Proceedings of the Institute of Acoustics

PROBABILISTIC MODELING OF SYNTHETIC APERTURE SONAR AND RADAR IMAGE TEXTURES

VL Myers NATO Undersea Resarch Centre, La Spezia, Italy, *myers@nurc.nato.int*

1 INTRODUCTION

Modeling of image textures using probabilistic representations provides information on the interactions and structures that compose an image. In particular, Markov Random Fields (MRF) using Gibbs probability distributions have been successful in modeling many types of natural and simulated image textures[6]. An accurate probabilistic representation of the background permits the evaluation of areas of imagery for such measures of effectiveness as the probability of detection and classification of certain targets. In addition, it allows for automatic target recognition algorithms to tune themselves and automatically adapt to changes in the environment.

In this paper pairwise pixel interactions using gray level difference histograms (GLDH), initially proposed by Gimmel'farb[4], are used as sufficient statistics to a Gibbs probability distribution. The parameters of the Gibbs distribution are found using a technique called controllable stochastic approximation. The method is extended to handle the case of non-uniformly distributed pixels that are statistically independent, a common case for synthetic aperture sonar and radar imagery, where independent Rayleigh or K-distributed noise often occurs. Some other models make for the use the output of filter banks as statistics to the Gibbs distribution, notably the FRAME model[9,10], or exponential families of probability distributions (auto-ising or auto-normal)[2,3]; many employ measures of pixel interactions such as co-ocurrence matrices or difference histograms, pairwise pixel interactions[5] or k-tuple interactions[7].

Having represented textures using the Gibbs model, focus is put on applying the technique for probabilistic modeling of sonar and radar data. Results show that such a representation is a reasonable method to model imagery, especially where no model is available. Finally, an approach to segmentation of imagery into areas of similar texture properties is presented. The technique identifies the appropriate Markov model for the textures to be segmented using the pairwise pixel interactions and the image is then divided into classes using a simple decision rule based on the Gibbs probabilities of the pixel amplitudes.

2 MARKOV RANDOM FIELDS

Markov random field (MRF) theory is used in probability theory to represent contextual or spatial dependencies of interacting data and has been used extensively in representing image features[6]. Let $\mathbf{X} = X(i), 1 \leq i \leq R$ be an $M \times N = R$ image which forms a 2D regular lattice and where pixel gray levels $X(i) \in [0 \ldots g_{max}]$ form a finite set of signal values. The neighbours of a pixel i are the set of pixels within a distance $dist(i,j) \leq \Delta$, where $dist(\cdot)$ is the Euclidean distance.

Given a few simple conditions (positivity, Markovianity)[1,6], the image $\mathbf{X}$ is a Markov Random Field. Due to a well-known theorem attributed to Hammersley and Clifford[1], a MRF is equivalent to a Gibbs Random Field (GRF). A GRF on $\mathbf{X}$ obeys a Gibbs distribution which has the form

$$\mathrm{P}(\mathbf{x}) = Z^{-1} \times \exp U(\mathbf{x}), \tag{1}$$

where

$$Z = \sum_{\mathbf{x} \in \mathbf{X}} \exp U(\mathbf{x}), \tag{2}$$

is a normalising constant called a partition function. The energy function

$$U(\mathbf{x}) = \sum V_a(\mathbf{x}) \tag{3}$$

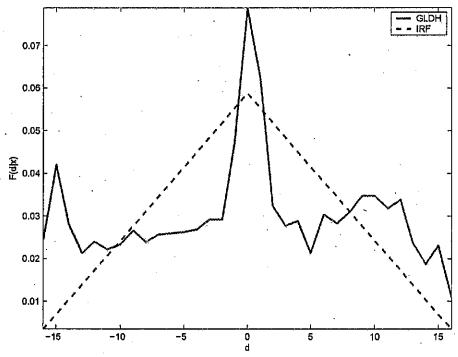


Figure 1: The relative gray level difference frequency histogram from the "sand ripples" image (Fig. 4) for $\{\Delta_v = 0, \Delta_h = 1\}$. The relative frequency distribution for pairwise pixel differences of the independent random field (IRF) is the triangle distribution, shown by the dashed line. When choosing characteristic interactions, it is the clique families which diverge the most from the IRF which are chosen.

