ELSEVIER

Journal of Structural Geology 28 (2006) 980–990

JOURNAL OF STRUCTURAL GEOLOGY

www.elsevier.com/locate/jsg

Paleostress analysis using the Hough transform for separating stresses from heterogeneous fault-slip data

Atsushi Yamaji *, Makoto Otsubo, Katsushi Sato

Division of Earth and Planetary Sciences, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

Abstract

This paper shows that the Hough transform, a basic technique of image processing, is useful for separating stresses from heterogeneous fault-slip data. The present method enumerates admissible stresses by evaluating their probability for a heterogeneous dataset. Using the results of the method applied to artificial data with known responsible stresses, it is argued that the criterion that has been generally used to judge the resolution and accuracy of a numerical technique for the separation is not appropriate. The result reveals the limitation of stress inversion based on the Wallace–Bott hypothesis. No matter how good an inverse method is, it sometimes inevitably yields unexpected stresses even from artificial data if the data have heterogeneity.

Keywords: Stress space; Hypersphere; Pattern matching; Stress tensor inversion; Fault-slip analysis

1. Introduction

Fault-slip data are called heterogeneous if the faults within the dataset were activated by multiple stress conditions. Given homogeneous data, it is straightforward to determine the responsible stress using mathematical inversion (Carey and Brunier, 1974; Angelier, 1979). Therefore, faults are often classified at outcrops by their apparent relative ages in order to divide the heterogeneous data into homogeneous subsets (e.g. Angelier et al., 1985; Bergerat, 1987; Choi et al., 2001; Vandycke and Bergerat, 2001). However, such fault classification is not always possible.

Accordingly, several researchers have attempted to develop computer-based methods for separating stresses from heterogeneous fault-slip data (Etchecopar et al., 1981; Armijo et al., 1982; Huang, 1988; Galindo-Zaldívar and González-Lodeiro, 1988; Angelier, 1994). Since the mid-1990s, numerical techniques developed in information engineering have been applied to this problem. Nemcok and Lisle (1995) adopted hierarchical clustering (e.g. Duda et al., 2001) to divide fault-slip data into homogeneous subsets. Shan et al. (2003, 2004) employed non-hierarchical clustering for fault sorting based on the σ-space proposed by Fry (1999).

This paper has two purposes. The first purpose is to present a new method to separate stresses from heterogeneous fault-slip data. The method utilizes the Hough transform, a technique of image processing (e.g. Leavers, 1992). The technique is briefly introduced in Appendix A.

The second aim of this article is more important. The methodological studies on this problem often use artificial data sets that are generated with known stresses for testing methods. Other testing datasets were prepared by combining natural datasets, each of which resulted from a single tectonic event (e.g. Liesa and Lisle, 2004). A good numerical technique is expected to detect only the assumed stresses. If unexpected stresses are obtained, they are regarded as spurious solutions and the employed method is judged unsuccessful.

Based on the geometrical interpretation of stress tensor inversion by Fry (2001) and Sato and Yamaji (2006a), it is shown in this paper that it is often inevitable that unexpected solutions are obtained among expected ones from artificial heterogeneous data. We argue that the criterion that has been used to judge the performance of a numerical technique in separating stresses from heterogeneous data is not appropriate. This is the second aim of this paper.

2. Method

The method that we propose in this article is an extension of the Ginkgo method (Yamaji, 2003). The improvement has been enabled by a geometrical formulation of the stress tensor inversion problem (Sato and Yamaji, 2006a) and by

* Corresponding author. Tel.: +81 75 753 4150; fax: +81 75 753 4189.
E-mail address: yamaji@kueps.kyoto-u.ac.jp (A. Yamaji).

0191-8141/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsg.2006.03.016

the derivation of a computational grid with uniform spacing (Sato and Yamaji, 2006b). Our method, which is an adaptation of the Hough transform to paleostress analysis, visualizes the object function of the inverse method by Angelier (1979).

2.1. Stress inversion and the Hough transform

The fitness of a stress to a fault-slip datum is evaluated in this article using d, the angular misfit between the observed and theoretical slip directions of the fault. Given the orientation of a fault surface and a reduced stress tensor, σ, the theoretical slip direction on the surface by the stress can be calculated with the Wallace–Bott hypothesis (Wallace, 1951; Bott, 1959). Namely, the slip direction of the fault, the surface of which is perpendicular to the unit vector $\vec{n}$, is given by $(\vec{n}\cdot\sigma\vec{n})\vec{n}-\sigma\vec{n}$. Therefore, d is a function of the orientation and the tensor. It is highly improbable that the stress tensor was responsible for the movement of a fault with a large d. Such a fault is said to be incompatible with the stress. We also use $\rho(d)$, a symmetric function, i.e. $\rho(-d)=\rho(d)$, with a unique maximum at $d=0$. Choi et al. (1996) called $\rho(d)$ a fitness criterion. The optimal stress for a dataset is determined by maximizing the function

$$M(\sigma)=\frac{1}{N}\sum_{k=1}^{N}\rho(d(k,\sigma)) \tag{1}$$

where N is the number of data, $d(k,\sigma)$ is the misfit angle of the kth fault for the trial tensor σ.

The specific form of $\rho(d)$ is arbitrary. The ordinary least-squares regression uses $\rho(d)\propto d^2$, which implicitly expects the residuals $d(1,\hat{\sigma}),\ldots,d(N,\hat{\sigma})$ for the optimal stress $\hat{\sigma}$ to obey Gaussian statistics. Such a robust statistical approach, which was introduced by the discussion between Fry (1992) and Will and Powell (1991, 1992), is suitable for obtaining the best-fit tensor for non-Gaussian data, because the technique suppresses the contamination of the optimal solution by outliers. However, the optimal solution is often spurious for heterogeneous data (Fig. 10a). The purpose of this study is not to determine the single tensor but to separate multiple tensors from a heterogeneous dataset.

$M(\sigma)$ in Eq. (1) is called an M-estimator (Rousseeuw and Leroy, 1986) and can make the inversion numerically robust for non-Gaussian residuals by choosing an appropriate function $\rho(d)$. The following functions have been used for stress inversion (Angelier, 1977, 1979, 1984):

$$(\cos|d|+1)/2 \tag{2}$$

$$1-\sin(|d|/2) \tag{3}$$

$$\begin{cases} 1-\tan^2|d| & \text{if } 0\le|d|\le 45^\circ \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\begin{cases} 1-\tan|d| & \text{if } 0\le|d|\le 45^\circ \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Fig. 1a shows the graphs of these functions. The choice of a form does not have a significant effect upon the optimal stress if homogeneous data are processed. Heterogeneity of data vitiates the solutions that are obtained with Eqs. (2) or (3) because the graphs of those functions have long tails. The summation in Eq. (1) involves fault-slip data that are incompatible with a test tensor σ. In contrast, the tails of Eqs. (4) and (5) are cut off at $d=45^\circ$ not to spoil the optimal stress by suppressing the contribution of the incompatible data to the summation. The stress inversion with such an elaborate function is robust to the heterogeneity and can determine the most significant stress for the heterogeneous data (Angelier, 1984). We evaluate the fitness of σ to a fault-slip datum with d and $\rho(d)$. For simplification of the following arguments, we use the function

$$\rho(d;\,T)=\begin{cases} 1-|d|/T & \text{if } 0\le|d|\le T \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where the semicolon in the left-hand side of Eq. (6) separates the different kinds of arguments: d depends entirely on tectonics, but T is an index of the method. Comparing Fig. 1a and b, it is obvious that $\rho(d;\,45^\circ)$ is an approximation of Eqs. (4) and (5). Likewise, $\rho(d;\,180^\circ)$ approximates Eqs. (2) and (3).

