GEOPHYSICS, VOL. 65, NO. 2 (MARCH-APRIL 2000); P. 656–667, 7 FIGS.

Recognition and reconstruction of coherent energy with application to deep seismic reflection data

Mirko van der Baan* and Anne Paul‡

ABSTRACT

Reflections in deep seismic reflection data tend to be visible on only a limited number of traces in a common midpoint gather. To prevent stack degeneration, any noncoherent reflection energy has to be removed.

In this paper, a standard classification technique in remote sensing is presented to enhance data quality. It consists of a recognition technique to detect and extract coherent energy in both common shot gathers and final stacks. This technique uses the statistics of a picked seismic phase to obtain the likelihood distribution of its presence. Multiplication of this likelihood distribution with the original data results in a "cleaned up" section. Application of the technique to data from a deep seismic reflection experiment enhanced the visibility of all reflectors considerably.

Because the recognition technique cannot produce an estimate of "missing" data, it is extended with a reconstruction method. Two methods are proposed: application of semblance weighted local slant stacks after recognition, and direct recognition in the linear τ-p domain. In both cases, the power of the stacking process to increase the signal-to-noise ratio is combined with the direct selection of only specific seismic phases. The joint application of recognition and reconstruction resulted in data images which showed reflectors more clearly than application of a single technique.

INTRODUCTION

As in exploration seismics, remote sensing or teledetection suffers from the fact that the vast amounts of data gathered require intensive manual labor before they are processed and interpreted (Richards, 1993). In addition, the obtained data may be distorted and may not always have the desired high quality. Therefore, attention in remote sensing has been focussed for a long time on the questions of how to reduce processing and interpretation time and how to enhance the obtained images using semiautomatic or interactive techniques. For example, a frequently encountered problem in remote sensing is the need to classify each pixel of an image into several classes to be monitored (e.g., crop growth or damages in forests due to summer fires). For this specific task, very often recognition and classification techniques are used.

In this paper, a standard classification technique in remote sensing is presented, and it is shown how this technique can be used to reduce the problem of low signal-to-noise (S/N) ratios in deep seismic reflection data. In conventional processing, two phenomena notably decrease stack quality. First, reflections tend to be visible on only a limited number of adjacent traces, and second, because of the large offsets (half-offsets of several kilometers are not unusual), moveouts very often exhibit static and even dynamic undulations on both large and small scales caused by the presence of overlying heterogeneities. Stack degeneration can readily be reduced by removing noncoherent energy along the stacking trajectories, thereby attacking the first problem. The second problem is beyond the scope of this paper.

Many different methods to increase the S/N ratio are available in exploration geophysics. Among these are the skeletonization method (Le and Nyland, 1990; Lu and Cheng, 1990), the statistical method of Hansen et al. (1988), methods using local τ-p analyses (Harlan et al., 1984; Kong et al., 1985), and spatial-prediction filtering (Hornbostel, 1991; Abma and Claerbout, 1995).

The simplest method would be to calculate a variable describing the local coherence between traces as is done, for example, in attribute maps such as the coherence cube (Bahorich and Fermer, 1995; Marfurt et al., 1998). A multiplication of the local coherence with the original traces emphasizes the coherent parts. However, the skeletonization method (Le and Nyland, 1990; Lu and Cheng, 1990) probably represents a

Manuscript received by the Editor February 18, 1998; revised manuscript received June 7, 1999.
*Formerly Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier, Grenoble, France; presently School of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK. E-mail: mvdbaan@earth.leeds.ac.uk.
‡Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier, B.P. 53, 38041 Grenoble Cedex 9, France. E-mail: apaul@ujf-grenoble.fr.

better alternative. This method reduces a seismic section to a line drawing by correlating waveforms of neighboring traces, thereby automatically tracking seismic reflectors. However, for these methods to work properly, a good S/N ratio must already be present in the data.

The method of Hansen et al. (1988) scans subsections over a range of slownesses and intercept times, and simply counts the number of positive and negative amplitudes in a line segment. Only if either number exceeds a certain threshold is the examined line segment passed as signal. Moreover, the method can assert both the probability that a line is passed containing only noise and the total number of noise-only segments.

The methods actually computing the local τ-p transforms use it either as a focusing measure (Harlan et al., 1984) or in combination with semblance (Kong et al., 1985). The first method uses the τ-p transform to obtain the desired amplitude distributions of noise and signals needed for the recognition and extraction of samples containing the highest percentage of signal, whereas in the second method semblance-weighted slant stacks are used to emphasize and extract laterally coherent signals.

Finally, spatial-prediction filtering is widely used in industry to improve data quality. This approach replaces samples by their predicted values based on information contained in adjacent traces. The laterally coherent signal energy can be extracted using two different techniques; least-mean-square adaptive filtering in the t-x domain (Hornbostel, 1991) or a Wiener filter in either the f-x or t-x domains (Hornbostel, 1991; Abma and Claerbout, 1995).

Unfortunately, all of the above-mentioned methods emphasize any laterally coherent energy, including coherent noise. Hence, even if such noise has a signature which is completely different from the desired reflections, it will be extracted. Therefore, a technique extracting preselected phases only is needed.

Moreover, most of the above-described methods assume that reflectors are at least locally linear (i.e., represented by a line in 2-D or a plane in 3-D data volumes), whereas it is well known that this is rarely the case in real data. Thus, in addition, a method is required which is still able to distinguish signal from noise, but does not impose these restrictions.

With these aims in mind, a standard supervised classification technique from remote sensing will be introduced. Supervised classification is done using some specific attributes describing a local part of the data (i.e., its signature) and consists basically of three stages (Richards, 1993): selection, training or learning, and classification. In the first stage the different classes are defined and a training set is selected for each class. Then, the characteristics of each class are estimated (the training stage). Finally, classification or clustering takes place. Each class represents a different category or type of signal (e.g., burnt and nonburnt woods or urban soil and water in remote sensing, or the different types of signals in seismic data: reflections, diffractions, ground roll, and noise). The characteristics of the training set are defined as the averages and variances of the attributes. With these, the statistical likelihoods of the presence of each class can be calculated, and the data can be clustered by means of a simple comparison. After clustering (i.e., classification), the desired signals are extracted, thereby enhancing the visibility of that specific signal.

What attributes best discriminate between different signals? If, in a first simplification, data are assumed to be composed of ground roll, reflections, and noise, then discrimination can be done using frequency content. Since reflections and diffractions differ only by a phase shift (Trorey, 1970), the most logical signature would be the local amplitude spectrum. Hence, the use of the local amplitude spectrum allows the classification technique to recognize signals displaying phase perturbations including, for example, diffractions. Most other techniques, however, demand phase consistency notably along the stacking trajectories, thereby excluding the detection of diffractions. This additional ability to detect diffracted signals sets the recognition method even further apart from the other methods.

On the other hand, the methods incorporating adjacent traces have the possibility to reconstruct "missing" energy, thereby improving data quality over mere classification. Fortunately, it is not difficult to adapt the recognition technique to estimate missing data. We propose two methods: (1) the application of the reconstruction method of Kong et al. (1985) using semblance-weighted slant stacks after recognition, and (2) direct application of the recognition technique in the τ-p domain. Unfortunately, these methods again require the assumption of locally linear events.

First, we describe the recognition method and show how it can be used to extract the "coherent" features in a seismic section. Then, the topic of reconstruction is addressed. The effectiveness of the recognition technique and some of the possibilities of reconstruction are both shown using data from the Ecors 2.5-D deep seismic reflection experiment (Marthelot et al., 1994). Finally, we discuss the combined recognition and reconstruction techniques in more detail, some possible modifications are given of the recognition method, and several further applications are indicated.

