Chapter 4

Texture features : review and selection

The two preceding chapters have used theory, simulation, and laboratory experiment, to investigate the way in which changes in illuminant direction affect image texture. For the test sets employed, it has been shown that variations in either illuminant slant or tilt affect image texture. It was suggested that normalisation may compensate for changes in the former but not the latter; as variation in tilt affects the directionality of image texture. Since directional features are used in texture classification and segmentation schemes; it is to be expected that variation in tilt may affect the performance of some of these schemes.

In reviewing texture features for use in classification and segmentation schemes this chapter therefore has two main objectives :

(i) to identify research on the effect of variation in illuminant direction on texture classification and segmentation, and

(ii) to select three sets of feature measures for further investigation as regards changes in illuminant slant and tilt.

However, it is not practical to provide an exhaustive survey of all texture measures here. Given the extent of the literature such a task is beyond the scope of this thesis. Rather this chapter reviews some of the more popular techniques. Concerning point (ii) above, the criteria used for the selection of the features were :

(a) popularity in the literature,

(b) ease of implementation and use, and

(c) reported performance.

Before describing feature measures in earnest, the meanings of three terms necessary to the following discussion : *segmentation, classification* and *feature measure,* are defined. The review itself starts by considering survey papers, which divide the subject into statistical and structural approaches. Only the former will be reviewed here. Within statistical methods two further groupings can be discerned. They are (a) model-based approaches (i.e. based upon parameter estimation techniques) and (b) non-model-based methods. Descriptions of these two feature measure types are followed by a resume of "rotation invariant" feature measures : the motivation for this discussion being that the rotation of subject texture under fixed lighting, would normally imply an effective change in illuminant vector relative to the texture. Finally the conclusion summarises the literature reviewed, with particular emphasis on the consideration given to lighting variations, and it identifies three types of feature measure for further investigation.

4.1. Definition of segmentation, classification and feature measure.

Before discussing the various texture features it is helpful to clarify what the terms *segmentation*, *classification* and *feature measure*, as used in this thesis, refer to. Texture *segmentation* is used to refer to the process of dividing an image up into homogeneous regions according to some homogeneity criteria. It is therefore intimately concerned with establishing the boundaries between these regions without regard to the type or class of the regions. For brevity the term *segmentation* on its own will normally be used to describe this process.

Texture *classification* refers to the process of grouping test samples of texture into classes, where each resulting class contains similar samples according to some similarity criterion. If the classes have not been defined a priori, the task is referred to as *unsupervised* classification. Alternatively, if the classes have already been defined (normally through the use of training sets of sample textures) then the process is referred to as *supervised* classification. In the following text nearly all the texture classification tests reported are of the supervised type and this process will be referred to simply as *classification.* Unsupervised classification will therefore be specifically identified as such. Note that classification tests may be performed on separate samples of texture, in which case the samples are presented as a number of separate, normally square, images and segmentation is not required. Alternatively, a single image containing multiple textures may be presented, requiring segmentation prior to classification. Note that if classification is performed on a pixel by pixel basis within a single image then, as a by-product, segmentation also occurs.

Before either segmentation or classification can take place, some homogeneity or similarity criterion must be defined. These criteria are normally specified in terms of a set of *feature measures,* which each provide a quantitative measure of a certain texture characteristic. These *feature measures* are alternatively referred to here as *texture measures* or just simply *features.* Groups of feature measures assembled for segmentation or classification purposes are often referred to as *feature vectors.*

Note that when the performance of feature measures are compared, it is misleading to compare classification and segmentation accuracies. The former normally refers to the percentage of correctly classified texture samples or regions, while the latter may refer to the number of correctly identified *pixels.*

4.2. Surveys

Haralick provided the classic survey of texture measures [Haralick79]. He listed and described a number of texture extraction methods which he divided into two types : structural and statistical, as did Wechsler [Wechsler80]. More recently Van Gool et al produced an excellent survey of texture analysis [VanGool85]. They again divided the field into structural and statistical camps. The former use primitives to describe texture elements and placement rules to describe the spatial relationship between elements. This approach is better suited to textures that exhibit a regular macro-structure, and will not be discussed further. The statistical approaches are better suited to micro textures, and Haralick identified techniques based upon auto correlation functions, frequency domain analysis, edge operators, grey-level co-occurrence matrices, grey-level run lengths, and autoregressive models. The taxonomy of statistical techniques due to Van Gool et al, is similar to Haralick's, but it also includes the use of filter masks such as Laws' energy features, and grey-level sum and difference histograms. In addition it provides a summary of reported performance.