is a sum of clique *potentials* over all cliques. A clique K is a subset of R consisting of either a single pixel i, pairwise neighbours i, j or triplets i, j, k and so on. Here, only the single-site clique K_0 and a specific set of pairwise cliques such that

$$K_a = (i, j) \in R^2 : (i - j) = (\Delta_v, \Delta_h), \tag{4}$$

where $a \in \mathcal{A}$, the set of all pairwise cliques, is an index and (Δ_v, Δ_h) are displacements in the vertical and horizontal directions of the image are considered. All pixel pairs in K_a are those with the same relative positions (Δ_v, Δ_h), differing only by their absolute position in the image.
The first-order potentials $V_0(x(i))$ are associated with the pixels themselves. The second-order potentials $V_a(x(i), x(j))$ associated with the clique family $K_a, a > 0$ depend only on the gray level differences, as shown in[4]:

$$V_a(x(i), x(j)) = (V_a(d = x(i) - x(j)) : d \in \mathcal{D} = \{-g_{max} \cdots g_{max}\}). \tag{5}$$

The Gibbs distribution for this problem has the form:

$$\mathrm{P}_{\mathbf{V}}(\mathbf{x}) = Z_{\mathbf{V}}^{-1} \exp\left[\sum_{g \in \mathcal{G}} V_0(g) H(g|\mathbf{x}) + \sum_{a \in A} \sum_{d \in \mathcal{D}} V_a(d)\, H_a(d|\mathbf{x})\right] \tag{6}$$

where $H(g|\mathbf{x})$ is the gray level histogram (GLH) of $\mathbf{x}$ and $H_a(d|\mathbf{x})$ is the gray level difference histogram (GLDH) of the clique family K_a. This is slightly different than the parameters found in[4] as $H(g|\mathbf{x})$ has been added in order to compute potentials for the pixel intensities and handle cases of independent but non-uniformly distributed noise. This case appears to have been addressed subsequently in[5].
Let V_0 and $\mathbf{V} = [V_a(d) : d \in \mathcal{D}, a \in \mathcal{A}]$ be the potentials associated with the gray level histogram (GLH) and GLDHs respectively. The task of estimating the potentials $\mathbf{V}$ is the fundamental problem of texture modeling with Gibbs distributions.

2.1 Initialising the potentials

The potentials will be refined using the controllable stochastic approximation method described below. However, an analytical solution for the initial estimates of the potential is shown in[4]. The distribution of the relative frequency distribution for pairwise signal differences of the independent random field

(IRF), the set of independent, equiprobable signals over R is described with a triangular distribution which has the form:

$$\mathrm{IRF}(d) = \frac{1 + g_{\max} - abs(d)}{(1 + g_{\max})^2}. \tag{7}$$

Fig. 1 shows the IRF for $g_{\max} = 16$. Next, a factor is computed which defines a step in the direction of the log-likelihood function of the potential values (see[4] for details). This factor is computed using

$$\lambda_{[0]} = \frac{\sum_{g \in \mathcal{G}} \left(F_0(g|\mathbf{x}_{\text{train}}) - u(g)\right)^2 + \sum_{a \in \mathcal{A}} \rho_a^2 \sum_{d \in \mathcal{D}} \left(F_a(d|\mathbf{x}_{\text{train}}) - \mathrm{IRF}(d)\right)^2}{\sum_{a \in \mathcal{A}} \rho_a^3 \sum_{d \in \mathcal{D}} \phi(d) \left(F_a(d|\mathbf{x}_{\text{train}}) - \mathrm{IRF}(d)\right)^2}, \tag{8}$$

where $F_0(g|\mathbf{x}_{\text{train}})$ and $F_a(d|\mathbf{x}_{\text{train}})$ are the relative frequencies of the GLH and GLDHs for the texture being modeled, i.e.:

$$F_0(g|\mathbf{x}_{\text{train}}) = H_0(g|\mathbf{x}_{\text{train}})/|R|, \tag{9}$$

