Identifying Reflectors in Seismic Images via Statistic and Syntactic Methods

Carlos A. PEREZ
Petroseismic Group, Industrial University of Santander
Bucaramanga, Santander, Colombia

and

German Y. OJEDA
Geophysics Group, Ecopetrol
Piedecuesta, Santander, Colombia

ABSTRACT

In geologic interpretation of seismic reflection data, accurate identification of reflectors is the foremost step to ensure proper subsurface structural definition. Reflector information, along with other data sets, is a key factor to predict the presence of hydrocarbons. In this work, mathematic and pattern recognition theory was adapted to design two statistical and two syntactic algorithms which constitute a tool in semiautomatic reflector identification. The interpretive power of these four schemes was evaluated in terms of prediction accuracy and computational speed. Among these, the semblance method was confirmed to render the greatest accuracy and speed. Syntactic methods offer an interesting alternative due to their inherently structural search method.

1. INTRODUCTION

Automation of reflector identification in seismic images based on statistical parameters has been the subject of many studies in pattern recognition. These methods include neural networks, coherence cube [2], fuzzy logic and genetic algorithms [7].

Besides the statistical methods, the syntactic pattern recognition theory has been argued as a useful approach [1]. Syntactic methods are based on the formal language theory and finite-state automata. This theory is widely used in structural pattern recognition problems like character recognition [8], texture analysis [3], among others [9].

However, no clear workflow has been proposed for the specific case of computer-assisted reflector identification through syntactic methods. In this study, simple syntactic workflows were developed using finite-state automata and Levenshtein criteria. Results from these workflows were compared to results obtained from semblance coefficient and temporal coherence. The performance of syntactic methods was acceptable relative to statistical methods in continuous reflectors, as well as in discontinuous reflectors that are not offset more than their wavelength.

2. THEORY

Statistical methods applied to seismic traces are based on measurements of the similarity between a pattern and a sample signal. As a result, these methods select the most correlated samples among the input and sample signals (Figure 1).

Measures of similarity considered in this work were temporal coherency *C(k)* [5] and semblance coefficient [6], defined in their discrete form by Equations (1) and (2) for *N* traces $s_i[n]$,

$$C(k) = \frac{\left\langle s_1[n] s_2^*[n+k] \right\rangle}{\sqrt{\left\langle \left| s_1[n] \right|^2 \right\rangle \left\langle \left| s_2[n+k] \right|^2 \right\rangle}} \tag{1}$$

$$R = \frac{\sum_{t=T1}^{T2} F^2[n] - \sum_{i=1}^{N} \sum_{n=T1}^{T2} s^2{}_i[n]}{(N-1) \sum_{i=1}^{N} \sum_{n=T1}^{T2} s^2{}_i[n]} \tag{2}$$

where *F[n]* is the stacked trace, that is, the sum of the *N* traces between samples *T1* and *T2*.

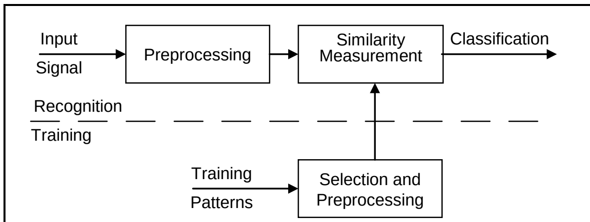

Figure 1: Block diagram of a statistic pattern recognition system

Syntactic methods take advantage of a specific use of the formal language theory. Those methods are based on treating numeric vectors (discrete signals) as character strings, and assuming rules for codification. Strings obtained in this manner are used to define grammars and languages that identify groups with defined similarities [1].

A typical syntactic recognition system, including the training and classifier networks, is showed in Figure 2. In the training network, the language associated with a pattern is inferred, and in the classifier network the sample signal is evaluated relative to the language with an automata or a distance criterion. Freeman's codification may be used to obtain the primitive representation of input signals [1]. Similarity is measured differently in the automata and Levenshtein methods. In the syntactic theory based on automata, similarity is measured as the path that renders the lowest cost -in production terms-, of a string from a language defined by a pattern string [1], [10]. On the other hand, since the minimum distance Levenshtein criterion can find the distance directly from a pattern string to the input string [1], [10] this approach does not need a grammar or language.

