Interpreter Driven, Interactive 3D Multi-Attribute Classification

J Henderson*, James Lowell, Steve Purves, and Gaynor Fisher, Foster Findlay Associates Ltd, UK
David Walraven, Anadarko Petroleum Corporation, USA

Abstract

This paper presents a novel 3D seismic attribute classification technique based on a framework in which the interpreter defines the data partitioning required to generate the classified data volume by directly interacting with a multi-attribute visualisation system. The classification process utilises the latest advances in graphics processing unit (GPU) parallel computation technology, enabling the classification to be achieved in real time on standard (non-cluster) desktop workstations.

The ability to deliver classification results in real time enables interpreter driven feature definition and introduces a new paradigm in multi-attribute analysis techniques, that can be applied so quickly and easily that they can be incorporated simply into any standard 3D seismic interpretation workflow.

Introduction

The use of 3D seismic to differentiate and understand patterns of similarity / variability within geological elements requires simultaneous consideration of multiple data characteristics or attributes. Techniques for doing this range from co-visualisation of multiple attribute volumes, through mathematical combination of attribute volumes to the use of sophisticated multi-attribute classification schemes.

Although these techniques have been available for some time and their power and capabilities demonstrated (Barnes 2001, Taner 2001, West et al 2002, Chopra et al 2005, Singh 2007), their use in operational seismic interpretation workflows is not as widespread as it ought to be. This is because, in general, the tools available to implement the methods are complex, difficult to use and non-interactive, producing results whose validity is often not easy to understand.

Using numerical classification to divide and label a dataset into distinct classes of similar properties is a widely applied signal/image processing method throughout a number of domains.

In general, the dataset to be analysed is first pre-processed to extract a number of feature signals or images that measure different characteristics of the original data. When used in combination these feature images represent an n-dimensional (for n images) "feature" space, which is subsequently analysed by a suitable classification algorithm. Critical to the overall process is appropriate selection of a set of feature images and appropriate construction of the feature space.

The analog of feature images in seismic data analysis are seismic attributes and effective seismic classification involves the use of a set of appropriate attributes. Classification thus involves examination of a feature space to determine groupings of property / attribute values that represent different elements within the data.

This process is generally termed cluster analysis and many techniques for performing this automatically have been defined for both geophysical and non-geophysical applications (Jain et al 1999). A comparison of many of the existing techniques, relevant to seismic attribute classification has been published recently by Marroquin et al. (2009).

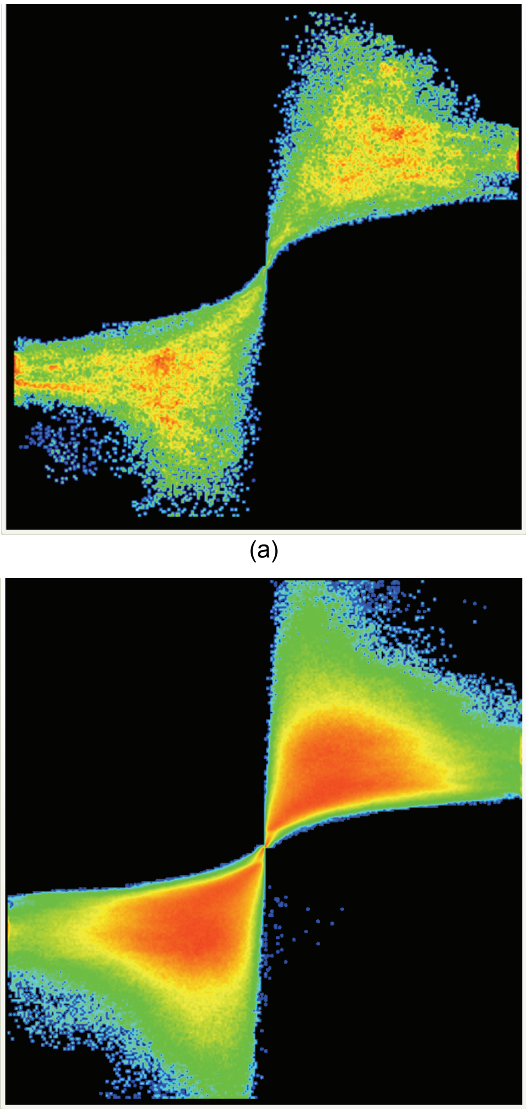
A fundamental issue with automated clustering techniques is that they are looking for structure within a multi-attribute space that either may not exist or may not capture the geological variations that the classification process is aiming to reveal. In addition, the structure of the attribute space is often highly dependent on the subset of data that is examined (figure 1).

