Bayesian Methods in Signal and Image Processing

W. J. FITZGERALD, S. J. GODSILL,
A. C. KOKARAM and J. A. STARK.
University of Cambridge, UK.

SUMMARY

In this paper, an overview of Bayesian methods and models in signal and image processing is given. The first part of the paper reviews some traditional classes of model employed for signal processing time series analysis. Marginal inference based upon analytic integration of hyperparameters is described for these models and illustrations are given for the problem of estimating sinusoidal frequency components in white Gaussian noise and for the general changepoint problem applied to digital communications. In the second part of the paper, state of the art applications are described which employ MCMC methods for the enhancement of noise-degraded audio signals, non-linear system identification and image sequence restoration. The complex modelling requirements and large datasets involved in these problems require sophisticated MCMC schemes employing efficient blocking schemes, model uncertainty strategies (both reversible jump and Gibbs variable selection) and non-linear/non-Gaussian models.

Keywords: SIGNAL PROCESSING, FREQUENCY ESTIMATION, CHANGEPOINT ESTIMATION, NON-GAUSSIANITY, HEAVY-TAILED NOISE, NON-LINEARITY, SCALE MIXTURES OF GAUSSIANS, IMAGE SEQUENCE, SPATIO-TEMPORAL MODEL.

1. INTRODUCTION

This paper is concerned with the processing of signals that have been digitised as a function of time, space or time and space. The basic theory behind Digital Signal Processing (DSP) has been in existence for decades and has extensive applications in fields such as speech and data communications, biomedical engineering, acoustics, sonar, radar, seismology, oil exploration, instrumentation, audio signal processing and image processing/compression.

Very often, the physical properties of the system being observed will be known and hence explicit mathematical models can be derived for the form of the signal and any observation noise sources. In these cases the behaviour of the signal is characterised by parameters whose values are usually unknown. Sometimes there can be many possibilities for the signal model. The aim is then to find the most appropriate model to describe the data (model selection) as well as to estimate the parameters. All observed signals are corrupted by noise. Noise may be the result of any number of different causes, such as damage to a storage medium, electrical or electromagnetic interference, material impurities and thermal effects. An important component of the model will thus be the characterisation of the noise statistics.

In this paper we discuss modelling and inference procedures for a range of signal processing problems. In section 2. we describe some basic models which have been extensively applied to many signals of interest. For these models analytic marginalisation of unwanted hyperparameters can be applied and successful results are obtained in a number of quite realistic settings. Examples are given for frequency estimation and changepoint detection. However, in order to achieve greater applicability and generality more sophisticated models must be investigated. Section 3. discusses some application areas in which MCMC simulation strategies are applied to non-Gaussian and non-linear models with very large datasets. Section 4. presents MCMC work applied to the restoration of film image sequences. We believe that numerical methods of the type used in sections 3. and 4. will soon find application in a very wide range of signal processing areas which have thus far been limited by considerations of computational tractability. Nevertheless, these methods are still highly computer intensive and efficient parallel and on-line implementations will be required for practical operation in many environments. More detailed background information about the application areas described in the paper can be found in (Ó Ruanaidh and Fitzgerald, 1996; Godsill and Rayner, 1998a; Kokaram, 1998).

2. LINEAR MODELS

The models reviewed in this section are examples of the general linear model in which the data may be described in terms of a linear combination of basis functions with an additive Gaussian noise component. Such models can be used as a reasonable approximation to many signals including speech, music and digital communications channels. We express the model in general form as:

$$d_n = \sum_{q=1}^{Q} b_q g_q(n) + e_n \quad \text{for} \quad 1 \leq n \leq N$$

where $g_q(n)$ is the value of a time dependent model function $g_q(t)$ evaluated at time t_n. Expressed in matrix-vector notation we have:

$$\mathbf{d} = \mathbf{G}\,\mathbf{b} + \mathbf{e} \tag{1}$$

where:

d is an $N \times 1$ matrix of data points

e is an $N \times 1$ vector of noise samples

G is an $N \times Q$ matrix whose columns are the basis functions evaluated at each point in the time series

b is a $Q \times 1$ linear coefficient vector.

Many standard signal processing structures are representatives of this general linear model structure and some examples are:

Sinusoidal Model:

$$d_n = \sum_{q=1}^{Q} \{a_q \, \sin(\omega_q t_n) + b_q \, \cos(\omega_q t_n)\} + e_n$$

Autoregressive (AR) Model:

$$d_n = \sum_{q=1}^{Q} a_q \, d_{n-q} + e_n$$

Autoregressive with Exogenous Input (ARX) Model:

$$d_n = \sum_{q=1}^{Q} a_q \, d_{n-q} + \sum_{q=1}^{Q} b_q \, u_{n-q} + e_n$$

where u_n is an observed system input.

Non-linear Autoregressive (NAR) Model:

$$\begin{aligned} d_n &= \sum_{q=1}^{Q} a_q \, d_{n-q} \\ &+ \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} a_{q_1 q_2} \, d_{n-q_2} d_{n-q_2} \\ &+ \ldots + e_n \end{aligned}$$

Volterra Model:

$$\begin{aligned} d_n &= \sum_{q=0}^{Q} a_q \, u_{n-q} \\ &+ \sum_{q_1=0}^{Q_1} \sum_{q_2=0}^{Q_2} a_{q_1 q_2} \, u_{n-q_2} u_{n-q_2} \\ &+ \ldots + e_n \end{aligned}$$

If the error term or innovations process $\{e_n\}$ is a zero-mean white Gaussian process, then the data likelihood may be written as:

$$p\left(\mathbf{d} \mid \mathbf{w}_m, \sigma, \mathbf{b}, M\right) = \left(2\,\pi\,\sigma^2\right)^{-\frac{N}{2}} \exp\left[-\frac{\mathbf{e}^T \mathbf{e}}{2\,\sigma^2}\right]$$

where $\mathbf{w}_m$ denotes the parameters of the matrix of basis functions $\mathbf{G}$ (this of course is only an approximate likelihood for models with an autoregressive component (Box, Jenkins and Reinsel, 1994)).

