Bayesian Inference and Sidescan Restoration

B R Calder L M Linnett

Department of Computing and Electrical Engineering
Heriot-Watt University
Riccarton
Edinburgh EH14 4AS
E-mail: (brc, lml)@cee.hw.ac.uk

Abstract

We consider a Bayesian approach to the problem of inferring parameters of the SONAR environment given only the gathered sidescan image. A simplified model of the process is developed along with suitable prior distributions on the parameters, and a sampling technique is utilised to estimate the parameters most likely given the data. As an example, we apply this technique to estimation of a step-gain TVG curve and use the results in restoration of legacy sidescan data.

1. Introduction

We attempt to draw a distinction between *processing* and *analysis* of SONAR data. Techniques for extracting information based on power spectra [1], fractals [2] and other statistical techniques [3] consider the surface statistics of the observed image, essentially treating the image as any other array of numbers with little regard to the creation process. Recently, attempts have been made by Beattie and Elder [4] and Dugelay *et al.* [5] to extract more data, although in Beattie's case the approach is motivated by visual arguments, and in Dugelay's examples, the effects are applied to very large scale processes where significant averaging has taken place.

We are more interested here in inferring parameters of the SONAR environment when the image considered was recorded, in an attempt to extract useful detail with which to further our understanding of the data. These parameters do not have to be recorded: anything which has a significant effect on the imaging process is implicit in the image created; all we have to do is devise a method with which to estimate the effect and hence infer the parameter value. Essentially, we are attempting to invert the SONAR imaging process, but given the complexity, a number of simplifications and assumptions have to be made. However, the assumptions are shown to be reasonable under normal operating conditions for typical surveys, and in particular for the dataset which we use for the example.

The process of inferring data in this manner is quite general; we choose a straightforward problem of estimating the step gains and durations of a staircase style TVG process in legacy data given no other information as an example in this case.

2. Data Description

The dataset considered consists of output from a high frequency sidescan operating in shallow coastal waters off the south coast of the UK. The electronics pack was an early model, and consequently used a simple staircase TVG, with consequent vertical striping of the record (figure 1). As well as being annoying on visual inspection, this process renders the data significantly non-stationary and complicates further analysis.

Unfortunately, this legacy data contains no record of the SONAR parameters, and even the first return time has been removed in an attempt to de-jitter the data and improve on storage. Thus, the task we consider is to estimate the relevant parameters, including the TVG step durations and gains with the intention of eventually removing the gain curve and replacing it with something more appropriate to avoid the observed artifacts.

The gains steps are observed to be of regular period. Figure 2(a) shows estimates of column means in the image of figure 1 using a sliding windows of width five columns. The periodic spikes are more easily seen in the FFT

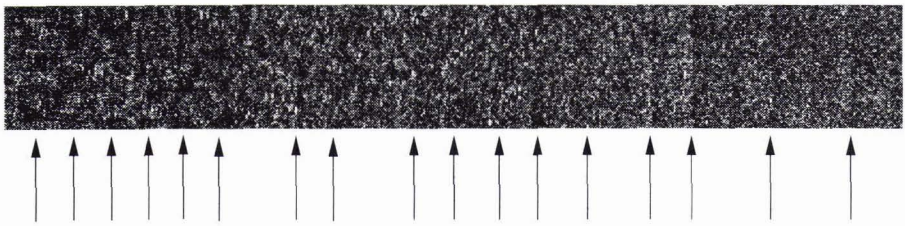
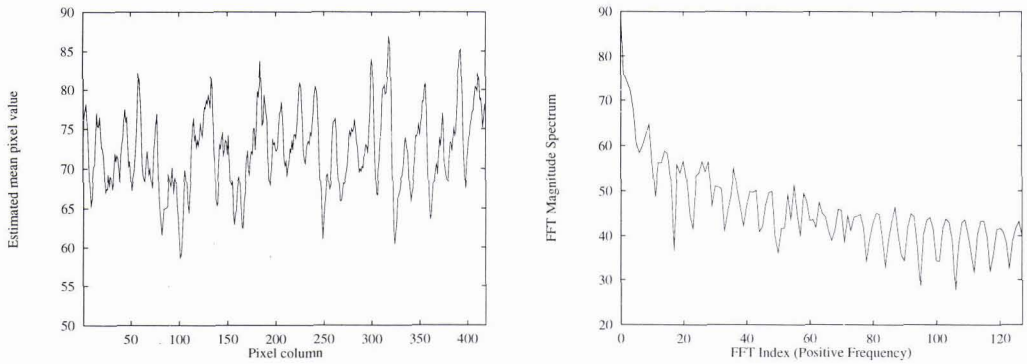


