Filtrage par reconnaissance des formes d'un profil sismique vertical

Abdelkader Benaissa, Zahia Benaïssa, Tahar Aifa, Amar Boudella

▶ **To cite this version:**

Abdelkader Benaissa, Zahia Benaïssa, Tahar Aifa, Amar Boudella. Filtrage par reconnaissance des formes d'un profil sismique vertical. Swiss Journal of Geosciences, 2009, 102 (2), pp.297-306. 10.1007/s00015-009-1324-2 . insu-00424567

HAL Id: insu-00424567

https://insu.hal.science/insu-00424567

Submitted on 27 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

1661-8726/09/020297-10
DOI 10.1007/s00015-009-1324-2
Birkhäuser Verlag, Basel, 2009

Swiss J. Geosci. 102 (2009) 297–306

Filtrage par reconnaissance des formes d'un profil sismique vertical
Vertical seismic profile filtering through the use of pattern recognition

Abdelkader Benaïssa [1], Zahia Benaïssa [1], Tahar Aïfa [2,*] & Amar Boudella [1]

Mots clés: reconnaissance des formes, ondelette, valeurs propres, vecteurs propres, matrice d'autocorrélation, PSV, onde P, rapport signal sur bruit, filtrage
Key words: pattern recognition, wavelet, eigenvalues, eigenvectors, autocorrelation matrix, VSP, P wave, signal to noise ratio, filtering

RÉSUMÉ

Pour améliorer le rapport signal sur bruit des enregistrements de profil sismique vertical, nous avons utilisé une méthode de filtrage basée sur la reconnaissance des formes. Elle consiste à reconnaître le long de la trace sismique, correspondant aux arrivées des divers événements, la forme de l'ondelette P considérée comme signal d'apprentissage. Cette reconnaissance s'effectue par des projections qui consistent à ne garder que les signaux dont la forme est identique à l'ondelette P et à atténuer les autres en fonction de leur degré de ressemblance à cette ondelette. L'étude sur des données synthétiques et réelles montre que cette méthode apparaît comme un filtre efficace. Cependant, elle reste tributaire d'un motif d'apprentissage qui doit être bien défini et identifiable.

ABSTRACT

To improve the signal to noise ratio of the vertical seismic profile recordings, we used a filtering method based on pattern recognition. It consists in recognizing along the seismic trace, corresponding to the arrivals of various events, the shape of the P wavelet considered as the training signal. This recognition is made of projections which retain only the signals similar in shape to the P wavelet, the others being attenuated according to their degree of resemblance to the training wavelet. The study undertaken on synthetic and real data shows that this method acts as an effective filter. However, it still depends on a training signal that must be well defined and identifiable.

Abridged English version

1. Introduction

In a practically zero offset vertical seismic profile (VSP), direct and reflected seismic arrivals on the Z component recordings, holding nearly all the information because of the very weak offset, are difficult to interpret when the signal to noise (S/N) ratio is weak. The interpretation can be improved by attenuating the noise component. However, in most cases, the noise is not localized in a given frequency band and can be present in the useful frequency fields, which makes its elimination difficult. Under these conditions, to improve the S/N ratio, processing is carried out which uses all that makes it possible to differentiate the required signal (Yilmaz 1987; Claerbout 1992; Mari et al. 2001; Corso et al. 2003). In the case of zero offset VSP vertical component, we are only dealing with the P waves. We suggest, in this study, to characterize the required signal (P wave) by its wavelet shape. The applied filtering method is based on pattern recognition. It consists in recognizing, along the seismic trace, the signal shape associated with the P wave.

2. Methodology

In pattern recognition applied to seismic, the object represents the seismic trace and the concept the wave we wish to extract or follow along this trace. The training stage aimed at recognizing a set of parameters or characteristics quite representative of the required signal shape. The wavelet corresponding to the arrivals of the signal to be extracted is used, in our case, as the training signal. It is identified with a vector of R^N. The autocorrelation matrix of this vector has N eigenvectors associated with N real and non negative eigenvalues. These eigenvectors constitute an orthonormal basis (Karhünen-Loeve basis) of signal space in which the training vector breaks up. The addition of standardized eigenvalues, classified by descending order, quickly tends towards the maximum percentage (100%) when one considers only the p first eigenvalues. Eigenvectors, corresponding to these p first eigenvalues, generate a vectorial

[1] Université des Sciences et de la Technologie Houari Boumediène, Faculté des Sciences de la Terre, BP 32, El-Alia, Bab-Ezzouar, 16000 Alger, Algérie.
[2] Géosciences-Rennes, CNRS UMR6118, Université de Rennes1, Bat.15, Campus de Beaulieu, 35042 Rennes cedex, France.
*Correspondance: E-mail: tahar.aifa@univ-rennes1.fr

subspace of p dimension, called eigensignals subspace (Cliet et al. 1986; Belaïd & Belaïd 1992; Ciarlet 2002; Lengellé 2002; Brezinski & Redivo-Zaglia 2004). The stage of interpretation and decision consists in a projection of the seismic trace on the subspace generated by the p eigensignals, in order to leave in this trace only what is connected with the training signal (Michaud 1985). We developed a filtering algorithm based on pattern recognition (Figure 1), and then wrote a corresponding Fortran code. This code requests subprograms for the input data to 1, 2 or 3 components, among which a Jacobi subprogram is used to compute eigenvalues and eigenvectors in descending order. As an example, using a Ricker wavelet as training signal (Figure 2), the correlation matrix of such a signal is of 21×21 dimension. The 21 eigenvalues of this matrix, sorted by descending order, vs. their amplitude, show that the first twelve values prevail (Figure 3). Tests were run to establish the optimum parameters and led to adopting the value of 90% for the projection threshold (Figures 3, 4), available for every sampling frequency used in well log VSP type. The projection threshold p is a significant parameter in the reproduction, after projection, of the training signal. Threshold choice directly influences projection subspace dimension.

3. Application to well seismic data zero offset VSP

3.1. Synthetic data

To show qualitatively how the non stationarity, in time and space, of the signal is apprehended by pattern recognition filtering, VSP data synthesized from a three-layered geological model is considered. The attenuated VSP is calculated in the frequency domain (Ganley 1981; Lavergne 1986) and then convoluted with 30 Hz centred Ricker wavelet (Figure 5a). The results obtained after application of the pattern recognition method, with the Ricker wavelet as training signal and a projection threshold of 90%, show a correct reproduction of the initial VSP. We added to the synthetic data a centred Gaussian random noise; the various events are recovered and the wavelet shape preserved (Figures 5b, c, d).

