Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Automatic Fault Detection for 3D Seismic Data

David Gibson[1], Michael Spann[1] and Jonathan Turner[2]

[1]School of Electronic Engineering, University of Birmingham,
Edgbaston, Birmingham, B15 2TT, U.K.
{gibsond, spannm}@eee.bham.ac.uk

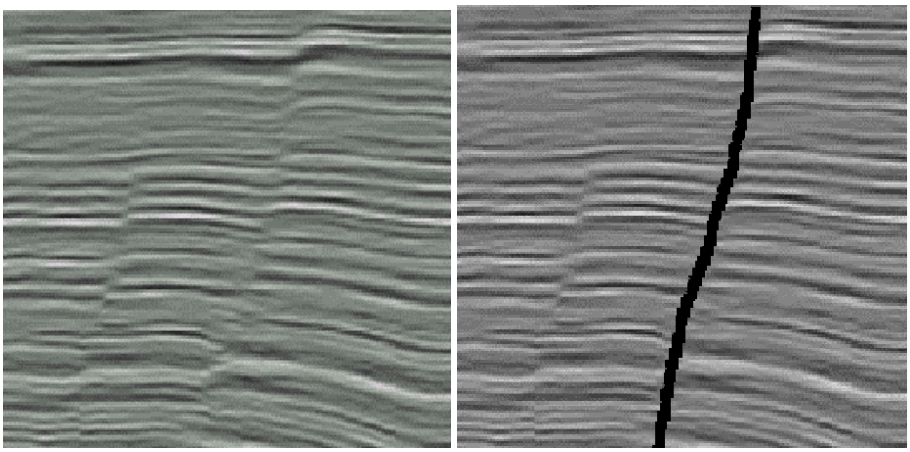
[2]School of Geography and Earth Sciences, University of Birmingham,
Edgbaston, Birmingham, B15 2TT, U.K.
J.P.Turner@bham.ac.uk

Abstract. A novel approach to the automatic detection of fault surface images in 3D seismic datasets is presented. Based on the premise that seismic faulting introduces discontinuities into the rock layering (that is, the horizons), a coherency measure is used to detect points of significant horizon discontinuity. A highest confidence first (HCF) merging strategy is then combined with a flexible surface model to estimate the 3D fault surfaces iteratively.

1. Introduction

Largely fuelled by the petroleum industry, major resources are targeted at the problem of imaging underground rock structures. Using acoustic reflection technology, large 3D datasets can be generated showing the changes in acoustic impedance as a function of depth. Most seismic reflections in a 3D seismic dataset are manifestations of rock layering, or horizons. These represent the subhorizontal sedimentary layering within petroleum-bearing sedimentary basins. A typical vertical 2D slice through a 3D data set is shown in Fig. 1. (a). In this, the depth below ground is shown vertically, with the horizontal axis representing a horizontal position on the ground. The *barcode-like* pattern clearly shows the sedimentary layers of the rocks. To the geologically-trained eye the patterns present in these 3D data sets provide useful insights into the rock formation and geological structure, as well as such things as the location and shape of geological structures in which hydrocarbons are trapped.

Due to the complexity and size of these data sets it is perhaps not surprising that image processing techniques have been brought to bear on a number of 3D seismic data interpretation problems. Most notable of these are *horizon-trackers* [1]. In this context, horizons are surfaces representing individual sedimentary layers of rock. Referring to Fig. 1(a), each black and white stripe represents an horizon (noting of course that the horizon is 3D not 2D). A number of horizon trackers have been successfully developed, and commercial products are available exploiting this technology. Research has now moved to the more difficult problem of seismic *fault-detection*. A fault is caused by the near-vertical, relative movement of adjacent rocks, resulting in the termination of horizons. This effect is shown in Fig. 1(b).



