M5.4

MARKOVIAN MODELING FOR BAYESIAN MULTI-CHANNEL DECONVOLUTION

J. Idier and Y. Goussard

Laboratoire des Signaux et Systèmes (CNRS / ESE / UPS)
École Supérieure d'Électricité, Plateau de Moulon
91192 Gif-sur-Yvette Cédex, France.

ABSTRACT

2-D Bayesian restoration of layered media and application to seismic deconvolution are addressed. The well known Bernoulli-Gaussian (B-G) approach [1] is suboptimal in a 2-D setting, since it does not account for spatial continuity. To make up for this deficiency, a new class of Markov random fields is built, which suitably extends 1-D Bernoulli processes to two dimensions. It provides a manageable model for 2-D stratified structures. Application to seismic deconvolution is considered: a *maximum a posteriori* detection-estimation algorithm is proposed. Compared to standard B-G algorithms, it performs much better without any increase in computational load.

I. INTRODUCTION

In this communication, we address the problem of estimating the shape and nature of homogeneous layers in a stratified structure, from acoustic measurements performed at its surface. The major fields of application are seismic exploration, non-destructive evaluation, biomedical imaging.

Under the assumption that the general orientation of the layers is horizontal, the propagation medium can be characterized by its reflectivity, i.e., the vertical logarithmic derivative of the acoustic impedance. In a one-dimensional (1-D) setting, the observed data can be considered as the noise-corrupted output of a linear system whose input is a vertical reflectivity sequence. When the layers are homogeneous, the incident wave only reflects at the boundaries between layers so that reflectivity appears as a sparse spike train. Restoration of pulse trains has been studied extensively. A Bayesian approach based on Bernoulli-Gaussian (B-G) modeling of the reflectivity proved to be very powerful. Throughout the paper, a B-G (λ, σ) random variable $X = (Q, R)$ is defined as *(1)* Q: 0-1 binary variable with $P(Q = 1) = \lambda$; *(2)* $R \mid Q = q$: Gaussian variable with zero mean and variance $q\sigma^2$. B-G modeling was introduced by Kwakernaak [2] in automatic control and Kormylo and Mendel [3, 1] in seismic deconvolution. In the last decade, several other works have been carried out in the field of B-G seismic deconvolution [4-6]. They provided efficient and rather costless detection-estimation algorithms to restore spike trains.

In most applications, one wishes to restore a whole section made of the juxtaposition of N 1-D reflectivity sequences, so as to obtain an image of the structure under investigation. Standard B-G methods are based upon independent processing of the N signals, which makes them clearly suboptimal since the general horizontal orientation of the layers induces a strong prior correlation between adjoining columns. To make up for this deficiency, a 2-D prior model should (i) account for spatial interactions between reflectivity columns, (ii) be a 2-D extension of the 1-D B-G representation and (iii) yield numerically tractable restoration methods.

Markov models naturally check constraint (i). However, the Markov random fields (MRFs) frequently used in image segmentation [7] neither check (ii), nor (iii). Therefore, it seems desirable to define a new class of Markov processes, specifically suited to model stratified media. A Markov-Bernoulli-Gaussian (M-B-G) model is presented in part II. It is shown to fulfill conditions (i) and (ii). In part III, a suboptimal algorithm for multi-channel seismic deconvolution is proposed. Its simple recursive structure is a direct consequence of the specific characteristics of the M-B-G model which is used as Bayes priors. A simulated example follows. Concluding remarks and some open issues are mentioned in part IV.

II. THE M-B-G MODEL

1. A hierarchical construction

Let $\Lambda^\circ = \{(m, n): m = 1, ..., M;\ n = 1, ..., N\}$ denotes a $M \times N$ rectangular grid. To each site (m, n) we assign a binary random variable Q_{mn} and a real-valued random variable R_{mn}. The *location* process $\mathbf{Q} = \{Q_{mn}: (m, n) \in \Lambda^\circ\}$ locates reflectors, whereas the amplitude *process* $\mathbf{R} = \{R_{mn}: (m, n) \in \Lambda^\circ\}$ models reflectivity values. For any site (m, n) with $Q_{mn} = 0$, we have also $R_{mn} = 0$. $\{\mathbf{Q}, \mathbf{R}\}$ is *hierarchically* organized: distribution of $\mathbf{R}$ is naturally defined *conditionally* to $\mathbf{Q}$ (see II.3.), whereas $\mathbf{Q}$ can be handled before (or apart from) the definition of $\mathbf{R}$.

