Tu P6 08

Channel Detection Using Unsupervised Learning Techniques

A.H. Mardan* (Amirkabir University of Technology), A. Javaherian (Amirkabir University of Technology), M. Mirzakhanian (NIOC)

Summary

Channel, an important geological facies for exploration and development of oilfields, has a narrow appearance. It means that the detection of this facies in the huge volume of seismic data and using numerous introduced seismic attributes is one of the most challenging tasks for interpreters. To address this difficulty, several computer-assisted learning techniques have been introduced. In recent years, the interpreters paid more attention to unsupervised learning techniques such as k-means, self-organizing maps (SOM), principal component analysis (PCA), and independent component analysis (ICA), because these learning techniques do not need to the knowledge of the interpreters about the studied area. In this study, to detect channel facies of one of the southwest oilfields of Iran, the mentioned algorithms have been used, and the results show that the studied area has two main channel branches those can be detected by all applied algorithms. Additionally, two narrow branches of the channel and some other geological feature are detected using ICA and PCA.

Introduction

In recent years, the size of 3D seismic data volumes and the number of seismic attributes have been increased, and it is difficult and time-consuming for an interpreter to examine every seismic line, and time slice to detect special features like channels. To address this problem, some automatic clustering algorithms such as the k-means, generative topographic mapping, and self-organizing maps have been successfully used. Additionally, principal component analysis and independent component analysis can propose a good insight to interpreters if it is visualized in red-green-blue (RGB) color system with proper situations.

After presenting seismic attributes in the 1970s (Balch, 1971), employing the computers in the interpretation of seismic data was considered by Sonneland (1983). The *k*-means algorithm (Jancey, 1966) is one of the first pattern recognition algorithms which was considered to analyze the seismic data, and it is used up to now (Coléou *et al.*, 2003; Sabeti and Javaherian, 2009; Zhao *et al.*, 2015). SOM (Kohonen, 1982), PCA (Pearson, 1901), and ICA (Comon, 1994) are other techniques which were used to analyze this data (Roy and Marfurt, 2010; Zhao *et al.*, 2015; Honório *et al.*, 2014). In this study, after introducing the used seismic attributes, the mentioned unsupervised learning techniques including *k*-means, SOM, PCA, and ICA are applied to a 3D seismic data volume acquired over the Abadan Plain to detect the buried channels in this area and facies analysis of these available data.

Method and/or Theory

The studied area relates to a time slice at 1.8 s from a seismic data set acquired over an oilfield in the southwest of Iran. The input of pattern recognition algorithms is a multi-attribute matrix so that each sample of input matrix has some seismic attributes and the amount of these attributes should be as least as possible (Zhao *et al.*, 2015). In this study, acoustic impedance (Figure 1(a)) and spectral decomposition (Figure 1(b)) were utilized to investigate the lithology and types of fluid, respectively. Also, root-mean-square amplitude and grey-level co-occurrence matrix were employed to determine the lateral variation (Figure 1(c)) and the distribution of data (Figure 1(d)), respectively.

As mentioned, the *k*-means method is one of the first and simplest clustering techniques and the goal its iterative process is minimizing the distance between the center of clusters (μ) and the data samples (x):

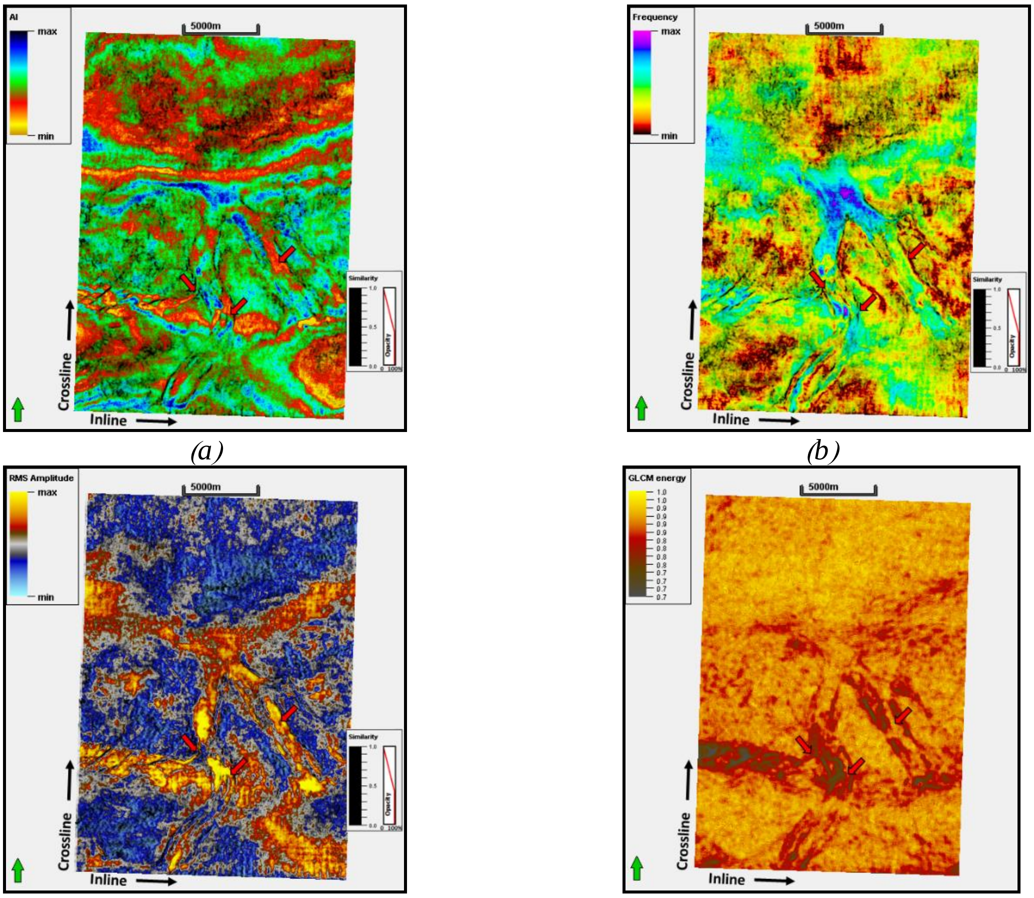
$$J = \frac{1}{m} \sum_{c(j)=1}^{k} \sum_{i=1}^{m} \left\| x_{(i)} - \mu_{c(i)} \right\|^2 , \tag{1}$$