Figure 1: Example of staircase TVG on SONAR data.



(a) Column mean estimates; sample mean estimates with five pixel sliding window

(b) FFT of column mean estimates; 512 sample radix-2 DIT algorithm

Figure 2: Observed periodic gain structure in SONAR image.

of the signal. This shows significant periodicity about the 14 sample period, equivalent to about 1.4 ms with the 10 kHz sampling rate used during digitisation. Such periodicity makes the effect unlikely to be natural, confirming that the observed artifacts are caused by the TVG.

The sections of data used in the analysis are 420 samples by 100 pings in size, and are extracted from a larger dataset essentially at random, under the constraint of satisfying the assumptions made in the next section.

3. Modelling Methodology

The technique used to extract information from SONAR images revolves around a simple model of the energy return expected from the seabed and a sampler technique to allow consistent estimates to be generated. In order to build a model which can be easily manipulated, a number of simplifying assumptions have to be made.

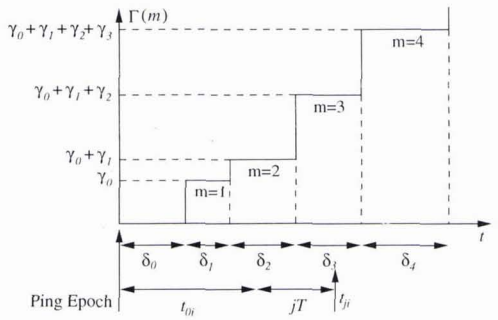
We assume firstly that the geometry of the system is simple. The towfish is maintained at the same height about the seabed, and is assumed to be sufficiently stable not to affect the images created; if fish motion is suspected, there are a number of techniques to improve the situation [6, 7]. To avoid complications with beam-pattern effects, we assume that the beam is sufficiently wide and narrow to avoid spotlighting and spreading in the assumed swathe width. The towfish for the dataset considered satisfies these conditions.

To further simplify the geometry, we assume that the towfish is of suffiently high frequency to allow a ray based solution to the propagation equations, and also that the speed of sound is constant so that the rays propagate in straight lines. We justify these latter two assumptions by noting that typical surveying with sidescan does not normally entail propagating through deep columns, and hence there is little chance of a thermocline being introduced, and that even if one is present, there is little chance of there being sufficient distance in propagation to allow the effects to become significant. Using a simple ray tracing model and temperature data from an instrumented range in a river estuary, typical rays even through strongly varying profiles show a maximum distance error of just over 1% for the longest ray examined. Consequently, the effect is insignificant compared to the other approximations made in the model.

The model includes spreading and absorbtion losses by assuming that the sound spreads spherically and that absorbtion is exponential, based on the propagation path length. The choice of spreading model is motivated purely by mathematical convenience, although the absorbtion is based on more accurate measurements [8].

Modelling of the seabed interaction is again motivated by mathematical convenience. We choose to simplify matters by using empirical justification of Lambert's Law (more correctly a "rule of thumb") [9] as a model, assuming that the seabed is a diffuse reflector, and considering only the monostatic scattering case for energy returned along the incident ray. More complex models such as those developed by Jackson *et al.* [10] could in theory be incorporated, but the manipulations would be much more difficult; in this pilot study, we opt to accept innacuracy in modelling as necessary to feasible implementation. A consequence of these modelling assumptions is that the point statistics of the observed image should be Rayleigh distributed [11], which can be shown by standard χ^2 testing on maximum likelihood fits of a Rayleigh distribution to estimated histograms from the dataset.

