's partial suspension of the criteria, and Zheng's rationalisation are both only partial solutions, and it seems unlikely that they will be effective in dealing with textures. Knill’s and Pentland's method make no smoothness assumption and will therefore be unaffected.

Knill’s and Pentland’s techniques represent the only techniques which we have been able to identify as being suitable for this problem. Although these schemes are both based on linear filtering, they do, however, differ in their derivation: Pentland’s scheme is deterministic, while Knill’s is probabilistic. Knill's technique is unsuitable for classification by definition: in order to employ *a priori* knowledge, the texture type must already be known.

In fact, closer examination of Pentland’s technique shows it to be almost identical to the independently developed Chantler’s filters. Although developed with different aims in mind, both techniques are derived from Kube and Pentland's linear model of the imaging process. The differences in the techniques are due to two factors: the different aims of the techniques and their treament of noise. In his F1 filter Chantler assumes noise to be white, though in the F2 filter he only assumes it to be isotropic and adopts an empirical approach to its radial frequency characteristics. Pentland assumes the noise spectrum is proportional to the image spectrum and develops his filter accordingly. The filters also differ in the purpose; Chantler only seeks to remove the directional effects of illumination and implicitly recovers the magnitude of the surface derivatives. Pentland aims to recover the surface height and therefore includes an $i\omega$ term to perform the integration of surface derivatives in the frequency domain. For our purposes these techniques are effectively the same, and will suffer from the same difficulties. We believe this equivalence to be quite revealing: Chantler's scheme is recovering a physically meaningful quantity— the magnitude of the surface derivative.

7.5 A Model-Based Approach Using Photometric Stereo

7.5.1 Model Based Classification

The model based approach is designed to anticipate the feature space distributions by modelling the underlying physical and analytical processes of imaging and feature extraction (*Figure 7.5.1*). This technique forms a spatial model of the training surfaces in the primary training stage. The classification process proper, begins with secondary training, when the recovered surfaces are synthetically rendered under the experimental conditions, and the resulting images form the basis of training. If the model’s components, the reflectance function and the surface description, are sufficiently accurate we should be able to obtain classification rates approaching those of the 'best case' classification, i.e. based on training on images obtained from surfaces illuminated at that tilt angle.

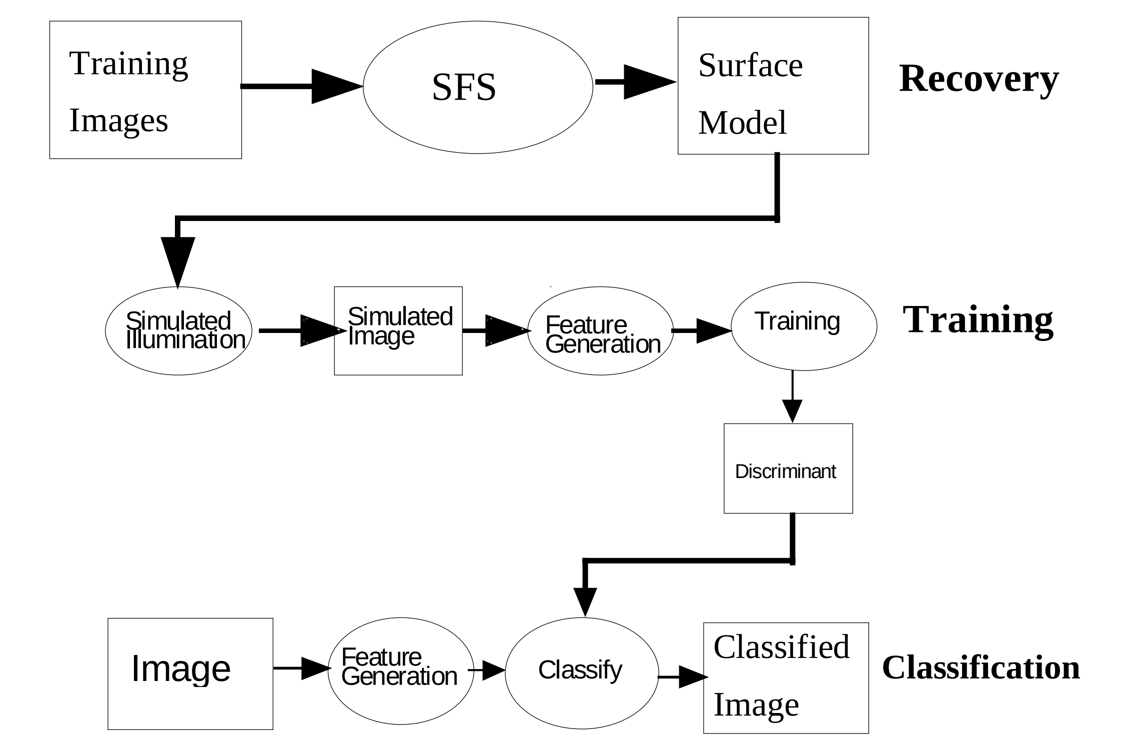


Figure 7.5.1 Simulation Based Classification

While we have an experimentally verified reflectance model for our textures (Chapter 3), we must still obtain the second component of the model: a description of the surface. To do this, a method of recovery must be adopted. As we discussed earlier, several cues to surface recovery have been investigated: focus [Noguchi94], binocular stereo [Papadimitriou95], and laser based approaches [Gross95], have been used. Let us assume that we will be able to invest more effort and exert more control in the recovery stage than in the classification stages. Ideally, we seek a technique which requires no additional hardware beyond that required for the classification.

The technique of photometric stereo allows us to form a surface description from several images of the same surface imaged under various illumination directions. It therefore seems ideally suited to our purposes since our problem is itself caused by variations in illuminant direction. In fact, surface representations acquired with photometric stereo has been used for modelling purposes by other authors. Russell

[Russell91] used photometric techniques to acquire depth maps which could then be synthetically illuminated to simulate aerial images.

7.5.2 Photometric Techniques

One approach to the SFS problem of under-determination is the use of photometric techniques [Woodham79]. These involve the use of several images of the same scene though under different illumination conditions. Each illumination condition will have its own unique reflectance map, and a given point's intensity will vary accordingly. Therefore each image will define a unique set of possible orientations for each point. If three or more images are used then the intersection of these solution sets will contain only one orientation; it is also possible to recover the albedo of the facet.

Consider a Lambertian surface illuminated from a given illumination direction, this defines a reflectance map *o(p,q), Figure 7.5.2*. If we are given a facet's intensity under these conditions, we may conclude only that its surface derivatives lie on a particular contour on the *p-q*-plane. This is essentially the fundamental problem of shape from shading; single image techniques use constraints in the spatial domain to resolve this ambiguity. Photometric techniques on the other hand use several images, imaged under different illumination conditions, with their own specific reflectance map. We therefore have a set of contours. The facet's derivatives are invariant and will lie at the intersection of these contours. Since two contours may overlap at more than one point we require three images to resolve ambiguities in all cases.

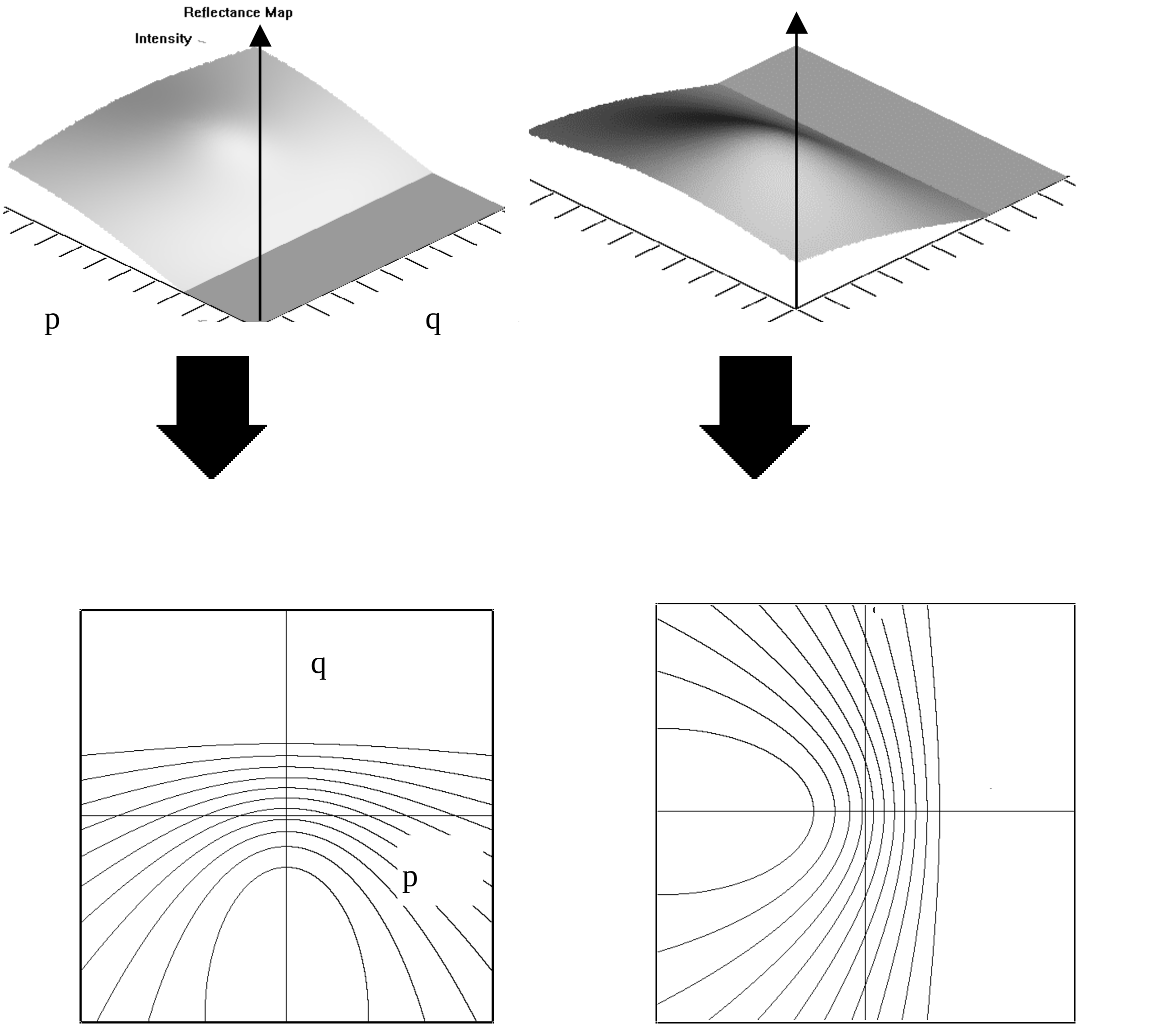


Figure 7.5.2 Reflectance Maps o(p,q) at τ 0° and 90°

More rigorously, consider Lambertian reflectance, where **L**,**S** and *I* represent the illuminant vector, normal vector and facet intensity respectively.

I=**L.S** where $\mathbf{L} = [l_1, l_2, l_3]$, $\mathbf{S} = [p, q, 1]$

Now, consider the same facet illuminated three times with different illuminant vectors: define: $\mathbf{I_{ph}} = [i_1, i_2, i_3]$ and the combined illuminant vector $L_{ph} = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$

Reiterating

$$I_{ph} = \mathbf{S}.\mathbf{L_{ph}} \qquad \text{and} \qquad \mathbf{I}_{ph}.\mathbf{L}_{ph}^{-1} = \mathbf{S}$$

While this may be solved numerically, most implementations use three dimensional lookup tables.

In this section we have proposed a simulation based scheme which is designed to counteract the effect of tilt on classification by predicting the location of feature space distributions for given tilt conditions. This is achieved by using a surface and a reflectance model to generate training data which is appropriate to classification under the specified illumination conditions.

The surface model is an important component in the scheme. We have identified, and given a brief description of, a technique which is capable of recovering the required model, with relatively little overhead in terms of additional hardware and training.

7.6 Conclusions

In this chapter we considered three techniques proposed by Chantler [Chantler94] for the reduction of tilt induced misclassification, as well as surveying the field of shape from shading and proposing a novel simulation based technique.

Chantler's first technique attempted to make the classifier robust by training the classifier over the range of illumination conditions which occur during the classification sessions. Examination of the feature means for two test textures showed that it is not safe to employ a single threshold to discriminate between textures throughout the tilt range. This result also showed that, in practice, the tilt angle must be known before the classification may be undertaken.

Chantler's second proposal advocated the use of a family of discriminants indexed by tilt angle. Application of this technique to our test montages showed that, while the technique is effective at reducing tilt degradation at the training angles, the misclassification rate quickly increases as the illuminant tilt moves away from these angles. From the performance of the classifier on our montages, we believe that a new discriminant must be designed at intervals not exceeding 20° of tilt.

Chantler also defined a system of filters which are based on inverting the directional effects modelled by Kube and Pentland, [Kube88]. We found application of this scheme to be limited for two reasons. The first is the requirement that test surfaces are isotropic; this condition is associated with the implementation of the filter estimation, and it is possible to suggest an alternative scheme using more than one estimation image to generalise the technique to directional surfaces. A more serious drawback is that of generality; the characteristics of a filter are specific to a particular surface and any general filter must be sub-optimal. In our experiments we found that, even with the isotropy restriction, the sub-optimal filter is unable to stabilise the features to a satisfactory degree. We do not pursue this approach further.

The field of single image shape from shading was reviewed. Mainstream SFS algorithms use a smoothness constraint and are therefore unsuitable for texture classification. Two techniques have been applied to fractal surfaces, however, implicit in both is the requirement for *a priori* information as to the nature of the surface. This, by definition, rules out these techniques from further investigation.

Finally a model based technique was proposed. This uses an estimate of the surface derivative field obtained using photometric stereo to predict the observed texture under specified illumination conditions. This prediction is then used to train the classifier for those illumination conditions. The next chapter will describe the evaluation of this technique.

Chapter 8

A Simulation-Based Approach To Tilt Effects

8.1 Introduction

In Chapter 6 it was shown that changes in illuminant tilt can induce classifier failure. In Chapter 7 we considered techniques to reduce the effect on classification: the field of shape from shading was surveyed for possible techniques, though none was found to be suitable. Also reviewed, and evaluated where appropriate, were the proposals advanced by Chantler [Chantler94]. From our experiments we concluded that tilt direction must be known *a priori* and that the linearity constraints of the frequency domain are too stringent for spectral compensation techniques to be applied to a wide range of textures. Finally a simulation based technique was proposed which is defined in the spatial domain and circumvents the constraints of the frequency domain. In this chapter, an implementation of the simulation-based approach will be proposed and evaluated.

We begin this chapter with a review of the literature describing photometric techniques, a simple method for surface derivative recovery under the Lambertian constraint is then described. Synthetic surfaces and the noise model derived in Chapter 4 are then used to assess the accuracy with which this technique can recover the surface derivatives. The effectiveness with which the system can simulate surfaces illuminated from arbitrary tilt angles is then assessed and finally the effectiveness of the system in dealing with classification under varying illuminant tilt angles is evaluated.

8.2 Structure of This Chapter

In Section 8.3 photometric techniques are reviewed as a mechanism for obtaining the surface derivative fields. A simple, though sub-optimal, technique is proposed. The structure of the remainder of this chapter is defined by the constraints on our experiments.

An investigation into the robustness of the model-based technique to different degrees of surface roughness, non-Lambertian reflectance characteristics, image blur, and white noise is most conveniently conducted using simulation. The effect of these variables is investigated in section 8.4.

The performance of the algorithm on *real* data is investigated in section 8.5. The model-based technique is evaluated at each stage of its operation. The accuracy of surface recovery is evaluated using a controlled, smooth surface. The accuracy of image prediction is evaluated both in terms of the "image signal to residue ratio" and in terms of the closeness of the predicted feature distributions to those obtained from the original data. Finally, the classification accuracy of the model-based technique is assessed in section 8.6 using three montages of real textures.

8.3 A Photometric Implementation

In the previous chapter photometric stereo was identified as a suitable approach to the recovery of surface characteristics and the reasoning behind the technique briefly outlined. In this section we review the literature associated with this area of work in order to identify appropriate implementations as well as the practical issues arising from the technique's use.

In the context of a model-based classifier, the accuracy of surface recovery is of secondary importance to the accurate prediction of training images. This relaxation of the requirement for surface accuracy allows other factors to be taken into account in the choice of recovery algorithm. In this case, the ease of implementation and the speed of operation become significant characteristics of the algorithm.

8.3.1 A Review of Developments in Photometric Stereo

Although several authors have attempted to reduce the number of images required [Onn90][Lee84], most work in photometric stereo has focused on the generalisation of the reflectance map to include specular and other non-Lambertian components. We begin by considering work that assumes calibration data is available from which a reflectance map may be calculated in a straightforward manner.

A Priori Known Reflectance Maps

Coleman and Jain [Cole80] used a fourth image allowing specular facets to be removed from the estimation process by thresholding albedo estimates. Using four images, they were able to obtain four estimates of both the facet normal and albedo. If none of the albedo values exceeds a threshold, the normal is taken to be the average of the four estimates. If, on the other hand, any of the albedos exceeded the threshold, the normal was based on the combination of images that gave the lowest albedo estimate.

Cho and Minamitani also uses an albedo thresholding scheme, though they do so with only three images [Cho93]. Assuming the surface is homogeneous, we would expect the reflectivity of facets to follow a normal distribution. Facets with estimated reflectivities greater than two standard deviations above the distribution mean are classified as being specular. The pixel in the image with greatest intensity is then readjusted by rescaling with a modified reflectivity.

Tagare and Figueirdo develop a photometric theory for the case of diffuse though not necessarily Lambertian reflectance functions [Tagare91]. Tagare defines a class of reflectance maps he denotes as ‘m-lobed’; in the model he discusses there are three lobes: the normal Lambertian component, the forescatter lobe corresponding to glossy specular, and a backscatter lobe. He also develops the photometric theory for this model.

Rajaram et al. adopt a more pragmatic approach and seek to develop a noise tolerant, stable and generalised photometric system based on a neural network [Rajaram95]. A back propagation network is trained on 5x5 patches of a Gaussian sphere. This empirical approach allows reflectance functions more complex than the Lambertian to be used; referring to Tagare’s m-lobed model, he states that the optimum complexity of the hidden architecture increases with the number of lobes present in the reflectance model. Rajaram presents experimental evidence that the technique is superior to Woodham’s analytical method [Woodham81].

Iwahori and Woodham [Iwahori95] applied PCA and two neural nets for the situation where illumination sources are close to the viewing direction. Both nets are trained on a calibration sphere: the first learns the intensity to derivative mapping and the second the derivative to intensity mapping. In operation the first net estimates the surface derivatives, while the second takes these estimates and forms an estimate of the intensity for the facet. Comparison of the real and estimated intensities gives a measure of the reliability of the estimate—a large error being indicative of effects such as cast shadows or interreflection.

Joint Estimation of the Surface and Reflectance Map

A more difficult problem is that of estimating a surface with an unknown reflectance map. Nayar et al use a linear combination of Lambertian and an impulse specular component [Nayar90].

In [Iwahori94] Iwahori and Woodham use the analysis/synthesis pairing of neural nets discussed earlier [Iwahori95]. However, if there is no calibration data available they adopt the following procedure. They define 25 instances of Phong's rendering model (see Watt pp.96-100 for a description of Phong's model) with various parameter values as starting points for the back-propagation algorithm and use the difference between the estimated and actual intensities as a penalty function. Finally, they select the net with the lowest mean square error and use its surface estimate [Iwahori94].

Kay and Caelli constrain the reflectance map to a simplified Torrance-Sparrow (TS) map with additional Lambertain and mirror-like specular terms, though the latter is in fact modelled as a Gaussian function [Kay95]. Using a large number of images (15-75), they use non-linear least squares to estimate the relative strengths of the Lambertian, glossy and mirror components, the roughness parameter associated with the TS model as well as the surface normal. Where the model is ill-defined, they describe a framework for the selection of an appropriate sub-model. Interestingly, Kay does not make any assumptions as to the homogeneity of surface properties, instead estimating the reflectance map for each facet.

Tagare and Figuierdo [Tagare90] use an 'm-lobed' reflectance function and concludes that of the order of ten illumination sources may be required for the joint estimation problem.

Solomon and Ikeuchi use a four light scheme similar to Jain's to identify specularities though they do discuss the case where a facet is illuminated by less than four sources [Solomon96]. Like Cho and Minamitani they opt for a statistically based albedo threshold ($\pm 3\sigma$ in this case) though as with Cole and Jain they then resort to the lowest albedo normal estimate. Using the surface normal estimate he is able to estimate the strength of the specular component for that facet, and also the nature of the reflectance map's specular lobe. Using the TS framework they are able to infer the sub-pixel roughness of the surface.

8.3.2 A Simple Photometric Scheme

The schemes discussed above obtain increasingly accurate approximations to the reflectance map of test surfaces and allow application of photometric techniques to a wider range of materials. However, we have already verified that our test surfaces can be accurately modelled as being Lambertian. In the context of this thesis, where surface reconstruction is a means to the end of image prediction, it may be viable to use a simple scheme. We note that such a scheme may be sub-optimal and almost certainly inferior to the schemes discussed above, however it is extremely simple to implement and may give acceptable results for the purposes of image prediction.

We therefore propose the following simple photometric scheme.