The surveys referred to above were performed in the late seventies and early eighties since which there has been an explosion of interest in model-based techniques (Markov fields, fractals etc.). These are well reviewed in a recent survey by Reed and du Buf [Reed93], which also covers feature-based methods (including statistical approaches) and structural methods.

The following review is therefore divided into two main groupings : model-based and non-model-based features.

4.3. Model-based features

A number of random field models (i.e. models of two-dimensional random processes) have been used for modelling and synthesis of texture. If a model is shown to be capable of representing and synthesising a range of textures, then its parameters may provide a suitable feature set for classification and/or segmentation of the textures. For a model based approach to be successful, there must exist a reasonably efficient and appropriate parameter estimation scheme, and the model itself should be parsimonious, i.e. use the minimum number of parameters. Popular random field models used for texture analysis include fractals, autoregressive models, fractional differencing models, and Markov random fields. These will now briefly be reviewed. A more extensive review of these approaches may be found in [Ahuja81] and [Reed93].

4.3.1. Fractal models

Fractals [Mandelbrot83] have, as discussed earlier, been used very successfully to synthesise natural looking textures [Voss88] [Saupe88]. Their use for synthesis together with their ability to characterise "roughness" [Pentland84] make their major parameter, fractal dimension, a natural candidate as a feature measure of texture. Many researchers have estimated the fractal dimension and used this directly as a texture measure. Such estimates are obtained either in the frequency domain, by estimating the gradient of the log-log plot of the power spectrum, or from the spatial domain by a variety of methods

[Voss88]. Note however, that as fractal dimension describes *scaling* behaviour, it is necessary to perform measurements over at least two scales, and that one would expect that the wider the variation in scales the more accurate would be the estimation procedure. Accurate estimation of fractal dimension therefore seems to be at odds with the accurate determination of texture boundaries — however, this is a trade-off that all texture segmentation schemes must make.

Pentland reported one of the earliest uses of fractal dimension estimates for segmentation purposes [Pentland84]. He used the power spectrum method to provide an omnidirectional estimate of the fractal dimension of 8x8 pixel blocks, which he used to segment a variety of indoor and outdoor scenes. From the results presented it appears that the scenes have all been coarsely segmented into textured and non-textured regions, something which could not have been achieved with straight grey-level thresholding.

Medioni and Yasumoto [Medioni84] also used a single omnidirectional fractal dimension estimate as a feature measure. They tried to segment an image containing multiple textures and obtained unsatisfactory results. They commented that *"it* (fractal dimension) *suffers the drawbacks associated with any single feature measurement space: it describes one aspect of texture and therefore can only separate textures which differ enough in roughness".* Keller and Chen [Keller89] similarly state that *"fractal dimension alone is not sufficient to characterise natural textures".*

Two techniques have been used to enhance the classification power of fractal dimension based methods. Firstly, additional parameters such as "lacunarity" have been utilised. Secondly, directionality, a key feature of most texture analysis schemes, has been employed by relaxing the assumption of isotropy and providing directional estimates of fractal dimension.

The term lacunarity was derived by Mandelbrot from the Latin term for gap (lacuna) and he used it to describe the size of holes in images of galaxies [Mandelbrot83]. Linnett [Linnett91a] likened this characteristic to "structure" in texture, which is dependent on phase information [Clarke92]. Keller and Chen [Keller89] used a measure of lacunarity together with an omnidirectional estimate of fractal dimension for segmentation purposes. They found that the use of this additional parameter, considerably improved the results achieved with a test image, consisting of a montage of Brodatz textures [Brodatz66].

Directionality is commonly exploited for texture classification and segmentation. It is not surprising therefore, that when Pentland [Pentland84] extended his study to compare the performance of his method against that of Laws, he used estimates of fractal dimension in x and y directions. Pentland reported a classification accuracy of 84.4% on a Laws test image (a Brodatz texture montage) which compared well with other techniques. Mosquera et al [Mosquera92] estimated fractal dimension in four directions (vertical, horizontal and the two diagonals) using a spatial domain method based on a 12x12 kernel. They achieved good segmentation results on synthetic and Brodatz textures.

