DOI 10.1007/s12182-012-0214-9

An application of matching pursuit time-frequency decomposition method using multi-wavelet dictionaries

Zhao Tianzi[1, 2*] **and Song Wei**[2]

[1] Library of China University of Petroleum, Beijing 102249, China
[2] CNPC Key Laboratory of Geophysical Exploration, China University of Petroleum, Beijing 102249, China

Abstract: In the time-frequency analysis of seismic signals, the matching pursuit algorithm is an effective tool for non-stationary signals, and has high time-frequency resolution and a transient structure with local self-adaption. We expand the time-frequency dictionary library with Ricker, Morlet, and mixed phase seismic wavelets, to make the method more suitable for seismic signal time-frequency decomposition. In this paper, we demonstrated the algorithm theory using synthetic seismic data, and tested the method using synthetic data with 25% noise. We compared the matching pursuit results of the time-frequency dictionaries. The results indicated that the dictionary which matched the signal characteristics better would obtain better results, and can reflect the information of seismic data effectively.

Key words: Matching pursuit, seismic attenuation, wavelet transform, Wigner Ville distribution, time-frequency dictionary

1 Introduction

Time-frequency analysis of seismic data is very important in processing and interpretation, such as measurements of seismic attenuation, detection of hydrocarbon, and improvement of resolution. The short time Fourier transform (STFT) is used widely in time-frequency analysis, but due to the limitation of Heisenberg's uncertainty principle, it is difficult to consider the needs of both time and frequency resolution. Liu et al (2011) proposed the local attributes time-frequency analysis method by shaping regularization on the basis of the conventional Fourier transform, and applied it successfully to detecting low-frequency shadows and ancient river sediments. Wavelet transform is also used widely in time-frequency analysis. Chakraborty and Okaya (1995) compared the difference between wavelet transform and Fourier transform in time-frequency analysis, and they pointed out that wavelet transform can improve the spectral resolution. Stockwell et al (1996) proposed the S transform. This method can overcome the shortcomings of STFT, at the same time it introduced multi-resolution analysis of wavelet transform. It has good time-frequency analysis capability, but it is limited in practical applications due to the fixed form of the basic wavelet function.

Matching pursuit has high time-frequency resolution and a transient structure with local self-adaption, and is used in many areas (Du and Shen, 2006; He et al, 2010; Fan et al, 2009; Wang, 2007; 2010; Zhang et al, 2010). It has the ability to extract more signal characteristics and is not affected by the noise in signals. It can overcome the shortcomings of traditional Fourier transform, windowed Fourier transform, wavelet transform and S transform (Li, 2006; Liu et al, 2004b; Xu, 2000; Zhang et al, 2006; Zou et al, 2004). In the seismic exploration, the time and frequency dictionaries of matching pursuit are limited. Castagna et al used matching pursuit to process 2D seismic signals, and further used the method to detect low-frequency shadows associated with hydrocarbons (Castagna et al, 2003; Nguyen and Castagna, 2000). Liu et al used matching pursuit based on Morlet and Ricker wavelet dictionaries to do time-frequency decomposition (Liu et al, 2004a; Liu and Marfurt, 2005). Wang and Yang (2010) used the improved Ricker wavelet dictionary to sparsely decompose the actual seismic signals. However, most of the actual seismic wavelets are mixed-phase, and matching pursuit needs to scan the whole time and frequency domains of the signal, so the atom dictionary must be over-complete in order to decompose and reconstruct the signal best. Therefore, we should add some new dictionaries which are mixed-phase to match the time and frequency characteristics of seismic signals better.

2 Matching pursuit

The principle of matching pursuit is to decompose any

* Corresponding author. email: 13626080@qq.com
Received March 24, 2011

signal into a linear expansion of waveforms that belong to a redundant dictionary of functions. These waveforms can match the signal structure best. The signal is decomposed into waveforms selected from a time-frequency atom dictionary, which is obtained by extension, translation, and modulation of signal window functions. By adding the Wigner distribution to the selected atom dictionary, we can obtain a time-frequency energy distribution. Because there are no interference terms in the distribution, we can obtain a clear picture in the time-frequency plane.

The $L^2(R)$ is the Hilbert space of complex valued function. The time-frequency atom dictionary is built by extending, translating, and modulating the signal window function such as $g(t) \in L^2(R)$. We suppose $g(t)$ is real, continuously differentiable and $O(\frac{1}{t^2+1})$, and suppose $\|g\| = 1$. So integration of $g(t)$ is nonzero, and $g(0) \neq 0$. For any scale s which is more than zero, frequency modulating coefficient ξ and translating coefficient u, we define $\gamma = (s, u, \xi)$, and let

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g(\frac{t-u}{s}) \mathrm{e}^{i\xi t} \tag{1}$$

We select an approximately countable subset of atoms $[g_{\gamma_n}(t)]_{n \in N}$, because the time-frequency atom dictionary is complete, matching pursuit will decompose every function which satisfies $f(t) \in L^2(R)$. After m iterations, a matching pursuit decomposes the signal f. The number of iterations is related to the decay rate of $\|R^n f\|$, and the norm of the remainder decays exponentially. When the value of $\|R^n f\|$ is less than the set accuracy, the decomposition will stop.