The TVG observed in the previous section is modelled by a base gain and a time varying component as illustrated in figure 3. We choose to parameterise in terms of step durations and additive gain for mathematical convenience in developing the sampler code. The step-like function $\lambda_\alpha(x)$ is used to allow differentiation to take place, but as $\alpha \to \infty$, this curve converges to the unit step (since $\lim_{\alpha\to\infty} d\lambda_\alpha(x)/dx|_{x\neq 0} \to 0$, and if $H(x)$ is the Hilbert step function, $\lim_{\alpha\to\infty} \int_{-\infty}^{\infty} |H(x) - \lambda_\alpha(x)|dx = 0$).



$$g_{ji} = \sum_{k=0}^{S} \gamma_{ki} \lambda_\alpha \left(t_{ji} - \sum_{m=0}^{k} \delta_{mi} \right) \tag{1}$$

$$\lambda_\alpha(x) = \frac{e^{\alpha x}}{1 + e^{\alpha x}} \tag{2}$$

Figure 3: Staircase TVG curve used in modelling.

We build the simplified model by concatenating all of these assumptions, modelling the expected return per pixel as μ_{ji} for the $(i,j)^{\text{th}}$ pixel. The return is a product of gain, absorbtion loss, spreading loss, and reflectivity at the seabed:

$$\mu_{ji} = \rho g_{ji} \cdot a(t_{ji}) \cdot s(t_{ji}) \cdot r(t_{ji}) \tag{3}$$

$$= \rho g_{ji} \cdot c^{vt_{ji}} \cdot \frac{1}{v^2 t_{ji}^2} \cdot \frac{t_{0i}^2}{t_{ji}^2} \tag{4}$$

$$= \frac{\rho c^{vt_{ji}} g_{ji} t_{0i}^2}{v^2 t_{ji}^4} \tag{5}$$

where $t_{ji} = t_{0i} + jT$ for samples of period T, and we allow that the first return time per ping may be variable. Some of the effects expected in the data cannot be distinguished from this model. For example, it is impossible to determine the reflectivity coefficient distinctly from the overall gain, or an increase in towfish height from a change in bottom depth. We assume, however, that these effects are not normally significant, and modelling potentially composite parameters is sufficient to gain further insight into the data.

4. Bayesian Posterior Analysis and Model Development

4.1. Sampler-based Approaches to Bayesian Analysis

Attempting to fit this model directly would be extremely difficult due to the interactions between the parameters, the non-linearity of the description and the possibility of multiple solutions. In an attempt to resolve these difficulties, we utilise the Bayesian approach to data analysis in order to incorporate prior knowledge of likely parameter values (and thus constrain possible parameter combinations), and use a sampler based scheme to deal with the complexities that this process generates.

In this application, the model developed forms the likelihood function, and we implement prior distributions on the parameters to complete the model. However, it would be almost certainly impossible to manipulate the posterior distribution analytically to determine the properties, and difficult even to implement a numerical attack on the integrations required for determination of the marginal distributions required. The alternative is to develop a Monte Carlo [12] scheme to estimate properties based on samples from the posterior.

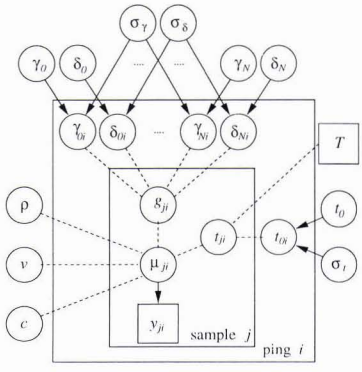
However, due to the complexity of the model, it is impossible to draw samples from the (very highly dimensional) posterior distribution. Therefore, we utilise the ideas of Monte Carlo Markov chain (MCMC) analysis [13] to develop a Markov chain which has, as its limiting distribution, the posterior distribution. This allows the samples to be generated by simple simulation of the chain, essentially swapping one large sampling task for a number of simple samplers. After an initial settling period, samples generated from the Markov chain can be considered to be correlated samples from the distribution and thus can be used to estimate any required property of the posterior.