Consider an illuminated surface, whose intensity corresponds to the following equation:

$$i(x,y) = \lambda\rho\left(\frac{-p\cos\tau\sin\sigma - q\sin\tau\sin\sigma + \cos\sigma}{\sqrt{p^2+q^2+1}} \right) \tag{8.3.2a}$$

If the surface is illuminated from τ =0°, 90° or 180° these simplify to equations 8.3.2b-d respectively:

$$i_0(x,y) = \lambda\rho\left(\frac{-p\sin\sigma + \cos\sigma}{\sqrt{p^2+q^2+1}} \right) \tag{8.3.2b}$$

$$i_{90}(x,y) = \lambda\rho\left(\frac{-q\sin\sigma + \cos\sigma}{\sqrt{p^2+q^2+1}} \right) \tag{8.3.2c}$$

$$i_{180}(x,y) = \lambda\rho\left(\frac{p\sin\sigma + \cos\sigma}{\sqrt{p^2+q^2+1}} \right) \tag{8.3.2d}$$

Adding equations 8.3.2b and 8.3.2d will produce a non-linear function of the surface derivatives (8.3.2e):

$$i_{NL}(x,y) = i_0(x,y) + i_{180}(x,y) = \frac{2\lambda\rho\cos\sigma}{\sqrt{p^2+q^2+1}} \tag{8.3.2e}$$

now dividing equations 8.3.2b and 8.3.2c by 8.3.2e we have two linear functions mapping surface slope to image intensity which are independent of albedo, ρ, and incident intensity, λ.

$$i_p(x,y) = \frac{i_0}{i_0 + i_{180}} = \frac{-p\tan\sigma + 1}{2} \tag{8.3.2.f}$$

$$i_q(x,y) = \frac{i_{90}}{i_0 + i_{180}} = \frac{-q\tan\sigma + 1}{2} \tag{8.3.2g}$$

these may be transposed to give:

$$p = \frac{1 - 2i_p}{\tan\sigma} \tag{8.3.2h}$$

$$q = \frac{1 - 2i_q}{\tan\sigma} \tag{8.3.2i}$$

The scheme therefore requires the capture of three images at tilt angles of 90° increments and the application of equations 8.3.2h and 8.3.2i to provide the estimates of the gradient field. The author notes that this simple technique is sub-optimal, and almost certainly inferior, in terms of accuracy, to the techniques reviewed in this section. The technique will perform comparatively poorly since:

- it assumes a Lambertian reflectance function,
- it ignores self and cast shadows, and
- it also ignores interreflection.

Set against these inadequacies, the scheme does provide a fast and simple implementation. It is attractive in an application, such as ours, where the absolute accuracy of surface recovery is of secondary importance to the accurate prediction of images. We conclude by noting that, where the reflectance function is near Lambertian, the technique could be used to provide an approximate initial estimate for an iterative scheme—similar to those discussed in section 7.4.2—which use more accurate reflectance models and surface constraints. We do not, however, pursue this approach as we do not believe the gains in the accuracy of image prediction would justify the effort.

8.4 Simulation based evaluation

In Chapter 7 we proposed a simple algorithm to reduce the effect of variations in illuminant tilt on classification, in the previous section we presented a simple implementation of the surface recovery component of the system. The remainder of this chapter is devoted to the assessment and verification of the technique. The primary objective of this section is to ascertain in what way, and to what degree, do departures from the ideal Lambertian model and the ideal imaging model, affect surface recovery and image prediction. The most practical method of achieving this, in a controlled and analytical manner, is the use of simulation. Simulation also represents the most convenient method of estimating the accuracy of surface recovery, the relationship between the estimated and the original surface, and how vital accurate surface recovery is to image prediction.

We investigate the technique's robustness with respect to three effects.

- *Non-Lambertian reflectance*. The recovery and prediction stages of the algorithm assume a Lambertian model; sub-section 8.4.2 considers the effect on surface recovery and image prediction of imposing a Lambertian model. Since the model's behaviour varies with facet gradient, the rms slope of the test surface will be varied and the effects notedwe will deal with the second condition;:
- *Blurring*. In sub-section 8.4.3 we introduce the first component of the noise model adopted in Chapter 4, blurring. In fact, this has much wider implications than might be immediately obvious. A surface is not bandlimited by the resolution of the imaging device and, if we assume that the surface is fractal, then we must accept that the recorded image will be a low pass approximation to the ideal image. The reflectance function is non-linear, therefore low pass filtering the image is not strictly equivalent to low pass filtering the surface. The 2.5D surface representation will be only an approximation to the low pass filtered surface. The accuracy of this approximation will have a bearing on the accuracy of image prediction.
- *Temporal noise*. In Chapter 4, we observed that if the image is sub-sampled, the time varying residue signal may be regarded as being white. In sub-section 8.4.4 we model the noise signal and observe its effect on surface recovery and image prediction.

Our approach, outlined in *Figure 8.4.1*, is incremental: investigating the effect of the empirical reflectance model in sub-section 8.4.1, we add the blur component in sub-section 8.4.3 before adding the final stage of the model (additive noise) in sub-section 8.4.4.

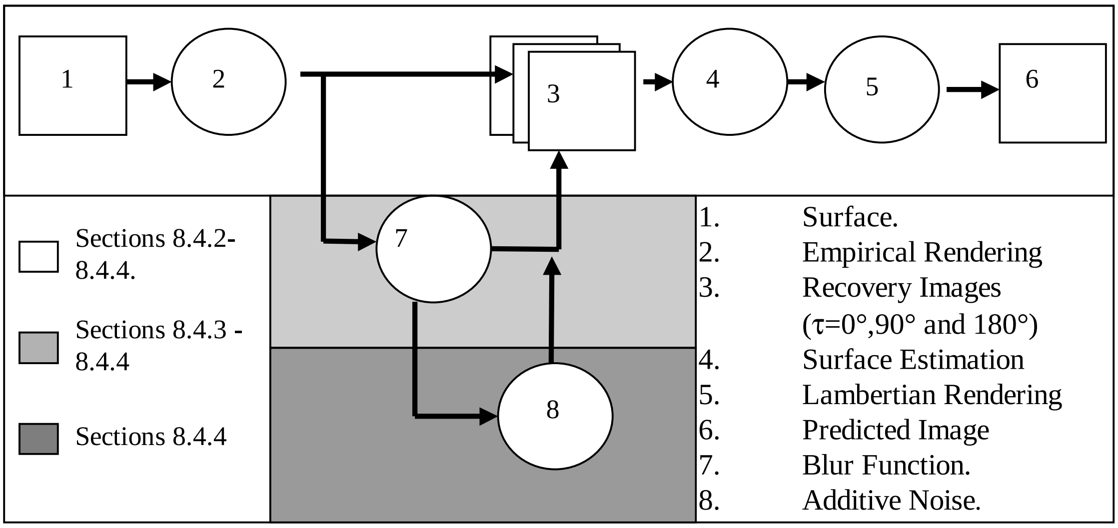


Figure 8.4.1 Experimental approach of section 8.4

While the simulation of the physical processes has been made as realistic as possible, we retain the Lambertian assumptions in the algorithm. The reasoning behind this is that the implementation of the algorithm should not require detailed knowledge of the reflectance function beyond the fact that the reflectance is predominantly diffuse in character.

8.4.1 Experimental Criterion

In this chapter we wish to assess both the accuracy of surface reconstruction and image prediction. Since reconstruction and prediction form consecutive stages in the algorithm it is of interest to see how the accuracy of the latter is related to that of the former. In order to allow comparison, a single generic criterion of accuracy is used in both cases: the signal to residue ratio. When the criterion is applied solely to the surface derivatives it will be denoted as either e_p or e_q depending on the field being observed. Where the criterion is used with images, it will be denoted as $e_{i.}$ Where the accuracy of both the image and the surface reconstruction are being assessed we denote the quantity as S/R. We apply it to the surface in the form below:

$$e_p = 10\log_{10}\left(\frac{Var(\mathbf{j.S}(x,y))}{Var(\mathbf{j.S}(x,y) - \mathbf{j.}\hat{\mathbf{S}}(x,y))}\right)$$

where

Var (x) is the variance of the process *x*.

$$\mathbf{j} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \hat{\mathbf{S}}(x, y) \text{ is the estimated gradient field}$$

Of more practical importance for our purposes is the technique's ability to predict the image of a surface illuminated from an arbitrary tilt angle, given a photometric estimate of the surface derivatives. We apply the criterion to measure the accuracy of the predicted image. For real textured surfaces we do not know the surface derivative field, so this is the first measure of accuracy we can apply to the process. In the case of the image quality measures, the noise effects will be applied to both the recovery and the evaluation images, again to allow comparison with the next section on real surfaces.

$$e_i(\tau) = 10\log_{10}\left(\frac{Var(i(x,y))}{Var(i(x,y) - \hat{i}(x,y))}\right)$$

8.4.2 Non-ideal reflectance

Implicit in the scheme is the assumption that the rendering can be described using equation 8.4.2.

$$i(x,y) = \frac{-p\cos\tau\sin\sigma - q\sin\tau\sin\sigma + \cos\sigma}{\sqrt{p^2 + q^2 + 1}} \qquad (8.4.2)$$

We make three criticisms of the scheme based on this fact:

1. although the reflectance function is certainly diffuse, in Chapter 3 we noted that it is not perfectly Lambertian,
2. by choosing to describe the reflectance function using equation (8.4.2), we imply that for certain facet orientations, a negative intensity will be observed; this is clearly not the case, and
3. equation (8.4.2) makes no allowance for the non-local effect of cast shadows.

The first effect is investigated using a synthetic fractal surface rendered with the empirical reflectance map. The degree to which the second and third effects occur is dependent on the roughness of the surface. These effects will be observed simultaneously, this sub-section corresponds to stages 1-6 of *Figure 8.4.1*.

Non-Lambertian Reflectance

As an initial test of the accuracy of this method we estimate the derivative of an isotropic fractal (β=3.0) surface with *rms slope* 0.35 illuminated from a slant angle of 60°.

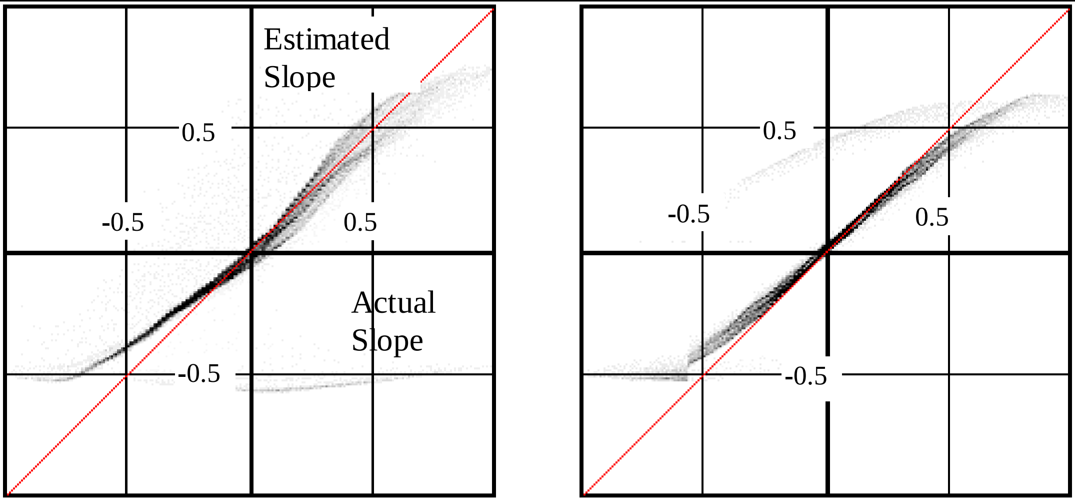


Figure 8.4.2 The accuracy of estimation of facet p-derivative (a) and q-derivative (b)

The actual and estimated derivatives of each facet are used to generate a scatter plot; the ideal case being a line of gradient 1 passing through the origin.

From *Figure 8.4.2* it is clear that there is an approximately linear relationship between the estimated and actual slopes. The estimated q-derivatives are more closely correlated with the actual derivatives than the p-derivatives. This is presumably due to the fact that the colinear recovery images were illuminated by sources parallel to the sense of the q-derivative, though the mechanism is unknown. As expected, inaccuracy increases for larger slopes; at both extremes of the range the magnitude of slopes is increasingly underestimated. However, for large positive p-derivatives there is a more striking effect: an increase in the spread of estimates, showing a much less deterministic relationship between estimate and actual slope as other factors come into play.

The aim of this technique is to predict the image, and ultimately the feature distributions, of a rough surface imaged under arbitrary tilt conditions. We now assess how the quality of the image prediction varies throughout the tilt range. That is the recovery images remain unchanged at 0°, 90° and 180°, but the tilt of the illumination used in image prediction (block 5 *Figure 8.4.1*) is varied between 0° and 360° in 10° steps. The predicted images are compared with those obtained by applying "empirical rendering" (block 2, *Figure 8.4.1*) using the signal to residue ratio (*Figure 8.4.3*).

Figure 8.4.3 The accuracy of reconstruction with varying tilt.

In fact, the variation is considerable, maxima occur at τ=0°,90° and 180°. In the work that follows two values of image S/R (e_i), at τ= 0° and 270°, will be quoted. Due to imaging constraints the experimental data images were only captured in the range 0° to 180° we would therefore expect the actual accuracy to be significantly better than that obtained by simulation at τ=270°. The selection of two tilt angles that are aligned to the grid axes allows shadowing to be implemented without supersampling and its associated artefacts.

Cast and Self Shadowing

The technique does not take into account the clipping of negative values returned by the Lambertian function; self-shadowed areas are, in effect, assumed to have negative intensities. The model-based technique, being inherently 2.5d, also fails to account for cast shadows in either the recovery or the prediction phases; we therefore incorporate cast shadows into the rendering algorithm. In Chapter 3 we saw that the shadowed areas do not have zero intensity—a value of 40 is more realistic in our scale.

It is interesting to consider the effect of shadowing in practice. *Figure 8.4.4* shows a stone surface, which was not considered in the main body of the thesis due to its violation of the random phase condition. The sample itself is sedimentary and formed from several layers of sandstone, which lie in the plane perpendicular to the camera axis.

Several layers are exposed and the regions of transition between the layers have an associated abrupt change in height. These abrupt changes are symptomatic of a phase rich surface. The texture displays two symptoms of non-random phase: asymmetry of the textural properties of the two images separated by 180°, and shadowing.

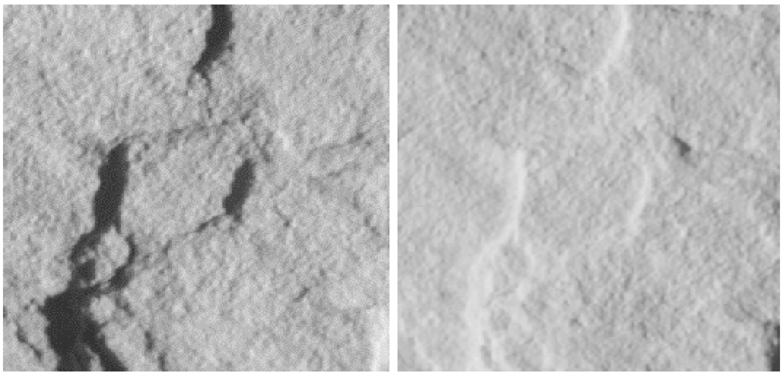


Figure 8.4.4 Lamina Texture Illuminated from Tau 90° and Tau 270°

Due to these characteristics, this sample provides an interesting test case to see how well the prediction algorithm performs in areas that are shadowed in one or more of the recovery images. Surprisingly, the image is modelled quite well for the shadowed image, however, when the shadowed areas are illuminated the algorithm performs poorly as shown in *Figure 8.4.5*.

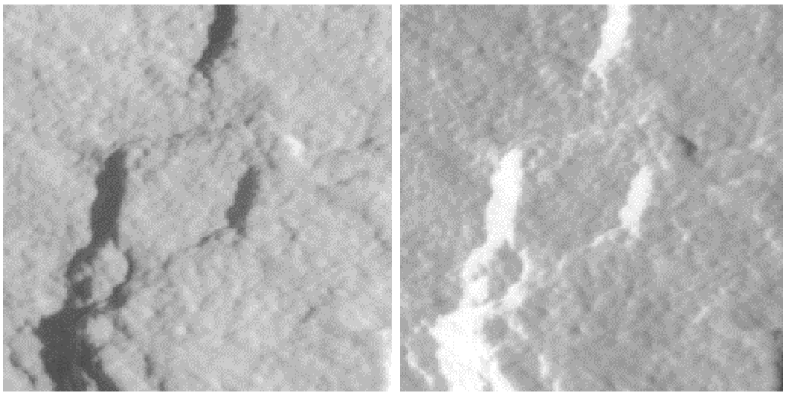


Figure 8.4.5 Simulated Lamina texture illuminated from Tau 90° and 270°

As *rms slope* increases, more and more facets are affected by cast or self-shadowing, (*Figure 8.4.6*) and the modelling error becomes more significant.

.

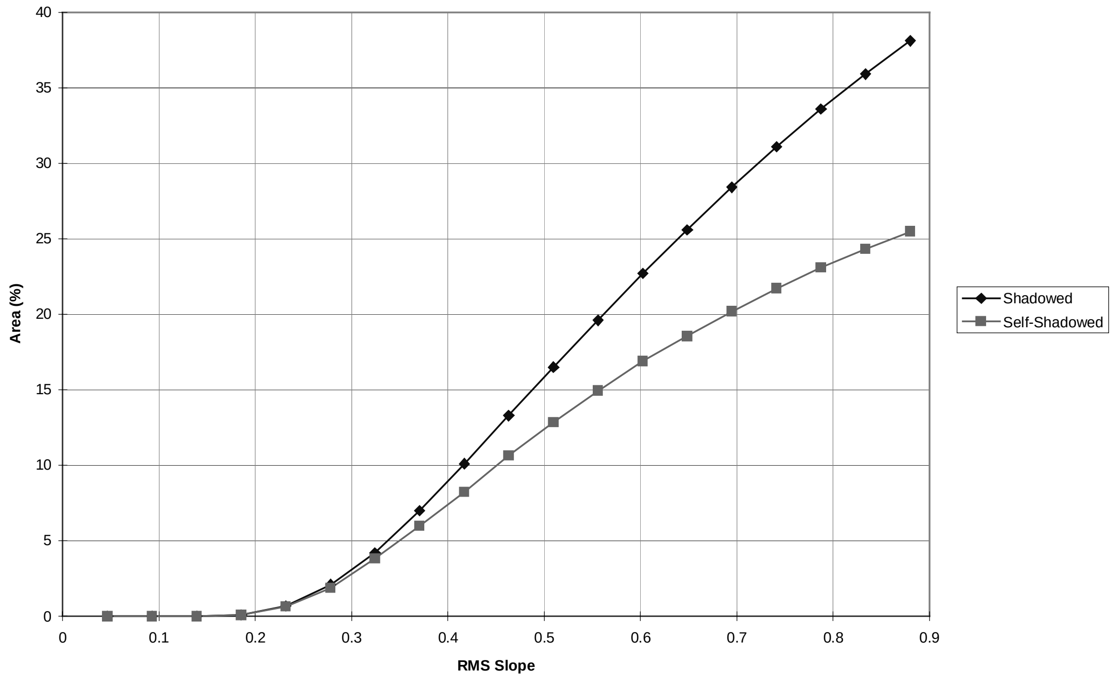


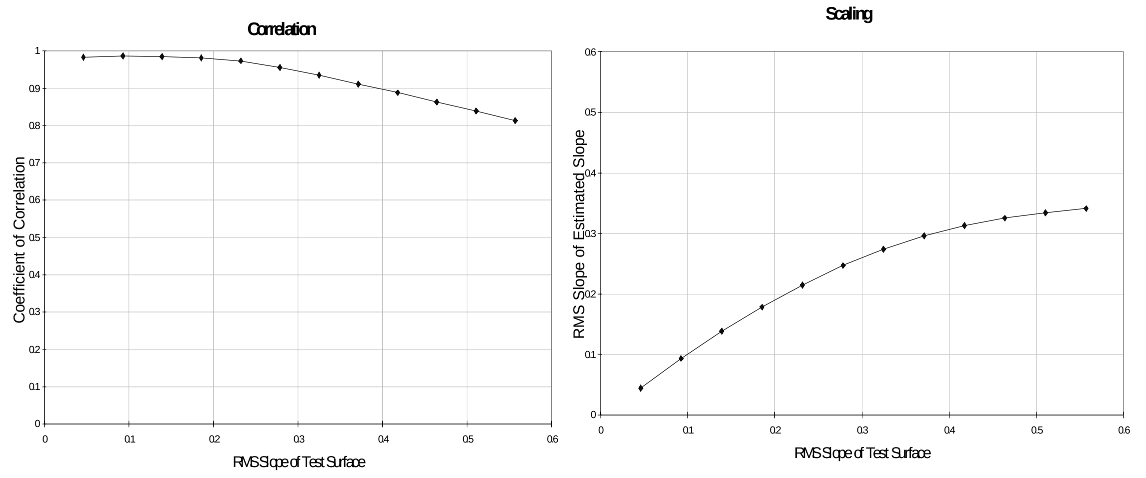
Figure 8.4.6 The effect of slope on the degree of shadowing of an isotropic fractal surface.

It follows that the effect of surface roughness on the accuracy of derivative estimation is significant and must be investigated. Since we are primarily concerned with the performance of the technique with textures, we assess the relationship between estimated and actual slopes using statistical techniques which take account of the distribution of slopes. At this stage we resolve the question of accuracy into two criteria:

(1.) There should be a linear relationship between the estimate and the actual slope, *and*

(2.) The ratio of estimate to actual slope should equal unity.

