Semi-automatic detection of faults in 3D seismic data

Kristofer M. Tingdahl[1]* and Matthijs de Rooij[2]

[1]*Department of Earth Sciences, Göteborg University, Box 460, 405 30 Göteborg, Sweden, and* [2]*dGB Earth Sciences, Nijverheidstraat 11-2, 7511 JM Enschede, The Netherlands*

ABSTRACT

The semi-automated detection of objects has been quite successful in detecting various types of seismic object, such as chimneys. The same technique can be applied successfully to detect faults in 3D seismic data. We show that several different attributes – among others, similarity, frequency and curvature, all of which potentially enhance the visibility of faults – can be combined successfully by an artificial neural network. This results in a fault 'probability' cube in which faults are more continuous and noise is suppressed compared with single-attribute cubes. It is believed that the fault-cube can be improved further by applying image-processing techniques to enhance the fault prediction.

INTRODUCTION

Traditionally, faults are interpreted by manual picking on vertical seismic sections and time-slices. This technique relies on the ability of the interpreter to create geologically consistent fault models. In particular, complex fault systems are not well represented on vertical seismic sections and the interpretation derived from these sections relies on subjective factors, such as the ability of the interpreter to imagine 3D objects and to define and adhere to consistent fault criteria. Coherence cubes (Bahorich and Farmer 1995; Marfurt *et al.* 1998) are often used to highlight reflector edges and to guide the picking process. Enhanced attribute cubes can be even more helpful and are used to automate parts of the manual workflow. In addition to the traditional coherence attributes and their derivatives, numerous other attributes for fault detection have been developed. We see this as a positive development, but it also poses new challenges for the interpreter. Most attributes are non-unique; one attribute can respond to other objects as well as just faults. There is no perfect fault attribute, but many attributes can contribute to fault interpretation. There are basically two strategies for handling this situation.

Firstly, the interpreter can display numerous attributes simultaneously, requiring a powerful imaging workstation. Secondly, the interpreter can combine the different attributes with some user-defined mathematical expression. The problem with the latter strategy is that no such mathematical expression exists, and if one is constructed for one data set and one type of fault, it will probably not be valid for another data set or other fault types. We present a way of automatically finding the optimal relationship between the presence of different types of fault at a location and multiple attributes.

The semi-automatic approach to seismic object detection, developed by Meldahl *et al.* (1999), uses directive attributes (i.e. attributes that are extracted in user- or data-steered directions) and neural networks to highlight objects of interest. The method was initially applied to highlight seismic chimneys for the interpretation of hydrocarbon migration paths (Heggland *et al.* 1999; Tingdahl, Bril and de Groot 2001). The same methodology can be used to detect other seismic objects, such as faults, reflectors and bright spots. This paper focuses on adapting the methodology to detect faults in 3D seismic data.

METHOD

The seismic-object detection method comprises several steps. First, the seismic data are scanned and faults are picked

---

*E-mail: kristofer.tingdahl@dgb-group.com

manually on a few sections and time-slices. Locations away from the faults are picked randomly by hand. All selections are made on raw seismic data and/or single-attribute sections and time-slices.

Next, a large set of seismic attributes is calculated at the picked fault and the non-fault locations. The selection of attributes and their parameters is based on experience, intelligent guesswork and visual inspection of the data.

Following this step, a multilayered perceptron artificial neural network is trained to give a high output value (1 in our case) at the fault locations and a low output value (0) at the non-fault data set. The training is performed on the major fraction of the fault/non-fault locations. Typically 10% to 25% of the data is used for this blind testing during training. This ensures the quality of the training and helps to determine the point at which the training must be stopped (e.g. de Groot 1999).

Finally, the trained neural network is applied to the entire data cube. The selected attributes are calculated at every sample position and from these the trained neural network calculates an output value. The output is a cube with numerical values indicating the presence of faults by aligned clusters of high probability values. The values range from 0 to 1; a high value indicates a high fault probability and a low value indicates a low fault probability.

Directionality

One of the key aspects of the applied seismic-object detection is the directionality principle, which states that attributes are preferably extracted along the events we seek to detect. This works very well while detecting objects with a known dip and azimuth, such as seismic chimneys (Heggland *et al.* 1999; Meldahl *et al.* 1999; Tingdahl *et al.* 2001). However, it can cause a problem with faults, as the directions of faults are generally unknown. On some data sets, only a few, if any, fault orientations prevail. The faults tend to have different directions in different areas of the data. An attempt could be made to construct a detection procedure in which faults in every major orientation are detected separately, but this would introduce a bias and would take more time than it saves.

A better approach is to run the detection process iteratively. In the first run, a fault probability is calculated using only the directionality of the seismic events. In an optional second pass, the dip and azimuth of the faults are calculated from the first-generation fault-probability cube. This dip and azimuth information is subsequently used to add directivity information to attributes used in the second pass. We show and discuss only the results of a so-called first-iteration fault-cube, since

this provides an adequate illustration of the use and value of the approach presented.