In order to see what happens by combining Eqs. (1) and (6), let us study the relationship between the Hough transform and stress inversion. Suppose that we have N data points on the xy-plane and that the pair $(x^{(k)},\,y^{(k)})$ indicates the coordinates of the kth point. Then, the Hough transform of the points is given by the equation

$$H(a,\,b)=\frac{1}{N}\sum_{k=1}^{N}\delta\left(ax^{(k)}+b-y^{(k)}\right) \tag{7}$$

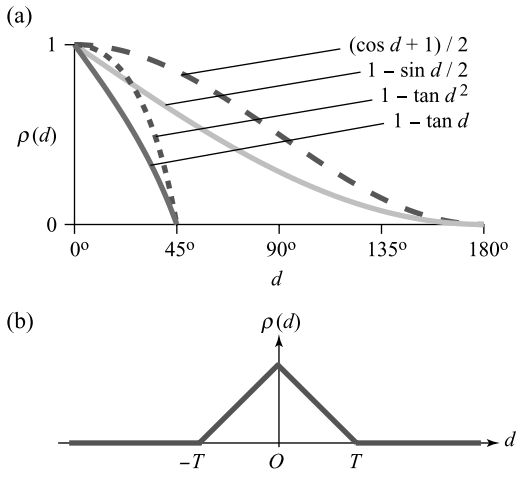

Fig. 1. (a) Examples of fitness criterion, $\rho(d)$. (b) Fitness criterion used in this article.

where

$$\delta(t) = \begin{cases} 1 & \text{if } t = 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

is a function of an arbitrary scalar variable t. The argument of this δ-function in Eq. (7), $ax^{(k)}+b-y^{(k)}$, comes from Eq. (A.1). Consequently, $H(a, b)$ has non-zero values only along lines on the ab-plane that correspond to the data points $(x^{(k)}, y^{(k)})$ (Fig. 10b).

Note that the peak of the function ρ in Fig. 1b narrows with decreasing $|T|$, but its height is kept constant. The relationship between the Hough transform and stress inversion turns out from the equivalence

$$\lim_{T\to 0} \rho(t;\ T) = \delta(t)$$

where ρ is the fitness criterion defined by Eq. (6). This equation allows us to rewrite Eq. (7) as $\lim_{T\to 0} H(a, b; T)$, where

$$H(a,\ b;\ T) = \frac{1}{N}\sum_{k=1}^{N} \rho\left(ax^{(k)} + b - y^{(k)};\ T\right)$$

When the Hough transform is employed for detecting lines, the results become numerically stable and robust to noise by casting votes not only to accumulation cells immediately under the lines but also to cells around the lines in the ab-plane (Thrift and Dunn, 1983). Such casting is indicated by $H(a, b; T)$ with $T>0$. As a result, a zone with a half width of T receives votes. Fig. 10c shows the 3D surface plot of $H(a, b; 0.1)$ for the data points in Fig. 10a. The surface has the prominent peaks corresponding to the two lines along which the data points are aligned. Increasing T amalgamates small peaks and ridges of the surface. At the same time, the two prominent peaks spread their bases. The peaks are eventually merged. Fig. 10d shows the surface $H(a, b; 1)$. This surface has a single peak near the point that corresponds to the line A. As T becomes larger, $\rho(d; T)$ approaches the L_1-estimator, meaning that the inversion maximizes the sum of the absolute values of residuals (Rousseeuw and Leroy, 1986).

The above arguments demonstrate that $M(\sigma)$ is comparable with the Hough transform of fault-slip data. $M(\sigma)$ has a peak every time σ has one or more compatible fault-slip data. Namely, the match of *each* fault-slip datum to various reduced tensors is represented by a spike in $M(\sigma)$, although $M(\sigma)$ is defined by the summation of ρ for the *entire* fault-slip data. Accordingly, significant stresses are separable as the prominent peaks of $M(\sigma)$. If a homogeneous dataset is given, $M(\sigma)$ has the only maximum at the stress responsible for the faults from which the data are obtained. We choose a value of $T=45°$ for the fitness criterion to approximate Eq. (5) to find out the reduced stress tensors that simultaneously match *several* fault-slip data.

2.2. Detection of peaks upon the five-dimensional hypersphere

Points upon the unit hypersphere in the five-dimensional space have one-to-one correspondence with reduced stress tensors and the Euclidean distance between points on the hypersphere equals Orife and Lisle's (2003) stress difference between the tensors corresponding to the points (Sato and Yamaji, 2006a). Accordingly, we regard $M(\sigma)$ in the same light as $M(\vec{x})$, where $\vec{x}$ is a unit vector with its initial point at the origin of the five-dimensional space (Appendix B). Because of the inequality $M(\sigma)\geq 0$, the values of the function

$$F(\vec{x}) = M(\vec{x}) + 1 \tag{9}$$

are always greater than or equal to 1, and represent a surface out of or tangent to the hypersphere. The peaks of this surface depicts the extremes of $M(\sigma)$, which further correspond to the responsible stresses for heterogeneous fault-slip data.

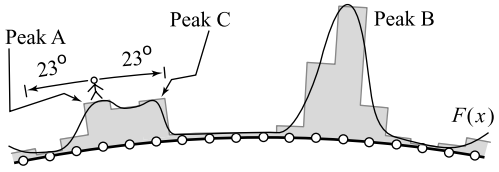

Fig. 2. Schematic illustration showing the topography of $F(\vec{x})$ (thin line) and the procedure to detect its peaks. Open circles are computational grid points with uniform intervals on the five-dimensional unit sphere produced by Sato and Yamaji (2006b). Stress tensors with maximal $F(\vec{x})$ are represented by the peaks of the topography, which can be detected by climbing the steps that are quantized from the topography. The grid points are also used as the initial positions of the 'climbers,' which have the field of view of a radius of 23°.

