Proceedings of the Institute of Acoustics

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

B R Calder, L M Linnett and S J Clarke (1)

(1) Department of Computing and Electrical Engineering, Heriot-Watt University, Edinburgh EH14 4AS

1. INTRODUCTION

Sidescan SONAR data is normally displayed as *images* of time against distance, but the data arrives at such a rate, and in such great volume that some form of automatic analysis is required. This paper describes an approach to sidescan SONAR data based on texture analysis of *images*, where the segmentation is constructed using stochastic descriptions of the data derived directly from the spatial domain.

The paper outlines the spatial analysis technique, and describes a number of models that may be used. The stages in model construction are considered with attention to estimates of reliability and accuracy in parameter estimation, and model selection. Finally, a segmentation technique based on generating a realisation of a random process is described, and a number of examples are considered.

2. SPATIAL MODELS

2.1. Conditional Probability models of texture

Texture analysis in general has a voluminous literature. Many statistical techniques have been used, along with fractals, power spectra and other ideas. Recently, analysis of texture in terms of a probability structure measured on the pixel values has become more common [14][13][10], and provides a formal model of texture with a very rich mathematical structure.

The authors have previously defined a segmentation model based on splitting an image into a number of binary planes and analysing each one separately [18]. The model here retains this idea but extends the models of the binary processes on each plane by describing the local conditional probability distribution structure explicitly, that is, the probability of a pixel being a particular class (or greylevel, etc.) given a finite set of neighbours.

Conditional probability as a model for spatial data was proposed by Besag [1] for analysis of lattice systems. The conditional distribution was more appealing that the equivalent joint distribution, he argued, partly because of its intuitively simple application to typical problems ("what is the probability that this pixel is any particular level, given its neighbours?").

Conditional distributions are also readily specified on a local neighbourhood set of pixels, so that the model has a ready intuitive interpretation as a local model of textural properties. This is easier to manipulate than the equivalent joint distribution which may be constructed (under some regularity conditions) from the conditional distribution. In addition, the flexibility in construction of the conditional distribution allows for a large degree of flexibility in the textural types that can be modelled by the process.

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

2.2. Proposed models of texture

To formalise the ideas of a neighbourhood system and a random field, consider a square finite toroidal lattice of pixels, each representing a particular realisation of a random variable. Assume that the alphabet of possible realisations of each variable is finite (normally representing the image intensity, or possibly the classification of any pixel when used in segmentation). Let the pixels be indexed by some (arbitrary) system, and use a single variable index for convenience. Let X_t be a random variable, with realisation x_t and let the neighbourhood of a pixel be denoted $X_{\delta t}$, with realisation $x_{\delta t}$. Then, the formal definition of the neighbourhood idea is Eqn 1, often referred to as the Markovian property.

$$P(X_t = x_t | X_s = x_s \; \forall s \neq t) = P(X_t = x_t | X_{\delta t} = x_{\delta t}) \tag{1}$$

The immediate advantage of using an explicit conditional distribution is the flexibility that this allows in specifying the distribution of the entire image. Besag [2], Geman & Geman [14], Derin & Elliot [13], Cressie & Lele [11] and Hu & Fahmy [17] among others all suggest variations on the theme. However, as shown in section 3, even a counted histogram approximation to the conditional distribution estimated from a sample of the texture may be used as a non-parametric estimate of the distribution. This flexibility is a significant advantage of this technique.

The formal MRF models used in this paper are the auto-binomial model proposed by Besag, and one based on the multinomial distribution, which is new here. The auto-binomial model is defined in four parts. Firstly, an interaction between symmetrically distributed pixels (about X_t) is calculated, Eqn 2. It is assumed implicitly in the notation that offsets $\pm j$ indicate pixel pairs with 180° rotation symmetry about the pixel at index t, and that pixels at increasing magnitude of offsets form a non-decreasing series in order of distance from the pixel at t. Then, an interaction sum is evaluated, Eqn 3 (where there are G greylevels in the image, and the β_j are the model parameters, of which there are $n_0 + 1$), which is mapped by a logistic transform, Eqn 4 (required for consistency) before the conditional distribution is applied, Eqn 5.