where *m* and *k* are the number of samples and the number of clusters, respectively. The competitive process of SOM is also based on the distances between some supposed neurons (*w*) and data samples. In this method, each data sample has a winner neuron which satisfies the following condition:

$$w_j(x_i) = \{w_j \mid j = \arg\min_s x_i - w_s , \forall w_s \in w\}, \tag{2}$$

where *s* is the number of the supposed neuron. In spite of *k*-means, the SOM method has two controlling parameters in its optimization path.

PCA and ICA are two visualization techniques, and they aim to extract the directions, which the analyzed data are most independent. PCA decorrelates the input data, but ICA also separates the remaining higher order dependencies (Draper *et al.*, 2003). Another important difference between PCA and ICA is their assumption about the input data. PCA assumes the signal as a Gaussian distribution, whereas ICA assumes that the signal is a non-Gaussian distribution (Lu, 2006). So, the non-Gaussian distribution of seismic data (Walden, 1985) fits the fundamental assumption of ICA.

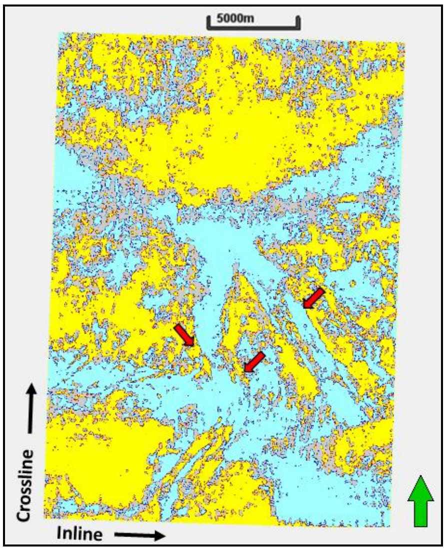


(a) *(b)*

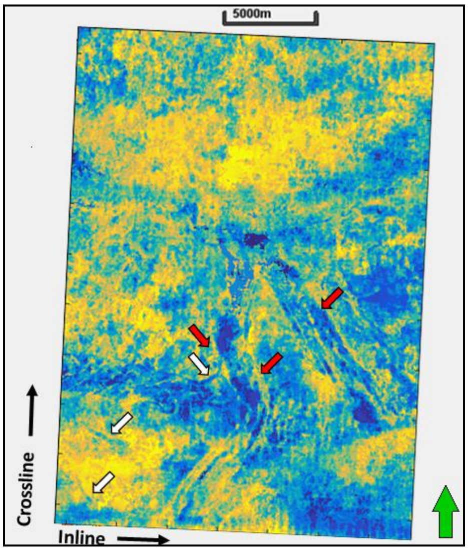
(c) *(d)*

Figure 1 Volume time slice at 1.8 s through (a) acoustic impedance, (b) spectral decomposition with the frequency of 15 Hz, (c) root-mean-square amplitude, the structure of channel facies have been marked with similarity attributes, (d) grey-level co-occurrence matrix energy attribute. Red arrows show two studied broad channels.