Peleg [Peleg84] extended the "sausage" method of measuring the length of a fractal curve [Mandelbrot83] to produce a "blanket" method of measuring the area of a fractal surface. In both cases the way in which the measurement (length L or area A) scales with the "measurement yardstick ε" is an exponential function of the fractal dimension D. Thus D can be estimated from the gradient of the log-log plot of A against ε. Peleg used an iterative method for the calculation of the area A at different values of ε — he simply calculated the position of the next blanket's surfaces by adding a radius onto the previous blanket's surfaces. In this manner he was able to generate 50 blankets, from which he obtained 48 estimates of the log-log gradient (each from a set of three blankets). The 48 gradients were then used as features to successfully classify a set of Brodatz textures.

Peleg also described the possible use of directional versions of his "blanket" based feature, a technique which Linnett and his colleagues [Linnett91a] [Linnett91b] [Linnett93] used to great effect in his segmentation scheme. As well as using directional operators Linnett

(i) used the blanket thickness as a feature measure directly, rather than the estimates of fractal dimension,

(ii) used the thickness at each position in the image for spatially accurate segmentations,

(iii) used only a low number of blanket iterations (typically two or three) thereby reducing the computational load considerably,

(iv) used a moving window averaging filter to improve the robustness of the classification (by reducing feature variances), and

(v) used an iterative statistical classification method [Linnett91a].

Linnett achieved impressive results on side scan sonar images of the sea bed and a Brodatz texture montage (4.3% classification error).

Clarke [Clarke92] further extended Linnett's techniques to provide an elegant method of *rotation invariant* classification. His scheme transforms the feature space of each segmented region using principal components analysis (PCA). The analysis is performed separately on each region in turn and the resulting principal components are used as features for classification.

The quality of these results inspired the author to apply Linnett's scheme to underwater images [Chantler91] and to embark on this research.

4.3.2. Autoregressive models

Autoregressive models have been used for spectral estimation [Marple87], coding [Kashyap80], segmentation and classification [Khotanzad87], and image restoration [Chellappa82].

A *time series* autoregressive model is a random process model in which the current value of the output is expressed as the sum of its mean value, the current value of a white noise process, and a linear aggregate of previous output values [Box76]. The number of output values used is known as the *order* of the model (p). An autoregressive model therefore has $p + 2$ parameters: p coefficients, the mean, and the variance of the white noise. These parameters may be estimated using either least square error or maximum likelihood techniques [Khotanzad89]. Autoregressive models have been used to model images as random fields (two-dimensional random processes) by a number of researchers [Kashyap83] [Kartikeyan91] [Mao92]. In the two-dimensional spatial case the "previous values" of the time series process are replaced by the grey values of local neighbourhood pixels. Unlike the temporal case there is normally no *preferred direction* in a lattice and neighbourhoods are therefore normally defined to consist of variables both "before" and "after" the variable being modelled, i.e. they are *non-causal* (two-sided). Again the parameters may be estimated either by using least square error or maximum likelihood techniques.

Khotanzad and Kashyap [Khotanzad87] used "simultaneous autoregressive models" for texture classification. They selected the order of the models to use via texture synthesis. They fitted simple local neighbourhood models to the test textures and the resulting parameter sets were used to synthesise textures. If the synthesised textures were dissimilar to one another, then their parameters were deemed to have good discriminatory powers over the test textures. However, if the converse were true, the process was repeated with a higher order model.

Kartikeyan & Sarkar [Kartikeyan91] reported a classification scheme in which they first identified the most suitable model parameters for *each* training class and then estimated the value of the parameters themselves. Thus each training class has its own feature space, consisting of the parameters of its model. Classification was performed by calculating a set of feature vectors (one for each of the feature spaces) for a test texture and determining likelihood of the texture belonging to a class by using the corresponding class feature space. This method achieved a miss-classification rate of "about 2.4%" on a set of four Brodatz textures.

As far as parameter estimation is concerned, Khotanzad and Chen [Khotanzad89] found little difference between least squares and maximum likelihood methods. They used a six parameter autoregressive model and edge detection in feature (parameter) space for segmentation of natural textures.

Kashyap and Khotanzad [Kashyap86] developed a *rotation invariant* classification scheme which uses two autoregressive models. Firstly they use a "circular symmetric autoregressive model" in which all the weights for the local neighbourhood are lumped together to give an isotropic (directionally insensitive) parameter. Secondly, a conventional autoregressive model is used to provide a measure of "directionality" : this feature is essentially the maximum difference between right-angled pairs of neighbourhood parameters e.g. (1,1) and (1,-1). Kashyap and Khotanzad tested their scheme on twelve Brodatz textures at seven different rotations and obtained an average classification accuracy of 91%.