(a) (b)

Fig. 1. (a) 2D slice of 3D seismic data set, (b) Illustration of a seismic fault (highlighted)

Seismic data sets typically contain a large number of faults at many different spatial scales. Knowledge of the location of the faults is critical to understanding a geological system. One effect that faults have, which is of real commercial significance, is that they act as membranes to the movement of hydrocarbons. Therefore having a good understanding of the fault positions is critical for the effective planning of drilling sites in order to maximise output efficiency. However, despite the significant progress in the development of horizon autotrackers, current approaches for ëpickingí faults are largely manual, and involve laborious hand-picking of discontinuities on a slice-by-slice basis, one fault at a time. This is time consuming resulting in hundreds of man-hours of work, performed by trained geologists. It is estimated that for every six months saved in the work leading up to the onset of production from a new oilfield, 5% will be saved from the total production bill. Hence, there is a strong financial imperative for this work. Our proposal is for an automatic strategy for fault surface detection [2,3]. In practice this proves to be an extremely difficult problem to solve due to noise, imaging artefacts, and large numbers of interacting faults at different spatial scales.

2. Proposed Approach

The proposed approach aims to exploit the idea that the presence of seismic faulting will result in discontinuities in the horizons. Detection of the consistency of the horizons constitutes the first step of the process. Following this, small planar patches representing small parts of the fault surfaces are generated. Finally, these small planar patches are merged into larger surfaces, using a highest confidence first (HCF) merging strategy. These steps are described in more detail in the following sections.

2.1 Semblance Matching

Firstly we make the assumption that seismic horizons can be modelled as locally planar, in the absence of seismic faulting. For a given point (x,y,z) on the image, this assumption is then tested. This is done by considering a series of h slices of seismic data, centred around (x,y,z), orientated with a normal **n**. This is illustrated in Fig. 2.

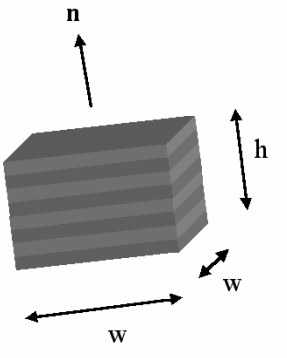


Fig. 2. Illustration of h slices of data, orientated with normal n, centred around point (x,y,z)

The data contained in the w by w array of interpolated data for slice r ($0 \le r \le h$-1) is represented by the vector $\mathbf{v}_r$. A normalised measure of the data variability averaged over the h slices is then computed using:

$$S = \frac{\sum_{r=0}^{h-1} \left(\sum \mathbf{v}_r \right)^2}{h \sum_{r=0}^{h-1} \sum \left(\mathbf{v}_r^2 \right)} \tag{1}$$

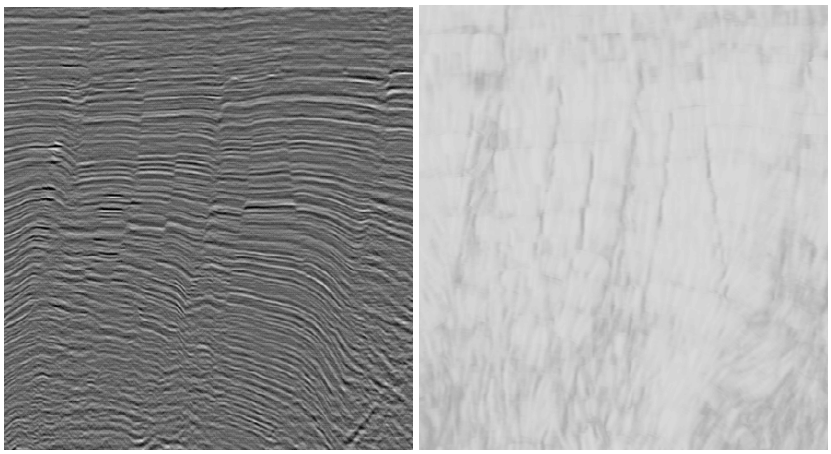
where the normal **n** is the estimated localised normal to the seismic horizons estimated from the structural tensor [5] as follows :

$$T = \begin{pmatrix} u_x^2 & u_x u_y & u_x u_z \\ u_x u_y & u_y^2 & u_y u_z \\ u_x u_z & u_y u_z & u_z^2 \end{pmatrix}$$

$$T_\sigma = T * G_\sigma \tag{2}$$

where u_x, u_y, u_z, are the respective seismic image gradients, G_σ is a Gaussian kernel, ë*í is the convolution operation, and **n** is the eigenvector corresponding to the largest eigenvalue of T_σ.