and

$$F_a(d|\mathbf{x}_{\text{train}}) = H_a(d|\mathbf{x}_{\text{train}})/|K_a|, \tag{10}$$

where $|\cdot|$ is the cardinality of a set. In (8), $\rho_a = |K_a|/|R|$, $u(g)$ is the uniform distribution on $\mathcal{G}$ and $\phi(d) = \mathrm{IRF}(d) \cdot (1 - \mathrm{IRF}(d))$ is the variance of the independent random field.
The initial estimate of the potential values, based on the log-likelihood function of **V** is:

$$V_{a[0]}(d) = \lambda_{[0]} \cdot \rho_a \cdot (F_a(d|\mathbf{x}_{\text{train}}) - \mathrm{IRF}(d)). \tag{11}$$

These are used a starting point to the stochastic approximation algorithm described below.

2.2 Characteristic pixel interactions

The choice of $\Delta_{h\max}$ and $\Delta_{v\max}$ determines the number of clique potentials to estimate, which is combinatorial and thus computationally prohibitive for large values. However, restricting $\Delta_{h\max}$ and $\Delta_{v\max}$ to small values limits the model's ability to capture pixel interactions whose relationship may be found in distances greater than the choice of parameters. To solve this problem, sufficiently large values for $\Delta_{h\max}$ and $\Delta_{v\max}$ are selected, after which the GLDHs are computed and the number of cliques reduced using a heuristic approach whose basis is that cliques which are close to the independent random field are less characteristic of the image than those who are not. This is accomplished by computing the distance between all of the relative frequency GLDHs for all cliques and the independent random field:

$$D_a = \sqrt{\sum_{d \in \mathcal{D}} \left(\left(F_a(d|\mathbf{x}_{\text{train}}) - \mathrm{IRF}(d)\right)^2\right)}. \tag{12}$$

The mean μ and standard deviation σ of these distances are found and then a threshold is chosen such that $\tau = \mu + k \cdot \sigma$ where k is a parameter chosen by the user that determines how strict the threshold τ will be. Finally, only those cliques families K_a whose distance $D_a > \tau$ are used in the stochastic approximation step, thus reducing the number of cliques to only those who are the "most" characteristic, meaning non-independent.

2.3 Stochastic approximation

Once the clique families have been chosen and the initial potential estimates found, a technique called stochastic approximation is applied to refine the potential values, which is similar to the algorithm of Geman and Geman[3]. In this algorithm, both the texture and the potentials are generated during the course of the approximation using the Metropolis algorithm[8]. The process begins with an image of independent pixels, uniformly distributed on $[0, \ldots, g_{max}]$. Each "sweep" is a visit of all the pixels in

the image in some random order, during which a new image is generated using pixel-wise stochastic relaxation. After the sweep, the estimates are refined using:

$$\mathbf{V}_{[t]} = V_{[t-1]} + \lambda_{[t]}(F(\mathbf{x}_{\text{train}}) - F(\mathbf{x}_{[t]})) \tag{13}$$

for all clique potentials whtv, Imgere $F(\mathbf{x}_{\text{train}})$ is the relative GLH or GLDH for the texture being modeled. The parameter $\lambda_{[t]}$ at iteration t is found using:

$$\lambda_{[t]} = \frac{c_0 + 1}{c_1 + c_2 \cdot t} \cdot \lambda_{[0]}. \tag{14}$$

The images $x_{[0]} \ldots x_{[t]}$ form a Markov chain whose GLH and GLDHs increasingly match those of the image being modeled.

3 MODELING SAR AND SAS TEXTURAL PROPERTIES

For these experiments, the values of the Metropolis algorithm are $c_0 = 0$, $c_1 = 1$ and $c_2 = 0.001$. Unless otherwise noted, the cliques are chosen by $k = 3$ and $\Delta_{h\,\max} = 60$, $\Delta_{v\,\max} = 60$ and $g_{max} = 32$. To check for algorithm convergence, at each iteration two values $s_1 = |F_0(g|\mathbf{x}_{\text{train}}) - F_0(g|\mathbf{x}_{\text{model}})|$ and $s_2 = |F_a(d|\mathbf{x}_{\text{train}}) - F_a(d|\mathbf{x}_{\text{model}})|$ are computed and compared to thresholds. The algorithm ends when both $s_1 < 0.03$ and $s_2 < 0.03$.