The first criterion represents the linearity and spread of the scatter plot distribution and will be tested by measuring the degree of correlation between the estimate and the actual slope. The second will be assessed by comparing the standard deviation of the derivative fields.



(a.) (b.)

Figure 8.4.7 The relationship between estimated and actual derivative fields of an isotropic fractal surface. Experiment corresponds to stages 1-4 of Figure 8.4.1

The correlation graph shown above (*Figure 8.4.7*a), shows the estimate and actual surface to be very well correlated for surfaces with low rms. slopes. The degree of correlation does fall significantly as the surface roughness increases, however, even at the extreme of the experimental range there is a significant level of correlation. The scaling function (*Figure 8.4.7*b) is close to unity for low slopes, however as surface roughness increases, the degree of underestimation also increases and is significant for even moderate slopes.

If we now phrase the question of accuracy in terms of S/R, we may consider the accuracy of both the surface and image. In these terms the surface is modelled very accurately for low slopes, however, the surface S/R ratio rapidly falls with increasing rms slope *Figure 8.4.8*.

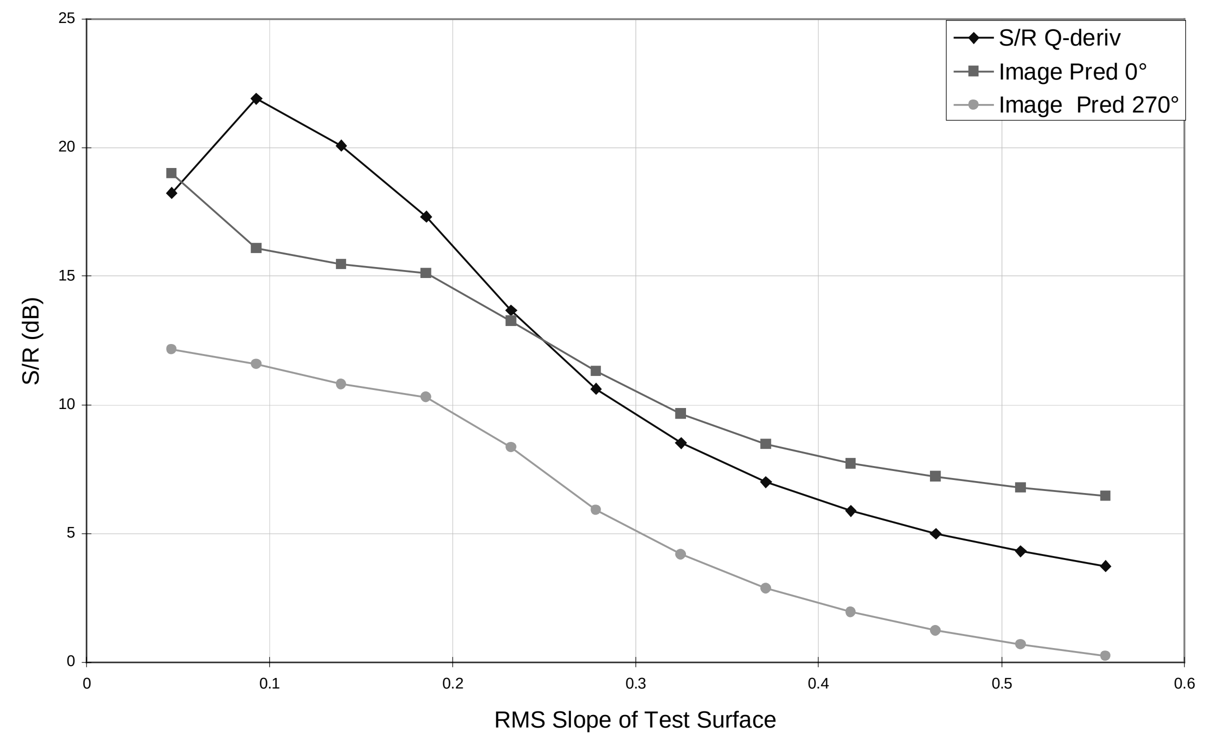


Figure 8.4.8 The Effect of Surface Roughness on Surface and Image Prediction.

The quality of the predicted image also undergoes a degradation with increasing slope, however, the degradation is more gradual, and for m_{rms}>0.25 the image S/R at τ=0° is actually higher than that of the surface; though the same is not true of the τ=270° prediction. The process of Lambertian rendering seems to de-emphasise the errors present in the surface estimate. We note two points from this:

(1.) the algorithm can predict images to an accuracy greater than 10dB for surfaces with rms. gradients less than 0.25 in spite of imposing the Lambertian model on a non-Lambertian system, and

(2.) accurate image prediction of a recovery image does not imply accurate surface recovery, although the accuracy of an image that is extrapolated out of the recovery range does give a more reliable indication

The surface slope clearly imposes limits on the surfaces to which the technique can be applied. However, the technique is able to maintain a prediction accuracy of greater than 4dB for even the roughest surfaces considered in this thesis. Another interesting point to emerge is that while surface recovery does seem to be sensitive to changes in the degree of surface roughness, image prediction, for recovery images at least, is more robust.

8.4.3 The Effect of Blurred Images

As stated earlier, any practical application involving the imaging of rough surfaces will almost certainly involve the loss of high frequency information. In this sub-section, we attempt to model this, in order to ascertain the effects this will have on surface recovery and image prediction. We now apply the Gaussian blur function adopted in Chapter 4 with the σ_b parameter set to 0.02 in accordance with the experimental findings.

The effect of blurring on surface and image prediction

We again plot the estimated and actual derivatives on a scatter plot (*Figure 8.4.9)*— the desired result being a linear mapping. The observed result shows there is not a simple linear mapping between the actual and estimated slopes for either of the derivative fields, though a degree of correlation is apparent. The general trend is an underestimation of the slope.

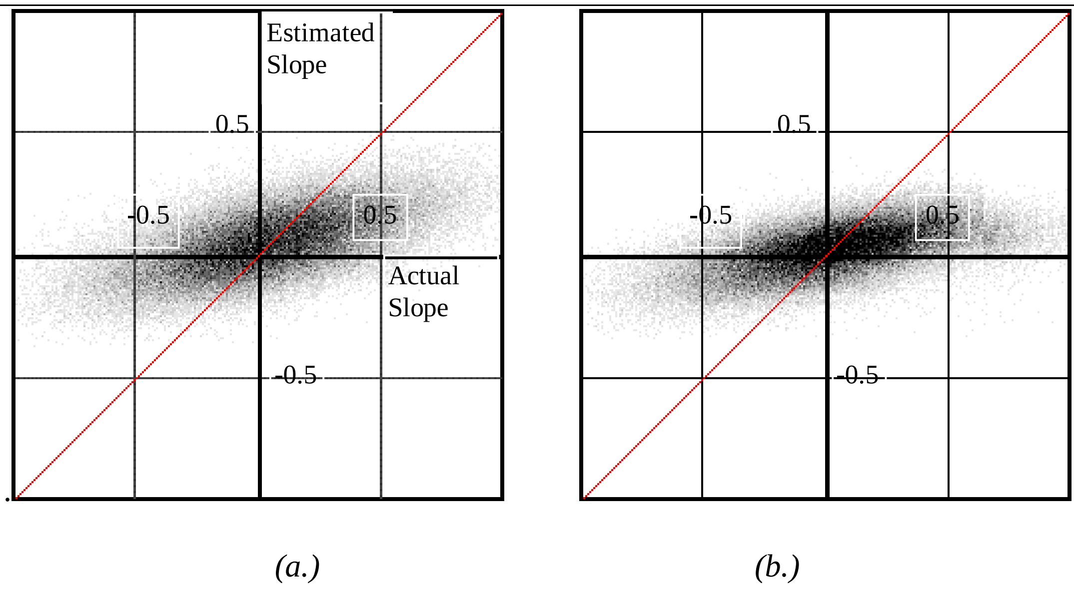


Figure 8.4.9 Scatter plot of actual and estimated (a) p-derivatives (b) q-derivatives estimated from blurred images.

Whereas blurring itself is modelled here as a linear operation, the effect of the non-linear reflectance function (block 2, *Figure 8.4.1*) means that the overall system is not linear. Using blurred, i.e. low pass filtered, recovery images will not necessarily lead to a simple low pass filtering of the surface estimate. Since the degree of non-linearity will vary with slope, we cannot easily predict the effect of scaling on either the surface estimate or the image prediction. We therefore consider this effect experimentally.

In *Figure 8.4.10* we plot the scaling and correlation coefficients for both the blurred and unblurred images. Comparison of the correlation functions shows that the surface estimates of the blurred case are much less related to the actual slopes than in the unblurred case. It is worth noting, however, that the decline of accuracy with increasing slope is less pronounced than with the unblurred version. The scaling coefficient for the blurred estimate is much lower than the unblurred case and the desired figure of unity. Recovery based on images blurred to the degree experienced in experiment will lead to noisy underestimates of the surface.

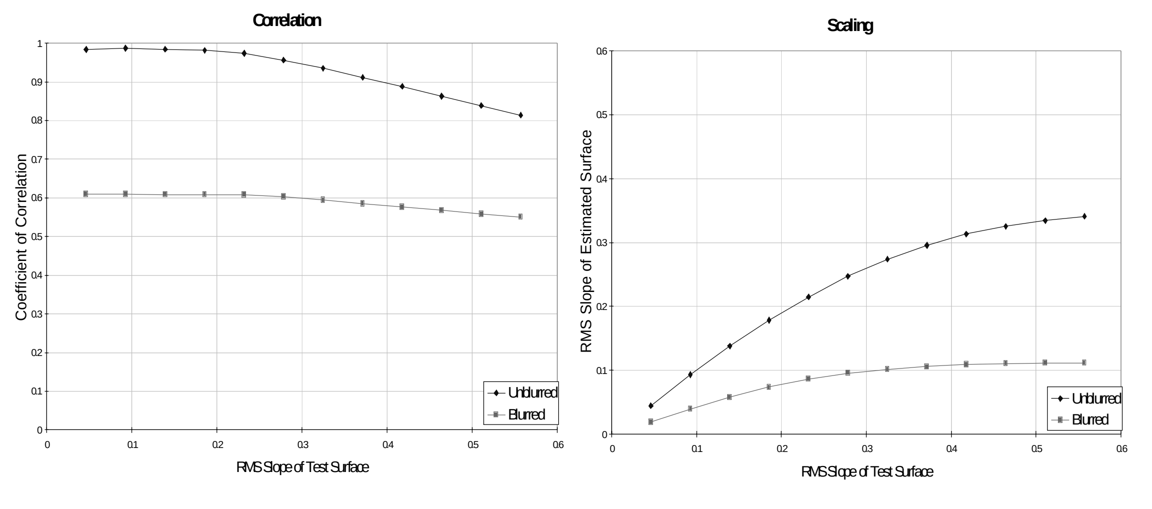


Figure 8.4.10 Relationship between actual and estimated rms slope from blurred images.

If we consider the effect of blurring on the S/R of the uncompensated surface, it is clear that this represents a very poor estimate of the actual derivative field, *Figure 8.4.11*. Blur is obviously a very significant obstacle to surface recovery of rough surfaces. In light of this, the performance of the image S/R is surprising. For low slopes the predicted image achieves an accuracy which is similar to the unblurred case, and while it is significantly lower for rougher surfaces, its performance is remarkable given the inaccuracy of the surface estimate.

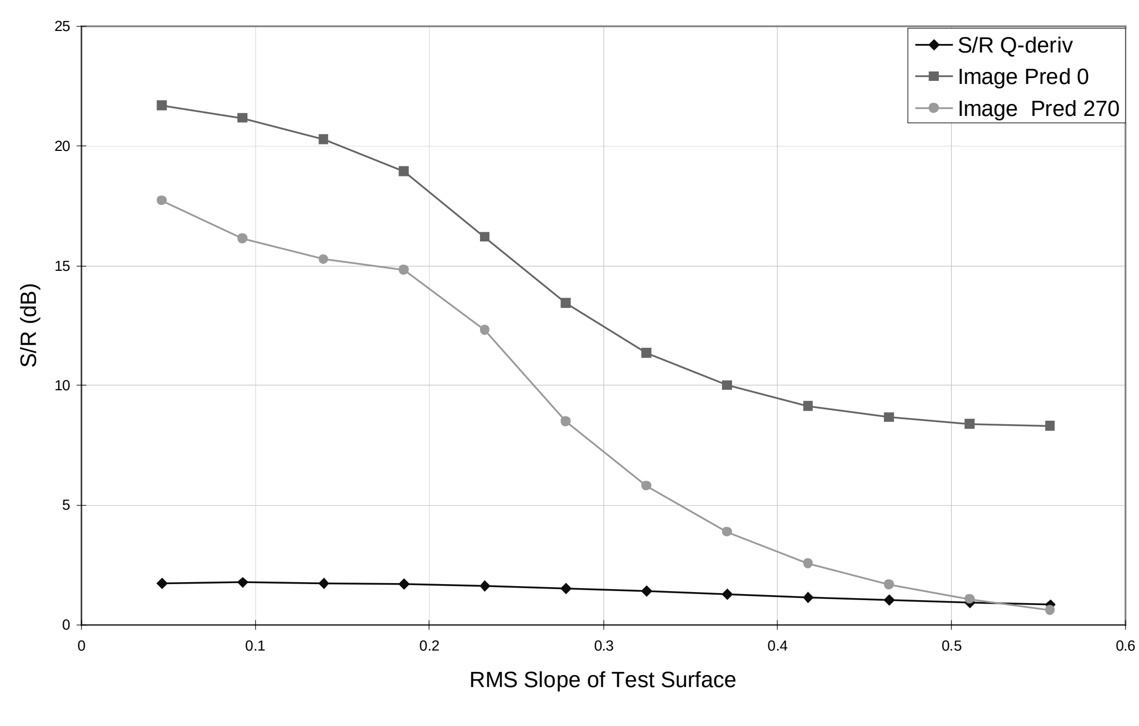
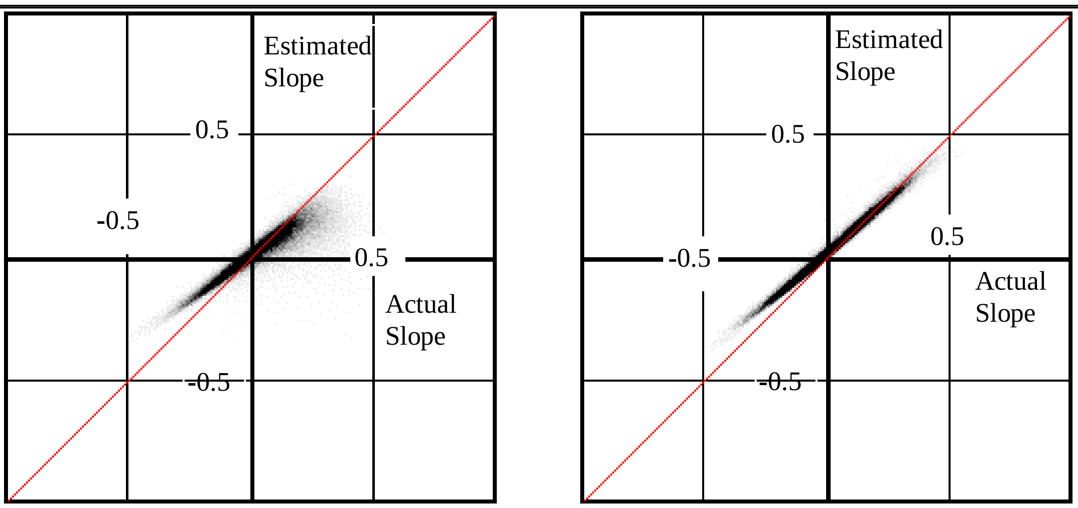


Figure 8.4.11 The effect of blurring on surface and image prediction.

This result leads us to make two statements:

1. A good image reconstruction does not imply an accurate model of the underlying surface if blur is present and uncompensated.
2. Conversely, an accurate surface model is not always necessary in order to accurately predict the image.

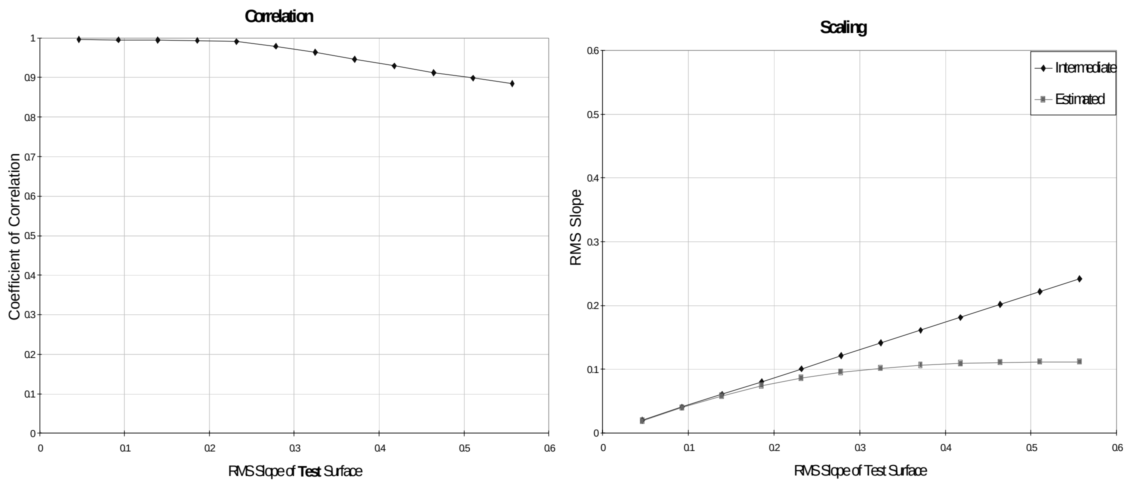


(a.) p-derivatives *(b.) q-derivatives*

Figure 8.4.12 The effect of image blur on surface reconstruction.

Does the estimated surface approximate a low pass filtered version of the true surface?

Earlier in this chapter we stated that due to non-linearity we could not guarantee that the surface estimated from blurred recovery images is a low pass filtered version of the original surface. We now question this statement. In *Figure 8.4.12* we plot the estimated facet slopes against the slope of the corresponding facets of a low pass (or intermediate) version of the original surface. The cluster shows a significant improvement in the degree of spread, underestimation and linearity compared with *Figure 8.4.9*. We therefore proceed to investigate the effect of surface roughness on this mapping.



(a.) (b.)

Figure 8.4.13 Comparison of statistics of a low pass filtered (intermediate) surface with those of a surface estimated from blurred images.

In *Figure 8.4.13(b)* we plot the rms slope of the intermediate and estimated frames. Unsurprisingly, both have much lower slopes than the original surface. However, since blurring is a linear operation, the rms slope of the intermediate frame is linearly related to that of the original test surface. In contrast, the estimated surface, which also includes the non-linear reflectance function, shows a much less linear relationship with the slope of the test surface and undergoes a more severe underestimate of slope.

In *Figure 8.4.13a* we plot the coefficient of correlation between the intermediate surface and the estimated surface slopes. The fields are very well correlated, even for rough surfaces. In fact, the degree of correlation is even higher than for the unblurred case, due to the fact that the mapping which occurs between the intermediate and estimated fields occurs over a range of much smaller slopes due to the underestimation shown in *Figure 8.4.13b*.

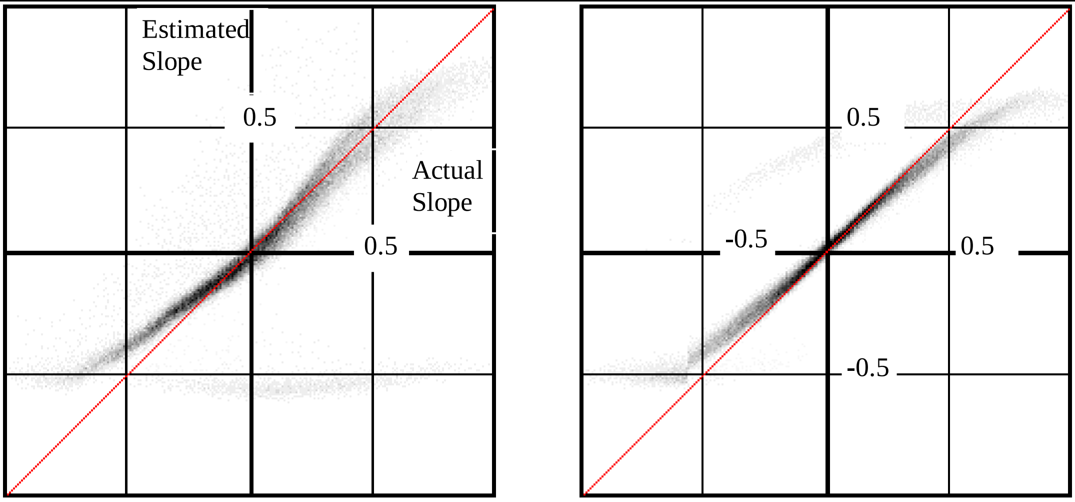
The effect of image blur on the accuracy of image prediction, while significant, is not critical. The effect on derivative recovery, as measured by our criteria, is so serious as to call into question our claim that our simulated images are estimated from a basis which is physically meaningful.

Interestingly, despite the non-linear character of the rendering function, the estimated derivative fields *do* approximate those obtained from a low pass filtered surface. We may therefore introduce the concept of a 'dual' of the surface. This intermediate, or pseudo-surface provides a link between the real surface and the perceived image. The low pass relationship of the intermediate to the real surface is analogous to the bandlimiting assumption made in Chapter 2; although in this case it forms a less arbitary, more physically-based alternative to the ideal filter assumed in the earlier theoretical work.