The net result is a semblance map S which gives a measure of horizon continuity, with $S(x,y,z)=0$ representing the worst case fit, and $S(x,y,z)=1$ indicating the ideal horizon continuity case. This semblance map, however, is only based on local constraints, given by the dimensions of the box used for the estimation ($h*w*w$). As fault surfaces, by definition, are not isolated phenomena, we can use this fact as an addition contextual constraint. In practice this can be achieved by applying a directed spatial filter to the semblance map S, as a post-processing step. The filter kernel is directed to concentrate the filtering action along the direction of the fault surfaces, which in turn is estimated using the same procedure as that used to estimate the normal to the horizons (using equation (2), but this time using the semblance map and not the raw seismic data). The filtered result, S_f is illustrated alongside the original seismic slice in Fig. 3, and shows good correlation between the areas of faultiness on the left image, with darkened areas (lower semblance match) on the right image.



(a) (b)

Fig. 3. (a) 2D slice of seismic data, (b) Equivalent semblance match (darker areas indicate increasing horizon discontinuity)

2.2 Patch Estimation

The semblance data is then thresholded against a pre-defined threshold. This is done on a point-by-point basis, with points having a semblance value below a predefined threshold being labelled as fault points. This 3D binary map of fault points is then spatially sub-sampled to give a set of seed points as shown in Fig. 4.

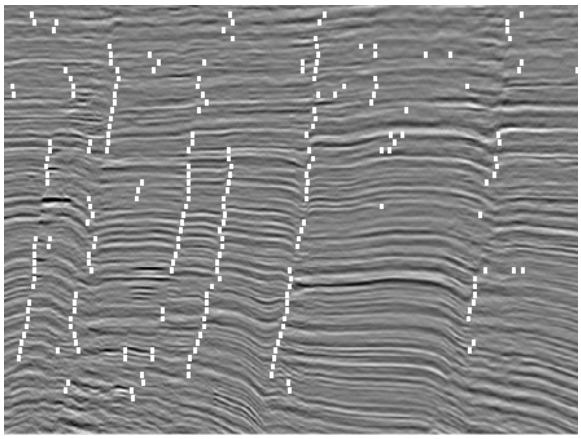


Fig. 4. Seed points overlaid onto a seismic slice

Given a set of seed points, these can be grouped into small planar patches, representing small sections of a fault surface. More formally, for a set of seed points **P**, a neighbourhood relationship is defined, so that two points p_1 and p_2 taken from **P** are neighbours if they fall within a spatial distance d.

$$p_1 \in \mathrm{N}(p_2) \text{ and } p_2 \in \mathrm{N}(p_1) \text{ if } \|p_1 - p_2\| < d \tag{3}$$

where $N(p_x)$ represents the set of points neighbouring p_x.

A proposed planar model centred around each point, $p_x \in \mathbf{P}$ is generated based on p_x and all of the neighbours of p_x. A compatibility function $C(p_x)$ gives an indication of how well a group of points, centred on p_x, would be represented by a planar model, where,

$$C(p_x) = \frac{\sum_{p_y \in N(p_x)} C(p_x, p_y)}{\# N(p_x)} \tag{4}$$

This results in the mean compatibility value between p_x and all of its neighbours, where $C(p_x, p_y)$ defines the compatibility of a pair of points p_x and p_y and can be calculated as:

$$C(p_x, p_y) = \left(1 - |\mathbf{n}_x.\mathbf{r}| - |\mathbf{n}_y.\mathbf{r}|\right)\left(|\mathbf{n}_x.\mathbf{n}_y|\right) \tag{5}$$

where $\mathbf{n}_x$ and $\mathbf{n}_y$ are the estimated normal unit-vectors to the points p_x and p_y respectively (as calculated earlier), and $\mathbf{r}$ is the unit position vector between p_x and p_y. $C(p_x, p_y)$ tends towards unity when the estimated surface normals to the points p_x, and p_y line up, and are perpendicular to the position vector between p_x, and p_y.