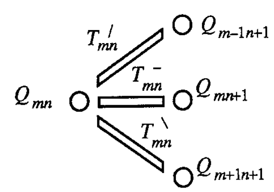
CH2847-2/90/0000-2001 $1.00 © 1990 IEEE

Such a structure justifies a two-stage definition: section II.2. is devoted to the construction of **Q**, which is the essential part and leads us to introduce an even higher level in the hierarchy (i.e. process **T**); **R** is defined afterwards in section II.3.

2. The binary Markov model

2.1. Introducing transition variables T_{mn}

Definition of **Q** requires the specification of the joint probability measure $P(\mathbf{Q} = \mathbf{q})$, for any $M \times N$ binary matrix $\mathbf{q}$. Let $\mathbf{Q}_n$ be the nth column in **Q**. From (i) and (ii), it seems natural to consider **Q** as a vector Markov chain, $\mathbf{Q}_n$ being the nth element of the chain. Such an approach is attractive, provided that we are able to specify transition probabilities $P(\mathbf{q}_{n+1} | \mathbf{q}_n)$ in accordance with (i) and (ii). Obviously, the huge number of possible transitions ($2^N \times 2^N$) precludes any explicit definition. Instead, local interactions must be utilized to give rise to a tractable factored form of the transition probabilities. To do so, the key idea is the introduction of some underlying binary variables T_{mn}, meant to link adjacent sites along layer boundaries. Since layers are roughly horizontal, it suffices to define junction sites between every pair of sites diagonally or horizontally adjacent according, to the following scheme:



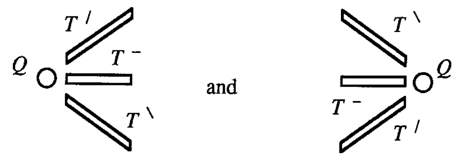
Transition process **T** is the random field $\{T^{/}_{mn}\}_{\Lambda^{/}} \cup \{T^{-}_{mn}\}_{\Lambda^{-}} \cup \{T^{\backslash}_{mn}\}_{\Lambda^{\backslash}}$. $\Lambda^{/}, \Lambda^{-}, \Lambda^{\backslash}$ are the three sets of junction sites. The key role of the *transition* process springs from the fact that reflectors are correlated (in location and value) if and only if they belong to the same layer boundary. Through *transition* variables, interactions between *location* variables are easy to introduce, whereas the joint probability of *location* variables (marginally from **T**) does not retain locality (see [8] and similar comments in [7] at the end of part IV).

2.2. Markov random field $\{\mathbf{T}, \mathbf{Q}\}$

We now turn to the joint characterization of $\{\mathbf{T}, \mathbf{Q}\}$ instead of **Q** alone. To begin with, we impose a Markov chain structure to $\{\mathbf{T}, \mathbf{Q}\}$:

$$P(\mathbf{t}, \mathbf{q}) = P(\mathbf{q}_1) \prod_{n=1}^{N-1} P(\mathbf{t}_n | \mathbf{q}_n) P(\mathbf{q}_{n+1} | \mathbf{t}_n) \tag{1a}$$

