Lecture Notes in Statistics

Edited by J. Berger, S. Fienberg, J. Gani, K. Krickeberg, I. Olkin, and B. Singer

74

P. Barone A. Frigessi M. Piccioni (Eds.)

Stochastic Models, Statistical Methods, and Algorithms in Image Analysis
Proceedings of the Special Year on Image Analysis held in Rome, Italy, 1990

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo Hong Kong Barcelona Budapest
BAYESIAN 3-D PATH SEARCH AND ITS APPLICATIONS TO FOCUSING SEISMIC DATA

R. Azencott
École Normale Supérieure (Paris) et Université Paris-Sud,
B. Chalmond
Université Paris-Sud,
Ph. Julien
TOTAL-CFP, Geophysical Research.

Abstract

The 3D-images studied here are essential to the analysis of cubes of seismic focalisation. In the detection of geological horizons, the improvement of migration techniques requires the construction of 3D “focal” paths. We start with blurred versions of (unknown) 3D-images consisting ideally of concentrated intensity spots which tend to lie on smooth isolated 3D-paths. The blur point-spread function is spatially dependent, roughly Gaussian in shape, and directly estimated on the blurred image. On the space of admissible paths, we describe the plausibility of a path by an energy function, using thus a 3D-Markov random field model. The adjustment of this Markov field model to the image data relies on an original interactive robust parameter localization approach.

Reconstruction of the original paths is based on a maximum (a posteriori) likelihood approach, implemented by a new variant of Besag’s ICM algorithm. Applications to actual 3D-seismic data are presented.

Acknowledgments

The authors gratefully acknowledge the financial support provided to this study by TOTAL-CFP (Recherche Géophysique), who has also supplied the seismic data and the geological know-how to handle them.

1 Introduction

1.1 Seismic data and horizon detection

The problem we have studied arises in the context of geophysics, and concentrates on the analysis of 3D-cubes of focalisation (cf. Faye and Jeannot [1], Yilmaz and Chambers [2]).

One of the main concerns in seismic data analysis is the detection of “geohorizons” separating “homogeneous” layers of rock, sediments, etc. (cf. Cheng at [3]). One starts with a 2D-seismic image (cf. Fig. 1), of a geological section, which composed of hundreds seismic traces. A seismic trace, a sinusoid-like waveform, is a set of reflectors of impedance contrasts between strata. If a cycle can be correlated laterally across many seismic traces on a seismic image, it is called a “seismic horizon.” On the seismic plot shown in Fig. 1, positive half cycles are darkened to highlight seismic horizons. A difficult problem is the determination of the correspondence between geological and seismic horizons (Fig. 2). The image shown in Fig. 2(a) is the 2D-image. For the propagation of seismic waves, one needs to know the speeds at which the waves propagate within the various geological strata. This knowledge is required to compute the geometric “time” axis of the image into a “depth” axis and to construct the mapping between geological and seismic horizons. The so-called “migration” technique creates a cube of focalisation data, using the 2D-image and an estimated model for subterranean wave propagation (cf. Clearbout [4]). These focalisation techniques allow the approximate positioning of the main geological horizons, which in turn paves the way for an improved speed model. This is the principle of an iterative algorithm for horizon detection on 2D-seismic images. Section 8 gives more details on this technique. It is not concerned with the treatment of such a cube without considering the 2D-image.

1.2 Focalisation cubes

Start with a 2D-seismic image $I^0$ (Fig. 1), whose points are denoted $(x, t)$ where $x$ is depth and $t$ is the (discrete) horizontal distance between the origin and the position of seismic sensors. The focalisation data associated to $I^0$ is a computed 3D-image of generic voxels $(x, y, t)$, where $y$ denotes the “depth error”.

The focalisation intensity $I(x, y)$ at voxel $(x, y, t)$ represents essentially an energy reflecting an horizon located at the point $(X, t)$, where $X$ can be expressed as a function of $(x, t)$, once the speed model is assumed to be known. For each fixed $t$-section of this 3D-cube $I$ is called the focalisation image $I_t$ associated to the $t$-trace $t$. Each 2D-image $I_t$ is perpendicular to the image $I^0$ through the $t$-column and consists essentially of reflectors more or less diffuse spots, possibly overlapping (cf. 2.2). If the speed model and the migration technique used to compute the cube.
Figure 1
A 2D-seismic image $I^0$.

Figure 2
(a) Example of seismic horizons detected on $I^0$. (b) This localization image is perpendicular to $I^0$ through the $\hat{t}$ column. (Black points in the blob centers emphasize detected blob centers.)
perfect, each strongly contrasted seismic horizon \((x_1 t_1) \cdots (x_p t_p)\) on the basic seismic image \(I^0\) would generate a sequence of theoretical focal points \(t.f.p.\) \(P_1 \cdots P_n\) of high
intensity in the focalization cube \(I, P_j = (x_j, 0, t_j) \in I_{t_j}\). Here, we shall define a \(t.f.p.\) as an isolated luminous point on a black background. The horizon itself would simply be the broken line linking these focal points in \(I^0\). But in practical and concrete data, the energy associated to a theoretical focal point \(P_j\) is diffused over an ill-defined focal spot located within \(I_{t_j}\), and roughly centered at \(P_j = (x_j, y_j, t_j)\) with \(y_j \neq 0\). Using the speed model, the error coordinate \(y_j\), and the erroneous depth \(z_j\), the migration technique gives a simple correction that will (later) yield a corrected depth value \(z_j\) hopefully closer to the unknown \(z_j\).

### 1.3 The 3D-image analysis task

Our data will be the 3D-focalisation cube \(I\) with generic voxel \((x, y, t)\) and intensity value \(I(x, y, t)\). We want to select a finite family of high intensity sequence \(Q_1 \cdots Q_p\) of points \(Q_j = (x_j, y_j, t_j)\) in \(I\), with \(t_{j+1} \geq t_j\) for \(j = 1 \cdots p\). Each such sequence will be called a focal path, and each \(Q_j\) will be a focal point.

