

SAPE: a program for semi-automatic parameter extraction for strain analysis

Kieran F. Mulchrone^{a,*}, Patrick A. Meere^b, Kingshuk Roy Choudhury^c

^aDepartment of Applied Mathematics, National University of Ireland, Cork, Ireland

^bDepartment of Geology, National University of Ireland, Cork, Ireland

^cDepartment of Statistics, National University of Ireland, Cork, Ireland

Received 8 October 2003; received in revised form 12 December 2004; accepted 23 May 2005

Available online 10 August 2005

Abstract

SAPE (Semi-Automatic Parameter Extraction) is a program for extracting information relevant to strain analysis from input digital images. Input images are manually produced by tracing the outlines of objects of interest. This is a less onerous, more objective and much faster task than manual measurement or digitisation of object data (i.e. aspect ratio R_i and orientation ϕ_i). SAPE rapidly extracts the required data by using a simple region-growing algorithm to identify regions of interest. Subsequently, the second moments are calculated for each region enabling the common strain analysis parameters to be readily computed. The performance of SAPE was tested on three samples (microphotographs of deformed sandstone and oolite as well as a meter scale photograph of a deformed conglomerate) using a non-overlap statistic. This statistic is a normalised measure of overlap between two regions and can be used to compare different fits applied to the same region. In each case SAPE out-performed the manual method. SAPE offers significant advantages in terms of speed, objectivity and robustness and should facilitate the collection of large datasets for strain analysis.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Software; Strain analysis; Parameter extraction; Semi-automatic; Comparison

1. Introduction

Strain determination using populations of suitable objects is an extremely useful technique for understanding the processes and products of deformation of the Earth's crust from the micro- to macro-scales. Based on the passive behaviour of elliptical objects during deformation Ramsay (1967) introduced the R_i/ϕ method for strain analysis. Since then there have been many graphical, approximate and algebraic methods developed, based on the same assumptions (Dunnet, 1969; Elliott, 1970; Dunnet and Siddans, 1971; Matthews et al., 1974; Borradaile, 1976; Shimamoto and Ikeda, 1976; Lisle, 1977a,b, 1985; Peach and Lisle, 1979; Yu and Zheng, 1984; Mulchrone and Meere, 2001; Mulchrone et al., 2003). Up until recently only the method

of Robin (1977) was considered general enough to work with non-elliptically shaped objects. However, Mulchrone and Roy Choudhury (2004) have demonstrated that if the objects used in an analysis can be identified using either boundary or region based ellipse-fitting methods, then most previously developed algebraic techniques can be applied to objects of arbitrary shape. An alternative class of methods based on object-object separations was initially championed by Fry (1979) and has undergone considerable development, spawning the Normalised Fry Method (Erslev, 1988) and the enhanced Normalised Fry Method (Erslev and Ge, 1990). Additionally, McNaught (1994) has extended these methods for application to aggregates of non-elliptical objects and Mulchrone (2003) applied Delaunay triangulation to resurrect the nearest neighbour method of strain analysis.

One disadvantage associated with all population-based strain analysis is the effort required to obtain the raw data. In a regional strain study of the order of 20–50 samples (preferably more) require analysis and for a 3D strain estimate this usually requires analysis of at least three 2D sections per sample. Given that Meere and Mulchrone

* Corresponding author. Tel.: +353 21 490 3411; fax: +353 21 42 708 13.

E-mail address: k.mulchrone@ucc.ie (K.F. Mulchrone).

(2003) have shown that a minimum of 150 measurements per 2D section are required for sufficient statistical accuracy, this amounts to approximately 10,000–20,000 individual object identifications. Manual data acquisition, i.e. grain identification and ellipse fitting will clearly be a time consuming and unenviable task, but has had to be done in the past. If the objects used are of arbitrary shape, then a numerical ellipse fitting technique needs to be applied (Mulchrone and Roy Choudhury, 2004). This is not practical to attempt manually. Clearly then, some form of automation is required to aid in the acquisition of raw data for strain analysis.

