
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 foreshatter 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 Lambertian 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 Figuerdo [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:

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

If the surface is illuminated from $\tau = 0^\circ$, 90° or 180° these simplify to equations 8.3.2b-d respectively:

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

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

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

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

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

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 .

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

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

these may be transposed to give:

$$p = \frac{1 - 2i_p}{\tan \sigma} \quad (8.3.2h)$$

$$q = \frac{1 - 2i_q}{\tan \sigma} \quad (8.3.2i)$$

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 noted we 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.

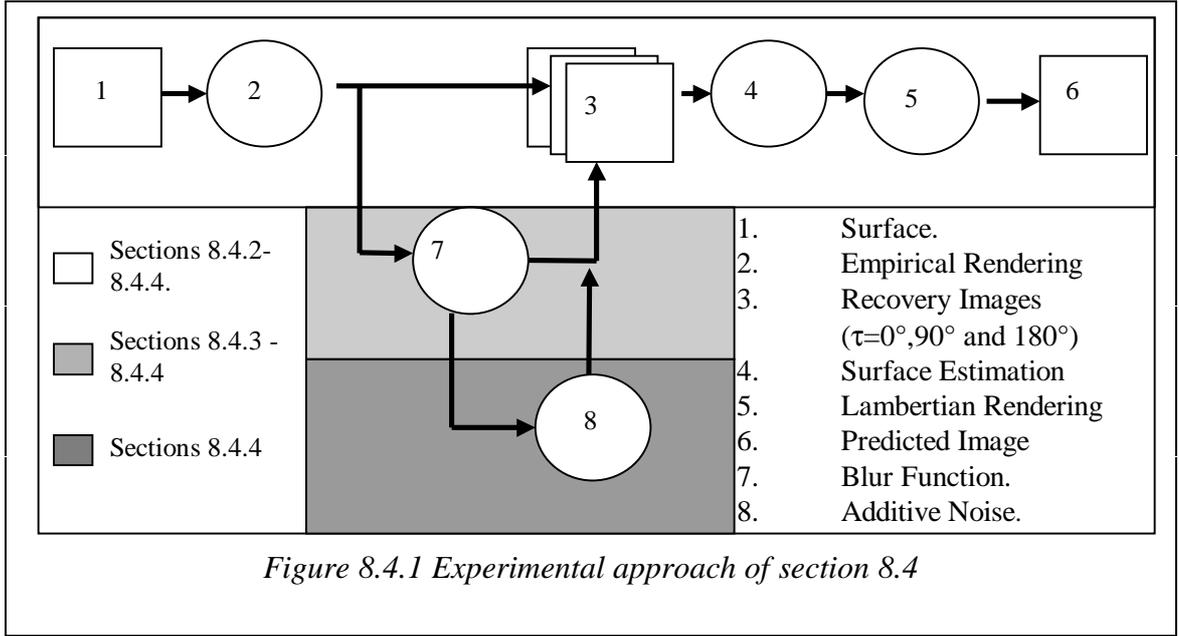


Figure 8.4.1 Experimental approach of section 8.4

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:

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

where

$\text{Var}(x)$ is the variance of the process x .

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

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.

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

8.4.2 Non-ideal reflectance

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

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

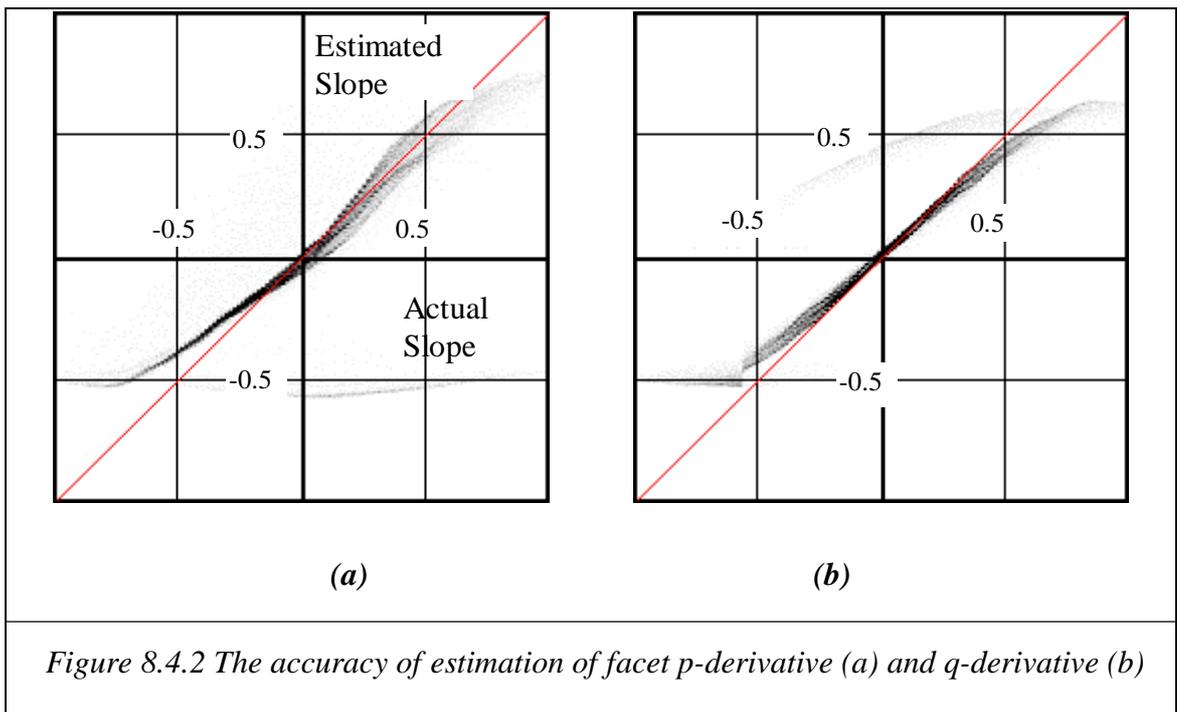
We make three criticisms of the scheme based on this fact:

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

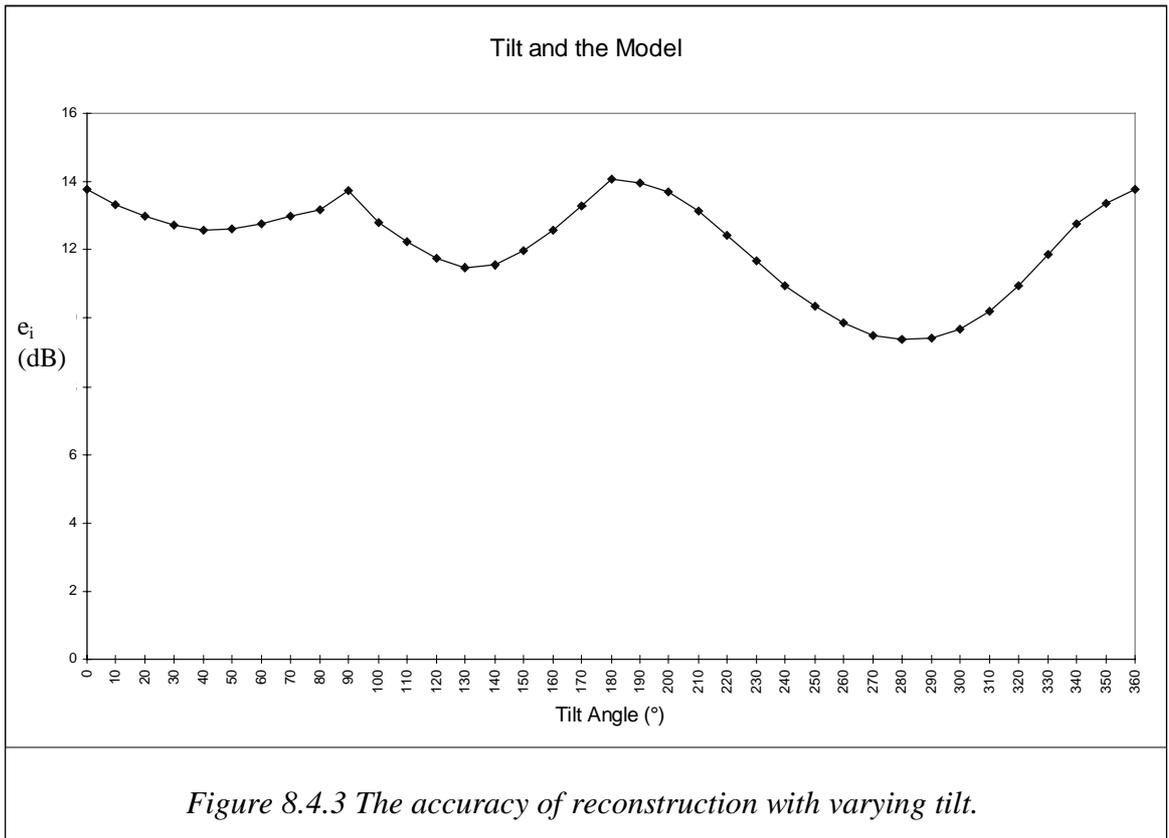
As an initial test of the accuracy of this method we estimate the derivative of an isotropic fractal ($\beta=3.0$) surface with *rms slope* 0.35 illuminated from a slant angle of 60° .



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*).



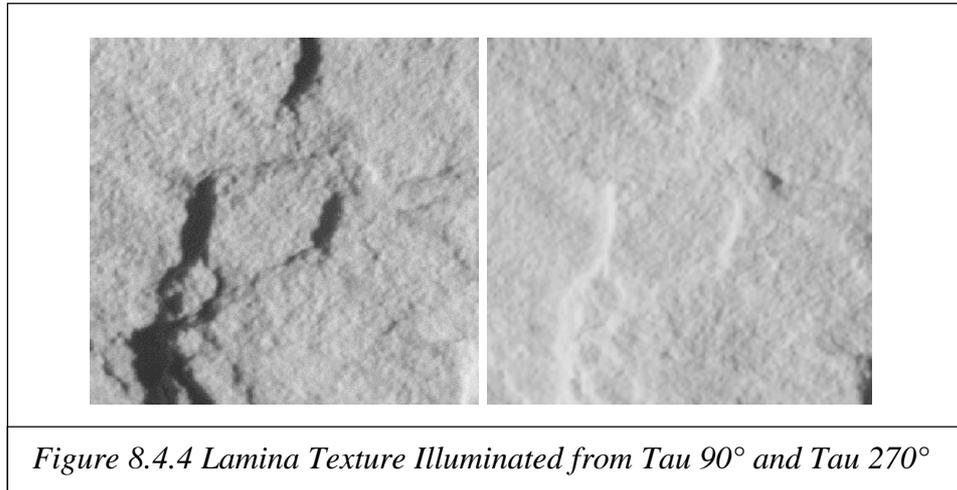
In fact, the variation is considerable, maxima occur at $\tau=0^\circ, 90^\circ$ and 180° . In the work that follows two values of image S/R (e_i), at $\tau=0^\circ$ and 270° , will be quoted. Due to imaging constraints the experimental data images were only captured in the range 0° to 180° we would therefore expect the actual accuracy to be significantly better than that obtained by simulation at $\tau=270^\circ$. 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

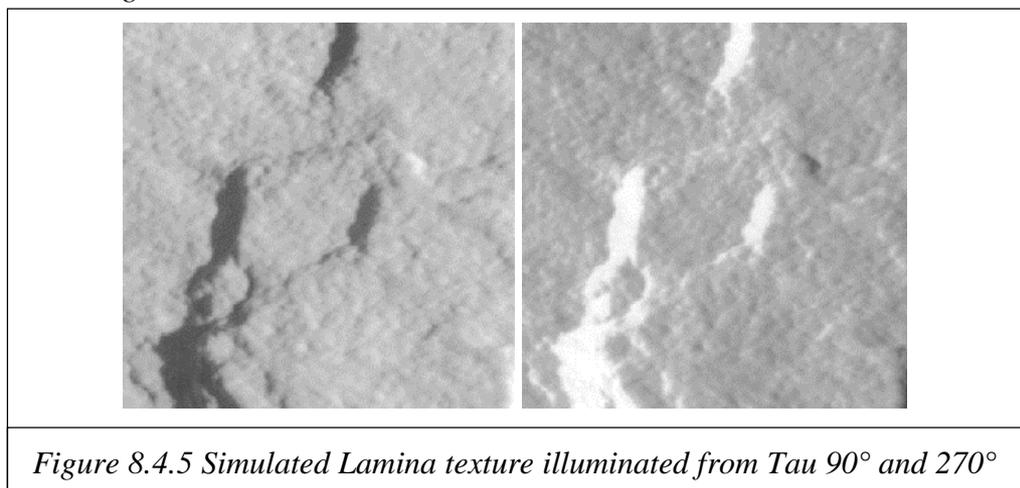
The technique does not take into account the clipping of negative values returned by the Lambertian function; self-shadowed areas are, in effect, assumed to have negative intensities. The model-based technique, being inherently 2.5d, also fails to account for cast shadows in either the recovery or the 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.