Starting with the planar model with the corresponding highest confidence value, the planar model is constructed, and the points removed from **P**. This is repeated until the confidence value of all remaining clusters is below a pre-specified threshold.

2.3 Surface Merging

Although a planar model can well describe a small patch of a much larger fault surface, it is inadequate for describing larger segments, or complete faults. To describe larger fault surfaces we make use of a combined parametric and residual field [4], the parametric model (in our case planar) models the basic structure of the fault surface, with the residual field modelling the surface irregularities and curvature.

More formally, a planar model is described by a set of vectors $\{\Omega_x, \Omega_y, \Omega_n, \Omega_o\}$, where Ω_x, Ω_y and Ω_n, are mutually perpendicular vectors, with Ω_n representing the normal to the plane. A known point on the plane is denoted Ω_o, and represents the approximate centre of the subsequent fault surface model. This is illustrated in Fig. 5(a).

In combination with this planar model is a residual field, constructed using a 2D mesh of vectors, normal to the plane and of varying lengths (see Fig. 5(b)).

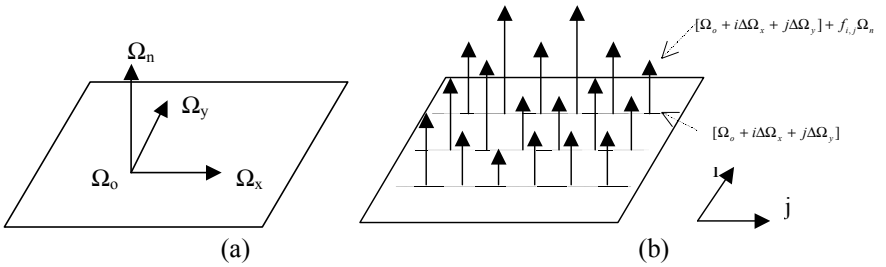


Fig. 5. (a) Illustration of the planar model, (b) Planar model plus residual mesh field

The resultant surface is then formed by interpolating this mesh, an illustrated in Fig. 6.

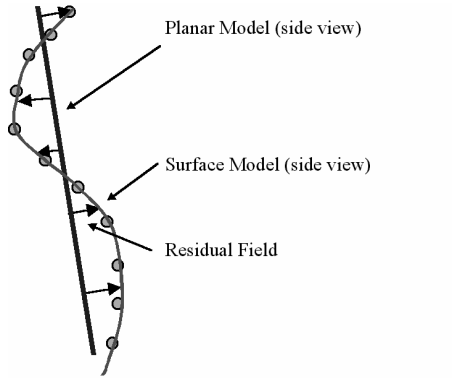


Fig. 6. Illustration of the surface formation as an interpolation of the residual field overlaid onto the planar model

Given a set a points, **P**, a surface model can be generated by firstly computing the planar model based on a simple least squares fit of the points in **P**. The distance of each point $p_x \in \mathbf{P}$ from the planar surface is then computed, and is denoted ε_x. The residual field can be described as a mesh of vectors **M**, as shown by the arrows in Fig. 5(b). The end position of the $(i, j)^{th}$ vector in **M** will then be of the form,

$$\mathbf{m}_{i,j} = [\Omega_o + i\Delta\Omega_x + j\Delta\Omega_y] + f_{i,j}\Omega_n \tag{6}$$

where $f_{i,j}$ represents the length of the vector $\mathbf{m}_{i,j}$, $[\Omega_o + i\Delta\Omega_x + j\Delta\Omega_y]$ is the start position, on the plane, of the vector corresponding to the index (i, j), and Δ is a constant which specifies the residual field mesh spacing. The values of $f_{i,j}$ for all values of i and j can then be calculated as a weighted average,

$$f_{i,j} = \frac{\sum_{p_x \in P} \varepsilon_x w(p_x)}{\sum_{p_x \in P} w(p_x)} \tag{7}$$

where $w(p_x)$ is a weighting function which diminishes with distance,

$$w(p_x) = \exp(-\omega^2 / \sigma^2) \tag{8}$$

for which ω is the distance between the point p_x and the point $[\Omega_o + i\Delta\Omega_x + j\Delta\Omega_y]$. The net result is a completely defined fault surface made up of a combination of the planar and residual models.