4.2. Sampler Implementation

We use the Gibbs sampler algorithm [14] to implement the MCMC analysis required. The MCMC technique originated in statistical physics [15] where many different sampler algorithms are used, but is now used extensively in both image reconstruction (e.g., [16]) and more mainstream statistical analysis (e.g., [17]), where variants of the Metropolis-Hastings algorithm are used. The Gibbs sampler is one such variant where the Markov chain is induced by sampling from the conditional distributions of the variables, i.e., $f(x_i|x_j \forall j \neq i)$.

However, determining the conditional distributions under general conditions is quite difficult. Here, we follow a simplification after Spiegelhalter *et al.* [18] where a graphical model is used to represent the relationship between the variables, and assumptions of marginal independence of the variables are made to allow development of the conditional distributions.

The graph representing the current model is shown in figure 4, where solid directed lines indicate statistical dependence of parameters in the indicated direction, and broken adirectional lines indicate a deterministic relationship. The relationships and distributions are summarised beside the graph, where $\mathcal{N}(\mu, \sigma^2)$ is a normal distribution, and $\mathcal{R}(\beta)$ is a Rayleigh distribution.



$$y_{ji} \sim \mathcal{R}(\mu_{ji})$$

$$\mu_{ji} = \frac{\rho c^{vt_{ji}} t_{0i}^2 g_{ji}}{v^2 t_{ji}^4}$$

$$t_{ji} = t_{0i} + jT$$

$$t_{0i} \sim \mathcal{N}(t_0, \sigma_t^2)$$

$$g_{ji} = \sum_{k=0}^{S} \gamma_{ki} \lambda_\alpha \left(t_{ji} - \sum_{m=0}^{k} \delta_{ki} \right)$$

$$\delta_{ki} \sim \mathcal{N}(\delta_k, \sigma_\delta^2)$$

$$\gamma_{ki} \sim \mathcal{N}(\gamma_k, \sigma_\gamma^2)$$

Figure 4: TVG estimation model and defining equations for stochastic and deterministic relationships

Use of rectangles in the graph indicates independence, and hence the graph here indicates that each ping is initially assumed to be marginally independent of all others (although, since they share a common ancestor in the t_0 and gain parameters, they are not necessarily marginally independent in the posterior distributions), as are the samples within each ping. This is required for consistency with the previous assumptions. The rectangles also represent a general variable, indicating that all of the pings have the same structure (so that, for example, the t_{0i} are marginally independent, but all depend on t_0 in the same manner).

The structures about t_{0i} are typical of models of this type (and are the same about the gain step durations, δ_{ki} and γ_{ki}). The structure indicates that some variation is expected, but the degree of variation is unknown. Therefore the variance of the parameter is included as another parameter of the sampler, and is also estimated as part of the fitting process. In this example, the value of σ_t could be expected to be reasonable, indicating small variation in the fish height or bottom bathymetry; the values for σ_γ and σ_δ might be expected to be smaller since they depend mainly on the stability of the electronics pack on the towfish, which should be good. They are, however, included for symmetry and so that this point can be determined by posterior marginal analysis.

Variable	Prior	Parameters		Support[i]
ρ^{ii}	$\mathcal{R}$	$\beta = 1.3 \times 10^6$		$[0, 3.95 \times 10^6]$
v^{iii}	$\mathcal{N}$	$\mu = 1500\,\text{ms}^{-1}$	$\sigma^2 = 600\,\text{m}^2\text{s}^{-2}$	$[1426, 1574]\,\text{ms}^{-1}$
$1 - c^{\text{iv}}$	$\mathcal{G}^{\text{v}}$	$\alpha = 11$	$\beta = 7 \times 10^{-4}$	$[0.98, 1.00]$
t_0	$\mathcal{R}$	$\beta = 10^{-2}\,\text{s}$		$[0, 30]$ ms
γ_k	$\mathcal{R}$	$\beta = 100$		$[0, 300]$
δ_k	$\mathcal{R}$	$\beta = 1.4\,\text{ms}$		$[0, 4.25]$ ms
$\sigma_t^{-2\,\text{vi}}$	$\mathcal{G}$	$\alpha = 0.01$	$\beta = 100$	N/A[vii]
σ_γ^{-2}	$\mathcal{G}$	$\alpha = 0.01$	$\beta = 10^{-6}$	N/A
σ_δ^{-2}	$\mathcal{G}$	$\alpha = 0.01$	$\beta = 100$	N/A