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*.



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.

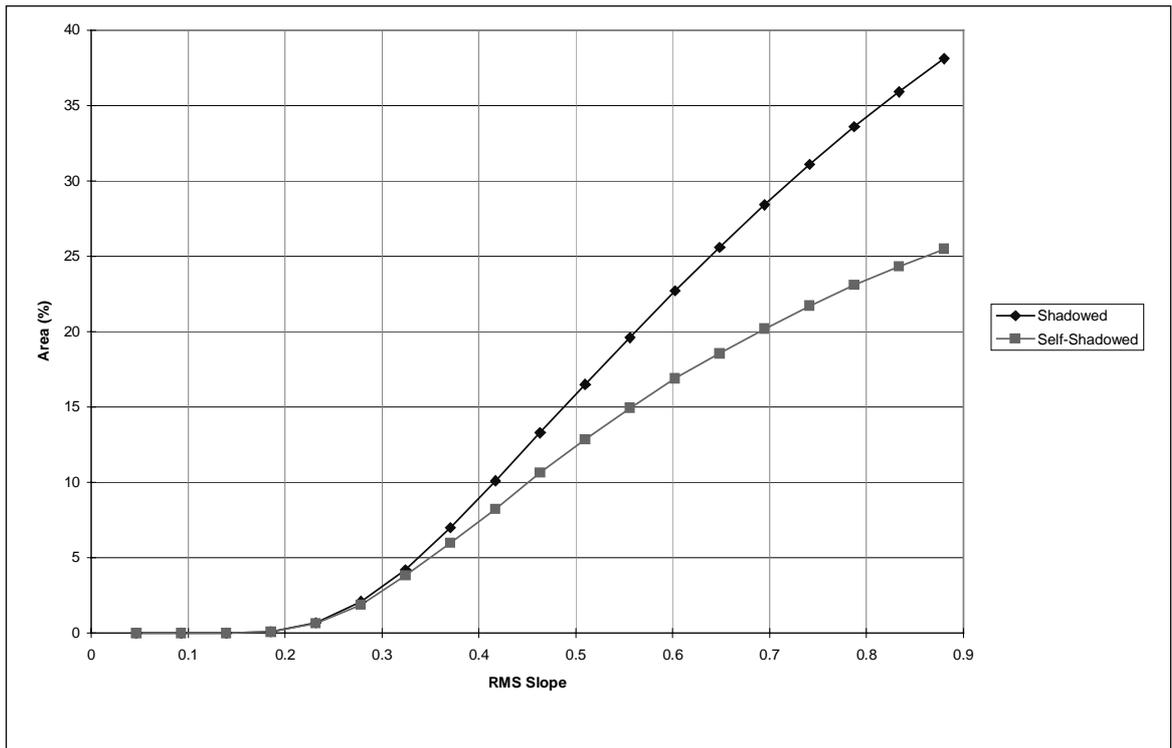


Figure 8.4.6 The effect of slope on the degree of shadowing of an isotropic fractal surface.

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.

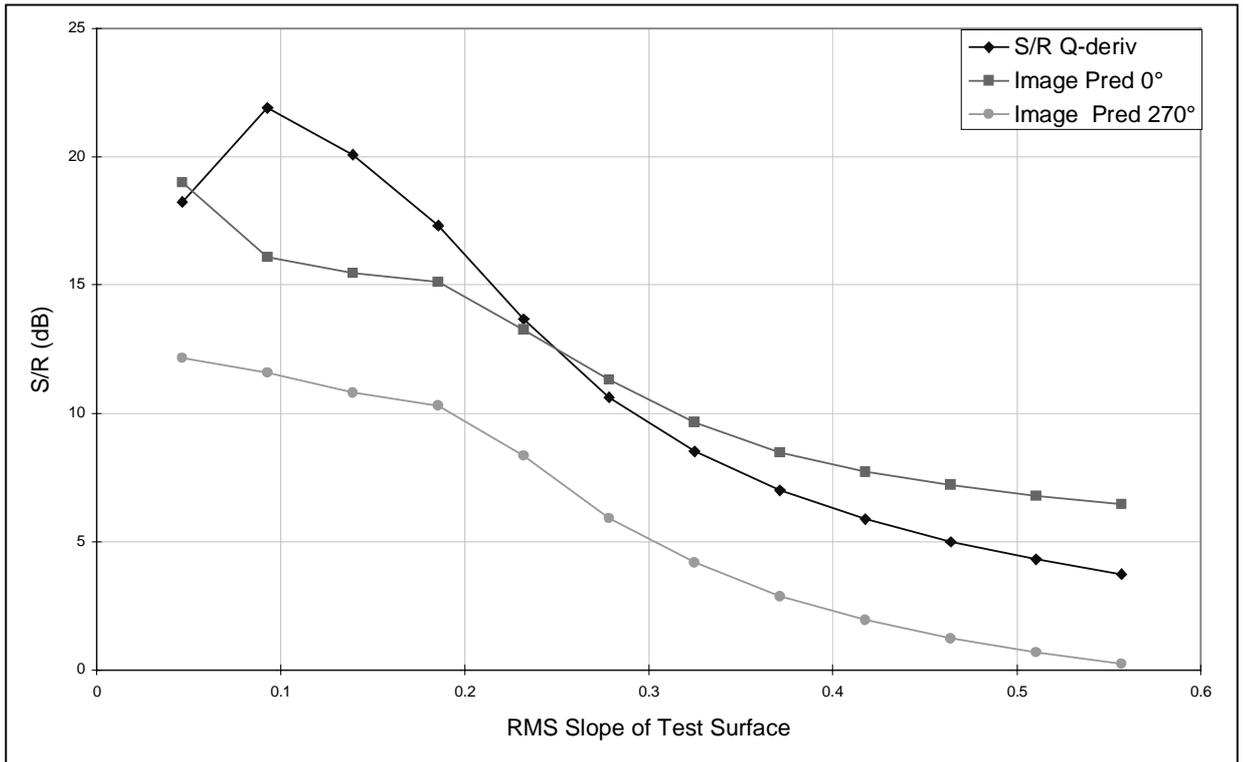


Figure 8.4.8 The Effect of Surface Roughness on Surface and Image Prediction.

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 $\tau=0^\circ$ is actually higher than that of the surface; though the same is not true of the $\tau=270^\circ$ 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

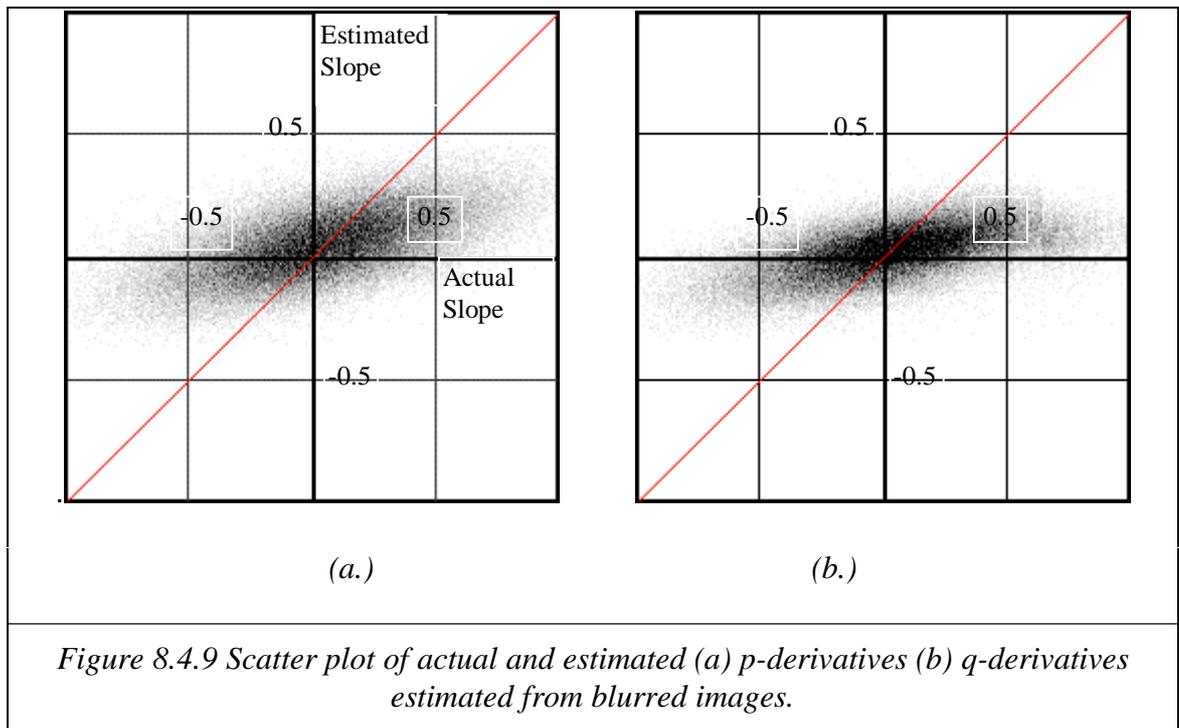
The surface slope clearly imposes limits on the surfaces to which the technique can be applied. However, the technique is able to maintain a prediction accuracy of greater than 4dB for even the roughest surfaces considered in this thesis. Another interesting point to emerge is that while surface recovery 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.



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.