3.2. Real data

Data used in this application are a zero offset VSP vertical component, acquired by "Baker Atlas Logging Algeria Services Company". The aim of the application is to filter these data (Figure 6a) using pattern recognition of the P wave. The P training signal is directly extracted from the first arrivals. Projection vectorial subspace is generated by the 9 eigenvectors of the P signal autocorrelation matrix. The result of filtering by projection of the vertical component, on the created subspace, shows an improved vertical resolution with a good S/N ratio in the higher part of the filtered VSP (Figure 6a, b). Most of the events in the lower part underwent attenuation, according to their degree of resemblance to the training wavelet (see Eq. 3) (Figure 7). The higher part of the corridor stacks resulting from horizontalized and deconvoluted upgoing wave fields, with and without filtering using pattern recognition (Figures 6c, c', d, d'), which is constituted mainly of primary upgoing events because of the narrow summation corridor, presents a better S/N ratio after the use of a pattern recognition filter. All events are correctly extracted. Beyond 2400 ms, several events are attenuated. In this part of the VSP log, summation is carried out in a window equal to the length of traces 1 to 7. Consequently, it can contain upgoing multiples (Hardage 2000; Mari & Coppens 2000) which we could identify as those which were attenuated after the use of a pattern recognition filter. The preserved events would correspond to primary upgoing reflections resulting from interfaces below the well.

4. Discussion and conclusions

The pattern recognition gave promising results. The threshold represents the confidence level allowed to eigenvectors to define this subspace. Signals are attenuated according to their degree of resemblance to the training signal. Deconvolution, which amplifies the noise, gave better results on frequency filtered data then after pattern recognition, suggesting hence that pattern recognition robustly filtered the data (Figures 6c, d). During a frequency filtering, if the noise is not localized in a given frequency band, its elimination becomes difficult. Moreover, a bad estimation of the frequency band leads to an insufficiently improved S/N ratio. In pattern recognition, the useful signal frequency band comes from the training wavelet. Within sight of the obtained results, pattern recognition appears as an effective filter which requires *a priori* neither on the frequency band of the received signal, nor on the wavelet phase. However, it remains dependent on the training signal which must be well defined and identifiable. This method can be applied in surface seismic to improve the resolution at great depths or each time one wishes to extract a type of waves well differentiated by its shape.

1. Introduction

En sismique de puits, de type profil sismique vertical (PSV) à déport nul, les arrivées sismiques directes et réfléchies sur les enregistrements de la composante Z, détentrice de la quasi-totalité de l'information à cause du très faible offset, sont délicates à interpréter quand le rapport signal sur bruit (S/B) est faible. Pratiquement, les données sismiques acquises peuvent être subdivisées en deux espaces: 1) un espace signal constitué des arrivées de type P et 2) un espace bruit renfermant les évènements parasites formés par les modes guidés (ou ondes de tube ou de Stoneley, liées à la présence du puits et de la boue), les résidus d'ondes converties, les multiples, etc. Dans les cas les plus favorables, ces évènements parasites correspondent à des hautes fréquences dans le signal reçu, ce qui permet de les éliminer facilement par des techniques de filtrage. Mais dans la plupart des cas, ces bruits, aléatoires ou

cohérents, ne sont pas localisés dans une bande de fréquence donnée et peuvent donc être inclus dans le domaine des fréquences utiles, ce qui rend leur élimination difficile par filtrage fréquentiel. Dans ces conditions, pour améliorer le rapport S/B, on passe, en général, par un traitement optimal qui utilise tout ce qui permet de différencier le signal recherché de l'espace bruit tels que la fonction de transfert, le temps d'arrivée, la bande passante, l'ondelette émise, etc. (Yilmaz 1987; Claerbout 1992; Mari et al. 2001; Corso et al. 2003). Dans le cas de la composante verticale d'un PSV à déport nul, le signal utile recherché est représenté par un seul type d'ondes, les ondes P.

Comme, dans le cas d'un PSV à déport nul, l'impulsion émise se propage et se réfléchit dans les séries géologiques en ne subissant que de faibles distorsions, on se propose donc de caractériser le signal recherché (onde P) par la forme de l'ondelette qui lui est associée.

La méthode de filtrage que nous utilisons est basée sur la reconnaissance des formes. Elle consiste à reconnaître, le long de la trace sismique issue de l'enregistrement de la composante Z d'un PSV à déport nul, la forme du signal associé à l'onde P. Cette méthode identifie tout signal composé de N échantillons à un vecteur de dimension N de l'espace vectoriel R^N. Elle consiste en une projection des signaux sur le sous-espace vectoriel de R^N, engendré par les premiers vecteurs propres (signaux propres) d'une matrice construite à partir d'un signal d'apprentissage (onde P), appelée matrice d'autocorrélation. Appliquée sur des enregistrements sismiques à trois composantes, cette technique permet également de filtrer les différents types d'ondes polarisées (Michaud 1985; Cliet et al. 1986).

2. Méthodologie

La reconnaissance des formes, utilisée dans différents domaines (médecine, électronique, linguistique [reconnaissance vocale], etc.) (Lengellé 2002), a pour but de reconnaître, parmi un ensemble de formes, une forme bien distincte que l'on désire extraire.

La structure générale d'un système de reconnaissance des formes est constituée de deux étapes: 1) une étape d'apprentissage dans laquelle on choisit un certain nombre de paramètres bien représentatifs du modèle à extraire (modèle d'apprentissage ou concept) et 2) une étape d'interprétation et de décision, dans laquelle le modèle donné est extrait de l'ensemble (objet) (Belaïd & Belaïd 1992). Appliquée à la sismique, l'objet représente la trace sismique et le concept, l'onde que l'on veut extraire ou suivre dans cette trace. Le principe de la méthode proposée ici consiste à ne garder, le long de la trace, que les signaux dont la forme est identique ou proche de celle de l'ondelette d'apprentissage. La sélection se fait par projection du signal objet sur un sous-espace créé à partir du signal d'apprentissage. Tout signal qui ressemble au signal d'apprentissage se projette bien et se retrouve donc conservé. Au contraire, un signal qui diffère de l'apprentissage se projette mal et est donc atténué ou éliminé.

2.1. Étape d'apprentissage

L'ondelette correspondant à l'arrivée du signal à extraire est appelé signal d'apprentissage. Ce dernier, composé de N échantillons, est représenté par un vecteur $\vec{S}$. La matrice d'autocorrélation de ce vecteur possède N vecteurs propres associés à N valeurs propres réelles positives ou nulles λ_i (Brezinski & Redivo-Zaglia 2004). Une valeur propre représente le terme pondérateur de la quantité d'informations ou d'énergie contenue dans le signal d'apprentissage et portée par le vecteur ou signal propre $\vec{V}_i$. Normalisée, la valeur propre représente le pourcentage, ou l'importance du signal propre, associé dans la description du signal d'apprentissage (Michaud 1985).

Les vecteurs propres forment une base orthonormée B_{KL} (base de Karhünen-Loeve) de l'espace des signaux. Le vecteur $\vec{S}$ se décompose dans cette base sous la forme:

$$\vec{S} = \sum_{i=1}^{N} (\vec{S} \cdot \vec{V}_i)\, \vec{V}_i$$

L'addition des valeurs propres normalisées λ_i', classées par ordre décroissant, tend rapidement vers la valeur maximale (100%) en considérant seulement les p premières valeurs propres. Ces dernières suffisent donc à décrire correctement le signal d'apprentissage avec une probabilité d'erreur P_r tel que:

$$P_r = 1 - \sum_{i=1}^{p} \lambda_i' \qquad (1)$$

Les vecteurs propres, correspondant à ces p premières valeurs propres et qui sont linéairement indépendants et orthonormés, engendrent un sous-espace vectoriel de dimension p, appelé sous-espace des signaux propres (Ciarlet 2002).