Mao & Jain [Mao92] developed a similar *rotation invariant* autoregressive model. It also uses the sum of unit circle based neighbourhood grey-levels, but in addition extends the order of the model to take in parameters based on wider radius circles. A directional measure is not used. Instead, to improve classification accuracy, a multi-resolution approach was adopted, in which the parameter estimation is repeated at different scales. By using a second order model at four different resolutions 100% classification accuracy was achieved. It should be noted however that, as in other "rotation invariant" tests, the test set was created by rotating *images* rather than rotating the natural textures themselves.

4.3.3. Fractional differencing models

Kashyap and his colleagues have extensively investigated the use of *fractional differencing models* (also termed *long correlation models*) for modelling, synthesis, classification, and segmentation, of texture [Kashyap84] [Kashyap89] [Choe91a] [Choe91b]. The one-dimensional fractional differencing model was suggested by Hosking [Hosking81] as a generalisation of Box and Jenkins ARMA(p,d,q) model, where p,d,q are the orders of the autoregressive (AR), differencing, and moving average (MA) parts of the model respectively [Box76]. The generalisation is a relaxation on the differencing parameter d (which is normally a low valued positive integer) to allow it to be real valued (i.e. fractional). Hosking defines the *fractional difference operator* ∇^d of order d using the binomial series :

$$\nabla^d = 1 - dB - \tfrac{1}{2}d(1-d)B^2 - \tfrac{1}{6}d(1-d)(2-d)B^3 \ldots\ldots$$

where B is the backward shift operator.

Thus for positive fractional values of *d* the ARMA(0,*d*,0) process is in fact equivalent to an infinite order AR(autoregressive) model. This explains its "long memory" characteristics [Choe91a].

Kashyap and Eom used this model to develop a texture segmentation scheme. The model's parameters are estimated in the frequency domain and used as features for texture boundary detection. They obtained reasonable results in tests using checker board images of Brodatz textures. Later, for classification and shape from texture purposes, Choe and Kashyap [Choe91a] [Choe91b] used both first and second order models. Their hierarchical approach performs a first level classification based on surface roughness. This uses a *rotation invariant* first order model in which the two directional fractal differencing parameters are lumped into a single omnidirectional measure. The second level uses a more complex model. It employs two additional directional parameters ω_1 & ω_2 to account for *pattern* (as opposed to *roughness*) and also embodies surface tilt, slant, and rotation parameters. A maximum likelihood estimate of these parameters is made given the subgroup of classes indicated by the level one classification. Feature measure means and variances for a test set of Brodatz textures are reported. They indicate that classification would be successful for the majority of the textures, and that feature measure estimates for rotated, tilted and slanted textures differ very little from those of the original textures. Note again that this "rotation invariant" test uses rotated images of texture; a more realistic test would be to use images of rotated texture.

4.3.4. Markov random fields

Markov random fields are a two-dimensional generalisation of Markov chains which are defined in terms of conditional properties. The conditional probabilities are the same throughout the chain (or field) and are dependent *only* on a variable's neighbourhood (the Markov assumption). The size and form of the neighbourhoods are defined by the order of the model. The first, second, third, and fourth order neighbourhoods of *x* are shown below, where the first order neighbourhood consists of the variables labelled with a "1", the second order consist of all "1"s and "2"s etc.

	4	3	4	
4	2	1	2	4
3	1	x	1	3
4	2	1	2	4
	4	3	4	

Figure 4.1 - Markov random field neighbourhoods

Hassner and Sklansky [Hassner80] adapted a Markov random field (MRF) simulation algorithm originally used for gas models. They used it to generate first order isotropic MRF textures which lacked realism because they were binary. They pointed out that the number of parameters needed rose roughly with the square of the number of grey-levels. They however, also suggested that parameter estimation of MRF models could be used for classification purposes. In the special case where the texture is Gaussian, an MRF model may be parsimoniously characterised by a linear model [Chellappa85a].