The joint posterior density becomes:

$$p(\mathbf{w}_m,\, \mathbf{b},\, \sigma|\mathbf{d},\, M) = \frac{p(\mathbf{d}|\mathbf{w}_m,\, \mathbf{b},\, \sigma,\, M)\, p(\mathbf{w}_m|M)\, p(\mathbf{b}|M)\, p(\sigma|M)}{p(\mathbf{d}|M)}$$

Making assumptions about the prior probabilities for the linear parameters (uniform) and the noise standard deviation (Jeffreys prior), one can easily obtain the following expression for the marginal posterior probability for the non-linear parameters:

$$p(\mathbf{w}_m \mid \mathbf{d},\, M) \propto \frac{\left[\mathbf{d}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{d}\right]^{\frac{-(N-Q)}{2}}}{\sqrt{\det\left(\mathbf{G}^T\mathbf{G}\right)}} \tag{2}$$

We now describe some applications of the linear model structure in frequency estimation and changepoint analysis.

2.1 *Frequency Estimation*

As an example of the application of the general linear model, consider the detection of a single frequency.

$$f(t) = A\cos(\omega t) + B\sin(\omega t)$$

The data are assumed to consist of samples from the signal $f(t)$ corrupted with independent white zero mean Gaussian noise samples with standard deviation σ. The signal model just described agrees with the general linear model and the structure of the is:

$$\mathbf{G} = \begin{bmatrix} \cos(\omega t_1) & \sin(\omega t_1) \\ \cos(\omega t_2) & \sin(\omega t_2) \\ \cos(\omega t_3) & \sin(\omega t_3) \\ \vdots & \vdots \\ \cos(\omega t_N) & \sin(\omega t_N) \end{bmatrix}$$

Under some simplifying assumptions, the marginal density for the angular frequency may be expressed in terms of the Schuster periodogram, $C(\omega)$, as:

$$p(\omega \mid \mathbf{d},\, \mathbf{I}) \propto \left[1 - \frac{2\,C(\omega)}{\sum_{i=1}^{N} d_i^2}\right]^{\frac{2-N}{2}}$$

Note that the term inside the square brackets is small (thus implying that the marginal density is large) if $2\,C(\omega) \approx \sum_{i=1}^{N} d_i^2$. This occurs if most of the data energy is concentrated around a single frequency ω. The Schuster periodogram, (and hence the Discrete Fourier transform), is designed to determine the value of a single frequency in white Gaussian zero mean noise. Therefore it should really only be used on data that satisfies the single frequency model, otherwise a more general $\mathbf{G}$ matrix must be used.

2.2 *General Linear Changepoint Detection*

A second application of the general linear model is to changepoint detection. The central concept behind the linear changepoint detector is that an observed time sequence is modelled by different models at different points in time. The models are known but the time instants at which the models change are not known. The various models may be expressed as the linear combination of some basis functions. For data with a single changepoint from one linear model to another, the problem may be formulated as:

$$d_i = \begin{cases} \sum_{q=1}^{Q} \alpha_q \, g_q(i) + e_i & \text{if } i \leq m \\ \sum_{q=1}^{Q} \beta_q \, g_q(i) + e_i & \text{otherwise} \end{cases}$$

where $g_q(i)$ is the value of a time-dependent model function evaluated at time t_i. This may be expressed in the general linear model form:

$$\mathbf{d} = \mathbf{G}\,\mathbf{b} + \mathbf{e}$$

The model matrix $\mathbf{G}$ now contains the known basis function terms $g_q(i)$ and the unknown changepoint m. Equation 2 may be used to express the posterior probability distribution for the changepoint (Ó Ruanaidh and Fitzgerald, 1996).

The changepoint framework can be used to detect discrete phase changes of a sinusoidal carrier wave, as one would obtain in a BPSK or QPSK signal, without prior knowledge of the carrier phase. In Binary Phase Shift Keying (BPSK) a carrier sinusoid is modulated with digital information by means of 180 degree phase changes in the carrier. In order to model the 180° phase change, the changepoint model includes a change in the basis functions at the point where the presence of a 180° changepoint is being tested. The changepoint is modelled as a 180° rotation of the axes. This method of rotation of the axes is also used in the changepoint models for detecting 90° phase changes.

$$\mathbf{G} = \begin{bmatrix} \sin(\omega t_1) & \cos(\omega t_1) \\ \sin(\omega t_2) & \cos(\omega t_2) \\ \sin(\omega t_3) & \cos(\omega t_3) \\ \vdots & \vdots \\ \sin(\omega t_m) & \cos(\omega t_m) \\ -\sin(\omega t_{m+1}) & -\cos(\omega t_{m+1}) \\ \vdots & \vdots \\ -\sin(\omega t_N) & -\cos(\omega t_N) \end{bmatrix} \tag{3}$$

As an example see Figure 1 and Figure 2. A block of 80 samples of a noisy BPSK signal was simulated using a sine wave with a period of 20 samples. A phase change of 180° occurred at sample number 25. White Gaussian noise was added to the signal to produce a signal to noise ratio of 7.2 dB. Figure 2 shows that the most probable position for the phase change based on the data block was at sample number 23.