$$\phi_{jt} = \begin{cases} G - 1 & j = 0 \\ x_{t+j} + x_{t-j} & j \neq 0 \end{cases} \tag{2}$$

$$\eta_t = \frac{1}{G-1} \sum_{j=0}^{n_0} \beta_j \phi_{jt} \tag{3}$$

$$p_t = \frac{e^{\eta_t}}{1 + e^{\eta_t}} \tag{4}$$

$$P(X_t = x_t | X_{\delta t} = x_{\delta t}) = \binom{G-1}{x_t} p_t^{x_t} (1 - p_t)^{G-1-x_t} \tag{5}$$

The scaling factor in Eqn 2 and Eqn 3 is not conventional. However, the models here deal with a differing numbers of greylevels, and unless this factor is included, the parameters which would generate (or represent) a particular texture vary (albeit linearly) with the number of symbols in the image alphabet. Adding the scaling removes this inconvenience, and also improves the stability of parameter estimation, since the parameters can remain at reasonable magnitudes even if the number of symbols increase.

The multinomial model is constructed from a local neighbourhood histogram. Let δt be the set of neighbours of t, and define the local histogram over this set as $n_t(g)$, Eqn 6, where $\delta(i, j)$ is the Kronecker delta function.

Then, the multinomial distribution may be written as Eqn 7, where the compliment in unity is added to ensure that the maximum probability occurs when all of the symbols in the local histogram are equal. Although the multinomial model could be applied to textures, its real purpose is as a model of clustering in the context of the segmentation process.

$$n_t(g) = \sum_{j \in \delta t} \delta(x_j, g) \qquad 0 \leq g < G \tag{6}$$

$$P(X_t|X_{\delta t}) = 1 - \frac{\left(1 + \sum_{g=0}^{G-1} n_t(g)\right)!}{\prod_{g=0}^{G-1}(n_t(g) + \delta(x_t, g))!} \prod_{g=0}^{G-1} p_g^{n_t(g)+\delta(x_t,g)} \tag{7}$$

Application of a specific conditional distribution implies assumptions about the data that may not be valid. Although these can be tested after fitting, this restricts the range of textures that can be represented. A non-parametric approach can remove this difficulty. One way of inferring data is a simple counted estimate of the conditional probability density based on an observed histogram. With a finite alphabet, there are only a limited number of distinct realisations of a fixed sized neighbourhood. Let $x_{\delta t}(i)$ be realisation i of the neighbourhood, $0 \leq i < G^{2n_0}$ (in general; with the pairing of pixels implied by Eqn 2, this can be reduced to $(2G-1)^{n_0}$). Then, the estimated conditional probability distribution is given by Eqn 8, where the notation $\#(\cdot)$ implies that the event is observed that number of times in the image as a whole.

$$P(X_t = g|X_{\delta t} = x_{\delta t}(i)) = \frac{\#(x_t = g, x_{\delta t}(i))}{\#(x_{\delta t}(i))} \tag{8}$$

This estimation operates effectively as long as there are sufficient observations in each neighbourhood group to make the estimates reliable. This restricts the size of neighbourhood that can be examined, but experiments indicate that the improvement in representation allows smaller sizes to be considered with similar effect to parametric models. This is followed up in section 4.

3. PARAMETER ESTIMATION

There have been a number of proposed parameter estimation systems for auto-model MRF structures. For multiple greylevel images, a maximum pseudo-likelihood estimation is normally used [2]. This is an effective technique, although it is often slow to converge to parameter estimates and requires the use of non-linear equation solution techniques, making it potential unstable.

Using binary fields, simpler estimation conditions apply, and multiple techniques have been proposed. Chen & Dubes [8] simplify conditions in an auto-binomial model and construct a linearisation of the model equations which allows easy solution; Derin & Elliot [13] apply a technique involving an estimate of the conditional distribution and a least-squares analysis; Gürelli & Onural [15] extend this with a technique to fill in missing data using a pre-computed look-up table. We consider here binary estimation since, combined with the analysis of images as binary process planes, this provides a convenient method for analysis of full images.