Then, we decompose each conditional probability in (1a) into a product of local probabilities defined by a unique measure $\tau(q, t^{/}, t^{-}, t^{\backslash})$ on two generic cells, symmetric to one another:



$$\begin{aligned} P(\mathbf{t}_n | \mathbf{q}_n) &= \prod_{m=1}^{M} \tau(t^{/}_{mn}, t^{-}_{mn}, t^{\backslash}_{mn} | q_{mn}) \\ P(\mathbf{q}_{n+1} | \mathbf{t}_n) &= \prod_{m=1}^{M} \tau(q_{mn+1} | t^{/}_{m+1n}, t^{-}_{mn}, t^{\backslash}_{m-1n}) \end{aligned} \tag{1b}$$

with adjustments at horizontal grid boundaries (undefined variables must be omitted). In the next section, attention will be focused on defining measure τ in accordance with (i) and (ii). Now for any τ, it is clear from (1) that $\{\mathbf{T}, \mathbf{Q}\}$ is a unilateral Markov random field (UMRF) on lattice $\Lambda = \Lambda^{\circ} \cup \Lambda^{/} \cup \Lambda^{-} \cup \Lambda^{\backslash}$. UMRFs form a very interesting subclass of general MRFs; for a comprehensive introduction to MRFs, UMRFs and some applications, we refer the reader to [9], [10], [7], and [11], [12]. The MRF property is straightforward to check;

$$\begin{aligned} \mathbf{N}^{\circ}_{mn} &= \{T^{/}_{mn}, T^{-}_{mn}, T^{\backslash}_{mn}, T^{/}_{m+1n-1}, T^{-}_{mn-1}, T^{\backslash}_{m-1n-1}\} \\ \mathbf{N}^{/}_{mn} &= \{Q_{mn}, T^{-}_{mn}, T^{\backslash}_{mn}, Q_{m-1n+1}, T^{-}_{m-1n}, T^{\backslash}_{m-2n}\} \\ \mathbf{N}^{-}_{mn} &= \{Q_{mn}, T^{/}_{mn}, T^{\backslash}_{mn}, Q_{mn+1}, T^{/}_{m+1n}, T^{\backslash}_{m-1n}\} \\ \mathbf{N}^{\backslash}_{mn} &= \{Q_{mn}, T^{-}_{mn}, T^{/}_{mn}, Q_{m+1n+1}, T^{-}_{m+1n}, T^{/}_{m+2n}\} \end{aligned} \tag{2}$$

are respectively the neighborhood sets of interior sites (m, n) in $\Lambda^{\circ}, \Lambda^{/}, \Lambda^{-}, \Lambda^{\backslash}$. In case of boundary sites, undefined variables appearing in (2) are simply omitted.

2.3. Defining the measure τ

From (1), the field $\{\mathbf{T}, \mathbf{Q}\}$ falls into the class of UMRFs, which yields a recursive representation for the joint probability measure $P(\mathbf{t}, \mathbf{q})$. This is a very positive feature with regard to the tractability (iii), but τ should now be specified precisely so as to meet successively property (i) and (ii).

The first point is to impose rules of dependency between **Q** and **T** according to their practical interpretation:

$$\tau(q = 1 | t^{/} = 1 \text{ or } t^{-} = 1 \text{ or } t^{\backslash} = 1) = 1 \tag{3a}$$

$$\tau(q = 1 | t^{/} = 0, t^{-} = 0, t^{\backslash} = 0) = \varepsilon \quad (\varepsilon << 1) \tag{3b}$$

Remark 1. In theory, (3a) breaks the positivity condition (in Hammersley-Clifford terminology [10]). However, this deficiency have no negative consequences in practice [8].

Remark 2. Not to impose $\varepsilon = 0$ may appear rather unfounded. In fact, several factors motivate a non-deterministic dependence in (3b) [8], the most intuitive being the possible existence of local discontinuities in the layered structure (for instance, geological faults in seismic exploration).

Then, it is clear from (1a) that $\mathbf{Q}$ is a vector Markov chain. A necessary condition for every $\mathbf{Q}_n$ to be Bernoulli distributed is the invariance of τ on $\{\mathbf{T}, \mathbf{Q}\}$. In [8], it is shown that τ is invariant if and only if $\tau(t^{/}, t^{-}, t^{\backslash}) = \tau(t^{/}).\tau(t^{-}).\tau(t^{\backslash})$. This property let us define a class of stationnary MRFs, namely Markov-Bernoulli random fields (MBRFs). It is described by four parameters: $\mu^{/} = \tau(t^{/} = 1)$, $\mu^{-} = \tau(t^{-} = 1)$, $\mu^{\backslash} = \tau(t^{\backslash} = 1)$ and $\varepsilon = \tau(q = 1 | t^{/} = 0, t^{-} = 0, t^{\backslash} = 0)$. Every $\mathbf{Q}_n$ is Bernoulli, with $\lambda = 1 - (1-\varepsilon)(1-\mu^{/})(1-\mu^{-})(1-\mu^{\backslash})$: property (ii) is fulfilled. Additional properties are investigated in [8].