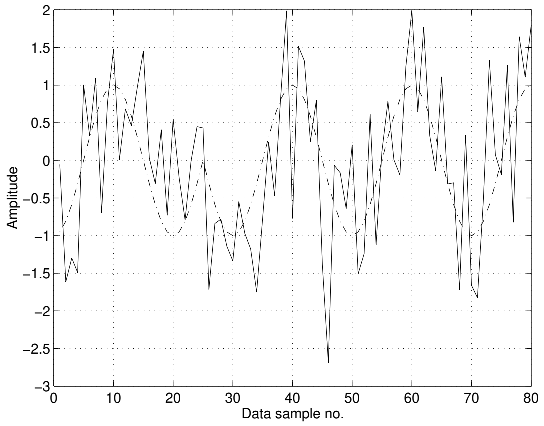


Figure 1: 180° phase change: noisy data (solid line) and clean signal (dash-dot line).

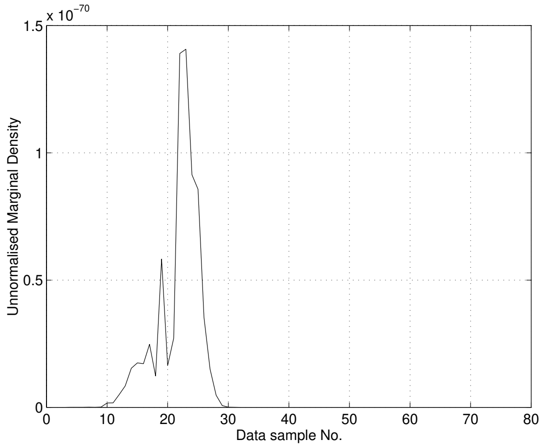


Figure 2: Plot of unnormalised marginal probability of a 180° phase change.

3. NON-GAUSSIAN AND NON-LINEAR MODELLING

Topics which are of key importance in Bayesian signal processing are the modelling of non-Gaussian and non-linear signals. The analytic methods reviewed in the previous section cannot in general be applied to the more elaborate models required here. Numerical analysis can be achieved using MCMC methods, and we review methods for the interpolation, enhancement and analysis of noise-degraded time series. The work finds application in the restoration of gramophone recordings and the enhancement of speech signals where noise degradation and the signals themselves are often highly non-Gaussian. In addition to non-Gaussian noise, audio signals are often distorted by non-linear processes such as groove deformation and tracking distortion in gramophone discs or non-linear communications channel characteristics. Recent work has also started to apply MCMC model uncertainty methods to the problem of detecting and correcting this type of non-linear defect.

3.1 *Enhancement of noise-degraded audio signals*

A classic example of heavy-tailed non-Gaussianity occurs in audio signals which have been recorded onto gramophone discs (Godsill and Rayner, 1998a) or passed through radio and other communications channels which are subject to atmospheric noise. See figure 3 for a typical example taken from a 78rpm gramophone recording. Note the characteristic 'spikes' which correspond to scratches or dust on the surface of the record. The approach taken here is closely related to MCMC work for time series in the presence of non-Gaussian noise and outliers, see e.g. (Shephard, 1994; McCulloch and Tsay, 1994; Carter and Kohn, 1994; Carter and Kohn, 1996). As a starting point we represent an audio signal $\{x_t\}$ as a scalar autoregressive (AR) process:

$$x_t = \sum_{i=1}^{P} x_{t-i}\, a_i \;+\; e_t$$

where $\{a_i\}$ are the Pth order AR coefficients. There is some physical basis for such a representation for vocal or musical instrument signals if $\{e_t\}$ is regarded as an excitation process (glottal waveform, instrument driving force, etc.) and $\{a_i\}$ is regarded as a resonant filtering opertation (vocal tract, instrument body, etc.), see e.g. Proakis, Deller and Hansen (1993). If we assume i.i.d. Gaussian noise for $\{e_t\}$ then a Gibbs Sampler can be devised to simulate from the unknown AR coefficients $\{a_i\}$, the residual variance σ_e^2 and any data points which are missing or heavily corrputed with noise. Such an approach has been investigated by (Ó Ruanaidh and Fitzgerald, 1994; Ó Ruanaidh and Fitzgerald, 1996) for the interpolation of long sections of missing data in musical recordings. A similar approach applied to the case of time varying AR models can be found in Rajan, Rayner and Godsill (1997). In cases where it is uncertain what the order P of the AR process should be, a reversible jump (Green, 1995) sampler can be used to incorporate model order uncertainty into the process, see (Troughton and Godsill, 1997b; Godsill, 1997b).

If we now assume an additive noise process, i.e. the observations are $y_t = x_t + v_t$, it is possible to perform outlier rejection and noise reduction. Outliers, or 'impulsive noise' can come from many possible sources in audio signals, including scratches or dust on a record surface and atmospheric interference in a radio environment. A model which has been applied successfully includes a binary 'indicator' process $\{i_t \in \{0,1\}\}$ to signify the presence of an outlier, and positive-valued auxiliary variables $\{g_t \in \mathbf{R}^+\}$, modelling heavy-tailed non-Gaussianity:

$$v_t \begin{cases} = 0, & i_t = 0 \\ \sim \mathrm{N}(0, g_t^2 \sigma_v^2), & i_t = 1 \end{cases}$$

Note that when $i_t = 1$ the noise is being modelled as a scale mixture of normals, a well known way of robustifying procedures (West, 1984; O'Hagan, 1988). $\{i_t\}$ is modelled as a Markov process and $\{g_t\}$ is assigned a prior appropriate to the type of non-Gaussianity present, e.g. Student or stable family distributions. Studies of the application of Gibbs Sampling methods to this model, which include an investigation of various blocking structures for speeding convergence can be found in (Godsill and Rayner, 1996b; Godsill and Rayner, 1998b).