The Wigner distribution of $f(t)$ is $Wf(t,\omega) = W[f,f](t,\omega)$. We can obtain $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Ef(t,\omega)\mathrm{d}t\mathrm{d}\omega = \|f\|^2$ by the energy conservation. So $Ef(t,\omega)$ can be considered as the energy density of f in time-frequency plane (t, ω), and there are no crossing terms. The time-frequency energy distribution $Ef(t,\omega)$ is a sum of single $g(t)$ energy distribution (Mallat, 1993).

3 Multi-wavelet time-frequency dictionary construction

Though matching pursuit has high resolution of time and frequency, it also has different time and frequency characteristics in the processing of different signals. So the decomposition results will be more accurate when we select the dictionary which best matches the signal. Based on the Gabor dictionary, we select several wavelet functions to expand the dictionary. These wavelet functions are selected according to the characteristics of seismic signal, so they have the same time and frequency characteristics as the seismic signal, and then the new dictionary can match the seismic signal better. We can use the Wigner Ville distribution function of the wavelet to build a multi-wavelet time-frequency dictionary. It can further improve the time-frequency resolution of the signal, and the decomposed atoms can maximally retain the original time-frequency characteristics.

Here, the function of seismic Ricker wavelet in time domain is $R(t) = [1-2(\pi g t)^2]\mathrm{e}^{-(\pi g t)^2}$ (g is the peak frequency), and the Wigner Ville distribution of it is

$$\begin{aligned} W_r(t,\omega) = 2\pi \mathrm{e}^{-2\pi^2 g^2 t^2 - \frac{\omega^2}{2\pi^2 g^2}} (2\sqrt{2\pi}\pi^2 g^3 t^4 - 3\sqrt{2\pi} g t^2 \\ + \frac{3\sqrt{2\pi}}{8\pi^2 g} - \frac{\sqrt{2\pi}\omega^2}{4\pi^4 g^3} + \frac{\sqrt{2\pi}\omega^4}{8\pi^6 g^5} + \frac{\sqrt{2\pi}\omega^2 t^2}{\pi^2 g}) \end{aligned} \tag{2}$$

The function of the Morlet wavelet is $\phi(t) = \pi^{-\frac{1}{4}}(\mathrm{e}^{-i\omega_0 t} - \mathrm{e}^{-\frac{\omega_0^2}{2}})\mathrm{e}^{-\frac{t^2}{2}}$, ω_0 is the angular frequency, and the Wigner Ville distribution of it is

$$W_\varphi(t,\omega) = 2\mathrm{e}^{-t^2-\omega^2-\omega_0^2}(\mathrm{e}^{-2\omega\omega_0}+1) - 4\mathrm{e}^{-t^2-\frac{\omega_0^2}{2}-\frac{(2\omega+\omega_0)^2}{4}}\cos(\omega_0 t) \tag{3}$$

The function of mixed-phase seismic wavelet (MSW) (Gao and Yang, 2007) can change the phase and amplitude of the wavelet through regulating the four parameters so as to match the seismic signal. Most of the deconvolution processing is based on the seismic wavelet being time-invariant and its phase being zero. Actually the seismic wavelet is time-variant, and most are mixed phase wavelets. So, using MSW to build a mixed phase time-frequency dictionary is very important for practical applications. The wavelet function is $s(t) = A\mathrm{e}^{-\tau(t-\theta)^2}\cos(2\pi f_m t)$, where parameters (A, τ, θ, f_m) are amplitude coefficient, attenuation factor, phase factor and peak frequency of the wavelet. Its Wigner Ville distribution is

$$\begin{aligned} W_M(t,\omega) = A^2 \frac{\sqrt{\pi}}{\sqrt{2\tau}} \mathrm{e}^{-\tau(2t^2-4\theta t+2\theta^2)-\frac{\omega^2}{2\tau}}[\cos(4\pi f_m t) \\ + \cos(2\pi f_m \frac{i\omega}{\tau})\mathrm{e}^{-\pi^2 f_m^2 \frac{2}{\tau}}] \end{aligned} \tag{4}$$

In these equations, t is time of the atom, and ω is frequency of the atom. We take the MSW dictionary as an example, and decompose the signal which is generated by convoluting the mixed phase wavelet and reflection coefficient using the Gabor, Morlet and Ricker wavelet dictionaries and MSW dictionary respectively. Fig. 1(a) is the signal, (b), (c), (d) and (e) are time-frequency spectrums which are the decomposition results of the signal using the Gabor, Morlet and Ricker wavelet dictionaries, and MSW dictionary respectively.

Because the signal has the same time-frequency characteristics as the MSW dictionary, so the best decomposition result should be achieved using the MSW dictionary in theory. Comparing (b), (c), (d) and (e), we can see that all of them can accurately reflect the time locations of signal components. (e) has the most concentrated time-frequency energy and the highest time resolution. The second is (d), and its energy distribution is poorer than (e) from 500 ms to 600 ms. (b) and (c) have poorer time-frequency energy distribution and lower time resolution than (d) and (e). After analysis, we conclude that in decomposition of the signal (a) we should select the MSW dictionary for the best result.