On our actual data, the intensity at each (unknown) \(t.f.p.\) \(P_j\) has been blurred over a focal spot located with \(I_{t_j}\) and roughly centered at \(P_j\). So the task at hand includes a combination of deblurring and selecting “continuous” paths within a 3D-cube of data. Note that the unknown blur characteristics (shape, spread, ...) vary quite a bit from point to point. Moreover, the overlap of several spots is quite frequent.

The dimensions of a typical focalisation cube being of the order of \((800, 500, 1000)\), while the number of spots fluctuates around \(20000\), it was natural to seek an automatization of the search for focal paths, as proposed in Julien et al. [6]. In this paper, we present a new algorithm to accomplish this task, using the conceptual framework of Markov Field Models. We have attempted to isolate the image analysis concepts from the one degree remote geophysics context.

### 1.4 A methodology for path search in blurred 3D-images

Start with an initial (unknown) 3D-image \(J\) consisting of the union of a finite family of fairly high intensity focal paths, on a dark background. We observed a degraded image \(I\) obtained by blurring each point of \(J\). The blur function at \((x, y, t)\) will be assumed to be Gaussian, position dependent, and concentrated over the 2D-dimensional t-sector \(I_{t_j}\) of \(I\). This flatness assumption on the blur function was strongly suggested by geophysics context. Thus \(I\) can be viewed as a sequence \(I_{t_j}, t_j = 1, 2, \cdots\) of 2D-images separately blurred. Our method extends easily to 3D-blurs. In the wide literature ab-deblurring (cf. Hummel and alt. [6], Chalmol [7], among many others), two problem are considered: identification-estimation of the blur kernel, and efficient deblurring w
known blur characteristics. Here the “blur kernels”, or “point spread functions” \(P\) can be estimated by considering (separately) each spot as the blurred image of a point source. This will be done by fitting a 2D-Gaussian distribution to the observed intensity distribution. Goodness of fit will be estimated (and optimized) as a Kullback distance. The difficulty here is that the spots are only spatially limited observations the assumed underlying Gaussians PSF. Moreover, the \(t.f.p.,\) intensity and position also unknown. We point out that this blur analysis will only be a pretreatment, since our goal here is not restoration by adequate identification of the \(t.f.p.\) and of its “physical characteristics.”

The goal is to select adequately a finite family of focal paths. Each focal path \(Q_1 \cdots Q_p\) where the \(Q_j = (x_j, y_j, t_j)\) are focal points and \(t_{j+1} \geq t_j\). Using a few qualitative notions about the “look” of good focal paths (suggested by natural geophysics assumptions), we shall define a (conditional) probability distribution \(P\) on the space of all admissible path family, quantifying, given the blurred data \(I\), the likelihood \(P(\omega)\) that an arbitrary path family \(\omega\) will be a focal path family. Actually \(P\) will be 3D-Markov field. The methodology falls thus within the Bayes/Markov approach 2D-image reconstruction launched by D. and S. Geman [8], which has since generate number of research papers (Azenecot [9], Besag [10], Bouthlom [11], Chalmol [12], Dubois and Konrad [14], Geman and alt. [15], Marroquin [16]).

To build our 3D-Markov model, we propose below (§4, 6) a practical approach based on three interactive steps: model selection, parameter estimation, validation. Our flexible methodology provides a good adequation between the model and a priori qualitative information.

Finally (§5) one has to select actual focal paths by determining good local maxima the a posteriori likelihood \(P\). This search optimization by local likelihood is accomplish by a new algorithm having similarities with Besag's ICM. Note however that the standa
version of ICM Besag [10] turns out to be seriously inefficient in our case.

In section 7, we present actual seismic data and comment the results obtained by applying the method described here. Our set of algorithms for the analysis of 3D-focalisation data has been implemented on a workstation SUN 4-370; this methodology is now currently being integrated (at TOTAL-CFP) as one step of a complete chain of seismic images analysis.

2 Mathematical model of the image analysis task

We are given a degraded 3D-image denoted by \( I_t \), indexed by the sampling grid \( G = \{(x,y,t), x = 1 \cdots n_1, y = 1 \cdots n_2, t = 1 \cdots n_3\} \). We shall study \( I_t \) as the union of the sequence of 2D-images \( I_t \) indexed by \( t \in T = \{1, \cdots, n_3\} \), where each \( I_t \) is indexed by the sampling grid \( S = \{(x,y), x = 1 \cdots n_1, y = 1 \cdots n_2\}, G = S \times T \). We denote respectively by \( s = (x,y) \) and \( g = (s,t) \) the generic pixel site in \( S \) and \( G \).

2.1 The degradation model

The unknown original 3D-image \( J \) is submitted to blurring with local characteristic dependent on the voxel site, and to additive noise with variance depending on local intensity. Any point of \( J \) with nonzero intensity will be called theoretical focal point (t.f.p.). So, \( J_t \) can be written as:

\[
J_t(s) = M_t(s) \quad \text{if } s \text{ is a t.f.p.}
\]

\[
J_t(s) = 0 \quad \text{otherwise.}
\]

Moreover, the t.f.p. are isolated points. We denote by \( u \) the generic t.f.p.