i Support ranges are for limits such that $\int_a^b p(x)dx = 0.99$.
ii Maximum gain is over 130 dB, based on deterministic estimates.
iii Medwin's formula [19] at $T = 4°\text{C}$, $S = 20$ ppt and $z = 40$m; variation of each parameter gives typical limits of $1420 - 1540\,\text{ms}^{-1}$.
iv Absorbtivity at 100 kHz at 4°C is 30 dB /km [8] ($c = 0.993$).
v i.e., a Gamma distribution, $\mathcal{G}(x; \alpha, \beta) = (\beta\Gamma(\alpha))^{-1}(x/\beta)^{\alpha-1}\exp(-x/\beta)$.
vi Inverse variance ("precision") is modelled to simplify the maths; the gamma distribution is conjugate to the Gaussian linking densities, giving a closed form conditional distribution.
vii Prior is intentionally vague since no information is available.

Table 1: Summary of prior distributions and parameters for TVG model.

The prior distributions used in the model are summarised in table 1, and while the parameters of the distributions are determined by physical constraints, the shape of the distributions are determined mainly for mathematical convenience. In most cases, the priors are fairly vague since only general constraints on the parameters are required; the purpose is to keep the sampler in a physically likely area of the parameter space, but to let it explore that region fairly freely.

The full conditional distributions required for sampling can be determined as follows. Let V be the set of all nodes in the graph (i.e., parameters), let c_v be the set of all *children* of node v (i.e., all those which depend on it, collapsing deterministic links as required), and let p_v be the set of parents of a node (i.e., those on which v depends). Then, Spiegelhalter *et al.* [20] show that the conditional is:

$$f(v|\cdot) = f(v|\mathsf{p}_v) \prod_{c \in \mathsf{c}_v} f(c|\mathsf{p}_c) \tag{6}$$

where $f(v|\cdot) = f(v|\mathsf{V}\backslash\{v\})$; nodes without parents use the prior for $f(v|\mathsf{p}_v)$. Thus, for example, the full conditional of t_{0i} can be read from the graph as:

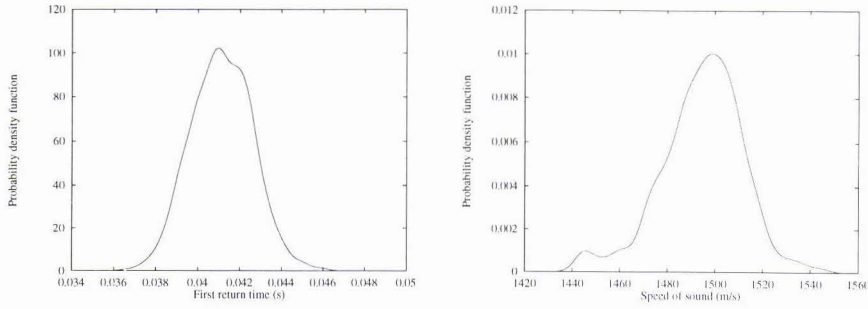
$$f(t_{0i}|\cdot) = f(t_{0i}|t_0, \sigma_t) \prod_j f(y_{ji}|\cdot) \tag{7}$$

$$\propto \exp\left(-\frac{(t_{0i} - t_0)^2}{2\sigma_t^2}\right) \prod_j \frac{y_{ji}}{\mu_{ji}^2} \exp\left(-\frac{y_{ji}^2}{2\mu_{ji}^2}\right) \tag{8}$$

This level of complexity in full conditional distributions is common. In some cases the function may be a particular distribution (with conjugate priors), but this is not usual. In the case of log-concave distributions, the ARS algorithm [21] may be used, which automatically constructs an accept-reject envelope [22] for the variable using a numerical technique. However, many of the full conditionals in this model are not log-concave (or, at least, cannot be *proved* to be unconditionally log-concave), and the more general ARMS algorithm is used [23]. Although conceptually simple, implementing ARMS is an exacting task; the authors are indebted to W. R. Gilks (contact: `wally.gilks@mrc-bsu.cam.ac.uk`) for providing the code to implement this.