A number of distinct clusters are visible in Figure 1a that are not apparent in figure 1b even though the data shown in figure 1b encompasses the data shown in figure 1a. In an intuitive sense, this agrees with the notions provided by the central limit theorem in statistics; as our ROI grows we have an increasingly large sample population, so our distribution will tend towards a smoother more generalised and in some cases normal distribution. In terms of seismic data analysis, where features of interest generally inhabit a particular layer or sequence within the dataset, this translates to ensuring that any training data is limited to the layer, sequence or object of interest.

This variation in topography is one of the reasons why (a) unsupervised classification techniques need to be constrained if they are to give repeatable results and (b) careful selection of training data is required if supervised classification methods are to be applied successfully.

A second reason automated clustering techniques are not widely accepted is that they are often presented as "black box" methods in which the decision criteria are not accessible to the user. This black box nature is not an intrinsic property of automated clustering methods, but stems from it being difficult to convey the internals of such techniques to non expert users in a meaningful way.

Applying these techniques arbitrarily to a large number of seismic attributes is another common pitfall. This can make it very difficult for a user to work backwards from a result to understand how each variable contributed.



(a)

(b)

Figure 1. Cluster space displays of event thickness vs amplitude for two different regions of interest from the same seismic data set showing a much higher degree of complexity in topography the small ROI (a) compared to the large ROI (b)

In addition, the discriminative ability of some classification schemes can be reduced if there is a high degree of correlation between the inputs. Ideally, therefore, in using any classification technique, attempts should be made to restrict the number of seismic attributes presented to the algorithm, based on an understanding of the possible relationships between the attributes and the properties of interest in the data.

Method

Consideration of these issues has led to the development of a very simple but novel methodology that uses interpretative knowledge to drive 3D seismic attribute classification. Using interpretrative knowledge to define classification constraints has several advantages over automated clustering techniques including an understanding of:

1. The geological context in which they are working.
2. The degree of variability in seismic signature that might be associated with a given geological element.
3. Variations in seismic properties due to non-geological factors.
4. The relationships between pairs or groups of seismic attributes and the geological elements that they are expected to relate to.

The new interpreter driven classification technique is based around a visualisation framework that has been designed to make the technique as responsive and interactive as possible and consists of 4 stages:

1. Attribute selection and computation.
2. Attribute normalisation.
3. Class membership selection.
4. Classification.

Stages 1 and 2 are the most computationally intensive; stages 3 and 4 are generally computed quickly enough to allow real time image updating.

The technique allows the use of any pre-computed volumetric seismic attribute or inverted seismic property data, therefore, stage 1 is not discussed further.

Attribute normalisation is required as the classification is based on an analysis of distances between clusters in a N dimensional attribute space where N is the number of input attributes. The normalisation process results in the histogram of each individual input attribute having zero mean and unit variance. This removes any bias on the classification process due to variations in the dynamic / data ranges of the input attributes. The ability to weight the influence of individual attributes on the classification is permitted by explicit selection of attribute weighting factors at the classification stage.

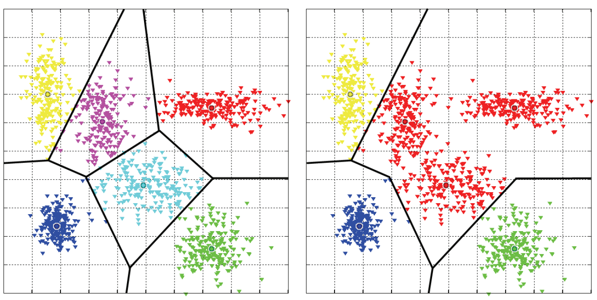
There is no theoretical limit on the number of input attributes that can be used although in the practical implementation this has been restricted to 5 to promote consideration of the most appropriate inputs prior to commencing the classification process.

The information generated by any classification process is dependent on how the points within the dataset are assigned membership to each class. These memberships are defined by a set of rules that determine how the attribute space is partitioned. One method of doing this is to determine a set of cluster “centres” and a metric that determines the position of the membership partitions in multi-attribute space that constrain each class. The simplest and one of the most commonly used measures for defining cluster boundaries is distance and this has been adopted in the first implementation of the novel classification technique described here.

In the new classification method, cluster centres are defined by direct interpreter interaction with the data using a visualisation system that enables the user to view composite displays of the original seismic data and a weighted mixture of the attributes on which the classification will be defined.

This composite mixing functionality allows a user to investigate the information content of the up to 5-dimensional attribute space used for classification, in a spatially and structurally consistent way that cannot be achieved through cross-plotting techniques.

Cluster centres are defined by user definition of individual points or a polygon over structural features of interest within the composite display. For individual points the cluster centre is defined by a 'feature vector' whose coordinates are the values of the input attributes at that point. When a polygon is used the cluster centre is defined as the position in feature space whose coordinates are given by the mean, median or mode of each of the attribute values within the region defined by the polygon. Use of first order statistics means that small polygon selections are the most meaningful. Multiple points & polygons can be "grouped" together providing great flexibility in how the attribute space is partitioned and hence how individual classes are defined (figure 3). This grouping functionality is significant, as it allows the distance classifier to actually discriminate non-linearly, a property normally exclusive to more numerically complex classification processes.