ATTRIBUTES

There is no general set of attributes or attribute parameters for fault detection using artificial neural networks, since the character of the faults of interest and the data quality differ between data sets. In the following sections, we describe a number of seismic attributes that define different characteristics of faults and are often included in the artificial neural network fault-detection scheme.

Energy

We define the energy E as the summed squared amplitude within a gate around the sample of investigation, i.e.

$$E = \frac{\sum_{i=0}^{N-1} f(t_0 + i \cdot dt)^2}{N}, \tag{1}$$

where f is the amplitude of the trace, t_0 is the upper limit of the gate, N is the number of samples in the gate and dt is the sample interval.

Similarity

The similarity attribute S between two trace-segments is defined as the distance in hyperspace between the vectors of the segments, normalized to the sum of the lengths of the vectors, i.e.

$$S = 1 - \frac{|\mathbf{v} - \mathbf{u}|}{|\mathbf{v}| + |\mathbf{u}|}, \tag{2}$$

where

$$\mathbf{v} = \begin{bmatrix} f(t_1, x_v, y_v) \\ f(t_1 + dt, x_v, y_v) \\ \vdots \\ f(t_2 - dt, x_v, y_v) \\ f(t_2, x_v, y_v) \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} f(t_1, x_u, y_u) \\ f(t_1 + dt, x_u, y_u) \\ \vdots \\ f(t_2 - dt, x_u, y_u) \\ f(t_2, x_u, y_u) \end{bmatrix}, \tag{3}$$

and t is the time-depth of investigation, dt is the sampling interval, t_1 and t_2 are the limits of the time gate, (x_v, y_v) and (x_u, y_u) are the two trace positions that are to be compared, and f is the amplitude value in the cube. Compared with the traditional coherence (Bahorich and Farmer 1995), similarity also takes into account the amplitude differences between the

two trace-segments. Faults are discontinuities in the data that give a low response to the similarity.

The problem with trace-to-trace measurements for fault detection is the sensitivity to the lateral positions of the trace-segments. The fault is detected only if the trace-segments are situated on either side of the fault. A solution is to calculate multiple similarity measurements with different orientations and different distances between the trace-segments. This yields a number of similarities, from which one single attribute is computed by taking the minimum, maximum or the average of the contributing similarities. We compute all possible combinations of similarities within a square of 3 × 3 traces (a total of 36 similarity-pairs) and the minimum of these is the output of the 'cube-similarity' attribute.

The group of 36 similarity-pairs will be quite similar in coherent areas away from faults. They will, however, be rather inhomogeneous close to the faults, since some of the pairs will cross the fault-border. This behaviour can be quantified by computing the statistical variance of the 36 pairs, which we call the 'cube-similarity-variance' attribute.

Dip-steering

Like most other trace-to-trace comparisons, similarity also gives a high response at high structural dips. This is sometimes desirable, but not when the objective is to detect faults. Using the directionality principle as described in the method above, we can decrease the similarity's sensitivity to structural dip by aligning the two trace-segments with a time lag. We calculate this time lag with a technique called dip-steering (Tingdahl 1999; Tingdahl *et al.* 2001). Dip-steering uses the local dip and azimuth to track the event locally to the trace-segments under investigation. Compared with techniques that only extrapolate the local dip plane all the way to the trace-segment of interest (Marfurt *et al.* 1998), the dip-steering technique follows the seismic event by using the local dip and azimuth information at every position along the track (Fig. 1).
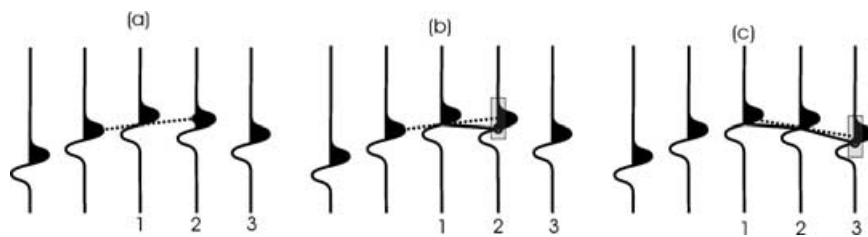

Figure 1 The steps of auto-adaptive dip-steering from trace 1 to trace 3. Dashed lines represent the local dip and azimuth; solid lines between the traces represent the path of steering. (a) The local dip and azimuth are followed from the starting trace to the next trace on the path towards the target trace. (b) Optionally, an aperture centred on the intercept-time (grey) can be searched for the same phase as at the starting point on trace 1. If the phase is found within the aperture, the time is adjusted to the time of equal phase. If not, the intercept-time is assumed to be the best available estimate. (c) The local dip and azimuth are calculated at the (optionally phase-adjusted) intercept-time and are followed further along the path towards the target trace. The procedure of calculating local dip and azimuth, following their direction and optionally the phase lock, is repeated along the path to the target trace (de Rooij and Tingdahl 2002).