Kashyap and Chellappa investigated parameter estimation schemes of, and texture synthesis with, both autoregressive and Markovian models. They concluded that selection between the two model types should depend entirely upon the data being considered [Kashyap83]. Chellappa et al further used MRF models for texture classification [Chellappa85a] and coding [Chellappa85b]. They tested a fourth order MRF model based feature set (i.e. 12 linear equation parameters). The test set consisted of 64x64 samples of Brodatz textures and they reported classification accuracies of 99% and 93% for the two feature sets respectively.

Cohen et al [Cohen91b] have also developed classifiers based on MRF models. Their application involves the detection of fabric defects, and in their tests a sixth order MRF parameter estimation scheme detected test defects with 100% accuracy.

Cohen also developed a *rotation and scale invariant* classification scheme using Markovian models [Cohen91a]. They developed a Gaussian MRF model that incorporates scale and rotation parameters, and showed that they could obtain estimates of these parameters given the normal MRF model. The classification scheme therefore first obtains maximum likelihood estimates for the scale and rotation factors of a test texture for each training class; and second, it determines the likelihoods of the test texture belonging to each class, given these estimated scalings and rotations. In a test using

Brodatz images they obtained good estimates of orientation and scaling factors, and 100% accurate classification results.

4.4. Non-model-based features

This section briefly reviews co-occurrence matrices, grey-level sum and difference histograms, Laws' masks, frequency domain methods, and Gabor filters.

4.4.1. Grey-level co-occurrence and other related features

Haralick [Haralick73] developed a set of fourteen feature measures based on "grey-tone spatial-dependence matrices", commonly referred to as grey-level co-occurrence matrices (GLCM). These matrices are essentially two-dimensional histograms of the occurrence of pairs of grey-levels for a given displacement vector. Typical displacement vectors include (1,0), (0,1), (1,1), (1,-1), (2,0). He achieved classification success rates of between 82% and 89% for photomicrographs of sandstones, aerial photographs, and satellite imagery.

Many researchers [Weska76] [Conners80] [Zucker80] [Davis81b] [Unser86] [Castrec88] [duBuf90] [Lovell92] [Shang93] have used Haralick's co-occurrence based features. The most popular features include Contrast, Angular Second Moment, Correlation, Inverse Difference Moment, and Entropy, with small displacement vectors e.g. (1,0) and (0,1) [Conners80].

Zucker [Zucker80] used a χ^2 test of independence for co-occurrence feature selection; the assumption being that the pairs of pixels would be independent of one another if the distance vector did not coincide with the structure of the texture. Lovell et al [Lovell92] used heuristic rules in a segmentation algorithm that combined Laws, co-occurrence, fractal dimension, and other feature measures; to produce good results on a diverse set of images which included an underwater image of an ROV (remotely operated vehicle).

Shang and Brown [Shang93] employed principal components analysis (PCA) to reduce the dimensionality of their co-occurrence based feature space. This improves the efficiency of training and classification sessions using neural network. They reported good results with Brodatz and side-scan sonar test sets.

Cheaper but related alternatives to grey-level co-occurrence features are those based upon grey-level *differences* [Weska76] [Conners80] [Davis81a] & [Castrec88]. These features are computed from histograms of differences between pairs of pixels (the pairs again defined by a displacement vector). Weska et al and Conners et al reported that performances from both feature sets are comparable, which is not surprising given their similarities. Their similarities are reinforced by the fact that the successful co-occurrence feature "contrast" can be computed directly from grey-level difference data. Unser went one step further and computed both sum and difference histograms [Unser86], from which he was able to calculate exact equivalents to nine of Haralick's co-occurrence features and estimate the remaining five. He used a test set of Brodatz images to compare his features against Haralick's. The results were almost identical.

Davis et al, generalised the grey-level co-occurrence matrix (GLCM) to take account of *any* features that may be generated from a pixel's neighbourhood (e.g. edge value and direction) rather than just their grey-levels [Davis81a]. The results of their experiments with a database of eight textures showed that the grey-level based descriptors (i.e. GLCM descriptors) gave the best results. Davis [Davis81b] also used "polarograms" as a way of showing the directional distribution of GLCM features, and proposed a set of polarogram statistics which are rotationally invariant. Haralick [Haralick73] had earlier suggested that *rotation invariant* features could be obtained from co-occurrence matrices by taking the average and range of each feature type over the four angles that he used. Rather than average existing features, Sun and Wee [Sun83] developed a directionally insensitive measure : the "neighbouring grey level dependence matrix". These matrices (indexed by grey-level k, and number of neighbours s) indicate the number of times a pixel of grey-level k, has s neighbours of a grey-level differing by less than a from k. Neighbourhoods are defined as all pixels within a specified radius. Hence the matrices, and features derived from them, contain no directional information. Despite this, tests using two features on Landsat images achieved percentage classification rates in the lower eighties.