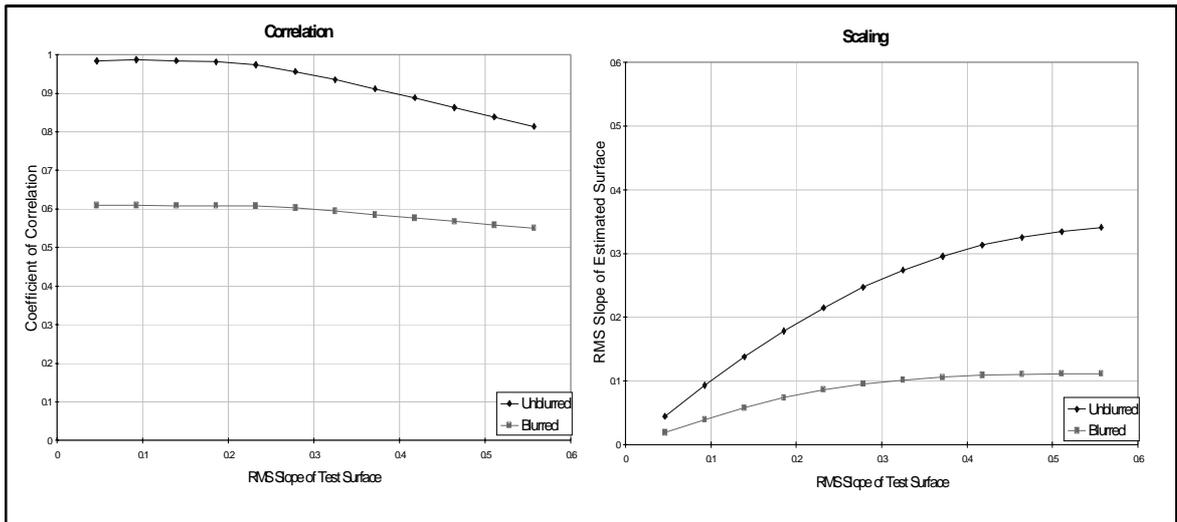


Figure 8.4.10 Relationship between actual and estimated rms slope from blurred images.

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.

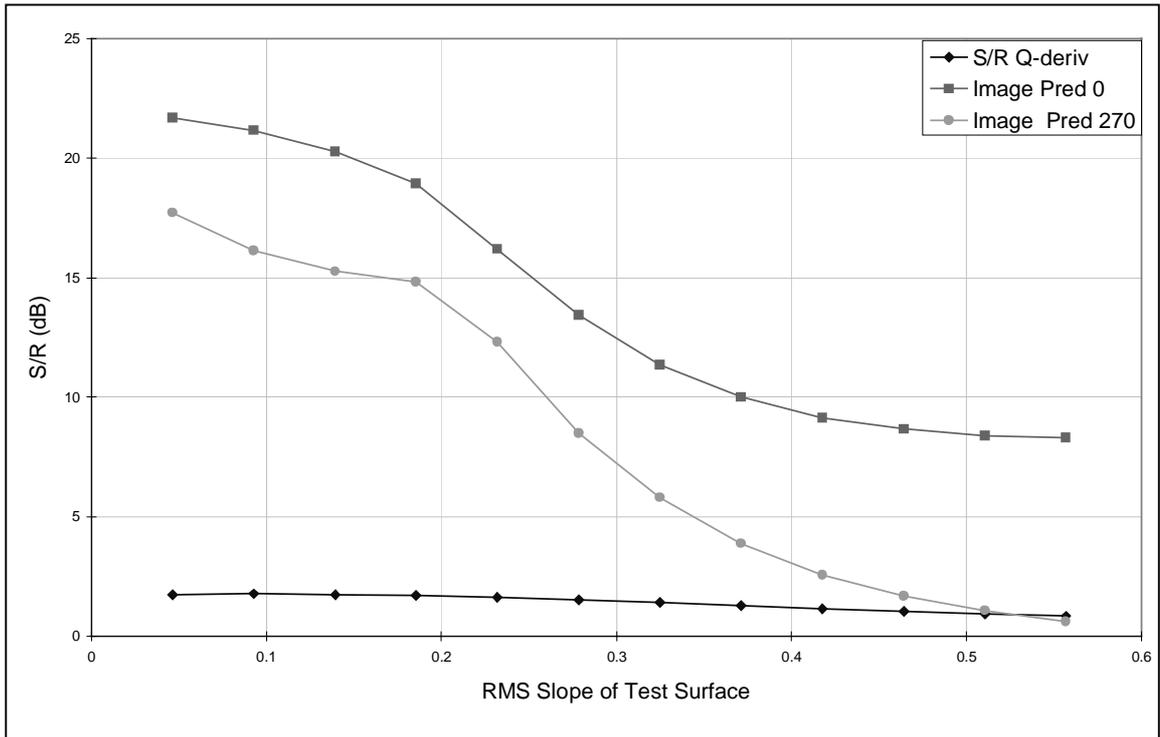


Figure 8.4.11 The effect of blurring on surface and image prediction.

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.

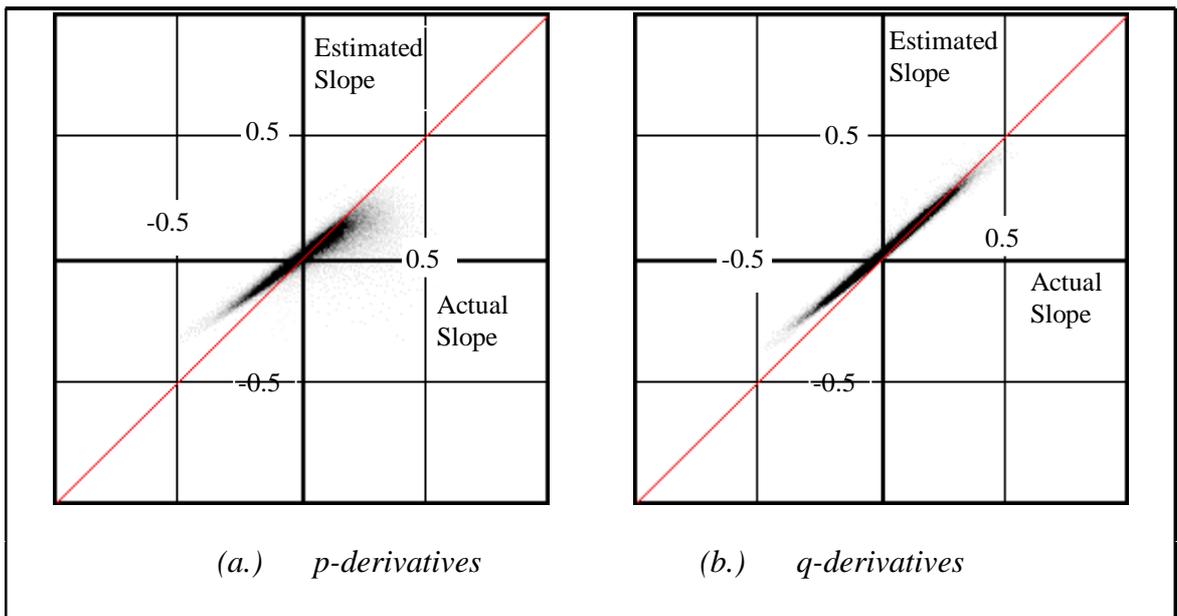


Figure 8.4.12 The effect of image blur on surface reconstruction.

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.

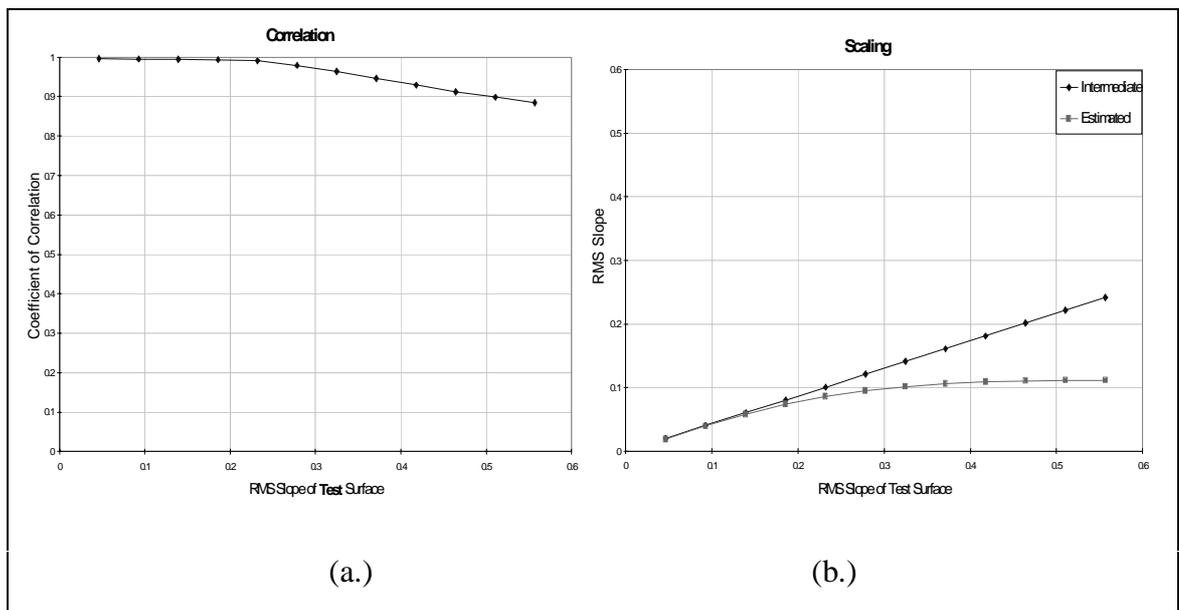


Figure 8.4.13 Comparison of statistics of a low pass filtered (intermediate) surface with those of a surface estimated from blurred images.

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*.