RECOGNITION METHOD

Unconditional maximum likelihood classification

The most common supervised classification method used for remote-sensing image data is maximum likelihood classification (Richards, 1993). Assuming that we have a vector $\mathbf{x}$ containing the attributes describing a local part of the data, classification only requires knowledge of the conditional probability density functions (pdf) $p(\omega_i \mid \mathbf{x})$, where ω_i are the I (predefined) spectral classes. Namely,

$$\mathbf{x} \in \omega_i \quad \text{if} \quad p(\omega_i \mid \mathbf{x}) > p(\omega_k \mid \mathbf{x}) \quad \forall k \neq i, \tag{1}$$

since the conditional pdf $p(\omega_i \mid \mathbf{x})$ represents the likelihood that ω_i is the correct class given the attribute vector $\mathbf{x}$. The spectral classes ω_i are mutually exclusive and span the spectrum Ω, such that always one and only one class must occur. Thus, mathematically, $\omega_1 \cup \omega_2 \cup \cdots \cup \omega_I = \Omega$ and $\omega_i \cap \omega_k = \emptyset \, \forall \, i \neq k$. These conditions are necessary to justify the use of Bayes' rule hereafter.

To be more specific, the spectral classes ω_i can be seen as the different types of anticipated signals and the attribute vector $\mathbf{x}$ as representing the local amplitude spectrum for a specific time and offset. In that case, the conditional pdf $p(\omega_i \mid \mathbf{x})$ will represent the probability that a particular signal type is present at a specific time and offset given the local amplitude spectrum.

The (yet) unknown conditional pdf $p(\omega_i \mid \mathbf{x})$ can be estimated using Bayes' rule, which states that

$$p(\omega_i \mid \mathbf{x}) = p(\mathbf{x} \mid \omega_i) p(\omega_i) / p(\mathbf{x}), \tag{2}$$

thus relating the conditional pdf $p(\omega_i \mid \mathbf{x})$ to the a priori pdf $p(\omega_i)$, to $p(\mathbf{x})$, and to the other conditional pdf $p(\mathbf{x} \mid \omega_i)$. The denominator $p(\mathbf{x})$ represents the probability distribution that any class is present for the attribute vector $\mathbf{x}$. Since it is independent of the classes ω_i, it is either neglected or treated as a normalization constant. The pdf $p(\omega_i)$ reflects our a priori estimation of the occurrence of class ω_i and is independent of the data. Finally, the conditional pdf $p(\mathbf{x} \mid \omega_i)$ describes the likelihood that class ω_i will be present at the data point with attribute vector $\mathbf{x}$ given the characteristics of class ω_i. The learning phase of the supervised classification method mainly consists of estimating this distribution function.

If we assume that the conditional pdf $p(\mathbf{x} \mid \omega_i)$ has a normal or Gaussian distribution, then equation (2) can be written as

$$p(\omega_i \mid \mathbf{x}) = cp(\omega_i)(2\pi)^{-J/2}\left|\mathbf{C}^{(i)}\right|^{-1/2} \times \exp\left\{-\tfrac{1}{2}\left(\mathbf{x}-\mathbf{m}^{(i)}\right)^t \mathbf{C}^{(i)^{-1}}\left(\mathbf{x}-\mathbf{m}^{(i)}\right)\right\}, \quad (3)$$

with c the normalization constant, $|\mathbf{C}^{(i)}|$ the determinant of the covariance matrix $\mathbf{C}^{(i)}$ for class ω_i, $\mathbf{m}^{(i)}$ the mean of the attributes for class ω_i, t the vector transpose, and J the size of vector $\mathbf{x}$.

If either the covariances are assumed to be negligible or not enough samples can be selected for each spectral class to justify their estimation, then equation (3) can be approximated by using only the variances $\sigma_j^{(i)^2}$ for each attribute j of class ω_i, i.e.,

$$p(\omega_i \mid \mathbf{x}) = cp(\omega_i)(2\pi)^{-J/2}\prod_{j=1}^{J}\sigma_j^{(i)^{-1}} \times \exp\left\{-\left(x_j - m_j^{(i)}\right)^2 \Big/ 2\sigma_j^{(i)^2}\right\}. \quad (4)$$

Equation (4) constitutes a close approximation to the exact equation (3) if covariances are negligible. However, if no accurate estimation of the covariance matrix can be obtained, then equation (4) may be a bad substitution for equation (3), since no guarantee exists that attributes are not highly correlated. In the next subsection, a standard technique to reduce existing correlations between attributes will be described.

Therefore, to perform supervised classification of the data, in the first stage a representative set of $K^{(i)}$ samples is selected, and their attribute vectors $\mathbf{x}^{(i,k)}$ $(k = 1, K^{(i)})$ are calculated for each class ω_i. Then in the learning stage, the means and covariance matrices or variances are calculated using

$$\mathbf{m}^{(i)} = \frac{1}{K^{(i)}}\sum_{k=1}^{K^{(i)}}\mathbf{x}^{(i,k)} \quad (5)$$

and

$$\mathbf{C}^{(i)} = \frac{1}{K^{(i)}-1}\sum_{k=1}^{K^{(i)}}\left\{\left(\mathbf{x}^{(i,k)}-\mathbf{m}^{(i)}\right)\left(\mathbf{x}^{(i,k)}-\mathbf{m}^{(i)}\right)^t\right\} \quad (6)$$

or

$$\sigma_j^{(i)^2} = \frac{1}{K^{(i)}-1}\sum_{k=1}^{K^{(i)}}\left(x_j^{(i,k)} - m_j^{(i)}\right)^2, \quad (7)$$

where j ranges over the elements of the attribute vector $\mathbf{x}$ (i.e., from 1 to $J^{(i)}$, the number of attributes selected for the ith class).

Finally, all conditional pdfs are calculated with either equation (3) or (4) and classified using equation (1).

As was mentioned above, in our case the attribute vector contains the amplitude spectrum of a local part of the data, and the characteristics are then calculated from representative picks of each spectral class (i.e., each signal type). To prevent detection problems due to amplitude fluctuations of reflection, energy contained in all amplitude spectra is normalized. Since classification occurs using a local f-t analysis, basically a flexible frequency analysis is used to detect coherent energy. With the amplitude spectra as attributes, equation (5) and (7) therefore represent the mean spectral amplitude and variance, respectively.

Correcting for correlations between attributes

A fundamental problem arises if equation (3) has to be approximated by equation (4) due to an insufficient number of picks to guarantee a close estimation of the covariance matrix $\mathbf{C}$. Namely, nonanticipated correlations between attributes may reduce the detection ability of the recognition technique. To give an idea of the number of picks required, it is often assumed that at least ten times more training samples are needed than the number of used attributes (Richards, 1993).

A principal component analysis is a standard technique in clustering or classification problems used to decrease existing dependencies between attributes and thereby reduce the number of required attributes for an accurate performance. This analysis is also known as the Karhunen-Loève transform and is most often used for image reconstructions and statistical analyses (Richards, 1993). It is closely related to the singular value decomposition (SVD), as shown by Freire and Ulrych (1988). In geophysics, it is most often used for wave-noise separation and wave-wave separation (Hemon and Mace, 1978; Freire and Ulrych, 1988; Glangeaud and Mari, 1994) or (multivariate) statistical analyses in reservoir characterization studies (Dumay and Fournier, 1988).

It is well known that a nonsquare matrix $\mathbf{X}$ can by rewritten in a semidiagonal form by means of an SVD (Lanczos, 1961), i.e.,

$$\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^t. \quad (8)$$

The columns of both matrices $\mathbf{U}$ and $\mathbf{V}$ contain orthogonal eigenvectors, and the matrix $\mathbf{\Lambda}$ is a diagonal matrix of which the elements λ_j are the singular values of $\mathbf{X}$. These singular values are the positive square roots of the eigenvalues of both $\mathbf{X}\mathbf{X}^t$ and $\mathbf{X}^t\mathbf{X}$. The Karhunen-Loève transform or principal component analysis consists of projecting data upon the eigenvectors contained in $\mathbf{U}$. Thus, for example, the principal component matrix $\mathbf{Y}$ is defined as (Freire and Ulrych, 1988; Richards, 1993)

$$\mathbf{Y} = \mathbf{U}^t\mathbf{X} = \mathbf{\Lambda}\mathbf{V}^t. \quad (9)$$

