Journal of Applied Geophysics 112 (2015) 52–61

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Seismic facies analysis based on self-organizing map and empirical mode decomposition

Hao-kun Du [a,b,*], Jun-xing Cao [a,b], Ya-juan Xue [a,b,c], Xing-jian Wang [a,b]

[a] State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, China
[b] School of Geophysics, Chengdu University of Technology, Chengdu 610059, China
[c] School of Communication Engineering,Chengdu University of Information Technology,Chengdu 610225, China

ARTICLE INFO

Article history:
Received 3 June 2014
Accepted 9 November 2014
Available online 15 November 2014

Keywords:
Neural networks
SOM
EMD
Seismic facies analysis
Seismic interpretation

ABSTRACT

Seismic facies analysis plays an important role in seismic interpretation and reservoir model building by offering an effective way to identify the changes in geofacies inter wells. The selections of input seismic attributes and their time window have an obvious effect on the validity of classification and require iterative experimentation and prior knowledge. In general, it is sensitive to noise when waveform serves as the input data to cluster analysis, especially with a narrow window. To conquer this limitation, the Empirical Mode Decomposition (EMD) method is introduced into waveform classification based on SOM. We first de-noise the seismic data using EMD and then cluster the data using 1D grid SOM. The main advantages of this method are resolution enhancement and noise reduction. 3D seismic data from the western Sichuan basin, China, are collected for validation. The application results show that seismic facies analysis can be improved and better help the interpretation. The powerful tolerance for noise makes the proposed method to be a better seismic facies analysis tool than classical 1D grid SOM method, especially for waveform cluster with a narrow window.

1. Introduction

The key to reservoir modeling is the effective evaluations of rock properties and accurate mapping of their heterogeneity. Various types of information, including core samples, well logs, production data, seismic data, and geological setting, are used in this model building (de Matos et al., 2007; Rezaee, 2002). Due to the heavy cost of drilling, there is usually no adequate number of wells to build the model for relatively large areas. Specifically, data from well logs and cores only represents local properties of the reservoir and it is unreliable to extrapolate these properties to the whole prospect with the absence of enough wells. In this case, 3D seismic data plays an important role in identifying the lateral changes of reservoirs and describing their geological features. Changes in lithology, porosity, and fluid content lead to changes in amplitude, frequency, lateral continuity and other seismic attributes (de Matos et al., 2007). Thus, if changes of seismic parameters can be identified and interpreted, some valuable information of reservoirs may be extracted and may help in the understanding of subsurface geology.

Seismic facies analysis aims to interpret the variations of seismic response parameters. The so-called seismic facies can be defined as groups of seismic traces; members of the same group possess similar wave sharp. They can be viewed as the manifestation of specific sedimentary facies or geologic bodies in seismic data (John et al., 2008). So far there have been several methods and techniques of pattern recognition applied to the classification of seismic facies with varying degrees of success (e.g., Jin et al., 2007; Li and Castagna, 2004; Saggaf et al., 2003). When the geological information is unavailable, unsupervised pattern classification has been demonstrated as a powerful method for seismic facies analysis (de Matos et al., 2007). The self-organizing map (SOM) (Kohonen, 2001) is certainly one of the most successful neural network algorithms applied to unsupervised classification (e.g., Roy et al., 2010, 2012; Singh et al., 2004; Taner et al., 2001).

Seismic waveform and attribute values are the most commonly used inputs to the classification process. Since the seismic waveform contains integrated information of multiple attributes such as amplitude, frequency and phase, it is more reliable to use this integrated data directly to analysis and classification (Singh et al., 2004). However, there also exist some unnecessary information such as noise and insensitive parameters to changing geologic structure in seismic waveform. Therefore, when we use seismic waveform as the input for classification, the outcomes are generally more susceptible to noise and lower in resolution. When we use attributes values as the input, if appropriate attributes can be extracted from seismic waveform to serve as the input, the results of classification may provide higher resolution and lower amount of noise (Kuroda et al., 2012). However, the seismic attributes selection is a difficult problem and requires utmost care because

* Corresponding author at: State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, China.
E-mail addresses: dhk20052005@163.com (H. Du), caojx@cdut.edu.cn (J. Cao), xueyj0869@163.com (Y. Xue), wangxj@cdut.edu.cn (X. Wang).

http://dx.doi.org/10.1016/j.jappgeo.2014.11.007
0926-9851/

of its evident effect on the result of classification (Raeesi et al., 2012). There is no criterion for the selection of which attributes can best represent the changes in rock property (Coléou et al., 2003). This leads to more uncertainty about whether the seismic attributes we used are the optimal one to the local structural feature and what is the relationship between them. Therefore how to enhance the accuracy of SOM clustering while preserving its reliability has been a research topic in seismic facies analysis.

Seismic data volumes are significantly noisy and greatly redundant. The classification can be greatly optimized by using appropriate preprocessing on seismic data (Coléou et al., 2003). de Matos et al. (2007) applied time-frequency techniques to the SOM clustering by using the WTMMLA curves derived from trace singularities as the input data. Saraswat and Sen (2012) employed Artificial Immune System (AIS) to the compaction of seismic data and use the reduced data for SOM processing. Both of them focus on the reprocessing of seismic data in order to obtain an excellent input to the SOM clustering.