For each \( t \in T, u \in S \), let \( h^t_u : \mathbb{Z}^2 \to \mathbb{R} \) be the discretized version of a 2D-Gaussian density function with mean \( \mu^t_u \) and covariance matrix \( \Sigma^t_u \). The kernel \( h^t_u \) will be the space-dependent blurring point spread function within the image \( J_t \), such that \( h^t_u = 0 \) if \( u \) is not a t.f.p. Let \( W_t, t = 1 \cdots r \) be a sequence of (independent) 2D-white noise images with equal variance and indexed by \( s \in S \). We state now the formal degradation model (slice by slice along the t-coordinate):

\[
I_t(s) = \sum_{u \in S} J_t(u) h^t_u(s - u) + \sum_{u \in S} J_t(u) W_t(u).
\]

For instance, if the unknown image \( J \) consists of one single luminous point \( P \) with intensity \( M > 0 \) (one single t.f.p.) located at \( (u, t) \) with \( u \in S \), then the degradation model generates an observed image \( I \) concentrated on \( I_t \), with intensities lying on a perturbed 2D-Gaussian shaped surface centered at \( u + m = \mu \), with total mass \( M \), covariance matrix \( \Sigma \). The additive perturbation is the 2D-white noise \( MW_t \), with mean zero and variance proportional to \( M^2 \). This elementary degraded image will be our model for a focal spot of global intensity \( M \), carried by the 2D-image \( I_t \).

For each observed focal spot, we denote by \( (M, \mu, \Sigma) \) the unknown parameters which describe the Gaussian model of its shape: \( M \) is the total intensity mass of the focal spot, \( \mu \in \mathbb{R}^2 \) is the location of its center within the \( t \)-slice \( I_t \) containing the focal spot, and \( \Sigma \) is the 2 \times 2 covariance matrix of the Gaussian model of its intensity distribution. Hence we write

\[
\mu = u + m, \quad \Sigma = \begin{pmatrix} \alpha & \rho \\ \rho & \beta \end{pmatrix},
\]

where \( u \in S \) is the location in \( I_t \) of the unknown t.f.p. generating the spot. Note that \( \mu = u + \Sigma m \) typically be of the order of a few pixels in our focalisation data.

2.2 The path space

A path \( \gamma \) within the original image \( J \) will be the continuous broken line linking by straight line segments a sequence of t.f.p. \( (s_1, t_1), \cdots, (s_p, t_p) \), where the \( s_j \) lie in \( S \), and \( t_j \in T \).

We impose the following constraints: \( \bullet \; t_{j+1} \geq 1 + t_j \) for \( j = 1 \cdots (p - 1) \) \( \bullet \) the \( p \) focal points lying on \( \gamma \) are the \( (s_j, t_j) \).

The 3D-vectors \( (s_{j+1} - s_j, t_{j+1} - t_j) \) are called the pointers defining the path \( \gamma \).

2.3 Pointer fields

Actually \( J \) is assumed to be the union of a finite family of paths. Any such family be coded by a pointer field \( \omega \) indexed by the 3D-sampling grid \( G \). We write \( \omega_g = (b_g, c_g) \) with \( b_g \in \mathbb{Z}^2 \) and \( c_g \in \mathbb{Z} \) for each \( g \in G \). We call \( c_g \) the width of the pointer. More exactly we impose for all \( g \in G \),

\[
0 \leq c_g \leq C \quad \text{where } C \text{ is a fixed bound}
\]

\[
c_g = 0 \quad \text{if and only if } \omega_g = 0
\]
(1.3) \[ J_{g} = 0 \implies \omega_{g} = 0. \]

We call \( \Omega \) the set of all pointer fields verifying (1). The associated family of paths is then constructed recursively (along the \( t \)-coordinate) using the following basic rules:

1. If \( g \) is a focal point lying on a path \( \gamma \), and \( \omega_{g} \neq 0 \), then \((g + \omega_{g})\) is also a focal point in \( \gamma \), namely the successor of \( g \) on \( \gamma \).
2. If \( J_{g} \neq 0 \) and if there is no \( g' \) in \( C \), such that \( \{g' + \omega_{g'} = g \land \omega_{g'} 
eq 0\} \), then \( g \) is the initial point of a path.

In fact we restrict further our families of paths (and hence our pointer fields) by forbidding extreme closeness of paths. More precisely, fixing a rectangular window \( R \) in \( \mathbb{Z}^{2} \), of size \((n, m)\), centered at zero, we impose for any two focal points \((s, l)\) and \((s', l)\) lying in the same \( t \)-slice the restriction \( s' \notin s + R \). We denote \( \Omega_{R} \) the set of pointer fields \( \omega \in \Omega \) which verify this restriction.

3 Blur estimation and pre-smoothing

For each \( t \)-slice \( I_{t} \) of our observed 3D-image, we set up an iterative estimation of the blur parameters, and use these estimates for a pre-smoothing of \( I_{t} \). We have used the following empirical guess: when "noiseless" blurring spots overlap, the local maxima of the resulting distribution of intensities give a good first approximation for the location of the centers of these spots.

3.1 Blur estimation

We begin by selecting all the local maxima of the observed intensity distribution \( s \rightarrow I_{t}(s) \) on slice \( I_{t} \). Here, the notion of local maximum is the usual one, relative to a spatial window of fixed size centered at the local maximum. This size is fixed by the user of the algorithm. All such local maxima will be called "rough spot centers". Since neighbouring spots overlap partially, we decided to observe each potential spot within a sequence of increasing windows \( D_{1}, D_{2}, \cdots \) centered at the spot center. For each such rough spot center \( u \) in \( S \), and each such window \( D \), we perform a local estimate of the parameters modelling the spot's shape, and test the quality of fit, according to the following procedure.