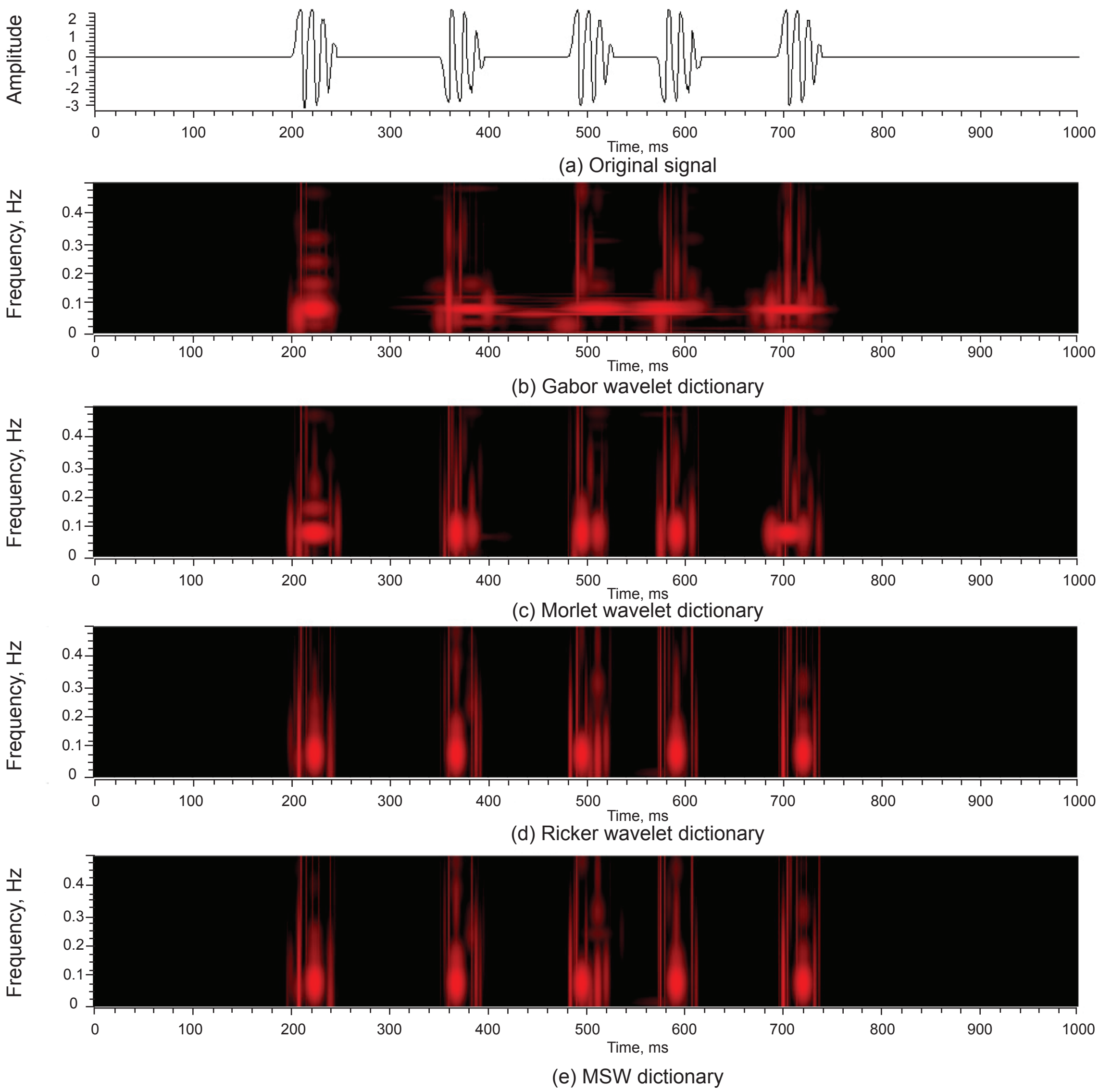


Fig. 1 Comparison of time-frequency spectrums using different dictionaries. (a) synthetic seismic signal using mixed phase wavelet, (b) decomposition result using Gabor wavelet dictionary, (c) decomposition result using Morlet wavelet dictionary, (d) decomposition result using Ricker wavelet dictionary, (e) decomposition result using MSW dictionary

Fig. 2 is the processing result of a simulated signal using this method. (a) is the simulated signal which is generated by mixed phase wavelet and its S/N value is two. We use the MSW dictionary to decompose this noisy signal and choose first ten atoms which have stronger energy to reconstruct the signal. (b) is the comparison of the reconstructed signal (blue) and original signal without noise (red). Their cross-correlation value is 0.86, and the average residual of every point is 1.8×10^{-5}. That is to say, the reconstructed signal through decomposition by the MSW dictionary can effectively reflect the original signal without noise.