Methods for automated image analysis have been presented in the past and have been successfully applied in geological analysis. Erslev and Ge (1990) developed the program INSTRAIN, which takes as input digitised object outlines and then calculates strain using the enhanced Fry, enhanced normalised Fry and mean object ellipse methods. McNaught (1994) extracts data for analysis by first producing an overlay tracing and the selecting boundary points using a digitising tablet for subsequent polygonal representation of objects. Masuda et al. (1991) used a LUZEX image analyser to measure grain areas in microphotographs and Ailleres et al. (1995) developed both hardware and third party software configurations for semi-automatic grain identification for Fry strain analysis (Fry, 1979). In applying the projection method, Panozzo (1984) utilised digitised sets of points representing linear or elliptical objects. Mukul and Mitra (1998) analysed an unprecedented 119 samples of quartzite using the modified normalised Fry method (McNaught, 1994) from an area of 200 km² around the Sheeprock Thrust Sheet, Sevier Fold-and-Thrust belt, Utah, USA. In order to achieve this high level of sampling a semi-automatic procedure for obtaining the raw data for strain analysis as described by Mukul (1998) was utilised. This involved taking microphotographs of thin sections and manually tracing quartz grain boundaries. Subsequently, third party image analysis software identified the centroids and areas of the grains. Panozzo Heilbronner and Pauli (1993) developed a ‘CIP’ method to identify *c*-axes of quartz grains. Heilbronner (2000) also developed a ‘Lazy Grain Boundary’ method that can be used in a fully automatic or semi-automatic mode to identify grain boundaries of various objects.

In this paper we present a method called Semi-Automated Parameter Extraction (SAPE). SAPE enables rapid parameter extraction from any material or surface that (a) can be photographed and (b) a human can distinguish the objects to be studied. This makes SAPE applicable to almost every type of lithology. In operational terms, SAPE is similar to the methods described in Erslev and Ge (1990) and Mukul (1998), in that all of them use manually drawn boundaries as input. The difference is in the output: Erslev and Ge (1990) and Mukul (1998) extract parameters necessary for centre-to-centre analysis and SAPE extracts parameters necessary for R_f/ϕ analyses.

2. SAPE methodology

2.1. Boundary/region identification

The most difficult part of extracting information about a population of objects in a digital image, from the perspective of computer implementation, is identification of boundaries or regions occupied by those objects. On the other hand, this is a trivial exercise for humans. Once the regions have been identified, parameters such as the aspect ratio and orientation of the object or second moments must be extracted. Humans can perform this task, but in a subjective manner that may introduce unwarranted biases into the data acquired (Mulchrone and Roy Choudhury, 2004). Additionally, accurate parameter estimation and recording of the same is time consuming. By contrast, parameter estimation is a trivial task for computers. The idea driving SAPE is to divide the tasks involved in parameter extraction according to human/computer aptitude.

Therefore, the task of identifying regions is relegated to the human user. The raw input for this activity is a digital image, which may be a microphotograph of the thin section or a field photograph of the rock surface containing the marker objects. This is referred to as the *raw image*. The end

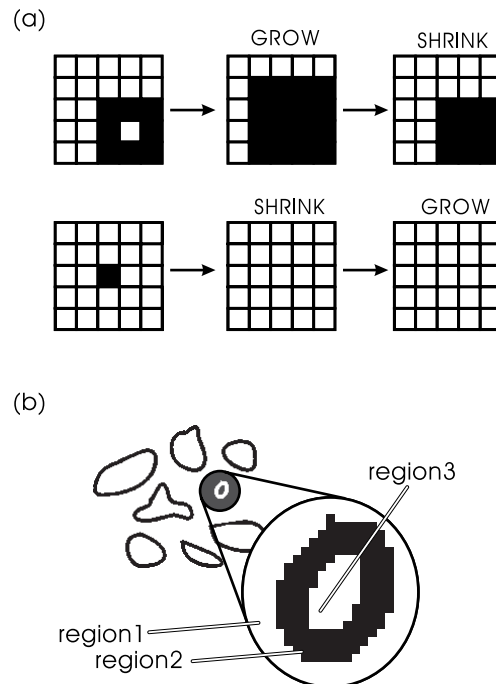


Fig. 1. (a) Examples of the simple growth and shrinking algorithms. In the upper case a spurious white spot is removed by a single growth followed by a single shrink. A shrink followed by a growth removes black spots. Spurious white inclusions were the most common and disruptive blemishes found in grain outlines as these could be erroneously recognised as grains. Therefore growth followed by shrink is implemented in SAPE. (b) Typical form of input image required by SAPE with white regions enclosed by black outlines on a white background. Also marked are the three regions referred to in the main text.