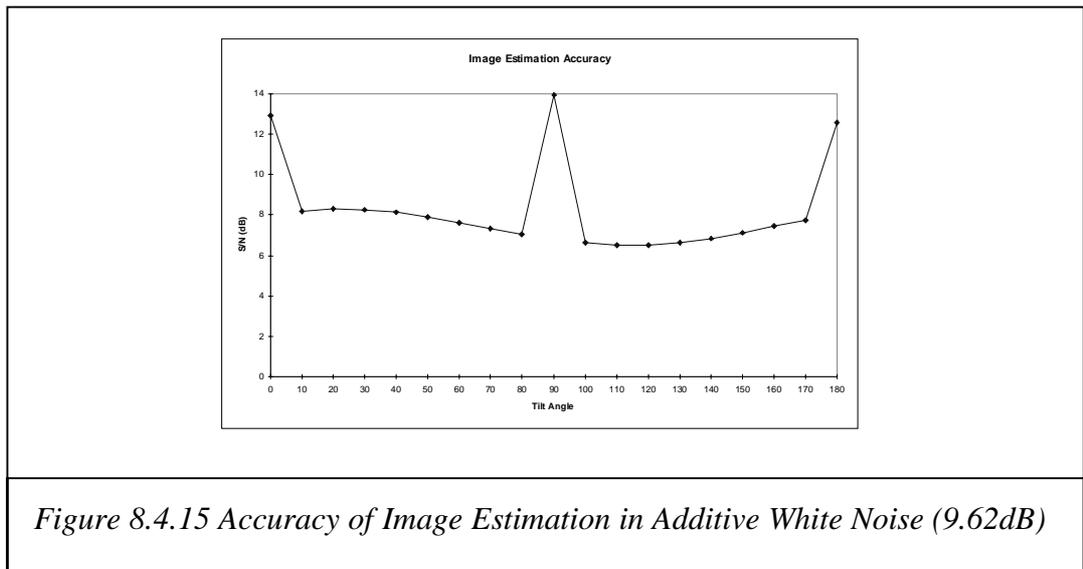
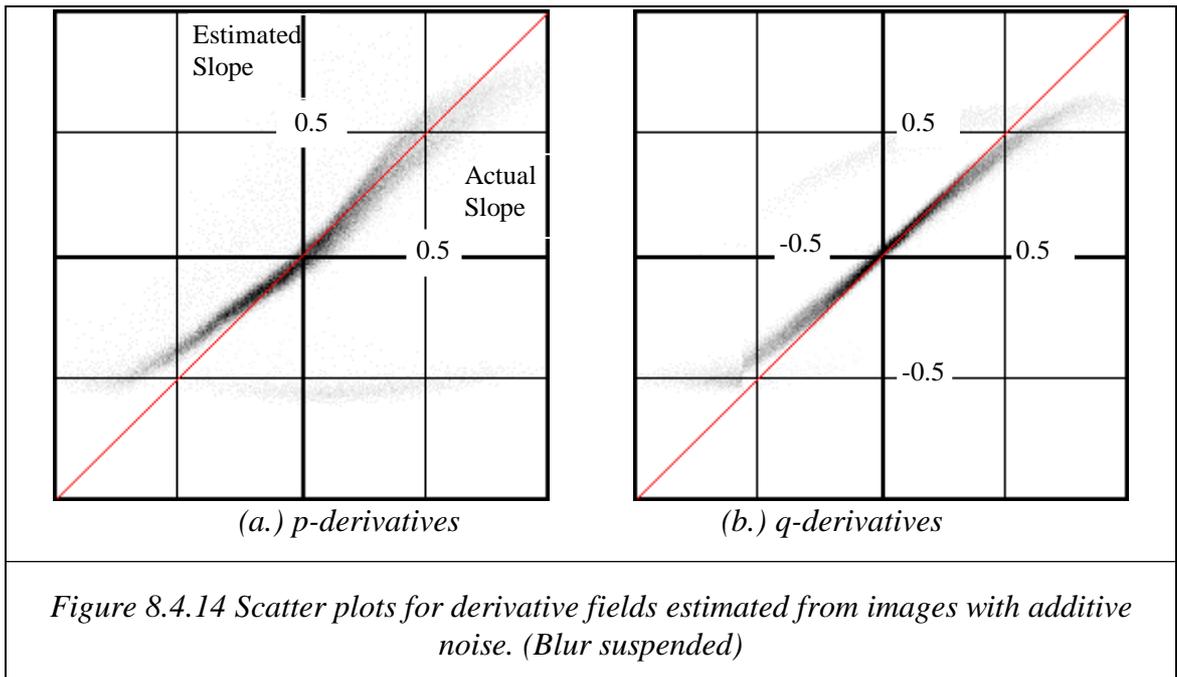
The effect of image blur on the accuracy of image prediction, while significant, is not critical. The effect on derivative recovery, as measured by our criteria, is so serious as to call into question our claim that our simulated images are estimated from a basis which is physically meaningful.

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 arbitrary, more physically-based alternative to the ideal filter assumed in the earlier theoretical work.

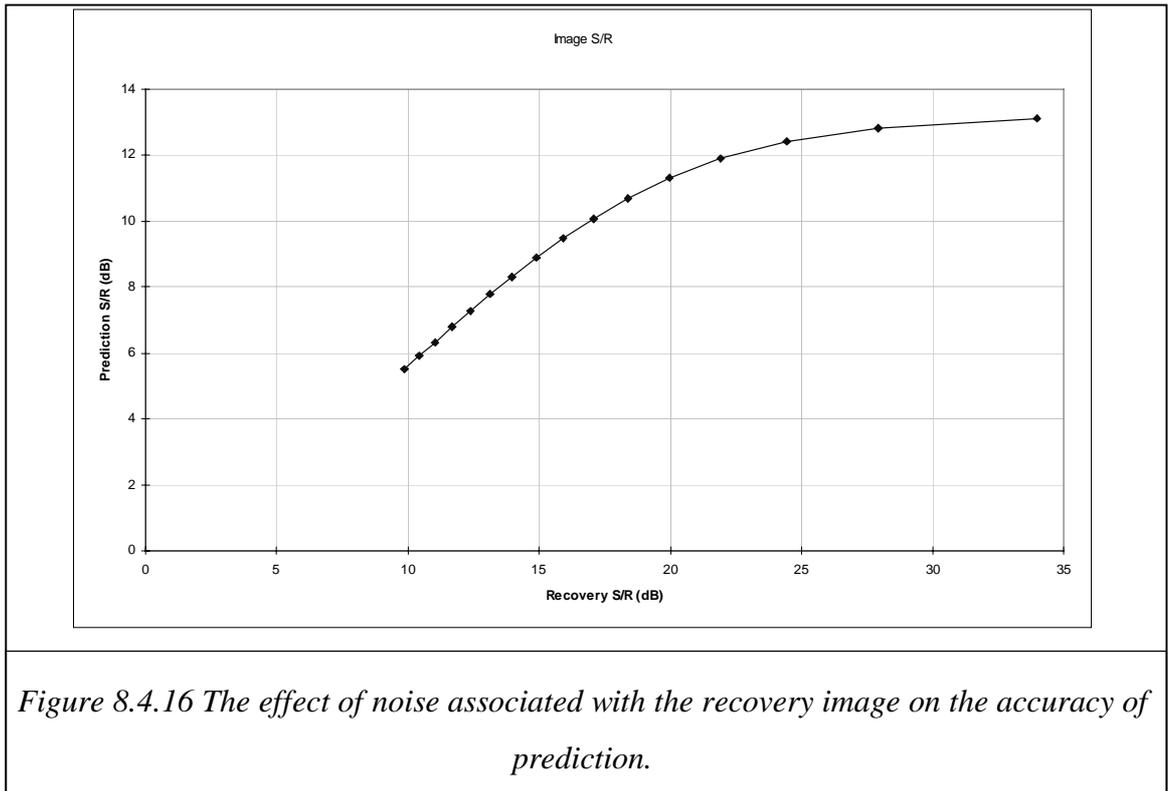
8.4.4 The Effect of Temporally Varying Additive Noise

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 it 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*



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 $\tau=50^\circ$ and $\tau=140^\circ$. The comparison images are corrupted with noise of the same variance as the recovery images.



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:

$$\text{Curl}\hat{\mathbf{S}} \neq 0$$

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.

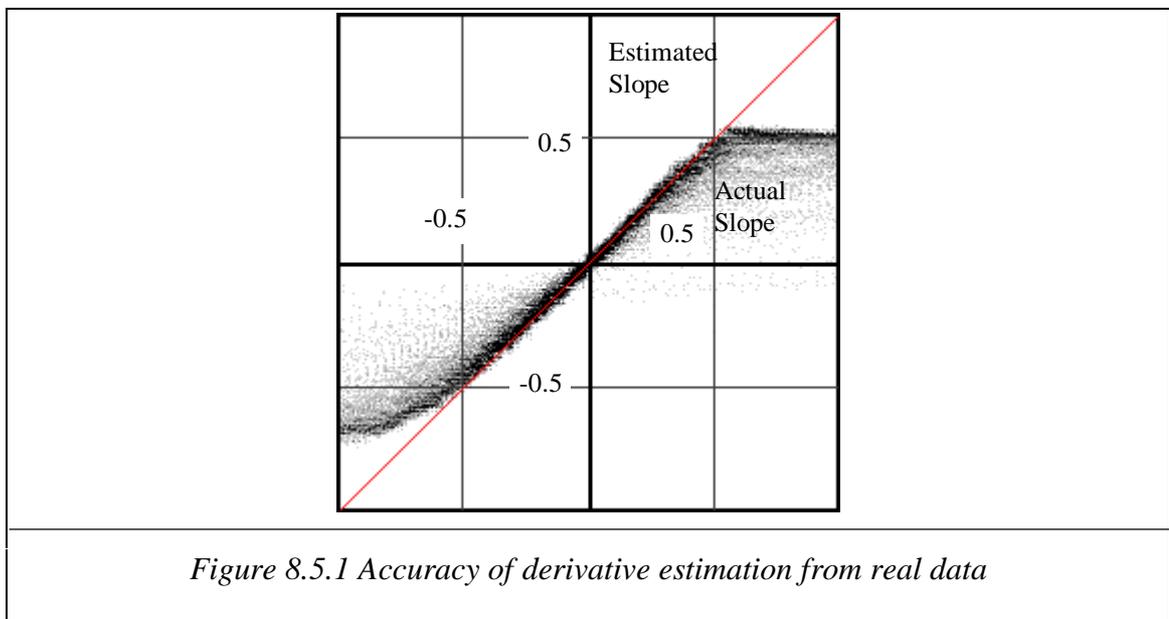
In fact, the difference between the two is academic, with one exception: the ability to generalise. Within the bounds of the work presented in this thesis this effectively means the algorithm must be able to maintain an adequate level of accuracy for all tilt conditions. By using the $\tau=270^\circ$ 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.

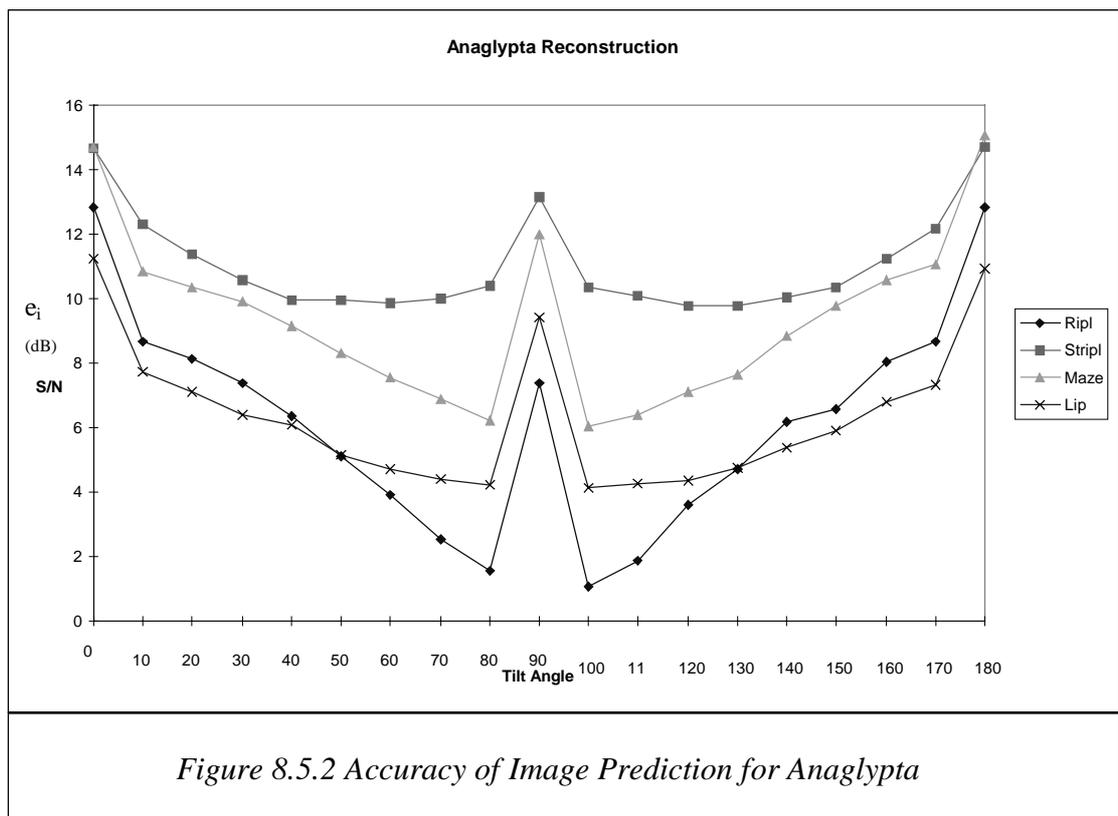


Both extrema show saturation, this being most prominent for large positive (i.e. self-shadowed) slopes. In fact within the slope range ± 0.5 , which is the region of relevance to texture modelling, the relationship is highly linear with little spread. The spread or 'vapour trail' effect corresponds to 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.