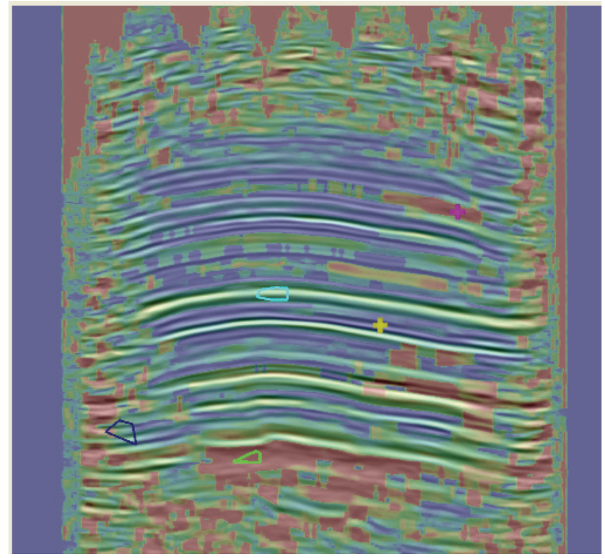
(a) (b)

Figure 2. Schematic representation of the partitioning of an attriute feature space in to 6 classes (a) individual classes; (b) effect on the partioning of grouping 3 classes.

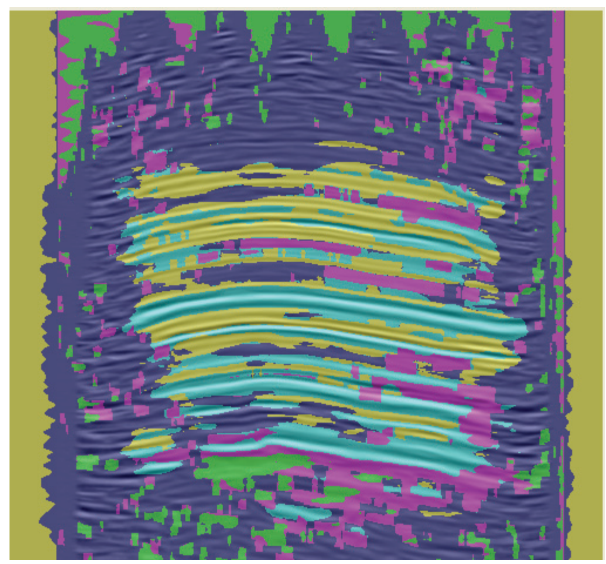
Crucial to the effectiveness of this process is that the classification is updated in "real time" as new cluster centres / classes are added, removed or grouped. Again, appropriate visualisation tools are the keystone to the technique so that the user can assess immediately whether the classification they have defined provides the degree of discrimination required (figure 4).

At this stage the influence of each of the selected input attributes can be varied through changing their individual weighting factors. The weighting factor is a linear scale factor that determines how much influence the attribute has on the minimum distance measures used to define the bounding surfaces separating each class in attribute space. As with adding or removing cluster centres, changing the weighting factors causes the classification result to be updated in real time.

Built into the classification system is a "degree of membership" estimation that gives an indication of the degree of confidence in the class assignment that has been given to each voxel. This is represented visually by varying the colour value assigned to each point dependent on the distance of the point from the class centre. Consideration can then be given to whether new clusters should be defined in areas where the degree of membership is low.



(a)



(b)

Figure 3. (a) Cross-line showing user definition of cluster centres on the composite attribute display, leading to (b) real time attribute classification.

The real time classification update is achieved through use of a Graphics Processing Unit architecture that allows the computation to be carried out in a highly parallel manner on the graphics hardware provided in standard commodity workstations. As the classification is carried out by analysis of an N dimensional attribute space whose size is dependent on the number of input attributes and not the size of the seismic volume under investigation, the computation process is largely independent of the size of the input data sets. To achieve the required performance, the process has been implemented so that it only produces the classification result in a visual format until the user is satisfied with the result and selects to compute a full 3D class volume.

The technique provides for two modes of volumetric computation, in the first each voxel is assigned a single class value and in the second each voxel is assigned both a class value and a degree of membership value. Computation of a classified volume is dependent on the size of the ROI or data set being analysed and can be done as a batch processing job that can be linked in with follow on analysis such as connected component labelling to generate sets of geobodies associated with particular classes.

Results

The new technique is undergoing thorough field testing and examples of its application to data from onshore US will be presented (e.g. figure 4).