result of this activity must be a white digital image with object boundaries identified in black (see Fig. 1), which is referred to as the *input image*. There are many possible approaches to producing the input image and we investigated two. First, we imported the raw image into popular image editing software packages such as Freehand or CorelDraw and traced out the object boundaries in the software to produce the input image. Second, we printed out a high quality copy of the raw image and, using an inking pen and tracing paper, produced a physical tracing. Subsequently, this was digitally scanned to produce the input image. For a large volume of work the second approach was found to be faster and mentally more bearable. The approach of tracing followed by image analysis was first advocated by Mukul (1998) leading to a high quality dataset (Mukul and Mitra, 1998). However, we have developed software specifically aimed at extracting data for the purpose of R_t/ϕ strain analysis.

2.2. Image analysis

SAPE takes the input image and extracts the required parameters with only minor further human intervention. SAPE expects the input image to be in standard bitmap format and utilises some standard image analysis techniques (see for example Seul et al., 2000, pp. 118–175) to identify the regions enclosed by the black boundaries in the input image. The software for image analysis is object-oriented and implemented in C++. There are a number of steps involved in the analysis:

1. Extract input image data.
2. Boundary smoothing.
3. Background identification.
4. Region identification.
5. Parameter extraction.

Access to the input image data is facilitated by the open-source library of image handling functions FreeImage (see <http://freeimage.sourceforge.net>). The input image data is copied into a two-dimensional array of data such that rows vary from 0 to $h-1$ (h , the height of the image in pixels) and columns vary from 0 to $w-1$ (w , the width of the image in pixels) and that the data at (0,0) corresponds to the top-left of the input image and that at ($h-1, w-1$) corresponds to the bottom-right of the image. Each data point consists of the following data elements (*value*, *label*, *grow*, x , y). The *value* parameter is 0 for a white pixel and 1 for black; *label* and *grow* are used in region identification, whereas x and y specify the location of the data point in the array. Initially, every data point has *label* set to 0 and *grow* set to 1.

Boundary smoothing is applied to reduce roughness at the pixel scale and to eliminate spurious spots and blemishes on the input image. Smoothing is accomplished by applying a pixel growth algorithm n -times followed by a pixel shrinking algorithm m -times, typically applied only once or

twice. The growth algorithm works by scanning through all the data points and if there are one or more black-valued data points adjacent to a white-valued data point (i.e. directly or diagonally), then the white data point is changed to black. The shrinking algorithm works in reverse by turning black points to white if they have one or more adjacent white points. Simple examples of these algorithms are provided in Fig. 1a.

Both background and region identification are based largely on a simple region growing (SRG) algorithm (Adams and Bischof, 1994). To explain the algorithm consider the situation shown in Fig. 1b and concentrate on the simplified problem of the magnified inset. There are three regions of data: (1) white data points outside the boundary, (2) black data points defining the boundary, and (3) white data points inside the boundary. For SRG it is assumed that one data point, termed the seed, in region 3 is known (it will be explained later how seed pixels are identified). A unique value for *label* (>0) is assigned to the seed. The GrowSeed function (see Fig. 2 for pseudocode) is created from the position in the array of the seed data point and the unique *label* value to be assigned to data points in

```

GrowSeed(int x, int y, int lab)
{
//x,y gives position of seed pixel in array
//lab is a unique label for pixels in the grown region
COBList border, nextborder; //empty linked lists
int count = 1, added, lvalue = data[x][y].value;
//pixels with value == lvalue are added to the region
border.Add(data[x][y]) //add seed to border
while(count==1) //stop when no pixels are added
{
    count = 0;
    for each border pixel bp
    {
        added = 0;
        if(bp.grow == 0) //i.e. pixel cant grow
            continue; //go to next border pixel
        for each neighbouring pixel np
        {
            if(np.val == lvalue and np.label != lab)
            {
                np.label = lab;
                nextborder.add(np);
                added++;
                count++;
            }
        }
        if(added == 0) //no pixels added for bp
        {
            nextborder.add(bp);
            bp.grow = 0;
            //i.e. this pixel can't grow
        }
    }
    remove all pixels from border;
    copy pixels from nextborder into border;
}
}