Redundancy in data can now be decreased by using only a limited number of eigenvectors $\mathbf{u}_j$, where $\mathbf{u}_j$ is defined as the jth column of $\mathbf{U}$, and singular values are sorted in decreasing order. Therefore, a vector $\mathbf{x}$ can be transformed to a vector $\mathbf{x}'$ of reduced length using only the first p eigenvectors $\mathbf{u}_j$:

$$x_j' = \mathbf{u}_j \cdot \mathbf{x}, \qquad j = 1, p. \quad (10)$$

Hence, a vector is projected upon a new set of axes, displaying a decreasing order of correlations between the training

attributes. Furthermore, a variable is needed to measure the perfectness of the projection (i.e., the information still contained in the new vectors). This can be done using the energy E_p contained in the first p eigenvectors, i.e.,

$$E_p = \sum_{j=1}^{p} \lambda_j^2 \Big/ \sum_{j=1}^{J} \lambda_j^2. \tag{11}$$

Thus, to decrease redundancy in the training samples and existing correlations between attributes, the columns of matrix $\mathbf{X}$ should contain the amplitude spectra of the training samples. The eigenvectors $\mathbf{u}_j$ are then used as a new set of axes upon which the attributes $\mathbf{x}$ can be projected using equation (10). Finally, note that the computation of the SVD is not strictly required, since $\mathbf{\Lambda}$ and $\mathbf{U}$ can be directly computed from $\mathbf{X}\mathbf{X}^t$. In our application, normally one or, in some rare cases, two eigenvectors will represent virtually 100% of the total energy, since the amplitude spectra of the picks are highly similar.

Classification then proceeds in the following way. First, examples for each single class are picked, and the needed eigenvectors $\mathbf{u}_j$ are calculated. Then, the samples are projected upon the eigenvectors with equation (10), and the averages and variances of the transformed attributes are calculated using equations (5) and (7). The number of required eigenvectors can be estimated using their energy distribution [equation (11)]. Finally, all samples to be classified are projected upon the same set of eigenvectors, and all conditional pdfs are calculated using equation (4) and classified using equation (1).

Unfortunately, this method requires the construction of a matrix $\mathbf{X}^{(i)}$ for each class ω_i, and projection has to be done upon the different sets of eigenvectors $\mathbf{u}_j^{(i)}$. If this is unwanted, then a different approach can be used, namely by means of a canonical or factor analysis. This method calculates the covariance matrices of training samples within an identical class and between different classes. It then tries to find a projection which minimizes distances within classes, while at the same time maximizing distances between classes, thus requiring only a single projection. For more details, we refer to Richards (1993) and for an application to Dumay and Fournier (1988).

Conditional maximum likelihood classification: Bayes' optimum rule

The classifier equation (1) labels each data part based on its spectral components only, without any consideration for the classification of the elements in its neighborhood. In general, however, it can be assumed that some correlation exists between neighboring points. That is, there will be some homogeneity in the spatial distribution of the different signal types. In a second iteration, this correlation can be incorporated into the classification algorithm.

The simplest method is to use some spatial averaging filter on all conditional pdf $p(\omega_i \mid \mathbf{x})$ and to reapply equation (1) on the new distributions. However, we choose a different, though related, approach using Bayes' optimum rule. Contrary to Richards (1993), we propose a method which allows only for spatial correlations between identical phases, thus neglecting the fact that some relationship may exist between the different spectral classes (e.g., a reflection is more easily misclassified as a refraction than as ground roll). By quantifying these correlations, this effect can be corrected for.

Thus, after the conditional pdf $p(\omega_i \mid \mathbf{x})$ have been calculated, they are multiplied with a weighting function incorporating the relations between different classes, $p(\omega_k \mid \omega_i)$, or the influence of a (predefined) neighborhood $A_\mathbf{x}$ surrounding the data point, $p(\omega_i \mid A_\mathbf{x})$. Since we neglect any relation between the different spectral classes, the optimum conditional pdf $L_\mathbf{x}(\omega_i)$ are obtained from Bayes' rule [equation (2)] using

$$\begin{aligned} L_\mathbf{x}(\omega_i) &= p(\omega_i \mid A_\mathbf{x}) p(\omega_i \mid \mathbf{x}) \\ &= p(\omega_i \mid A_\mathbf{x}) p(\mathbf{x} \mid \omega_i) p(\omega_i) / p(\mathbf{x}), \end{aligned} \tag{12}$$

where we define

$$p(\omega_i \mid A_\mathbf{x}) = E\{p(\omega_i \mid \mathbf{y}),\ \mathbf{y} \in A_\mathbf{x}\}, \tag{13}$$

yielding that the conditional pdf $p(\omega_i \mid A_\mathbf{x})$ is the expectancy or mean of the likelihoods of class ω_i for all data points $\mathbf{y}$ in the neighborhood $A_\mathbf{x}$. If necessary, some weighting can be incorporated or several iterations can be used to update the likelihood distributions, since after each iteration the neighboring likelihood distributions have changed.

To reclassify each data point, equation (1) changes to

$$\mathbf{x} \in \omega_i \quad \text{if} \quad L_\mathbf{x}(\omega_i) > L_\mathbf{x}(\omega_k) \quad \forall k \neq i. \tag{14}$$

Extraction

After classification has taken place, "cleaned-up" versions of the seismic section $\phi(x, t)$ can be obtained by extracting the different classes. To extract these features, a representation $R(x, t)$ of their likelihoods for each data point (i.e., for each offset x and time t) has to be obtained, which is then multiplied with the original section, yielding

$$\phi'(x, t) = \phi(x, t) R(x, t). \tag{15}$$

Depending on the desired result, this representation may yield the conditional pdf of either equation (2) or (12) for a particular class only. That is, for class ω_i,

$$R_i(x, t) = p(\omega_i \mid \mathbf{x}(x, t)) \text{ or } L_{\mathbf{x}(x,t)}(\omega_i). \tag{16}$$

In addition, the classifier conditions (1) or (14) can be used as thresholds. In this case, the likelihood of a specific class remains unchanged if the condition is satisfied, but is set to zero if not, that is, for class ω_i,

$$R_i(x, t) = \begin{cases} p(\omega_i \mid \mathbf{x}(x, t)) \text{ or } L_{\mathbf{x}(x,t)}(\omega_i) & \text{if } \omega_i \text{ is selected class} \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

where $\mathbf{x}(x, t)$ is the attribute vector obtained at a specific time and offset.

RECONSTRUCTION METHOD

Although many reconstruction techniques exist in remote sensing or more accurately in image processing (see, e.g., Stark, 1987), we resorted to a technique which has already been used for a long time in geophysics and is both easy to implement and relatively fast. The local τ-p (intercept time–apparent dip or slowness) transform has already been used several times to

extract laterally coherent signals (see, e.g., Harlan et al., 1984, and Kong et al., 1985).

The τ-p transform or slant stack is known in mathematics as the Radon transform and consists of a summation along "slant lines" described by a specific slowness and intercept time. Thus, adjacent traces are incorporated in the extraction process to increase and facilitate detectability. This is in contrast to the recognition method which does not make use of the information contained in the adjacent traces, but treats each trace separately. The drawback of the use of the τ-p transform for reconstruction, however, is that no phase changes are allowed for. That is, it is assumed that a signal has constant amplitude and phase along the slant lines.

To extract and reconstruct the coherent signals, the locally τ-p transformed section is multiplied with a coherency measure along the same slant lines before being inverted back to the offset-time domain. This method allows for the interpolation and extrapolation of the coherent signals.

Following Kong et al. (1985), the seismic section $\phi(x, t)$ is divided into N partly overlapping subsections $\phi_n(x, t)$ of K traces each, which are transformed to the τ-p domain:

$$\psi_n(p_i, \tau_j) = \sum_{k=1}^{K} \phi_n(x_k, \tau_j + p_i x_k). \tag{18}$$

Since the value $\tau_j + p_i x_k$ will not necessarily be represented by a discrete time point, a two-point linear interpolator in time will be used for its approximation. Moreover, edges may introduce artifacts in the transformed section (Yilmaz, 1987). Therefore, all edges are progressively weighted down. The exact weighting is of minor importance for the performance of the recognition and reconstruction algorithms.