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

3.1. Weighted Least Squares parameter estimation

The estimation of binary field parameters can also be cast as a multiple linear regression problem due to the linear nature of Eqn 3. Let ϕ_t be the column vector of interaction pairs, Eqn 2, and let β be the vector of parameters; then, $\eta_t = \phi_t'\beta$. Since the field is binary, there are a finite number of realisations of ϕ, and for $n_0 + 1$ parameters, there are 3^{n_0} which may be distinguished. Hence, solution of Eqn 4 for η_t results in a set of 3^{n_0} equations linear in the parameters, Eqn 9, which may be solved for β (where the subscript has been dropped since the variables are constant where the given conditions apply). The value of $\hat{p}(j)$ may be estimated from the image using an approximation similar to that of Eqn 8.

$$\phi'(j)\hat{\beta} = \hat{\eta} = \ln\left\{\frac{\hat{p}(j)}{1-\hat{p}(j)}\right\} \qquad 1 \le j \le 3^{n_0} \tag{9}$$

This set of equations is overdetermined, and normally also inconsistent due to sampling errors, estimation bias and mismatch of model and image. To deal with this inconsistency, we propose a Multiple Linear Regression (MLR) formalism, in which we minimise the residual between the expected mean and the estimated value of η, Eqn 10, constructed from Eqn 9. In this equation, $w(j)$ is a weighting function, used to convert the variables so that they more closely follow the assumptions made by the least-squares technique.

$$r^2 = \sum_{j=1}^{3^{n_0}} r_j^2 = \sum_{j=1}^{3^{n_0}} w(j)\left\{\phi'(j)\beta - \hat{\eta}(j)\right\}^2 \tag{10}$$

In practice, two weighting functions are utilised. The first compensates for problems with limited observations in particular neighbourhoods, Eqn 11, and the second compensates for the non-linear nature of the transformation function, Eqn 12. With these weighting functions, the authors have previously shown that the Weighted Least Squares technique is normally as stable as Maximum Pseudo-Likelihood, and in some cases can have better worse case performance [5].

$$w_1(j) = \frac{\#(x_{\delta t} = j)}{\sum_k \#(x_{\delta t} = k)} \qquad 1 \le j \le 3^{n_0} \tag{11}$$

$$w_2(j) = \frac{1}{2}\left(1 + \exp\left\{-\frac{(\hat{p}(j) - 1/2)^2}{0.3^2}\right\}\right) \qquad 1 \le j \le 3^{n_0} \tag{12}$$

3.2. Goodness-of-fit and parameter accuracy measures

The significant advantage of the WLS technique is that it facilitates, as part of the estimation process, a measure of model/data fit, i.e., it performs an integral goodness-of-fit test. Goodness-of-fit has been a neglected problem in analysis of MRFs. Cross & Jain proposed a test based on a χ^2 test of significance [12], which, although it would in theory work for any number of greylevels or neighbourhood size is in practice limited to binary fields of small size due to the demands on data to enable reliable estimation.

Chen [7], in a full examination of various techniques suggests a scheme which is practical for multiple greylevel images since it is based on a number of features derived from the image, independent of the type and size of model used. However, the technique adds another set of unrelated features to the testing process, and the computational burden of the test is very high.