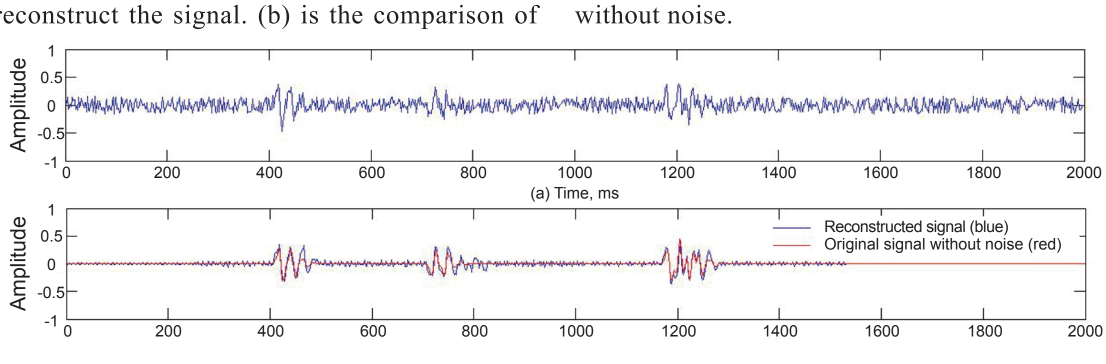


Fig. 2 (a) The original signal which is generated by mixed phase wavelet (S/N ratio is 2), (b) comparison of the original signal without noise (red) and reconstructed signal using the MSW dictionary (blue)

4 Applications

In order to verify the applicability of this method in analysis of seismic wave attenuation, we built a seismic record model which is generated by the mixed phase wavelet (Fig. 3).

There is a sand formation in this model, whose thickness is 10 m, and velocity is 3200 m/s. The quality factor of the traces from 11 to 25 which indicate the gas-containing sand body varies from 30 to 5 evenly. From trace 25 to trace 40, the quality factor increases to 30 evenly. Both ends are the tight sandstone, and the quality factor is 150. The upper and lower formations are mudstone, the velocity is 2400 m/s, and the quality factor is 150. We add 25% Gaussian random noise in the model (the noise frequency band is 0-120 Hz). The reflection layer F is equivalent to the reflection near the bottom of the sand, and there are 50 seismic traces. No denoising method is used in calculating the seismic attributes of the model.

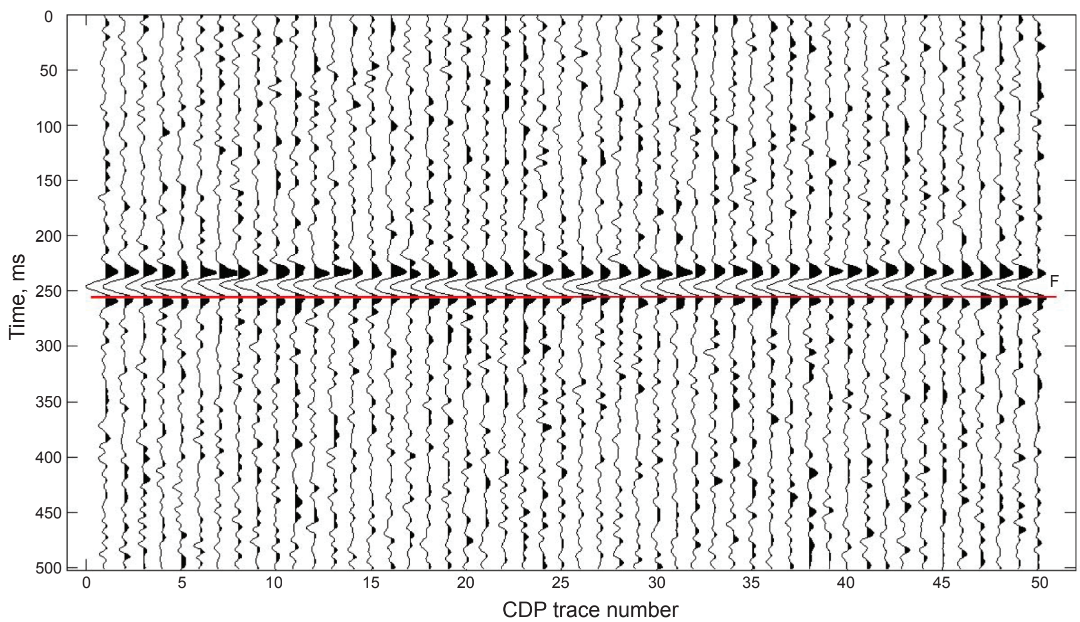
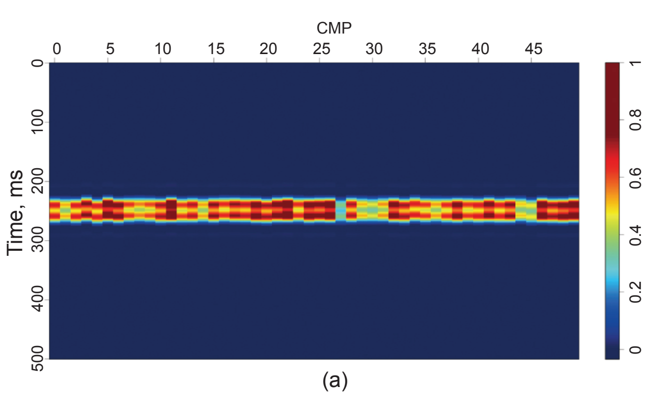