A more general model has been investigated in Godsill and Rayner (1996a) and extended in Godsill (1997a) to the case of ARMA signals, in which a pure Gaussian noise component is included in the model in order to perform some background noise removal in addition to just impulse noise removal:

$$v_t \sim \begin{cases} \mathrm{N}(0, \sigma_v^2) & i_t = 0 \\ \mathrm{N}(0, g_t^2 \sigma_v^2), & i_t = 1 \end{cases}$$

Efficient Kalman filter based simulation methods (Frühwirth-Schnatter, 1994; Carter and Kohn, 1994; de Jong and Shephard, 1995) are used for drawing jointly all the elements $(x_1, \ldots, x_N)$ in this method. See figures 3-5 for a typical restoration example. Figure 3 shows the corrupted input record, digitised directly from a 78rpm disc at a 44.1kHz sampling rate. Figure 4 shows the MCMC estimated posterior mean for the signal $\{x_t\}$, while figure 5 shows the estimated posterior probabilities for the observational outlier process $\{i_t\}$ and an innovational outlier process which is modelled in an exactly similar way (innovational outliers allow for non-Gaussian 'voicing' pulses in the signal excitation process). It can be seen that the procedure successfully identifies and removes outliers from the waveform. When the procedure is applied to all the sub-frames within a particular recording a good degree of noise reduction and click removal can be perceived.

3.2 *Correction of non-linear distortion in audio signals*

If we now suppose that the signal is observed through some non-linearity which might be caused by saturation of electronics or magnetic recording media, groove deformation in gramophone recordings or non-linearity in a communications channel. This will give rise to unpleasant artefacts in the sound of the recording which should ideally be corrected. In general we will not have any very specific knowledge of the distortion mechanisms, so we adopt a fairly general modelling approach to the problem. Many models are possible for non-linear time series, see e.g. Tong (1990). In our initial work (Mercer, 1993; Troughton and Godsill, 1997a; Troughton and Godsill, 1998b; Troughton and Godsill, 1998a) we have adopted a cascade model in which the undistorted audio $\{x_t\}$ is modelled as a linear autoregression (as above) and the non-linear distortion process as a 'with memory' polynomial non-linear autoregressive (NAR) filter containing terms up to a particular lag p and order q, so that the observed output y_t can be expressed as:

$$\begin{aligned} y_t = x_t + \sum_{i_1=1}^{p} \sum_{i_2=1}^{i_1} b_{i_1 i_2}\, y_{t-i_1}\, y_{t-i_2} + \sum_{i_1=1}^{p} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} b_{i_1 i_2 i_3}\, y_{t-i_1}\, y_{t-i_2}\, y_{t-i_3} \\ + \ldots + \sum_{i_1=1}^{p} \sum_{i_2=1}^{i_1} \cdots \sum_{i_q=1}^{i_{q-1}} b_{i_1 \ldots i_q}\, y_{t-i_1} \, \ldots \, y_{t-i_q} \end{aligned} \tag{4}$$

The model is illustrated in figure 6. The problem here is to identify the significant terms which are present in the non-linear expansion of equation 4 and to neglect insignificant terms, otherwise the number of terms in the model becomes prohibitively large. This is dealt with using a model uncertainty framework in which each coefficient $b_{i_1 \ldots i_n}$ in the nth order expansion is allocated a binary indicator variable in much the same way as for the click removal procedures described above. In this way many of the terms in the expansion can be explicitly excluded from the model, according to their posterior probability. The problem is then essentially one of Bayesian variable selection, for which Gibbs sampling methods are well suited (Kuo and Mallick, 1997; George and McCulloch, 1993). Initial results show promise (see figures 7 and 8 for typical MCMC output from a non-linearly distorted AR process) and current work is focussing upon model elaborations which will make the scheme realistic for a variety of non-linear audio environments.