2.2. Étape d'interprétation et de décision

Il s'agit d'une projection de la trace sismique sur le sous-espace engendré par les p signaux propres, de façon à ne conserver dans cette trace que les signaux s'apparentant au signal d'apprentissage (Michaud 1985). En se fixant un seuil de confiance s, p sera le nombre minimal de valeurs propres normalisées λ_i' qui vérifie la condition suivante:

$$\sum_{i=1}^{p} \lambda_i' \geq s \qquad (2)$$

A tout signal $\vec{U}$ représenté dans l'espace des signaux de dimension N, on associe sa projection $\vec{U}_P$ tel que: $\vec{U}_P = P\vec{U}$ où P est la matrice de projection: $P = V_P V_P^t$. La matrice V_P^t possède des colonnes portant les vecteurs propres orthogonaux entre eux. La proximité géométrique β de ces deux vecteurs traduit la ressemblance formelle entre les deux signaux:

$$\beta = \frac{\left\| \vec{U}_P \right\|}{\left\| \vec{U} \right\|} \qquad (3)$$

Ainsi plus β est proche de 1, plus le signal ressemble à l'apprentissage. La projection du signal $\vec{U}$ sur le sous-espace vectoriel, engendré par les p premiers signaux propres de

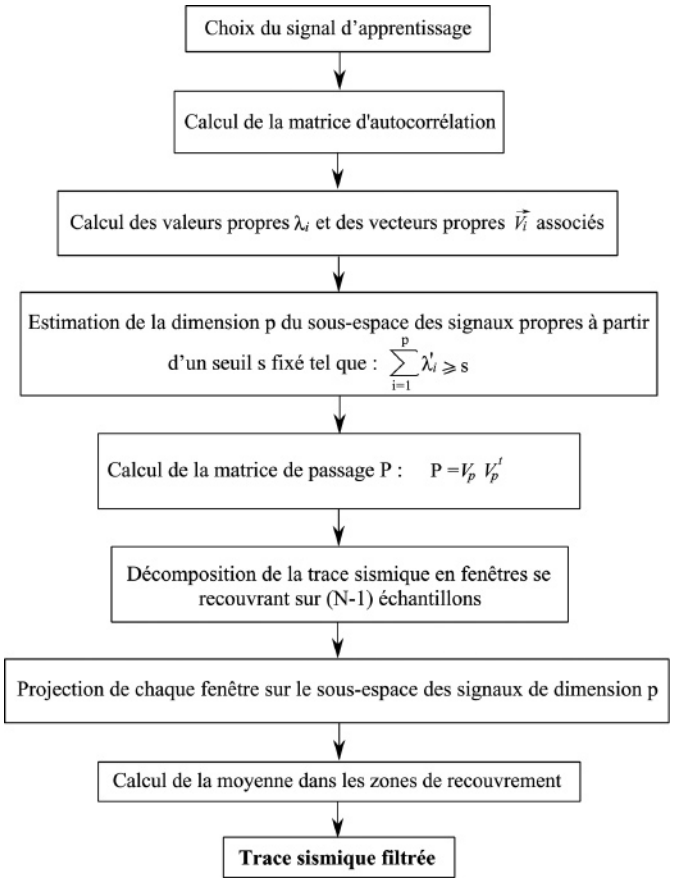

Fig. 1. Représentation de l'algorithme de filtrage utilisant la reconnaissance des formes.
Fig. 1. Representation of the filtering algorithm using pattern recognition.

la matrice d'autocorrélation, revient à n'en conserver que l'information qu'il partage avec le signal d'apprentissage $\vec{S}$, si elle existe. Ainsi, si l'on considère une trace sismique comme une suite de signaux U(t) de longueur N échantillons, se recouvrant chacun sur (N – 1) échantillons, sa projection sur le nouveau sous-espace revient donc à projeter séparément les signaux consécutifs qui la constituent. La trace projetée est reconstruite par une opération de moyenne arithmétique dans les zones de recouvrement (Michaud 1985). Nous avons développé un algorithme de filtrage par reconnaissance des formes (Figure 1), puis écrit un code Fortran correspondant. Ce code fait appel à des sous-programmes pour des données d'entrée à 1, 2 ou 3 composantes, dont le sous-programme Jacobi pour le calcul des valeurs et vecteurs propres dans l'ordre décroissant.

2.3. Le paramètre seuil

Le seuil de projection, s, est un paramètre important dans la reproduction, après projection, du signal d'apprentissage. Le choix du seuil influe directement sur la dimension du sous-espace de projection, p. A titre d'exemple, en considérant comme signal d'apprentissage l'ondelette de Ricker de fréquence centrale 30 Hz, de longueur 80 ms avec un pas d'échantillonnage de 4 ms (Figure 2), la matrice d'autocorrélation de ce signal est de dimension 21 × 21. Les 21 valeurs propres de cette matrice, rangées par ordre décroissant, en fonction de leur amplitude, montrent que les 12 premières valeurs sont dominantes (Figure 3). Elles contiennent, par conséquent, la quasi-totalité de l'information relative au signal d'apprentissage. Pour plusieurs valeurs du seuil de projection, s, et par conséquent de la dimension du sous-espace de projection, p (cf Eq. 2) (Figure 3), la reproduction du Ricker est effective

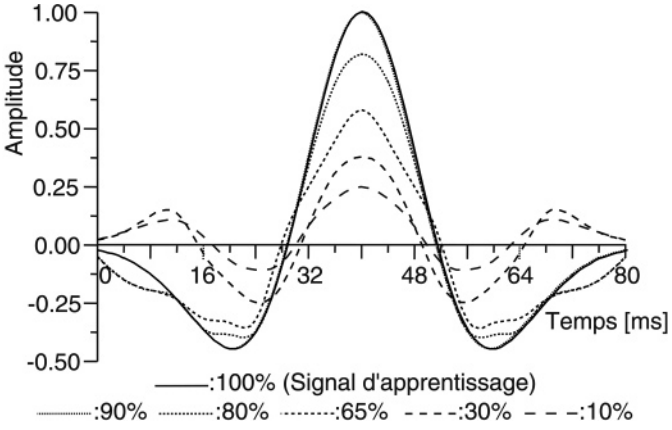

Fig. 2. Signal d'apprentissage (Ricker de fréquence centrale 30 Hz) et sa reproduction après projection avec différentes valeurs du seuil (voir Figure 3). N = 21.
Fig. 2. Training signal (of 30 Hz mid-band frequency Ricker) and its reproduction after projection, with various threshold's values (see Figure 3). N = 21.