Conclusions

A multi-attribute classification technique has been developed that is unique in both the degree of control that the interpreter has over the classification process and the speed and interactivity with which a classification is achieved. The technique addresses some of the fundamental issues that have hampered the acceptance of attribute classification schemes within mainstream interpretation workflows.

The technique described is extremely flexible, requires no separate training data or training cycle, has no theoretical constraints on the number and types of attributes used and is scale independent meaning it can be utilised to partition whole 3D seismic volumes or individual seismic events. It can be constrained to a variety of 3D regions of interest enabling a hierarchical analysis methodology.

By putting the interpreter at the heart of the classification process and providing direct and instantaneous feedback on the consequence of varying the classification parameters removes the concerns about the “black box” nature of many automated clustering or neural network architectures. The speed and ease of use of the new technique has the potential to allow multiple realisations to be created very rapidly enabling the robustness of the classification result to be evaluated very easily.

Acknowledgments

Data courtesy of RMTOC and US Dept of Energy.

References

Jain, A.K., M.N. Murty and P.J. Flynn, 1999, Data clustering: A review. ACM Computing Surveys, 31, 264-323.

Marroquin, I.D., J-J. Brault and B.S. Hart, 2009, A visual data-mining methodology for seismic-facies analysis: Part 1 – Testing and comparison with other unsupervised clustering methods, Geophysics, 74, P1-P11.

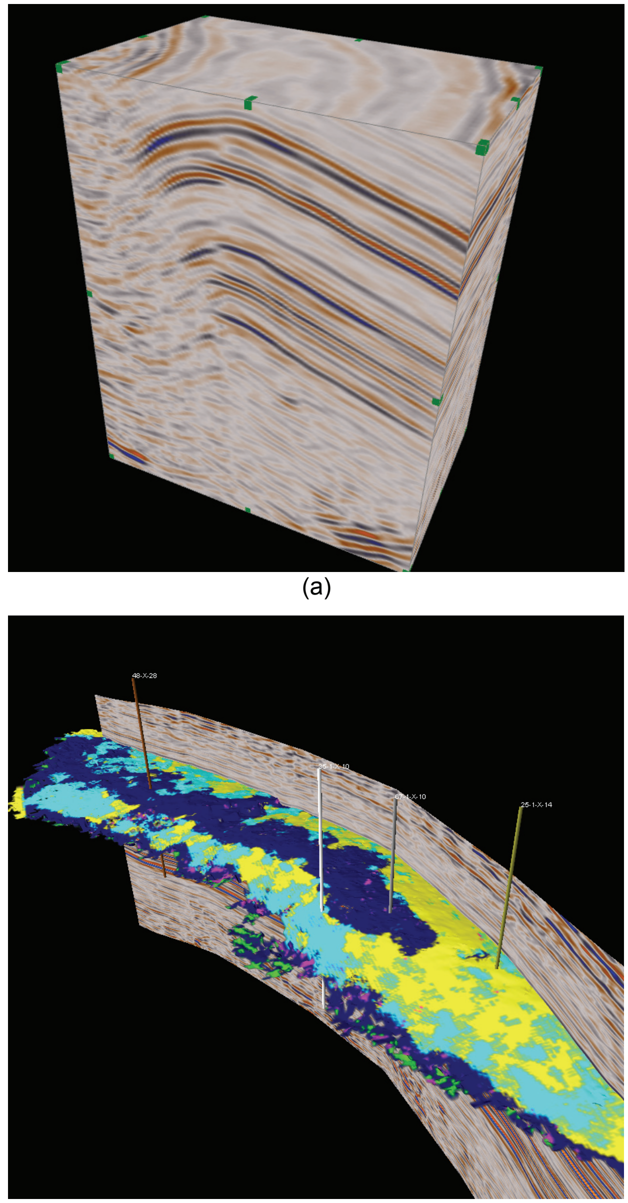
Barnes, A.E, September 2001, Seismic attributes in your facies, CSEG Recorder, P41-P47.

Taner, M.T., September 2001, Seismic attribute, CSEG Recorder, P49-P56.

West, B.P., May, S.R., Eastwood, J.E., Rossen, C., October 2002, Interactive seismic facies classification using textural attributes and neural networks, The Leading Edge, P1042-P1049.

Chopra, S., Alexeev, V., 2005, Application of texture attribute analysis to 3D seismic data, CSEG Recorder, 2005, P28-P32.

Singh, Y.P, 2007, Translation invariant attributes for effective lithofacies discrimination, Geophysics, 72, No.6, P57-P66.



(a)

(b)

Figure 4. Example of the application of the new Interpreter Driven 3D volume classification technique to differentiate continuous parallel facies from chaotic facies and potential small scale faulting (a) Input reflectivity cube, (b) Horizon extraction from the classified volume.