3. *Amplitude* process R

Conditionally to $\{\mathbf{T}, \mathbf{Q}\}$, it is natural to write the probability measure of $\mathbf{R}$ in a Markov product form:

$$P(\mathbf{r} | \mathbf{t}, \mathbf{q}) = P(\mathbf{r}_1 | \mathbf{q}_1) \prod_{n=2}^{N} P(\mathbf{r}_n | \mathbf{q}_n, \mathbf{t}_{n-1}, \mathbf{q}_{n-1}, \mathbf{r}_{n-1}) \qquad (4)$$

where every conditional probability in the r.h.s. factors into a product of M locally dependent terms $P(r_{mn} | \ldots)$. The r_{mn}'s are assumed to be sampled from one of two possible Gaussian distributions, depending whether r_{mn} has a *predecessor* $r_{m+dm,n-1}$ or not (by *predecessor* we mean a reflector belonging to the same boundary; dm can take values $-1, 0, 1$) :

1. If r_{mn} has a *predecessor*, then r_{mn} is sampled from a first-order AR process: $r_{mn} = (1-a^2)r_{m+dm,n-1} + n_r$, where n_r is Gaussian with zero mean and variance $a\sigma^2$ and a is adjustable between 0 and 1 to control the similarity between spikes along the same boundary.

2. Otherwise, r_{mn} is simply sampled from the Gaussian distribution with zero-mean and variance σ^2.

Such specifications are compatible with constraints (i), (ii) and (iii). Finally, the joint probability $P(\mathbf{x}) = P(\mathbf{r} | \mathbf{t}, \mathbf{q})P(\mathbf{t}, \mathbf{q})$ of the M-B-G model $\mathbf{X} = \{\mathbf{T}, \mathbf{Q}, \mathbf{R}\}$ is fully factorable into locally dependent probabilities.

III. MULTI-CHANNEL SEISMIC DECONVOLUTION

1. From B-G to M-B-G deconvolution

Application of M-B-G modeling to Bayesian seismic deconvolution is considered now, in order to extend (and surpass) the B-G approach.

B-G deconvolution consists of performing MAP estimation of a single B-G(λ, σ) column $\mathbf{x}_1 = (\mathbf{q}_1, \mathbf{r}_1)$ given noisy blurred observations $\mathbf{z}_1 = \mathbf{h} * \mathbf{x}_1 + \mathbf{n}_1$. Blur function $\mathbf{h}$ is assumed to be known, as well as λ, σ and the signal to noise ratio (SNR); $\mathbf{n}_1$ is a Gaussian white noise. Rather than maximizing the *a posteriori* likelihood $P(\mathbf{x}_1 | \mathbf{z}_1)$, it has been empirically shown [1], [6] that the following detection-estimation procedure is preferable:

$$\begin{aligned} &1. \text{ Detection of the reflectors: } \hat{\mathbf{q}}_1 = \arg\max_{\mathbf{q}_1} P(\mathbf{q}_1 | \mathbf{z}_1) \\ &2. \text{ Amplitude estimation: } \hat{\mathbf{r}}_1 = \arg\max_{\mathbf{r}_1} P(\mathbf{r}_1 | \mathbf{z}_1, \hat{\mathbf{q}}_1) \end{aligned} \qquad (5)$$

The first step is a combinatorial exploration problem, suboptimally performed by a deterministic iterative search (e.g. SMLR [1]), while the second step consists of optimal minimum-variance estimation.