Fig. 2 illustrates the procedure to detect the peaks of the surface $F(\vec{x})$ in the five-dimensional space. Sato and Yamaji (2006b) distribute 60,000 points with uniform intervals on the unit hypersphere to furnish computational grid points to stress tensor inversion. First, we evaluate the function $F(\vec{x})$ at the points. Second, the F value of a grid point is compared with the values at its neighbouring grid points. If the value is the greatest among those of its neighbours, a peak is identified at the grid point. In this study, the neighbour is defined as the grid points within a Euclidean distance of 0.4 in the parameter space, which corresponds to a stress difference (Orife and Lisle, 2003) of the same value (Sato and Yamaji, 2006a). The angular distance between two points on the hypersphere is calculated from the Euclidean distance between the points. The chord with the length of 0.4 subtends the arc with a length of $2\tan^{-1}(0.4/2)\approx 23°$. If two peaks exist within this angular distance, the lower peak is ignored. For example, the distance between peaks A and B in Fig. 2 is greater than this threshold distance. So, our procedure can recognize both peaks A and B, but peak C is not separated from peak A.

3. Test

3.1. Artificial data

For testing of the present method, a set of heterogeneous fault-slip data was generated with two stresses of strike-slip faulting regimes. The set contains 60 faults with various orientations. Forty faults were thought of as those activated by Stress A, which has σ_1 and σ_3 orientations of 035°/20° and 125°/00°, respectively (Fig. 3a). Twenty faults were assumed to

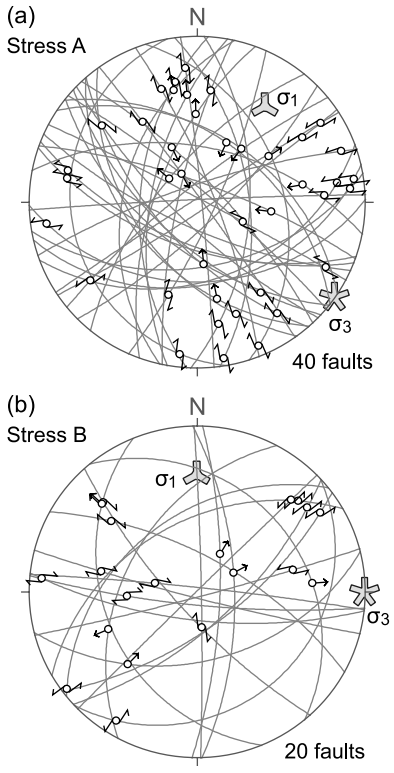

Fig. 3. Artificial data for testing. Faults activated by Stress A (a) and Stress B (b) were mixed to form a heterogeneous data set. Both the stresses have a stress ratio of 0.5. Lower-hemisphere, equal-angle projection.

be activated by Stress B, where σ_1 and σ_3 orientations were 000°/20° and 090°/00°, respectively (Fig. 3b). Both the stresses had the stress ratio $\Phi \equiv (\sigma_2-\sigma_3)/(\sigma_1-\sigma_3)=0.5$, where we use the sign convention that compression is positive. The angular distance between the σ_1 orientations is 33° and the stress difference (Orife and Lisle, 2003) between the stresses is 1.10. A heterogeneous data set was made by the concatenation of those fault-slip data.

3.2. Result

The present method was applied to the heterogeneous data set. Five peaks were detected and the corresponding solutions were labelled as Stresses I–V (Fig. 4). Parameters of the stresses are listed in Table 1. Both Stresses A and B were successfully detected as the first and second highest peaks, which correspond to Stresses I and II. The principal orientations of these stresses were rotated from the assumed ones by 1–4°, because our computational grid points have intervals of a few degrees. The stress ratios of I and II were approximately equal to the assumed value of 0.5.

However, three unexpected solutions (Stresses III–V) were also obtained. Fault sorting by the present method was good but imperfect. The column headed m in Table 1 indicates the number of faults with $d<20°$. If a fault has a small d, the fault is said to match the stress. This is a key to fault sorting. Several values in the range from 10 to 30° have been used for the misfit threshold (e.g. Etchecopar et al., 1981; Casas et al., 1990; Nemcok and Lisle, 1995; Ramsay and Lisle, 2000). We use 20° here but the choice of the value does not affect the conclusions of this paper, if it is in the range.

Fig. 5 shows the fault sorting based on Stresses I and II. The faults activated by Stress A were completely classified to be so, but three of those by Stress B were mixed in this class. The faults activated by Stress B were correctly related to Stress II except for one fault, because the fault was nearly parallel to the assumed σ_1-axis. This reflects the instability of d for the faults that are nearly parallel to principal stress planes (Gephart, 1990). Among the faults that should have been related to Stress I, six faults were compatible with Stress II.

3.3. Comparison with other methods

The present technique is an extension of the Ginkgo method (Yamaji, 2003). The difference is in the techniques for identifying the peaks of the object function $M(\sigma)$. The Ginkgo

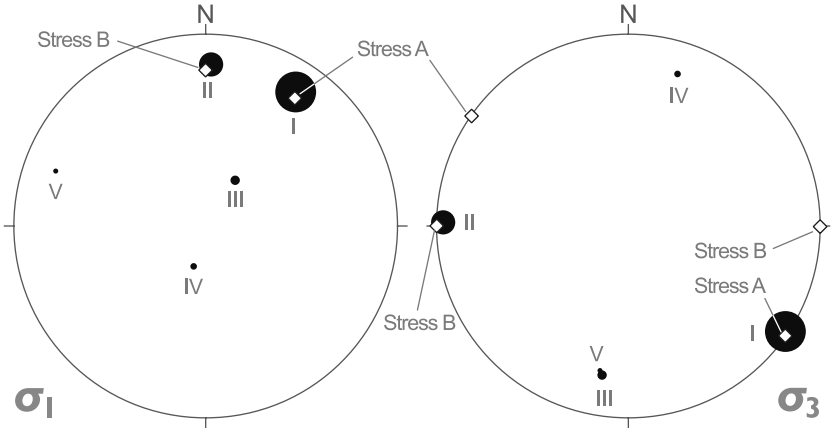

Fig. 4. Paired stereonets showing the result of the present method applied to the artificial fault-slip data (Fig. 3). The left and right stereonets show the σ_1 and σ_3 orientations, respectively. Closed circles are plotted at the principal orientations of Stresses I–V, which correspond to the peaks of $F(\vec{x})$. Diameter of the circles is proportional to the height of the peaks. Diamonds indicate the principal orientations used for production of the test data. Stresses III–V are unexpected solutions. Lower-hemisphere, equal-angle projection.

Table 1
Five stresses detected by the present method from the heterogeneous fault-slip data shown in Fig. 3. The second and third columns in this table show the azimuth/plunge of σ_1 and σ_3 orientations, respectively. Further explanation in text

	σ_1	σ_3	Φ	$F(\vec{x})$	m
Stress I	034/17	124/01	0.52	1.675	43
Stress II	002/17	271/04	0.50	1.400	25
Stress III	033/67	190/22	0.72	1.155	10
Stress IV	196/72	018/18	0.62	1.109	8
Stress V	290/18	191/24	0.62	1.083	6

method visualizes the peaks on paired stereograms. The resolution of the new technique is decisively better than the original method. Specifically, it was difficult for the Ginkgo method to detect Stress B, the second significant stress.