Fig. 3 Synthetic seismic record with a sand formation about 10 m thick

Fig. 4(a), (b), and (c) are time-frequency energy profiles obtained from matching pursuit decomposition using the Morlet wavelet, Ricker wavelet, and MSW dictionaries of the signal in Fig. 3 respectively. From Fig. 4(a) and (b), we can see that the overall signal energy changes irregularly because the time-frequency dictionary does not match the wavelet of the seismic data, and the influence of noise is significant. The time resolution is low, and the sand layer in Figs. 4(a) and (b) is thicker than that in the true model. It is difficult to recognize the actual information of the sand layer. From Fig. 4(c), we can see that the top and bottom interfaces of the sand layer are clear, and the signal energy decays near the reflection layer F, which indicates gas-containing. Using the matching pursuit method, we can obtain a high time-



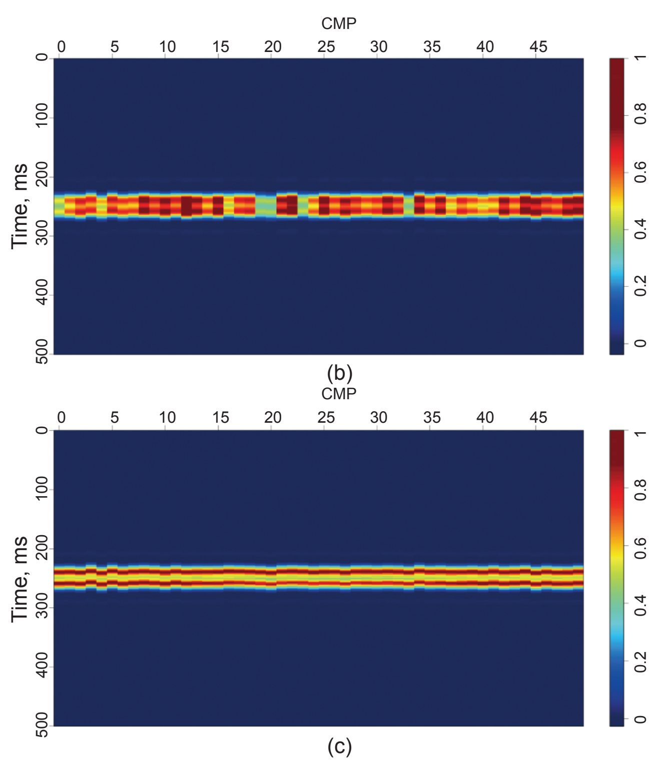


Fig. 4 (a) Matching pursuit time-frequency energy profile using the Morlet wavelet dictionary, (b) matching pursuit time-frequency energy profile using the Ricker wavelet dictionary, (c) matching pursuit time-frequency energy profile using the MSW dictionary

frequency resolution, and the sand layer in Fig. 4(c) reflects its actual position in the model. For this example, it is feasible to identify a 10 m-thick layer using the wavelet whose frequency is 30 Hz.

Fig. 5 is comparison of the change rate of matching pursuit time-frequency energy at the reflection layer F using MSW dictionary, Ricker wavelet dictionary, and Morlet wavelet dictionary respectively. The change rate is the ratio of time-frequency energy of every trace to that of tight sandstone. The blue one represents the energy change rate using MSW dictionary, the green one represents the energy change rate using Ricker wavelet dictionary, and the red one represents the energy change rate using Morlet wavelet dictionary. Near the bottom of the gas-containing sand, the energy change using MSW dictionary is more significant than that using other dictionaries. Therefore, when the sand layer is 10 m thick, with 25% noise in the records, the time-frequency energy of matching pursuit using MSW dictionary which matches the characteristics of the signal is the most sensitive to the quality factor. In other words, when the formation thickness is 10 m, noise will not affect the time-frequency

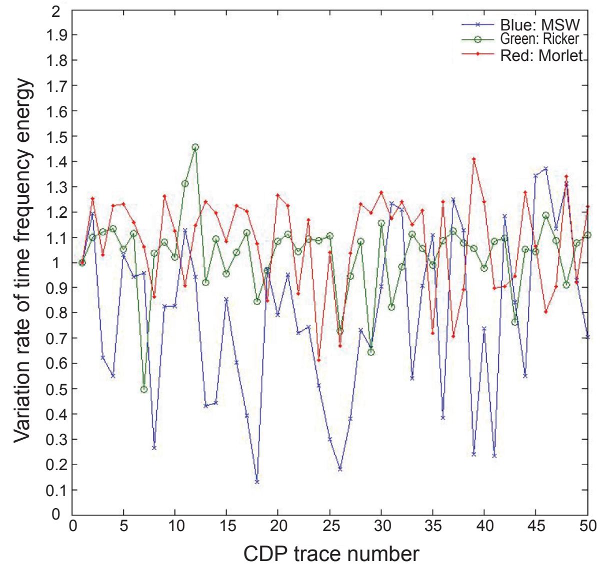


Fig. 5 Comparison of change rate of matching pursuit time-frequency energy at the reflection layer F using the MSW dictionary, Richer wavelet dictionary, and Morlet wavelet dictionary, respectively