4.4.2. Laws' texture energy filters

Laws [Laws79] [Laws80] investigated three texture feature generation methods in detail : co-occurrence, correlation, and spatial-statistical techniques. From the myriad of spatial statistical texture measures — essentially a set of statistical moments of a very wide and often ad hoc set of masks — and a desire to produce a computationally efficient method; Laws developed a coherent set of "texture energy" masks. All the masks were derived from three simple one-dimensional non-recursive filters. These may be convolved with each other to give a variety of one and two-dimensional filters. Laws found the most useful to be a set of seven bandpass and highpass directional filters, implemented as 5x5 masks. The outputs from these masks are passed to "texture energy" filters. These consist of a moving window calculation of variance (hence justifying the term "energy filter") or, more cheaply, a moving window average of absolute values. Laws used 15x15 windows, as a compromise between classification accuracy and computational cost. Texture energy images are used either directly, or via principal components analysis, as feature images for segmentation and/or classification.

Laws used Brodatz textures and other images to compare his masks with co-occurrence and correlation based features. He achieved pixel classification success rates of 94%, 72%, and 65% respectively. Castrec [Castrec88] however, found grey-level sum and difference based features to be superior for segmentation of side scan sonar images. Pietikainen et al [Pietikainen82] [Pietikainen83] tested Laws, co-occurrence contrast, and "edge per unit area" operators on Brodatz and geological terrain types. They found that the Laws operators performed consistently better than either edge or co-occurrence based features. Miller & Astley [Miller91] used Laws masks and morphological operators to detect glandular tissue in breast X-rays. They found that Laws masks R5R5, L5L5, and S5R5[1]; combined with a 31x31 local variance filter gave good results.

Greenhill and Davies [Greenhill93] employed 3x3 Laws' masks in conjunction with a neural network classifier. The output of the neural net is passed through a mode filter to

[1]Laws' masks are defined in chapter 5.

remove small areas which have been incorrectly classified. They reported the results of a set of experiments on a Brodatz montage that used a variety of sizes of averaging and mode filters. They concluded that the optimum sizes for these two filters are 11x11 and 13x13 respectively, and that mode filters represent a valuable but underused technique.

Harwood et al [Harwood85] reported 92% and 94% success rates for classification of 120x120 samples of six Brodatz textures using L5E5 and L5S5 masks respectively. DuBuf et al [duBuf90] used the variances of the masks' outputs within a relatively small 7x7 window and concluded that Laws' features were among the best tested out of a wide variety of texture measures.

In summary therefore a number of researchers in addition to Laws himself have found these easily implementable "texture energy measures" to compare very well with alternative approaches.

4.4.3. Frequency domain methods

Two-dimensional power or magnitude spectra provide information on texture coarseness and directionality from their radial and angular distributions respectively [Weska76]. The most commonly extracted features consist of sums of coefficients within *wedges, rings,* or *sectors* of two-dimensional power spectra [Lendaris70] [Kruger74] [Weska76] [He88]. Weska et al [Weska76] and Conners and Harlow [Conners80] found these features to give inferior results to those obtained using grey-level co-occurrence or difference based features. Other frequency domain measures include those derived from the characteristics of "spectral peaks". D'Astous and Jernigan [D'Astous84] used features which included the frequency (f), direction (θ), area, and relative power of spectral peaks. They tested their features against co-occurrence matrices and concluded that the former provided a higher level of discrimination between a test set of Brodatz textures. He et al [He88] later used the same spectral peak features together with co-occurrence and power spectra sector based measures on a Brodatz test set. However, when they used stepwise forward feature selection to select ten texture measures, they found six of the first seven were co-occurrence based and the remainder were derived from power spectra sectors, i.e. none of the first ten were derived from spectral peaks. The disappointing results of such frequency based techniques are not surprising given the difficulty of obtaining reliable spectral estimates from small samples of random signals [Marple87].