3.1 Real and synthetic aperture sonar images

Fig. 2 shows modeling results for a Klein 5400 sidescan sonar over different seabeds comprising changing textural properties. In (a) the model is able to capture the structure of the sand ripples texture using a few important cliques. The cliques themselves provide information on the components of the structure. In (b), an image of the sea grass "posidonia" is shown. Accurate physical modeling of sea grass is impractical, with the physical movement of the vegetation creating many difficulties in estimating the scattering strength. However, this probabilistic approach can at least give visually acceptable results for simulating such a texture. Finally, a texture of flat, sandy seabed is modeled. Typically, such an area would be considered as comprising independent pixels following Rayleigh-distributed noise. However, the characteristic interactions show some correlation between a pixel and its 8-neighbourhood, resulting in non-independent pixels and a textured image.

A SAS image is shown in Fig. 3. The sensor operates at a center frequency of 150 kHz with a 60 kHz bandwidth. The characteristic interactions shown here have a much larger number of non-independent pixels. This may be caused in part by the texture of the image itself, or perhaps by the image being slightly unfoussed.

Imagery from a simultaneous dual-frequency (HF = 105-135 kHz, LF = 8-52 kHz) synthetic aperture sonar mounted on an autonomous underwater vehicle is shown in Fig. 4. Figs. (a) and (b) show high and low frequency images for the same patch of sandy seabed, with both exhibiting generally the same characteristic interactions. Figs. (c) and (d) contain images at both frequencies for an area of sand ripples. The high-frequency image is represented reasonably well, however the structure is not captured for the low frequency image. This may be due to the presence of a constant, high level of noise in this image which worsens the image contrast. By lowering the threshold for the characteristic interactions to $k = 2$, the structure is recovered in (e).

Finally, imagery is shown from an EdgeTech 4400 SAS can be see in Fig. 5. In (a) and (b) are images from two areas of flat seabed with different bottom properties (fine sand and mud, respectively) gathered under the same conditions. The characteristic interactions are the same for both, suggesting that those interactions somehow describe the effective footprint of the beamformed image. A modeled field of sand ripples for this sensor is shown in (c).

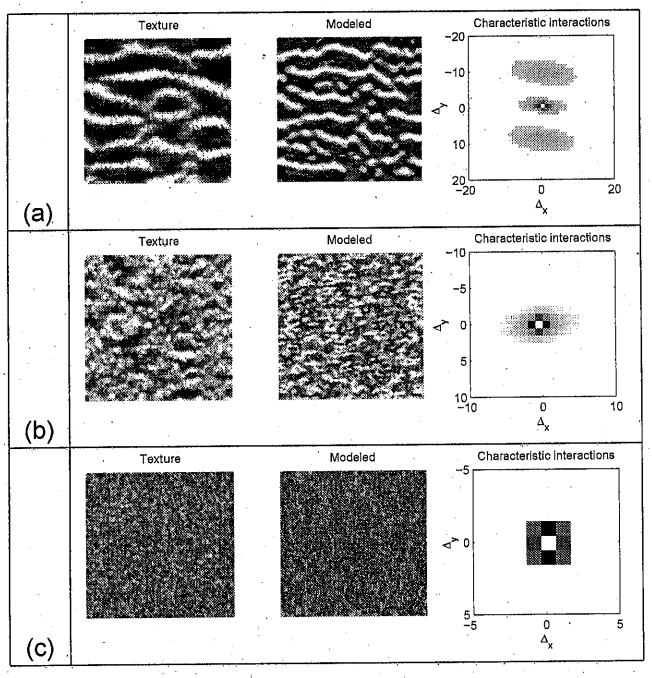


Figure 2: Real , simulated and the characteristic interactions for (a) sand ripples, (b) sea grass and (c) flat sandy seabed. The darker the pixel in the charactersitic interactions, the greater the deviation from the IRF.