Call \( f(s) = M h(s) \) the unknown distribution of intensities for the spot centered at \( u \), where \( h \) is the discretized version of a Gaussian density with mean matrix \( \mu \) and covariance matrix \( \Sigma \). Let

\[
M_D = \sum_{s \in D} M h(s), \quad m_D = \sum_{s \in D} I_t(s).
\]

We now adjust the unknown parameters \( \theta = (M, \mu, \Sigma) \) to minimize the Kullback distance \( d \) between the two probability distributions on \( D \) defined by \( M^2 D f \) and \( M^2 D I_t \). Let \( u \) be the point of interest (the exact pixel location). Recall that if \( f_1, f_2 \) are two probability distributions on \( D \), the Kullback distance

\[
d(f_1, f_2) = \max \{K(f_1, f_2), K(f_2, f_1)\}
\]

where \( K(f_1, f_2) = \sum_{s \in D} f_1(s) \log(f_2(s)/f_1(s)) \). We refer to Basseville [17], Kullback [18] for standard statistical consideration justifying its use in parameter estimation problems.

Call then \( \hat{\theta} \) this estimate

\[
\hat{\theta} = \text{Arg min } d\left(\frac{f}{M_D}, \frac{I_t}{m_D}\right)
\]

and \( \hat{M}_D \) the corresponding estimate of \( M_D \). We accept this estimate of \( \theta \) on \( D \), if

\[
d\left(\frac{\hat{f}}{\hat{M}_D}, \frac{I_t}{m_D}\right) < d_0 \quad \text{and} \quad \hat{M}_D > M_0
\]

where \( d_0 \) and \( M_0 \) are empirical thresholds. When (2) is verified, we stop increasing the size of our estimation window \( D \), and accept the current estimate \( \hat{\theta} \) as our final one (for the particular spot centered at \( u \)). To construct \( d_0 \), we compute the actual Kullback distance between two Gaussian distributions for which the covariance matrix are submitted to given distortions, using a classical formula of Kullback [18].

Now if these criterions are not satisfied, we increase the window \( D \) and compute the new Kullback estimate etc. Obvious bounds on the size of \( D \) are imposed by the locality of the neighbouring local minimum, as well as by an \( a \) priori inspection of the data by the user. If the window \( D \) reaches this maximal size without a successful verification of criteria (2), we decide that no "real" spot is centered at the currently studied local maximum \( u \).
3.2 Pre-smoothing

Call SPOT; the subset of the sampling grid $S$ constituted by the local minima $u \in S$ which survived the previously described estimation and pruning process on the image $I_t$. Of course, here $t$ is fixed and SPOT$; is relative to this particular $t$-slice $I_t$. We now decide to truncate each Gaussian density $h$, with mean $\mu$, covariance matrix $\Sigma = \begin{pmatrix} \alpha & \rho \\ \rho & \beta \end{pmatrix}$ by restricting $h$ to a window centered at $\mu$, and of size $(2\sqrt{\alpha}, 2\sqrt{\beta})$. Denote by $\tilde{h}$ the truncation of $h$. To each surviving local maximum $u$, we have thus attached a truncated modeled spot, having the intensity distribution $\tilde{M}^u \tilde{h}^u$, with an obvious indexation by $u$ of the parameters $M, h$. We now replace our observed $t$-slice $I_t$ by the following smoothed version

$$\tilde{I}_t = \max_{u \in \text{SPOT}_t} \tilde{M}^u \tilde{h}^u.$$  

We have essentially pruned out the “weak” spots and enhanced the visualisation of the “strong” spots by forcing gaps between neighbouring strong spots. Recall that this whole operation has to be repeated separately for each $t$-slice, and is of course $t$-dependent.

For each $g = (s, t) \in G$, such that $\tilde{I}_t(s)$ is nonzero, we have singled out a unique spot active at $g$; we shall denote by $\theta(g)$ its estimated parameters, and we shall say that $g$ lies within the spot centered at $\mu_t(s)$, with parameters $\theta_t(s) = (M_t(s), \mu_t(s), v_t(s))$, that we denote more briefly by $\theta_g = (M_g, \mu_g, v_g)$.

The smoothed 3D-image intensities $\tilde{I}_t(s)$ where $s \in S$, $t \in T$ will be recoded more simply by $F_g = \tilde{I}_t(s)$ where $g = (s, t) \in G$. We call FOCS the subset of the sampling $S$ constituted by the estimated exact centers of the surviving spots in $I_t$. Each such center will be called a potential focal point of the 3D-image $I$.

3.3 An example of pre-smoothing

Using actual focalisation data provided by a well studied benchmark 3D-cube, we present the results of such a pre-smoothing in illustrations 3a, b, c, d. Sketches 3a and 3c represent the intensity distribution $I_t$ for a given focalisation plane (or $t$-slice), using standard perspective and intensity level curves respectively. Sketches 3c, 3d represent the smoothed $t$-slice $\tilde{I}_t$, using the same convention.

For some spots, the pre-smoothing modelizes only the central peak of the spot. This is quite enough in this context since we only want to locate the focal points. Hence
there would have been little to gain if we had tried to build a more precise smoothing: modelizing carefully each spot by a sum of Gaussian intensity distributions, since a rougher smoothing preserves almost all significant focal points anyway.

4 The Markov field model

We now start with the collection \(\text{FOC} = \{\text{FOC}_t, t \in T\}\) of focal point candidates obtained by pre-smoothing, the associated estimated spot parameter field \(\theta = (\theta_g, g \in G)\) and the smoothed 3D-image \(P = (P_g, g \in G)\). To each possible pointer field \(\omega \in \Omega_\mathcal{F}\), we are going to attach a plausibility grade \([-U(\omega)]\) which will be high when \(\omega\) is a quite plausible pointer field given \(F, \theta\), and low in the converse situation. The function \(U\) which we call the (conditional) energy of \(\omega\) given \(F, \theta\), will be a linear combination of four particular energies \(U_i\)

\[
U = \sum_{i=1}^{4} \beta_i U_i
\]

where the \(\beta_i > 0\) are unknown parameters.