5. Experimental Analysis

Analysis with the sampler is a process of setting a suitable initialisation scheme, and then running the Gibbs sampler algorithm sufficiently long to ensure that the induced Markov chain has settled to its final distribution. In this example, the data set shown in figure 1 is processed.



(a) First return time density estimates

(b) Speed of sound density estimate

Figure 5: Kernel density estimates of first return time and speed of sound. Estimates from last 1000 samples of 3000 in run.

The initialisation scheme is not critical; in theory, any parameter set should be sufficient under very weak conditions on the sampler. Experience with sampler algorithms suggests that starting the chain about the marginal posterior modes of the variables is as good an initialisation strategy as any, although it does not guarantee any better convergence rate for the sampler. The derived parameters (i.e., t_{0i}, γ_{ki} and δ_{ki}) are sampled initially according to their defining equations, ensuring that the values generated are physically possible (this is guaranteed on later samples by the construction method for the ARMS algorithm envelope).

The question of how many samples to allow for settling of the sampler (the "burn-in" time) and how many samples from the sequence to use is still uncertain except in special cases. Our (very conservative) scheme is to simulate 3000 samples from the chain and use only the very end of the run for estimation of properties. Judging from experience of other researchers (particularly [24]) and from traces of the parameter behaviour during development, convergence appears rapid and the estimates generated are stable. The only difficulty in this experiment is in the gain parameters for the last few stages of the TVG; the step parameters tend to settle sequentially from the first return side of the image, and longer runs are required to ensure sufficient time for all of the parameters to settle.

However, kernel estimates of the first return time and speed of sound, figure 5, show stable monomodal densities within the correct parameter range. Confirmation that the sampler is not simply sampling from the prior distributions is found in the fact that the first return time is significantly higher than first expected, and indeed is distinctly in the upper tail of the prior distribution. In addition, the posterior variances are much smaller than the prior variances, indicating that the vague priors used do not affect the sampler significantly. The mean values of the two parameters give $t_0 = 41.1\,\mathrm{ms}$ and $v = 1491.2\,\mathrm{ms}^{-1}$, leading to a towfish height of 30.6 m, typical of 100 kHz surveys.

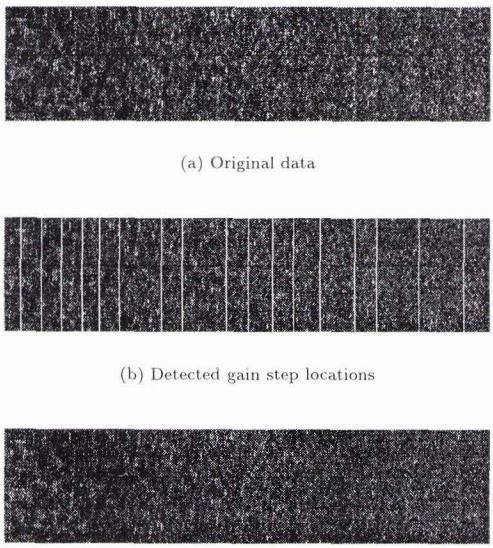
Analysis of the sampling variances indicates that there is sufficient evidence in the t_{0i} to conclude that the first return time varies on a ping by ping basis, but that there is no significant variation between the gain duration and amplitudes. In subsequent runs, the model could be simplified accordingly.

The major test of this model, however, is in restoring the data set test image. The output from the last 1000 samples of the chain were averaged to give posterior estimates of the required variables, and these were used to remove the step gains and replace them with the TVG assumed in the model. The results, figure 6, show signifcant improvement from the original data, which is confirmed by the column mean estimates shown for original and processed images.