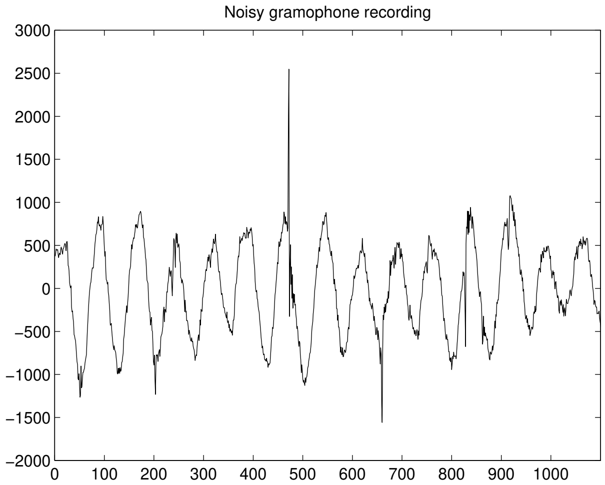


Figure 3: Noisy gramophone recording

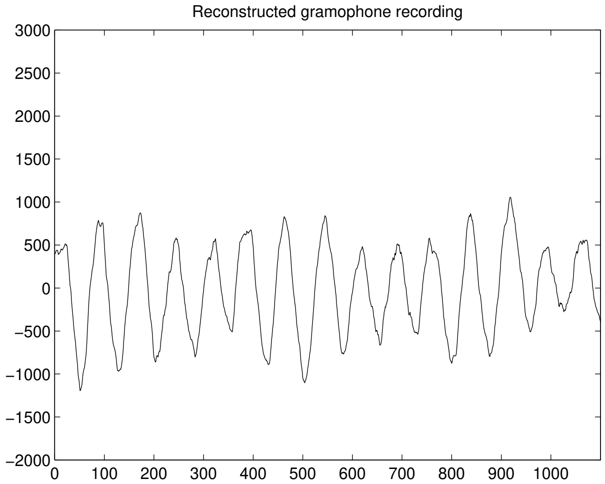


Figure 4: Reconstructed gramophone recording

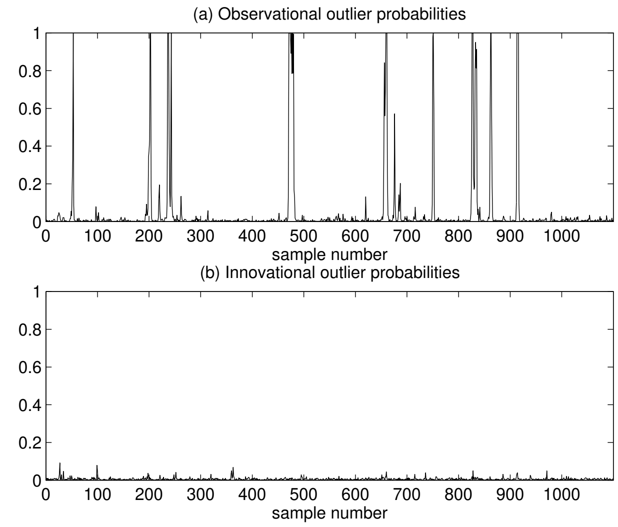


Figure 5: Outlier probabilities for gramophone recording

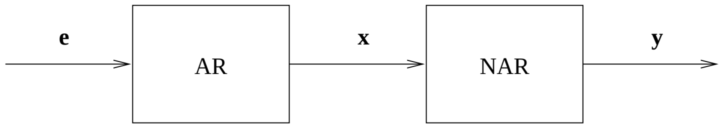


Figure 6: Block diagram of AR-NAR model

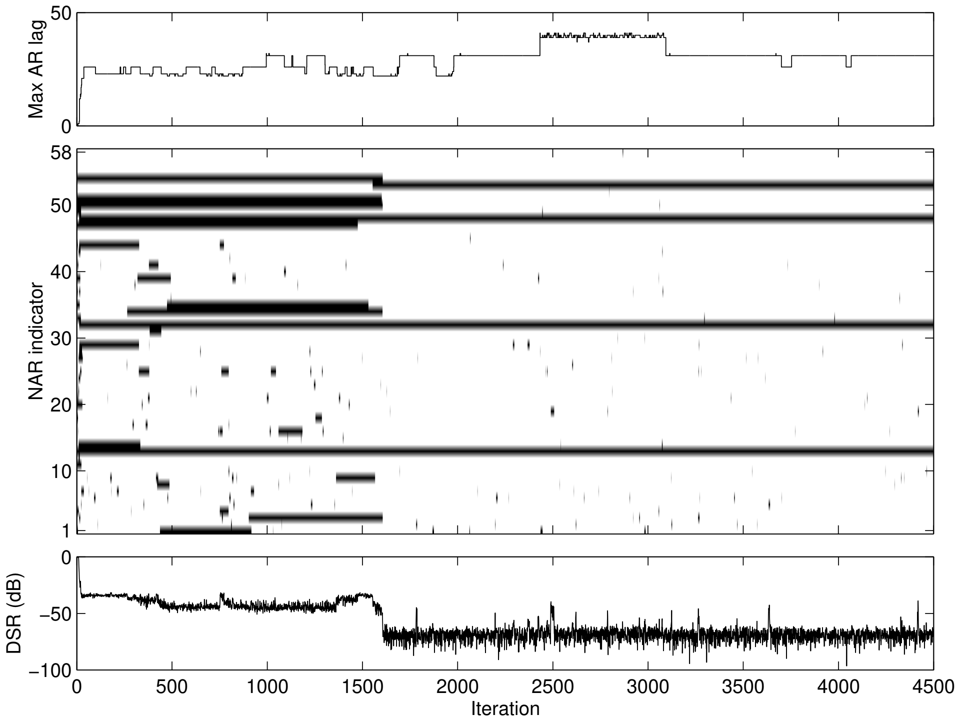


Figure 7: MCMC Output for AR-NAR restoration: Top - Linear AR model order (using a reversible jump scheme (Troughton and Godsill 1997b)); Middle - Indicator variables for non-linear model terms (black indicates a term is switched 'on'); Bottom - Mean squared error between restored and true AR data

4. IMAGE SEQUENCE ENHANCEMENT

A further area of importance to signal processing is image enhancement, and here we consider the enhancement of images from film and video footage which is degraded with dropouts and other replacement noise effects, leading to the 'speckled' and lined appearance of old film. These defects can be treated in a manner analogous to the audio restoration methods described above. Now, instead of the time series autoregression models which were assumed for audio signals we can represent the images in a sequence with spatio-temporal models which incorporate both the spatial and temporal continuity of the sequence. A suitable model is a spatio-temporal autoregressive model with motion offset:

$$x_s = \sum_{q \in \mathcal{S}} a_q x_{s-q-d(s_n, s_n - q_n)} + e_s, \qquad (s \in Z) \tag{5}$$

where $s = (s_i, s_j, s_n)$ is the location of the pixel (i.e. s represents co-ordinate (s_i, s_j) in