The Ginkgo method has a mechanism, similar to the right-dihedra method (Angelier and Mechler, 1977). The latter can discriminate a few stresses on a stereonet, if their principal orientations are clearly separated. The Ginkgo method also uses stereonets on which the principal orientations and stress ratios are distinguished by eye. However, the reduced stress tensors to be discriminated are entities in the five-dimensional parameter space. The method projects them onto a stereonet on a 2D plane for visualization. The shadows of the projection are superimposed many times so that the discrimination of significant stresses on the stereonet is difficult (Lisle and Vandycke, 1996).

In contrast, the peaks of $F(\vec{x})$ are enumerated by the present method within the five-dimensional parameter space. Therefore, the peaks and corresponding stresses can be clearly separated.

On the other hand, the multiple inverse (MI) method (version 5) (Yamaji, 2000) separated both the assumed stresses, though Stress B was not as evident as the result of the present technique. Significant stresses detected from the data by the MI method are indicated by clusters on paired stereograms in Fig. 6a. Each member in the clusters represents a reduced stress tensor. Therefore, the tensors were classified into two groups on the five-dimensional unit sphere using the k-means

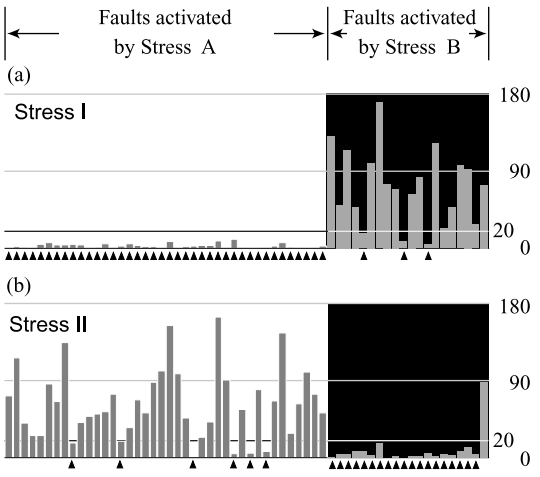

Fig. 5. Misfit angles (d) of the 60 faults in Fig. 3 for Stresses I and II. Triangles indicates the faults with small angular misfits ($d<20°$).

clustering by Otsubo et al. (2006) for determining the averages of the clusters. Fig. 6b shows the result. The representative stresses determined through this technique for the two clusters were approximately equal to the assumed ones. The principal orientations were rotated only by a few degrees. Both the representative ones had a stress ratio of 0.41.

Although the comparative study of the present and MI methods is not comprehensive at the moment, we think that both the methods have shown similar performance in separating stresses from heterogeneous data sets. However, the time of computation is different. The computational complexity of the MI method inflates as N^5 or so, but that of the present method is proportional to N. Therefore, the present method usually spends a much shorter time than the MI method in computation.

4. Unexpected solutions

The tangent-lineation diagrams in Fig. 7 illustrate why the unexpected solutions appeared. They show the match or mismatch of each fault-slip datum to the detected stresses. White thin arrows in the diagrams indicate the theoretical slip directions of the footwall blocks, the fault planes of which are perpendicular to the orientations designated by the positions of the arrows on the stereonets (Twiss and Moores, 1992). The patterns made by the arrows have one-to-one correspondence with reduced stress tensors. Therefore, the separation of stresses from heterogeneous fault-slip data involves finding the patterns that match data. Therefore, the present technique is an application of pattern matching. The fact is that detection of objects using the Hough transform is equivalent to template or pattern matching technique (Stockman and Agrawala, 1977). Consequently, several thick arrows match thin white arrows by chance in Fig. 7c–e. Unexpected stresses reflect such coincidences.

Those solutions are comparable with the accumulation cells that got fewer votes than the significant two peaks in Fig. 10c. We judged that those cells corresponded to spurious alignments because we knew the correct solutions. Can we judge the unexpected stresses spurious?

Stresses III–V were regarded as unexpected solutions because they were neither of the assumed stresses. However, we do not know correct ones when we deal with natural data. Accordingly, there would be no reason to reject these three stresses, when they are obtained from natural heterogeneous data. The Wallace–Bott hypothesis places loose constraints on this problem.

These observations encourage us to abandon the criterion that has been often employed to judge the resolution and accuracy of a method in separating stresses from heterogeneous fault-slip data. Namely, methods are judged unsuccessful if they yield unexpected solutions. Instead, all we can do is to enumerate stresses matching or compatible with parts of the data. Good methods for this problem should do this. The present method fills this demand.

The present method satisfactorily detected the two tested stresses as prominent peaks. However, three unexpected stresses also appeared. In addition, fault sorting based on

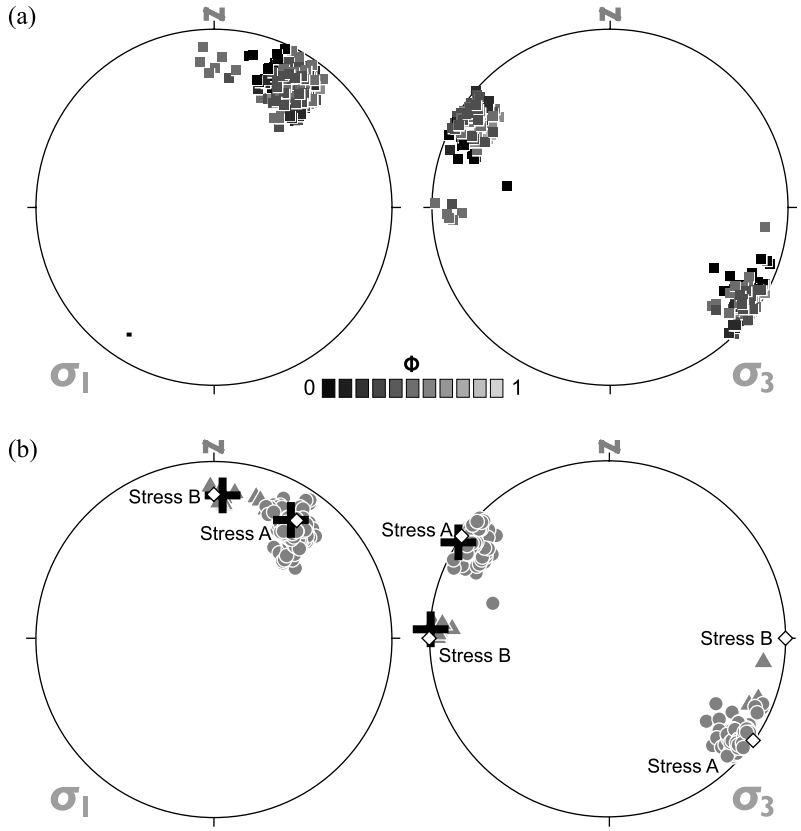

Fig. 6. (a) Result of the multiple inverse method applied to the artificial data in Fig. 3. Fault combination number (k) and enhance factor (e) are 5 and 6, respectively. See Yamaji (2000) for details of these parameters. Each square plotted on the stereograms represent a reduced stress tensor, the stress ratio of which is indicated by a grey scale. Lower-hemisphere, equal-area projection. (b) Result of automatic clustering of the reduced stress tensors indicated in (a) using the technique of Otsubo et al. (2006). Crosses indicate the representative principal orientations of the clusters. Diamonds depict the principal orientations of the assumed stresses for the data.