Given a surface model definition, a merging strategy based on the HCF principle is adopted, in much the same manner as in the last section. Firstly, all of the planar patches from the previous section are *promoted* into surfaces. All possible pairs of surfaces are considered for merging, with a compatibility function for two surfaces s_1 and s_2 defined as,

$$C(s_1, s_2) = \frac{K + mean(|\varepsilon(s_1)|) + mean(|\varepsilon(s_2)|)}{mean(|\varepsilon(s_1 \,\&\, s_2)|)} \tag{9}$$

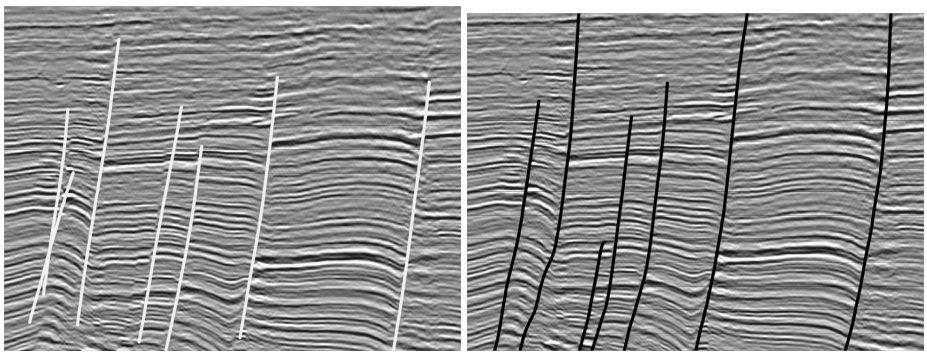
where $mean(|\varepsilon(s_x)|)$ is the mean residual field magnitude value, $(s_1 \,\&\, s_2)$ represents the combination of surfaces, and K is a constant which encourages small surfaces to merge.

The merging of surfaces is continued iteratively until the confidence value falls below a predefined value.

3. Results

Natural fault systems are present at a wide range of spatial scales, from faults that cover the entire data set under consideration, to faults that are smaller than the spatial resolution of the data. Our aim is to estimate the set of large faults that make up the general structure of the fault field. In order to evaluate the performance of the method a manually labelled set of fault surfaces have been identified by a geologist. An example set or results showing vertical 2D slices of the seismic data are presented

providing a comparison between the automatically generated results produced using the proposed method, with manually labelled images. These results are shown in Fig. 7, with the automatically generated and manually generated results overlaid in white and black respectively. It should be noted, of course, that the lines shown on the figures are the result of slicing the surfaces. In effect, the surfaces can be imagined coming out of, and going into the page.



(a) (b)

Fig. 7. (a) Estimated faults, (b) Equivalent, manually labelled faults

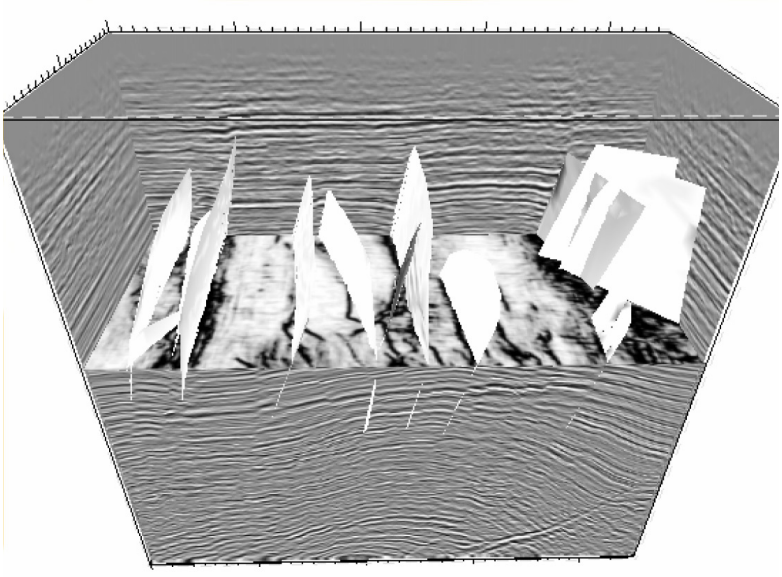
As can be seen from the comparative results there is a broad agreement between the placement of many of the faults from the automatic detection and manually labelled methods. A 3D view of the detected surfaces along with the equivalent manually labelled faults is shown in Fig. 8.