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

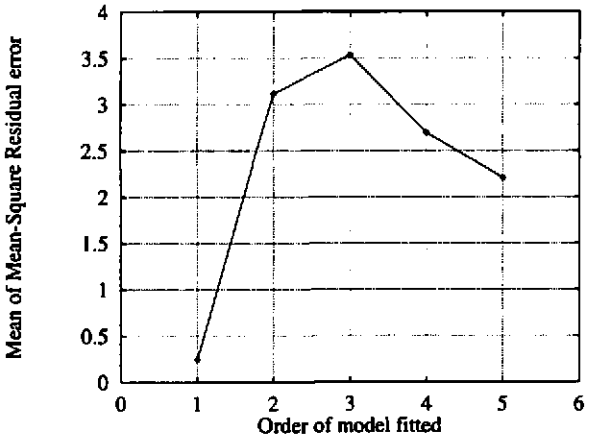

Figure 1: Mean of mean square residual after model fitting using WLS

Least-squares analysis measures the quality of the fit by the mean square residual after fitting (quotient of residual in Eqn 10 and the number of degrees of freedom, in this case $3^{n_0} - (n_0 + 1)$), which is calculated as a byproduct of the analysis; experiments indicate that this is a sensitive measure of fit. Figure 1 shows the results of an experiment where the order of model fitted to a sequence of test images is varied, and the mean of the mean square residual after each fit is calculated. The images are all first order, which is clearly shown in the reduced mean square residual at that order.

Standard analysis [19] indicates that the mean square error is only an unbiased estimate if the model actually fits the data (in other cases it is positively biased). An estimate of the true variance which is unbiased under all conditions can be calculated, however, by considering an estimate of pure error. Pure error can be calculated as the sample variance estimate of the calculated interaction sums extracted from the image. Unfortunately, the number of degrees of freedom associated with such an estimate are usually small, leading to a less powerful test.

Consequently, the image is split into ρ sections so that the sections are independent in the Markov sense. Under these conditions, estimates in each section are independent, and may be pooled in order to provide an estimate of pure error which has significantly higher degrees of freedom. From the constructed neighbourhood equations, the sample variance for each neighbourhood realisation is calculated, Eqn 13, and the pooled variance is calculated as Eqn 14.

$$s^2(j) = \frac{1}{\rho - 1} \sum_{i=1}^{\rho} w(j)(\eta_i(j) - \bar{\eta}_i(j))^2 \qquad 1 \leq j \leq 3^{n_0} \tag{13}$$

$$s^2 = \frac{\sum_{j=1}^{3^{n_0}} (\rho - 1) s^2(j)}{\sum_{j=1}^{3^{n_0}} \rho - 1} = 3^{-n_0} \sum_{j=1}^{3^{n_0}} s^2(j) \tag{14}$$

The ratio of the mean square residual to s^2 is a valid F-test under the null hypothesis that the model fitted correctly represents the data with numerator degrees of freedom associated with the mean square residual, and $(\rho - 1)3^{n_0}$ degrees of freedom in the denominator. A test against a suitable significance in the upper

tail of the F distribution can be used to measure the quality of fit of the model, not only in terms of several competing models, but also for different sizes of the same model, as in Fig 1.

Additionally, this unbiased estimate of the common variance can be used to construct confidence interval estimates for the parameter values. Let $\phi = [\phi(1) \ldots \phi(3^{n_0})]$ be the matrix of possible interaction vectors; the variance-covariance matrix associated with the parameters is given by Eqn 15. The Scheffé joint confidence limits at significance α are given Eqn 16, where $s^2(\beta_i)$ is the ith diagonal element of the variance-covariance matrix, and $F(\alpha, \nu_1, \nu_2)$ is the critical F value to give α error in the upper tail of the distribution with ν_1 degrees of freedom in the numerator, and ν_2 degrees of freedom in the denominator.

$$s^2(\beta) = \left(\phi \mathrm{diag}(w^2(j))\phi'\right)^{-1} \frac{s^2}{\rho} I \tag{15}$$

$$\Delta\beta_i = \pm\sqrt{(n_0+1)F(\alpha, n_0+1, (\rho-1)3^{n_0})}s(\beta_i) \qquad 0 \le i \le n_0 \tag{16}$$

The addition of this extra "instrumentation" to the estimation process significantly improves the technique. Not only does it check whether the model is in fact suitable for the data, it also gives an absolute measure of fit for use in selection of models, and provides information for use in evaluating the suitability of the model for further work. The estimates of parameter accuracy are another useful indicator on the model quality (in fact, the F-test significance and the parameter accuracy should be considered an information pair).