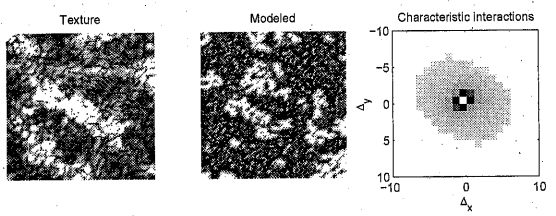


Figure 3: INSAS synthetic aperture sonar imagery.

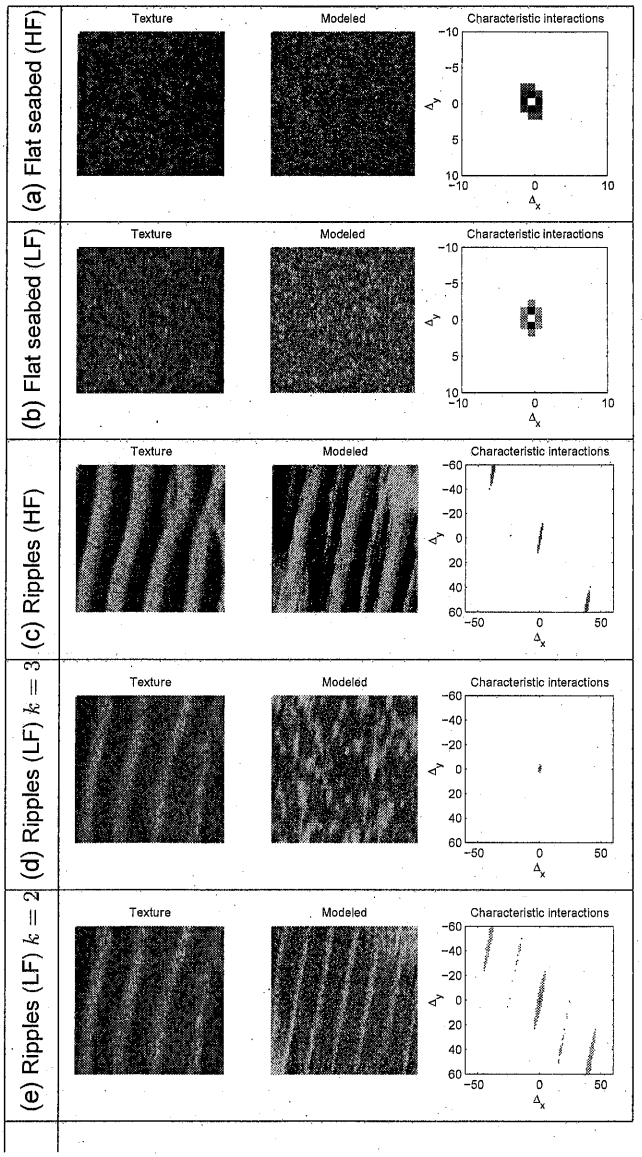


Figure 4: Dual-frequency SAS data. See text for details.