Given a pointer field \(\omega\), we define for all \(g\) in \(G\) the indicator function \(\delta\) of its support by

\[
\delta_g = 1_{\{\omega \neq \theta_g\}}.
\]

The first energy \(U_1\) will force pointer fields to build focal paths of high global intensity:

\[
U_1(\omega) = -\sum_{g \in G} \delta_g \tau_1(P_g)
\]

where \(\tau_1 : \mathbb{R}^+ \rightarrow \mathbb{R}\) is an increasing scale transformation to be optimally selected later.

The second energy \(U_2\) will force homogeneity of the shapes of focal spots along for paths, by minimizing the average Kullback distortion \(d\) between successive Gaussian shapes on paths

\[
U_2(\omega) = \sum_{g \in G} \delta_g \tau_2 [d(\theta_g, \theta_{g+})]
\]

where \(g+ = g + \omega_g\) is the successor of \(g\), \(\tau_2\) is an increasing scale transformation to be selected later, and \(d\) is the Kullback distance between the Gaussian models \(\theta_g, \theta_{g+}\) of the spots centered at \(\mu_g\) and \(\mu_{g+}\).
The third energy $U_3$ tends to regularize the change of "speed" along focal paths:

$$U_3(\omega) = \sum_{g \in G} \delta_g \delta_{\eta^+} \tau_3 \|\ell_g(\omega) - \ell_{\eta^+}(\omega)\|$$

where $\tau_3$ is an increasing scale transformation to be selected later, and where for any pointer field $\omega$, the "speed" at $g$ is defined by $\ell_g(\omega) = b_g/c_g$ whenever $c_g = 0$ with $b_g \in \mathbb{Z}^+, c_g \in \mathbb{Z}^+$. The norm $\|\|$ on $\mathbb{Z}^2$ is the $L_1$-norm.

The fourth energy $U_4$ forces the focal paths to be rather long and sparse:

$$U_4(\omega) = \sum_{g \in G} \delta_g (1 - \delta_{\eta^+}) - \tau_4 \sum_{g \in G} \delta_g \delta_{\eta^+}$$

where the first sum is simply the total number of focal paths created by the pointer field $\omega$. The second sum in $U_4$ is the total length of all these focal paths and $\tau_4 \in ]0, 1[.$

We point out that $U_3$ and $U_4$ do not depend on the smoothed data $F$ nor on the estimated shape parameters $\theta$. On the other side, $\beta_1 U_1 + \beta_2 U_2$ represents the constraints imposed on the focal paths by the output $(F, \theta)$ of our data smoothing.

The parameters to be estimated below (in §6) to finalize the specification of $U$ are the weights $\beta = (\beta_1 \cdots \beta_6)$ and the scaling factors $\tau = (\tau_1 \cdots \tau_4)$.

We adopt a now well grounded probabilistic setup by imposing on the pointer fields $\omega$ in $\Omega_R$ the Markov field distribution associated to the (conditional) energy $U$ given $F, \theta,$ by the Gibbs formula:

$$P(\omega) = \frac{1}{Z} \exp [-U(\omega)] \quad \text{with} \quad Z = \sum_{\omega \in \Omega_R} \exp [-U(\omega)].$$

For all $g \in G$, let $N_g$ be the parallelepipedic 3D-window in $G$, centered at $g$, of size $(n, m, 2C+1)$ and deprived of its center $g$. Recall that in §2.4 we introduced the window $K$ of size $(n, m)$ and the bound $C$ on pointers widths. As is easily seen $N_g$ is the set of sites in $G$ which are in interaction with $g$, either through the actual expression of $U$ or due to the implicit constraints of §2.4, imposed on widths for pointers (1) and closeenss for paths given by the restriction to $\Omega_R$. Actually it is easily proved that, given $F, \theta$, we have

$$P(\omega_\gamma | \omega_{N_\gamma}) = P(\omega_\gamma | \omega_{G - \gamma})$$

where $\omega_A = \{\omega_{g'}, g' \in A\}$ and hence $P$ is a Markov field with respect to the neighborhood system $(N_g, g \in G)$ (cf. Besag [10]).

5 Estimation of focal paths

We assumed for the moment that after pre-smoothing of our focalisation data, the energy parameters $\beta_i, \tau_i, i = 1 \cdots 4$ have been adjusted as described in section §6 below. The question tackled in the current section is the search of a good feasible estimate $\hat{\omega}$ for the unknown actual pointer field.

The standard maximum likelihood estimator maximizing $P(\omega)$ in $\omega \in \Omega_R$ is unrealistic, given the huge size of the data. We have introduced instead a local optimization algorithm which resembles the ICM estimator introduced by Besag [10]. For comparison purposes, let us recall the standard ICM algorithm: one starts with an initial estimate $\omega^1$, and then $\omega^k$ is iteratively modified by refreshing one site at a time in a given arbitrary periodic order. Call $g_{n+1}$ the site modified at the time $n + 1$, and $\omega^n$ the existing pointer field at time $n$. Then the standard ICM selects $\omega^{n+1}$ as follows:

$$\omega^{n+1} = \omega^n \quad \text{for} \quad g \neq g_{n+1}$$

$$\omega^{n+1} = \text{Arg max}_{\omega_g} P(\omega_g | \omega_{N_g}) \quad \text{for} \quad g = g_{n+1}.$$}

This algorithm converges after a few complete sweeps of the 3D-grid $G$, and the li configuration $\omega^\infty$ is a local maximum of $P(\omega)$. How good this estimate $\omega^\infty$ can strongly depends on the context.