8.4.4 The Effect of Temporally Varying Additive Noise

We now model the effect of temporal noise by adding different realisations of a white noise process to each recovery image to give a signal to temporal noise ratio of approximately 25dB. In *Figure 8.4.14* a scatter plot of the estimated derivatives against the actual derivatives is plotted. In the interests of clarity, the blur component is temporarily suspended for the scatter plot only. The derivative estimate is relatively unaffected by this level of temporal noise, though comparison with *Figure 8.4.2* does show a slight increase in the spread of the clusters. However is clear that this does not have the same magnitude of effect on slope estimation as the blurring function.

In section 8.5 we will measure the S/R ratio for real images. The recovery images will also be used for evaluation. We incorporate this into our simulations, so that random noise effects will be duplicated within each of the recovery/evaluation pairs. Our simulations of S/R consequently predict a characteristic three peak variation or "W" waveform with tilt, *Figure 8.4.15*



(a.) p-derivatives *(b.) q-derivatives*

Figure 8.4.14 Scatter plots for derivative fields estimated from images with additive noise. (Blur suspended)

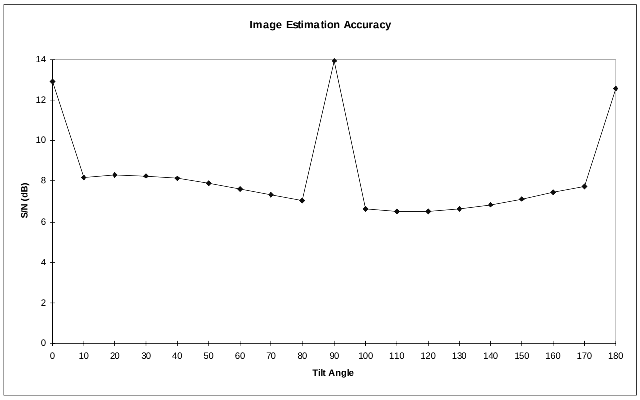


Figure 8.4.15 Accuracy of Image Estimation in Additive White Noise (9.62dB)

We now observe the effect additive white noise in the recovery images has on the closeness of the simulated images to images of surfaces illuminated at τ=50° and τ=140°. The comparison images are corrupted with noise of the same variance as the recovery images.

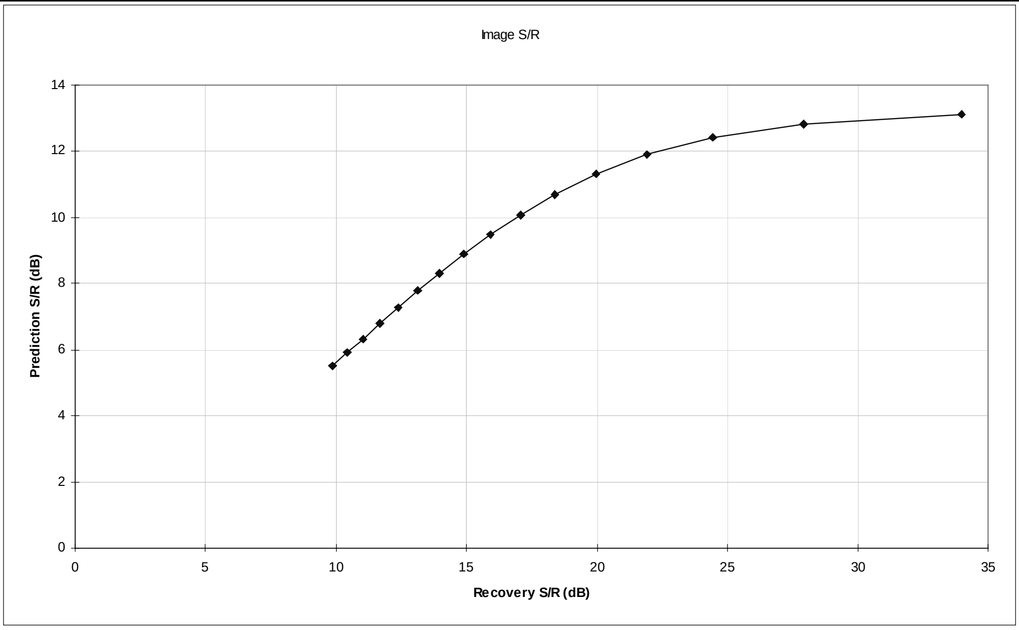


Figure 8.4.16 The effect of noise associated with the recovery image on the accuracy of prediction.

In *Figure 8.4.16* we plot the S/R ratio of the simulated image against that of the recovery images of an isotropic fractal surface with an rms slope of 0.23. The temporal S/R ratio associated with the real images is in the range 18-30dB suggesting that the prediction error, e_i, associated with this type of surface; due to the combination of non-Lambertian reflectance, blurring and temporal noise would be in the range 11-13 dB *Figure 8.4.2* shows that temporal noise of the magnitude observed in the images used in this report is much less significant than the degradation caused by blurring.

8.4.5 Discussion

In this section we have considered the effects due to the imposition of the Lambertian model, the effect of surface roughness, blur and additive noise. Of these effects blur was identified as being the most significant to rough surface recovery. That is it was found to that at levels of noise and blurring under which the real images were obtained, blurring had the most significant effect on our error measure. It is arguable that this will not be case for data obtained under different conditions, or where accuracy is measured under different criterion, e.g. errors in surface height. This notwithstanding, our investigation into both the deterministic effects, i.e. non-Lambertian behaviour and blur, lead us to make two statements:

(1.) A good image prediction does not imply an accurate surface reconstruction;

and conversely,

(2.) Accurate image prediction does not always require an accurate surface reconstruction.

Blur can be compensated for; non-Lambertian photometric estimation can be carried out, at the expense of increased memory, computational and experimental expense. However, these are not investigated since the uncompensated estimate gives image predictions of sufficient accuracy for training purposes over a wide range of surfaces.

The experimental findings show that the surface derivative vector field estimated from a series of blurred images is more directly related to the low pass filtered surface than to the original surface. The estimated vector field may be described as being the low pass filtered image of the original surface with additive, signal dependent noise.

$$\hat{S}(u,v) = G(u,v).S(u,v) + N$$

We also note that, in general, the estimated vector field will not be conservative, i.e. in general:

$$Curl\hat{\mathbf{S}} \neq 0$$

The algorithm represents a mapping from three images to a *dual space* and an inverse mapping from the dual space to an image. In the definition of the algorithm we stated that the dual had a physical meaning, i.e. as the derivative fields of the surface. The simulations show this to be a reasonable description under certain conditions. However, it was found that under other conditions it formed a poor approximation, *yet still formed a good basis for image prediction*. We may therefore say that the algorithm maps intensity triplets into a 2D space which may be said to lie between two descriptions:

1. A physically meaningful model of the surface derivatives
2. An arbitrary dual with an associated interpolation mechanism.

The point at which the algorithm lies is dependent on the accuracy of the forward and reverse models.

In fact, the difference between the two is academic, with one exception: the ability to generalise. Within the bounds of the work presented in this thesis this effectively means the algorithm must be able to maintain an adequate level of accuracy for all tilt conditions. By using the τ=270° image for comparison we have shown that this is the case for low slope surfaces.

8.5 How well does the technique work on real data?

In the previous section simulation was used to investigate the algorithm's performance on surface recovery and image prediction. In this section we evaluate the algorithm on a real data set. The accuracy of surface recovery is briefly investigated with a Gaussian sphere, however, the main thrust of this section is to investigate the accuracy of image and feature prediction. The performance of image prediction is assessed using the S/R metric introduced previously. The accuracy of feature prediction is assessed in terms of the feature space. The same classifier used in Chapter 6 for the main classification tasks is now used to discriminate between the predicted and the actual images. If the classifier is unable to reliably discriminate between the two images this will provide convincing evidence that the algorithm is able to predict the image to the required level of accuracy.

8.5.1 Surface Recovery

The Gaussian sphere used in Chapter 3 to evaluate the reflectance map is now used to assess the accuracy of recovery. Since the sphere is smooth, blur is less significant and we would expect the accuracy of recovery to be much greater than that of a rough surface.

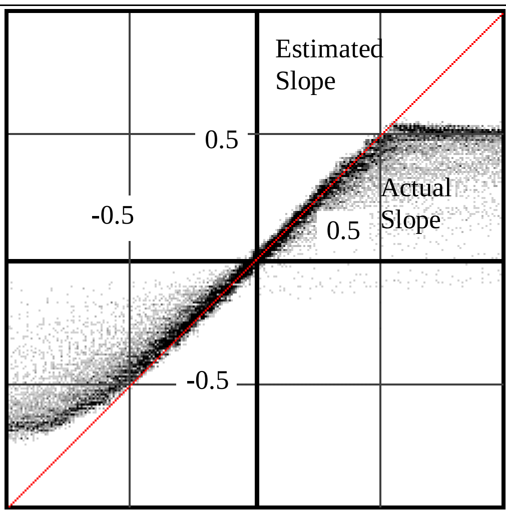


Figure 8.5.1 Accuracy of derivative estimation from real data

Both extrema show saturation, this being most prominent for large positive (i.e. self-shadowed) slopes. In fact within the slope range ±0.5, which is the region of relevance to texture modelling, the relationship is highly linear with little spread. The spread or 'vapour trail' effect corresponds to a mapping of the actual slope to a lower estimate. However, assessment of the algorithm's ability to recover the surface is of limited importance, since the real issue is whether the technique is able to simulate the observed image with sufficient accuracy and we will investigate this in the next section.

8.5.2 Image Prediction

Having obtained the surface derivatives of a field, it is possible to render these derivatives to predict the appearance of the original images. We measure the S/R ratio, as defined in the previous section, to assess the accuracy with which we can predict the appearance of members of the *Stone1*, *Stone2* and *Anaglypta* montages. The results are plotted in *Figure 8.5.2* and *Figure 8.5.7* and all display the characteristic three peak traces predicted in *Figure 8.4.15*. The anaglypta results are remarkably variable: the 'stripl' texture remains above 11dB while the 'ripl' texture falls below 2dB. We suggest that the poor results associated with the anaglypta textures, relative to the "Stone" montages, are due to the steepness of many of the facets associated with this type of surface.

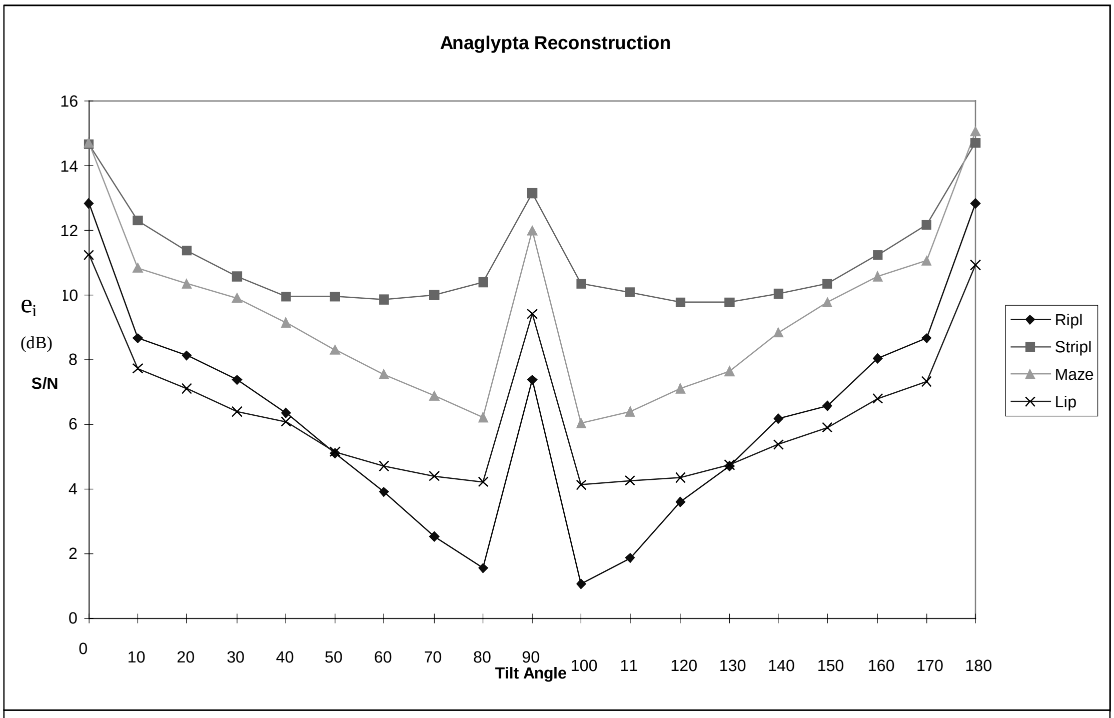
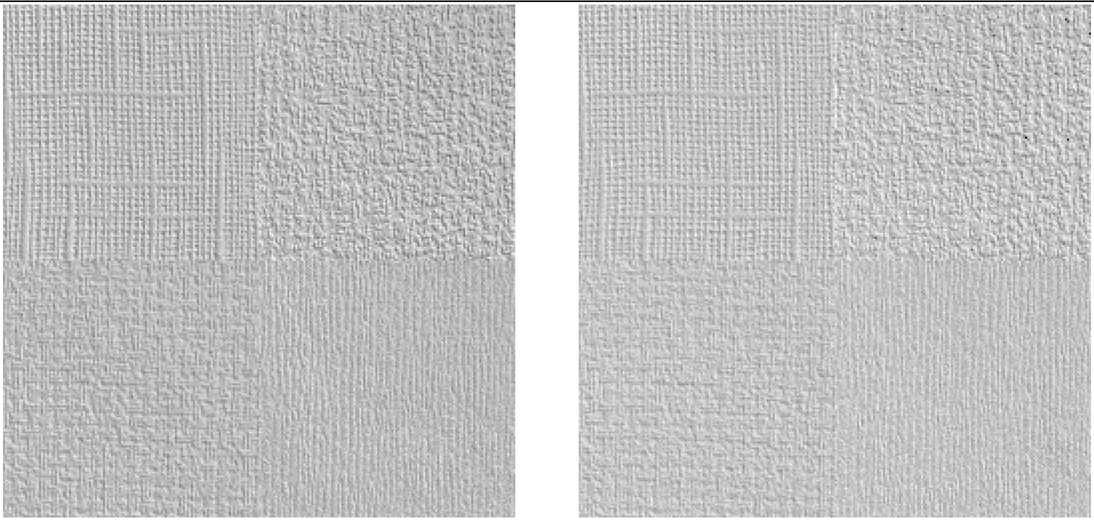


Figure 8.5.2 Accuracy of Image Prediction for Anaglypta

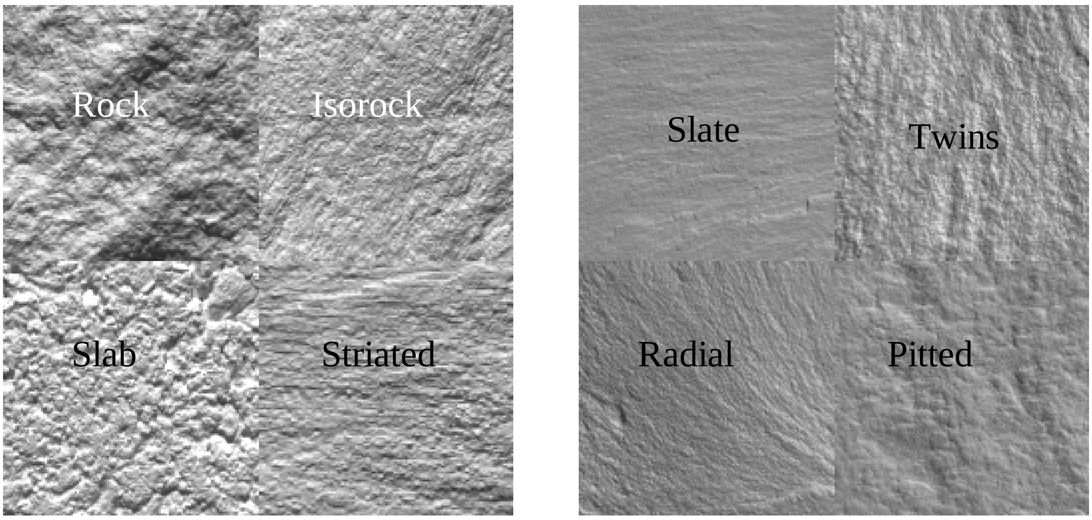
The real anaglypta surfaces illuminated from τ=50° and the simulated image are shown in *Figure 8.5.3*. While there is some difference in the average levels of the individual textures, the image textures have been predicted well. It is interesting to note that the algorithm is able to model both the vertical and horizontal directionalities, since the recovery images were illuminated in these directions they tended to have one of the directionalities, but not both. This result has shown that the algorithm is able to effectively integrate information from the recovery images.



(a) *(b)*

Figure 8.5.3 Real(a) and model(b) anaglypta montages illuminated from τ= 50°

We use two montages of rock textures, *Figure 8.5.4* and we reiterate the temporal S/R figures (e_i) for these textures see *Table 8.5.1*. The values tabulated form a loose upper bound on the level of accuracy with which we can expect the system to predict.



(a.) *(b.)*

Figure 8.5.4 Stone1(a) and Stone2(b) Montages

Texture	e_t (dB)	Texture	e_t (dB)
Rock	28.90	Slate	18.90
Striate	27.08	Pitted	12.63
Isorock	25.67	Twins	23.76
Slab	29.23	Radial	20.20

Table 8.5.1 Signal to temporal residue ratios for captured images (copied from Chapter4)

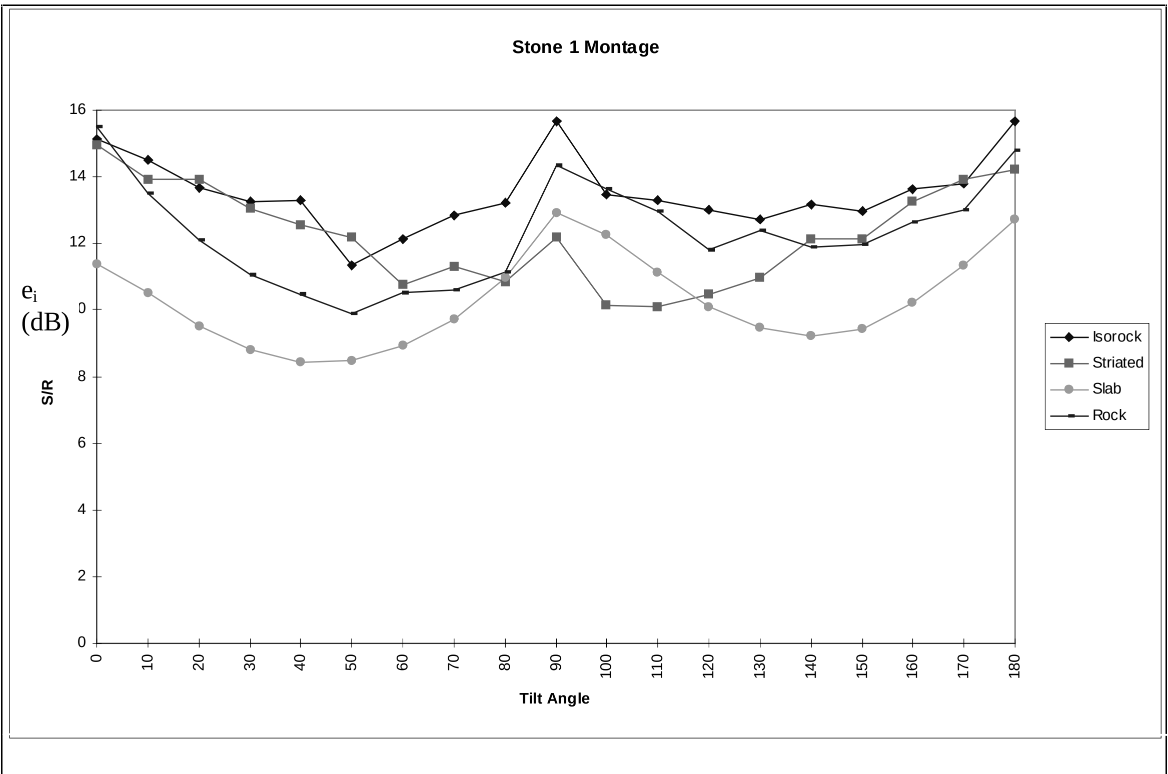
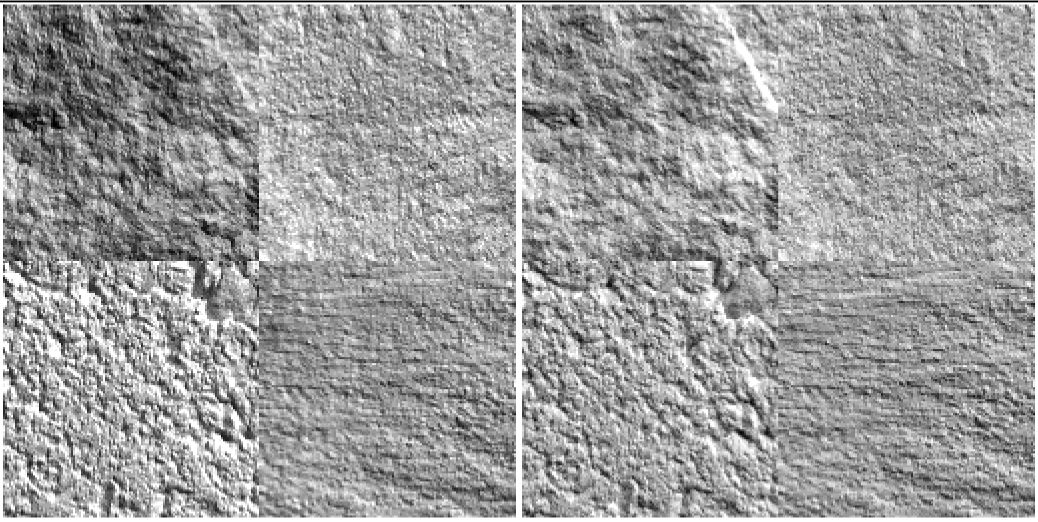


Figure 8.5.5 Prediction accuracy for Stone 1 textures.