```

Fig. 2. Pseudocode for the simple region-growing algorithm.

the grown region. This algorithm works by progressively building a list of pixels on the border of the growing region that grows until it hits black pixels in region 2. Two lists of data points are maintained called *border* and *nextborder*. Note that the seed point has a unique *label* and all other points have *label*=0, and from initialisation all points have *grow*=1 meaning that they can grow, if necessary. To begin, the seed point is added to *border*. Next, for each data point in *border* (initially just the seed) the adjacent data points (eight in total) are checked. Any point with *grow*=1, and the same *value* as the seed and a *label* different from that of the seed will be added to *nextborder*. If no adjacent data points were added then the current data point can no longer grow and *grow* is set to 0 and the current point is added to *nextborder*. This procedure is repeated until such time as no more new data points are added to *border*, i.e. the region has grown to its full extent.

The program assumes that the data point at (0,0) represents a white pixel and the background region is identified by using this pixel as a seed and growing a region with label 1. Subsequently a scan of all the data points is performed. Any data point encountered with *value*=0 and *label*=0 is not in the background region (i.e. *label*=1) or another grown interior region (i.e. *label*>1) or a boundary (boundary pixels have *value*=1) and so must be a seed. Once a seed pixel is encountered, the value of *label* is incremented by one to get a unique value and the GrowSeed function is created for the seed and new unique *label*.

After image processing is complete (i.e. steps 1–4) the array of data points contains background pixels with *label*=1, region pixels with *label*≥2 and border pixels with label 0. The final step of parameter extraction simply consists of sequentially selecting the pixels that belong to each region with *label*≥2 and calculating the second moments of each such region using the methods of Mulchrone and Roy Choudhury (2004). More formally, a set of functions can be defined on the rectangular region *G*, which encloses the entire input image, i.e. in terms of pixels *G* goes from 0 to *w*−1 and from 0 to *h*−1:

$$b_k = (i, j) = \begin{cases} 1 & \text{if data}[i][j].\text{label} = k \\ 0 & \text{if data}[i][j].\text{label} \neq k \end{cases} \quad (1)$$

where $k \geq 2$. The $(p+q)$ th moment of the k th region is then given by:

$$(m_{pq})_k = \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} i^p j^q b_k(i, j) \quad (2)$$

The central second moments for each region are given by:

$$(u_{20})_k = (m_{20})_k - \frac{(m_{10})_k^2}{(m_{00})_k} \quad (3)$$

$$(u_{02})_k = (m_{02})_k - \frac{(m_{01})_k^2}{(m_{00})_k} \quad (4)$$

$$(u_{11})_k = (m_{11})_k - \frac{(m_{10})_k(m_{01})_k}{(m_{00})_k} \quad (5)$$

and, shown in Mulchrone and Roy Choudhury (2004), the aspect ratio (R_k), orientation (ϕ_k), length of major (a_k) and minor (b_k) axes of the best fit ellipse of the k th region are then given by:

$$\Delta_k = \sqrt{4(u_{11})_k^2 + ((u_{20})_k - (u_{02})_k)^2} \quad (6)$$

$$\phi_k = \frac{1}{2} \tan^{-1} \left(\frac{2(u_{11})_k}{(u_{20})_k - (u_{02})_k} \right) \quad (7)$$

$$a_k = \sqrt{\frac{2((u_{20})_k + (u_{02})_k + \Delta_k)}{(u_{11})_k}} \quad (8)$$

$$b_k = \sqrt{\frac{2((u_{20})_k + (u_{02})_k - \Delta_k)}{(u_{11})_k}} \quad (9)$$

$$R_k = \sqrt{\frac{(u_{20})_k + (u_{02})_k + \Delta_k}{(u_{20})_k + (u_{02})_k - \Delta_k}} \quad (10)$$

3. Using SAPE

The program has a simple menu/toolbar driven interface and is written in C++ (Visual C++ version 6.0) and is designed to minimise user error. An overview of the program options is given in Table 1. By choosing File → Import Image, an input image file for processing can be selected through a standard file open dialog interface. This option is always available. File → Save Bitmap enables the current bitmap image on screen to be saved to a bitmap file. File → Export Parameters allows the extracted data to be saved to a tab delimited text file, in several formats, for further processing. File → Export Parameters → All Data exports the region centroid x and y -coordinates, the length of the major and minor best fit ellipse axes, ellipse long axis orientation and ellipse area for each fitted ellipse on each