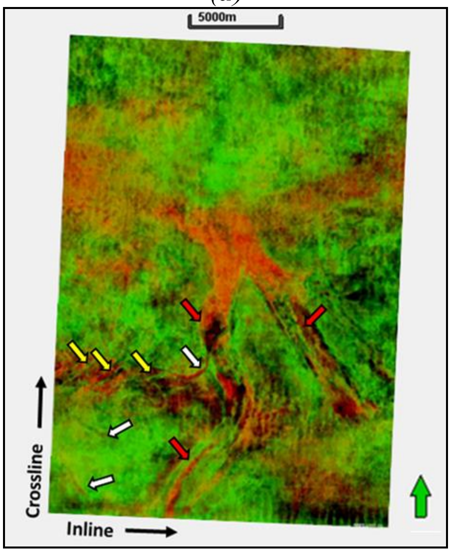
As the outputs of the used algorithms are shown, *k*-means (Figure 2(a)) is not appropriate as other methods to detect the narrow facies. Although the SOM (Figure 2(b)) and PCA (Figure 2(c)) have relatively identical results, SOM is more time-consuming than PCA, and this is the superiority of PCA. In addition to detected features from previous algorithms, the output of ICA (Figure 2(d)) clearly shows a narrow channel in the east of the studied area (blue arrows) which cannot be determined using other methods and it shows, ICA is more accurate than other used methods.



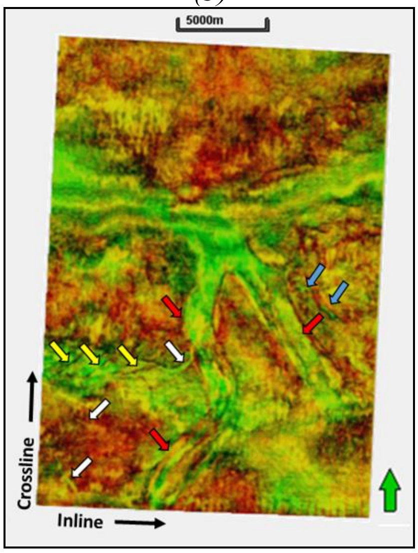
(a)



(b)



(c)



(d)

Figure 1 Volume time slice at 1.8 s through (a) k-means, (b) SOM, (c) PCA, (d) ICA. Red arrows show two studied broad channels. The white arrows show a narrow branch of channel which is noticeable in SOM, PCA, and ICA result. The yellow arrows indicate some extracted linear feature like a fracture in the west of the studied area, and finally, the blue arrows show a narrow channel which is noticeable in ICA result.

Conclusions

In this study, four unsupervised pattern recognition techniques, *k*-means, SOM, PCA, and ICA, were used to identify the buried channels within the Sarvak Formation in Abadan Plain, SW Iran. The *k*-means result is not appropriate as other used techniques. SOM and PCA identify a narrow channel branch in the west of the studied area. Additionally, some small fractures are identified in that portion by using PCA, where the fractured zone behaves like the channels, but the fracture lines are different from the channels system. This phenomenon is the cause of selected attributes in this study. Another channel is identified in the east by using ICA and this is the superiority of ICA over the other used techniques.

Due to the dip of the structure, the channels are prolonged from south to north, and in this way, four branches of channels are joined together.

Acknowledgements

The authors thank National Iranian Oil Company (NIOC) Exploration Directorate for making available the required data in this research.

References

Balch, A.H. [1971] Color Sonagrams: A New Dimension in Seismic Data Interpretation. *Geophysics*, **36**, 1074-1098.
Coléou, T., Poupon, M. and Azbel, K. [2003] Unsupervised Seismic Facies Classification: A Review and Comparison of Techniques and Implementation. *The Leading Edge*, **22**(10), 942-953.
Comon, P. [1994] Independent Component Analysis—a New Concept? *Signal Processing*, **36**, 287–314.
Draper, B.A., Baek, K., Bartlett, M.S. and Beveridge, J.R. [2003] Recognizing Faces with Pca and Ica. *Computer Vision and Image Understanding*, **91**(1), 115-137.
Honório, B.C.Z., Sanchetta, A.C., Leite, E.P. and Vidal, A.C. [2014] Independent Component Spectral Analysis. *Interpretation*, **2**(1), SA21–SA29.
Jancey, R. [1966] Multidimensional Group Analysis. *Australian Journal of Botany*, **14**(1), 127-130.
Kohonen, T. [1982] Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, **43**(1), 59-69.
Lu, W. [2006] Adaptive Multiple Subtraction Using Independent Component Analysis. *Geophysics*, **71**(5), S179-S184.
Pearson, K. [1901] Liii. On Lines and Planes of Closest Fit to Systems of Points in Space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**(11), 559-572.
Roy, A. and J.Marfurt, K. [2010] Applying Self-Organizing Maps of Multiple Attributes, an Example from the Red-Fork Formation, Anadarko Basin. *SEG Technical Program Expanded Abstracts 2010***,** 1591-1595.
Sabeti, H. and Javaherian, A. [2009] Seismic Facies Analysis Based on K-Means Clustering Algorithm Using 3d Seismic Attributes. *1st EAGE International Petroleum Conference and Exhibition*, Iran.
Sonneland, L. [1983] Computer Aided Interpretation of Seismic Data. *SEG Technical Program*, Expanded Abstracts.
Walden, A. [1985] Non-Gaussian Reflectivity, Entropy, and Deconvolution. *Geophysics*, **50**, 2862–2888.
Zhao, T., Jayaram, V., Roy, A. and Marfurt, K.J. [2015] A Comparison of Classification Techniques for Seismic Facies Recognition. *Interpretation*, **4**(3), SAE29-SAE58.