Table 1 All attributes used and their parameter settings in the artificial neural network

| Attribute	Time gate (ms)	Lateral position	Other settings
Energy	[−32,32]	–	–
Similarity	[−24,24]	(0,1) and (0,−1) 90° rotated	Minimum similarity is output
Steered similarity	[−24,24]	Positions (0,1) and (0,−1) 90° rotated	Minimum similarity is output
Cube similarity	[−24,24]	Full block	Full steering, minimum similarity is output
Cube similarity variance	[−24,24]	Full block	Full steering
Frequency-AFS	[−36,36]	–	–
Frequency-FSF	[−36,36]	–	–
Dip angle	–	–	–
Dip variance	[−32,32]	Stepout 1	–
Most positive curvature	–	Stepout 2	–
Most negative curvature	–	Stepout 2	–
Disc variance	–	Stepout 3	Full steering
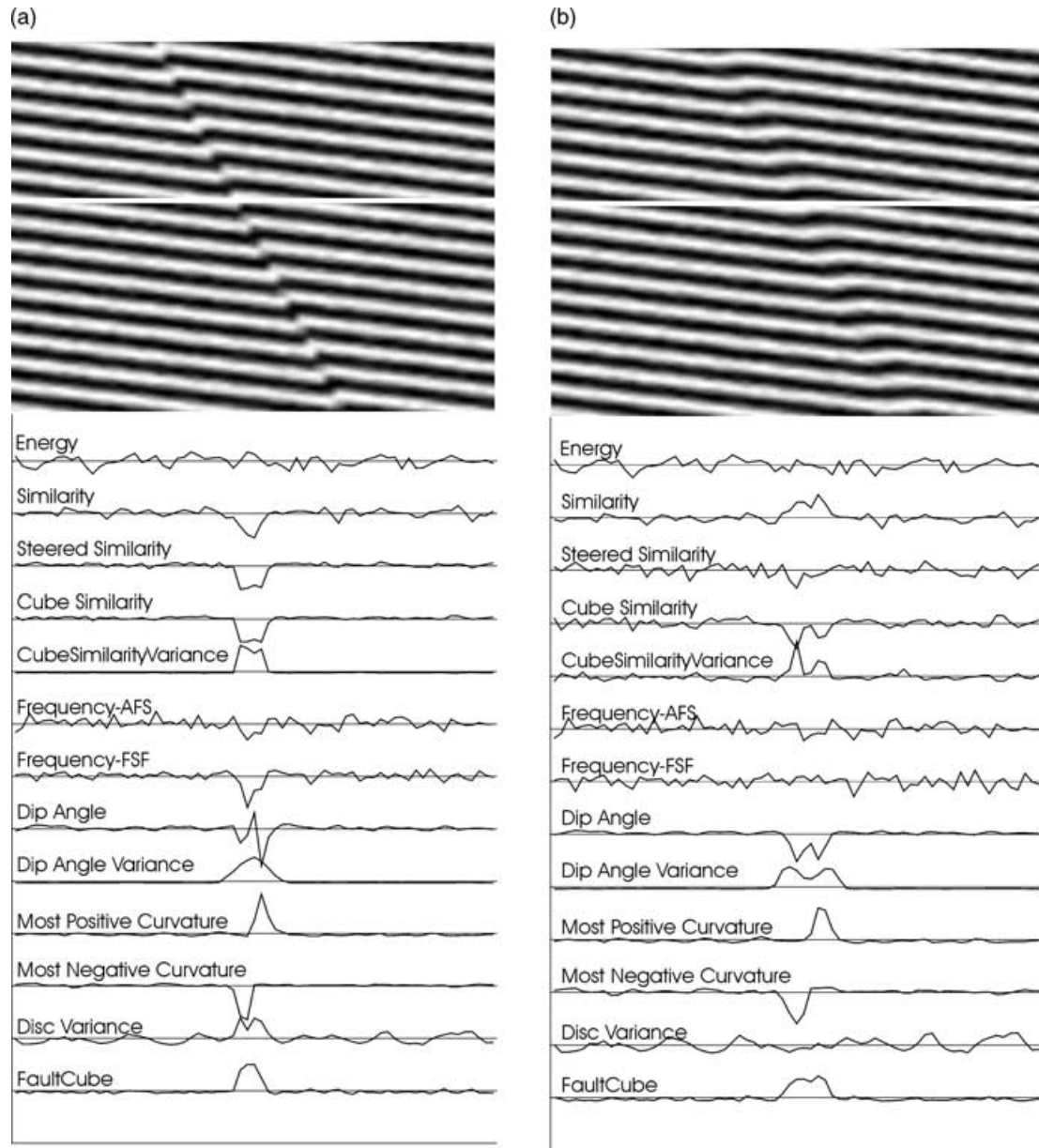
Figure 2 (a) and (b) Synthetic data with faults. The data are generated using a 40 Hz sine signal with 10% random noise added. The dipping layers have a dip of 75 μs/m. The section is 1000 m wide. The 12 attributes (Table 1) are computed along the white line through the faults. The attributes are normalized to a zero-mean and a standard deviation of 1.