6. Discussion and Extentions

The results in figure 5–6 show that the system proposed can detect features of the SONAR environment from the images, and (since the values generated restore the image correctly) that the model developed has to be reasonably accurate for the data considered. The aim of this experiment, and similar techniques, is not necessarily to give a perfectly accurate physical model, merely one which is sufficiently close to approximate the data but also sufficiently flexible in support by its sampler to allow the imperfections to be compensated in fitting.

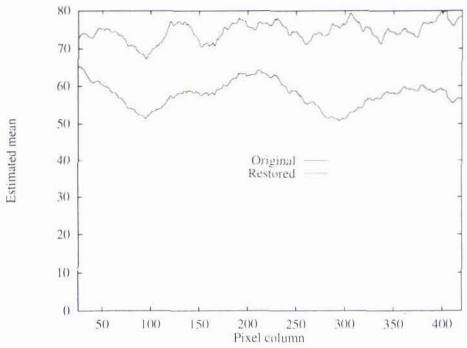
Close examination of figure 6 shows that much significant detail is obscured in the original data, particularly the gentle bathymetry in the middle of the swathe, which is very obvious in the restored image. Comparison of the before and after images shows that the detail is present in the original, but is obscured by the step-gain effects.



(a) Original data

(b) Detected gain step locations

(c) Restored data



(d) Original and restored data column mean estimates

Figure 6: Original and restored data, with column means. Column means, estimated with a window width of five samples show significantly less periodicity after restoration, corresponding to the smoother appearance of the restored image.

This technique has a number of advantages over other analysis systems, not least that it can easily attack problems which would be difficult by other methods. The MCMC process makes it a simple task to specify the relationships between variables and then estimate their properties given the data, even when the variables were never measured. Although implementation may be computationally demanding, this has to measured against the observation that there often is no alternative estimation technique.

The process is also intuitively satisfying. Many other techniques attempt to infer properties of the seabed simply from the surface statistics, treating the data as simply another image. Here, we have the ability to infer probability densities of the SONAR environment parameters, and the understanding of the dataset which is implicit in this process is very powerful. Although fairly limited through the assumptions made in developing the pilot model, this technique is a first step towards infering real knowledge from the images.

Work on these techniques is continuing. We envision that the major improvements required initially are inclusion of more general seabed surfaces, both in terms of bathymetry and multiple sediment types. Extention to self-shadowing of surfaces, discrete objects and targets, and better models of the physical processes are also being considered in order to infer properties useful for further analysis.

7. Conclusions

A method has been proposed for structuring the problem of inference of parameters in SONAR. We construct a simple model of SONAR physics, use the Bayesian method to incorporate prior information on the parameters (to improve the stability of the system), and utilise a MCMC technique to numerically estimate features of the (high dimensionality) Bayesian posterior distribution. These parameters are useful in their own right, infering information on the SONAR environment when the image was gathered, and may also be used to manipulate the image whence they were generated.

In this particular example, a simple TVG restoration problem has been posed, and an example of the dataset has been processed. The sampler outputs can be summarised to present probability density functions for variables of interest, or can be used to remove and replace the step TVG curve with a more appropriate smooth function which reduces the artifacts. The consequence of this is much better visual appearance, and better visibility of significant effects in the SONAR trace.

It is possible to include a great deal of structuring information into a difficult problem, and hence to include complex statistical models without very complex analysis techniques. The specification of the model is simple, and although computationally demanding, implementation of the sampler system is quite straightforward. As well as adopting a more rigorous "investigative data analysis" approach to the problem, the system also builds inference on parameters that are otherwise difficult to obtain and leads to much better understanding of the process of image gathering.

The model is being actively developed. Planned improvements include a better model of image formation including self-shadowing and irregular seafloors, and the inclusion of multiple sediment types. The ultimate goal is to develop a sampler which will infer relative reflectivities of each sediment, taking into account shadowing effects, and hence give sediment rather than texture maps of the surveyed area.

References

[1] N. G. Pace and H. Gao. Swathe Seabed Classification. *IEEE Oceanic Eng.*, 13(2), 1988.

[2] D. R. Carmichael, L. M. Linnett, S. J. Clarke, and B. R. Calder. Seabed Classification Through Multifractal Analysis of Sidescan Sonar Imagery. *Proc. IEE*, 143(3), 1996.