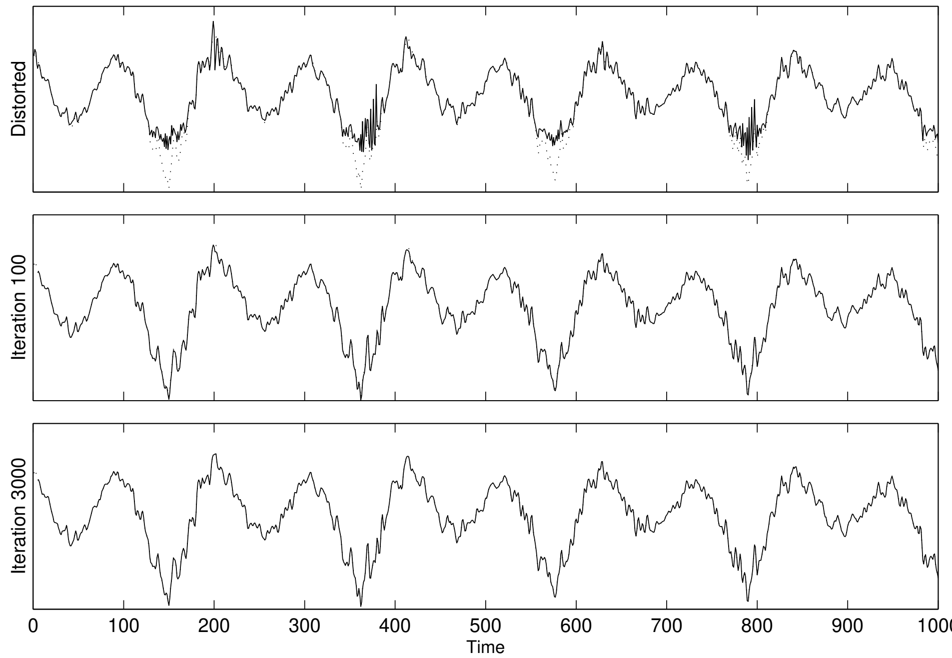


Figure 8: Example of AR-NAR restoration: Top - Non-linearly distorted input (solid), Undistorted input (dotted); Middle - Restored data (100 iterations); Bottom - Restored data (3000 iterations)

frame s_n of the sequence) and x_s is the continuous grey-scale intensity at that coordinate (we neglect the effects of quantisation). Z is an integer lattice of pixels (i, j, n) which will typically index an $N \times M$ sub-block within a particular frame of the sequence. $\mathcal{S}$ defines a spatio-temporal neighbourhood of support for each pixel which we choose to be *causal* or *unilateral* (Chellappa, 1985). a_q is a weighting factor for support element q and $d(n, m) = (d_i(n, m), d_j(n, m), 0)$ is a spatial offset, or *motion vector*, which corrects for translational motion of the image patch Z between frames n and m. e_s is a prediction error which is assumed to be white and Gaussian with variance σ_e^2. The restriction to a causal region of support, although less general than a full non-causal support (Chellappa, 1985), facilitates estimation of the parameters a_q, since the conditionals for the $\{a_q\}$ are then approximately Gaussian, and appears general enough to model the textures found in most film image sequences.

The degradation in image sequences typically replaces the image information for small 'patches' of the frame. We can represent this in the following way:

$$y_s = \begin{cases} x_s, & i_s = 0 \\ c_s, & i_s = 1 \end{cases}$$

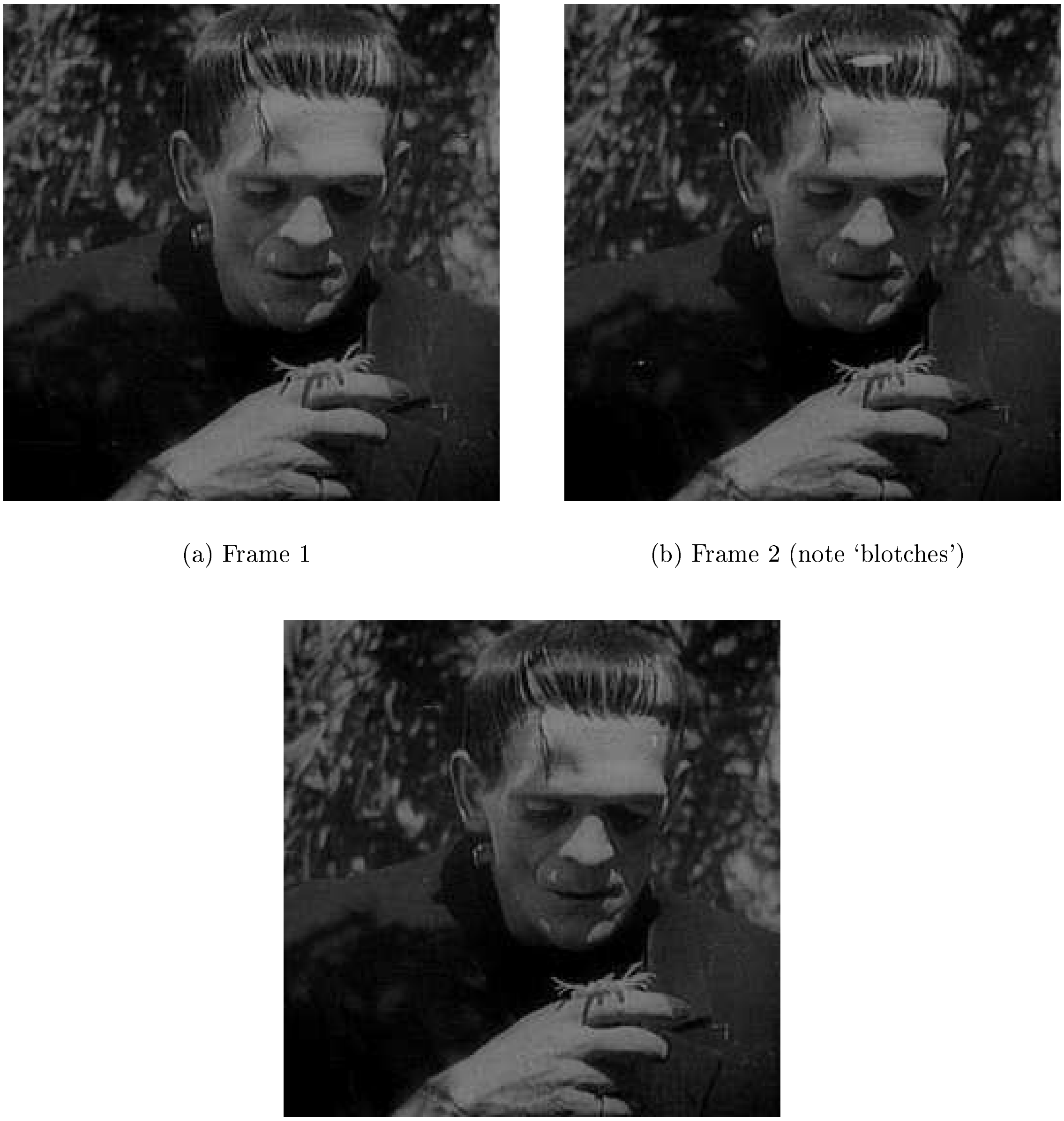
where once again i_s is a binary indicator variable which indicates whether the replacement noise process $\{c_s\}$ is present or not. Except in the special case of line scratches (see Morris (1996)) there will be no temporal continuity in the degradation. Hence a purely spatial Markov random field prior is used to model just spatial continuity in both the amplitude corruption process $\{c_s\}$ and the binary indicator process $\{i_s\}$.