Dip variance

At fault locations, the dip often varies more than elsewhere. This behaviour is quantified by the dip variance and is calculated as the statistical variance of the dip values inside a small subcube of the dip data:

$$\mathrm{var}\,(p_x) = \frac{1}{n-1} \sum_{\beta=-x_s}^{x_s} \sum_{\alpha=-y_s}^{y_s} \sum_{\tau=a}^{b} (p_x\,(x+\alpha,\,y+\beta,\,t+\tau) - \bar{p}_x)^2, \quad (4)$$

where

$$\bar{p}_x = \frac{\sum_{\beta=-x_s}^{x_s} \sum_{\alpha=-y_s}^{y_s} \sum_{\tau=a}^{b} p_x\,(x+\alpha,\,y+\beta,\,t+\tau)}{n}, \quad (5)$$

and n is the total number of terms in the triple summation, x_s and y_s are the maximum trace stepout values from the centre of the small subcube in the x- and y-directions, respectively, a
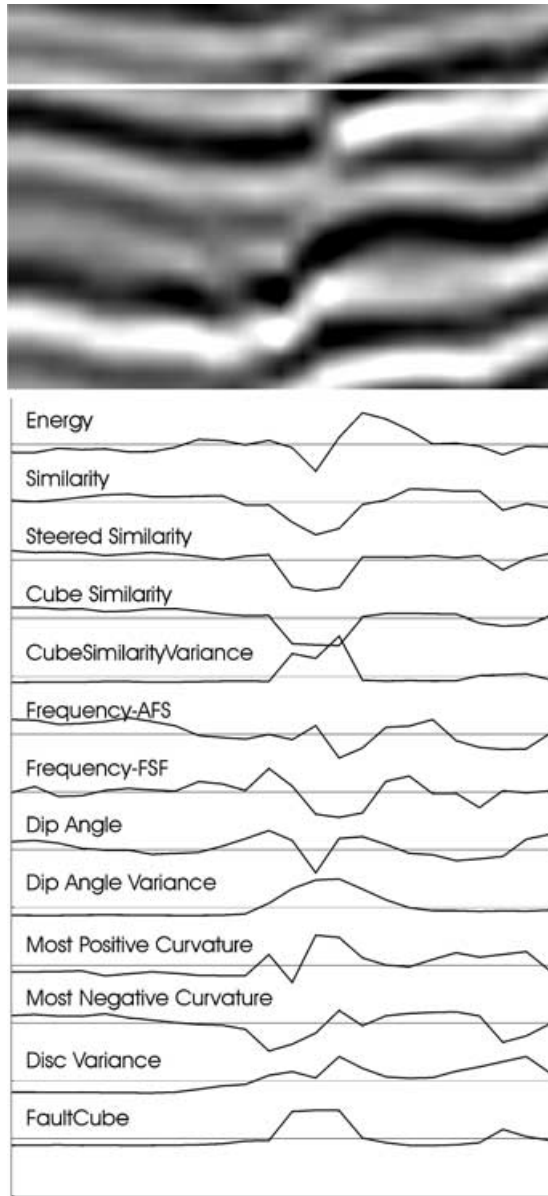
Figure 3 Detail from the real data set. The 12 attributes (Table 1) are computed along the white line through the faults. The attributes are normalized to a zero-mean and a standard deviation of 1.

and b are the upper and lower limits of this subcube, and p is the dip cube.

In our data volume, the 'dip' p is expressed in seconds per metre, so it has the same appearance as slowness. It is not constructive to calculate the variance in slowness directly. At high dips, the slowness values will tend to positive or negative infinity. A dip varying by $\pm 5°$ between $5°$ and $10°$ will yield a slowness variance that is completely different from a

$5°$ dip variation between $75°$ and $80°$. It is better to perform variance analysis on slowness converted to dip, where the above-mentioned problem does not exist. The use of one single uniform velocity suffices in the conversion to dip, since it is the variance that we are interested in, not the absolute values.

Disc variance

If a fault is sharp, the events or horizons will be cut off by the fault. If the horizon is horizontal, the amplitudes on a small, local time-slice will have a higher variance at the fault locations than elsewhere. We can therefore use the 'horizon' variance as a fault indicator. A horizon is rarely horizontal, so we use the dip-steering technique to follow the local dip and azimuth information. The success of the latter strategy depends on the character of the faults. If the fault is broad, the dip-steering might connect the horizon on both sides of the faults and the desired effect will therefore not be found.