The algorithm performs well for all the textures in the Stone 1 montage—in several instances exceeding the S/N range predicted using simulations. Of the samples used in this montage the slab texture is the least well modelled; this is noteworthy for two reasons: firstly this is the texture which had the highest S/R ratio in *Table 8.5.1*; secondly it is the texture which has the steepest facets and the most shadowing. This indicates that, in agreement with the simulation section, temporal noise is less important than surface characteristics in determining the quality of image prediction.



(a) Actual texture montage (b) Simulated montage.

Figure 8.5.6 Comparison of real and simulated Stone 1 montages at τ=50°.

Visual comparison of real and simulated textures again shows some variation in the mean levels of the real and simulated textures; this being most apparent for the *Rock* surface. It should be noted that, even for the *Rock* surface, the characteristic features, or 'landmarks' have been preserved. The second class of error which is apparent on inspection is the underestimation of intensity for those slopes on the slab surface oriented in the direction of the illuminant. Since this is the roughest surface, this suggests that either the slopes are being underestimated, or the rendering function is inaccurate for large slopes.

The algorithm's performance on the *Stone2* montage is much less impressive: firstly the majority of samples lie below the predicted S/R ratio; secondly two textures, *Radial* and *Twins*, do not display the usual three peak waveform. The *Slate* texture is also poorly modelled, though as with the other textures in this group this may be partially explained by the poor S/R rate of the original images

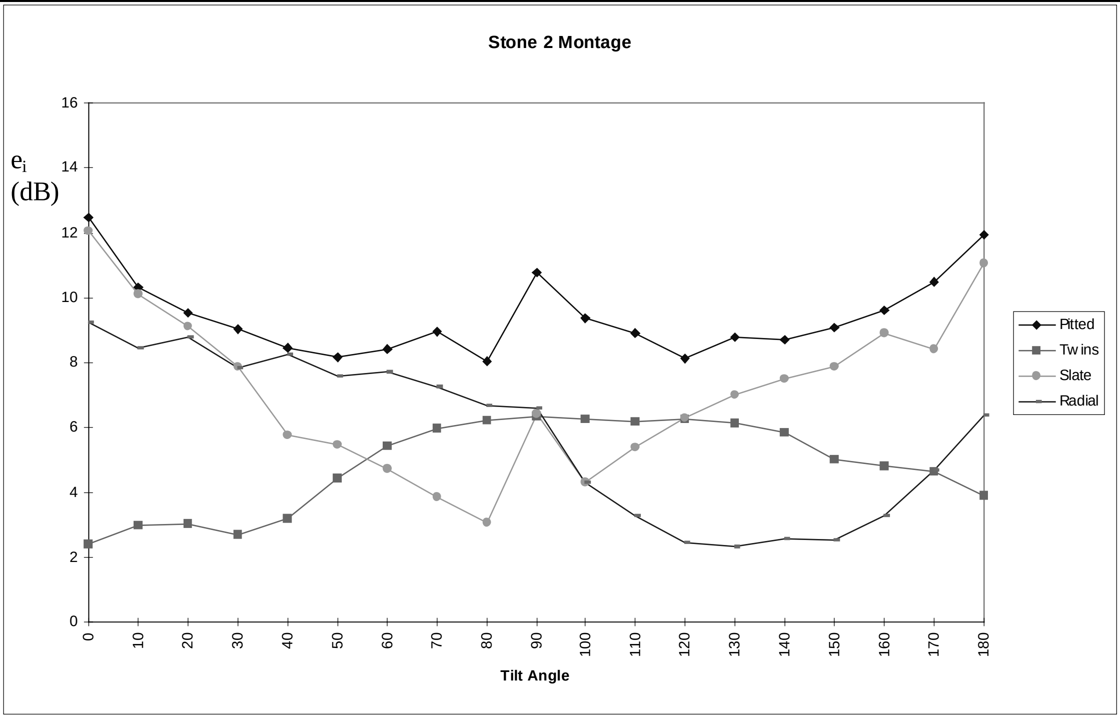
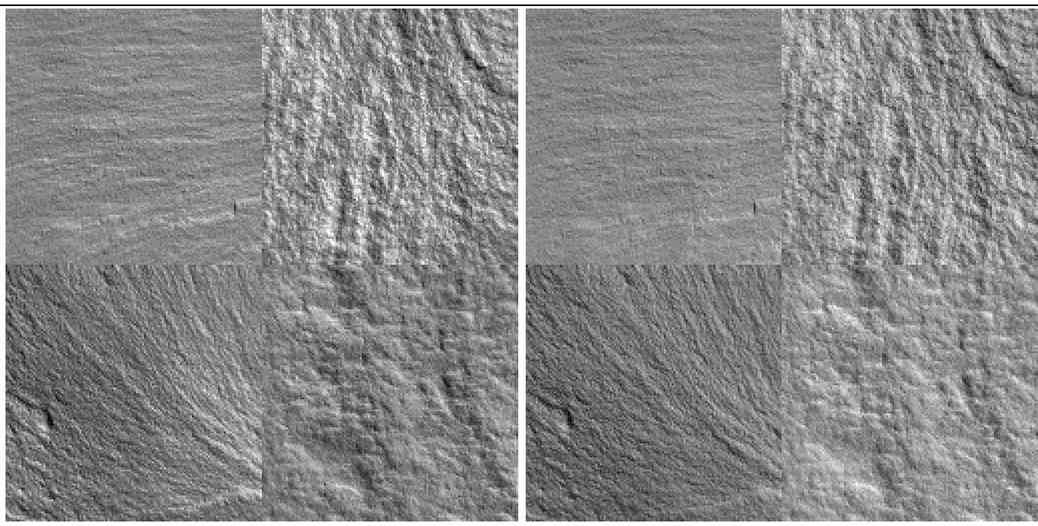


Figure 8.5.7 The Accuracy of Image Prediction for Rock Textures (Recovered and Simulated at σ=60°)

.



(a) Actual Stone2 Montage (b) Simulated Montage

Figure 8.5.8 Comparison of real and simulated Stone2 montages

As with the *Stone1* montage, the visually apparent errors in the *Stone2* montage (*Figure 8.5.8*) appear to fall into two classes. The errors in the mean level of textures (most apparent in the 'Pitted' texture) and an underestimation of the intensity of steep slopes, notably in the *Twins* texture. It is worth reiterating that even in these cases, the spatial relationships of intensities appear to be accurately predicted.

The accuracy of image prediction was shown to vary widely within the data set, ranging from almost 16dB to just over 2dB, this variation being related to the surface characteristics of the textures. In the previous section simulations indicated that the accuracy of recovery would vary over an even wider range, depending on the rms slope. The large range of the estimate, and our inability to characterise the rms slope of the surfaces do not permit any meaningful conclusions to be drawn from this overlap. All bar two of the textures displayed the characteristic W-waveform noted in the simulation section.

8.5.3 Accuracy in the Feature Domain

As has been stated elsewhere in this thesis, although the final arbiter of the scheme is the ability to reduce tilt-induced misclassification, the interdependency of the test set presents an obstacle to the analysis of the process. While section 8.5.2 treated the surfaces individually, it has given little idea of how the scheme would perform in a classification task. This section attempts to reconcile the contradictory demands of (1) a measure of the model accuracy in the feature domain stated in terms of classification accuracy, and (2) an analytical approach that treats each texture in isolation.

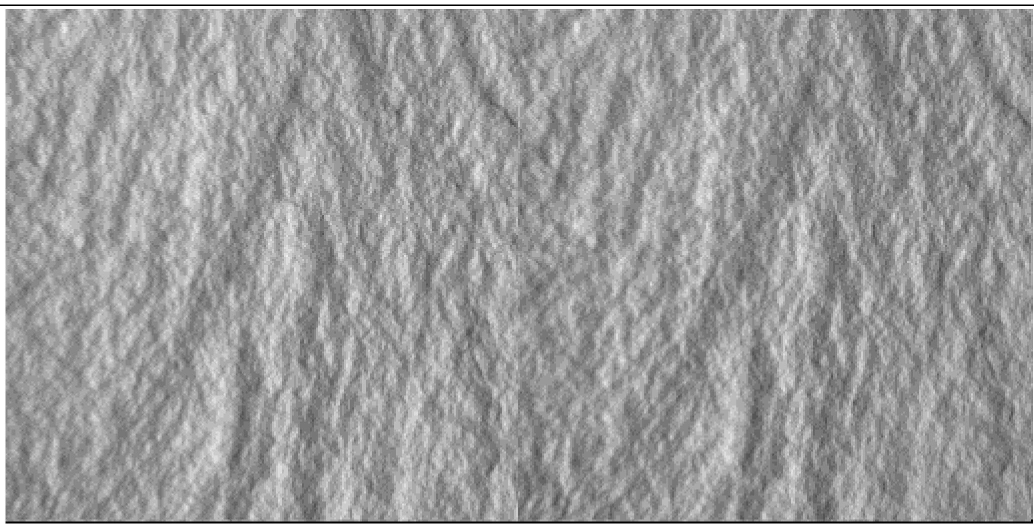


Figure 8.5.9 Montage comprising real (left) and simulated(right) 'Twins' textures.

In this section we will ask how well are the textures modelled in the feature domain. We adopt the following procedure: a two class montage consisting of the captured and synthesized images of a texture is defined (*Figure 8.5.9)*. The entire feature set used in the classification task, and a quadratic classifier trained on the montage is brought to bear on the task of discriminating between the model and the texture. The misclassification rate will vary from 0% where the synthesized and the actual texture are easily seperable, to 50 % where they are indistinguishable on the basis of the feature set. We note that even a misclassification rate approaching zero does not imply that the technique is useless for our purposes, since this will also depend on the closeness of the other textures in feature space.

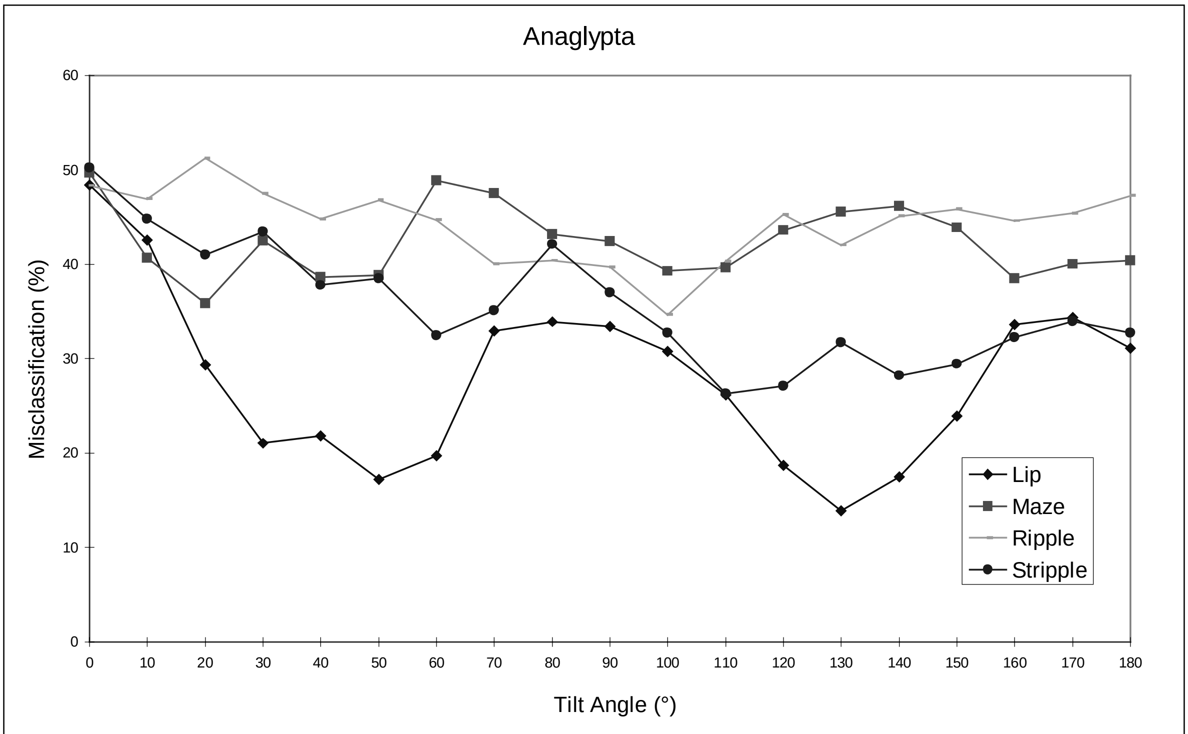


Figure 8.5.10 Classification accuracy for two class anaglypta montages.

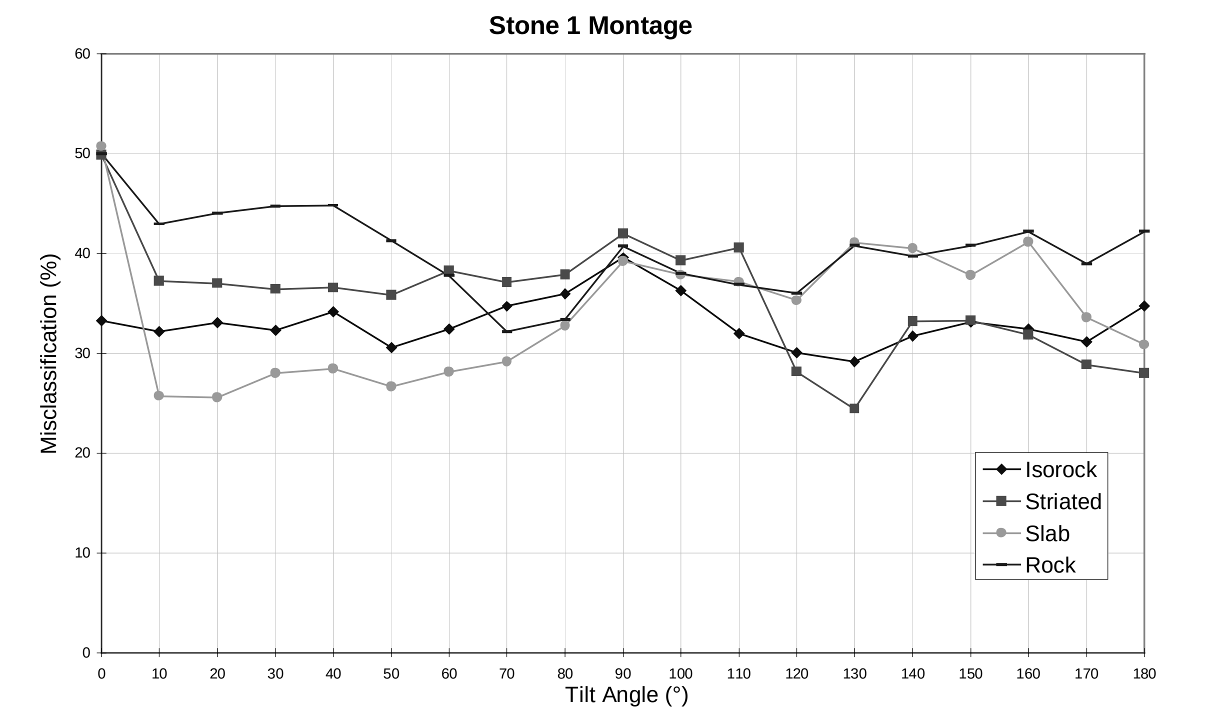


Figure 8.5.11 Classification accuracy for 2 class Stone1 texture montages.

Figure 8.5.12 Classification accuracy for Stone 2 two class montages.

The results show much less variance between the textures than might be expected from the previous section. The *Stone 1* montage is modelled most effectively, though again the performance of the members of the *Stone 2* montage is not as poor as we might expect from the image accuracy measure. Several textures have misclassification rates approaching 50% for at the τ= 0° recovery image, however this quickly falls to a more typical value. Most textures give a misclassification rate of around 30%. As a purely illustrative exercise, we show the degree of overlap necessary for a single feature with two normal distributions of equal variance to give this degree of overlap in *Figure 8.5.13*. The degree of overlap in the individual features will be greater or equal to this.

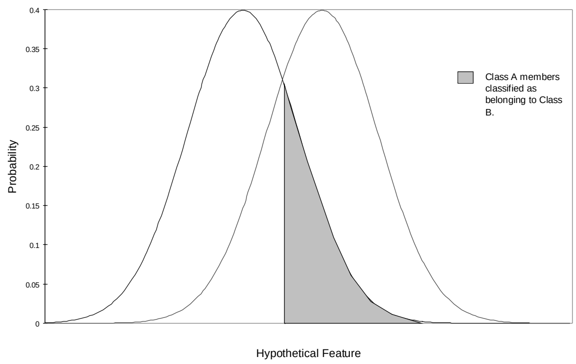


Figure 8.5.13 Degree of overlap (of a single feature) required to cause a 30% misclassification rate.

While the textures are not perfectly modelled, the fact that a classifier using a full feature set and a discriminant specially trained for the illuminant conditions cannot reliably discriminate between the model and the texture is a good indication that the textures will be adequately modelled for accurate and consistent discrimination between textures.

8.5.4 Discussion

This section has shown that there exists a wide variation in the accuracy of image prediction between different surfaces. This is typified by the members of the anaglypta montage where the S/R ratio for a given tilt angle may vary from 1 to 10 dB. Performance for the *Stone1* montage is generally good; all bar one texture, maintain a S/R greater than 10dB. The algorithm's performance for the textures in the *Stone2* Montage is much inferior, while two textures actually diverge from the W waveform, which was first noted using simulation and subsequently found in all the other textures.

In view of these results, the performance in feature space is surprising. The 'ripl' texture, which was the least accurately modelled of the anaglypta textures in terms of the S/R ratio, was the texture most accurately modelled in feature space. Similarly, although there was an obvious difference in the image accuracy of the *Stone1* and *Stone2* montages, the difference between the feature space accuracy is less significant. This suggests that the

S/R ratio is heavily influenced by high frequency noise which is ignored by a classifier that is biased towards low frequencies.

This inconsistency not withstanding, the algorithm's performance in feature space is accurate enough to prevent the classifier from reliably discriminating between the actual and simulated textures. This gives strong supportive evidence that this is an effective approach.

8.6 The Accuracy of Classification

The goal of this thesis is to develop a surface classifier that is able to maintain a level of classification accuracy despite changes in illuminant tilt. In this section we evaluate the proposed algorithm within the terms of the thesis goal.

This chapter has been concerned with the evaluation of the model-based algorithm. A reductionist approach was adopted, evaluating the performance of the system after each operation. In this section we evaluate the final stage in the process: classification. The experimental results established in this section represent the most convincing evidence that the model-based approach is appropriate for the problem of tilt dependency. However, the interdepency of samples during classification means that this stage is the least amenable to analysis, and the results do not generalise.

The approach of this section is firstly to establish that the model-based technique does represent an effective mechanism for the stabilisation of misclassification at a reasonable level. The model-based approach is designed to combat the problem of tilt dependency by generating training data appropriate to the classification task. Misclassification may be due to either:

(1.) inadequate image prediction,

(2.) the limitations of the classifier.

The former case is of more relevance to this thesis; in the second part of this section we therefore evaluate this quantity by comparing the model based classifier with a classifier based on real training data imaged under the appropriate illumination condition. This is referred to in the text as the "best case" classifier. The difference between the model-based misclassification rate and that of the "best case" classification will be due solely to the inaccuracy of image prediction.

Anaglypta Montage

Figure 8.6.1 shows the performance of a naive, the best-case and a model-based classifier on the anaglypta montage shown in *Figure 8.5.3*. The naive classifier (developed in Chapter 5) was trained at τ=0°. The full feature set was used with a quadratic discriminant. The model-based classifier used the same classifier and recovery images obtained at τ=0°,90° and 180°.