In our case, the initial configuration is defined as follows: for every $g$ in FOC (potential focal points), we define $\omega_g = \bar{g} - g$ where:

$$\bar{g} = \text{Arg min}_{g' \in N_g} \{||g' - g||, g' \text{ posterior to } g\}.$$}

By convention $g'$ is posterior to $g$ if the $t$-coordinates $t', t$ of $g', g$ verify $t' > t$. $g \notin$ FOC, we let $\omega_g = 0$. However the standard ICM is not efficient enough here: indeed due to the constraint preventing extreme closeeness of paths, the ICM algorithm (7) never modify the support of $\omega^1$, and hence will not prune the initial $\omega^1$, although it reorganize it a bit.

We thus prefer to rely on a local optimization algorithm which we call BLOCK-I and which we now describe. Fix a current pointer field $\omega$. For every $g \in G$, such that $g$ lies on a path $\gamma$ defined by $\omega$, and such that $g$ is not the starting point of $\gamma,$
let $g_-$ be the predecessor of $g$, so that $g = g_- + \omega_{g_-}$. Fix $n' < n$, $m' < m$, and let $g = (s, t) \in G = S \times T$, call $\Delta_g$ the 2D-window (within $S \times I$) centered at $g$, and with size $(n', m')$. If $g$ lies on a path defined by $\omega$, define:

$$B_g = \{g_-\} \cup \Delta_g \quad \text{if } g_- \text{ exists}$$

$$B_g = \Delta_g \quad \text{if } g \text{ is a starting point.}$$

Call $NB_g$ the union of all $N_{g'}$ for $g' \in B_g$. In our BLOCK-ICM algorithm, we fix arbitrary infinite periodic sequence $g_n$, visiting all sites $g \in G$. At time $n$, the current configuration $\omega^n$ will be modified only on the sites which belong to $B_{g_n}$. Let then $\omega^n = \omega$ and $g_n = g$ to simplify the notations and select for the modified block configuration $\omega$ replacing $\omega_{B_g}$ as the block configuration realizing:

$$\max_{\omega_{B_g}} P[\omega'_{B_g} | \omega] = P[\omega'_{B_g} | \omega_{B_g}].$$

Then define $\omega^{n+1}$ by

$$\omega^{n+1}_B = \omega_B, \quad \omega^{n+1}_{\Delta_g} = \omega^{n}_{\Delta_g}.$$

Due to the constraint on mutual distances between paths and on the size of the wind $\Delta$, then whenever $g_-$ exists, the possible paths created by a single block of pointers $\omega$ can intersect $\Delta_g$ at only one (arbitrary) point $g' \in \Delta_g$, which restricts immediately the list of acceptable $\omega'_{B_g}$. Sketch 4 above represents, in 2D-projection, three instances of local path configurations branching out of a tentative particular choice for $g'$ in $\Delta_g$.

6 Actual Markov field modelization

6.1 Methodology

The first step is the selection of an energy $U$, which was described in section §4. The second step is the adjustment of the parameters of this particular energy, taking account of the specifications of performance for the estimator $\hat{\omega}$ (of the unknown point field $\omega$) produced by our BLOCK-ICM algorithm.

This has been achieved by the following interactive scheme:
We point out that at the end of this interactive process, the parameters \( \beta \) and \( r \) are selected once and for all, and used to handle incoming new focalization data. The parameters left to the choice of the actual user of the algorithm are the integers \( (n, n) \) which determine the dimensions of \( N_x \), and may be selected according to the "look" data.

Initially, we tried linear scale changes for \( \tau_2 \tau_3 \tau_4 \) and a logarithm scale change for \( \tau_1 \). We had to give up this combination in the validation study and our final (and efficient) choice has been to use only linear scale changes for \( \tau_1 \cdots \tau_4 \).

6.2 Determination of the weights \( \beta \)

Assume the scale parameters \( \tau \) are already estimated. We shall constrain \( \beta = (\beta_1 \cdots \beta_4) \) by imposing \textit{a priori} bounds on conditional probabilities \( P(\omega | \omega_{N_x}) \), using "extreme configurations" for which such probabilities should clearly be quite close to 0 or 1. This yields for \( \beta \) a system of linear equations which are easily solved. This approach is in spirit of the "qualitative boxes" approach sketched by Azencott [9]; we also refer on point to Possolo [20] and Derin [21].

Up to a normalizing constant, the local specifications are given by:

\[
P(\omega | \omega_{N_x}) \propto \exp - \nabla_\omega(\omega)
\]

\[
\nabla_\omega(\omega) = U(\omega) - U(\bar{\omega})
\]

\[
\bar{\omega}_{G^{-}} = \omega_{G^{-}} \text{ and } \bar{\omega} = 0.
\]

For a given pointer field \( \omega \), let \( \xi_\rho(\omega) = 1 \) if \( \rho \) is a focal point of \( \omega \), having a predecessor \( g_- \), and let \( \xi_\rho(\omega) = 0 \) otherwise. Rewrite the energies \( U_1, U_2, U_3, U_4 \) as \( U_i = \sum_{g \in \mathcal{C}} \) with obvious local components \( U_{i, g} \) specified by formulas (3-6) in section §4. Then immediate computation yields:

\[
\nabla_\omega = \sum_{i=1}^{4} \beta_i Z_{i, g} \quad \text{with}
\]

\[
Z_{1, g} = U_{1, g}
\]

\[
Z_{2, g} = U_{2, g}
\]

\[
Z_{3, g} = \xi_\rho U_{3, g^-} + U_{3, g}
\]

\[
Z_{4, g} = \xi_\rho U_{4, g^-} + U_{4, g}.
\]
Locally, \( Z_1 \ldots Z_4 \) quantify respectively the path intensity, the non homogeneity of the spots, the speed irregularity, and the presence/absence of path break. We may thus rewrite (9) as:

\[
\sum_{i=1}^{4} \beta_i Z_i(\omega) = - \log \frac{P(\omega \mid \omega_{N_x})}{P(\omega \mid \omega_{N_x} = 0)}. 
\]

Let us now select a finite set ELC of "extreme local configurations". Each such configuration selects an arbitrary set of sites \((y \cup N_y)\), and specifies arbitrarily on this set the values taken by the fields \(F, \theta, \omega\) (which represent resp. the smoothed image, the shape parameters, the pointer field).