The entire collection of unknowns includes the weighting coefficients $\{a_q; q \in S\}$, the motion offsets $d(.,.)$, the reconstructed data $\{x_s; s \in Z\}$, the replacement process

$\{c_s; s \in Z\}$ and the excitation variance σ_e^2. These can be simulated from the joint posterior using the Gibbs sampler, and the scheme adopted uses the following blocking structure to take advantage of analytic structure within the model:

$$(\mathbf{a}, \sigma_e^2, \mathbf{d}) \sim p(\mathbf{a}, \sigma_e^2, \mathbf{d}|\mathbf{y}, \mathbf{c}, \mathbf{x})$$
$$(x_s, c_s, i_s) \sim p(x_s, c_s, i_s|\mathbf{y}, \mathbf{i}_{-s}, \mathbf{x}_{-s}, \mathbf{a}, \sigma_e^2, \mathbf{d})$$

where an obvious vector notation has been adopted for groups of unknowns. Full details of the various approaches can be found in (Godsill and Kokaram, 1996; Kokaram and Godsill, 1996; Godsill and Kokaram, 1997; Kokaram and Godsill, 1997; Kokaram, 1998).



(a) Frame 1

(b) Frame 2 (note 'blotches')

(c) Frame 3

Figure 9: Three frames from degraded movie sequence

(a) Frame 2 of degraded sequence (as above)

(b) MCMC estimated detection field $\{i_s\}$ for frame 2 of sequence ($i_s = 1$ corresponds to white pixels)

(c) MCMC estimated data $\{x_s\}$ for frame 2 of sequence

Figure 10: MCMC detection and restoration

Three frames from a movie sequence can be seen in figures 9(a)-9(c). We proceed to restore the second of these frames, which has some noticeable 'blotches'. The method described above is applied to the whole frame, allowing for different sets of AR coefficients and motion vectors over a grid of sub-patches in the image, since the image characteristics and motion are spatially varying. Figure 10(b) shows the MCMC estimate of the detection field for corrupted regions, $\{i_s\}$, which has clearly identified the main regions of corruption. Figure 10(c) shows the corresponding restoration, in which all the visible 'blotches' have been successfully removed from the image.

As for the audio case, it is straightforward to extend the methods to the case where $\{y_s\}$ is observed in additive Gaussian or non-Gaussian noise, and this is one of the topics of our current work (Kokaram and Godsill, 1998)

5. CONCLUSIONS

In this paper we have described how modern Bayesian methods can be applied to a few of the important problems that occur in signal and image processing. By its very nature, signal processing is an enormous field with many problems that have, as yet, not been approached from the Bayesian perspective. However, our current research is addressing many other topics such as sequential methods for tracking, detection and identification, digital communications in non-Gaussian and time-varying channels, radar clutter removal, image segmentation, source separation (independent factor analysis) and beam-forming. A current area of interest is the modelling of physical (heavy-tailed) noise processes encountered in atmospheric interference and underwater acoustics. These are well modelled by the stable law distribution family, for which MCMC methods are well suited.

The material presented in this paper presents only a small fraction of the possible areas to which modern Bayesian methodology can be successfully applied, and we believe that the next few years will witness a proliferation of these techniques throughout the signal processing community, in a similar fashion to the revolution recently seen in the statistical community with the introduction of MCMC and other sophisticated numerical techniques. In particular the signal processing community will require the development of computer efficient real time (on-line) methods which can readily be applied inexpensively to real world problems.

6. A NOTE ABOUT REFERENCES

In this paper we include citations to many papers from outside the main statistical literature. Preprints of many of these papers can be found on the following websites: http://www-sigproc.eng.cam.ac.uk/, http://www-sigproc.eng.cam.ac.uk/ sjg and http://www-sigproc.eng.cam.ac.uk/ ack.

REFERENCES

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall 3rd edn.

Carter, C. and Kohn, R. (1996). Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika* **83** 589–601.