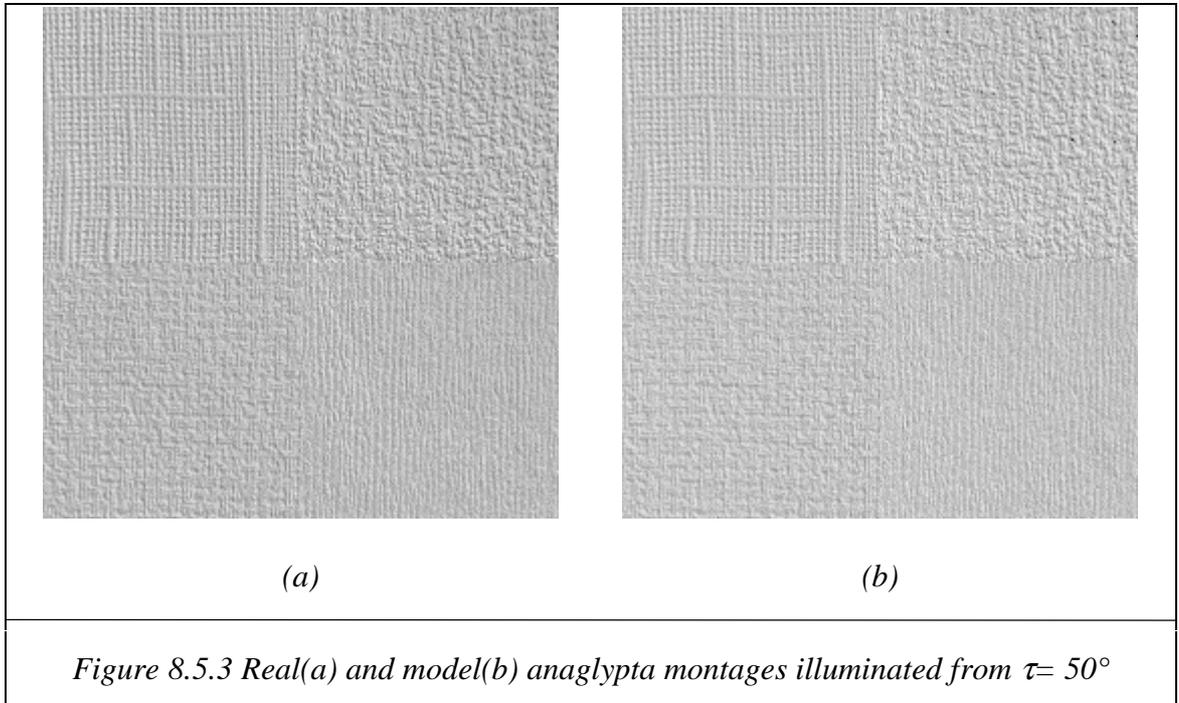
8.5.2 Image Prediction

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.

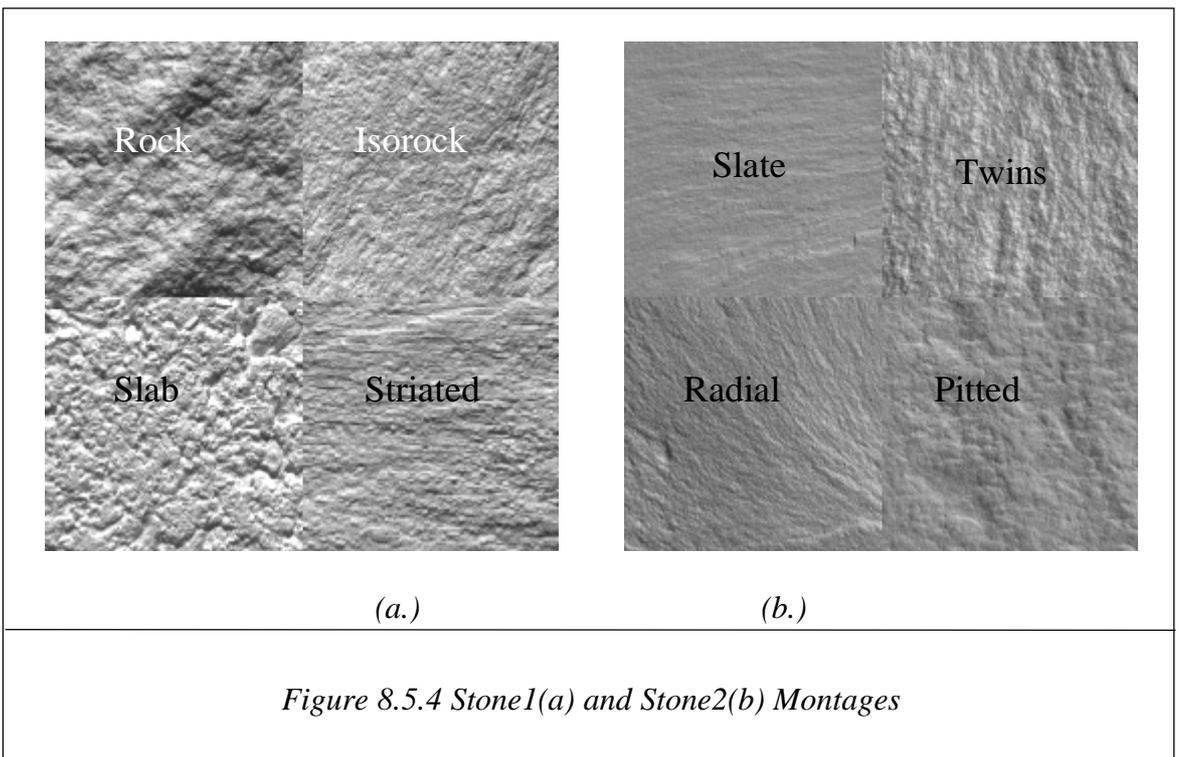


The real anaglypta surfaces illuminated from $\tau=50^\circ$ and the simulated image are shown in *Figure 8.5.3*. While there is some difference in the average levels of the individual textures, the image textures have been predicted well. It is interesting to note that the algorithm is able to model 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.



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.



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)

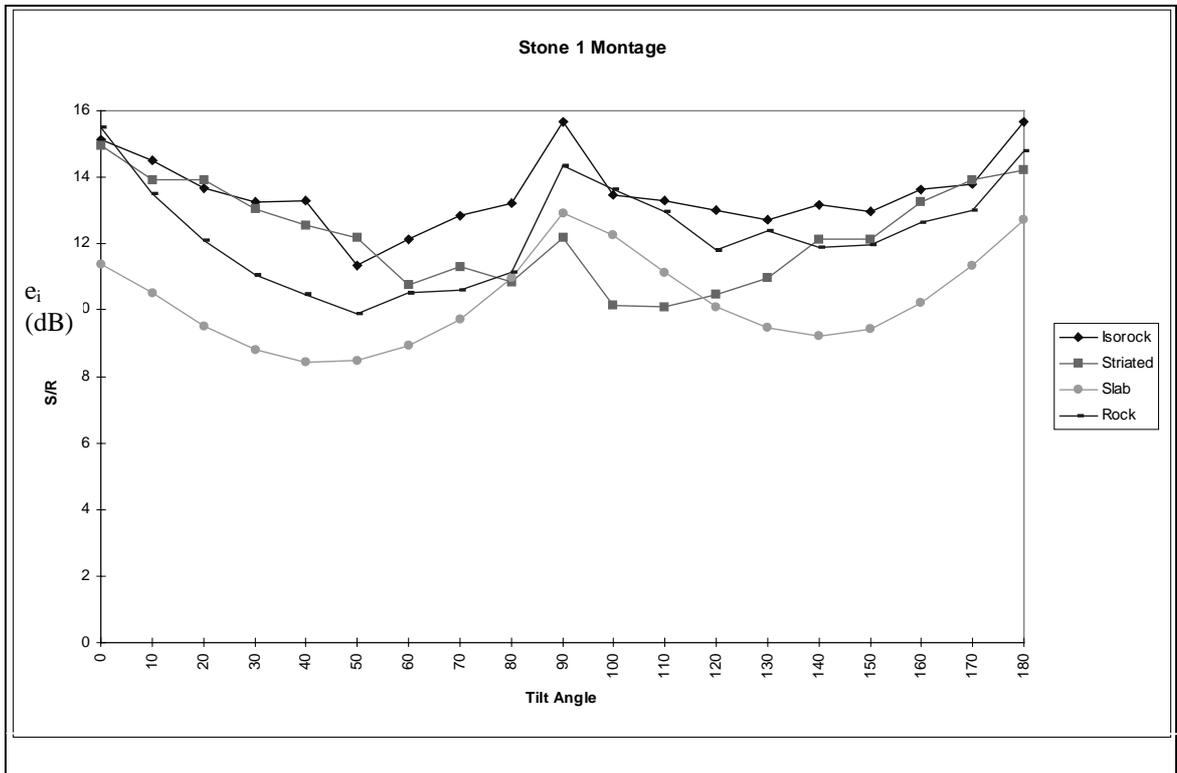
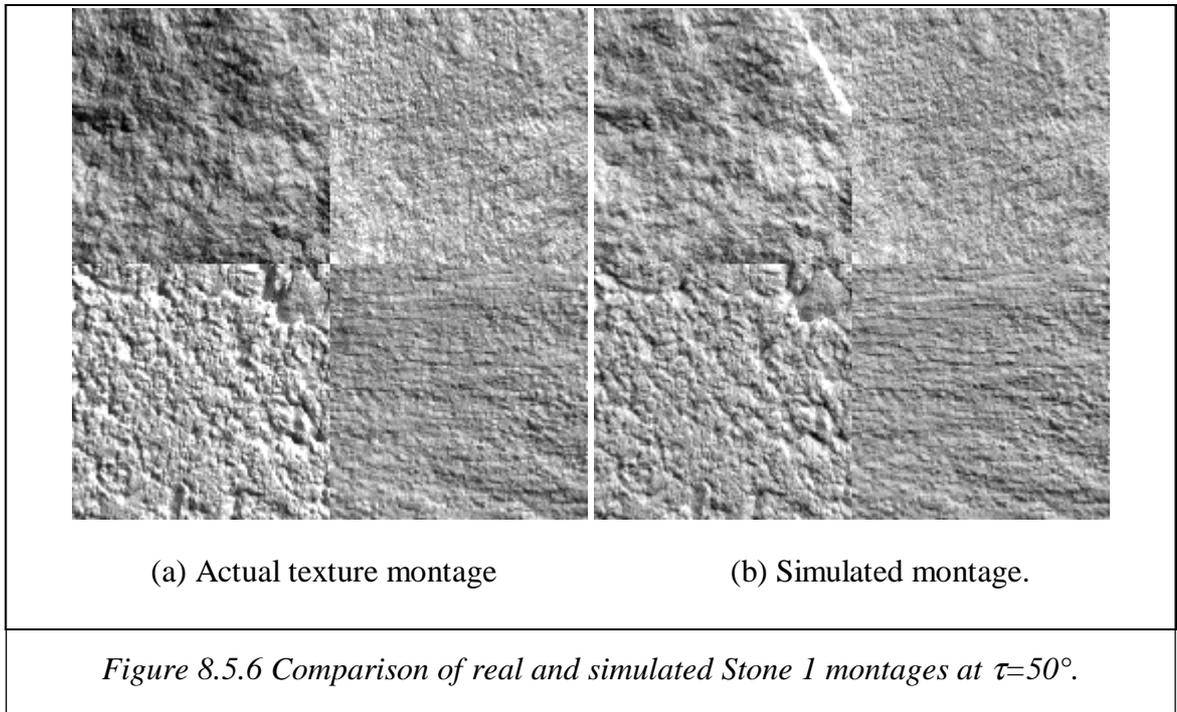


Figure 8.5.5 Prediction accuracy for Stone 1 textures.