Curvature

Curvature (Roberts 2001) is a family of attributes that characterizes the local shape of a horizon. Curvature analysis is normally limited to manually picked or auto-tracked horizons. We use dip-steering to track a horizon segment locally from the point of investigation. Since we have dip and azimuth information at every sample position in a cube, we can track a horizon segment at every position. Consequently, the curvature-attribute family is extended to an application in full 3D mode. Curvature attributes require nine points on a horizon to be defined. We utilize dip-steering to find the eight positions around the central position, in order to define the required 'horizon' (Tingdahl and de Groot 2003).

Frequency

The local frequency response is likely to be affected when the seismic records a sharp change caused by a fault. We include the average frequency squared (*AFS*) attribute and the frequency slope fall-off (*FSF*) to capture this response. The *AFS* is the amplitude-weighted mean squared frequency, given by

$$AFS = \frac{\sum\limits_{i=1}^{N} p\,(\omega_i)^2\,\omega_i^2}{\sum\limits_{i=1}^{N} p\,(\omega_i)^2}, \qquad (6)$$
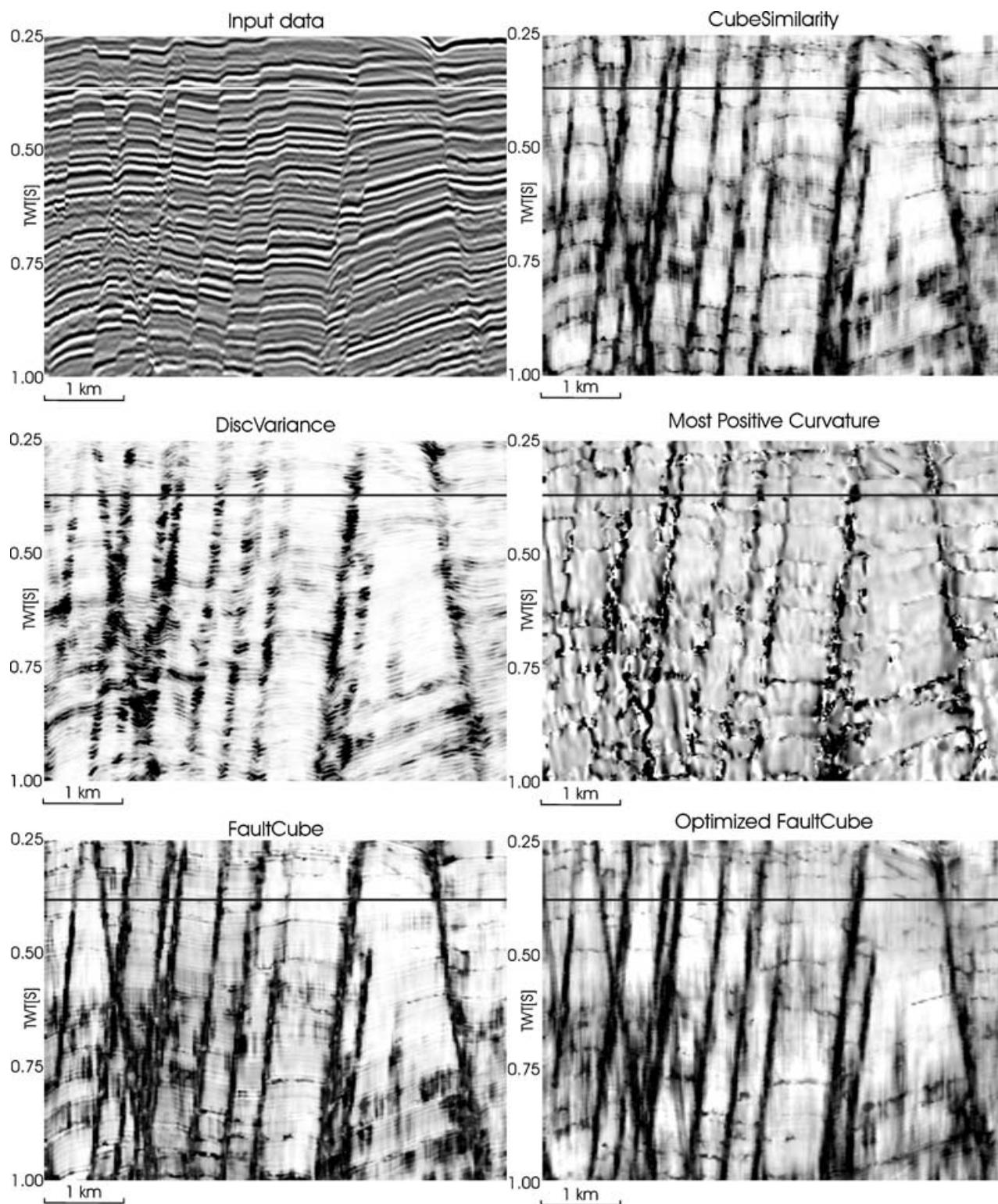
Figure 4 Section showing the raw seismic data, some of the attributes used and two different fault-cube results. The positions of the sections are indicated on the time-slices in Fig. 5.
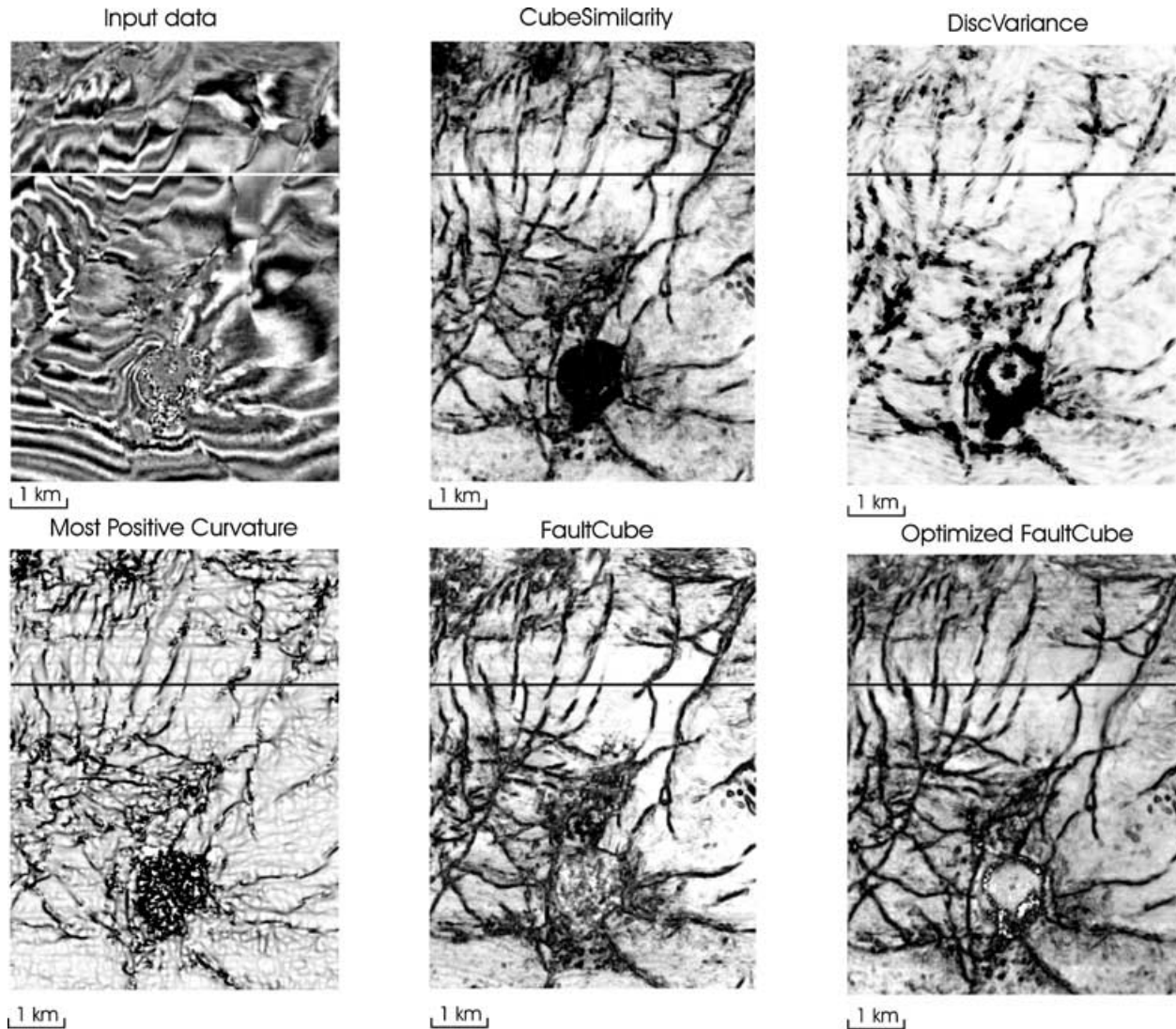
Input data