4. IMAGE SEGMENTATION

4.1. Conceptual model of segmentation

Following the work of Cohen & Cooper [9], we cast the segmentation problem as equivalent to realisation of an MRF which describes the regional clustering process corresponding to texture patches. We thus build a hierarchical segmentation system, where we consider the segmentation map to be an MRF with as many symbols in its alphabet as there are texture classes, and model the underlying textures with independent MRF models of suitable type.

Let $\boldsymbol{S}$ be the segmentation map, where each pixel value indicates the most probable texture for the corresponding pixel in the original image. Let the image be denoted $\boldsymbol{Y}$, and let there be C texture classes, with parameter vectors $\beta(c)$. Then, the joint distribution can be expressed as the product of conditional distributions as given in Eqn 17. The second and fourth terms are constant with respect to the segmentation map point in question, and hence may be eliminated if ratios of this function are taken (as occurs in the realisation algorithm used to construct the segmentation map). Consequently, define L_{tc} as given in Eqn 18, where c is the assumed classification at any time.

$$\begin{aligned} p(Y_t S_t) &= p(Y_t|S_t, S_{\delta t})p(S_t, S_{\delta t}) \\ &= p(Y_t|Y_{\delta t}, S_t, S_{\delta t})p(Y_{\delta t}|S_t, S_{\delta t})p(S_t|S_{\delta t})p(S_{\delta t}) \quad 1 \le t \le N^2 \end{aligned} \tag{17}$$

$$L_{tc} = p(Y_t|Y_{\delta t}, \beta(c))p(S_t = c|S_{\delta t}) \tag{18}$$

Generation of the segmentation map is similar to the Monte Carlo-Markov Chain spin-flip algorithm proposed by Hammersley & Handscomb [16], as modified by Chen [7]. The only modification is to replace the

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

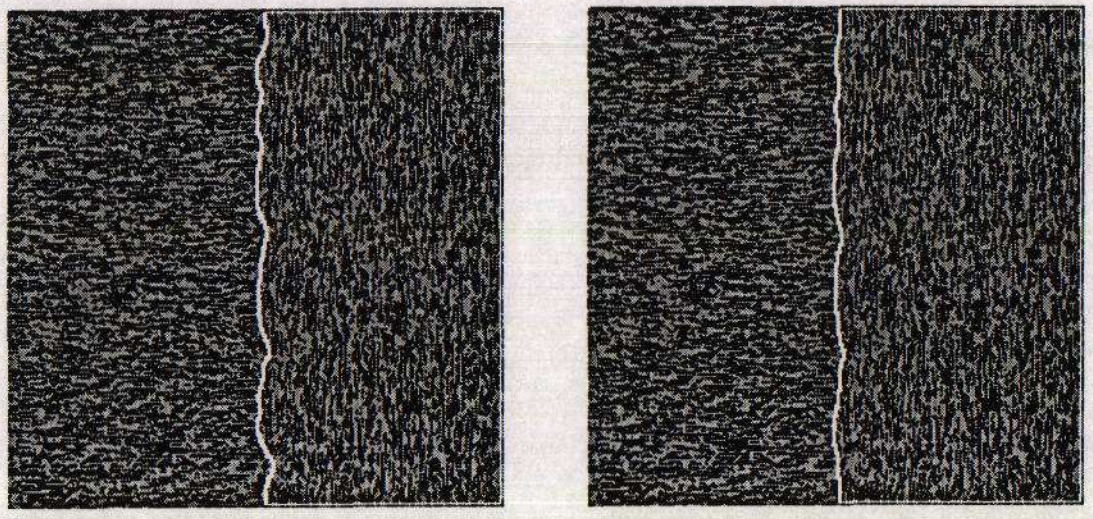

Figure 2: Comparison of parametric and non-parametric analysis of synthetic texture types

conditional distribution function of the MRF with L_{tc} so that joint function is maximised. This technique suffers from the problems associated with MCMC systems, and in particular, slow convergence. This can to some extent be ameliorated by careful consideration of the scan path used to update the image [4]. Choice of the correct update path can have up to 6% difference in the convergence time, which is often significant; fortunately, experiments on the subject indicate that the most robust path with respect to these effects (which depend on the texture properties as well as other artifacts) is the simple raster scan, for the type of image of interest, making the application simple.