Now let us consider the extension to the deconvolution of N successive columns $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ given the N corresponding observation vectors $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$. Let $\mathbf{X} = \{\mathbf{T}, \mathbf{Q}, \mathbf{R}\}$ be a M-B-G model as defined in the previous section. Optimal MAP estimation would involve a Viterbi algorithm. Unfortunately such a procedure is intractable, considering the huge size of the corresponding state space. Instead, we propose to compute $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N$ recursively according to the two-step MAP procedure:

$$\begin{aligned} &1. \; (\hat{\mathbf{t}}_{n-1}, \hat{\mathbf{q}}_n) = \arg\max_{(\mathbf{t}_{n-1}, \mathbf{q}_n)} P(\mathbf{t}_{n-1}, \mathbf{q}_1 | \mathbf{z}_n, \hat{\mathbf{q}}_{n-1}, \hat{\mathbf{r}}_{n-1}) \\ &2. \; \hat{\mathbf{r}}_n = \arg\max_{\mathbf{r}_n} P(\mathbf{r}_n | \mathbf{z}_n, \hat{\mathbf{t}}_{n-1}, \hat{\mathbf{q}}_n, \hat{\mathbf{r}}_{n-1}) \end{aligned} \qquad (6)$$

for $n = 2, \ldots, N$. The first step locates reflectors in column $\hat{\mathbf{x}}_n$, whose values are estimated by the second step. The algorithm is not self-starting: for $n = 1$, apply (5) instead. Implementation of (6) is very similar to repeated applications of (5), and it requires neither more memory nor processing time. Finally, let us note that the whole B-G approach appears as a degenerate case of M-B-G restoration, for which the underlying MBRF is characterized by $\mu^{/} = \mu^{-} = \mu^{\backslash} = 0$ and $\varepsilon = \lambda$.

2. Simulation results

Figure 1 provides a comparison between B-G (fig.1d) and M-B-G (fig.1e) performance. The signal to restore (fig.1a) is a synthetic reflectivity section. Fig.1b shows the source wavelet (i.e. vertical blur function) and fig.1c the noisy measurements (SNR = 10dB) from which fig.1d and fig.1e are processed. For both methods, the blur function and the hyperparameters are known or empirically chosen. Comparison between the two clearly shows the superiority of the M-B-G approach and validates the MBRF model.

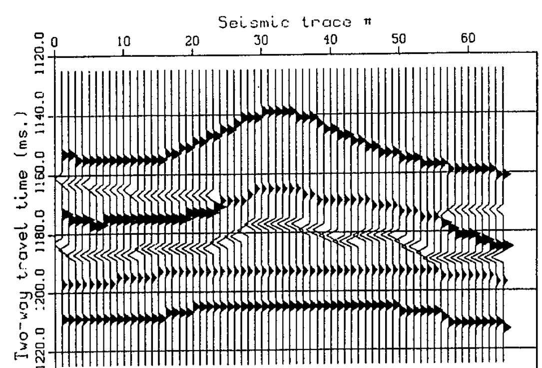


Fig.1a: Actual reflectivity

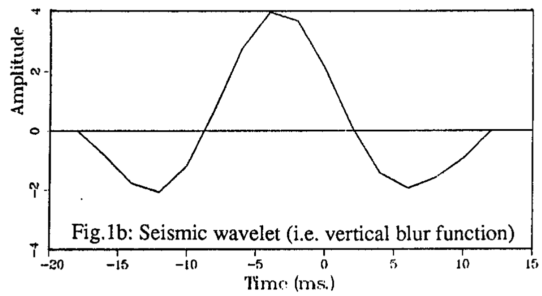


Fig.1b: Seismic wavelet (i.e. vertical blur function)