CubeSimilarity

DiscVariance

Most Positive Curvature

FaultCube

Optimized FaultCube

Figure 5 Time-slices of the same attributes as in Fig. 4, at the depth indicated in Fig. 4.

where $p(\omega)$ is the complex response of a trace-segment in the Fourier domain and N is the length of the signal.

The *FSF* measures the peak shape of the response of the trace-segment in the Fourier domain and is given by

$$FSF = 1 + \frac{p_{max}^2 - \sum\limits_{i=1}^{N} p(\omega_i)^2}{p_{max}^2 + \sum\limits_{i=1}^{N} p(\omega_i)^2},\tag{7}$$

where p_{max} is the highest peak in the frequency domain. If all energy is concentrated in the highest peak, the *FSF* value will be close to 1; if the peak is very broad, the *FSF* response will be close to 0.

RESULTS

We apply these types of attribute to two different synthetic seismic cubes with a single fault and to a real data volume with numerous faults. To facilitate comparison, and to demonstrate the added value of the combination of attributes, we first construct a single-attribute set. Based on a parameter scan and the behaviour of the parameters at the fault, the 'best compromise' attribute set is constructed, giving good results for all three data sets simultaneously. The attribute parameters are shown in Table 1.

This is not the standard procedure; it is performed in order to show that different attributes make different contributions to different data sets. Normally, each data set requires its own attribute and parameter selection, and less significant

Table 2 All used attributes and their parameter settings in the optimized artificial neural network

Attribute	Time gate	Lateral position	Other settings
Reference time	–	–	–
Energy	[−80,80]	–	–
Energy	[−16,16]	–	–
Energy average	[−16,16]	Stepout 2	No steering
Dip variance	[−60,60]	Stepout 1	–
Dip variance	[−16,16]	Stepout 6	–
Cube similarity	[−40,40]	Stepout 1	Full steering; minimum similarity is output
Similarity	[−60,60]	(−1,−1) and (1,1) 90° rotated	Full steering; minimum similarity is output
Similarity	[−32,32]	(−1,−1) and (1,1) 90° rotated	Full steering; minimum similarity is output
Similarity	[−60,60]	(−1,0) and (1,0) 90° rotated	Full steering; minimum similarity is output
Similarity	[−32,32]	(−1,0) and (1,0) 90° rotated	Full steering; minimum similarity is output
Similarity	[−32,32]	(−1,0) and (1,0) 90° rotated	No steering; minimum similarity is output
Frequency	[−32,32]	–	*FSF*
Most positive curvature	–	Stepout 2	–
Most negative curvature	–	Stepout 2	–
Disc variance	–	Stepout 1	Full steering
Disc variance	–	Stepout 2	Full steering
Disc variance	–	Stepout 3	Full steering

attributes are removed to speed up the calculation. Therefore, for comparison we also show the result obtained from the real data set using an optimized set of attributes.

Synthetic data sets

Figure 2 shows the results of all individual attributes and the combined fault-cube result on both synthetic data sets. The non-steered similarity attributes have a noisier background level. Consequently, the contrast at the fault is not as good as with the steered similarities. However, in the broad fault (Fig. 2b) we observe a drawback of steered similarity. The fault is so broad that the steering manages to follow the horizon through the fault, resulting in a high similarity across the fault, compared with the non-steered similarity.

The success of the disc-variance attribute also depends on the fault character. If the fault is broad, the dip-steering might connect the horizon on both sides of the faults and the desired effect will therefore not be recorded.

We also observe that the 'most positive' curvature attribute detects the positively curved structure directly to the right of the faults, while the 'most negative' curvature detects the negative curvature directly to the left of the faults.

In general, the individual attributes in the sharp fault (Fig. 2a) provide a clear definition of the fault. The similarity attributes in particular show a good resolution. In the broad fault, however, only few of the attributes clearly indicate the location of the fault, although the dip-angle attribute does indicate the correct location of the fault.

Real data

The real data set was obtained from offshore Nigeria. Figure 3 shows profiles of the same attributes as in Fig. 2. In this case, the dip angle and the similarity attributes all clearly highlight the fault, but the fault-cube result derived from all attributes is superior to any of the individual attributes.

Figures 4 and 5, respectively, show a number of sections and time-slices with various attributes. The last slice and section in these figures also show the result using an attribute set optimized for this data volume. The complete attribute set used to generate this fault-cube is summarized in Table 2.