4.2. Non-parametric Segmentation

The non-parametric model of texture implied by Eqn 8 may be used in the segmentation process, although it has significant difficulties in estimation where the number of neighbourhoods becomes large. Although suitable hashing techniques can be used [6] to avoid extremely large neighbourhood observation spaces, the major problem is in having sufficient data to reliably estimate the conditional distribution for each neighbourhood type. In practice, this limits the technique to small models.

However, these models can have surprisingly subtle effects. In the two results shown in Fig 2, the task is to segment the two synthetically generated textures which have identical parameters, except that they are rotated by 90°. This test is artificial, but is designed to give a testbed case where the models are known to be perfectly suited to the data. The segmentation on the left was generated by an MRF model of the two textures, and that on the right was generated by a non-parametric model of the textures; in both cases, the regional clustering process is multinomial based. Although both segmentation results are very good (99.1% and 99.6%, respectively), the important difference is at the boundary between the two textures, where the non-parametric method produces a much cleaner edge. This suggests that it is capturing more accurately the information essential to represent the textures, and hence may more accurately detect the discontinuous change in parameter vectors caused by the sharp boundary.

4.3. Parametric Segmentation

In the following results, the textures are expressed as auto-binomial models, and the regional clustering is represented by the multinomial model. It is assumed that the *a priori* probabilities in the multinomial model are equal (given no information to the contrary), which allows a significant simplification of the multinomial model, essentially reducing to the multinomial coefficient.

Figure 3 shows the result of the analysis of some standard texture patterns (in this case taken from the Brodatz album [3]), designed to allow for groundtruthed results on real textures. The image on the left is a checkerboard of two textures with different directionality properties and basic texture size (D68 (woodgrain) top left, D77 (cotton canvas) top right). The result shows clearly that the segmentation, which uses a ninth order MRF model (25 parameters, 7×7 region of support) and a multinomial with a nine pixel square support region, successfully distinguishes between the two textures. Furthermore, the accuracy where the four quarters meet in the centre is very good, indicating that the discrimination of sharply varying parameter vectors is quite precise. This effect is due to a combination of the size of the MRF model and the multinomial support region; if the model is too small, there is insufficient information for the multinomial model to work with, and disjoint regions and ragged borders occur. If, on the other hand, the multinomial model is too large, it forms regions too assiduously, resulting in over-smooth boundaries. The final misclassification error in this case is 1.63%.

The image on the right is a four-class problem using (clockwise from top left), D11 (woolen cloth), D16 (herringbone weave), D77 (cotton canvas), and D68 (woodgrain). The analysis uses a fifth order MRF (13 parameters, 5×5 region of support) and a 15×15 region of support multinomial model. In this case, the smaller MRF size is possible due to the increased multinomial size (results of experiments conducted with a ninth order MRF on the same data indicate similar results). The larger multinomial is possible due to the fact that this problem has four classes rather than just two; addition of extra classes ensures that the distribution of error pixels in any region of support that the multinomial considers favours any one class less than for a two-class problem, resulting in a smaller probability that anomalous regions will be formed within a supposedly solid class.

Again, however, the edges are sharp and the centre region where the four classes meet is dealt with accurately and reasonably. The overall pixel error rate in this case is 1.47%.

Two extracts from a sidescan SONAR survey of Bigbury Bay are shown in Fig 4, segmented using the same techniques. Both analyses use a fifth order MRF model, and a 11×11 region of support for the multinomial model. Since no ground-truth for this exists, no numerical segmentation accuracy can be calculated, but the segmentation boundaries are observed to follow the limits that a human observer might draw, especially for the left hand case. In the right hand image, some anomalous regions are observed, mainly on the left and bottom edges. These are thought to be due to the fact that the statistics of the texture are non-stationary, probably caused by the capture process at the fish (vertical banding on the right hand side is probably time varying gain steps). The majority of the image is correctly segmented, however, especially the major boundary between the two textures. In fact, this is a difficult problem, since neither of the textures has a very strong distinguishing feature, and the segmentation is probably relying significantly on average value.