Table 1
Main menu options available in SAPE

File	Process	View
Import bitmap	Set parameters	Raw data
Save bitmap	Run analysis	Processed data
Export parameters		Extracted data
All data		Processed data and fitted ellipses
R_i/ϕ data		
Moment data		
Print		
Print preview		
Exit		

line of text. File → Export Parameters → R_i/ϕ Data exports only the best fit ellipse aspect ratio and long axis orientation per best fit ellipse per line of text. File → Export Parameters → Moment Data exports the region centroid x and y -coordinates, and the following moments $(u_{00})_k$ $(u_{02})_k$ $(u_{20})_k$ and $(u_{11})_k$. The File → Print and Print Preview options send the current screen content to the printer and File → Exit ends the application.

Under the Process menu Set Parameters brings up a dialog box enabling customisation of parameters affecting the outcome of the main processing algorithm. The data reduction parameters are applied after the analysis is performed and allow for automatic exclusion of processed regions by area. For example, setting the bottom tail% to five means that regions with an area less than the 5% percentile are automatically excluded from the results. This exclusion may be manually overridden at a later stage as described below. In the smoothing section the shrink and grow numbers refer to the number of times the preprocessing shrinking and growth algorithms are applied. By default each of these values are set to 2. In order to apply the changes made in this dialog, the analysis must be re-run. Selecting Process → Run Analysis kicks off the main processing algorithm and is only available when an input image has been previously selected. During processing a progress bar is displayed in the status bar to indicate how much time is remaining before completion.

The View menu enables the user to specify whether or not they want to see the toolbar or status bar. More importantly, it allows the user to select between different views of the data. All views are available only after first selecting an image and then processing it. There are four available views (see Fig. 3): (1) a view of the raw unprocessed image (Raw Data), (2) a view of the processed image including identified regions in turquoise and white labels for each region and excluded regions are labelled in red (Processed Data), (3) a list of the extracted data from the image, again excluded regions are identified (List of Extracted Data), and (4) a view of the labelled processed data with best fit ellipse superimposed in red (Processed Data and Fitted Ellipses). In both of the processed data views the colour of the region label indicates inclusion (white) or exclusion (red). The inclusion or exclusion of a particular region can be toggled by clicking on its label.

There are a number of potential problems associated with SAPE, most of which can be solved with foresight and manual intervention. SAPE works best and without manual intervention if the input image consists of separable regions, i.e. between any two regions that are in contact. However, in practice this may not be the case. For example, in Fig. 4a, a cluster of regions enclose an area of background. SAPE will mistakenly identify the surrounded background as an additional region. However, this problem is easily solved by manual intervention. SAPE cannot cope with incompletely outlined regions (see Fig. 4b) and these will usually be consigned to the background region. Another possible

problem occurs if the top-left portion of the image is not white, which causes the background identification part of the algorithm to fail, which usually means that it is mistakenly identified as a large region.

4. A validation and comparison study

4.1. Introduction

The effectiveness of the SAPE software was tested by comparing data acquired using SAPE with those using fully manual methods. Three naturally deformed lithologies, a sandstone, oolite and conglomerate were selected for this purpose. These samples represent three distinct lithological types and involve acquisition of data from both the microscopic and mesoscopic scales of observation. The sandstone is a deformed continental quartz wacke from the Upper Devonian of southern Ireland. Data were collected from a photomicrograph of a section cut orthogonal to the cleavage fabric. Oolite data were obtained from a negative print photomicrograph of a deformed ironstone oolite in fig. 5.7 from Ramsay and Huber (1983, p. 79). The conglomerate data was collected from an outcrop scale field photograph of a deformed Silurian cobble conglomerate (quartzite and psammite clasts) from Croagh Patrick, south Mayo, Ireland.