Figure 6 Neural-network topology of the artificial neural network from the real data, using the attributes in Table 1. The darkness of the nodes indicates the relative weight assigned to the input attributes: a dark node has more weight than a light node.
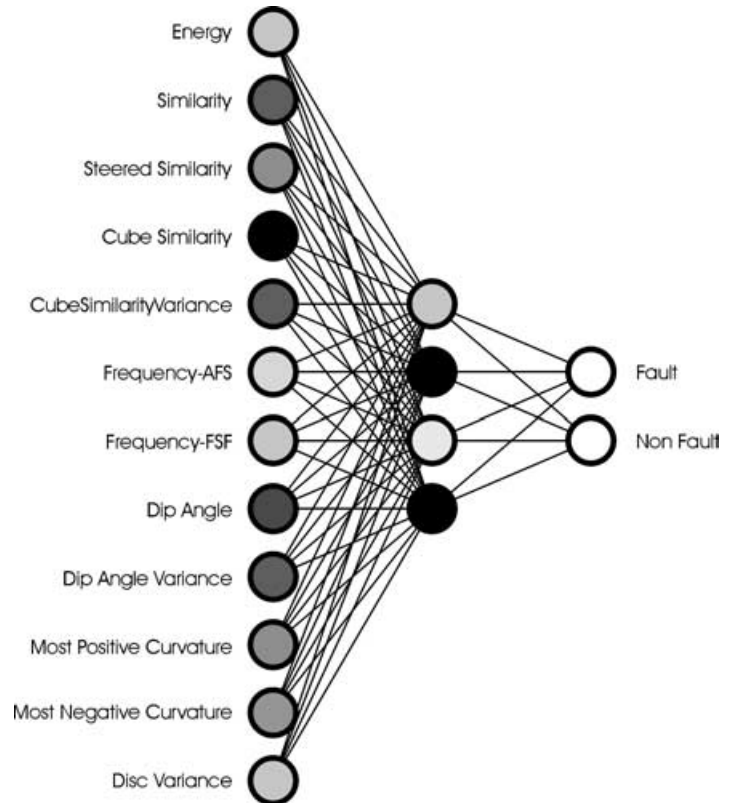
For the fault-cube with attributes shown in Table 1, 359 positions were picked to represent the faults and 285 positions were picked on non-fault locations. For the optimized fault-cube, 4534 and 5314 positions were picked, respectively, while the attributes were calculated according to Table 2. In both cases, the artificial neural network was trained on 75% of the example locations, while the remaining 25% were used for blind testing. The darkness of the nodes in Fig. 6 shows the relative contribution of the individual attributes to the fault-detection output of the artificial neural network with the attributes presented in Table 1.

DISCUSSION AND CONCLUSION

The analysis presented here shows that faults can be detected in a semi-automated way with the seismic-object detection method developed by Meldahl *et al*. (1999). In all three data sets, the results of the neural-network-predicted fault 'probability' cube compare favourably with the single-attribute results in terms of fault continuity and contrast. The optimized fault-cube is able to separate individual faults around the chimney that are not visible in the non-optimized fault-cube.

Comparing the cube similarity for the two fault-cubes (Fig. 5), we see that continuity has increased in general, and the big gas chimney does not show up in the fault-cubes, thus cube similarity could not distinguish between faults and chimneys. Also, the area northwest of the chimney is less noisy and exhibits more continuous faults. Similar improvements are observed along the northern edge of the survey area. The section views in Fig. 4 show similar improvements. The two faults running from top to bottom in the middle of the section show significant improvements in continuity.

Fault-cubes produced by neural networks sometimes exhibit quite broad 'cigar-shaped' faults. This leaves room for improvement in the technology presented here. For example: the process of picking and attribute selection can be repeated with the 'first-generation' result to generate a second-generation fault-cube. It is then possible to use the principle of directionality and calculate the dip and azimuth of the faults from the first-generation fault-cube and to define new attributes that follow the fault direction.

Filtering techniques can be applied to enhance the first-pass fault-cube: for example, the Canny non-maximum suppression ridge detection algorithm (Canny 1986) or the data-adaptive azimuth ridge enhancement filter (de Rooij and

Tingdahl 2002). Current research is focused on developing a 'thinning' algorithm to improve resolution and increase the performance of poor data.

REFERENCES

Bahorich M. and Farmer S. 1995. 3-D seismic discontinuity for faults and stratigraphic features: the coherence cube. *The Leading Edge* **14**, 1053–1058.

Canny J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**, 679–698.

de Groot P.F.M. 1999. Seismic reservoir characterisation using artificial neural networks. 19th Mintrop-Seminar, Münster, Germany.

Heggland R., Meldahl P., Bril A.H. and de Groot P.F.M. 1999. The chimney cube, an example of semi-automated detection of seismic objects by directive attributes and neural networks: Part II; interpretation. 69th SEG meeting, Houston, USA, Expanded Abstracts, 935–937.

Marfurt K.J., Kirlin R.L., Farmer S.L. and Bahorich M.S. 1998. 3-D seismic attributes using a semblance-based coherency algorithm. *Geophysics* **63**, 1150–1165.

Meldahl P., Heggland R., Bril A.H. and de Groot P.F.M. 1999. The chimney cube, an example of semi-automated detection of seismic objects by directive attributes and neural networks: Part I; methodology. 69th SEG meeting, Houston, USA, Expanded Abstracts, 931–934.

Roberts A. 2001. Curvature attributes and their application to 3D interpreted horizons. *First Break* **19**, 85–100.

de Rooij M. and Tingdahl K.M. 2002. Meta-attributes – the key to multivolume, multi-attribute interpretation. *The Leading Edge* **21**, 1050–1053.

Tingdahl K.M. 1999. *Improving seismic detectability using intrinsic directionality*. Paper B194, Earth Sciences Centre, Göteborg University.

Tingdahl K.M., Bril A.H. and de Groot P.F.M. 2001. Improving seismic chimney detection using directional attributes. *Journal of Petroleum Science and Engineering* **29**, 205–211.

Tingdahl K.M. and de Groot P.F.M. 2003. Post-stack dip- and azimuth processing. *Journal of Seismic Exploration* **12**, 113–126.