5. CONCLUSIONS

The idea of using spatial models as a representation of texture has been introduced, and a number of example models have been proposed. The parameter estimation problem for a class of spatial model has

SPATIAL INTERACTION MODELS FOR SONAR IMAGE DATA

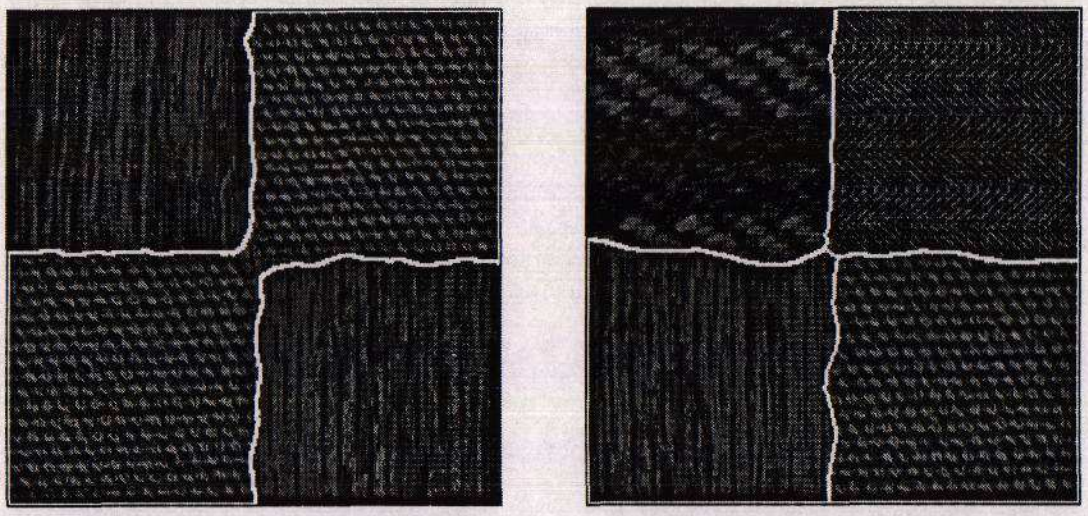

Figure 3: Segmentation of Brodatz textures. Two-class (D68/D77), left, and four-class (D11, D16, D77 and D68), right.