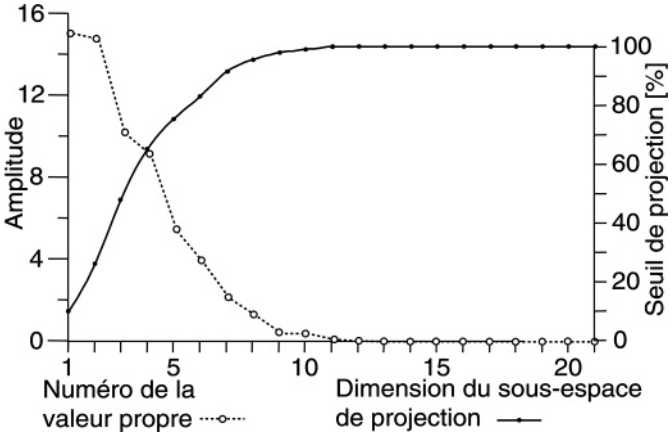

Fig. 3. Variation de la dimension du sous-espace des signaux propres correspondant à la matrice d'autocorrélation du signal d'apprentissage (voir Figure 2) en fonction du seuil de projection. Classement par ordre décroissant des valeurs propres de la matrice d'autocorrélation du signal d'apprentissage.
Fig. 3. Variation of the eigensignals subspace dimension corresponding to the autocorrelation matrix of the training signal (see Figure 2) vs. projection threshold. Sorting by decreasing order of the autocorrelation matrix eigenvalues of the training signal.

(Figure 2). On peut donc distinguer son rapprochement du signal d'apprentissage, au fur et à mesure que la valeur du seuil augmente. Une reproduction fidèle du Ricker n'est atteinte que pour des valeurs de seuil supérieures ou égales à 90%. Il apparaît ainsi qu'un seuil de 90%, équivalent à une dimension égale à 7 du sous-espace de projection (Figure 3), suffit pour reproduire correctement, après projection, le signal d'apprentissage.

La figure 4 présente les résultats du filtrage par projection, avec un seuil de 90%, d'une trace sismique synthétique constituée de trois arrivées Ricker de fréquence centrale 30 Hz et d'amplitudes différentes (Figures 4a, a'), bruitée à 10% (Figures 4b, b') et 50% (Figures 4e, e'). Les traces obtenues par projection sur la base propre du Ricker (Figures 4d, g) montrent une bonne extraction des trois ondelettes dans les deux cas. Ces résultats sont comparables à ceux obtenus après application d'un filtre en fréquence trapézoïdal [2–10, 50–80] Hz (Figures 4c, f), avec cependant un rapport S/B légèrement supérieur. L'examen des spectres en amplitude correspondants (Figures 4c', d', f', g') montre, en effet, que le filtrage par re-

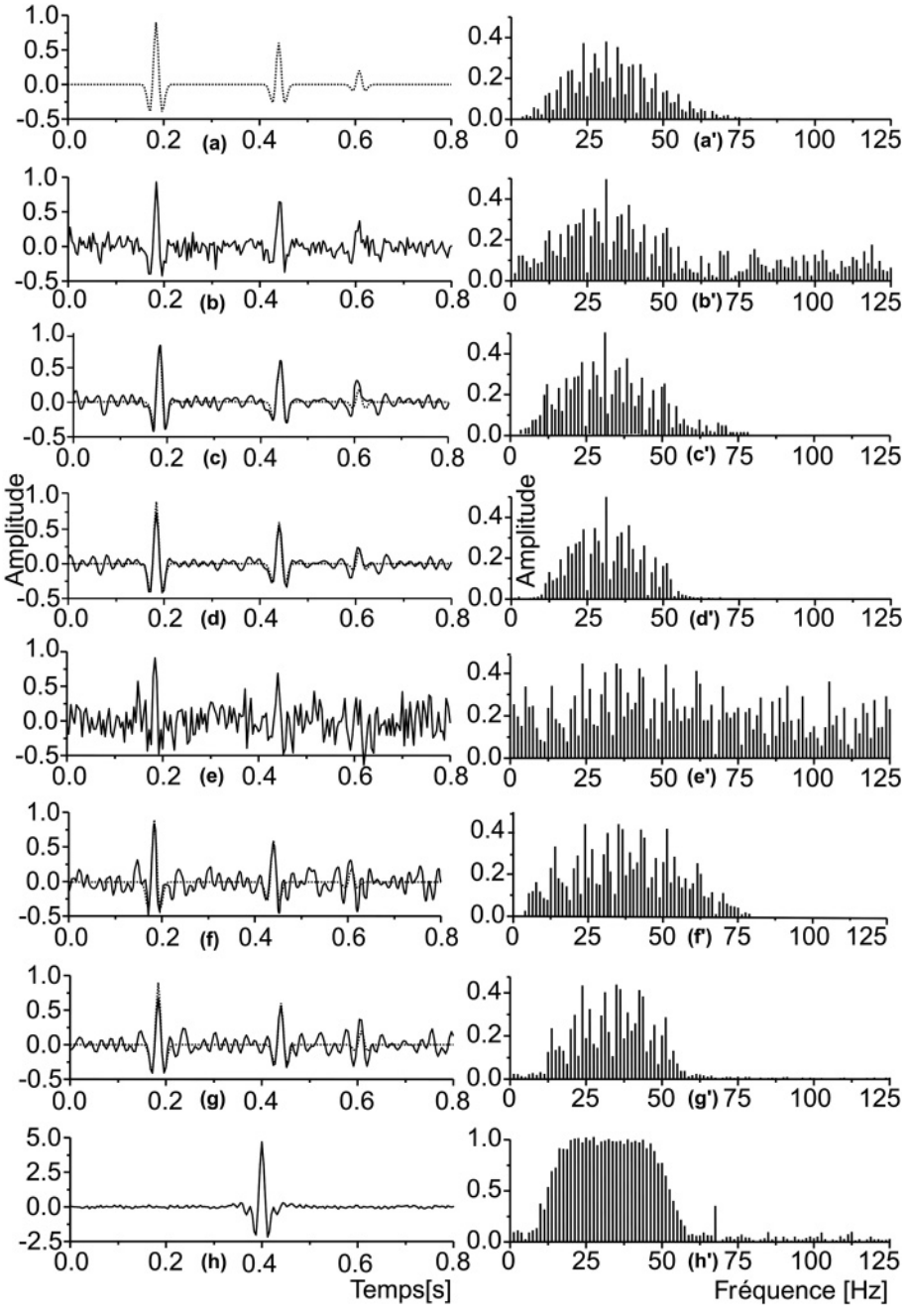

Fig. 4. Exemple de filtrage par reconnaissance des formes avec un seuil de projection de 90%; (a)- trace initiale (reproduite en tiretés sur les traces filtrées); (a')- spectre en amplitude de la trace (a); (b)- trace initiale (a) avec 10% de bruit; (b')- spectre en amplitude de la trace (b); (c)- trace (b) après filtrage en fréquence trapézoïdal [2–10, 50–80] Hz; (c')- spectre en amplitude de la trace (c); (d)- trace (b) après filtrage par reconnaissance des formes (l'apprentissage est l'ondelette de Ricker); (d')- spectre en amplitude de la trace (d); (e)- trace initiale (a) avec 50% de bruit; (e')- spectre en amplitude de la trace (e); (f)- trace (e) après filtrage en fréquence trapézoïdal [2–10, 50–80] Hz; (f')- spectre en amplitude de la trace (f); (g)- trace (e) après filtrage par reconnaissance des formes (l'apprentissage est l'ondelette de Ricker); (g')- spectre en amplitude de la trace (g); (h)- réponse impulsionnelle du filtre par reconnaissance des formes; (h')- fonction de transfert du filtre par reconnaissance des formes.