Note that for each extreme local configuration \(c \in \text{ELC}\), the \(Z_i(\omega)\) are easily computed and yield numbers \(Z_i\), \(i = 1 \ldots 4\). Moreover, the conditional probabilities appearing on the right hand side of (10) will be a priori set at values:

\[
p^c = P(\omega \mid \omega_{N_x}), \quad p^c_0 = P(\omega \mid \omega_{N_x} = 0),
\]

which will be very close to zero or one. Of course, this requires an adequate choice of the extreme configuration \(c\). We now have a linear system:

\[
\sum_{i=1}^{4} \beta_i Z_i^c = - \log \frac{p^c}{p^c_0}
\]

of cardinal equal to \(\text{Card} (\text{ELC}) = 4\), to have more equations than unknown. The determination of \(\beta\) is then obtained by standard least squares resolution of the system (11).

6.3 Examples of extreme local configurations

We have actually used 12 such configurations. The following table lists the values of the \(Z_i^c\) and the \(p^c\). Actual values of the \(Z_i^c\) have been replaced by qualitative evaluations \(H = \text{High}\) or \(L = \text{Low}\) to enhance readability. On Fig.5, we present the first two extreme configurations.

<table>
<thead>
<tr>
<th>configuration</th>
<th>(Z_1^c) intensity</th>
<th>(1/Z_2^c) homogeneity</th>
<th>(1/Z_3^c) regularity</th>
<th>(Z_4^c) path break</th>
<th>(p^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>N (\bar{\alpha})</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>N (\bar{\alpha})</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>N (\bar{\alpha})</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>N (\bar{\alpha})</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>N (\bar{\alpha})</td>
<td>0.9</td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>N (\bar{\alpha})</td>
<td>0.1</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>N (\bar{\alpha})</td>
<td>0.1</td>
</tr>
<tr>
<td>8</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>N (\bar{\alpha})</td>
<td>0.01</td>
</tr>
<tr>
<td>9</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>Yes (\bar{\alpha})</td>
<td>0.1</td>
</tr>
<tr>
<td>10</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>Yes (\bar{\alpha})</td>
<td>0.01</td>
</tr>
<tr>
<td>11</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>Yes (\bar{\alpha})</td>
<td>0.99</td>
</tr>
<tr>
<td>12</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>Yes (\bar{\alpha})</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 2

Extreme configurations \((H = \text{High}, L = \text{Low})\)

6.4 Determination of the scale parameters

Start with an "arbitrary" guess for the scale parameters \(\tau\). Apply the method just described in 6.2 to estimate \(\beta\) by \(\hat{\beta}\). The overdetermined linear system (11) above yields for each extreme local configuration \(c\) in ELC, a residual error term \(\varepsilon_c\) given by

\[
\varepsilon_c = - \log \left( \frac{p^c}{p^c_0} \right) - \sum_{i=1}^{4} \beta_i Z_i^c.
\]

We want these residuals \(\varepsilon_c\), \(c \in \text{ELC}\) to be small. If not we attempt to modify the scale parameters \(\tau\). This is done by inspecting the large residuals \(\varepsilon_c\) and more precisely of their corresponding components \(\beta_i Z_i^c\). This suggests immediate qualitative modifications of the \(Z_i^c\), which are achieved by changing the \(\tau\).

This methodology, which can easily be formalized, is analogous to interactive modification in multiple linear regression (Draper-Smith [22]). In our context, we complete this examination of large residuals by one last validation step. Whenever all residuals \(\varepsilon_c\) are small, we actually check how the \(\hat{\beta}\) estimate of \(\beta\) performs on another control set of local configurations, using a similar point of view. Thus we are validating our linear system by an adjoined extra set of linear equations of the same type, with less tightly specified right hand sides.
7 Applications to actual 3D-focalisation data

The data are real seismic data produced by the algorithmic chain provided in the MIGPACK software package [1]. The actual 3D-image has dimensions (800, 160, 400). We observe roughly 20 spots per t-slice \( I_t \). Our selected dimensions for the \( N_f \) window are \( n = 17, m = 8, C = 8 \), and the BLOCK-ICM algorithm uses \( \Delta_y \) windows of size \( n' = 9, m' = 5 \).

On sketch 6a, we represent the projections on the \((x, t)\) plane of the focal paths detected by the algorithm developed by Julien et al. [5]. Clearly, the paths appear to be noisy, essentially due to irregular local slopes. Sketch 6b on the other hand represents the projections on the \((x, t)\) plane of the focal paths detected by the Markov field approach developed in this paper. We notice an evident smoothing of the paths, and the erasing of most of the noise. Let us emphasize that the initial configuration \( \omega^1 \), used by the BLOCK-ICM algorithm, is very close to the configuration of Fig. 6a. Computations using the BLOCK-ICM estimation have required 6 complete sweeps of the 3D-image, which amounted to about one hour of CPU time on a CONVEX. The total number of detected focal paths is of the order of 200. Fig. 2a depicts the focal paths obtained with another parameters \( n, m, C \).

The second advantage of this refined automatic picking is that it provides an interpretable image directly comparable to the seismic section shown on Fig. 1. One can notice that the chief horizons, in terms of focused energy, have been determined correctly. A lot of qualitative and pseudo-quantitative information has been calculated as well. This complementary knowledge refers to the shape and the amplitude of the spots, the length of the paths and better describes the seismic data and further the geological subsurface. This kind of information was of course already contained in the raw dataset but it was not explicitly computed.