Similar to equation (15), extraction of the desired features consists of the multiplication with some representation $R(p_i, \tau_j)$ of coherency:

$$\psi'_n(p_i, \tau_j) = \psi_n(p_i, \tau_j) R_n(p_i, \tau_j). \tag{19}$$

Applying the inverse τ-p transform on the summation of all subsections $\psi'_n(p_i, \tau_j)$ then yields the "cleaned-up" seismic section $\phi'(x_k, t_j)$:

$$\phi'(x_k, t_j) = \frac{d}{dt} H \left[\sum_{i=1}^{NP} \sum_{n=1}^{N} F_w(p_i, t_j - p_i x_k) \psi'_n(p_i, t_j - p_i x_k) \right]. \tag{20}$$

The time derivative and the Hilbert transform H are required to prevent spectral distortion of the signal. NP represents the total number of slowness traces in the τ-p domain. The weighting function $F_w(p_i, \tau_j)$ equals the reciprocal of the number of times a specific (p_i, τ_j) has been calculated. It is used to calibrate the different subsections $\psi'_n(p_i, \tau_j)$ and to correct for artifacts due to partial overlap. However, it is only of importance at the edges, since it is otherwise uniform. Moreover, to prevent the loss of important information at the edges, very often nonrelevant traces and time samples are added at the edges.

What measure can now be best used to represent or quantify coherency? Kong et al. (1985) used smoothed semblance. Semblance is equivalent to the ratio of the coherent over the total energy:

$$S_n(p_i, \tau_j) = \frac{\left[\sum_{k=1}^{K} \phi_n(x_k, \tau_j + p_i x_k) \right]^2}{K \sum_{k=1}^{K} \phi_n^2(x_k, \tau_j + p_i x_k)}. \tag{21}$$

Smoothing is necessary to stabilize the semblance map and can be achieved by a rolling average along the τ-axis of length W, representing the longest principal period present in the data. The smoothed semblance is then used as the representation of coherency in equation (19). Kong et al.'s method therefore uses semblance-weighted slant stacks for the reconstruction of laterally coherent signals. For more details, examples, and optimization of parameter settings, see Kong et al. (1985). Yilmaz (1987) provides background on the τ-p transform.

It is, however, also possible to use different measures of coherency. In fact, it is possible to perform recognition directly in the τ-p domain. In that case, the attribute vector in the conditional pdf $p(\omega_i \mid \mathbf{x}(\tau, p))$ or $L_{\mathbf{x}}(\tau, p)(\omega_i)$ will represent the attributes (i.e., the local amplitude spectrum) of a data point in the τ-p domain. The representation $R_n(p_i, \tau_j)$ will then be equal to the conditional pdfs for a specific spectral class for all intercept times and slownesses with possibly a conditional threshold, analogous to the extraction method in the x-t domain discussed above [equations (16) and (17), respectively, where $\mathbf{x}$ now equals $\mathbf{x}(\tau, p)$]. In this case, the recognition method can be seen as a flexible frequency analysis performed in the τ-p domain instead of in the x-t domain.

DATA

The data used to show the effectiveness of the new method comes from the Ecors 2.5-D deep seismic reflection experiment (Marthelot et al., 1994) conducted within the western North Pyrenean thrust zone, an area of complex surface tectonics. The experiment was performed to test the feasibility of determining the 3-D geometry of deep structures within a crustal volume using current acquisition techniques from exploration geophysics. It is called 2.5-D because it consists only of a single swath of a true 3-D exploration configuration.

The experiment included six parallel receiver lines of 120 receiver groups each with 14 shot lines perpendicular to the receiver lines using a symmetrical split-spread configuration. The distances between lines, receivers, and shots in a column were, respectively, 1600, 100, and 400 m. The lines and columns were moved by steps of 2 km in the direction of the lines. In total, 292 shots were recorded (Marthelot et al., 1994; Siliqi, 1994).

The North Pyrenean Zone (NPZ) is limited on the north by the North Pyrenean Frontal Thrust and on the south by the steep North Pyrenean Fault. This zone extends along the whole length of the Pyrenean Range with widths from less than 20 km (central and eastern part) to 60 km (western Pyrenees). During the Early Cretaceous, the NPZ was thinned and affected by significant subsidence, whereas in the Late Cretaceous and Tertiary, it was compressed in response to the opening of the Bay of Biscay, the relative motion between Iberia and Europe, and the formation of the Pyrenees. More geological background and maps are provided in Marthelot et al. (1994), Siliqi (1994), and Lecerf (1996).

The experiment resulted therefore in a large dataset recorded in a region of complex tectonics. The size of the dataset made manual processing techniques nearly impossible or at least very expensive. As a consequence, any technique to increase the S/N ratio should be automatic or at least semiautomatic.

DATA APPLICATION: RECOGNITION

Strategy

To detect and extract the coherent energy present in the seismic sections, the following strategy was employed. Since we were mostly interested in reflections from the deeper parts of the crust, just a single spectral class was used, namely a picked reflection from the Moho.

Thus, in the first stage, the Moho reflection is selected on several adjoining traces of a high-quality common shot gather. Then, in the second (learning) stage, the characteristics (i.e., the means and variances of the attributes) of this specific phase are calculated using equations (5) and (7). We do not use the covariance matrix, since it will rarely happen that sufficient picks can be done to justify its estimation and thus its use. The attribute vectors contained the amplitude spectra of a 25% cosine-tapered window centered around the picks (i.e., 25% of the total window length is tapered). Energy contained in the amplitude spectra was normalized to correct for possible different amplitudes of signals. To prevent a degradation of detection results due to unaccounted for correlations between attributes, a principal component analysis was used following equation (10). Since the amplitude spectra of the picked reflection displayed a very high similarity, only a single eigenvector was needed.

Next, likelihood estimation or "classification" takes place. That is, the likelihood of the presence of the picked phase is calculated with equation (4) for all times and offsets using a sliding window to obtain the local amplitude spectra. In a second iteration, the contributions of a predefined neighborhood is incorporated using equation (12) to readapt the likelihoods. This neighborhood has the shape of the wings of a butterfly centered around the considered sample to allow for static shifts and time dips.

Finally, the obtained likelihoods are used as a representation of coherency in equation (15) to "clean up" the seismic section and thus to extract the coherent energy. Before extraction, the likelihood distributions are corrected for the employed windowing taper by means of a convolution of this taper with the likelihood distributions. Since only a single spectral class was used, all amplitude factors in equation (4) were neglected [including the a priori distribution $p(\omega_{refl}) = 1$]; only the product of exponentials was retained.

Results

Figure 1a shows common shot gather (CSG) 8184 after frequency filtering and application of automatic gain control to increase visibility. Clearly visible are the top of the basement at 4 s and the Moho just below 9 s. Figure 1b shows a zoom-in on the same CSG from 8.5 to 10 s and offsets 40.0 to 45.0 km. It displays the positions of the 14 picks used to describe the Moho. A 128-ms (16-points) window is used for the calculation of the means and variances of the frequencies contained in the amplitude spectrum.

Picks were done on the Moho reflection because our initial aim was to extract it automatically in the whole dataset. However, the computation of the unconditional likelihood distributions of equation (4) without a principal component analysis had two unexpected results. Firstly, the method seemed to work too rigorously, since only the picked parts of the Moho reflection were detected. Secondly, small parts of overlying reflections showed up in addition to large parts of the refraction branches. Seemingly, the spectral signature of deep reflections and refraction arrivals have a large similarity.