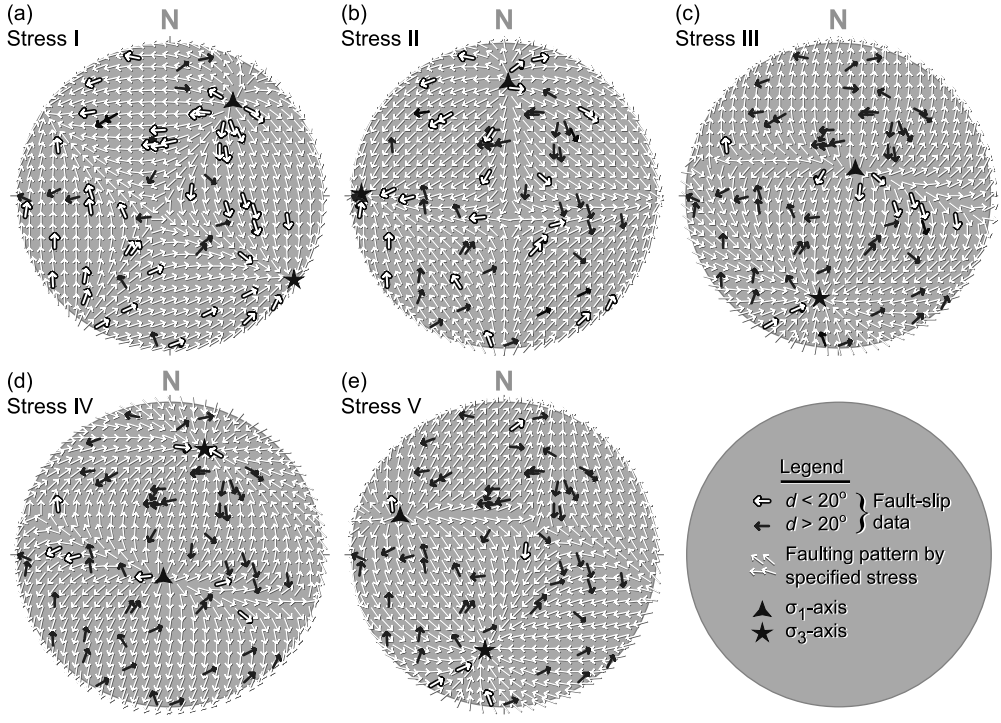

Fig. 7. Tangent-lineation diagrams showing the matches of the artificial fault-slip data (Fig. 3) to the faulting patterns (white thin arrows) produced by Stresses I–V. Thick arrows indicate the data. The data with $d<20^\circ$ are denoted by thick white arrows edged with black. Those with $d>20^\circ$ are shown by thick solid arrows.

the detected stresses was not perfect. In the next section, the limitation of stress tensor inversion is discussed using me geometrical formulation by Sato and Yamaji (2006a).

5. Heterogeneous data in the parameter space

In this section, we consider how heterogeneous fault-slip data are expressed in the high dimensional parameter space and show the inevitability of obtaining 'unexpected' solutions from those data sets. The following argument is entirely based on the theory by Sato and Yamaji (2006a), who embed Orife and Lisle's (2003) stress difference as a metric in the parameter space proposed by Fry (2001). Data points on the physical space (*xy*-plane) correspond to lines in the parameter space (*ab*-plane) in the detection of lines (Appendix A). Then, what kind of figure in the five-dimensional parameter space does correspond to a fault-slip datum?

Points on the unit hypersphere can represent both fault-slip data and reduced stress tensors (Sato and Yamaji, 2006a). A fault-slip datum is indicated by the paired unit vectors $\vec{\varepsilon}$ and $\vec{\varepsilon}'$ in the five-dimensional parameter space (Appendix B). The points P, Q, R and S in Fig. 8 are on the hypersphere and the rays OR and OS are parallel to $\vec{\varepsilon}$ and $\vec{\varepsilon}'$, respectively. The reduced stress tensors that are indicated by the points along the great circle arc PRQ in the figure cause slips with a null deviation angle ($d=0$) on the fault plane that has the paired unit vectors. An analytical equation for the tensors was derived by McKenzie (1969). Sato and Yamaji (2006a) demonstrate that a stress tensor indicated by a vector $\vec{x}$ results in the misfit angle

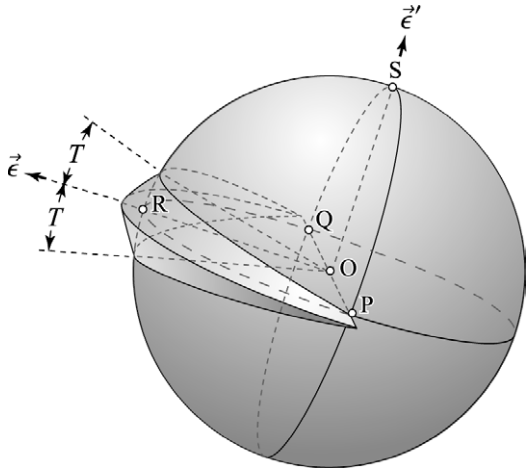

Fig. 8. Schematic illustration showing the unit hypersphere centred at point O in the five-dimensional parameter space. The rays OR and OS are parallel to the unit vectors $\vec{\varepsilon}$ and $\vec{\varepsilon}'$, respectively, which are perpendicular to each other. The admissible reduced stress tensors for the datum are represented by the points along the great circle perpendicular to $\vec{\varepsilon}'$ and on the hyper-hemisphere containing point R. The line PQ is defined to be perpendicular to the vectors. There are three orientations perpendicular to the vectors in this five-dimensional space. However, they are represented by the single line PQ for illustration purposes. The ridge along the great circle arc PRQ depicts the topography of $F(\vec{x})$ for the case of the function $p(d)$ in Eq. (6). The ridge has the maximum width of $2T$ at point R, but narrows toward its ends at P and Q. The height of the ridge is constant along the great circle arc PRQ.

$$d = \tan^{-1}\frac{|\vec{\varepsilon}'\cdot\vec{x}|}{\vec{\varepsilon}\cdot\vec{x}}$$

where $0 \le d \le 180°$. Namely, d equals the angle between $\vec{\varepsilon}$ and the vector that is the shadow of the orthogonal projection of $\vec{x}$ onto the plane spanned by the paired vectors.