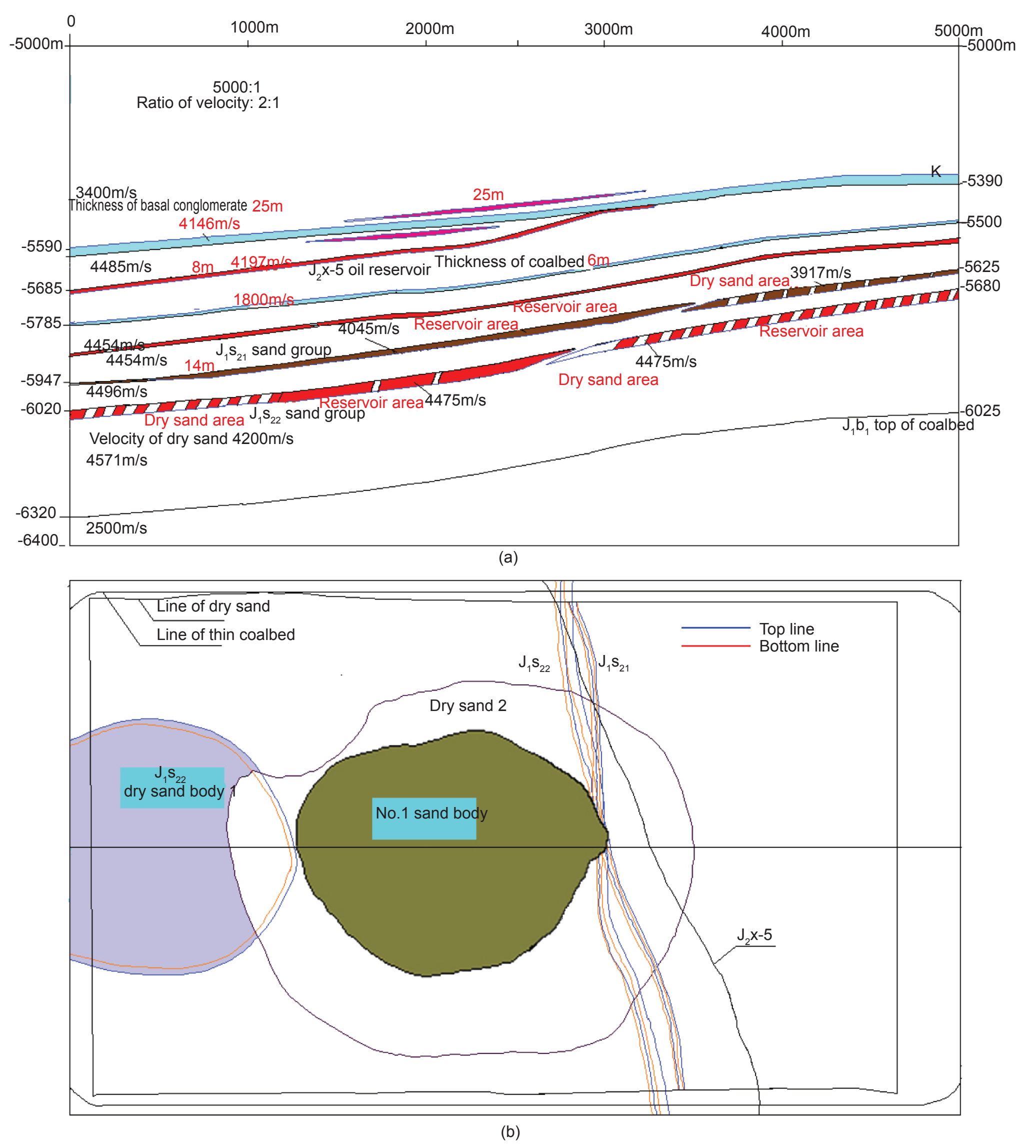


Fig. 6 (a) Geological cross-section of physical model, (b) No.1 sand body shape

energy using appropriate matching pursuit dictionary. The reason is that when calculating the time-frequency energy, the dictionary can best match the signal, and the noise is not calculated in the reconstruction. We can say this method improves the S/N ratio.