Adapting the parameters of window length and tapering did not change the results significantly. To detect more energy,

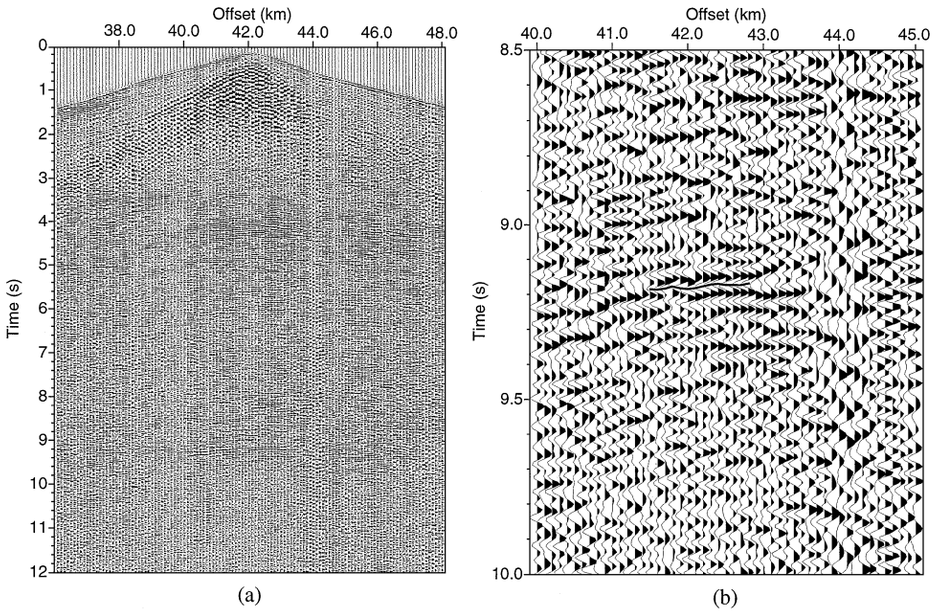

FIG. 1. CSG 8184. (a) Raw data. (b) Zoom-in showing the positions of the used picks (drawn line).

more, but questionable, picks of the Moho had to be included, though this was not the way we wanted to proceed.

To relax the strictness of the method, we applied a principal component analysis, thereby reducing the number of frequencies in the amplitude spectra to a single attribute, since the first eigenvector represented virtually 100% of the total energy. After projection of the local amplitude spectra of the Moho picks upon the first eigenvector and recalculating the mean and variance, this single attribute was used for likelihood estimation or "classification" and thereafter extraction. Figures 2a and 2b show the results. Much more energy is now passed as being "coherent." In the extracted section, nearly all reflections are visible.

In Figures 3a and 3b the results are displayed after incorporation of the surrounding likelihoods. A neighborhood of twice two adjacent traces was used. Static shifts of ±40 ms and linear move outs up to ±0.25 s/km were allowed for, resulting in a neighborhood with the shape of the wings of a butterfly. Any lateral coherency is now more evident in both the likelihood distribution and the extracted section.

If the method has to be semiautomatic, then the selected training set has to be representative not only for a single CSG,

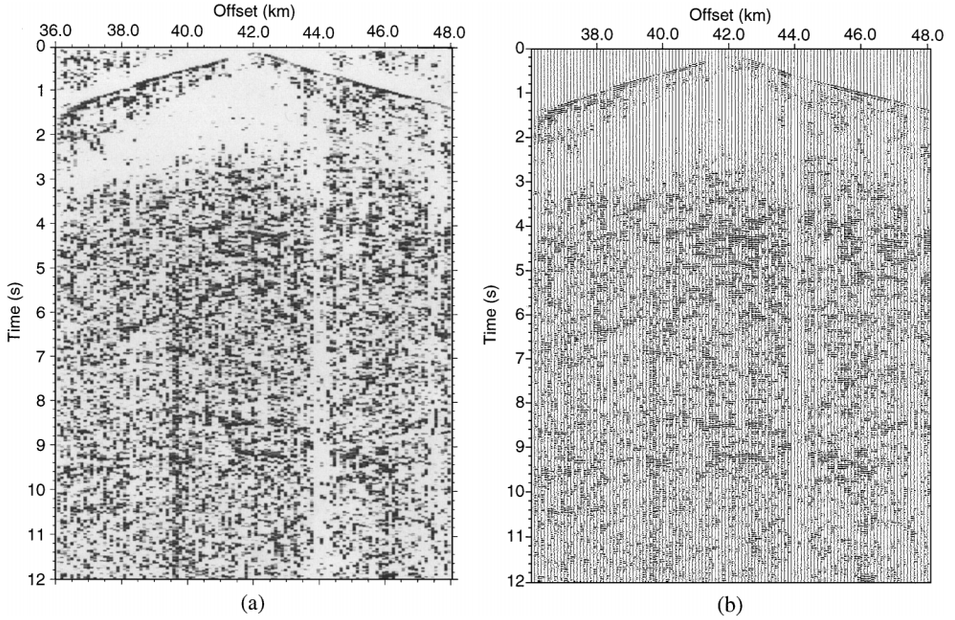

FIG. 2. CSG 8184. (a) Likelihood distribution of the presence of reflection energy using a principal component analysis to reduce the number of attributes. (b) Extracted data.

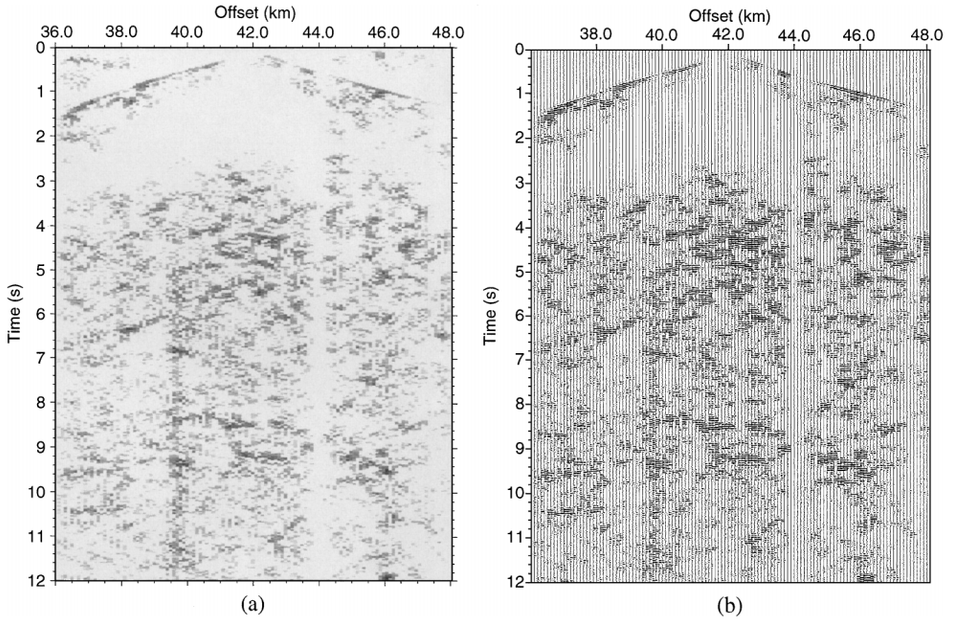

FIG. 3. CSG 8184. (a) Likelihood distribution of reflection energy after incorporation of surrounding probabilities. (b) Extracted data.

but also for a smaller or larger part of the stack profile. From extensive testing using CSGs up to several kilometers apart, several conclusions could be drawn.

First, the calculated characteristics proved to be very robust. That is, there is not much need to repick a phase to extract the coherent energy of other CSGs. However, if the coherent signals are already hard to detect with the eye, the algorithm will not detect them either, due to the fact that the trained eye of an interpreter will always be better able to detect coherent signals.

Nevertheless, trial and error tests have to be performed to check whether another set of picks can be obtained which results in better signal extractions and to verify to what extent the training set can extract signals from neighboring common midpoints (CMPs) or CSGs.

Finally, extracted reflectors may exhibit a more or less fragmentary character, since no reconstruction of "missing" energy can be done by the recognition technique. This topic will be addressed in the next section.

DATA APPLICATION: RECONSTRUCTION

After having detected and extracted coherent energy, we want to explore some of the possibilities of reconstruction. As a strategy, we adopt the philosophy that it is not correct to reconstruct energy directly over long offsets (i.e., many traces), but that an iterative approach is recommended. Thus, the number of traces will be gradually increased with iteration to interpolate small gaps first. Because of the enormous number of possible combinations only a few selected examples will be shown. However, which combination or strategy works best for a specific data set can only be found out by trial and error.

To begin, we show the results of the algorithm of Kong et al. (1985) using smoothed semblance as a representation of lateral coherence. This technique emphasizes any energy which is spatially coherent. The only tunable parameters are the range of slownesses, the number of traces used in the local τ-p analysis, and the width W of the averaging filter. To prevent reconstructions of ground roll energy, slowness ranges of ± 0.2 s/km were used and W was put to 64 ms, although W has only a minor influence. Two consecutive passes were performed using, respectively, 5 and 10 traces in the τ-p transformation. Subsections were taken with overlaps of 50%.