Manual measurement of strain marker parameters (i.e. aspect ratio/long axis orientation R_i/ϕ_i) is very subjective and there is no general consensus on methodology (Mulchrone and Roy Choudhury, 2004). The manual approach adopted here is to first identify the longest possible line through the target shape and take this to be the long axis. The short axis is then taken to be the longest possible line normal to the long axis direction. Alternative approaches might include a simultaneous visual fit of both axes to the marker. For SAPE, digital photographs of the samples were imported into Freehand and manual measurements were made using the line tool. In addition, for each sample a line trace of marker boundaries was made from the digital sample photographs and subsequently scanned into a bitmap image, and used as the input image for SAPE.

A detailed example is illustrated for the sandstone sample in Fig. 5. In Fig. 5a, the original microphotograph is shown and the corresponding tracing (see Fig. 5b) is used as the input for SAPE. Meere and Mulchrone (2003) have estimated that at least 150 grains are required for reliable strain estimation from a microphotograph. Accordingly, we have chosen a selection of 150 grains for this tracing. We also made an effort to choose a representative selection of grains from the microphotographs. The SAPE processed image containing identified regions, labels and best fit ellipses is illustrated in Fig. 5c. To produce this image the growth and shrink algorithms were applied once each and the 5% tails are excluded automatically (i.e. those with red labels). Clearly most of the excluded regions should actually