Fig. 4. Pattern recognition filtering example with a projection threshold of 90%; (a)- initial trace (reproduced in dashed lines on the filtered traces); (a')- amplitude spectrum of trace (a); (b)- initial trace (a) with 10% of noise; (b')- amplitude spectrum of trace (b); (c)- trace (b) after trapezoidal filtering [2–10, 50–80] Hz; (c')- amplitude spectrum of trace (c); (d)- trace (b) after filtering using pattern recognition (training signal is Ricker's wavelet); (d')- amplitude spectrum of trace (d); (e)- initial trace (a) with 50% of noise; (e')- amplitude spectrum of trace (e); (f)- trace (e) after trapezoidal filtering [2–10, 50–80] Hz; (f')- amplitude spectrum of trace (f); (g)- trace (e) after filtering using pattern recognition (training signal is Ricker's wavelet); (g')- amplitude spectrum of trace (g); (h)- impulse response of pattern recognition filter; (h')- transfer function of pattern recognition filter.

connaissance des formes a été plus sévère, dans ce cas, pour les fréquences inférieures à 10 Hz et celles supérieures à 50 Hz. La réponse impulsionnelle ainsi que le module de la fonction de transfert de ce filtre sont présentés respectivement en figures 4 h et 4h'.

Des tests effectués, en faisant varier la fréquence centrale et le pas d'échantillonnage, ont montré que ce seuil reste suffisant dans tous les cas. Dans la littérature, bien souvent, ce seuil de projection est fixé à 90%, valeur que nous utilisons par la suite dans ce travail.

3. Application au filtrage de données de sismique de puits type PSV à déport nul

3.1. Données synthétiques

Pour montrer qualitativement comment la non stationnarité, dans le temps et dans l'espace, de l'ondelette recherchée est appréhendée par le filtrage par reconnaissance des formes, nous considérons des données PSV synthétisées à partir d'un modèle géologique à 3 couches. Les vitesses sont fixées à 2.5, 3 et 3.5 km/s et les épaisseurs à 800 m, 700 m et l'infini pour les couches 1, 2 et 3 respectivement. Pour introduire le facteur de distorsion, facteur qui rapprocherait des conditions réelles de terrain, le PSV impulsionnel est calculé dans le domaine fréquentiel (Ganley 1981) en tenant compte de l'atténuation du signal due aux phénomènes d'absorption et de dispersion. Pour simuler une forte atténuation, nous avons fixé le facteur de qualité Q, qui caractérise l'aptitude que possède un milieu à propager une onde et qui est lié au coefficient d'absorption α par la relation suivante:

$$\alpha = \frac{|\omega|}{2 \cdot C \cdot Q} \tag{4}$$

où $|\omega|$ est la valeur absolue de la pulsation angulaire et C la vitesse de phase (Ganley 1981; Lavergne 1986), à 20, 40 et 80 dans les couches 1, 2 et 3 respectivement. Ce PSV impulsionnel est ensuite convolué à l'ondelette de Ricker, de fréquence centrale 30 Hz, que nous avons choisi pour représenter la signature de la source. Il est composé de 34 traces s'échelonnant de 50 m à 1700 m de profondeur, avec un pas de 50 m. La longueur de l'enregistrement sismique est restreinte à 2000 ms, avec un pas d'échantillonnage de 4 ms. Le PSV ainsi calculé (Figure 5a), après filtrage par reconnaissance des formes utilisant l'ondelette de Ricker comme signal d'apprentissage, est restitué en dépit des distorsions subies par l'ondelette, dues aux phénomènes d'absorption et d'interférence (Figure 5b). Même en présence d'un taux de bruit de 50% (Figure 5c), le signal est correctement extrait avec le même signal d'apprentissage que précédemment et le rapport S/B amélioré (Figure 5d). Ce taux est évalué à partir de la puissance du signal sans bruit. Les enregistrements PSV sont représentés après égalisation sur une fenêtre de longueur 250 ms (Yilmaz 1987).

3.2. Données réelles

Les données utilisées dans cette application sont extraites de la composante verticale d'un PSV à déport nul, acquis par la compagnie "Baker Atlas Logging Services (Algeria)". La source, positionnée à 60 m de la gueule du puits, est un Vibroseis à deux vibrateurs, générateur d'ondes P et le signal émis est un sweep [8–72] Hz, avec un balayage de 12 s. Les traces, de longueur 4000 ms, ont été enregistrées avec un pas d'échantillonnage de 2 ms sur 110 côtes et un pas de remontée de la sonde variable d'une côte à l'autre. La côte la plus proche de la surface du sol est de 1280 m, la plus profonde, associée à la trace 1, est de 3470 m. L'enregistrement de la composante Z de ce PSV, après filtrage en fréquence et égalisation (Figure 6a), est composé essentiellement d'ondes de volume de type P. S'y ajouteraient des résidus d'ondes de volume de Stoneley insuffisamment atténuées par filtrage médian et des résidus d'ondes de volume de type S, en raison du déport de la source certes faible mais suffisant pour générer des ondes SV de conversion. Ces résidus sont considérés dans la suite comme faisant partie du bruit. Le but de cette application est de filtrer cette composante verticale par la reconnaissance de la forme de l'onde P (filtrage en ondes P).

La source en ondes P utilisée pour l'acquisition de cette composante est un Vibroseis, source répétitive par excellence. L'autocorrélation du sweep émis constituerait donc le signal d'apprentissage pour l'ensemble des traces du PSV. Mais, dans le but d'avoir un motif d'apprentissage le plus proche possible de celui du signal émis, nous avons extrait le signal d'apprentissage P directement des arrivées premières du PSV, étant donné que celles-ci ne contiennent que des ondes P. En effet, le signal émis subit de fortes distorsions lors de son passage dans la zone altérée. Par conséquent, l'ondelette P prélevée au niveau de l'arrivée première sur la trace, après corrélation, n'est pas celle de l'autocorrélation du sweep mais celle de la corrélation entre le sweep émis et le signal à la sortie de la zone altérée, plus représentative du signal réellement envoyé dans le sous-sol. L'idéal serait de considérer un signal d'apprentissage par trace afin de tenir compte des conditions de trou variables à chaque côte. Dans le cas présent, nous nous sommes restreints au signal extrait de la trace 15, correspondant au temps 1008 ms et de pseudo-période 52 ms, choisi parmi les 110 arrivées premières comme ondelette d'apprentissage de l'ensemble des traces, pour son motif peu bruité. Le pas d'échantillonnage étant égal à 2 ms, la dimension de l'espace des signaux est donc fixée à 26. La matrice d'autocorrélation correspondante possède 26 valeurs propres. Les 14 premières valeurs propres normalisées sont dominantes. Le seuil de 90% est atteint à la 9ème valeur (cf Eq. 2). Le sous-espace vectoriel de projection est donc de dimension 9 et la probabilité d'erreur d'environ 10% (cf Eq. 1). Le résultat du filtrage par projection de la composante verticale sur le sous-espace créé (Figure 6b) montre une résolution verticale améliorée et un bon rapport S/B visible dans la partie supérieure du PSV filtré. La majorité des évè-