4.4.4. Gabor filters

Related to the Fourier based techniques described previously, are those that use banks of filters to highlight sections of two-dimensional spectra. Unlike the pure Fourier techniques however, the output is a set of images (one for each filter) that retain spatial information, and can therefore be used for segmentation purposes. Gabor filters are popular because the human vision system is also thought to employ similar banks of directional bandpass filters [Jain91]. Bovik et al provided an accessible description of Gabor filters in [Bovik87] where they also describe the separate use of magnitude and phase outputs for segmentation. Jain and Farrokhnia [Jain91] used banks of Gabor filters, followed by energy filters similar to those used by Laws [Laws80]. Segmentation tests on a variety of texture combinations, including Brodatz and MRF textures, gave good results, with *pixel* miss-classification rates ranging between 0.5% an 13%.

4.5. Comparative studies

This section collects together various comparative studies that have been mentioned in the above review. Its purpose is to allow a comparison to be made from "independent" tests of feature measures. Unfortunately these comparative studies do not cover all the features described above and in particular model-based features are poorly represented.

Weska et al [Weska76] investigated three types of texture measure for the classification of aerial and Landsat images. They used features based upon angular and radial power spectra, grey-level co-occurrence, grey-level difference, and grey-level run length. From the results of their experiments they concluded that

(i) Features based on grey-level difference or co-occurrence measures gave similar performances. Both gave better results than features derived either from power spectra or run length statistics.

(ii) The computationally cheapest feature, the mean of the grey-level differences, did about as well as other grey-level difference and co-occurrence measures.

Conners and Harlow [Conners80] reported a theoretical comparison of four types of texture measure that Weska et al investigated empirically. They measured the "amount of texture-context information contained in the intermediate matrices of each algorithm", using a set of synthetically generated Markovian textures. Their conclusions were similar to those of Weska et al.

Du Buf, Kardan & Spann [duBuf90] tested the ability of seven types of feature measure (computed in a 7x7 mask) to segment a set of synthetic test textures. They concluded that co-occurrence and Laws gave among the best results. The grey-level co-occurrence features were calculated using images requantised to 64 grey-levels. Distance vectors of (0,1) and (1,0) and features *contrast* and *difference variance* were found to give good results for single feature segmentation, while a combination of (1,0), (0,1), (1,1) and (1,-1) directions for the *contrast* feature gave the best multi-feature segmentation results. Of the Laws operators R5R5, E5L5, E5S5, and L5S5, managed to segment most of the test images. It is interesting to note that they describe these 5x5 masks as low cost Gabor functions. They also used fractal dimension but obtained "disappointing" results which they attributed to the coarse estimation methods they employed.

Castrec and Kernin [Castrec88] investigated the use of grey-level co-occurrence matrices, grey-level sum and difference histograms, and Laws' texture energy filter features. They applied these measures to the task of side scan sonar image segmentation. They found that the Laws features did not perform as well as the other two techniques, and that the co-occurrence contrast, correlation and variance features could be calculated 60 times more efficiently from grey-level sum and difference histograms. Their conclusion therefore was that features based on the latter technique were the most promising.

He et al [He88] did not perform an explicit comparison of texture feature performance, rather they used a standard feature selection procedure (forward stepwise) to select ten feature measures out of a mix of co-occurrence, PSD, and spectral peak based texture measures. Six out of the first seven selected were derived from co-occurrence matrices, while the rest were PSD sectors. None were based upon spectral peaks. This confirms the results reported by Weska et al, and Conners and Harlow, that Fourier spectrum based features do not seem to perform as well as their co-occurrence based competitors.

4.5.1. A league table of feature measures

The table below summarises the findings of comparative studies described above. The numbers indicate the relative order, in terms of classification and/or segmentation performance, that each study placed the particular feature concerned. Note that there are several *joint* placings and that 1 = first place (i.e. best). Blanks indicate that the feature was not investigated by the researchers.

	Co-occurrence	Sum and difference	Laws' masks	Run length	Fractal dimension	PSD wedges and rings	PSD peaks
Weska	1	1		2		2	
Conners	1	1		2		2	
duBuf	1		1		2		
Castrec	1	1	2				
He	1					2	3

Table 4.1 - Comparative studies of texture features

As the educational institutions are well aware, league tables should be treated with caution. However, from the above two points are immediately apparent. Comparative studies have tended to concentrate on non-model-based approaches, and of the features tested, co-occurrence matrices and their cheaper alternatives, the sum and difference based features, are reportedly better than other approaches.