Fig. 6(a) is the geological cross-section of a physical model, (b) is the shape of No.1 sand body, which is in the middle of J_1s_{22} sand body group. Fig. 7 is the migration profile of the physical model. Fig. 8(a) is the single frequency slice of No.1 sand body (22 Hz). The red area is in the middle, which means the area is thick, and when the frequency is 22 Hz the tuning energy is a maximum. (b) is the single frequency slice of No.1 sand body (30 Hz). In this figure the red area is small, which means the middle of the layer is thin. In this area, when frequency is 22 Hz, the thickness of the sand layer can be recognized. (c) is the single frequency slice of No.1 sand body (50 Hz), where the red area looks like a circle. There is a gap in the southwest, which means the layer is thick in the middle and thin in the surroundings, and the shape of the sand body is approximately oval. (d) is the peak amplitude of No.1 sand body. The boundary of the red area can reflect the shape of No.1 sand body. From the single frequency slices with different frequency, we can estimate the thickness, location and shape of the sand, and especially identify the thin layer. However, a higher frequency does not produce a clearer geological body distribution, so we should select an appropriate frequency to analyze the geological body.

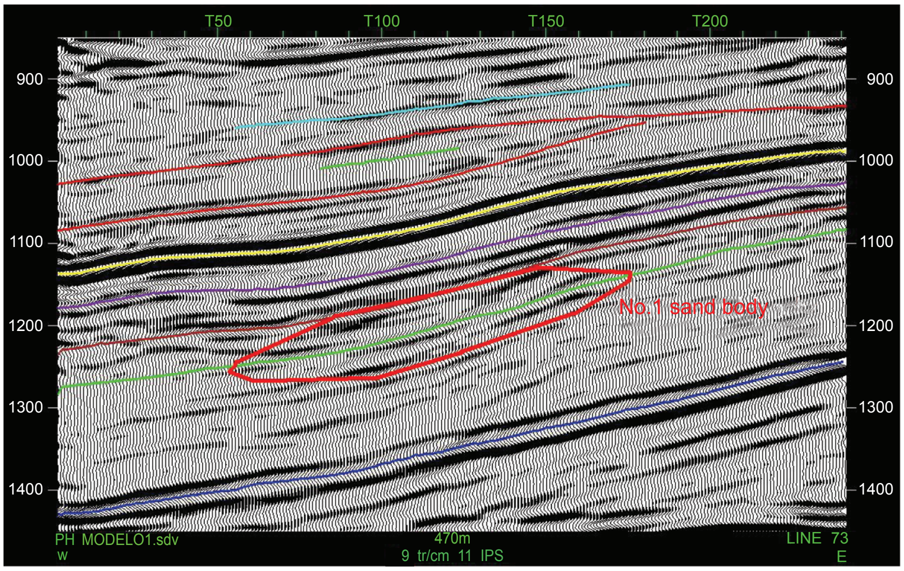


Fig. 7 Migration profile of physical model

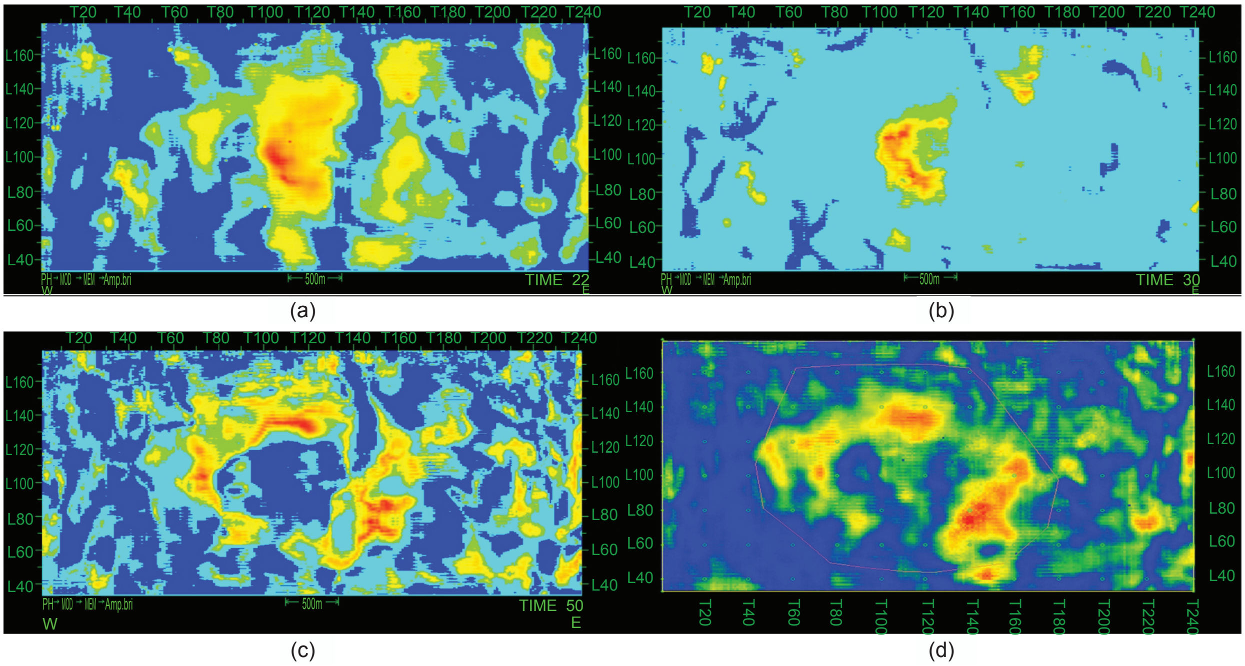


Fig. 8 (a) Single frequency layer slice of No.1 sand body (22 Hz), (b) single frequency layer slice of No.1 sand body (30 Hz), (c) single frequency layer slice of No.1 sand body (50 Hz), (d) peak amplitude of No.1 sand body