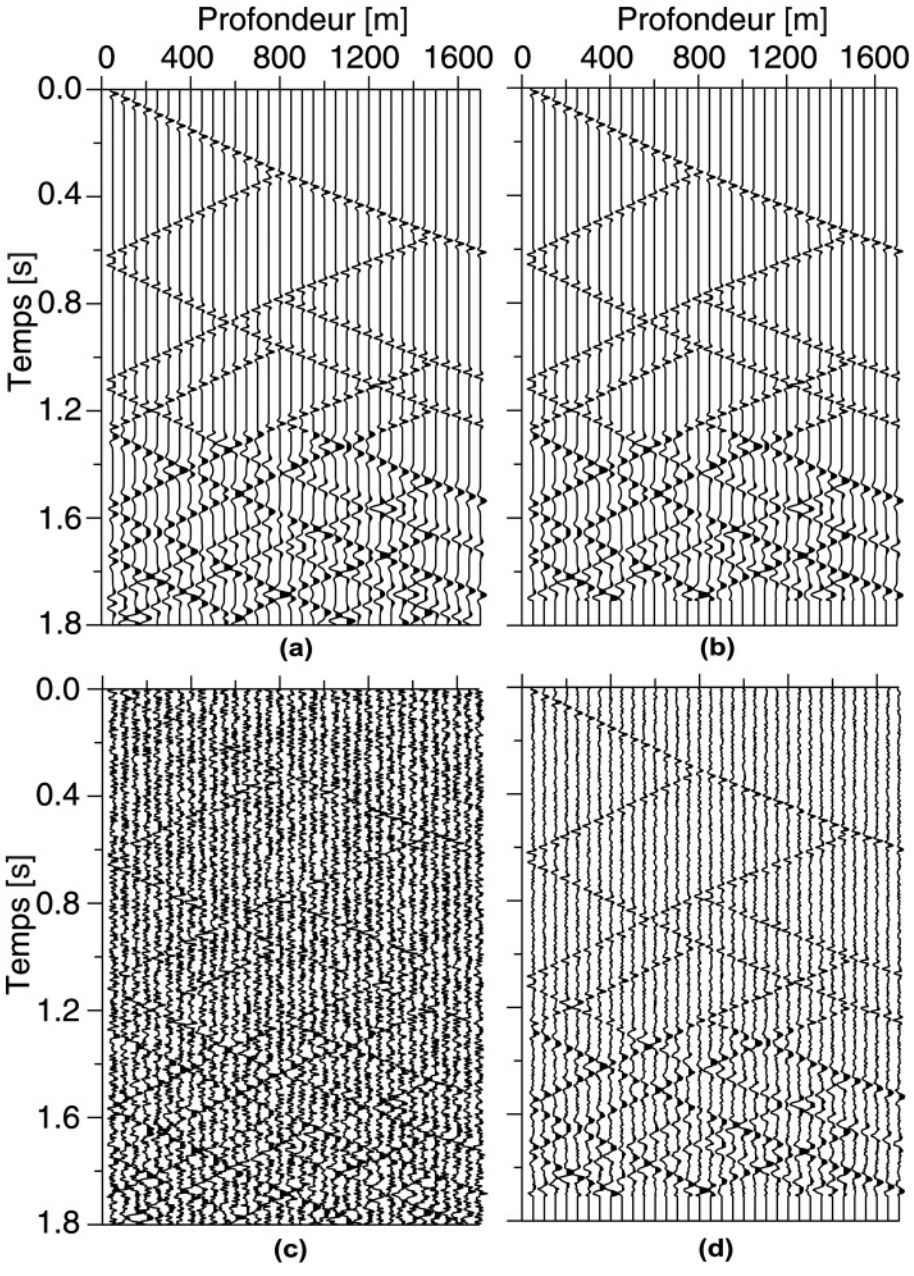

Fig. 5. Filtrage par reconnaissance des formes des données PSV synthétiques atténuées et en présence de bruit; (a)- PSV synthétique calculé; (b)- PSV synthétique (a) après filtrage par reconnaissance des formes; (c)- PSV synthétique (a) avec 50% de bruit; (d)- PSV synthétique (c) après filtrage par reconnaissance des formes.
Fig. 5. Filtering using pattern recognition of attenuated synthetic VSP data with noise; (a)- calculated synthetic VSP; (b)- synthetic VSP (a) after filtering using pattern recognition; (c)- synthetic VSP (a) with 50% of noise; (d)- synthetic VSP (c) after filtering using pattern recognition.

nements de la partie inférieure a, en revanche, subi une forte atténuation. L'onde descendante directe est préservée et son amplitude renforcée. Les évènements multiples descendants, parallèles à l'arrivée directe, sont bien extraits entre 500 et 2000 ms car ils ressemblent encore au signal d'apprentissage en raison des faibles trajets parcourus. Au delà, ils sont de plus en plus atténués par la projection car leurs formes diffèrent de plus en plus de l'apprentissage à cause des longs trajets parcourus. Il en est de même pour les évènements multiples montants. Les ondes montantes primaires, seules ondes montantes qui intersectent les arrivées directes, sont mises en évidence en raison de l'atténuation des résidus d'ondes S et de tube qui subsistent après traitement spécifique, et des bruits aléatoires stationnaires divers. L'atténuation est fonction du degré de ressemblance du signal au motif d'apprentissage. Plus le signal diffère de l'apprentissage, plus il est atténué. Par exemple, en figure 7, le signal d'apprentissage (indiqué par une flèche noire verticale en figures 6a, b) est bien reproduit après filtrage (Figure 7a), avec une proximité géométrique (cf Eq. 3) de 0.97. Néanmoins, le signal prélevé à 1210 ms sur la trace 108 (indiqué par une flèche blanche horizontale en figures 6a, b), dont la forme diffère de l'apprentissage, est atténué par l'opération de filtrage; la proximité géométrique n'est plus que de 0.41 (Figure 7b).