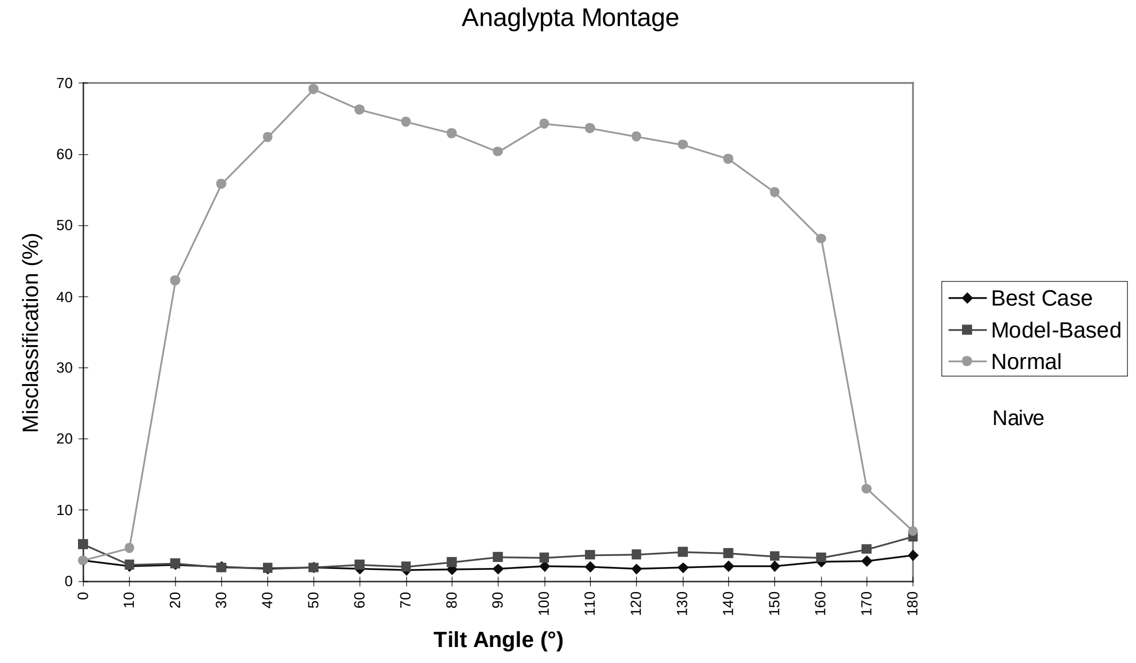


Figure 8.6.1 Classification of Anaglypta Textures

As was shown in chapter five, the misclassification rate of the naive classifier quickly rises as the tilt angle is moved away from that used during training, before returning to the original rate as the tilt angle approaches 180°. While the model based technique initially performs more poorly than the naive classifier, it does maintain a stable level of misclassification throughout the tilt range. Assessed over the tilt range the model-based technique is superior in performance to the naive classifier.

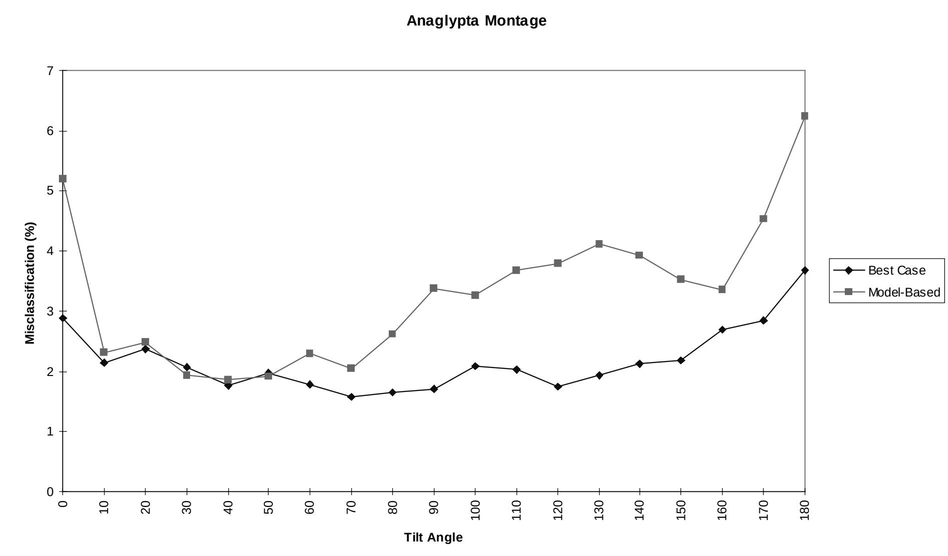


Figure 8.6.2 Comparison of Model Based and Best Case Classification For Anaglypta Montage

In *Figure 8.6.2* we look more closely at the relationship between the model-based and the Best Case classifiers. With the exception of the images at either extreme of the tilt range, the anaglypta model-based approaches the best case error rate. The fact that the model performs least well on the recovery images is particularly puzzling since these are the images for which we would expect the algorithm to be most effective.

Stone Montages

Both the Stone montages are inherently more difficult to classify; a fact reflected in the relatively poor classification rate obtained with the naive classifier at the training angle. Nevertheless, in both cases the misclassification rates follow a similar pattern to that exhibited for the anaglypta montage. The error rate of the naive classifier quickly rises as tilt is varied, falling again as tilt approaches 180°. While the model-based error is worse than the naive classifier for the training angle, it is superior over the rest of the tilt range. For the *Stone1* data set the model-based technique performs significantly worse than the best case classifier, though for the *Stone2* montage its error rate approaches the lower bound.

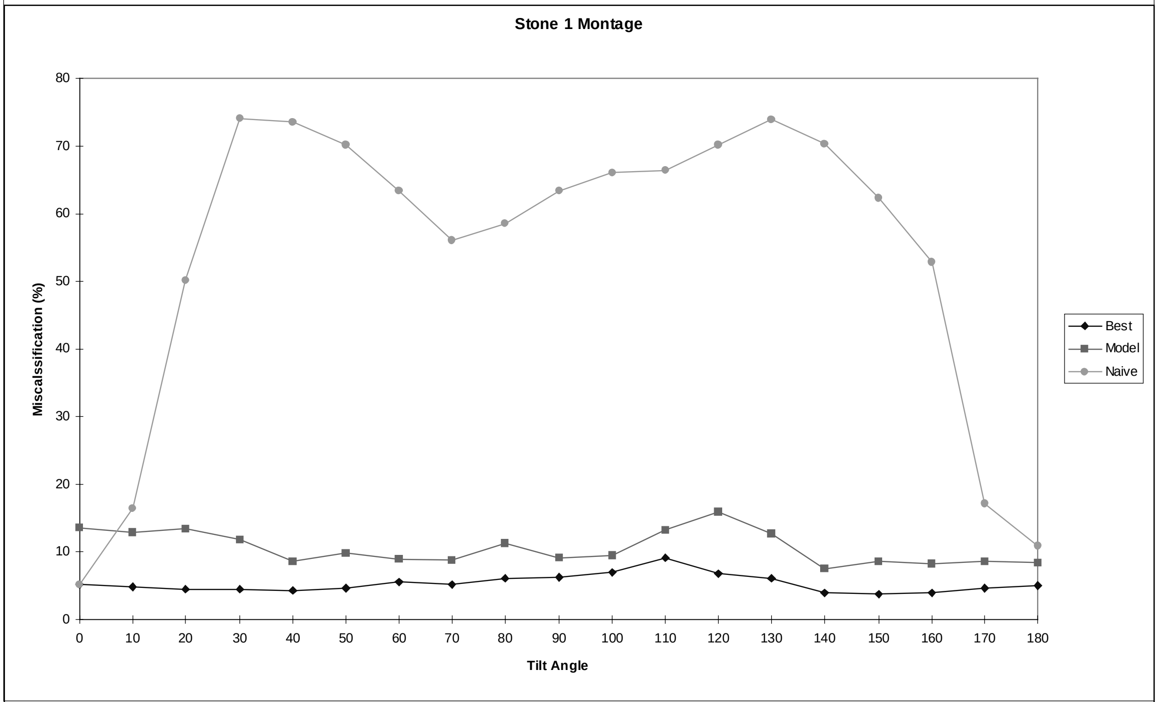


Figure 8.6.3 Classification of Stone1 Montage

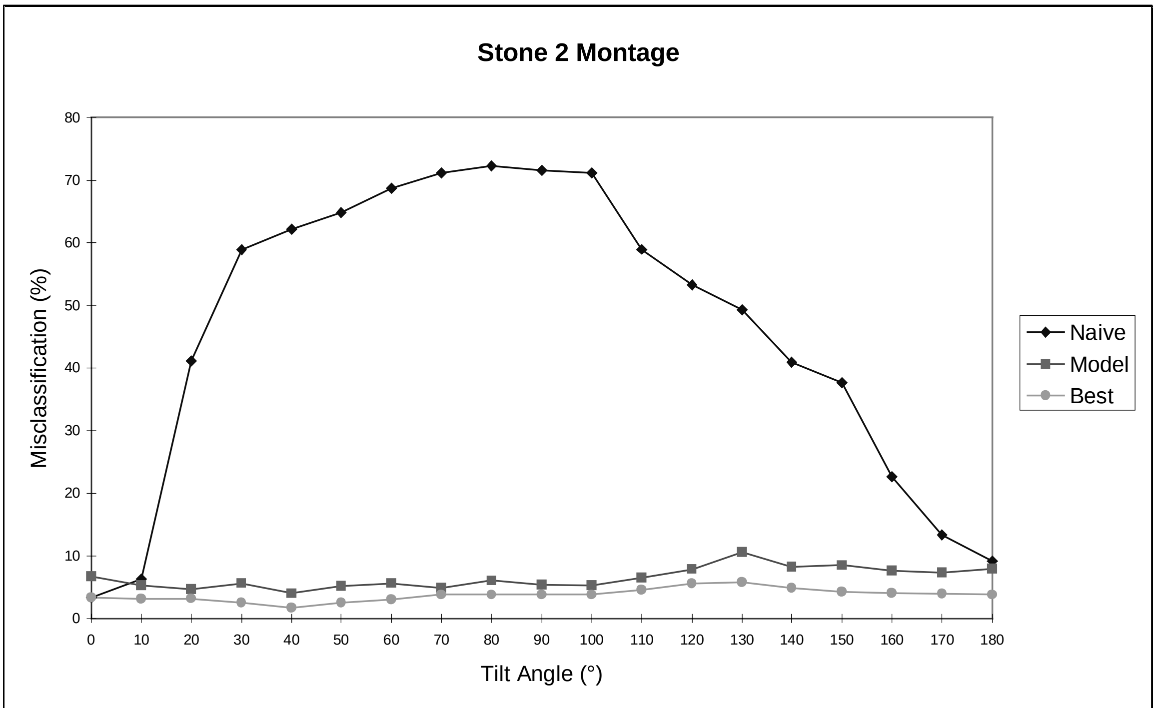


Figure 8.6.4 Classification of Stone 2 Montage

Neither of the model-based performance rates obtained for the Stone montages approach the best case rates as closely as the results obtained for the anaglypta montage, however, set against the naive approach, the model-based results are more impressive. Comparison of the waveforms of the best case and model-based classifiers shows that the model-based misclassification rate 'shadows' the best case rate, *Figure 8.6.5*. This leads to an interesting question: does the model-based classifier make the same errors as the best case as well as some additional errors, or does it make a different set of errors.

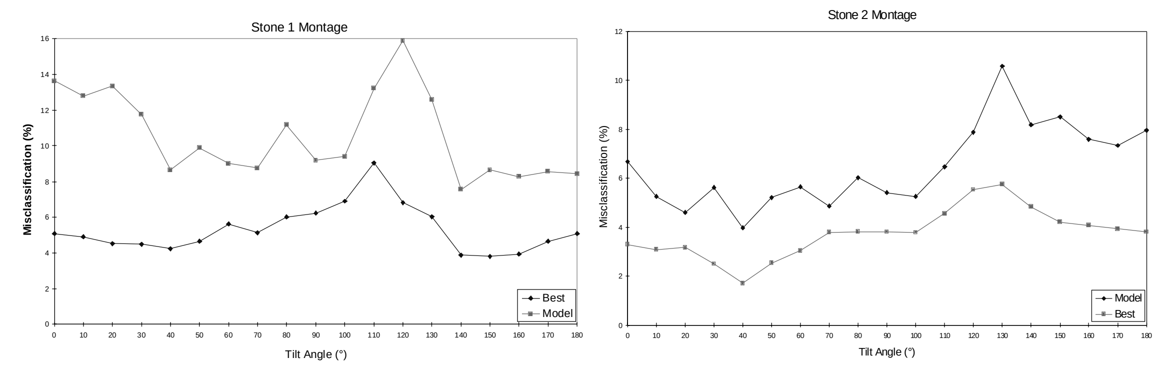


Figure 8.6.5 Comparison of model-based classifier with best case classifier.

To answer this question, we define two new error rates: the first takes the best case classification as the correct, or ground case, and calculates the model's error rate from that point; the second error rate is the simple arithmetic difference between the misclassification rates. If the classification errors of the model-based classifier include all of the errors of the best case classifier then we would expect the two values to be equal.

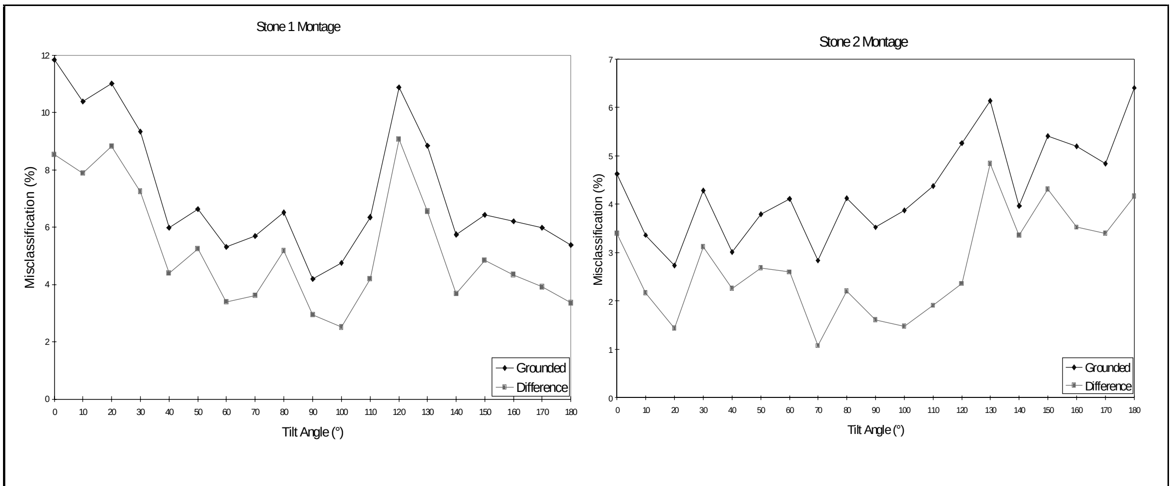


Figure 8.6.6 Comparison of the 'grounded' and 'difference' error rates.

The misclassification rates do shadow each other very closely, and the difference between the two traces is relatively small, especially for the Stone 1 classifier *Figure*

8.6.6. We therefore conclude that the model-based technique does share many of the same errors as the best case classifier as well as 2-3% which are attributable to the model-based method.

For all three data sets, the simulation-based technique clearly offers a significant reduction in tilt-induced misclassification. For the montages used in this research we have established that the model-based classifier maintains the misclassification rate at a level which is comparable with, if not actually approaching, that of the best case classifier. In terms of training requirements, the model-based classifier is most similar to the multiple discriminant technique evaluated in chapter six. The ability of the model-based technique to maintain a broadly constant rate of misclassification shows that it is more effective than this technique. The results show that in all cases the technique has successfully stabilised the misclassification rate at a low level and is, with the exception of the 'best case' classifier, the most effective solution to the problem of tilt dependency.

8.7 Discussion

While the experimental results of the simulation-based scheme are extremely promising, there are some points which require discussion.

- *Reflectance assumptions*. In the experimental work carried out here, all surfaces were of approximately Lambertian reflectance and uniform albedo. Suspension of the former condition will require more complex surface recovery and rendering algorithms, but will not of itself affect the validity of the simulation/training concept. The uniformity condition on the other hand is more serious: although the albedo variation may be isolated from the training data, it will be difficult to generalise the demodulation process to the classification data.
- *Rotation-Invariance*. This is probably the most fundamental and intractable limitation of the model-based scheme. The model-based system as discussed in this chapter is unable to deal with the rotation of directional surfaces. However simulation-based techniques may still have a significant role as a means of economically generating large volumes of training data for the design of classifiers that are invariant both to the rotation of textured surfaces and the associated variation in relative illuminant tilt.
- *Limited Data Set*. The experiments described in this thesis were carried out on twelve real surfaces. Images of each surfaces were obtained over a tilt range of 0-180°, at 10° increments, giving a total of two hundred and sixteen test images. Despite the large

number of images we note that the restricted number of surfaces used limits the scope of our results.

8.8 Conclusions

Our first conclusion is that even the simple shape recovery system described in this thesis is able to form the basis of image prediction to a good degree of accuracy for almost all the textures considered here. Using simulation we have shown the technique to be robust to noise in the recovery images. This notwithstanding, we do express reservations as to the accuracy of derivative recovery. Inaccuracies at the observed levels do not appear to seriously affect image simulation. However, we caution against blind application of surface height recovery where errors are cumulative. Nevertheless, the level of accuracy obtained is adequate for the purposes of training a classifier to discriminate between the members of our data sets. More sophisticated models are available and are an area of continuing research within the machine vision community and application of these techniques will allow the relaxation of the Lambertian requirement.

The second conclusion is that, of the techniques described in this thesis, a simulation-based system forms the best approach to consistently good classification regardless of illuminant tilt. The misclassification rate is consistent and in most cases approaches the ‘best case’ level. The concept of simulated training data does represent a powerful tool for the development of systems robust to noise and illuminant changes for little overhead at the training stage.

Our final conclusion is that for the surface estimation technique used in this report, the simulation technique may be safely applied only to cases that have the same slant angle as the classification image. This effectively means that the scheme is best thought of as an interpolation scheme, though a more effective photometric technique may be more open to extrapolation.

Chapter 9

Summary and Conclusions

9.1 Summary

The aim of this thesis is the development of a classifier that is capable of discriminating between rough surfaces on the basis of their visual appearance, yet which is robust to illuminant tilt. The first part of this aim; that is the development of a rough surface classifier, was undertaken in chapters 2 to 6. A physically-based approach was adopted, which placed the algorithmic classifier in the context of the surface and the intervening processes in the measurement of the data set.

Chapter 2 surveyed methods of surface description adopted in scattering theory and tribology. The power spectrum, in association with a phase condition, was adopted as the principal means of surface description. Three models, framed in terms of this description were adopted. These form the starting point of the simulation work performed in this thesis.

The aim of this thesis requires the classification of rough surfaces on the basis of their visual appearance, and the formation of the image was considered in Chapter 3. Image formation is considered in two distinct stages: the local intensity of a surface facet; and the interaction of facets to form a global image. The local interaction is a function of the surface and illuminant properties, but also of their relative geometeries. Three models of diffuse reflection were compared with the reflectance function for the test surfaces and the classical Lambertian model was adopted.

The global interaction of surface with the illuminant was first modelled by Kube [Kube88]. In the second part of Chapter 3 the accuracy and scope of application of a linear model relating surface and image spectra was considered. The parameter behaviour of the optimal least squares linear filter was compared with that predicted by Kube's model. The form of Kube's predictions was found to be accurate for isotropic surfaces and directional surfaces of moderate slope. However, it was noted that there exists an additional scaling relationship, which is related to the degree of surface roughness.

The algorithms adopted in the later chapters of this thesis are not applied to the incident image, rather they are applied to a data set that is subject to noise and distortion. Chapter 4 quantified the nature and magnitude of the imaging process that can be accommodated into the existing model.

Chapter 5 describes the final stage of the surface classifier: the application of the classification algorithms to the data set. This chapter begins by introducing the feature measures and discriminants. An analytical model incorporating feature extraction into the existing model from the previous chapters is developed in Chapter 6. This model predicts that the features will be a function of the illuminant tilt, which implies that the classification will also be affected. The effect of tilt on feature measures, as well as classification was then verified experimentally.

In Chapter 6 the tilt dependency of the classifier was shown; in Chapter 7 several strategies for removing, or at least reducing, this dependency were considered. Three proposals made by Chantler were investigated. The first proposal attempts to cope with tilt induced feature variation by training the classifier over a range of tilt conditions. It was shown that tilt induced variation is much larger than the variation between features of different textures at any given tilt for the test set. This is therefore not an appropriate strategy for our test set. Chantler’s second proposal also trains the classifier under several tilt conditions, however, in this case a series of discriminants are developed which are indexed by tilt angle. The classifier therefore employs whichever discriminant function is closest to the illuminant tilt. While this technique was found to reduce misclassification at the training angles, the misclassification rate quickly rose as the illuminant moved away from this angle. To maintain a reasonable level of classification, up to nine training images would be required. Chantler’s third and favoured approach was based on the inversion of Kube’s frequency domain model. Experimental work in this chapter and in Chapter 3 suggests that the optimal form of the filter varies widely from surface to surface, and the general filter, while reducing the tilt effect to a degree, does not do so to a satisfactory extent.

Since this thesis is concerned with the classification of surfaces from their visual appearance, there would appear to be a great deal of overlap with the field of shape from shading. The central problem of single image shape from shading is that of underdetermination. The majority of algorithms counter this by imposing a smoothness constraint. This is clearly not compatible with our interest in rough surfaces. We have, however, uncovered two techniques in the literature that are more suited to rough surfaces. Pentland's frequency-based scheme [Pentland90] is similar to Chantler's inverse filter, though it recovers a height map rather than a map of the derivative magnitudes. However, this similarity means that the technique will suffer from the same problems as Chantler's technique. Knill uses a pair of linear filters to recover the surface derivatives [Knill90]. Unlike either Pentland's or Chantler's technique, he uses an adaptive algorithm to train the filters. In doing so, he requires *a priori* knowledge of the surface the technique is applied to. By definition this is not available in a classification task.

Our final operation in Chapter 7 was to propose a model-based technique for the suppression of tilt-effects. This used photometric techniques to obtain a model of the surface derivative fields for each training sample. The model was synthetically rendered under the illuminant conditions in which the classifier was to be applied in order to generate appropriate training data.