5 Conclusions

Theoretical analysis and model test results indicated that when the time-frequency dictionary which matched the seismic data was used to decompose the signal, we could obtain good decomposition and reconstruction results. We assumed that the time-frequency dictionary matched the wavelet of seismic data, so for the MSW dictionary, the wavelet of seismic data should be mixed phase. If the wavelet of seismic data is not mixed phase, we should choose other dictionaries, such as zero phase Ricker wavelet dictionary and minimum phase wavelet dictionary. Therefore, we first need know the wavelet type of the seismic data, then select the appropriate time-frequency dictionary for matching pursuit to achieve optimal decomposition and reconstruction results. The more the time frequency characteristics of the dictionary matches the signal, the better the effect of decomposition and reconstruction. With the advantage of high time-frequency resolution of matching pursuit, the method can be used in seismic denoising, S/N ratio improving, and oil and gas prediction. The use of matching pursuit in quantitative analysis of the thickness of thin layers is a further research direction.

References

Castagna J P, Sun S J and Siegfried R W. Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons. The Leading Edge. 2003. 22(2): 120-127

Chakraborty A and Okaya D. Frequency-time decomposition of seismic data using wavelet-based methods. Geophysics. 1995. 60(6): 1906-1916

Du X L and Shen Y. Denoising of ultrasonic testing signals by matching pursuits. Nondestructive Testing. 2006. 28(8): 409-415 (in Chinese)

Fan H, Meng Q F, Zhang Y Y, et al. Matching pursuit based on nonparametric waveform estimation. Digital Signal Processing. 2009. 19(4): 583-595

Gao J H and Yang S L. On the method of quality factors estimation from zero-offset VSP data. Chinese Journal of Geophysics. 2007. 50(4): 1198-1209 (in Chinese)

He H J, Wang Q Y and Cheng H M. The application of wavelet decomposition technique based on matching pursuit algorithm in thin interbedded reservoir prediction. Computing Techniques for Geophysical and Geochemical Exploration. 2010. 32(6): 641-644 (in Chinese)

Li H B. Characteristics of seismic attenuation in wavelet domain and the detection of gas. China University of Mining & Technology Ph.D. Thesis. 2006 (in Chinese)

Liu G C, Fomel S and Chen X H. Time-frequency characterization of seismic data using local attributes. Geophysics. 2011. 76(6): P23-P34

Liu J L, Wu Y F, Han D H, et al. Time-frequency decomposition based on Ricker wavelet. Expanded Abstracts of 74th SEG Annual International Meeting. 2004a. 1937-1940

Liu J L and Marfurt K J. Matching pursuit decomposition using Morlet wavelets. Expanded Abstracts of 75th SEG Annual International Meeting. 2005. 786-789

Liu Q Y, Li Z S and Liu Z H. Time-frequency analysis and its present situation. Computer Engineering. 2004b. 30(1): 171-184 (in Chinese)

Mallat S G and Zhang Z F. Matching pursuits with time-frequency dictionaries. IEEE Transactions on Signal Processing. 1993. 41(12): 3377-3415

Nguyen T and Castagna J. Matching pursuit of two-dimensional seismic data and its application. SEG Annual Meeting Expanded Technical Program Abstracts with Biographies. 2000. 2067-2069

Stockwell R G, Mansinha L and Lowe R P. Localization of the complex spectrum: The S transform. IEEE Transactions on Signal Processing. 1996. 44(4): 998-1001

Wang C M and Yang S L. Seismic signal's MP decomposition based on Ricker wavelet. Computer Knowledge and Technology. 2010. 6(4): 976-978 (in Chinese)

Wang Y H. Seismic time-frequency spectral decomposition by matching pursuit. Geophysics. 2007. 72(1): 13-21

Wang Y H. Multichannel matching pursuit for seismic trace decomposition. Geophysics. 2010. 75(4): 61-66

Xu K S. Time-frequency analysis of the earthquake signals. Northwestern Seismological Journal. 2000. 22(4): 479-482 (in Chinese)

Zhang F C, Li C H and Yin X Y. Seismic data fast matching pursuit based on dynamic matching wavelet library. Oil Geophysical Prospecting. 2010. 45(5): 667-673 (in Chinese)

Zhang F, Zhong Y Y, Zhu X Y, et al. Application of time-frequency analysis to seismic wave spectrum research. Seismological and Geomagnetic Observation and Research. 2006. 27(4): 17-22 (in Chinese)

Zou W, Chen A P and Gu H M. Joint time-frequency analysis and its application in seismic prospecting. Progress in Exploration Geophysics. 2004. 27(4): 246-250 (in Chinese)

(Edited by Hao Jie)