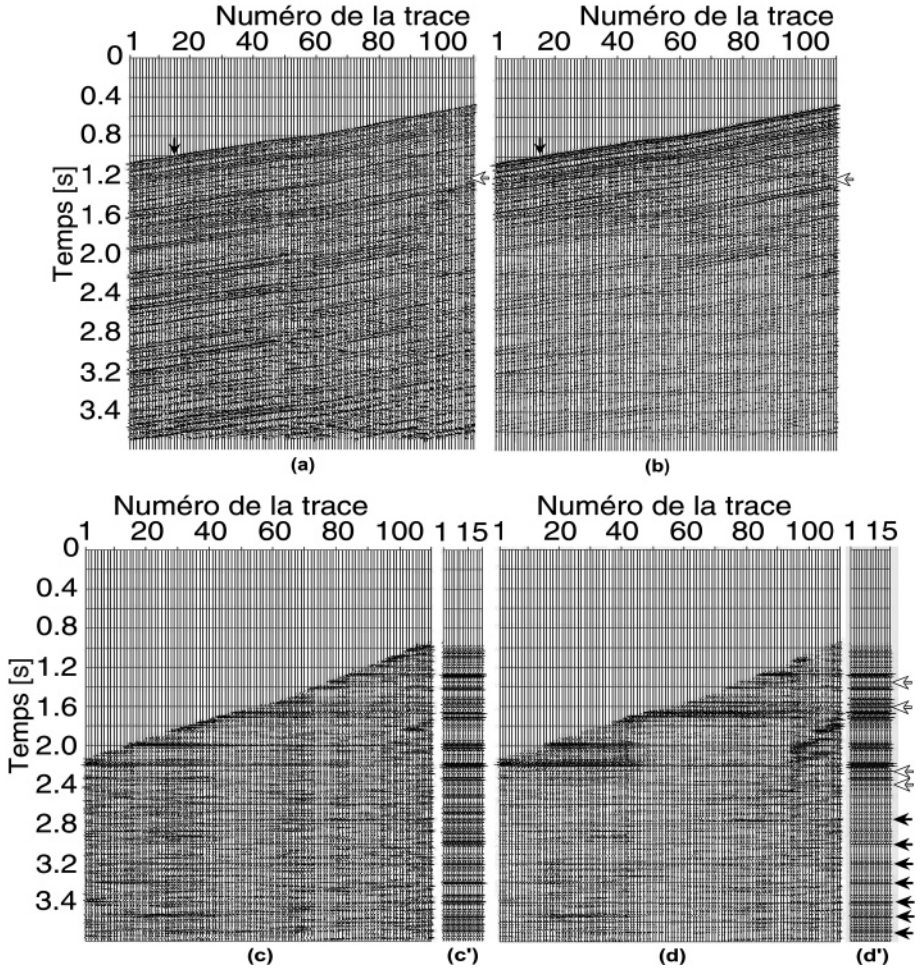

Fig. 6. Filtrage en ondes P, par reconnaissance des formes, d'un PSV réel; (a)- PSV réel (composante Z); données initiales filtrées en fréquence et égalisées; (b)- PSV (a) après filtrage par reconnaissance des formes; (c)- section en ondes montantes, horizontalisées et déconvoluées, du PSV (a); (c')- trace somme correspondant à (c); (d)- section en ondes montantes, horizontalisées et déconvoluées, du PSV filtré (b); (d')- trace somme correspondant à (d).

Fig. 6. P waves filtering using pattern recognition of a real VSP; (a)- real VSP (Z component); initial data filtered in frequency and equalized; (b)- VSP (a) after filtering using pattern recognition; (c)- horizontalized and deconvoluted upgoing waves section of the VSP (a); (c')- stack corridor corresponding to (c); (d)- horizontalized and deconvoluted upgoing waves section of the VSP (b); (d')- corridor stack corresponding to (d).

Dans le cas du PSV, le signal désiré est le champ des ondes montantes. Il est extrait du PSV par filtre médian, déconvolué par le champ des ondes descendantes pour s'affranchir du signal source et des multiples descendants puis horizontalisé (Figures 6c, d). Les ondes montantes sont ensuite sommées dans un couloir suivant immédiatement les arrivées premières, en empruntant à chaque trace les soixante premiers échantillons, en-dessous de la position du géophone de puits au-delà de la première arrivée, sauf pour les traces 1 à 7 où l'ensemble de la trace est considéré. On distingue une nette amélioration sur la trace somme obtenue avec les données filtrées (Figure 6d') par rapport aux données non filtrées (Figure 6c'). La partie supérieure du film (entre 1000 et 2400 ms), constituée essentiellement d'évènements montants primaires en raison de l'étroitesse du couloir de sommation, présente un meilleur rapport S/B après filtrage par reconnaissance des formes. Les évènements sont tous correctement extraits, avec renforcement de l'amplitude des évènements proches de la surface (indiqués par des flèches blanches sur la figure 6d'). Au-delà de 2400 ms, plusieurs évènements ont été atténués. Dans cette partie du log PSV, la sommation est réalisée non pas dans un couloir étroit mais sur une fenêtre égale à la longueur des traces 1 à 7. Par conséquent, elle peut renfermer des multiples montants (Hardage 2000; Mari & Coppens 2000) que l'on pourrait identifier comme ceux atténués après filtrage par reconnaissance des formes, car, à ces profondeurs, les signaux relatifs à ces événements, à cause des longs trajets parcourus, sont fortement distordus malgré le traitement appliqué aux données pour tenter de corriger les effets dus à l'absorption, la divergence sphérique, etc. Les évènements préservés (indiqués par des flèches noires en figure 6d') correspondraient à des réflexions montantes primaires issues d'interfaces en-dessous du puits.

4. Discussion et conclusions

La méthode de filtrage par reconnaissance des formes consiste: 1) à garder, par projections successives, le long de la trace sismique, uniquement les arrivées dont la forme ressemble à celle du signal d'apprentissage et 2) à rejeter ou atténuer relativement les autres. Elle a donné des résultats prometteurs moyennant certains réglages de paramètres. Le seuil qui fixe la dimension du sous-espace de projection, pris égal à 90%, valeur confirmée