In Chapter 8 the model-based system was evaluated. This was performed in two stages. The first stage used the models developed in the previous chapters to implement a simulation-based investigation into how well the system recovered rough surface derivatives and predicted the image in the presence of imaging artefacts. The second stage dealt with the evaluation of the system on real data. Although we do not have the means to estimate the accuracy of recovery of rough, real surfaces, it was possible to assess the accuracy for a smooth real surface. It is also possible to measure the accuracy of image and feature prediction. Finally, the performance of a classifier based on the model-based technique was assessed against that of the naive classifier proposed in Chapter 5 and the 'best case' classifier which was retrained under each tilt condition. The model-based technique was found to suppress the effect of tilt variation on classification and in several cases had a mis-classification rate approaching the 'best-case' lower bound.

9.2 Future Work

Slant Invariance

This thesis has dealt only with the problem of tilt invariance and has neglected the effect on classification of variations in slant angle. In theory, a model based scheme should be equally adept at dealing with variations in slant. However, this represents a much more demanding task for the technique since image prediction now requires extrapolation from, rather than interpolation between, the training images. The task is complicated by the fact that the mean level of the texture image will vary with tilt, making camera non-linearity a much more significant problem.

Surface Recovery

In the first chapter of this thesis we proposed a hypothetical inspection task, which requires rough surfaces to be identified. It is not difficult to imagine applications where, having identified an anomaly or other region of interest, it would be useful for a human operator to be able to examine the three dimensional structure of the anomaly. In Chapter 6, we concluded that the single image, rough surface recovery techniques required *a priori* knowledge of the surface type in order to perform well. The identification of the region of interest using classification techniques would allow application of this knowledge to the recovery of surface topography.

Photometric Classification

This thesis has considered the case where the illuminant tilt was not under the control of operator during the classification stage. However, in many cases it will be possible to control the direction of illuminant. Where this is the case, we advocate the integration of photometric techniques into the classification process.

A hypothetical system could use photometric techniques to estimate the surface derivative fields and apply texture analysis techniques to the estimated fields. If the derivative fields could be consistently estimated this would remove the effects of illuminant direction. Furthermore, it would allow the use of much more surface information than is present in the image since the directional filtering effect of the illuminant would be avoided. The removal of the directional filtering would also allow the application of rotation invariant techniques to rough surface textures.

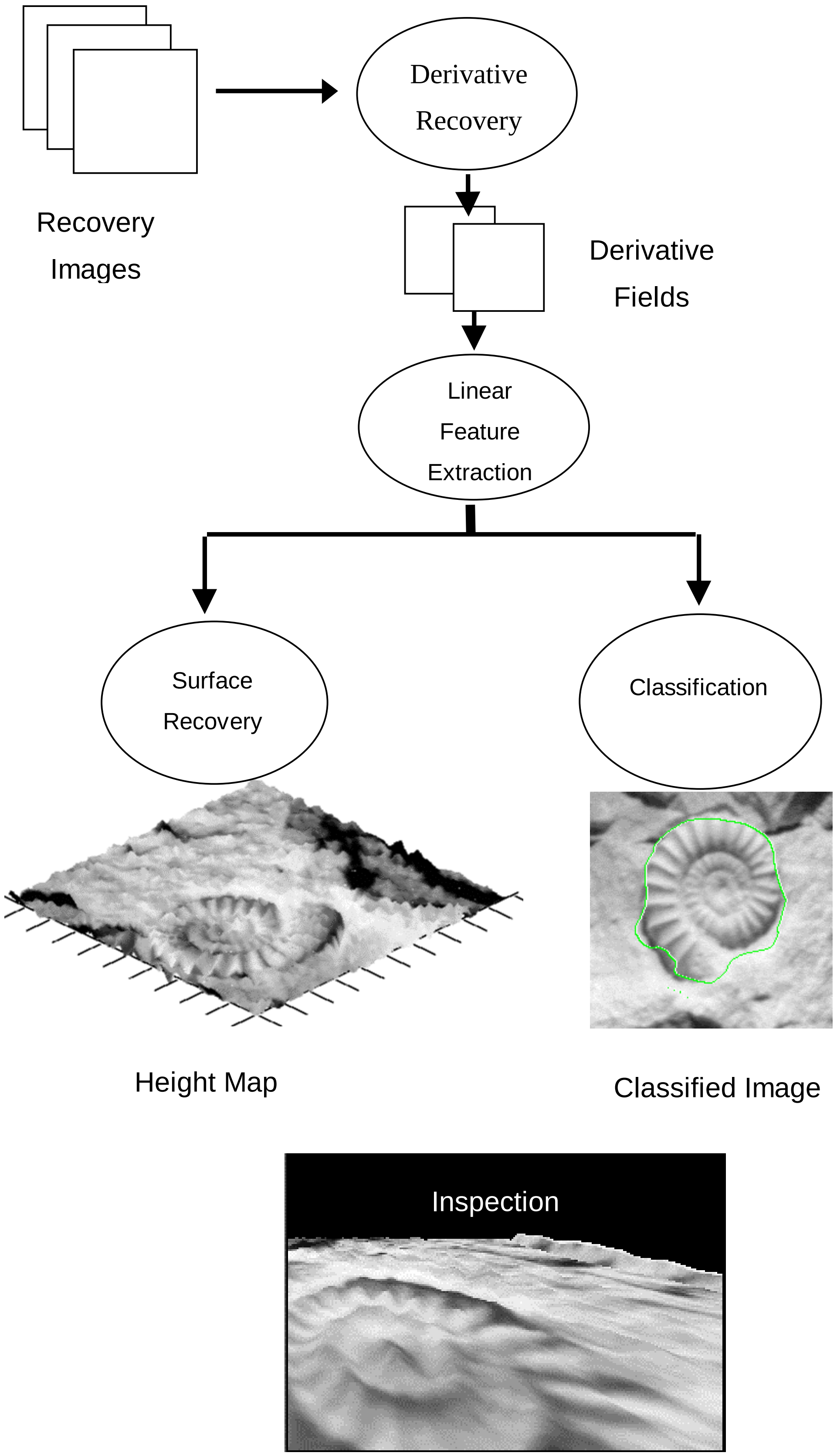


Figure 9.2.1 A proposed scheme integrating classification with surface recovery.

The recovery of derivative fields also suggests that surface recovery is a possibility. This would require the integration of the derivative fields. One significant obstacle is lack of integrability, although this can be overcome relatively easily for smooth surfaces, it is likely to present more serious difficulties for the reconstruction of rough surfaces. While integration can be carried out in the spatial domain, it can also be undertaken in the frequency domain [Frankot88], by weighting the spectral components and swapping the real and imaginary components. Pentland noted that filters localised in space and frequency form an alternative to the Fourier Transform in his frequency based shape from shading algorithm [Pentland90]. Integration of this approach into a photometric classifier would allow the recovery of surface height with little computational cost beyond that already incurred for the classification while avoiding the restrictions on albedo, linearity and noise associated with a single image scheme. The surface type classification, and the use of spectrally defined measures would also allow domain knowledge to be used by an iterative technique. A hypothetical system is shown in *Figure 9.2.1*.

9.3 Conclusions

This thesis has developed a series of models of the process of classifying rough surfaces on the basis of their visual appearance. These models are used in the description, analysis and simulation of the process.

Three spectral models of rough surfaces were adopted from the literature, and their interaction with incident light to form an image was shown in a spectrally definable form. It was found that under certain conditions imaging artefacts could be considered as consisting of a Gaussian blur transfer function, and additive white noise. The combination of the imaging model with the spectral definition of the features showed that the features on which classification is based are functions, not only of the surface, but also of the illuminant tilt angle. This was confirmed by simulation and experiment. The ability of tilt variation to induce classifier failure was demonstrated with both simulation and three montages of real textures.

Several schemes proposed in the literature for reducing the tilt induced effect were evaluated. We concluded that none of these was sufficiently effective for our data set and proposed a new model based scheme. This technique used photometric techniques to estimate the derivative fields of the training surface which could then be rendered under the appropriate illumination conditions to produce training data suited to the prevailing conditions. Evaluation of this scheme showed that it was able to accurately model real textures and to significantly reduce the effect of tilt variation on classification—easily outperforming the other schemes considered.

Appendix A: Principal Symbols

β	Power roll-off
ω	Radial frequency
θ	Polar frequency angle
σ	Slant angle
τ	Tilt Angle
ϕ	Direction of propagation ofGabor Filer
λ_0	Correlation length of an isotropic surface
ω_0	Fundamental frequency
λ_1	Correlation length in the x-direction
λ_2	Correlation length in the y-direction
σ_b	Blur
γ_b	Exponent of camera amplification.
ω_c	Cut-off frequency
$\sigma_{disparity}$	Standard deviation of the difference between two images
σ_m	Standard deviation of measure image.
σ_n	Standard deviation of overall noise process.
Σ_n	Covariance matrix of class n
σ_p	Measured polar bandwidth
σ_p	Standard deviation of the Gabor filter spectrum in the x-direction.
μ_r	Mean of Feature image
σ_r	Standard deviation of Feature image.
σ_s	Rms roughness
σ_t	Standard deviation of temporal noise.
σ_x	Standard deviation of the Gabor filter envelope in the x-direction.
σ_y	Standard deviation of the Gabor filter envelope in the y-direction.
σ_y	Standard deviation of the Gabor filter spectrum in the y-direction.
a	Parameter of feature/tilt model.
a b c	Parameters of optimal linear model.
b	Parameter of feature/tilt model.
B_ϕ	Polar bandwidth of the filter.
B(u,v)	Blur function
B_r	Radial frequency bandwidth of the filter.

$c(t)$	Autocovariance function
$d(x,y)$	Measured image data set
$D(x,y)$	Filter Outputs Vector
$d_{f\phi}(x,y)$	Output of Gabor filter f,ϕ
$e(x,y)$	Residue Process
$F(x,y)$	Feature Vector
$f_{f\phi}(x,y)$	Feature Response derived from filter f,ϕ
$G_{\omega\phi}(u,v)$	Gabor filter
$H(u,v)$	Combined filter function
$i(x,y)$	Incident image
i_0,i_{90},i_{180}	images obtained from τ=°,90° and 180° respectively.
i_d	Desired image
i_{NL}	Non-linear component of surface to image mapping.
i_p	Image which is a linear function of p-derivative field only.
i_q	Image which is a linear function of qderivative field only.
k	Topothesy
$k(F\|l_i)$	Probability that a vector x belongs to class n over the entire tilt range.
k_1 k_2 k_3	Parameters of Kube's linear model.
$L(x,y)$	Illuminant vector
$l(x,y)$	Label field
m_{fg}	f^{th} and g^{th} order statistical moment.
M_n	Mean vector of class n
m_{rms}	Rms Slope
$n(x,y)$	Noise process
$o(p,q)$	Reflectance function
$p_{\mathbf{F}\|l_i}(\mathbf{F}\|l_i)$	Probability that a vector x belongs to class n
p	Facet slope in the x-direction
p_{rms}	Rms slope in the x-direction
p_x	Second derivative of surface, in the x-direction.
q	Facet slope in the y-direction
q_{rms}	Rms slope in the y-direction
q_x	Second derivative of surface, in the x-direction.
R	Correlation matrix of the surface
$R(u,v)$	Illumination function
$r_c(t)$	Autocorrelation function
R_{cla}	Centre line average
$s(x)$	Surface height profile.

$s(x,y)$	Surface height
$S(x,y)$	Surface derivatives
t	Lag
u	Horizontal frequency index
u_0	Centre frequency of filter in the x-direction.
v	Vertical frequency index
$V[a\ b\ c]$	Least squares linear model of the illumination process.
v_0	Centre frequency of filter in the x-direction.

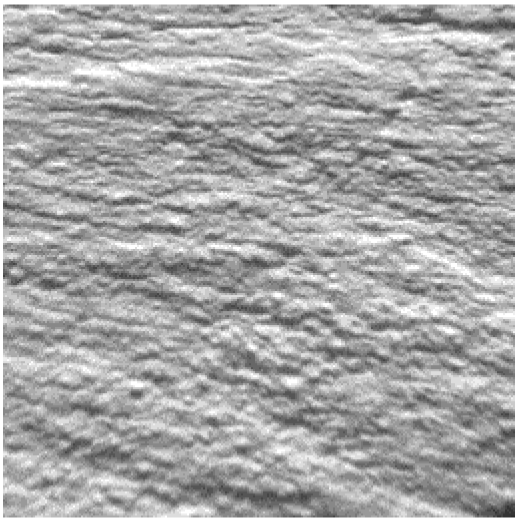
Appendix B

Textures used in Stone 1 Montage

Isoroc

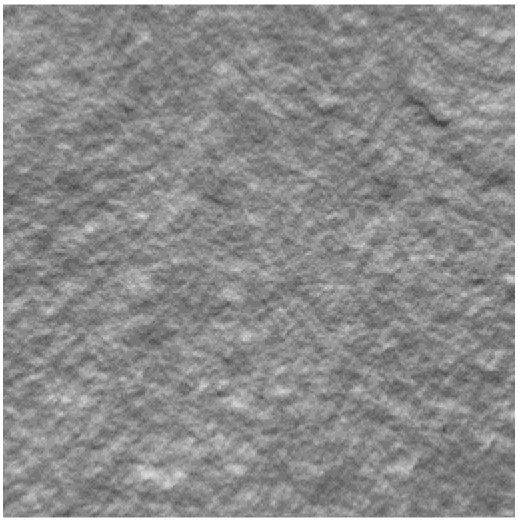
Rock

Slab

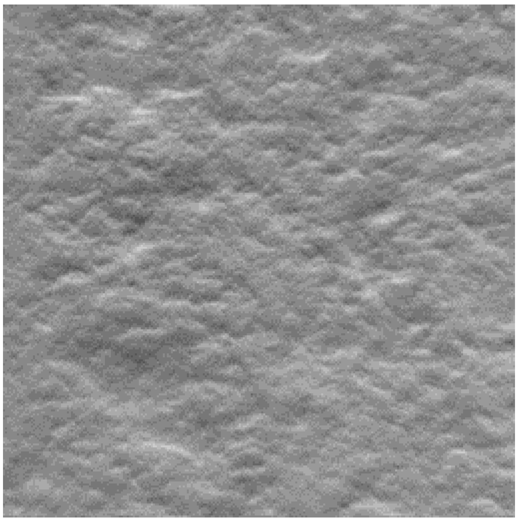


Striate

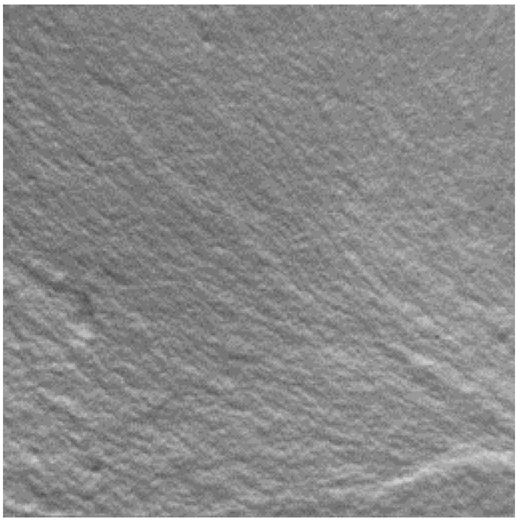
Textures used in Stone 2 Montage



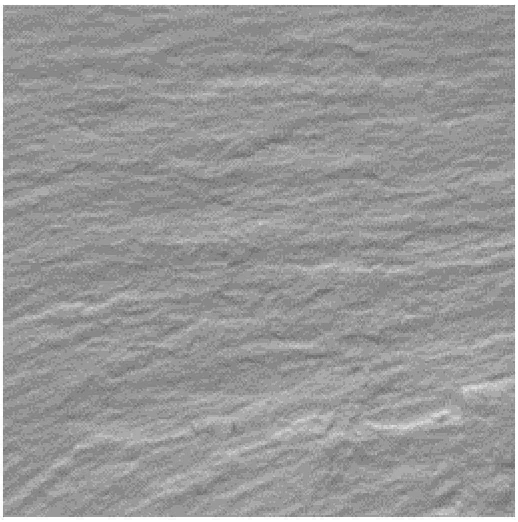
Twins



Pitted



Radial



Slate

References

Aach88 T. Aach, R. Mester & U. Franke, "From Texture Energy Measures to Quadrature Filter Pairs-A System Theoretical View of Texture Feature Extraction", Proceedings EUSIPCO88 Signal Processing Theories and Applications, 4th European Conf. Grenobel, France, Sept. 1988, pp.827-830.

Aach95 T. Aach, A. Kraup & R. Mester, "On texture analysis: Local energy transforms versus quadrature filters", Signal Processing 45 (1995) pp. 173-181.

Ade83 F. Ade, "Characterization of Textures by 'Eigenfilters'" Signal Processing 5 (1983) pp.451-457.

Antoine93 J.-P. Antoine, P. Carrette, R.Murenzi & B. Piette, "Image analysis with two-dimensional continuous wavelet transform", Signal Processing 31 (1993) pp.241-272.

Austin94 R.T. Austin, A.W. England & G.H. Wakefield, "Special Problems in the Estimation of Power-Law Spectra as Applied to Topographical Modeling", IEEE Transactions of Geoscience and Remote Sensing, Vol.32, No.4, July 1994, pp.928-939.

Beckmann63 P. Beckmann & A. Spizzichino, "The Reflection of Electromagnetic Waves from Rough Surfaces" Basingstoke, Macmillan.

Bennett89 J.M. Bennet & Lars Mattson "Introduction to Surface Roughness and Scattering", Optical Society of America, Washington D.C. 1989.

Berry M.V. Berry & J.H. Hannay, "Topography of Random Surfaces", Nature Vol.273 15 June 1978, p.573.

Blinn77 J.F. Blinn, "Models of light reflection for computer synthesized pictures", Computer Graphics, 11 (2), pp.192-198.

Boie92 R.A. Boie & I.J. Cox An Analysis of Camera Noise, IEEE Transactions of Pattern Analysis and Machine Intelligence Vol.14 No.6 June 1992, pp. 671-686.

Bovik91 A.C. Bovik "Analysis of Multichannel Narrow-band Filters for Image Texture Segmentation",IEEE Transactions on Signal Processing, 1991, Vol.39, No.9, pp.2025-2043.

Bovik96 A.C. Bovik, J.P. Havlicek, D.S. Harding "Multicomponent AM_FM Image Representation" IEEE Transactions on Image Processing 1996 Vol.5 No.6 pp.1094-1110.

Bruce92 V. Bruce & P.R. Green, "Visual Perception, Physiology, Psychology and Ecology", 2nd Ed.,Lawrence Erlbaum Associates, Publishers.

Bush79 A.W. Bush, R.D. Gibson & G.P. Gibson, "Strongly Anisotropic Rough Surfaces", Trans. of ASME Journal of Lubrication

Technology. Jan 1979, Vol. 101, pp15-20.

Chantler94a M.J. Chantler, “The effect of illuminant direction on texture classification”, Ph.D Thesis, Dept. of Computing and Electrical Engineering, Heriot-Watt University, 1994.

Chantler94b M.J. Chantler, "Why illuminant direction is fundamental to texture analysis", IEE Proc. Visual Image and Signal Processing, 1995, Vol. 142 No.4 pp. 199-206.

Chellappa85 R. Chellappa & S. Chatterjee, "Classification of textures using Gaussian Markov random fields", ASSP V.33, Aug. 1985, pp.959-963.

Chellappa92 R. Chellappa, R.L. Kayshap & B.S. Manjanuth, “Model-Base Texture Segmentation and Classification”,appears in "Handbook of Pattern Recognition and Computer Vision", Eds. Chen, Pau and Wang, 1993, pp. 277-310.

Chen93 "Handbook of Pattern Recognition and Computer Vision", Eds. Chen, Pau and Wang, World Scientific,1993.

Chen95 J.L. Chen & A. Kundu, "Unsupervised Texture Segmentation using Multichannel Decomposition and Hidden Markov Models" IEEE Trans. Image Processing Vol.4 No.5 May 1995, pp.603-619.

Cho93 C. Cho & H. Minamitani, "A New Photometric Method Using 3 Point Light Sources" IEICE Trans. Inf. & Syst. V.E76-D, No.8, August 1993, pp. 898-904

Christobal91 G. Christobal, C. Gonzalo & J. Bescos, "Image filtering and analysis through the Wigner Distribution", Advances in Electronics and Electron Physics Series, (Hawkes, P.W.,ed) Vol.80, pp 309-397 (1991) Academic Press, Orlando, Florida

Church82 E.L. Church & H.C. Berry, "Spectral Analysis of the Finish of Polished Optical Surfaces", Wear,83 (1982) pp. 189-201

Church88 E. Church "Fractal Surface Finish", Applied Optics, Vol.27, No.8 15 April 1988 pp1588-1526

Clarke92 S.J. Clarke, "The Analysis and Synthesis of Texture in Sidescan Sonar Data", PhD Thesis. Dept. of Electrical and Electronic Engineering, Heriot-Watt University, 1992

Cohen91 F.S. Cohen, Z. Fan & M.A.S. Patel, "Classification of rotated and scaled textured images using Gaussian Markov field models", Trans. of Pattern Analysis and Machine Intelligence Vol.13 February 1991, pp.191-202.

Cole82 E.N. Coleman & R. Jain, "Obtaining 3-Dimensional Shape of Textured and Specular Surfaces Using Four-Source Photometry", Computer Graphics and Image Processing 18 pp. 309-328 (1982).

Cook82 R.L. Cook & K.E. Torrance, "A Reflectance Model for Computer Graphics", Computer Graphics 15 (1) pp.51-72.