Empirical Mode Decomposition (EMD), as a new decomposition method for analyzing nonlinear and non-stationary data introduced by Huang et al. (1998), has been increasingly used for seismic signals analysis and demonstrated great potential in this application. (e.g., Battista et al., 2007; Wen et al., 2009; Xue et al., 2013, 2014). In this method, complicated seismic signals can be decomposed into a series of Intrinsic Mode Functions (IMFs) in the temporal domain (Xue et al., 2014). In essence, this decomposition method can be considered as a dyadic filter bank that serves a similar function to wavelet transform (Flandrin et al., 2004). But unlike wavelet transform required for pre-set base function, it is an adaptive data-driven method based on the local characteristic of data in time scale. Therefore, EMD method offers some distinct advantages over wavelet transform used for noise reduction and resolution enhancement of seismic data (Battista et al., 2007; Ehrhardt et al., 2012; Huang et al., 2011; Xue et al., 2013).

In this study, we introduce EMD method into unsupervised seismic facies analysis based on SOM. The seismic signals are decomposed by EMD to obtain the featured subsignals reconstructed by a number of IMFs that highlight the fine details and smooth noise. This step is to remove spikes, reduce noise and improve resolution of seismic data. Then, the reconstructed seismic subsignals are allowed as the input to SOM clustering and the useful geological information can be extracted from the results of clustering. In this paper, we will first introduce the concept of SOM and EMD, and then test our method on synthetic and real data.

2. Principle and methods

2.1. Empirical Mode Decomposition (EMD)

EMD aims to obtain IMF which is a monofrequency signal. Thus IMF has well-behaved Hilbert transforms and the physical meaning. Each IMF is defined to meet the following conditions (Huang et al., 1998):

(1) The number of zero-crossings and extrema is the same;
(2) The mean value of the upper envelops and the lower envelops is equal to zero.

EMD is carried out by a sifting process. The decomposition of a signal into IMFs by EMD is performed as follows:

1. Find out the maxima and the minima of the original signal.
2. Construct the upper envelops and the lower envelops of the signal. Generally the cubic spline method is used.
3. Obtain the mean values by averaging the upper and the lower envelopes.
4. Subtract the mean values from the original signal. Ideally, the first IMF component is produced. If the signal obtained by subtracting the mean values from the original signal does not meet the IMF condition, repeat the steps 1–4 to this signal until the first IMF component is obtained.
5. Subtract the first IMF component from the original signal and carry out the steps 1–4 to the residual component until all the IMFs are obtained.

After decomposition, the original signal can be expressed as a sum of IMFs and the residual component which is usually a monotonic function.

From the sifting process of EMD, we can find that the different IMF has different frequency ranges and probably highlights different details. For seismic data, the subsignal reconstructed by the selected IMFs which is the main component of the original seismic signal can highlight the fine details and reduce the noise.

2.2. Classification using SOM and EMD

Seismic facies analysis is an efficient measure of predicting the underlying structure and depositional environment by recognizing and analyzing the characteristics of a group of seismic reflections. The SOM, as a type of unsupervised learning (Kohonen, 2001), has been widely used in seismic facies analysis (de Matos et al., 2007; Roy et al., 2010; Saraswat and Sen, 2012; Taner et al., 2001). The EMD method is an excellent data reconstruction algorithm that removes the noise and preserves essential features of original seismic data. In our method, we introduce the EMD method into seismic facies analysis based on SOM to obtain an enhancement.

The workflow of proposed method is illustrated in Fig. 1. We first decompose seismic data into IMFs using EMD method. Then, the IMFs which contain the fine information and smooth noise are chosen to reconstruct a new seismic data. Next, we allow the reconstructed data as the input to SOM training and clustering with different numbers of facies. Comparing the facies maps, the optimal facies number can be defined.

In this workflow, the selection of the IMFs and facies numbers should be noticed. The proper selection of IMFs usually results in a better de-noising effect on seismic data. Thus it is not hard to see that the selection of IMFs directly affects the quality of the classification based on our method. Each original seismic trace usually produces a series of IMFs. Then, we should analyze the features of the IMFs displayed as seismic sections and the correlation coefficients between each IMF and original seismic trace. The IMFs which have higher correlation coefficients and greater similarity to the original seismic trace on sections will be selected for reconstruction. The deserted IMFs tend to have low correlation and a lot of noise which is evident from the seismic sections.

The number of classes, i.e., the number of seismic facies, has important effects on the classification process and should be estimated very carefully (Raeesi et al., 2012). A low number of classes can only offer a very rough classification in which some important information may be obscured and the facies of interest usually cannot be identified. Conversely, the high number of classes can enhance details and accuracy of the classification; however, it also produces lots of redundant facies which might complicate the interpretation. In unsupervised seismic facies analysis, one can use more seismic facies than the number of expectant geofacies in the researched region (Raeesi et al., 2012). Because the additional facies could be required to represent the noise (Roy et al., 2012), including background noise and horizon interpretation noise caused by interpretation errors or time displacement, and the transition zone between the main facies. The noises are represented by one or several facies with haphazard distribution; while the transition zone's facies usually distribute along the periphery of the main facies. The best situation is that with the increasing number of classes, the main facies remain about the same and only the number of unnecessary facies is in the growth. Therefore, the estimation of the number of classes in study area requires iterative testing and prior knowledge.