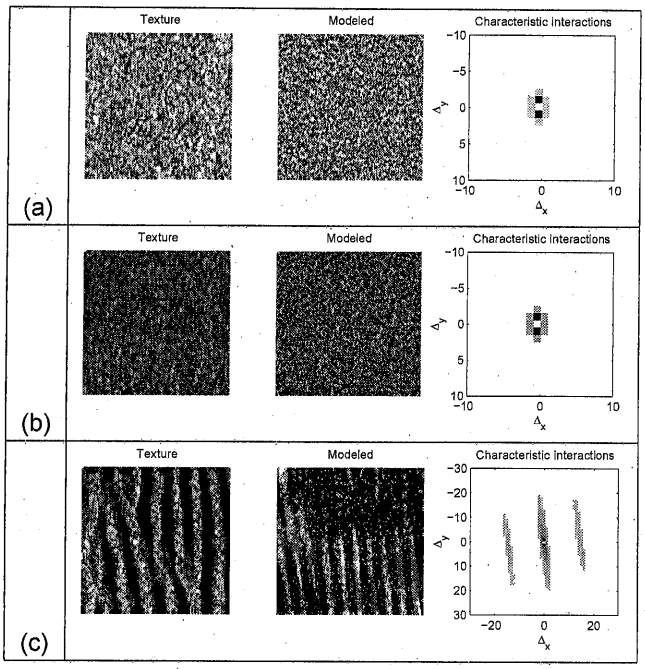


Figure 5: EdgeTech SAS data. See text for details

3.2 SAR images

The modeling technique can also be applied to synthetic aperture radar images. Fig. 6 shows textures extracted from ENVISAT imagery for the sea surface and desert sand formations. The textures are modeled, visually, realistically. The image of the sea surface shows some correlation between adjacent pixels, however close examination shows these to be very weak, and may in fact be related to the bicubic interpolation used to subsample the image. The weaker pixel interactions of the desert topology required a threshold computation of $k = 1$ to find a sufficient number of cliques to describe the texture.

3.3 Segmentation

Gibbs potentials can also be used to to segment images based on texture properties. The first step is to find the potentials for the training textures separately as in Section 2, with the total number of clique families being the union of all the characteristic pixel interactions of all training images, i.e.:

$$K_a = K_{a(1)} \cup K_{a(2)} \cup \cdots \cup K_{a(n)}, \tag{15}$$

where $K_{a(i)}$ are the cliques for the ith training image texture. Using these potentials, the Gibbs distribution for a given pixel amplitude x in an image can be evaluated using the conditional probability:

$$\mathrm{P_V}(\mathbf{x}) = Z_{\mathbf{V}}^{-1} \exp\left[\sum_{i=1}^{n}\sum_{g\in\mathcal{G}} V_0(g)H(g|\mathbf{x}) + \sum_{i=1}^{n}\sum_{a\in A}\sum_{d\in\mathcal{D}} V_a(d)\,H_a(d|\mathbf{x})\right], \tag{16}$$

where the potentials are summed over the i training textures. One of the advantages of this method is that the Gibbs potentials can be verified by a simple visual comparison of the modeled textures and the training textures prior to performing the segmentation. Once the potentials have been found,

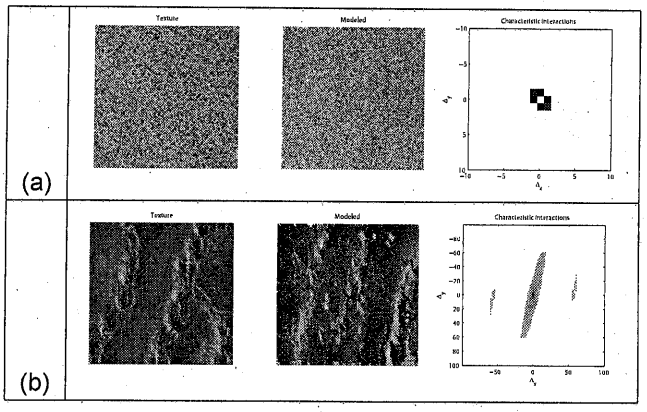


Figure 6: ENVISAT data of (a) the surface of the Mediterranean sea to the west of Crete and (b) sand formations in the Western Sahara desert. Images have been subsampled to 60 times.

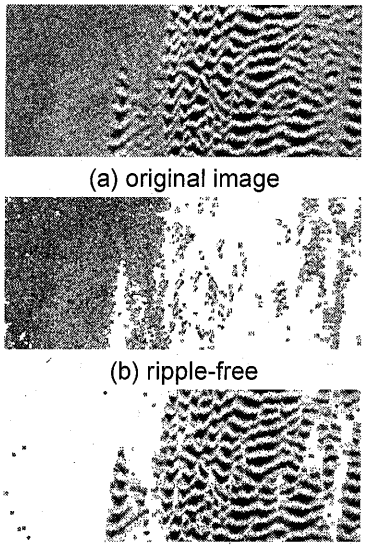
segmentation is simply a matter of computing the probabilities of all the classes and classifying the image pixel as the one which is the highest *a posteriori* probability.
As an example, an image containing sand ripples and areas of "ripple-free" seabed was segmented using this technique (see Fig. 8). Stochastic approximation was used with $\Delta_{h\max} = 60, \Delta_{v\max} = 60$ and $k = 2$ for both ripple and ripple-free seabed textures. After the segmentation, some simple morphological operators The given model is able to represent the main visual components of an image texture and the Gibbs probability is able to were applied to eliminate isolated clusters of pixels. Fig. 9 shows a more detailed close-up of the segmented boundary of the image, showing good visual agreement with the true contour.