Combining Eqs. (1) and (9), we obtain

$$F(\vec{x}) = \frac{1}{N}\sum_{k=1}^{N}\rho(d(k,\vec{x})) + 1 \tag{10}$$

This is the superposition of the undulations corresponding to fault-slip data. The topography for a single datum is represented by

$$F(\vec{x}) = \frac{\rho(d(k,\vec{x}))}{N} + 1$$

This is designated by the ridge along the arc (Fig. 8). Analogous to the shape of the graph in Fig. 1b, the ridge has triangular cross-sections. Its width has a maximum of $2T$ at point R and narrows towards both ends. The ridge has an arc length of 180° (Fry, 2001). There are three lines perpendicular to the vectors $\vec{\varepsilon}$ and $\vec{\varepsilon}'$ in the five-dimensional space. Therefore, there are actually three ridges that meet at right angles at the point R. Eq. (10) denotes the Hough transform of fault-slip data. The local maxima of $F(\vec{x})$ appear at the loci where ridges meet on the hypersphere.

If the entire fault-slip data match to the only one reduced stress, the ridges cross each other at the point that corresponds to the stress. Heterogeneous data are represented by ridges that meet at multiple points. The global maximum of $F(\vec{x})$ represents the most probable stress for the data. Every local maximum is a possible solution and their heights designate the probability of the corresponding stresses for the data.

The local maxima appear at the loci where ridges meet. Fig. 9 shows a 'Mercator chart' of those ridges on

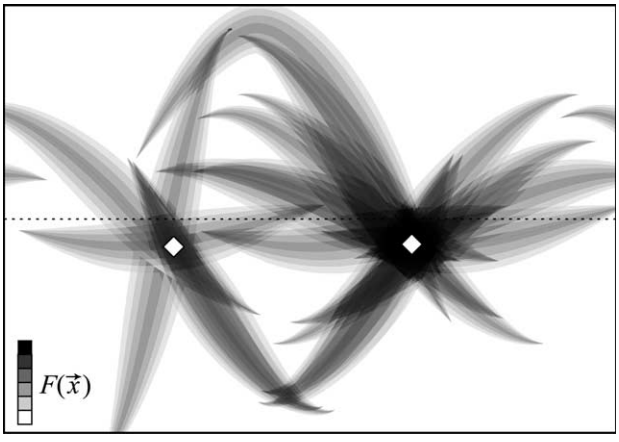

Fig. 9. Analogue of 'Mercator chart' showing the topography $F(\vec{x})$ for the heterogeneous fault-slip data that are generated with two stress tensors. Diamonds designate the stresses. The value of $F(\vec{x})$ is indicated by a grey scale. A single fault-slip datum is represented by a worm-like pattern, the axis of which corresponds to the arc PRQ in Fig. 8. There are local maxima that appear where the patterns cross each other. Those loci manifest themselves as 'unexpected' solutions in the result of stress tensor inversion.

the five-dimensional hypersphere projected onto a four-dimensional hyperplane. The chart is an analogue of the Hough transform in Fig. 10c. A ridge is schematically illustrated in the chart as a sinusoidal long grey zone looking like a striped worm, both ends of which correspond to points P and Q in Fig. 8. The line PQ represents three lines that are perpendicular to the couple of vectors $\vec{\varepsilon}$ and $\vec{\varepsilon}'$. Therefore, a starfish with three pairs of arms, which meet at right angles at point R, is actually a more suitable simile than a 'worm.' However, there is no way to draw the three pairs, so the 'worm' is intended to represent one of the pairs. The local maxima appear where the 'worms' lie one on the other.

All sinusoidal 'worms' in the chart have an arc length of 180°, long enough to allow them to overlap by chance. Therefore, when we generate artificial data with heterogeneity, it is not surprising if the ridges not only meet at the loci corresponding to the assumed stresses but also meet at other places. Those crossings give rise to 'unexpected' stresses.

Unexpected stresses determined from heterogeneous fault-slip data are not always spurious solutions for stress tensor inversion.

6. Summary

A numerical technique based on the Hough transform is presented. The present method enumerates the possible reduced stress tensors for given heterogeneous fault-slip data and evaluates the probability of the tensors.

There is no reason to reject the possible stress tensors if we deal with natural data. Any method often yields unexpected stresses from artificial data if they have heterogeneity. The appearance of those stresses is not a sign of the defect of a method. In this respect, the criterion for judging the performance of the method for separating stresses from heterogeneous data should be changed. Stress tensor inversion based on the Wallace–Bott hypothesis constrains or enumerates possible stresses that are compatible with the parts of heterogeneous data.

Acknowledgements

Comments by C.L. Liesa, R.J. Lisle and an anonymous reviewer considerably improved the manuscript. Financial support came from JSPS (C2-14540423).

Appendix A. Hough transform

The Hough transform is a technique for identifying the locations and orientations of certain types of features in a digital image. The transform explains the mechanics of our method for detecting reduced stress tensors from heterogeneous data. Instead of the locations and orientations, we are

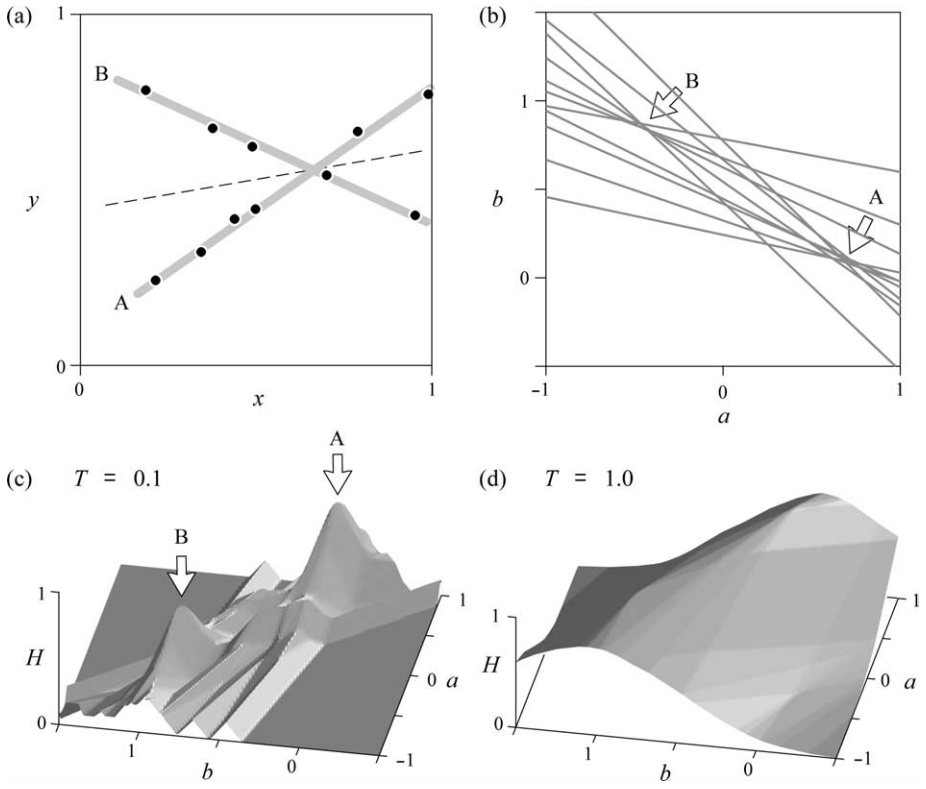

Fig. 10. Detection of lines using the Hough transform. (a) Data points indicated by closed circles on the *xy*-plane. The points make up two lines A and B. Dashed line shows the line determined by the ordinary least-square regression. (b) Grey lines showing the Hough transform of the points in (a). The lines meet at or near the two points A and B. The two points represents the two lines in (a). Surfaces showing $H(a,b)$ for the data points for the cases of $T=0.1$ and 1.0, respectively, in (c) and (d). The prominent peaks correspond to lines A and B in (a), but there are many other peaks and ridges in (c).

determining the shapes and orientations of stress ellipsoids compatible for the data.