Figure 4a shows the result of the second pass. It is clearly visible that any lateral coherency has been stressed. Unfortunately, much undesired energy has remained or, worse, has been created. For example, some low frequency ground roll has remained. The worst examples of false reconstruction, however, can be found near the refraction branches in Figure 4a. There, energy has been created before the actual first arrivals.

Nevertheless, many of the artifacts can be removed if reconstruction is done after recognition. Figure 4b contains the reconstructed image of the extracted section displayed in Figure 2b using the same parameter settings as before. Clearly, the occurrence of many artifacts has been prevented, though some remain around the refraction branches. Applying the reconstruction technique on the extracted section of Figure 3b, in which the influence of the surrounding likelihoods has been accounted for, gave even better results, since even more background noise has been filtered out of the data (not shown). However, changing the order and applying recognition after reconstruction resulted in considerably worse images. Although it removed some of the artifacts, many remained.

As an alternative, recognition and reconstruction can be performed in a single step by directly calculating the local amplitude spectra and the likelihood distributions in the τ-p domain and regarding these as the desired representation of coherency. Multiplying the corrected likelihoods with the τ-p transformed subsection and inverse transforming to the x-t domain resulted in Figure 5, where subsections consisting of 5 traces were used.

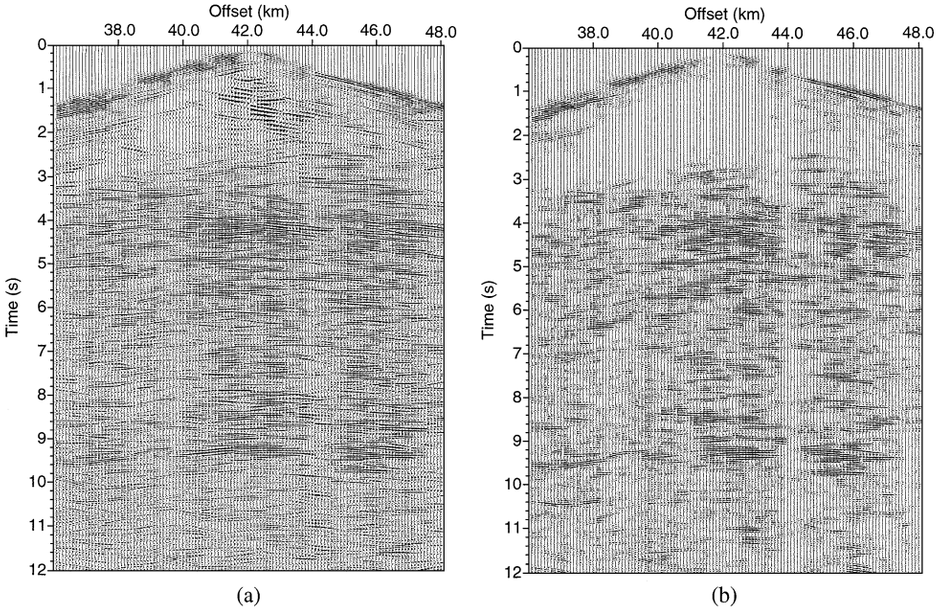

FIG. 4. CSG 8184. Reconstructed data with semblance weighting after two consecutive passes (with, respectively, 5 and 10 traces) using (a) raw data (Figure 1a) and (b) extracted data (Figure 2b).

This method produces high-quality reconstructions which remain close to the original seismic section. Thus, nearly no artifacts have been introduced. A second pass using 10 instead of 5 traces per subsection did not produce significantly different results. The image had become only slightly clearer. However, further reconstructions can be done using again the semblance technique of Kong et al. (1985).

A last conclusion which could be drawn from all reconstructions is the tendency of the repeated passes to emphasize more laterally coherent structures in the data than just a single pass does. On the other hand, results from these consecutive passes tend to exhibit more artifacts due to the exaggerated extrapolations, such as those displayed around the refraction branches.

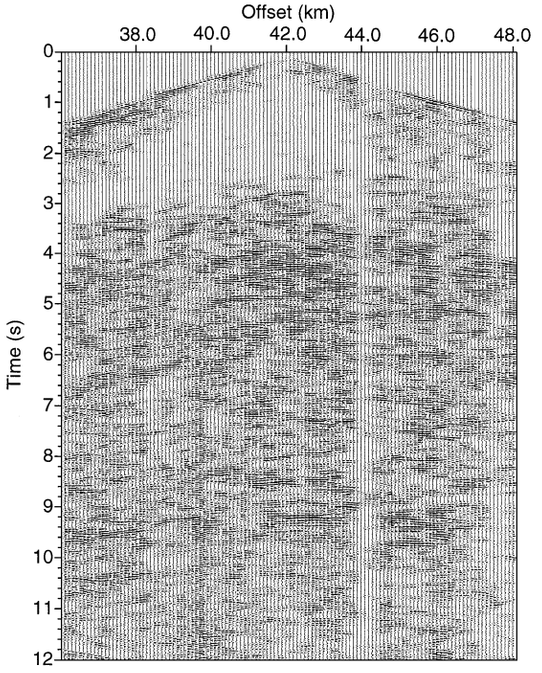

FIG. 5. CSG 8184. Reconstructed data using selection in the τ-p domain. One pass: 5 traces.

DATA APPLICATION: STACKED SECTIONS

As a second example, we will show how the combined methods significantly enhance the S/N ratio of a stacked section. The stacked section is situated at exactly the same position as the common midpoints of CSG 8184 (Figure 1a). In general, stacked sections from deep seismic reflection data have a lower quality than those frequently encountered in exploration seismics. However, the Ecors 2.5-D case had an extraordinarily low stack quality. This can be partly attributed to the particular acquisition configuration, which resulted in low fold stacks over large offsets (in excess of 4 km). In addition, the complexity of the surface tectonics has also contributed to the stack degeneration (Lecerf, 1996). This led to fragmentary and nonhyperbolic moveouts of the reflections.

To increase stacking quality, half-offsets were limited to 3000 m, and the bin size perpendicular to the receiver lines was reduced (Siliqi, 1994). Nevertheless, reflections are barely visible (Figure 6a). Though a semblance-weighted reconstruction did improve the stacked profile (Figure 6b), it did not result in an easily interpretable section, since details remain unclear. Only a very global idea of the existing geology can be obtained. Slant stacks were done in a single pass on subsections consisting of 20 traces with 50% overlap to emphasize the large structures in the data.

To increase the quality of the stacked section, recognition was applied before stack on all CMPs contributing to the stacked profile. To characterize the desired reflections, the same Moho picks (Figure 1b) were used throughout the whole stacking analysis, showing again the stability of the recognition method. The resulting profile displayed, however, a very fragmentary image of the reflector positions. This can easily

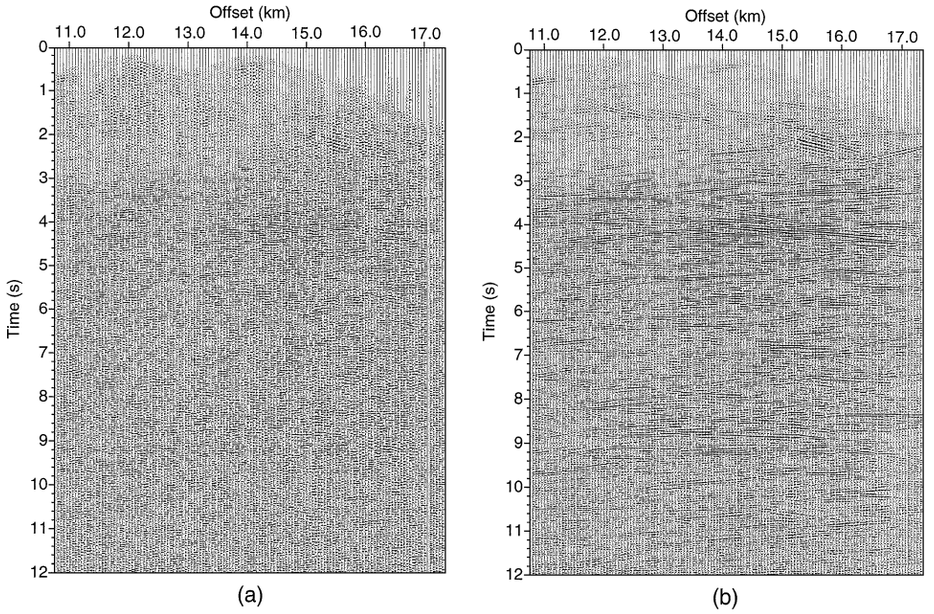

FIG. 6. Stack. (a) Raw data. (b) Associated reconstruction using semblance weighting.

be explained by the fact already discussed that the extracted reflections in CSGs tend to have a more or less fragmentary character (see, e.g., Figure 2b). Although already a clearer image of the reflector positions was obtained by applying reconstruction after stack on the stacked section, the clearest results were obtained using both recognition and reconstruction before stack.