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.



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

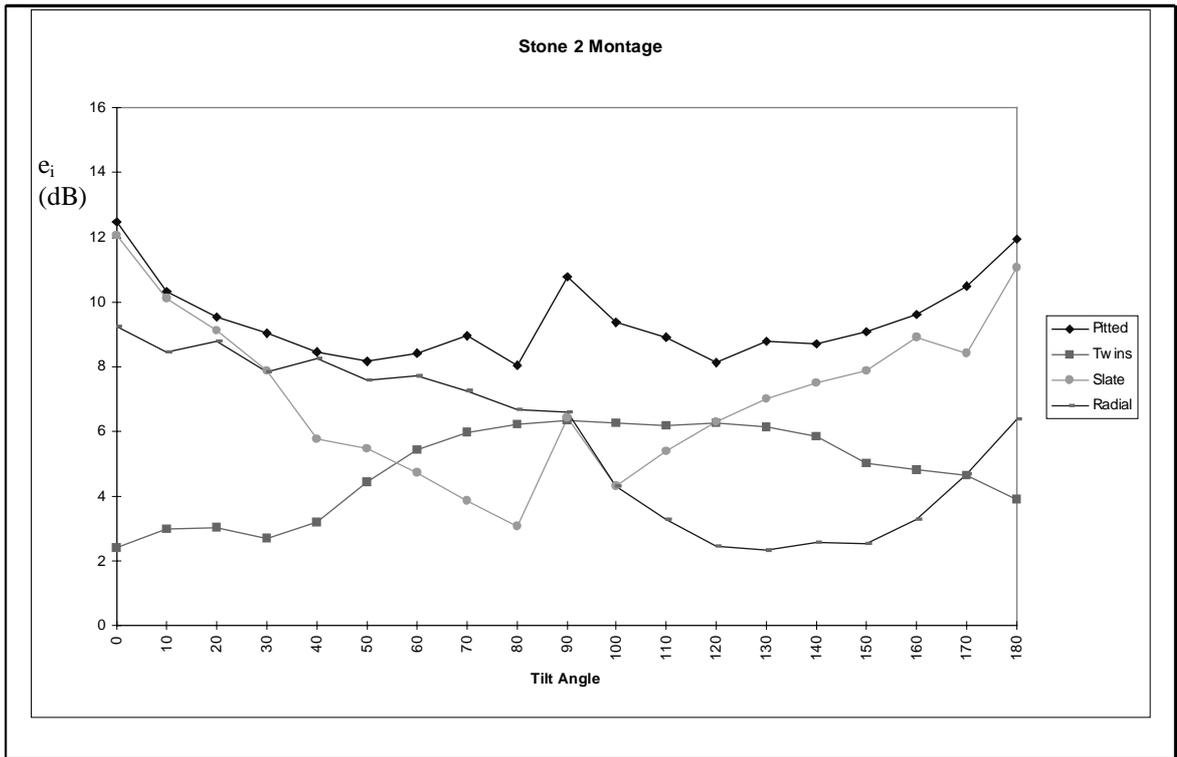


Figure 8.5.7 The Accuracy of Image Prediction for Rock Textures (Recovered and Simulated at $\sigma=60^\circ$)

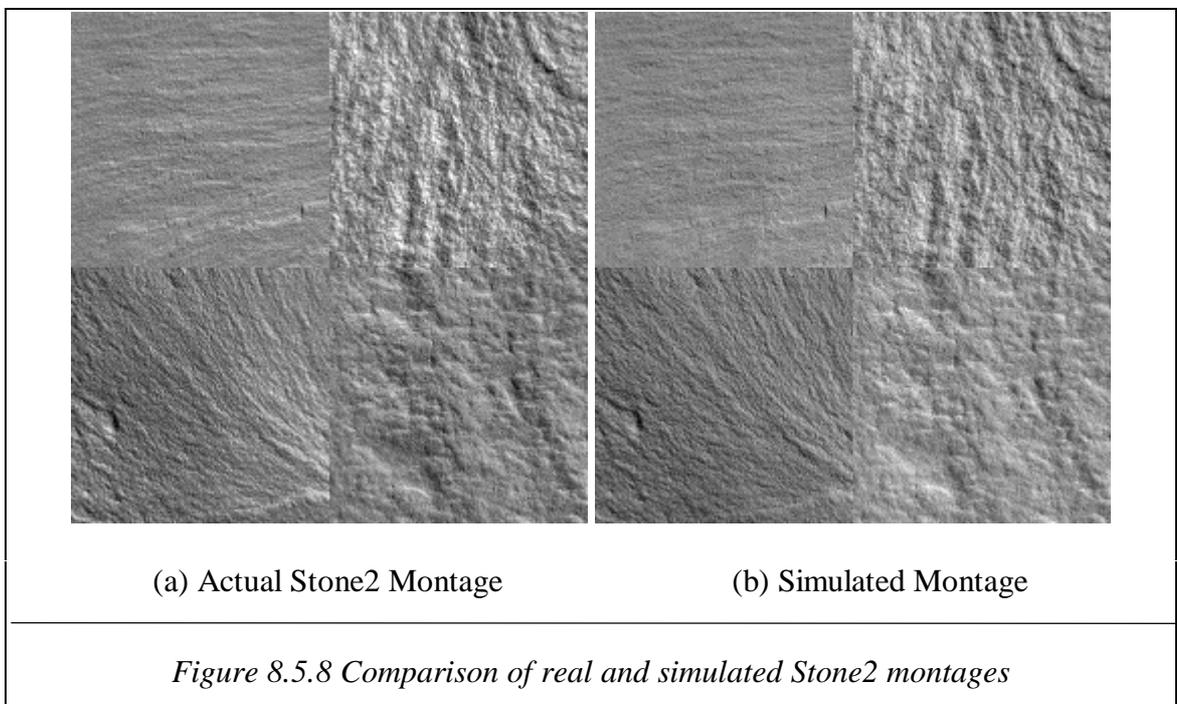


Figure 8.5.8 Comparison of real and simulated Stone2 montages

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 but 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.

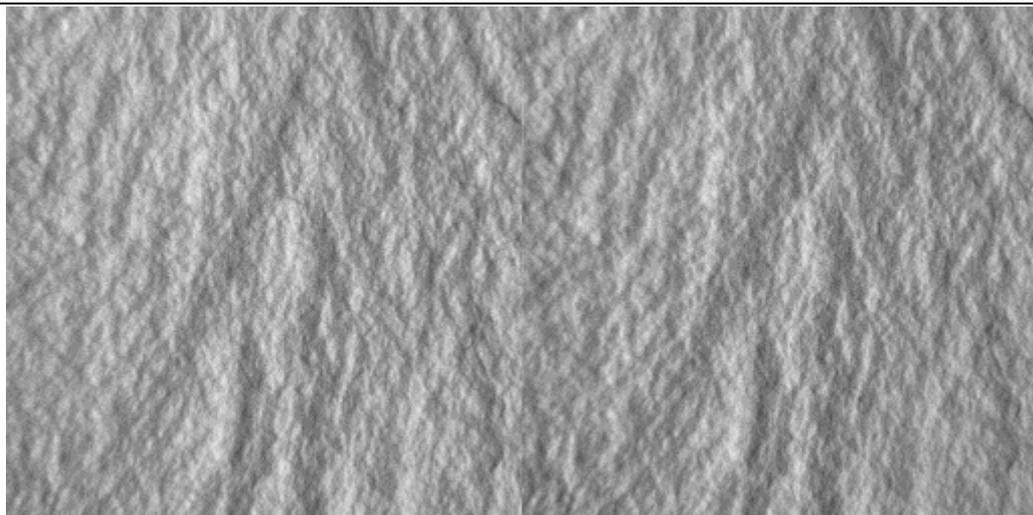
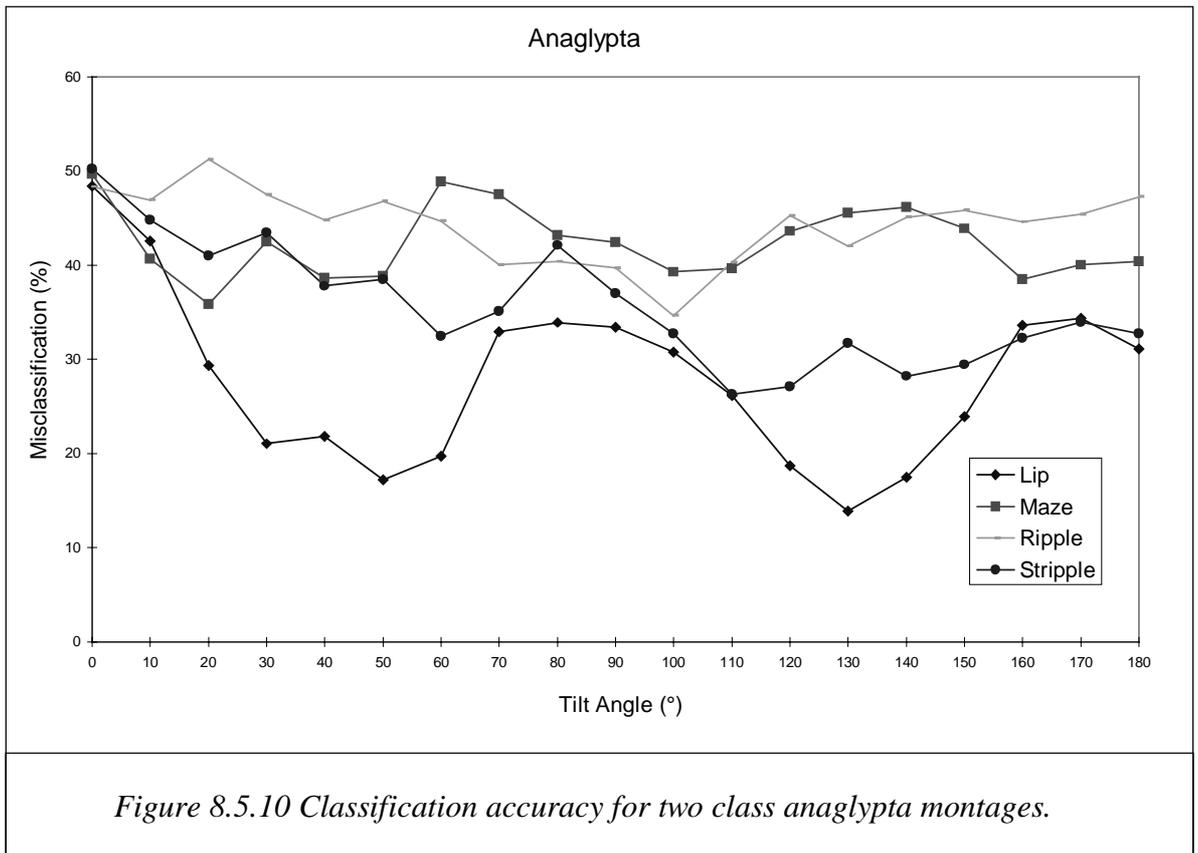


Figure 8.5.9 Montage comprising real (left) and simulated(right) 'Twins' textures.