Suppose that we have n data points that lie on plural lines on the xy-plane (Fig. 10a). Linear regression can determine the best-fit line for all the points, but the line in this case is meaningless. In contrast, the Hough transform can detect the lines. The simplest form in which to express a line on the plane is

$$y = ax + b \tag{A.1}$$

If the n points from $(x^{(1)}, y^{(1)})$ through $(x^{(n)}, y^{(n)})$ are co-linear, all these couples of x and y satisfy Eq. (A.1) with the same a and b. The ab-plane is the parameter space for the detection of lines. The xy- and ab-planes have duality: every point on the ab-plane corresponds to a line on the xy-plane and every point on the xy-plane corresponds to the line

$$b = -ax + y \tag{A.2}$$

on the ab-plane. This line is the Hough transform of the point (x,y). The line denoted by Eq. (A.2) has the slope $-x$ and the intercept y. Therefore, the lines indicating the Hough transform of the n points meet at the same point on the ab-plane (Fig. 10b). Therefore, the detection of a line on the xy-plane results in the detection of a point on the ab-plane where multiple tines meet. The next problem is to detect those points on the ab-plane.

The ab-plane is subdivided into 'accumulation cells,' which represent different lines on the xy-plane. The Hough transform works by letting each data point on the xy-plane 'vote' to the cells that corresponds to the lines passing through the data point. In consequence, the cells with the maximum number of votes represent the linear patterns that are made up of the points in the xy-plane. In this way, the line detection reduces to an optimization problem, i.e. the search for the cells that have got the maximum votes.

The Hough transform is useful for detecting multiple lines. The cell with the global maximum corresponds to the line along which the maximum number of data points are aligned, but a local maximum indicates another possible alignment. There are two lines in Fig. 10a: lines A and B consist of 7 and 5 points, respectively. Corresponding to the lines, the Hough transform of the points intersect at or near the two points A and B on the ab-plane (Fig. 10b). More lines meet around point A than point B, because line A has more points than line B.

However, there are problems in this technique. There are always local maxima other than the significant ones in the parameter space and most local maxima are spurious. Accordingly, the peaks standing out from others should be identified. In addition, measurement errors or noise disperse the data points on the xy-plane and further scatter the intersection points in the parameter space. The number of votes is, therefore, affected by the noise. A noise reduction technique is usually required for the Hough transform (Hunt et al., 1990). A simple way for stabilizing the voting is carried out by casting votes not only to the cells along the line denoted by Eq. (A.2) but also to the cells in a zone along the line (Thrift and Dunn, 1983). One more modification for the stabilization is to replace the parameter space. For this example, the equation $x\cos\theta + y\sin\theta = \rho$ is another expression of a line on the xy-plane. It is known that ρ and θ constitute a better parameter space than a and b (Duda and Hart, 1972). The choice and discretization of parameter space are important.

Appendix B. Unit vectors in the five-dimensional parameter space

The five-dimensional unit vector denoting a reduced stress tensor $\sigma = (\sigma_{ij})$ has the components

$$\vec{x} = \begin{pmatrix} -\left(\frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{12}\right)\sigma_{11} + \left(\frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{12}\right)\sigma_{22} + \frac{1}{\sqrt{6}}\sigma_{33} \\ \left(\frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{12}\right)\sigma_{11} - \left(\frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{12}\right)\sigma_{22} + \frac{1}{\sqrt{6}}\sigma_{33} \\ \sigma_{12} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} \tag{B.1}$$

The tensor is assumed here to be deviatoric:

$$\sigma_{11} + \sigma_{22} + \sigma_{33} = 0 \tag{B.2}$$

Accordingly, σ has one-to-one correspondence with $\vec{x}$ via Eqs. (B.1) and (B.2).

On the other hand, a fault-slip datum is represented by the unit vectors, $\vec{n}$ and $\vec{v}$, i.e. $\vec{n} = (n_1, n_2, n_3)^{\mathrm{T}}$ is the fault normal and $\vec{v} = (v_1, v_2, v_3)^{\mathrm{T}}$ indicates the slip direction of the hanging-wall block. We also use the auxiliary unit vector $\vec{b} = (b_1, b_2, b_3)^{\mathrm{T}} = \vec{n} \times \vec{v}$. The datum is represented in the five-dimensional parameter space by the paired unit vectors

$$\vec{\varepsilon} = \begin{pmatrix} \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}}\right)v_1 n_1 - \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\right)v_2 n_2 + \sqrt{\frac{2}{3}}v_3 n_3 \\ -\left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\right)v_1 n_1 + \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}}\right)v_2 n_2 - \sqrt{\frac{2}{3}}v_3 n_3 \\ -v_2 n_3 - v_3 n_2 \\ -v_3 n_1 - v_1 n_3 \\ -v_1 n_2 - v_2 n_1 \end{pmatrix}$$

and

$$\vec{\varepsilon}' = \begin{pmatrix} \left(\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{6}}\right)b_1n_1 - \left(\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{6}}\right)b_2n_2 + \sqrt{\frac{2}{3}}b_3n_3 \\ -\left(\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{6}}\right)b_1n_1 + \left(\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{6}}\right)b_2n_2 - \sqrt{\frac{2}{3}}b_3n_3 \\ -b_2n_3 - b_3n_2 \\ -b_3n_1 - b_1n_3 \\ -b_1n_2 - b_2n_1 \end{pmatrix}$$

See Sato and Yamaji (2006a) for further explanations.

References

Angelier, J., 1977. La reconstitution dynamique et géométrique de la tectonique de failles à partir de mesures locales (plans de faille, stries, sens, de jeu, rejets): quelques précisions. Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences D285, 637–640.

Angelier, J., 1979. Determination of the mean principal directions of stresses for a given fault population. Tectonophysics 56, T17–T26.

Angelier, J., 1984. Tectonic analysis of fault-slip data sets. Journal of Geophysical Research 89, 5835–5848.

Angelier, J., 1994. Fault slip analysis and paleostress reconstruction. In: Hancock, P.L. (Ed.), Continental Deformation. Pergamon Press, Oxford, pp. 53–101.