Thus, recognition and thereafter reconstruction were applied on all CSGs contributing to the stacked profile. Prestack reconstructions were done in a single pass using subsections consisting of 10 traces with 50% overlap. Since the stack tended to reintroduce noncoherent energy, the recognition technique was reapplied after the stacked section had been obtained (see Figure 7a for the resulting profile). If required, the reconstruction technique can also be reapplied on the obtained profile with the same parameter settings as used for Figure 6b, which resulted in Figure 7b.

Figures 7a and 7b give clear indications of the locations of the reflectors, thereby rendering interpretation much easier compared with the raw stacked sections (Figures 6a and 6b). It leaves no doubt that the S/N ratio has been increased significantly.

It should be noted that applying reconstruction both before and after stacking without the application of any recognition also improved results, but to a much lesser extent because a large amount of background noise was left in the data. This conclusion could already be expected from a comparison of Figures 4a and 4b.

Finally, as a last test, reconstructions were done prestack on the CSGs using recognition in the τ-p domain (analogous to Figure 5). Although final reconstruction results were better than Figure 6b, more background noise remained than in Figure 7b, and some unexplained strongly dipping features could be seen interfering with the Moho reflection (just below 9 s). It remained unclear if these strongly dipping features were actually present in the data and caused by geology or were caused by nonremoved noise interference.

As a final remark, the strong resemblance between Figures 7b and Figure 4b should be noted, indicating the excellent reconstruction of CSG 8184.

DISCUSSION

The combined recognition and reconstruction methods result in a very powerful technique, since it combines the power of the stacking process to increase the S/N ratio with the possibility to directly select a specific seismic phase.

It is this selection possibility that sets the method apart from nearly all other techniques to increase the S/N ratio. The skeletonization technique (Le and Nyland, 1990; Lu and Cheng, 1990) has nearly no selection possibilities at all. It will simply try to correlate waveforms laterally. Both the statistical method of Hansen et al. (1988) and the methods using the local τ-p transform (Harlan et al., 1984; Kong et al., 1985) can only select by changing the range and length of the slant lines. However, this will not always be adequate as shown in Figure 4a, where, for example, some surface wave energy remains. On the other hand, the performance of all techniques can be enhanced by means of accurate data processing.

At a first glance, it may seem remarkable that amplitude spectra of reflections picked at 9 s display a large similarity with those of reflections at 3 or 4 s. However, existing estimates of Q values in the region suggest that attenuation is mainly due to the upper sedimentary cover with values of Q of 300 and less in the subsurface and 700 and higher in the underlying crust (Rappin et al., 1991; Rappin, 1992). Hence, spectra of reflections from the lower part of the sedimentary cover and below it do not evolve much with depth. Moreover, it explains why the refraction branches are better detected than the direct arrivals close to the source (e.g., Figure 2b).

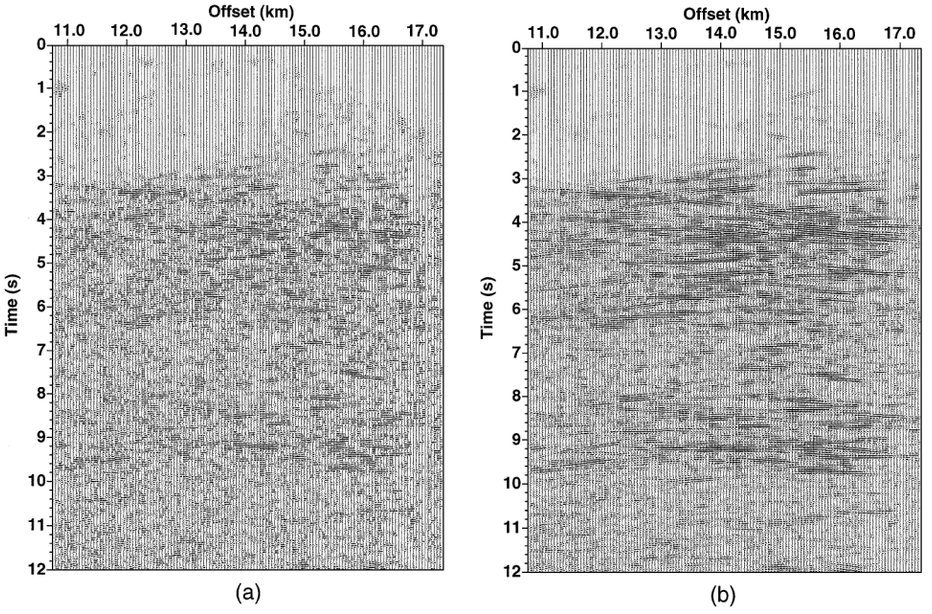

FIG. 7. Stack. (a) Prestack recognition and reconstruction on CSGs and a second application of recognition after stack. (b) Associated reconstruction using semblance weighting.

An interesting side effect is that in regions with strongly varying attenuation values, the recognition method is able to detect and extract a specific reflection only, thus enabling the extracting of specific reflection horizons only. Otherwise, to extract reflections at all depths and offsets, several different reflections have to be picked to correct for changes in the characteristics.

The influence of some parameters has not yet been accounted for. The most important step in any supervised classification method is the selection phase and only to a lesser extent the learning phase. The selection of a representative set of picks is of vital importance, especially if it also has to represent the coherent energy of other CSG or CMP gathers. Clearly, there will be a trade-off between the internal consistency of the picked set and the amount of extracted energy. Or to speak in the terms of the statistical method of Hansen et al. (1988), there will be a trade-off between the "false alarm rate" and the probability of a "miss." The best way to obtain a representative set is simply by trial and error.

The length of the sliding window in the local f-t analysis for calculation of the attributes has to represent the characteristic scale length of the selected seismic phase, which is at least equal to its largest principal period. We found, however, that the results remained quite stable for larger windows. As an alternative, a change of attributes can be considered; namely, a wavelet transform may provide the desired characteristic scale-lengths.

The application of a principal component analysis greatly improved the detection results. It reduced the number of attributes to a single one, thereby correcting for possibly existing correlations between the attributes. Moreover, it produced superior detection results over other than Gaussian probability distributions which we tested. The simplest of these consisted of dividing each term within the exponentials of equation (4) with the number of used attributes J to correct for the influence of different window lengths and thus the number of attributes. This method may be applicable if more than a single eigenvector is used in the principal component analysis.

Combined recognition and reconstruction in the x-t domain gives the best results. However, the reconstruction technique explicitly demands phase consistency along the stacking trajectory, whereas the recognition technique allows for random perturbations of the phase. Thus, the reconstruction technique seems to limit the possibilities of the more general recognition technique to some extent.

If reflection energy is not distributed on a straight line (e.g., due to statics), then semblance weighting may not be the optimum method, since a large amount of energy will be suppressed. In that case, the method of Harlan et al. (1984) as implemented by Bano (1989) may be a better solution. This method estimates the standard deviation of noise from histograms of a τ-p transformed part of the data containing noise only, and uses thereafter twice this standard deviation as a threshold for all τ-p transformed data.