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 separable, 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.



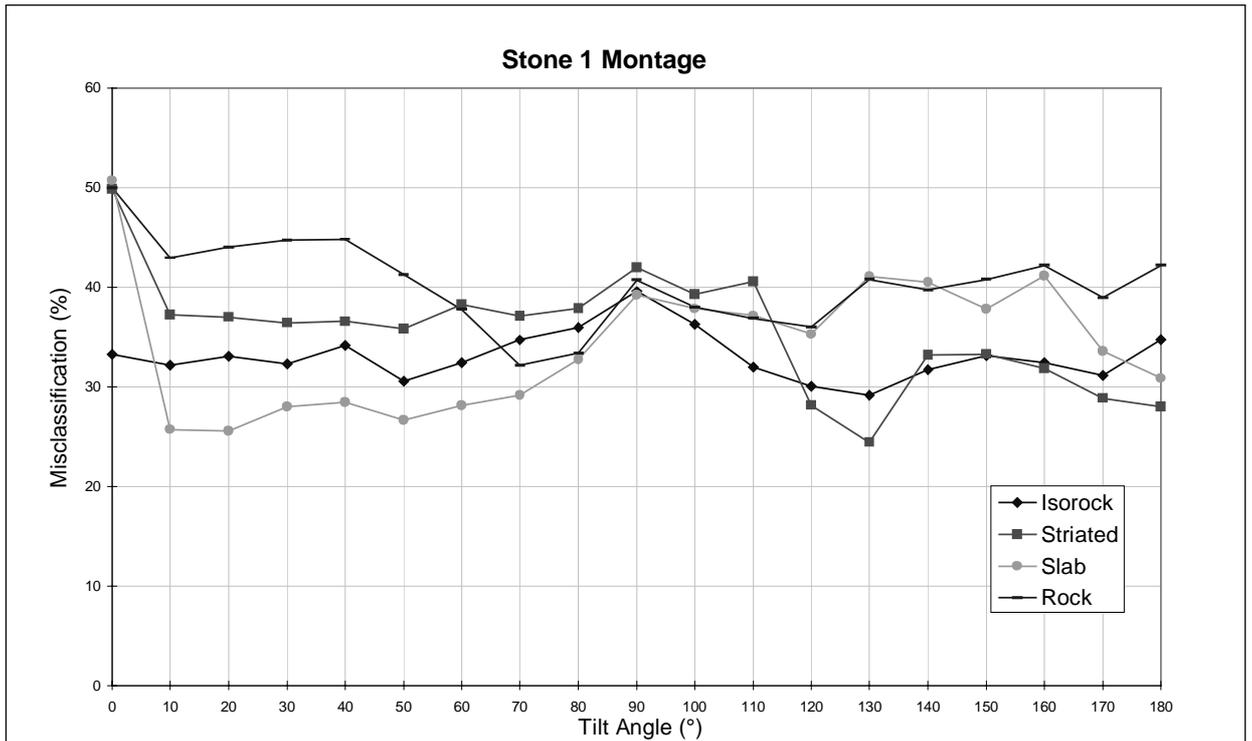


Figure 8.5.11 Classification accuracy for 2 class Stone1 texture montages.

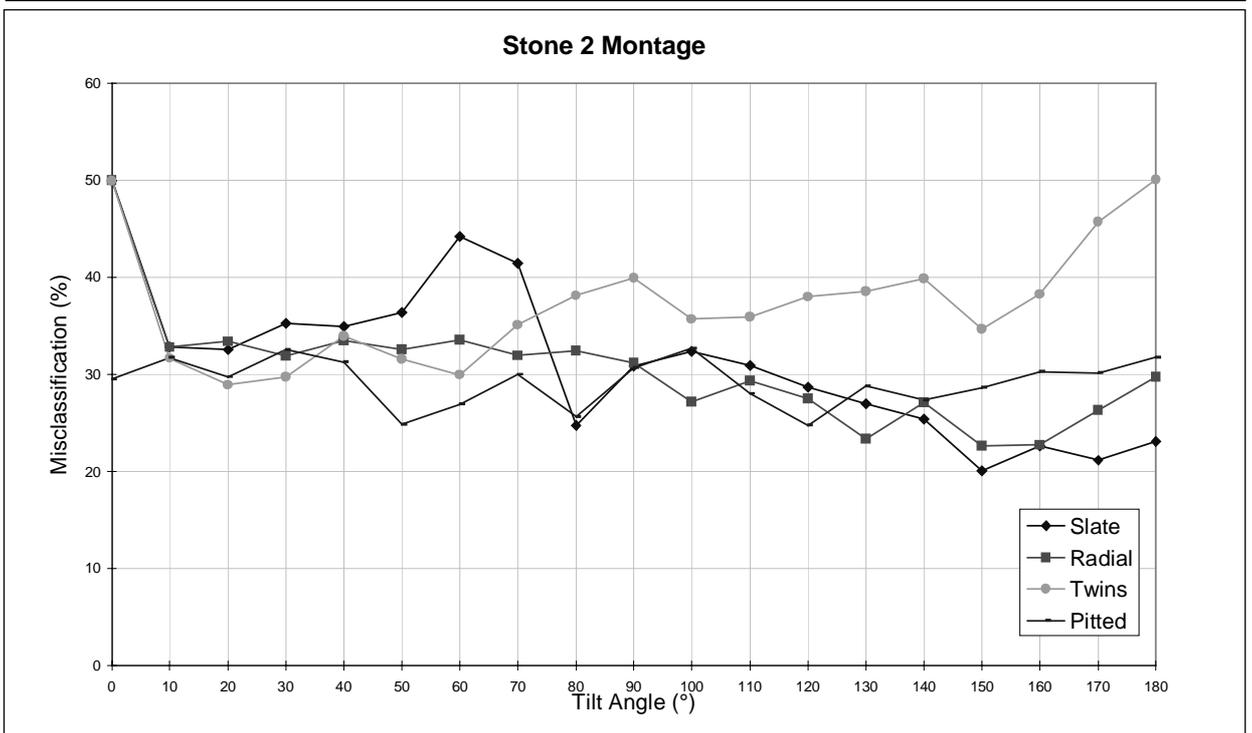
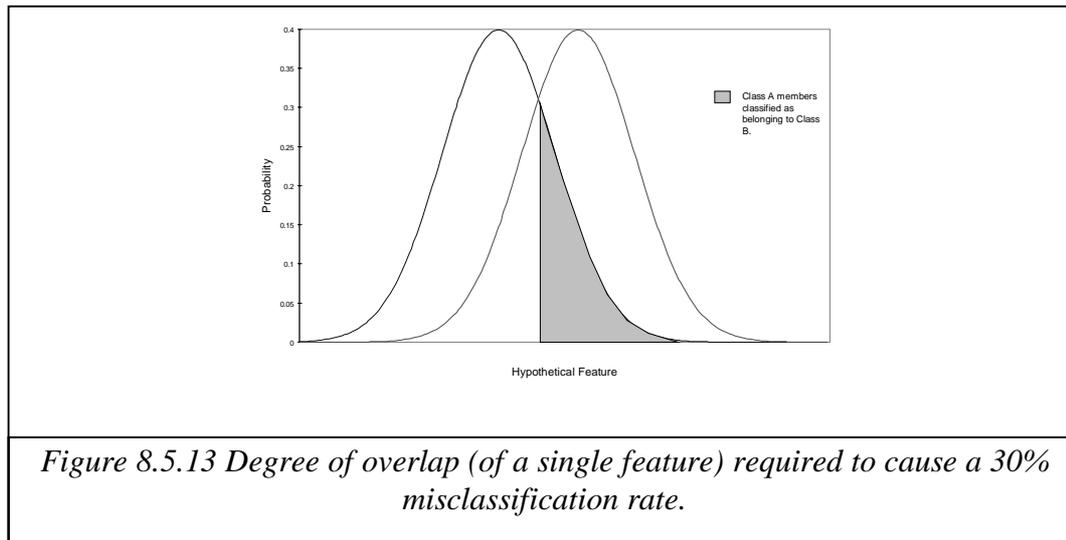


Figure 8.5.12 Classification accuracy for Stone 2 two class montages.

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 $\tau=0^\circ$ 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.



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 notwithstanding, 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.

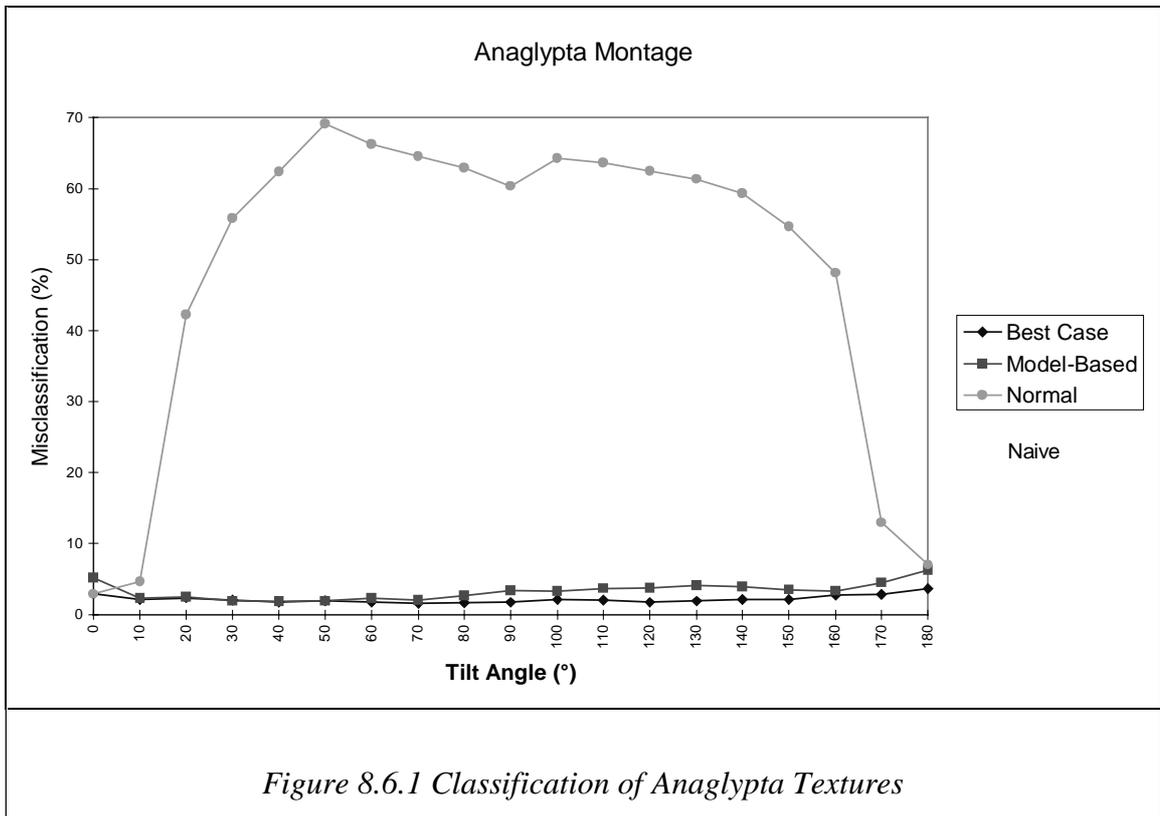
This chapter has been concerned with the evaluation of the model-based algorithm. A reductionist approach was adopted, evaluating the performance of the system after each operation. In this section we evaluate the final stage in the process: classification. The experimental results established in this section represent the most convincing evidence that the model-based approach is appropriate for the problem of tilt dependency. However, the interdependency 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.