4 CONCLUSION

A method for modeling sonar and radar image textures using pairwise pixel interactions, in the form of gray-level difference histograms, as statistics to a Gibbs distribution has been demonstrated. Beginning with an initial estimate of the Gibbs potentials and using stochastic approximation to refine this estimate, the parameters of the Gibbs distribution are refined and several image textures can be successfully reproduced using this model. A method which uses the Gibbs model for the segmentation of sonar images based on texture is also shown.
The Gibbs distribution derived from the image provides a powerful way to probabilistically describe a remotely sensed surface area which can be described as a texture. By integrating this distribution, it is possible to determine a texture-based performance evaluation given a target type which can then be used to plan future missions over the area. In addition, automatic target recognition algorithms designed with such a description for the distribution of pixels, rather than the usual independent pixels assumption, should be able to exploit this contextual awareness for improved performance.

ACKNOWLEDGEMENTS

Klein 5400 sonar data was gathered by GESMA during the BP02 trial in La Spezia, Italy in May 2002. SAS imagery from Figure 3 was gathered during the InSAS trial in collaboration with Qinetiq and FFI; thanks to Andrea Bellettini of NURC for his assistance. The dual-frequency SAS imagery was gathered during MX3 trial in Oct. 2005 and is developed by the Woods Hold Oceanographic Institute and Hydroid, Inc; thanks to Edoardo Bovio of NURC, Jose Fernandez of the Naval Surface Warfare Centre in Panama City and Kerry Commander from ONR for the use of this data. The EdgeTech SAS



(a) original image

(b) ripple-free

(c) sand-ripples

Figure 7: Segmentation using the maximum *a posteriori* Gibbs probability.

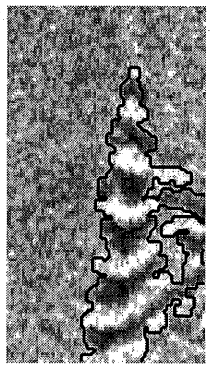


Figure 8: Segmentation detail.

data was gathered during the SWIFT sea trial in Dec. 2005 by the Royal Norweigian Navy; thanks to Roy Hansen and Per-Espen Hagen of FFI, Norway for processing this data. Thanks to Lt(N). Roger Stenvoll for processing of ENVISAT data.

References

1. J.E. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, B36:192–236, 1974.
2. Ibrahim M. Elfadel and Rosalind W. Picard. Gibbs random fields, co-occurrences, and texture modeling. *IEEE Transactions on Pattern Analysis and Machine Learning*, 16(1):24–37, January 1994.
3. Stuart Geman and Donald Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, November 1986.
4. G.L. Gimmel'farb. Texture modelling by multiple pairwise pixel interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1110–1114, November 1996.
5. G.L. Gimmel'farb. *Image Textures and Gibbs Random Fields*. Kluwer Academic Publishers, 1999.
6. S.Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer, 1995.
7. R. Paget and I.D. Longstaff. Texture synthesis via a noncausal nonparametric multiscale Markov random field. *IEEE Transactions on Image Processing*, 7(6):925–931, June 1998.
8. Chib Siddhartha and Edward Greenberg. Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49(4):327–335, November 1995.
9. S.C. Zhu, Y.N. Wu, and D. Mumford. FRAME: Filters, Random fields and Maximum Entropy - towards a unified theory for texture modeling. *Int'l Journal of Computer Vision*, 27(2):1–20, 1998.
10. Song Chun Zhu, Xio When Lui, and Ying Nian Wu. Exploring texture ensembles by efficient Markov Chain Monte Charlo - toward a Trichromacy theory of texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(6):554–569, 2000.