Carter, C. K. and Kohn, R. (1994). On Gibbs Sampling for state space models. *Biometrika* **81**(3) 541–553.

Chellappa, R. (1985). Two-dimensional discrete Gaussian Markov random field models for image processing. In L. Kanal and A. Rosenfeld (eds.), *Progress in Pattern Recognition* 79–112. Elsevier.

de Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. *Biometrika* **82**(2) 339–350.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis* **15** 183–202.

George, E. I. and McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of American Statistical Association* **88**(423) 881–889.

Godsill, S. and Rayner, P. (1998a). *Digital Audio Restoration*. Berlin: Springer.

Godsill, S. J. (1997a). Bayesian enhancement of speech and audio signals which can be modelled as ARMA processes. *International Statistical Review* **65**(1) 1–21.

Godsill, S. J. (1997b). Robust modelling of noisy ARMA signals. In *Proc. International Conference on Acoustics, Speech and Signal Processing*.

Godsill, S. J. and Kokaram, A. C. (1996). Joint interpolation, motion and parameter estimation for image sequences with missing data. In *Proc. EUSIPCO*.

Godsill, S. J. and Kokaram, A. C. (1997). Restoration of image sequences using a causal spatio-temporal model. In *Proc. 17th Leeds Annual Statistics Research Workshop (The Art and Science of Bayesian Image Analysis)*.

Godsill, S. J. and Rayner, P. J. W. (1996a). Robust noise reduction for speech and audio signals. In *Proc. International Conference on Acoustics, Speech and Signal Processing*.

Godsill, S. J. and Rayner, P. J. W. (1996b). Robust treatment of impulsive noise in speech and audio signals. In J. Berger, B. Betro, E. Moreno, L. Pericchi, F. Ruggeri, G. Salinetti and L. Wasserman (eds.), *Bayesian Robustness - proceedings of the workshop on Bayesian robustness, May 22-25, 1995, Rimini, Italy* vol. 29 331–342. IMS Lecture Notes - Monograph Series.

Godsill, S. J. and Rayner, P. J. W. (1998b). Robust reconstruction and analysis of autoregressive signals in impulsive noise using the Gibbs sampler. *IEEE Trans. on Speech and Audio Processing* Previously available as Tech. Report CUED/F-INFENG/TR.233.

Green, P. J. (1995). Reversible jump Markov-chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**(4) 711–732.

Kokaram, A. (1998). *Motion Picture Restoration*. Berlin: Springer.

Kokaram, A. and Godsill, S. (1998). Joint noise reduction, motion estimation, missing data reconstruction, and model parameter estimation for degraded motion pictures. In *Proc. SPIE, San Diego*.

Kokaram, A. C. and Godsill, S. J. (1996). A system for reconstruction of missing data in image sequences using sampled 3D AR models and MRF motion priors. In *Computer Vision - ECCV '96* vol. II 613–624. Springer Lecture Notes in Computer Science.

Kokaram, A. C. and Godsill, S. J. (1997). Detection, interpolation, motion and parameter estimation for image sequences with missing data. In *Proc. International Conference on Image Applications and Processing, Florence, Italy*.

Kuo, L. and Mallick, B. (1997). Variable selection for regression models. *Sankhya* (to appear).

McCulloch, R. E. and Tsay, R. S. (1994). Bayesian analysis of autoregressive time series via the Gibbs sampler. *Journal of Time Series Analysis* **15**(2) 235–250.

Mercer, K. J. (1993). *Identification of signal distortion models*. Ph.D. thesis University of Cambridge.

Morris, R. (1996). A sampling based approach to line scratch removal from motion picture frames. In *Proc. IEEE International Conference on Image Processing, Lausanne, Switzerland*.

O'Hagan, A. (1988). Modelling with heavy tails. *Bayesian Statistics 3* 345–359.

Ó Ruanaidh, J. J. K. and Fitzgerald, W. J. (1994). Interpolation of missing samples for audio restoration. *Electronic Letters* **30**(8).

Ó Ruanaidh, J. J. K. and Fitzgerald, W. J. (1996). *Numerical Bayesian methods applied to signal processing*. Springer-Verlag.

Proakis, J., Deller, J. and Hansen, J. (1993). *Discrete-Time Processing of Speech Signals*. Macmillan, New York.

Rajan, J. J., Rayner, P. J. W. and Godsill, S. J. (1997). A Bayesian approach to parameter estimation and interpolation of time-varying autoregressive processes using the Gibbs sampler. *IEE Proc. Vision, image and signal processing* **144**(4).

Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika* **81**(1) 115–131.

Tong, H. (1990). *Non-linear Time Series*. Oxford Science Publications.

Troughton, P. and Godsill, S. (1998a). MCMC methods for restoration of nonlinearly distorted autoregressive signals. In *Proc. European Conference on Signal Processing*.

Troughton, P. T. and Godsill, S. J. (1997a). Bayesian model selection for time series using Markov chain Monte Carlo. In *Proc. International Conference on Acoustics, Speech and Signal Processing*.

Troughton, P. T. and Godsill, S. J. (1997b). A reversible jump sampler for autoregressive time series, employing full conditionals to achieve efficient model space moves. Tech. Rep. CUED/F-INFENG/TR.304 Cambridge University Engineering Department.

Troughton, P. T. and Godsill, S. J. (1998b). Bayesian model selection for linear and non-linear time series using the Gibbs sampler. In J. G. McWhirter (ed.), *Mathematics in Signal Processing IV*. Oxford University Press.

West, M. (1984). Outlier models and prior distributions in Bayesian linear regression. *Journal of the Royal Statistical Society, Series B* **46**(3) 431–439.