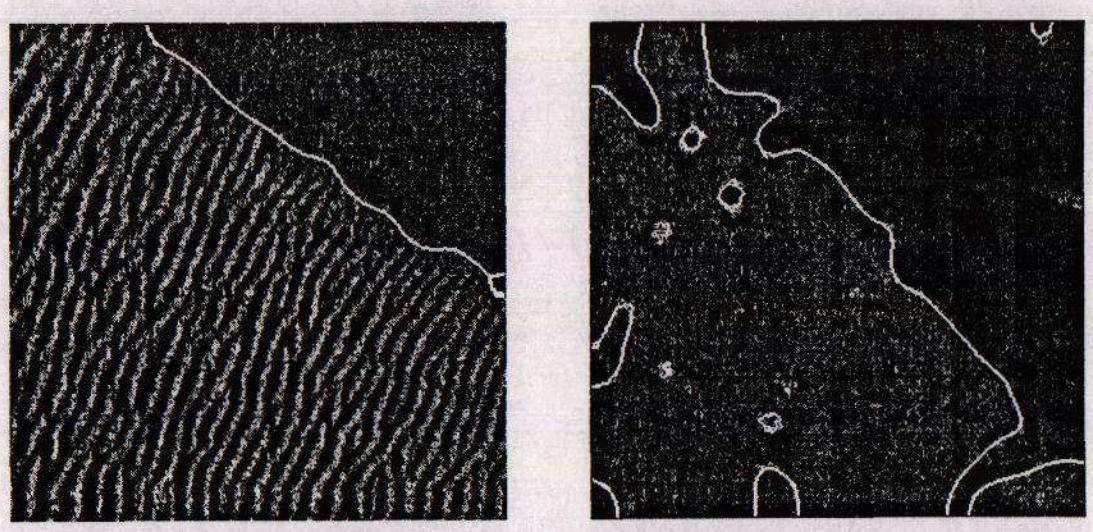

Figure 4: Segmentation of Bigbury Bay sidescan SONAR sweeps

been outlined, and a technique for extracting parameters based on a least-squares criterion fit has been described. This technique provides the basis for a powerful "goodness-of-fit" test to check on model quality and appropriateness after fitting, and a measure of parameter estimation accuracy. These extra "instrumentation" steps provide for a more reliable estimation algorithm, where estimation quality and stability can be assessed.

A segmentation technique based on a hierarchical model of the segmentation processes has been outlined. Examples of this technique using the models presented have been given for parametric and non-parametric analyses, for a number of different texture types. On examples where the ground-truth texture is known, average performance is 98.9%. On typical SONAR data, observed edges coincide with those presented by the segmentation algorithm. The techniques and models presented appear to have scope for the reliable automatic (supervised) segmentation of natural textures, including sidescan SONAR sweeps.

6. REFERENCES

[1] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Royal Stat. Soc. Series B*, 2, 1974.

[2] J. Besag. On the statistical analysis of dirty pictures. *J. Royal Stat. Soc. Series B*, 48(3), 1986.

[3] P. Brodatz. *Textures.* Dover, 1966.

[4] B. R. Calder. First year interim report, project CB/USL/5013. Technical report, Department of Computing and Electrical Engineering, Heriot-Watt University, Edinburgh, 1995.

[5] B. R. Calder, L. M. Linnett, S. J. Clarke, and D. R. Carmichael. Improvements in Markov Random Field parameter estimation. *IEE Colloquium on Multiresolution Modelling and Analysis in Image Processing and Computer Vision*, 1995.

[6] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *J. Comput. and Sys. Sci.*, 18, 1979.

[7] C.-C. Chen. *Markov Random Fields in Image Analysis.* PhD thesis, Michigan State University, 1988.

[8] C.-C. Chen and R. C. Dubes. Experiments in fitting discrete Markov Random Fields to textures. In *IEEE Conf. on Comp. Vision and Pattern Recog.*, San Diego, CA., 1989.

[9] F. S. Cohen and D. B. Cooper. Simple parallel hierarchical and relaxation algorithms for segmenting non-causal Markov Random Fields. *IEEE Trans. PAMI*, 9(2), 1987.

[10] F. S. Cohen and Z. Fan. Maximum likelihood unsupervised textured image segmentation. *CVGIP: Graph. Models and Image Proc.*, 54(3), 1992.

[11] N. Cressie and S. Lele. New models for Markov Random Fields. *J. Appl. Stats.*, 29, 1992.

[12] G. R. Cross and A. K. Jain. Markov Random Field texture models. *IEEE Trans. PAMI*, 5(1), 1983.

[13] H. Derin and H. Elliot. Modelling and segmentation of noisy and textured images using Gibbs Random Fields. *IEEE Trans. PAMI*, 9(1), 1987.

[14] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Baysian restoration of images. *IEEE Trans. PAMI*, 6(6), 1984.

[15] M. İ. Gürelli and L. Onural. On a parameter estimation method for Gibbs-Markov Random Fields. *IEEE Trans. PAMI*, 16(4), 1994.

[16] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods.* Methuen, 1964.

[17] R. Hu and M. M. Fahmy. Texture segmentation based on a hierarchical Markov Random Field model. *Signal Processing*, 26, 1992.

[18] L. M. Linnett, S. J. Clarke, and D. R. Carmichael. Texture classification using a spatial point process model. *IEE Proc. Vis. Image Signal Proc.*, 142(1), 1995.

[19] J. O. Rawlings. *Applied Regression Analysis.* Wadsworth & Brooks/Cole, California, 1988.