A very important characteristic of such an automatic picking is that it is continuous and, as a consequence, it leads to the dip of focusing paths. This new knowledge is of great interest. It allowed us to derive a new equation for computing the velocity correction in the case of dipping layers. Until now, the velocity updating was only carried out in the simple horizontal case. The new velocity field varies continuously, laterally and in depth. In this way, non-linear variations of the velocity field may be described within a layer.

![Figure 6](image)

*Figure 6* Projection on the \((x, t)\) plane of the focal paths. (a) detected by the standard algorithm of [5]. (b) detected by our Markov field approach.
The last topic which may be pointed out is the problem of multiples (Wiggins [23]; Julien and Raoult [24]). Multiple reflections are composed of energy which has been upward reflected more than once. Multiples constitute a coherent noise which often presents a lower apparent velocity than the one of primary events. Presently, no general method is able to suppress all kinds of multiples. If focusing paths are projected and displayed on the plane of the seismic section, at their right (corrected) locations with a color depending on the velocity model, multiples and primary events which present a differential velocity may be easily discriminated just by looking at the colors of the paths.

8 More details on the seismic problem

The aim of seismic processing is to help the geologist to gain a better (and easier) understanding of the subsurface from seismic data study. Seismic sources located at the surface of the earth generate a wavefield which is reflected by the interfaces between layers of different kinds. The reflected wavefield is recorded at the surface, as a function of time, by a network of receivers. These seismic data provide an image of the subsurface reflectivity (see Fig. 1).

In the seismic processing approach, a very important stage is probably 2D migration [4], a process which reduces the distance between a 2D seismic section and the corresponding geological section. The problems which are solved or reduced by migration are those connected with wavefield propagation effects. If, for instance, a seismic (stack) section is migrated, time diffraction hyperbolas (due to diffracting points such as extremities of technical faults) disappear and geological horizons are then located at their right positions on the processed result.

Pre-stack depth migration, which processes directly the raw dataset recorded on the field, is the ultimate in structural 2D processing. It is a good way to produce a correct image of a complex subsurface in the presence of strong lateral velocity variations. The shot-geophone pre-stack depth migration (Denelle et al. [25,26]) involves two stages. First, the pressure field is downward extrapolated. Given the recorded wavefield at the surface and the acoustic wave equation, the pressure is computed for all depths by using a finite differences algorithm (Clearbout and Doherty [27]; Schultz and Sherwood [28]). Each elementary downward continuation consists in applying two operators: the phase-shift operator which corrects the pressure field for the vertical wave propagation an focusing operator which focuses the diffracted energy to the scattering points. In second stage (imaging step), the migrated result is obtained by mapping the wave at a propagation time equal to zero. Really, when the propagation is null, the source and the receiver are located at the same place. The energy that is recorded by receiver is exactly the energy emitted by the source times the local reflectivity of the medium. Hence, at a null time, the downward continued recorded wavefield is a direct measurement of the local reflectivity of the medium.

The shot-geophone prestack depth migration proves to be effective and affords from time recorded seismic data, it leads to a depth image. That means that it requires prior knowledge of the propagation velocities i.e. the shape and interval velocities for each layer.

Now, all authors agree to say that the most important input information for migration procedures is the velocity distribution. Generally, this information is approximately known and errors in the migration output due to errors in the velocity input occur in a lot of practical applications. So the following paradox has to be performed: a correct pre-stack depth migration requires that the main features of the resulting seismic section in terms of layer geometry plus the interval velocities w each layer be input as processing parameters. Faye and Jeannot [1] proposed a simple procedure to solve this problem. This procedure refines the velocities using depth focus analyses. Such an approach had already been introduced by Doherty and Clearbout and by Yilmaz and Chambers [2] for time migration. The influence of velocity errors on the focusing aspects of migration has been theoretically studied by De Vries [30]. More recently, some new developments have been derived by Julien et al. [5] who proposed a depth continuous velocity analysis method based on an automatic picking of all the relative maxima of the whole focusing 3D block computed over an entire seismic line. MacKay and Abma [31] proposed to obtain directly a refined prestack migration image from the focusing block.

The general principle of the focusing analyses is the following. When the velocity is exact, all the diffracted energy focuses, for a null propagation time (imaging criterion) at the location of the scattering point. If the migration velocity is greater than the exact velocity, the energy focuses at a greater rate than expected, at a positive time.
each shot point, a focusing plane is the result of the extrapolation of seismic data at every depth and every time for the zero offset i.e. for a geophone located at the source. Each interface leads to a focusing point. The location of the points on a focusing plane is related to the used velocity model. If a focusing point is not located on the imaging curve (corresponding to a null propagation time), the velocity model is changed to bring this point on this curve. In fact, due to numerical problems, the focusing pick is not exactly a point but is rather a spot with a certain length, width and amplitude. To provide the explorationist with a good tool for imaging structures in such hard conditions, a full prestack depth migration program combined with an interactive focusing analyses based software, has been developed and made operational in an industrial environment in house (Julien et al. [32]).

Until now, due to the manual picking of relative maxima, few focusing planes were processed in a seismic line. Hence, it was difficult to achieve geological coherency from one focusing plane to the next one. Too simple velocity models were obtained compared to what was expected from the prestack depth migration ability to take into account large velocity variations. If we process all the focusing planes (a focusing plane for each shot point), we get a 3D focusing block. With such a volume, it becomes possible to carry out a continuous velocity analysis. For a real case, this data volume is very large (about four hundred megabytes) and there are several thousands of spots. The goal of this presented paper is to present a refined automatic picking. This picking leads to focusing paths of relative maxima. These paths represent geological interfaces and the associated depth errors. Generally, these paths must be smooth and continuous. With such an automatic picking, lateral velocity variations will be faster and better described.

References