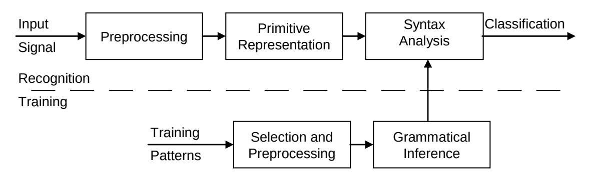

Figure 2: Block diagram of a syntactic pattern recognition system

3. RESULTS

Results were found by evaluating the two statistical (semblance and coherence) and two syntactic (automata and Levenshtein) methods on a database of 5 synthetic seismic images and 1 real image previously interpreted by hand. An example of our image database may be seen in Figure 3, showing the principal reflectors (identified by red circles) and main faults (interpreter solution). A given test was considered successful if a reflector was identified left and right of a fault.

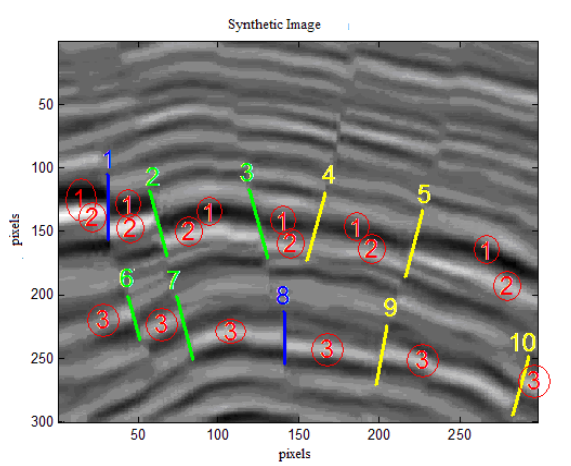

Figure 3: Example image of the database used in our tests.

The total number of tests was 42, 14 of which included faults dipping to the left ("negative), 14 with faults dipping to the right ("positive"), and 14 with vertical faults. Results are showed in Tables 1 and 2. Note that the semblance coefficient method rendered the best overall accuracy (Table 2).

Algorithm	Fault Dip	Correct Results	Accuracy
Coherence	Positive	14/14	100%
Semblance	Positive	14/14	100%
Automata	Positive	12/14	85.15%
Levenshtein	Positive	12/14	85.15%
Coherence	Negative	11/14	78.57%
Semblance	Negative	14/14	100%
Automata	Negative	12/14	85.15%
Levenshtein	Negative	12/14	85.15%
Coherence	Vertical	14/14	100%
Semblance	Vertical	13/14	92.86%
Automata	Vertical	12/14	85.15%
Levenshtein	Vertical	13/14	92.86%

Table 1: Results of the algorithms versus different fault dips.

Algorithm	Correct	# of tests	Accuracy
Coherence	39	42	92.86%
Semblance	41	42	97.62%
Automata	36	42	85.72%
Levenshtein	35	42	83.33%

Table 2: General results of the algorithms.

For the computational speed test, the seismic image on Figure 3 was evaluated, taking window sizes of 3 up to 21 samples (*T1* to *T2*). This test showed the dependence of the algorithm on window size. The test was done using a 64 bit processor running at 1.8 GHz with 512 MB RAM. Our speed test indicates that statistic methods have an approximately linear increase in time, while syntactic methods have an approximately exponential icrease in time. The method with the lowest computational time was the semblance coefficient (300 traces of 20 samples each, processed in 0.27 sec), while the slowest method was the automata.

In addition to our 6 test images, coherence and Levenshtein algorithms were applied on a seismic image from the Colorado oil field in Colombia. Figure 4 shows the interpretation by hand and two reflectors identified by a combined syntactic - statistical approach.