As an alternative to semblance-weighted slant stacks, spatial prediction filtering (Hornbostel, 1991; Abma and Claerbout, 1995) in either t-x or f-x domain can be considered for signal reconstruction, since these techniques are also able to give an estimate for "missing" energy. However, it remains explicitly assumed that events are at least locally linear; furthermore, it is unclear if these techniques are able to deal with signals displaying phase perturbations. Thus, final results will probably display large similarities.

Some of the limitations implicit to the recognition method have to be thoroughly realized. First of all, the recognition algorithm is basically a classification technique. This implies that, in the case of superposed signals, the algorithm will only be able to recognize the strongest signal or none at all and, thus, no reflections hidden by ground roll or excessive noise. Also very close to overlapping superpositions of two reflections may pose problems (e.g., in pinch outs). On the other hand, reconstruction methods will probably solve some of these problems.

In addition, amplitudes are both adjusted by the recognition and reconstruction techniques, which may pose a problem if amplitude information becomes of fundamental importance for interpretation (e.g., amplitude variation with offset). Moreover, reconstructions may remove some of the existing small-scale features of reflections. On the other hand, the original sections always remain available for a joint interpretation.

Finally, we want to state that many of the artifacts visible in the reconstructed seismic sections can be prevented if additional processing is done (e.g., muting of traces and removal of surface waves). The results, however, show that the recognition technique is very stable even on raw data, because any processing was only done to increase visibility. Nevertheless, results will improve with some fine tuning.

CONCLUSIONS

We have presented a new method to enhance the S/N ratio of a seismic section in a semiautomatic way using a supervised classification technique from remote sensing. The employed maximum likelihood technique allows for the detection and extraction of any energy that has the same spectral attributes as a selected seismic phase and should therefore be seen as a recognition technique.

Because the method is insensitive to phase perturbations of a signal, both diffractions and waves with polarity changes can be detected by the technique. Moreover, the recognition method can be applied both prestack and poststack.

A shortcoming of the recognition method is the fact that it does not use the information contained in adjacent traces, but treats each trace separately. Thus, it cannot estimate "missing" energy. Therefore, the new method was combined with two reconstruction techniques using local τ-p transforms to combine the S/N ratio enhancement of the stacking process with the possibility to directly select only particular phases.

The first reconstruction method, using semblance-weighted slant stacks, gave good results for both stack profiles and CSGs after extraction had been done. The second method, applying recognition directly in the τ-p domain, resulted in particular good reconstructions for CSGs, but worked less well for stacks. Both reconstruction techniques demand, however, phase consistency of the signal energy, thereby sacrificing some of the possibilities of the recognition technique.

To test the method, we applied it to the Ecors 2.5-D dataset, which has a very low data quality as far as the lateral coherency of reflections and their ability to be stacked correctly are concerned. Our method succeeded in enhancing the S/N ratio in both individual shot gathers and in the stacked section. Figure 7b shows much more clearly than the brute stack of Figure 6a that the only laterally coherent reflections correspond to the fold and thrust system of the sedimentary cover (3–5 s) and the base of the crust at 9 s. In between, the whole crust

(sedimentary cover excepted) is characterized by a large number of short reflection segments indicative of a strong heterogeneity at small scale.

The potential applications of the described method are manifold.

The recognition technique can be used as an interactive picking tool. That is, after picking only a part of a reflection, the interpreter can pick more easily the remaining part and other reflections in both common shot gathers and final stacks with the aid of the extracted sections.

Interpretation of stacked seismic sections can be facilitated, since the coherent energy (i.e., the reflectors) can be identified more easily in noisy data. In addition, it can be used for data segmentation (i.e., reduction of the data volume for storage).

Finally, the removal of noncoherent energy in common midpoint gathers prevents stack degeneration due to noise interference, and in stacked sections it will prevent the occurrence of migration artifacts such as, for example, smiles (Kong et al., 1985; Yilmaz, 1987). A limitation of the method is its amplitude fidelity. Care should be taken in using the method if the interpretation of true amplitude data is needed because the recognition technique adjusts the amplitudes; after reconstruction, small-scale features of reflectors may have disappeared.

ACKNOWLEDGMENTS

The first author wishes to thank Kabir Roy-Chowdhury for discussions and Frédérik Pivot and Jean-Michel Marthelot for giving him an overview of all the difficulties associated with the dataset and for allowing him free access to their software. In addition, especially Georges Poupinet is thanked for urging each time we find a more generally applicable method and for numerous discussions. Finally, we are grateful for the constructive reviews of Michel Dietrich, Jérôme Mars, Clark Trantham, and especially Warren S. Ross.

REFERENCES

Abma, R., and Claerbout, J., 1995, Lateral prediction for noise attenuation by t-x and f-x techniques: Geophysics, **60**, 1887–1896.

Bahorich, M., and Farmer, S., 1995, 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube: The Leading Edge, **14**, no. 10, 1053–1058.

Bano, M., 1989, Extraction automatique des réflexions, modelisation des diffractions et migration des données de sismique profonde ECORS: Ph.D. thesis, Université Louis Pasteur, Strasbourg.

Dumay, J., and Fournier, F., 1988, Multivariate statistical analyses applied to seismic facies recognition: Geophysics, **53**, 1151–1159.

Freire, S., and Ulrych, T., 1988, Application of singular value decomposition to vertical seismic profiling: Geophysics, **53**, 778–785.

Glangeaud, F., and Mari, J., 1994, Wave separation: Editions Technip.

Hansen, K., Roy-Chowdhury, K., and Phinney, R., 1988, The sign filter for seismic event detection: Geophysics, **53**, 1024–1033.

Harlan, W., Claerbout, J., and Rocca, F., 1984, Signal/noise separation and velocity estimation: Geophysics, **49**, 1869–1880.

Hemon, C., and Mace, D., 1978, Essai d'une application de la transformation de Karhunen-Loève au traitement sismique: Geophys. Prosp., **26**, 600–626.

Hornbostel, S., 1991, Spatial prediction filtering in the t-x and f-x domains: Geophysics, **56**, 2019–2026.

Kong, S., Phinney, R., and Roy-Chowdhury, K., 1985, A nonlinear signal detector for enhancement of noisy seismic record sections: Geophysics, **50**, 539–550.

Lanczos, C., 1961, Linear differential operators: D. Van Nostrand Co.

Le, L., and Nyland, E., 1990, Pattern analysis of seismic records: Geophysics, **55**, 20–28.

Lecerf, D., 1996, Apports des données sismiques E.C.O.R.S. 2,5D à la connaissance de la structure superficielle de la croûte dans la zone nord pyrénéenne: Ph.D. thesis, Université Louis Pasteur, Strasbourg.

Lu, S.-Y., and Cheng, Y.-C., 1990, An iterative approach to seismic skeletonization: Geophysics, **55**, 1312–1320.

Marfurt, K., Kirlin, R., Farmer, S., and Bahorich, M., 1998, 3-D seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, 1150–1165.

Marthelot, J.-M., Siliqi, R., Bitri, A., Paul, A., Hirn, A., Daignières, M., Damotte, B., Specht, M., de Bazelaire, E., Lortscher, A., and Rappin, D., 1994, Three-dimensional imaging of the crust using a sparse land acquisition grid: The ECORS 2.5-D experiment: Tectonophysics, **232**, 365–377.

Rappin, D., 1992, Apport des analyses d'amplitude et temps-fréquence à l'exploitation des données de sismique profonde: Ph.D. thesis, Université Louis Pasteur, Strasbourg.

Rappin, D., Marthelot, J.-M., and de Bazelaire, E., 1991, Amplitude and time-frequency analysis of ECORS deep reflection seismic data: 53rd Mtg. Eur. Assoc. Expl. Geophys., Abstracts, 384–385.

Richards, J., 1993, Remote sensing digital image analysis, an introduction: Springer-Verlag.

Siliqi, R., 1994, Elaboration du traitement et interpretation d'une expérience de sismique réflexion 3D à l'échelle de la croûte dans les pyrénées (ECORS 2.5D): Ph.D. thesis, Université Louis Pasteur, Strasbourg.

Stark, H., Ed., 1987, Image recovery, theory and application: Academic Press.

Trorey, A., 1970, A simple theory for seismic diffractions: Geophysics, **35**, 762–784.

Yilmaz, O., 1987, Seismic data processing: Soc. Expl. Geophys.