Couch L.W. Couch, "Digital and Analog Communication Systems" Fourth Edition, Macmillan.

Cryer95 J.E. Cryer, P.S. Tsai & M. Shah, "Integration of Shape from Shading and Stereo", Pattern Recognition, Vol.28 No.7, pp 1033-1043, 1995.

Daugman85 J.G. Daugman "Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two dimensional visual cortical filters" J. Opt. Soc Am A Vol. 2 No.7 July 1985 pp.1160-1169.

Dunn95 D. Dunn & W.E. Higgins, "Optimal Gabor Filters for Texture Segmentation", IEEE Trans. Image Processing, Vol.4 No.7 July 1995, pp. 947-964.

Ehrich78 R.W. Ehrich & J.P. Foith, "A View of Texture Topology and Texture Description", Computer Graphics and Image Processing 8, pp.174-202 (1978).

Frankot88 R.T. Frankot & R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms", IEEE Trans. Pattern Analysis and Machine Intelligence Vol.10, No.4, July 1988, pp.439-451.

Freeman91 W.T. Freeman & E.H. Adelson, "The Design and Use of Steerable Filters", IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 13, No.9, September 1991, pp. 891-906.

Fukunaga90 Keinosuke Fukunaga, "Introduction to Statistical Pattern Recognition" Second Edition, Academic Press 1990.

Gibson50 J.J. Gibson, "The perception of the visual world." Boston: Houghton Mifflin 1950.

Gibson79 J.J. Gibson, "The ecological approach to visual perception." Boston: Houghton Mifflin 1979.

Greenhill93 D. Greenhill & E.R. Davies, "Texture analysis using neural networks and mode filters", Proceedings of BMVC93 , September 1993, University of Surrey, pp.509-518.

Greenspan94 H. Greenspan, R. Goodman, R. Chellappa & C.H. Anderson, "Learning Discrimination Rules in a Multiresolution System", IEEE Trans. Pattern Analysis and Machine Intelligence Vol.16, No.9 Sept. 1994, pp.894-901.

Gross95 M.H. Gross & R. Koch, "Visualization of Multidimensional Shape and Texture Features in Laser Range Data Using Complex-Valued Gabor Wavelets", IEEE Trans. Visualization and Computer Graphics, Vol.1 No.1 March 1995, pp.44-59

Hall95a P. Hall & S. Davies, On direction invariance of fractal dimension on a surface" Applied Physics A (Materials Science & Processing) Vol.60, pp. 271-274 (1995).

Hall95b T.E. Hall & G. B. Giannakis, "Bispectral Analysis and Model Validation of Texture Images", IEEE Transactions on Image Processing Vol. 4 No.7 July 1995, pp.906-1009.

Hasegawa93 M. Hasegawa, J. Liu, Y. Konishi, "Characterization of Engineering Surfaces by Fractal Analysis", Int. J. Japan Soc. Prec. Eng., Vol.27 No.3 (Sept 1993), pp.192-196.

Healey89 G. Healey, "Using color for geometry-insensitive segmentation", Journal of the Optical Society of America, 1989, Vol.6, No.6, pp.920-937.

Healey94 G.E. Healey & R. Kondepudy, "Radiometric CCD Camera Calibration and Noise Estimation",IEEE Transactions of Pattern Analysis and Machine Intelligence Vol. 16 No.3 March 1994, pp.267-276.

Healey95 G. Healey, L. Wang, "Illuminant-Invariant Recognition of Texture in Color Images" Journal of the Optical Society of America, 1995, Vol. 12, No.9, pp.1877-1883.

Hlawatshch92 F. Hlawatsch & G.F. Boudreaux-Bartels, "Linear and Quadratic Time-Frequency Signal Representations", IEEE Signal Processing Magazine, April 1992, pp.21-67.

Horn90 B.K. Horn, "Height and Gradient from Shading", International Journal of Computer Vision, 5:1, pp 37-75 (1990)

Imre92 A. Imre, T. Pajkossy & L. Nyikos. "Electrochemical Determination of the Fractal Dimension of Fractured Surfaces". Acta metall. mater. Vol.40, No.8 pp1819-1826 1992.

Iwahori95 Y. Iwahori, R. Woodham & A. Bagheri, "Principal Components Analysis and Neural Network Implementation of Photometric Stereo" 1995 IEEE pp117-125.

Jain91 A.K. Jain & F. Farrokhnia, "Unsupervised texture segmentation using Gabor Filters" Pattern Recognition, Vol.24, No.12, pp1167-1186, 1991.

Jain96 A.K. Jain & K. Karu, "Learning Texture Discrimination Masks", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No.2 February 1996, pp.195-205.

Kay95 G. Kay, T. Caelli, "Estimating the Parameters of an Illumination Model Using Photmetric Stereo" Graphical Models and Image Processing, Vol.57, No. 5, September 1995 pp. 365-388.

Kayshap83 R.L. Kayshap & R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images", IT Vol.29 , January 1983, pp. 60-72.

Keshner82 M. Keshner, "1/f Noise", Proceedings of the IEEE, Vol.70 No.3, March 1982, pp212-218.

Kiernan95 "Implementation and Design of Discrete Gabor Filter for Sonar Texture Classification" PhD Thesis, Dept. Computing and

Electrical Engineering, Heriot Watt University, 1995.

Knill90a D.C Knill, D. Field & D. Kersten, "Human discrimination of fractal images" J. Optical Soc. of America, Vol.7 No.6 June90, pp.1113-1123.

Knill90b D.C. Knill & D. Kersten, "Learning a Near-Optimal Estimator for Surface Shape from Shading" Computer Vision, Graphics and Image Processing, 50 pp.75-100 (1990).

Kolb95 C. Kolb, D. Mitchell & P. Hanrahan, A Realistic Camera Model for Computer Graphics SIGGRAPH95 Los Angeles, California, August 6-11 1995, pp. 317-324.

Kube88 P. Kube & A. Pentland, "On the Imaging of Fractal Surfaces" IEEE Trans. on Pattern Analysis and Machine Intelligence Vol.10 No.5 Sept 1988, pp.704-707.

Laws79 K.I. Laws, "Texture Energy Measures", Proc. Image Understanding Workshop, No. 1979 pp.47-51.

Lee84 H.H. Lee & A. Rosenfeld, "An approximation technique for photometric stereo", Pattern Recognition Letters 2 (1984) pp.339-343.

Lee96 T.S. Lee, "Image Representation using 2D Gabor Wavelets" IEEE Transactions on Pattern Analysis and Machine Intelligence", Vol.18, No.10 October 1996, pp.959-971.

Lim J. S. Lim, "Two-dimensional Signal and Image Processing" Prentice Hall.

Linnett91 L.M. Linnett, “Multi-Texture Image Segmentation”, PhD Thesis, Dept. of Electrical and Electronic Engineering, Heriot Watt University, 1991.

Liu90 Song-Sheng Liu & M.E. Jernigan, Texture Discrimination in Additive Noise", Computer Vision, Graphics and Image Processing 49, pp.52-67 (1990).

Livens97 S. Livens, P. Scheunders, G. van de Wouwer & D. Van Dyck, "Wavelets for texture analysis, an overview" IPA97, Dublin, 15-17 July 1997, pp581-585.

Longuet-Higgins56 M.S. Longuet-Higgins, "The Statistical Analysis of a Random, Moving Surface",Philosophical Transactions of the Royal Society, Vol.249, Series A, 1957, pp.321-387.

Majumdar90 A. Manjumdar, C.J. Tien. "Fractal Characterization and Simulation of Rough Surfaces". Wear, 136, 1990. pp. 313-327

Miller91 P. Miller & S. Astley, "Classification of Breast Tissue by Texture Analysis" Proceedings of BMVC91, Glasgow, Sept. 1991.

Mulvanney89 D.J Mulvanney, D.E. Newland & K.F. Gill. "A Complete Description of Surface Texture Profiles", Wear 132 (1989) pp. 173-182.

Namuduri94 K.R Namuduri, R. Mehrotra & N. Ranganathan, "Efficient Computation of Gabor Filter Based Multiresolution Responses", Pattern Recognition, Vol. 27, No.7 pp.925-938, 1994.

Navarro95 R. Navarro, A. Tabernero & G. Christobal, "Image Representation with Gabor Wavelets and its Applications", Advances in Imaging and Electron Physics, Ed. P.W. Hawkes, Academic Press Inc, Orlando, FL.

Nayak71 P.R Nayak, "Random Process Model of Rough Surfaces", Transactions of ASME Journal of Lubrication Technology, July 1971 pp.398-407.

Nayar90 S.K. Nayar, K. Ikeuchi, T. Kanade."Determining Shape and Reflectance of Hybrid Surfaces by Photometric Sampling", IEEE Transactions on Robotics and Automation, 1990, Vol.6, No.4, pp.418-431.

Nayar91 S.K. Nayar, K. Ikeuchi, & T. Kanade, "Surface Reflection: Physical and Geometric Perspectives" IEEE Trans. Pattern Analysis And Machine Intelligence Vol.13, No.7 July 1991.

Nayar94 S.K. Nayar, Y. Nakagawa "Shape from Focus", IEEE Transactions on Pattern Analysis and Machine Intelligence,1994, Vol.16, No.8, pp.824-831.

Nestares96 O. Nestares, R. Navarro, J. Portilla & A. Tabernero, "Automatic computation of the area irradiated by ultrashort laser pulses in Sb materials through texture segmentation of TEM images" Ultramicroscopy, Vol. 66 No.1-2, pp. 101-115, October 1996.

Neubauer92 C. Neubauer, "Segmentation of Defects in Textile Fabric", Proceedings of 11th IAPR International Conference on Pattern Recognition. Vol. 1 Conference A : Computer Vision and Applications. pp. 688-91.

Nikias93 C.L. Nikias & J.M. Mendel, "Signal Processing withe Higher-Order Spectra", IEEE Signal Processing Magazine, July 1993, pp.10-30.

Noguchi94 M. Noguchi & S.K. Nayar, "Microscopic Shape from Focus Using Active Illumination", Proceedings of 12th IAPR International Conference on Pattern Recognition-Conference A: Computer Vision and Image Processing, Jerusalem 9-13 October 1994, Ch.200, pp. 147-152.

Oden92 P.I Oden, A. Majumdar, B. Bhushan, A. Padmanabha && J.J. Graham, "AFM Imaging, Roughness Analysis and Contact Mechanics of Magnetic Tape and Head Surfaces", Transactions of ASME, Journal of Tribology, October92, Vol.114, pp.666-674.

Ogilvey91 J.A. Ogilvy, "Theory of wave scattering from Random rough surfaces", Publisher Adam Hilger.

Ogilvy89 J. Ogilvy & J. Foster "Rough surfaces: gaussian or exponential statistics? "J. Phy. D Appl. Phys 22 (1989) pp.1243-1251.

Ohanian92 P.P. Ohanian & R.C. Dubes, "Performance evaluation for four classes of textural features", Pattern Recognition, Vol.25, No.8, pp. 819-833, 1992.

Onn90 R. Onn & A. Bruckstein, Integrability Disambiguates Surface Recovery in Two-Image Photometric Stereo", International Journal of Computer Vision 5:1 pp.105-113 (1990).

Oren95 M. Oren & S.K. Nayar, "Generalization of the Lambertian Model and Implications for Machine Vision", International Journal of Computer Vision 14, 227-251 1995.

Papadimitriou95 P. Papadimitriou & T. Dennis, " A stereo disparity algorithm for 3D model construction", IEE Conference on Image Processing and Its Applications, 4-6 July 1995, pp.178-182.

Parsons S.J. Parsons, "Metrology and Gauging", 2nd Edition. Macdonald, 1970.

Peebles P.Z. Peebles Jr. "Probability, Random Variables & Random Signal Principles" 2nd Edition McGraw Hill International.

Pentland84a A.P. Pentland, "Local Shading Analysis", IEEE Trans. Pattern Analysis And Machine Intelligence Vo.6 No.2 March 1984, pp.170-187.

Pentland84b A.P. Pentland, "Fractal-Based Description of Natural Scenes", IEEE Trans. Pattern Analysis And Machine Intelligence, Vol.6, No.6 November 1984, pp.661-674.

Pentland86 A.P. Pentland "Shading into Texture", Artificial Intelligence, 29 (1986) pp.147-170.

Pentland90 A. Pentland, "Linear Shape from Shading", International Journal of Computer Vision, 4, pp.153-162 (1990).

Pentland91 A. Pentland "Photometric Motion", IEEE Trans On Pattern Analysis And Machine Intelligence Vol.13 No.9 September 1991, pp. 879-890.

Picard96 R.W. Picard, “A society of models for video and image libraries”, IBM Systems Journal, Vol.35, Nos.3&4, 1996. pp. 292-312.

Porter97 R. Porter & N. Canagarajah, "Robust rotation invariant texture classification: wavelet, Gabor filter and GMRF based schemes", IEE Proc. Vis. Image Signal Process. Vol. 144, No.3, June1997

Quian92 Shie Quian & Dapang Chen "Joint Time Frequency Methods & Applications" Prentice Hall 1992.

Raghu96 P.P. Raghu & B. Yegnanarayana, "Segmentation of Gabor-Filtered Textures Using Deterministic Relaxation" IEEE Trans. on Image Processing, Vol.2 No.12 December 1996, pp.1625-1635.

Rajaram95 K.V. Rajaram, G. Parthasarthy & M.A. Faruqui, "A Neural Network Approach to Photometric Stereo Inversion of Real-

World Reflectance Maps for Extracting 3-D Shapes of Objects" IEEE Trans. on Systems, Man & Cybernetics, Vol.25, No.9, Sept 1995, pp. 1289-1300.

Randen94 T. Randen & J.H. Husoy, "Multichannel filtering for texture image segmentation", Opt. Eng. Vol. 33 No.8 pp2617-2625, August 1994

Randen95 T. Randen & J.H. Husoy, "Optimal Texture Filtering", International Conference on Image Processing - Proceedings, Vol. I-III, A 1994 pp.374-377.

Randen96 T. Randen, V. Alvestad, J.H. Husoy, "Optimal filtering for unsupervised texture feature extraction" Proc. SPIE Vol.2727 - Visual Communications and Image Processing, March 1996, Pt 1-3 pp.441-452

Reed90 T.R. Reed & H. Wechsler, "Segmentation of Textured Images and Gestalt Organization Using Spatial/Spatial-Frequency Representations" IEEE Trans. Pattern Analysis And Machine Intelligence Vol.12, No.1 January 1990, pp.1-11.

Russell91 K.L. Russell, J.E. McFee & M.R. Ito, "Simulation of images by photometric stereo modeling" Optical Engineering, September 1991, Vol.30 No.9, pp1337-1345.

Sayles78 R.S. Sayles & T.R. Thomas, "Surface Topography as a nonstationary random process", Nature Vol.271, 2 Feb. 1978 pp.431-434.

Sivia D.S. Sivia, "Data Analysis: A Bayesian Tutorial", Oxford University Press 1996

Solomon96 F. Solomon & K. Ikeuchi, "Extracting the Shape and Roughness of Specular Lobe Objects Using Four Light Photometric Stereo", IEEE Trans. Pattern Analysis And Machine Intelligence, Vol. 18, No. 4, April 1996, pp.449-454.

Song92 K.Y. Song, M. Petrou & J. Kittler, "Wigner based crack detection in textured images." International Conference on Image Processing and its Applications 7-9 April 1992, Maastricht, Netherlands, pp.315-318.

Stone94 R.A Stone & S.A. Shafer "Determination of surface roughness from reflected step edges" J. Opt. Soc. Am. A Vol. 11, No.11, Nov. 1994, pp.2969-2980.

Strat79 T. Strat, "A Numerical method for shape from shading for a single image" S.M. Thesis, Dept. of Electrical Engineering and Computer Science. MIT Cambridge, MA.

Subbarao94 M. Subbarao & M.C. Lu, "Image sensing model & computer simulation for CCD camera systems" Machine Vision and Applications, Vol.7, No.4 pp.272-289, 1994.

Sziranyi96 T. Sziranyi "Robustness of Cellular Neural Networks in Image Deblurring and Texture Segmentation.", International Journal of

Circuit Theory and Applications, Vol. 24 pp381-396 (1996)

Tagare90 H.D. Tagare and R.J. deFigueirdo, "Simultaneous Estimation of Shape and Reflectance Maps from Photometric Stereo", ICCV 90 pp. 340-343

Tagare91 H.D. Tagare & R.J. deFigueirdo, "A Theory of Photometric Stereo for a Class of Diffuse Non-Lambertian Surfaces", IEEE Trans Pattern Analysis And Machine Intelligence, Vol.13 No.2, February 1991, pp.133-151.

Tagare93 H.D. Tagare & R.J. deFigueirdo, "A Framework for the Construction of Reflectance Maps for Machine Vision" CVGIP Image Understanding, Vol. 57 No.3 May pp. 265-282, 1993.

Talmi80 Y. Talmi & R.W. Simpson, "Self-scanned photodiode array: a multichannel spectrometric detector", Applied Optics, Vol.19, No.9, 1 May 1980.

Tang95 H.W. Tang, V. Srinvasan & S.H. Ong, "Texture segmentation via nonlinear interactions among Gabor feature pairs", Opt. Eng. Jan. 95, Vol. 34 No. 1 pp. 125-134

Tatsanis92 M.K. Tsatanis & G.B. Giannakis, "Object and Texture Classification Using Higher Order Statistics" IEEE Pattern Analysis And Machine Intelligence Vol. 14 No.7 July 1992, pp. 733-749.

Terrillon96 J.C. Terrilon "Image Processing for rotation-invariant pattern recognition in the presence of signal-dependent noise", Applied Optics 10 April 96/ Vol. 35 No.11, pp.1879-1893.

Thau90 R.S. Thau "Illuminant Precompensation for Texture Discrimination using Filters" Image Understanding Workshop (DARPA) Pittsburg PA, Sept 1990 pp. 179-184. Publ. Morgan Kaufmann ISBN 1-55860-140-6.

Therrien C.W. Therrien "Discrete Random Signals and Statistical Signal Processing" Prentice-Hall 1992.

Thomas G.B. Thomas & R.L. Finney, "Calculus and Analytical Geometry", 7th Ed. Addison-Wesley.

Torrance67 K.E. Torrance "Theory of Off-Specular reflection from roughened surfaces", J. Opt. Soc Am. 57 1105 (1967) pp.32-41.

Tuceryan92 Tuceryan & A.K. Jain, "Texture Analysis", appears in "Handbook of Pattern Recognition and Computer Vision" Eds Chen, Pau and Wang, World Scientific, 1993 pp.235-276.

Turing52 A. Turing, "The Chemical Basis of Morphogenesis," Philosophical Transactions of the Royal Society (B), Vol.237, 1952, pp.37-72.

Turk91 G. Turk, "Generating Texture on Arbitrary Surfaces Using Reaction/Diffusion" Computer Graphics Vol. 25 No.4 July 1991, pp. 289-298

Underwood86 E.E. Underwood & K. Banerji, "Fractals in Fractography" Materials Science and Engineering. Vol.80 (1986) pp. 1-14

Unser90 M. Unser & M. Eden, "Nonlinear operators for improving texture segmentation based on features extracted by spatial filters" IEEE Transactions on Systems, Man and Cybernetics, Vol.20 No.4 July/August 1990, pp. 804-815

Unser92 M. Unser, A. Aldroubi and M. Eden "On the Asymptotic Convergence of B-Spline Wavelets to Gabor Functions" IEEE Transactions on Information Theory, Vol.38 No.2 March 1992, pp.864-872.

Unser95 M. Unser, "Texture Classification and segmentation using wavelet frames", IEEE Trans. on Image Processing, 1995, V.4 No.11, pp.1549-1566.

Wang96 C.C. Wang & C.C. Chen "Low-cost subpixel detection method for phase drift of camera/grabber signal conversion" Opt. Eng. 35 (12) pp. 3385-3391 (December 1996).

Watt94 Alan Watt "3d Computer Graphics" 2nd Edition, Addison-Wesley.

Weldon96 T.P. Weldon, W.E. Higgins & D.F. Dunn, "Efficient Gabor Filter Design for Texture Segmentation", Pattern Recognition Vol.29, No.12, pp. 2005-2015, 1996.

Witkin91 A. Witkin & M. Kass, "Reaction Diffusion Textures" Computer Graphics Vol. 25 No.4 July 1991, pp. 299-308

Wolff94 L.B. Wolff "Diffuse-reflectance model for smooth dielectric surfaces" J. Opt. Soc. Am A/ Vol.11, No.11/November 1994.

Wolff94a L.B. Wolff, "Relative brightness of specular and diffuse reflection" Optical Engineering, January 1994, Vol.33 No.1 pp285-293.

Woodham80 R.J. Woodham, "Photometric method for determining surface orientation from multiple images" Optical Engineering Jan./Feb. 1980 Vol. 19 No.1 pp.139-144.

Woodham94 R.J. Woodham, "Gradient and curvature from the photometric-stereo method, including local confidence estimation", J. Opt. Soc. Am Vol. 11 No. 11 November 1994, pp.3050-3068.

Zheng91 Q. Zheng & R. Chellappa, "Estimation of Illuminant Direction, Albedo, and Shape from Shading." IEEE Trans. Pattern Analysis And Machine Intelligence Vol.13, No.7, July 1991.

Zhu93 Y.M. Zhu, R. Goutte & M. Amiel, "On the use of two-dimensional Wigner-Ville distribution for texture segmentation", Signal Processing 30 (1993) pp. 329-353