Angelier, J., Mechler, P., 1977. Sur une méthode graphique de recherche des contraintes principales également utilisable en tectonique et en séismologie: la méthode des dièdres droits. Bulletin de la Societé Géologique de France 19, 1309–1318.

Angelier, J., Colletta, B., Anderson, R.E., 1985. Neogene paleostress changes in the Basin and Range: a case study at Hoover Dam, Nevada-Arizona. Geological Society of America Bulletin 96, 347–361.

Armijo, R., Carey, E., Cisternas, A., 1982. The inverse problem in microtectonics and the separation of tectonic phases. Tectonophysics 82, 145–160.

Bergerat, F., 1987. Stress fields in the European platform at the time of Africa–Eurasia collision. Tectonics 6, 99–132.

Bott, M.H.P., 1959. The mechanics of oblique slip faulting. Geological Magazine 96, 109–117.

Carey, E., Brunier, B., 1974. Analyse théorique et numérique d'un modèle mécanique élémentaire appliqué à l'étude d'une population de failles. Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences D279, 891–894.

Casas, A.M., Gil, I., Simón-Gómez, J.L., 1990. Los métodos de análisis de paleoesfuerzos a partir de poblaciones de fallas: sistemática y técnicas de aplicación. Etudios Geologicos 46, 385–398.

Choi, P.-Y., Angelier, J., Souffaché, B., 1996. Distribution of angular misfits in fault-slip data. Journal of Structural Geology 18, 1353–1367.

Choi, P.-Y., Kwon, S.K., Hwang, J.H., Lee, S.R., An, G.O., 2001. Paleostress analysis of the Pohang–Ulsang area, southeast Korea: tectonic sequence and timing of block rotation. Geoscience Journal 5, 1–18.

Duda, R.O., Hart, P.E., 1972. Use of the Hough transform to detect lines and curves in pictures. Communications of the Association for Computing Machinery 15, 11–15.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification, 2nd ed John Wiley & Sons, New York.

Etchecopar, A., Vasseur, G., Daignieres, M., 1981. An inverse problem in microtectonics for the determination of stress tensors from fault striation analysis. Journal of Structural Geology 3, 51–65.

Fry, N., 1992. A robust approach to the calculation of paleostress fields from fault plane data: discussion. Journal of Structural Geology 14, 635–637.

Fry, N., 1999. Striated faults: visual appreciation of their constraint on possible paleostress tensors. Journal of Structural Geology 22, 441–452.

Fry, N., 2001. Stress space: striated faults, deformation twins, and their constraints on paleostress. Journal of Structural Geology 23, 1–9.

Galindo-Zaldívar, J., González-Lodeiro, F., 1988. Faulting phase differentiation by means of computer search on a grid pattern. Annales Tectonicae 2, 90–97.

Gephart, J.W., 1990. Stress and the direction of slip on fault planes. Tectonics 9, 845–858.

Huang, Q., 1988. Computer-based method to separate heterogeneous sets of fault-slip data into sub-sets. Journal of Structural Geology 10, 297–299.

Hunt, D.J., Nolte, L.W., Reibman, A.R., Ruedger, W.H., 1990. Hough transform and signal detection theory performance for images with additive noise. Computer Vision, Graphics, and Image Processing 52, 386–401.

Leavers, V.F., 1992. Shape Detection in Computer Vision Using the Hough Transform. Springer-Verlag, London.

Liesa, C.L., Lisle, R.J., 2004. Reliability of methods to separate stress tensors from heterogeneous fault-slip data. Journal of Structural Geology 26, 559–572.

Lisle, R.J., Vandycke, S., 1996. Separation of multiple stress events by fault striation analysis: an example from Variscan and younger structures at Ogmore, South Wales. Journal of Geological Society, London 153, 945–953.

McKenzie, D., 1969. The relation between fault plane solutions for earthquakes and the directions of the principal stresses. Bulletin of the Seismological Society of America 59, 591–601.

Nemcok, M., Lisle, R.J., 1995. A stress inversion procedure for polyphase fault/slip data sets. Journal of Structural Geology 17, 1445–1453.

Orife, T., Lisle, R.J., 2003. Numerical processing of paleostress results. Journal of Structural Geology 25, 949–957.

Otsubo, M., Sato, K., Yamaji, A., 2006. Computerized identification of stress tensors determined from heterogeneous fault-slip data by combining multiple inverse method and *k*-means clustering. Journal of Structural Geology, this issue, doi.:10.10106/j.jsg.2006.03.008.

Ramsay, J.G., Lisle, R.J., 2000. The Techniques of Modern Structural Geology. Volume 3: Applications of Continuum Mechanics in Structural Geology. Academic Press, London.

Rousseeuw, R.J., Leroy, A.M., 1986. Robust Regression and Outlier Detection. John Wiley & Sons, New York.

Sato, K., Yamaji, A., 2006. Embedding stress difference in parameter space for stress tensor inversion. Journal of Structural Geology, this issue, doi.:10.10106/j.jsg.2006.03.004.

Sato, K., Yamaji, A., 2006. Uniform distribution of points on a hypersphere for improving the resolution of stress tensor inversion. Journal of Structural Geology, this issue, doi.:10.10106/j.jsg.2006.03.007.

Shan, Y.H., Suen, H.B., Lin, G., 2003. Separation of polyphase fault/slip data: an objective-function algorithm based on hard division. Journal of Structural Geology 25, 829–840.

Shan, Y.H., Li, Z., Lin, G., 2004. A stress inversion procedure for automatic recognition of polyphase fault/slip data sets. Journal of Structural Geology 25, 829–840.

Stockman, G.C., Agrawala, A.K., 1977. Equivalence of Hough curve detection to template matching. Communications of the Association for Computing Machinery 20, 820–822.

Thrift, P.R., Dunn, S.M., 1983. Approximating point-set images by line segments using a variation of the Hough transform. Computer Vision, Graphics and Image Processing 21, 383–394.

Twiss, R.J., Moores, E.M., 1992. Structural Geology. W.H. Freeman and Co., New York.

Vandycke, S., Bergerat, F., 2001. Brittle tectonic structures and paleostress analysis in the Isle of Wight, Wessex basin southern U.K.. Journal of Structural Geology 23, 393–406.

Wallace, R.E., 1951. Geometry of shearing stress and relationship to faulting. Journal of Geology 59, 111–130.

Will, T.M., Powell, R., 1991. A. robust approach to the calculation of paleostress fields from fault plane data. Journal of Structural Geology 13, 813–821.

Will, T.M., Powell, R., 1992. A robust approach to the calculation of paleostress fields from fault plane data: reply. Journal of Structural Geology 14, 639–640.

Yamaji, A., 2000. The multiple inverse method: a new technique to separate stresses from heterogeneous fault-stop data. Journal of Structural Geology 22, 441–452.

Yamaji, A., 2003. Are the solutions of stress inversion correct? Visualization of their reliability and to separate stresses from heterogeneous fault-slip data. Journal of Structural Geology 24, 241–252.