[3] L. M. Linnett, D. R. Carmichael, and S. J. Clarke. Texture Classification using a Spatial Point Process Model. *IEE Proc. Vis. Image Signal Proc.*, 142(1), 1995.

[4] R. S. Beattie and S. C. Elder. Sidescan SONAR Image Restoration using Simulated Annealing and Iterative Conditional Modes. In *Sonar Signal Processing*, volume 17 of *Proc. Inst. Acoustics*, December 1995.

[5] S. Dugelay, X. Lurton, and J. M. Augustin. A New Method for Seafloor Characterisation with Multibeam Echosounders: Image Segmentation using Angular Backscattering. In *Proc. 3rd European Conference on Underwater Acoustics*, June 1996.

[6] D. T. Cobra, A. V. Oppenheim, and J. S. Jaffe. Geometric Distortions in Sidescan Sonar Images: A Procedure for Their Estimation and Correction. *IEEE J. Ocean. Eng.*, 17(4), 1992.

[7] R. S. Beattie and S. C. Elder. MAP Sidescan SONAR Motion Distortion Correction using Iterative Conditional Modes and Simulated Annealing. In *Sonar Signal Processing*, volume 17 of *Proc. Inst. Acoustics*, December 1995.

[8] F. H. Fisher and V. P. Simmons. Sound Absorption in Sea Water. *J. Accoust. Soc. Am.*, 62(3), 1977.

[9] S. Stanic, K. B. Briggs, P. Fleischer, R. I. Ray, and W. B. Sawyer. Shallow water high frequency bottom scattering off Panama City, Florida. *J. Accoust. Soc. Am.*, 83(6), 1988.

[10] D. R. Jackson, D. P. Winebrenner, and A. Ishimaru. Application of the composite roughness model to high frequency bottom backscattering. *J. Accoust. Soc. Am.*, 79(5), 1986.

[11] J. A. Ogilvy. *Theory of Wave Scattering from Random Rough Surfaces.* Adam Hilger, Bristol, 1991.

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

[13] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov Chain Monte Carlo in Practice.* Chapman and Hall, 1996.

[14] A. E. Gelfand and A. F. M. Smith. Sampling-Based Approaches to Calculating Marginal Densities. *J. Am. Stat. Assoc.*, 85(410), 1990.

[15] K. Binder, editor. *The Monte Carlo method in Condensed Matter Pyhsics*, volume 71 of *Topics in Applied Physics.* Springer-Verlag, 1992.

[16] D. Geman and C. Yang. Nonlinear Image Recovery with Half-Quadratic Regularization. *IEEE Trans. Image Proc.*, 4(7), 1995.

[17] B. P. Carlin, A. E. Gelfand, and A. F. M. Smith. Hierarchical Bayesian Analysis of Changepoint Problems. *Appl. Stat.*, 41(2), 1992.

[18] D. J. Spiegelhalter, N. G. Best, W. R. Gilks, and H. Inskip. Hepatitis B: a case study in MCMC methods. In Gilks et al. [13].

[19] H. Medwin. Speed of sound in water: A simple equation for realisitic parameters. *J. Accoust. Soc. Am.*, 58(6), 1975.

[20] D. J. Spiegelhalter, A. P. Dawid, S. L. Lauritzen, and R. G. Cowell. Bayesian analysis in expert systems. *Stat. Sci.*, 8, 1993.

[21] W. R. Gilks and P. Wild. Adaptive Rejection Sampling for Gibbs Sampling. *Appl. Stat.*, 41, 1992.

[22] A. M. Law and W. D. Kelton. *Simulation Modeling and Analysis.* McGraw-Hill, 1991.

[23] W. R. Gilks, N. G. Best, and K. K. C. Tan. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics*, 44(4), 1995.

[24] A. E. Rafferty and S. M. Lewis. How many iterations of the Gibbs sampler? In J. M. Bernardo, J. Berger, A. P. Dawid, and A. F. M. Smith, editors, *Bayesian Statistics 4.* Oxford University Press, Oxford, 1992.