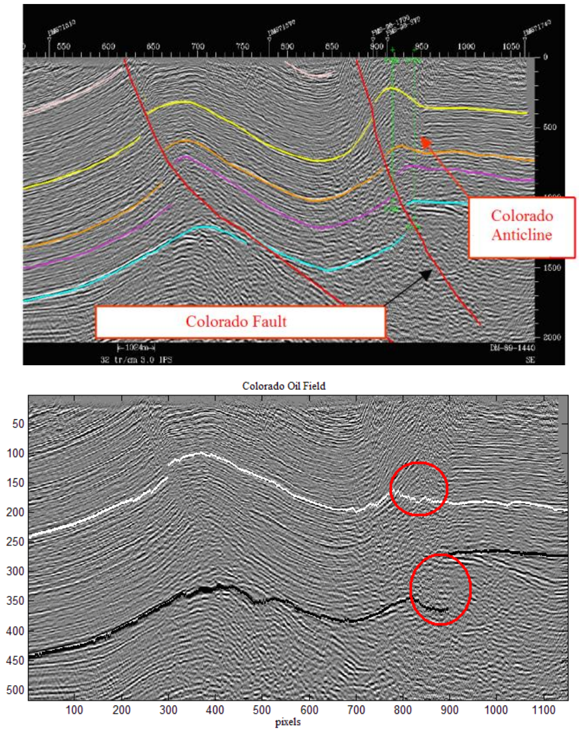

Figure 4: Test with a combined syntactic – statistical approach on real data from Colombia. Top: interpretation by hand. Bottom: semi-automatic identification by combined approach.

Real data do not allow an exact validity test, since the actual subsurface structure is never perfectly known. However, perusal of results in Figure 4 indicates great coincidence among the manual and semi-automatic interpretations. Red circles in Figure 4 show areas of discrepancy among the two interpretations, which is due to complex structure.

4. CONSLUSIONS

The semblance algorithm provided the greatest overall reflector recognition accuracy (97.62%). Coherence, automata and Levenshtein algorithms showed a relatively good performance (over 83.3%). The computational speed tests confirm that the semblance coefficient algorithm offers the best performance. Results confirmed that reflectors identified by the proposed methods coincide with interpretations done by experts. Combining statistical with syntactic approaches may offer an interesting alternative.

5. ACKNOWLEDGEMENTS

The authors would like to thank Ecopetrol for their cooperation in providing the data and support, and the help received in Universidad Industrial de Santander.

6. REFERENCES

[1] K. Y. Huang. **Syntactic Pattern Recognition for Seismic Oil Exploration**. Series in Machine Perception and Artificial Intelligence – Vol. 46, 2002.

[2] M. Bahorich, S. Farmer. **The Coherence Cube**. The Leading Edge 14(10). 1053-1058, oct 1995.

[3] K. S. Fu. **Syntactic Pattern Recognition and Applications**. Prentice-Hall, Englewood Cliffs, NJ, 1982.

[4] N. Neidell, M. Taner. **Semblance and other coherency measures for multichannel data.** Geophysics 36(3). 482-497, jun 1971.

[5] W. A. Gardner. **Exploitation of spectral correlation in cyclostationary signals**. Spectrum Estimation and Modeling, Fourth annual ASSP workshop on. Published: 1988. Pages: 1-6.

[6] M. Tanner. **Semblance and other similarity measurements.** Rock Solid Images, nov 1996.

[7] M. Aurrnhammer and K. Tönnies. **A genetic algorithm for automated horizon correlation across faults in seismic images**. IEEE Transactions on evolutionary computation, Vol. 9, No. 2, April 2005.

[8] F. Ali and T. Pavdilis. **Syntactic recognition of handwritten numerals.** IEEE Transactions on Syst. Man Cybern. 7 (1977) 537-541.

[9] H. Bunke and A. Sanfeliu. **Syntactic and structural pattern recognition – Theory and applications**. World Scientific, 1990.

[10] C. A. Pérez. **Identificación de reflectores en imágenes de sección sísmica aplicando métodos estadísticos y métodos sintácticos**. Undergraduate thesis. Universidad Industrial de Santander, January 2008.