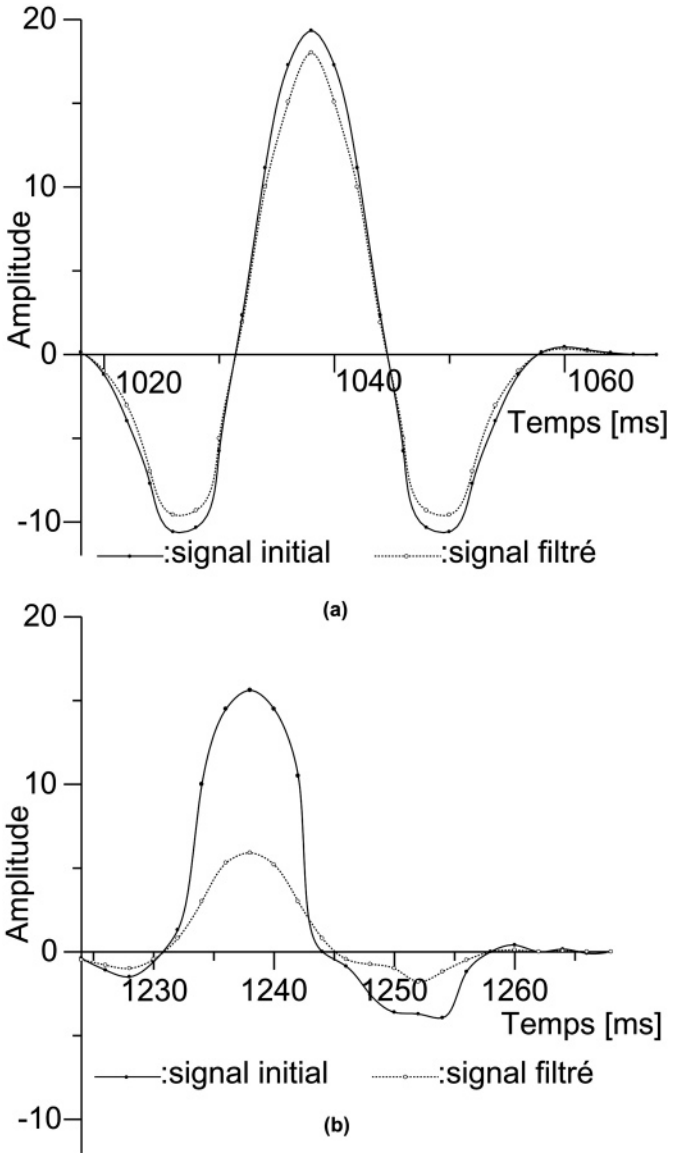

Fig. 7. Deux signaux du PSV réel, avant et après filtrage par reconnaissance des formes; (a)- le signal d'apprentissage P enregistré sur la trace 15 à 1008 ms; (b)- signal enregistré sur la trace 108 à 1210 ms.

Fig. 7. Two signals of the real VSP, before and after filtering using pattern recognition; (a)- training signal P recorded on trace 15 at 1008 ms; (b)- recorded signal on trace 108 at 1210 ms.

par l'étude sur les données synthétiques, représente le degré de confiance attribué aux vecteurs propres qui définissent ce sous-espace.

Les signaux sont atténués en fonction de leur degré de ressemblance au signal d'apprentissage. En sismique de puits type PSV, les arrivées utiles P primaires subissent, en général, de faibles distorsions à cause, d'une part, des rais qui traversent une seule fois la zone altérée et, d'autre part, de la proximité du géophone de puits et du réflecteur (en sommant dans le couloir, on ne considère que les réflexions provenant de réflecteurs juste en-dessous de la position d'ancrage du géophone). Elles sont donc correctement extraites par reconnaissance de la forme de l'arrivée première. L'effet des signaux parasites dont la forme diffère fortement du motif d'apprentissage est annihilé. L'étude qualitative menée sur les données synthétiques a montré que le signal d'apprentissage est bien extrait même en présence d'un bruit élevé (50%).

Le choix du motif d'apprentissage est très important. Il conditionne la qualité des résultats. Si la source est répétitive, un signal d'apprentissage pour toutes les traces du profil suffit. Dans le cas d'une source non répétitive, comme la dynamite, un apprentissage par trace, issu de l'enregistrement du géophone de surface ou de l'arrivée première de la trace, est nécessaire.

Après déconvolution, le PSV filtré en fréquence puis par reconnaissance des formes (Figure 6d) présente un rapport S/B plus élevé, comparé au PSV filtré en fréquence seulement (Figure 6c). La déconvolution, qui amplifie le bruit même très faible sur la trace, a donné de meilleurs résultats sur les données filtrées en fréquence puis par reconnaissance des formes. La reconnaissance des formes aurait donc davantage filtré les données. Lors d'un filtrage en fréquence, si le bruit n'est pas localisé dans une bande de fréquence donnée, son élimination devient difficile. De plus, une mauvaise estimation de la bande de fréquence peut conduire à un rapport S/B insuffisamment amélioré. La connaissance du contenu fréquentiel du signal utile n'est pas nécessaire dans le cas du filtrage par reconnaissance des formes. Ce dernier est véhiculé par le signal d'apprentissage qui, indirectement, fixe la bande de fréquence.

La reconnaissance des formes apparaît comme un filtre efficace qui ne nécessite pas d'*a priori* sur la bande de fréquence du signal reçu ni sur la phase de l'ondelette. Cependant, elle reste tributaire d'un signal d'apprentissage qui doit être bien défini et identifiable.

Cette méthode, relativement simple quant à sa mise en oeuvre, est, jusqu'à maintenant, appliquée en sismique de puits, mais elle peut très bien s'utiliser en sismique de surface pour améliorer la résolution aux grandes profondeurs ou à chaque fois que l'on désire extraire un type d'ondes bien différencié par sa forme.

Remerciements

Nous remercions vivement L. Baron (Université de Lausanne), et un rapporteur anonyme pour les remarques constructives qui ont permis d'apporter des améliorations significatives à ce travail.

RÉFÉRENCES

Belaïd, A. & Belaïd, Y. 1992: Reconnaissance des formes. Méthodes et Applications, InterEditions, Paris, 429 p.

Brezinski, C. & Redivo-Zaglia, M. 2004: Méthodes numériques directes de l'algèbre matricielle, Ellipses Ed., Paris, 133 p.

Ciarlet, P.G. 2002: Introduction à l'analyse numérique matricielle et à l'optimisation, Dunod Ed., Paris, 279 p.

Claerbout, J.F. 1992: Earth soundings analysis: processing versus inversion, Blackwell Scientific Publications, Inc., 304 p.

Cliet, C., Dubesset, M. & Michaud, O. 1986: Pattern recognition: an application to wave separation in three-component offset VSP, SEG 56th

annual international meeting Huston nov 2–6, expanded abstract, 308–310.
Corso, G., Kuhn, P.S., Lucena, L.S. & Thomé, Z.D. 2003: Seismic ground roll time-frequency filtering using the gaussian wavelet transform, Physica Acta 318, 551–561.
Ganley, D.C. 1981: A method for calculating synthetic seismograms which include the effects of absorption and dispersion, Geophysics, 46 (8), 1101–1107.
Hardage, B.A. 2000: Vertical seismic profiling: principles (third updated and revised edition). Seismic Exploration, vol.14, Elsevier Science Ltd, Oxford, UK, 580 p.
Lavergne, M. 1986: Méthodes sismiques, Edition Technip, Paris, 207 p.
Lengellé, R. 2002: Décision et reconnaissance des formes en signal, Hermès Science Publications, Paris, 319 p.
Mari, J.L. & Coppens, F. 2000: Sismique de puits, Edition Technip, Paris, 240 p.
Mari, J.L., Glangeaud, F. & Coppens, F. 2001: Traitement du signal pour géologues et géophysiciens. 1- prospection sismique. Edition Technip, Paris, 213 p.
Michaud, O. 1985: Utilisation de la reconnaissance des formes en sismique trois composantes, Document IFP, nr. 33585, Paris, 93 p.
Yilmaz, O. 1987: Seismic data processing, Society of Exploration Geophysicists, Tulsa (USA), 526 p.

Manuscript received May 27, 2008
Revision accepted March 3, 2009
Published Online first July 27, 2009
Editorial Handling: F. Marillier & S. Bucher