4.6. Rotation invariance

A number of the papers discussed above have reported the development of *rotation invariant* texture features. It might be expected that such research would encompass the effects of variation in illuminant vector. However, few of the papers reviewed in this chapter discuss this topic and none investigate it in detail. They therefore implicitly assume that

(i) the textures concerned consist of surface markings only (albedo texture), or that

(ii) the illuminate vector is perpendicular to the surface of the texture, or that

(iii) the illumination is omnidirectional (flat).

The use of Brodatz textures as test cases reinforces the above assertion, as the only practical way of obtaining *rotated* examples of these textures is to rotate Brodatz's album before scanning, or alternatively to rotate the images once they have been scanned in. In either case this is *clearly not the same as rotating the physical textures themselves,* as the illumination is effectively rotated with the texture. Note that many of the Brodatz textures were photographed using directional-lighting to highlight surface relief.

For ease of reference the "rotation invariant" feature measures discussed above will now be summarised. They naturally fall into two camps. Firstly there are those that ignore directional information completely, i.e. they only consider omnidirectional measures or averages of directional features. Secondly there are those that exploit *relative* directional information, i.e. directional information that is independent of absolute angle, such as the angle *between* the two major directions in the texture etc.

4.6.1. Omnidirectional feature measures

Haralick [Haralick73] suggested computing omnidirectional features from directional measures by averaging his co-occurrence based features over the four directions. Sun and Wee [Sun83] took this to its logical conclusion and computed omnidirectional features directly from an omnidirectional matrix : the neighbouring grey-level dependence matrix whose entries depend on the values of neighbours in all directions. Kashyap and Khotanzad [Kashyap86] developed a *circular* autoregressive model which simply averaged all the pixels on the unit circle neighbourhood into a single value associated with a single parameter — producing a model containing no directional information. Mao and Jain [Mao92] took this one stage further by increasing the order of this rotation invariant autoregressive model to take in neighbourhood pixels on larger radii.

4.6.2. Rotation invariant directional feature measures

This section briefly describes texture features that measure directional characteristics of texture, yet are "rotation invariant". Such features include *relative* angular measures, such as the angle *between* the two major directions in the texture etc. Eichmann [Eichmann88] performed a Hough transform and then extracted rotation invariant features from parameter space. Although they were rotationally invariant they did exploit directional information : the angles between sets of lines, the spacing between parallel lines, and the number of principal line directions. Davis [Davis81b] constructed "polarograms" from GLCM features, from which he computed rotation invariant moments which he used as feature measures. Kashyap and Khotanzad [Kashyap86] augmented the omnidirectional "circular" autoregressive model with a directional one, from which they extracted a measure of directionality. Cohen et al [Cohen91a] integrated scaling and rotation parameters into an MRF model. Choe and Kashyap [Choe91a] [Choe91b] used a two level classification system : firstly an omnidirectional fractional differencing model was used for coarse classification according to texture *roughness*, and secondly, a second order directional model with rotation, tilt and slant parameters was used to refine the classification. Clarke [Clarke92] used principal components analysis on each texture region separately and then performed classification on the resulting principal components.

4.7. Conclusion

The preceding sections have reviewed a number of model-based and non-model-based feature measures, their purpose being to facilitate the selection of three types of texture measure for investigation as regards the effects of variation in illuminant vector. Papers dealing with illuminant vector effects are therefore particularly relevant to this selection. Unfortunately the author has not been able to find investigations of such effects in any of the literature reviewed above. This is particularly surprising in the case of rotation-invariant schemes.

The selection of texture measures for investigation has therefore used the following rather pragmatic criteria :

(i) popularity in the literature,

(ii) ease of implementation and use, and

(iii) performance.

From the table summarising the results of comparative studies the co-occurrence based features stand out as being prime candidates due to their popularity and performance. They were therefore selected for investigation. Texture measures based on Laws' masks were also selected. These features are particularly simple and efficient, are popular in the literature, and have the added bonus that they may be easily implemented in hardware. The third feature set was selected on a mixture of pragmatics and performance : Linnett's fractal inspired operator had been used by the author for segmentation of underwater images [Chantler91], and perhaps more importantly had achieved good results at the hands of Linnett and his colleagues e.g. [Linnett91a] [Clarke92] [Linnett93].

To summarise — the features selected for investigation as to the effects of illuminant vector variation are :

(i) Laws' texture energy masks [Laws80],

(ii) co-occurrence features [Haralick73], and

(iii) Linnett's fractal based operator [Linnett91a].

12/08/94, v7, CH4.DOC