The proposed method performs well at detecting the larger faults, although it misses some of the smaller, less well-defined faults. Problems can also occur if two faults are spatially very close together. This can result in over-merging of faults, the net result being one surface poorly representing two faults.

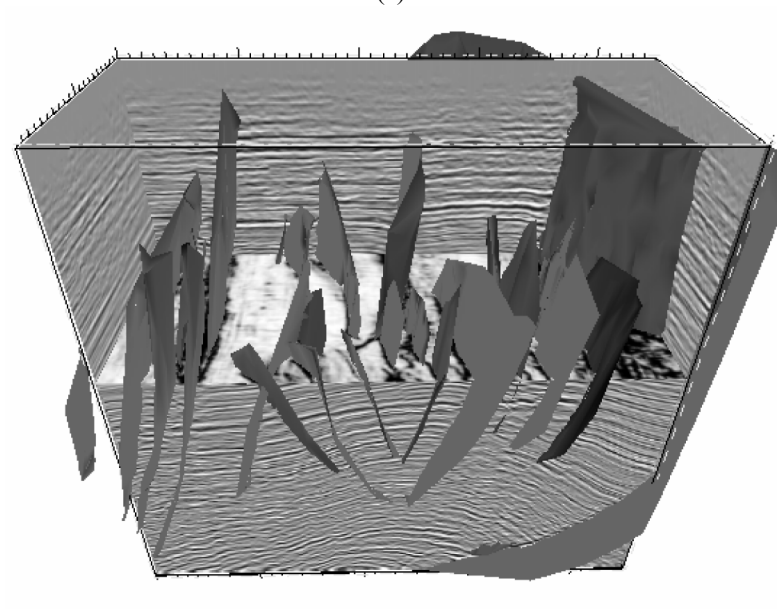
4. Conclusions and Future Work

A method has been presented to tackle the difficult and resource consuming task of fault detection in 3D seismic datasets. Based on a multi-stage approach, it first detects points of horizon discontinuity, and progressively groups these points into larger surfaces. The final surface representation is a combined parametric and residual field model, which allows for a highly flexible surface representation. Comparative results with manually labelled faults show promising results.

One of the key questions not addressed in this work is that of combining the automatic fault detection approach with some element of human input, to give a semi-automatic fault detection tool. We believe that such an ability to edit, refine, direct, or impose prior constraints on the problem could provide considerable productivity gains over and above a manual approach, whilst allowing the flexibility required for complex data sets with low signal to noise ratio and multiple interacting faulting.



(a)



(b)

Fig. 8. (a) Estimated faults, (b) Equivalent, manually labelled faults

References

1. Aurnhammer, M., Tˆnnies, K.D., Mayoral, R.: A Genetic Algorithm for Constrained Seismic Horizon Correlation. Proceedings of the International Conference on Computer Vision Pattern Recognition and Image Processing (CVPRIP 2002), pp. 859-862, Durham, North Carolina USA, 8-14 March, 2002.
2. Randen, T., Pedersen, S.I., Signer, C. , S□nneland, L.: Image Processing Tools for Geologic Unconformity Extraction. IEEE SYMPOSIUM, Sem GjestegÂrd, Asker, Norge, 9-11, September, 1999.
3. Tingdahl, K.T., Steen, ÿ ., Meldahl, P., Herald, J.: Semi-automatic detection of faults in 3-D seismic signals. SEG/San Antonio 2001.
4. Black, M.J., Jepson, A.: Estimating Optical Flow in Segmented Images using Variable-Order Parametric Models with Local Deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 10, Oct. 1996, pp. 972-986.
5. Fehmers, G., Hocker, F.W.: Fast Structural Interpretation with Structure-orientated Filtering. Geophysics, Vol. 68, No. 4, August, 2003, pp. 1286-1293.