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 $\tau=0^\circ$. The full feature set was used with a quadratic discriminant. The model-based classifier used the same classifier and recovery images obtained at $\tau=0^\circ, 90^\circ$ and 180° .



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.

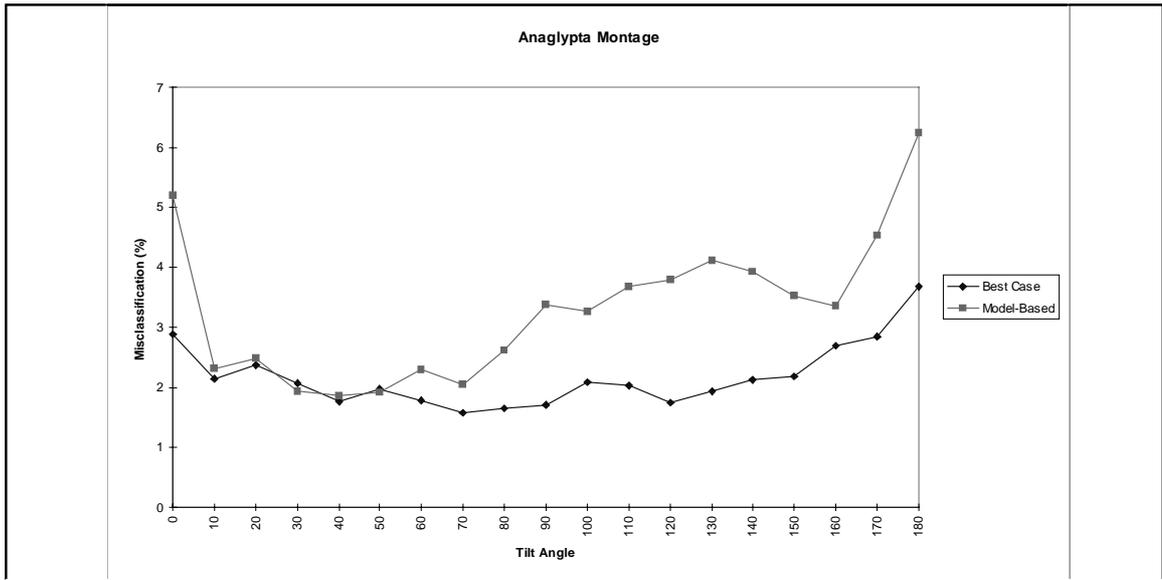


Figure 8.6.2 Comparison of Model Based and Best Case Classification For Anaglypta Montage

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.

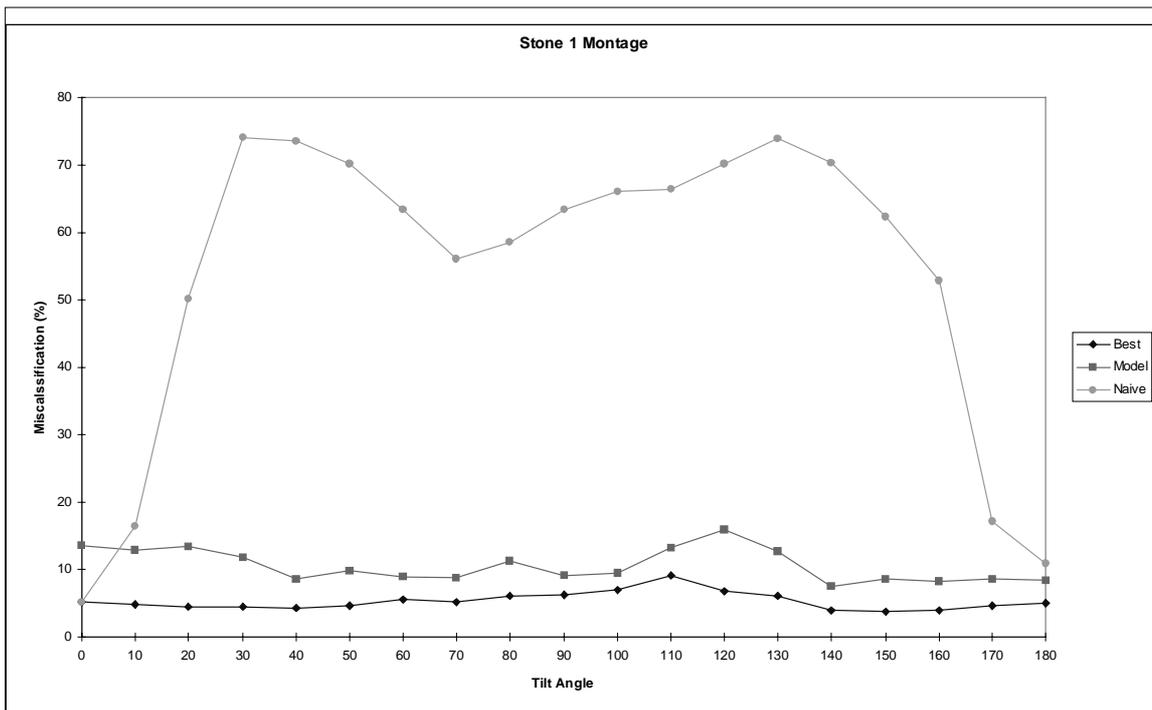


Figure 8.6.3 Classification of Stone 1 Montage

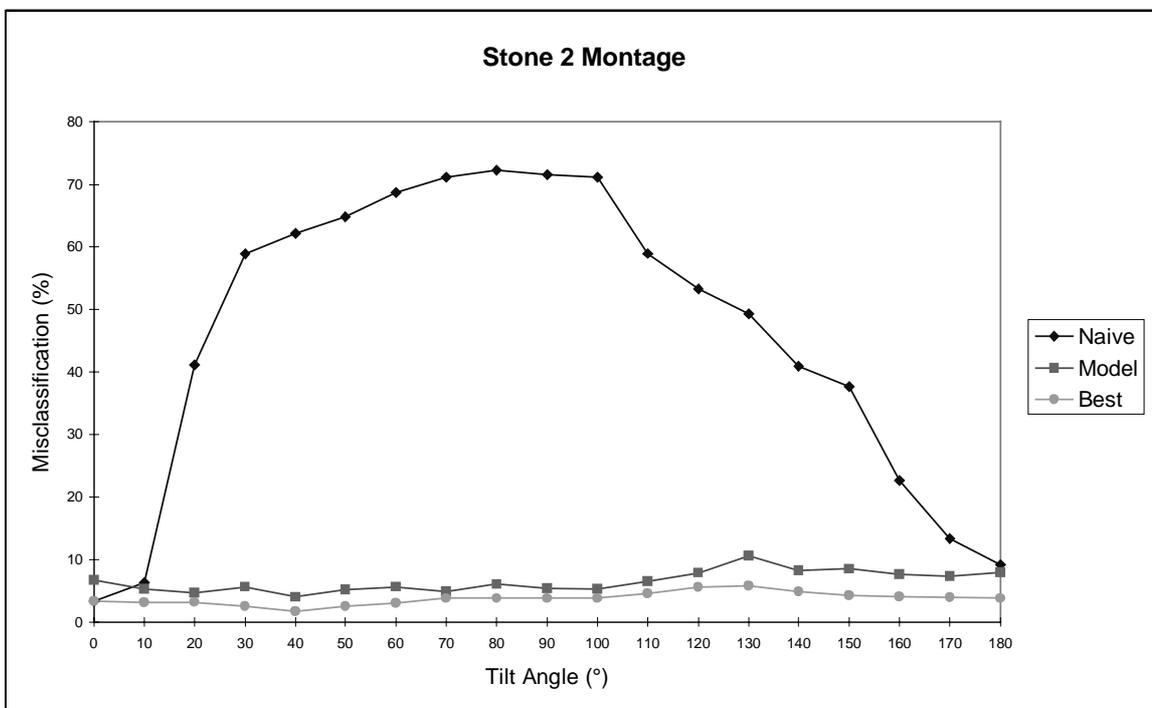


Figure 8.6.4 Classification of Stone 2 Montage

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.

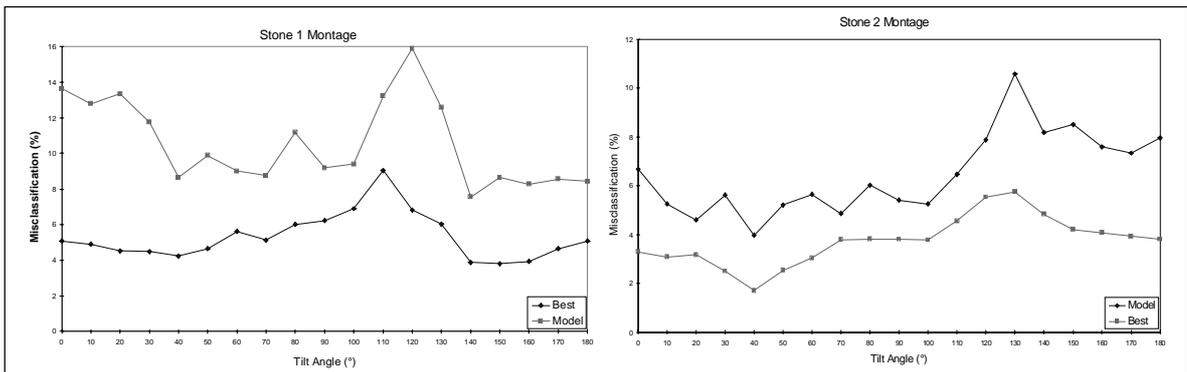


Figure 8.6.5 Comparison of model-based classifier with best case classifier.

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.

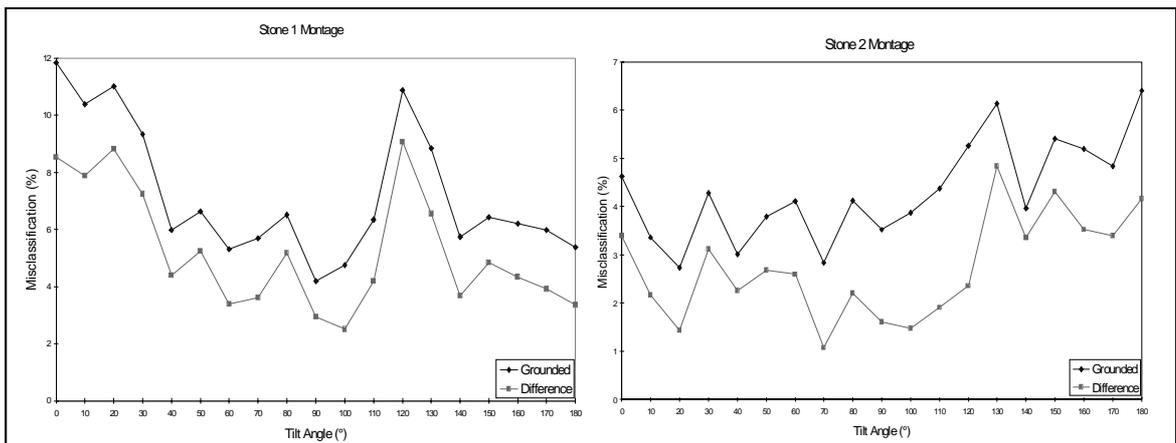


Figure 8.6.6 Comparison of the 'grounded' and 'difference' error rates.

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.