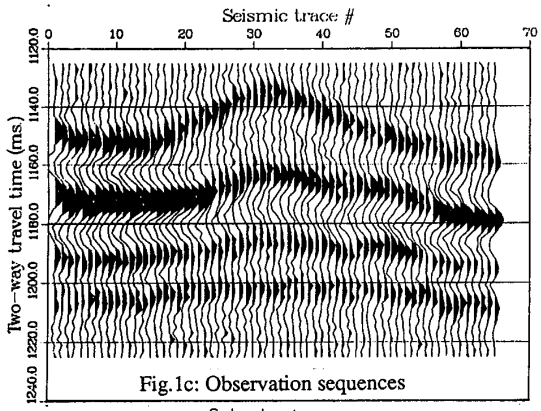


Fig.1c: Observation sequences

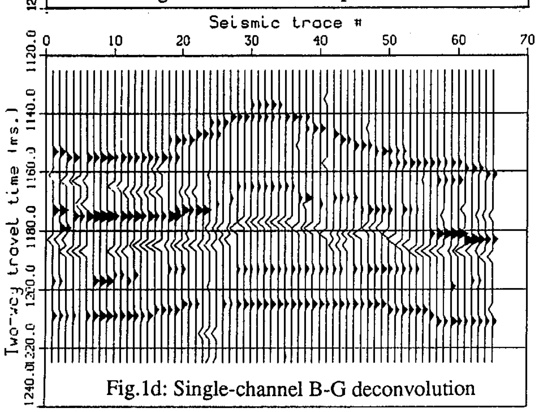


Fig.1d: Single-channel B-G deconvolution

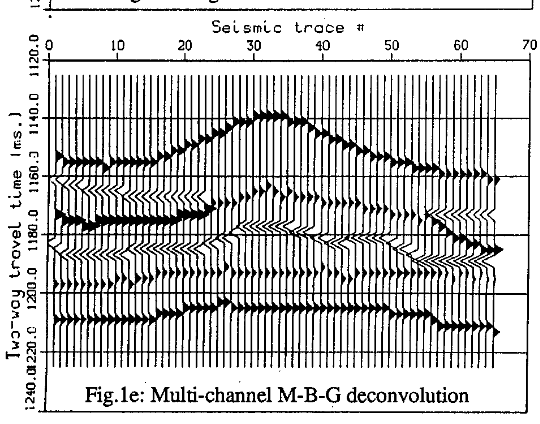


Fig.1e: Multi-channel M-B-G deconvolution

IV. CONCLUSION

In this paper we have proposed the construction of Markov-Bernoulli random fields (MBRFs), which provides a general frame for 2-D Bayesian restoration of layered structures. Application to seismic deconvolution has been tackled, and an original multi-channel algorithm has been designed and validated. Further development includes the use of MRF restoration techniques in conjunction with MBRF modeling to restore full layers instead of reflectors. Finally, the problem of estimating the hyperparameters and the blur function is a general open issue, for which very few results are available yet.

REFERENCES

[1] J.M. Mendel, "Optimal seismic deconvolution," Academic Press, New York, 1983.

[2] H. Kwakernaak, "Estimation of pulse heights and arrival times," *Automatica*, **16**, 367-377, 1980.

[3] J. Kormylo and J.M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Information Theory*, **IT-28**, 482-488, 1982.

[4] A.K. Mahalanabis, S. Prasad and K.P. Mohandas, "Recursive decision-directed estimation of reflection coefficients for seismic data deconvolution," *Automatica*, **18**, 721-726, 1982.

[5] S.D. Kollias and C.C. Halkias, "An instrumental variable approach to minimum-variance seismic deconvolution," *IEEE Trans. Geoscience & Remote Sensing*, **GE-23**, 778-788, 1985.

[6] Y. Goussard and G. Demoment, "Recursive deconvolution of Bernoulli-Gaussian processes using a MA representation," *IEEE Trans. Geoscience & Remote Sensing*, **GE-27**, 384-394, 1989.

[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis & Machine Intelligence*, **PAMI-6**, 721-741, 1984.

[8] J. Idier and Y. Goussard, "Markov modeling for Bayesian restoration of 2-D layered structures," *LSS Internal Report # GPI-89/06*, 1989.

[9] R. Kinderman and J.L. Snell, "Markov random fields and their applications," Amer. Math. soc., Providence, RI, 1980.

[10] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. R. Statist. Soc. B*, **36**, 192-235, 1974.

[11] J. Goutsias, "Mutually compatible Gibbs random fields," *To appear in IEEE Trans. Information Theory*, 1990.

[12] H. Derin, H. Elliott, R. Cristi and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields," *IEEE Trans. Pattern Analysis & Machine Intelligence*, **PAMI-6**, 707-720, 1984.