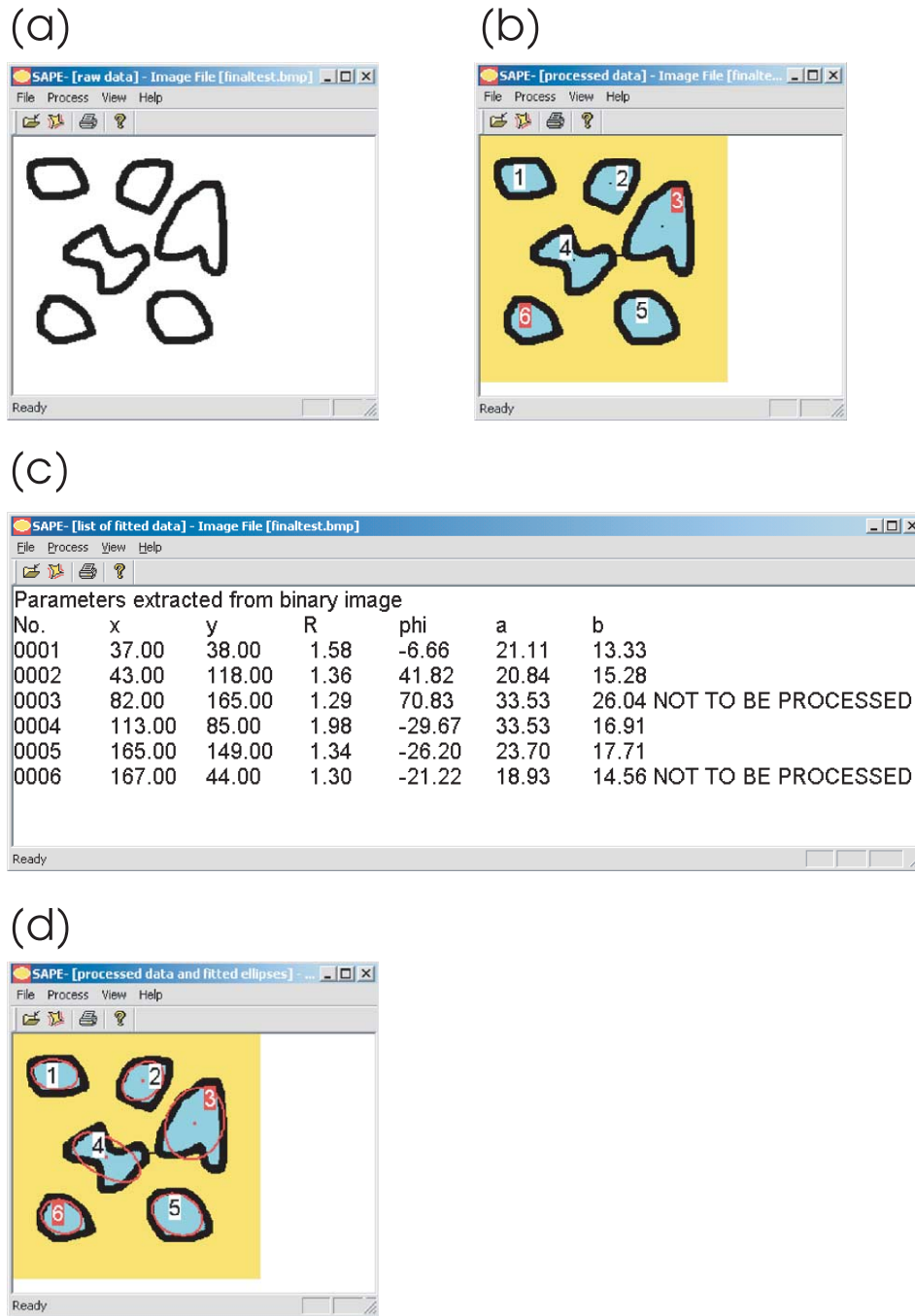


Fig. 3. Views available in SAPE. (a) Unprocessed input image, (b) processed image, (c) list of extracted data, and (d) processed data and fitted ellipses.

be included. This could be remedied either by manually clicking the labels or by re-running the analysis and setting the tails to 0%. Also, in the bottom right of the processed image, a clast outline has not been identified due to a hole in the boundary.

4.2. SAPE versus manual method

Bivariate plots of SAPE versus manual for R_i and ϕ_i data demonstrate a good overall correlation between the manual

and SAPE methods for all samples studied (see Fig. 6). In addition, the fact that the line $y=x$ passes through the middle of the data clusters indicates that there is no systematic under- or overestimation in SAPE measurements relative to manual measurements. Considering the R_i data, it is evident that the correlation between the methods deteriorates for increasing R_i values, a trend best illustrated by the sandstone data (Fig. 6a). This trend is not as apparent for the oolite data (Fig. 6b), which results in a good correlation value ($r=0.95$), primarily because R_i does not

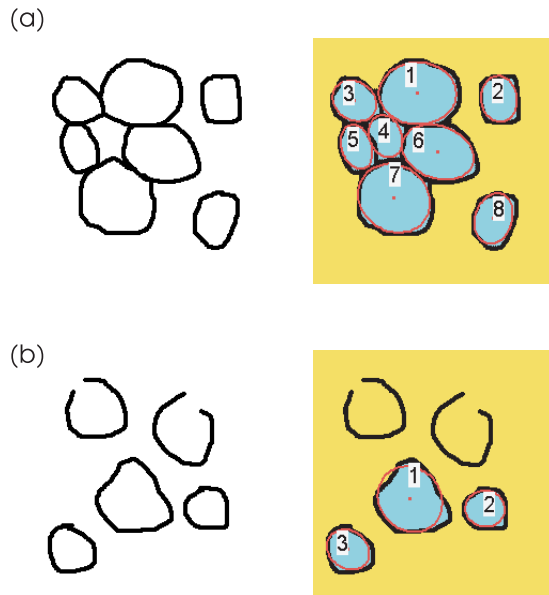


Fig. 4. Two examples of potential problems for SAPE.

take large values. This effect can also be seen in the conglomerate data set that contains R_i values of up to 5.0 and higher (Fig. 6c). However, there is an additional complication with the conglomerate data in that there is a consistent overestimation of R_i by the manual method in comparison with SAPE and hence the data set shows the poorest correlation value of $r=0.77$. One likely explanation for this is related to the actual shape of the conglomerate clasts and the settings used in SAPE. This lithology has undergone significant flattening producing elongated clasts with narrow ‘peninsula-like’ projections in the flattening plane. The manual method takes account of these projections while the SAPE software has a tendency to cut off and exclude them from the main body of the clast due to a combination of narrow region widths in pixels and the growth and shrink settings (both set to 2 for this analysis). This ‘cut-off’ problem is best avoided by scaling the input image such that the width in pixels of all target regions are large, i.e. greater than around 10 pixels and may also be reduced by reducing the growth and shrink settings.

It is interesting to note that the ‘cut-off’ problem does not show up in the ϕ_i data for the conglomerate where a relatively good correlation is observed. Overall the ϕ_i data shows better correlation values than the R_i data and there are no trends indicating more or less correlation along any particular direction.

4.3. Relationship between error and strain marker size

The error associated with fitting ellipses to strain markers is estimated by the difference between values calculated using the manual and SAPE methods. In Fig. 7, plots of the

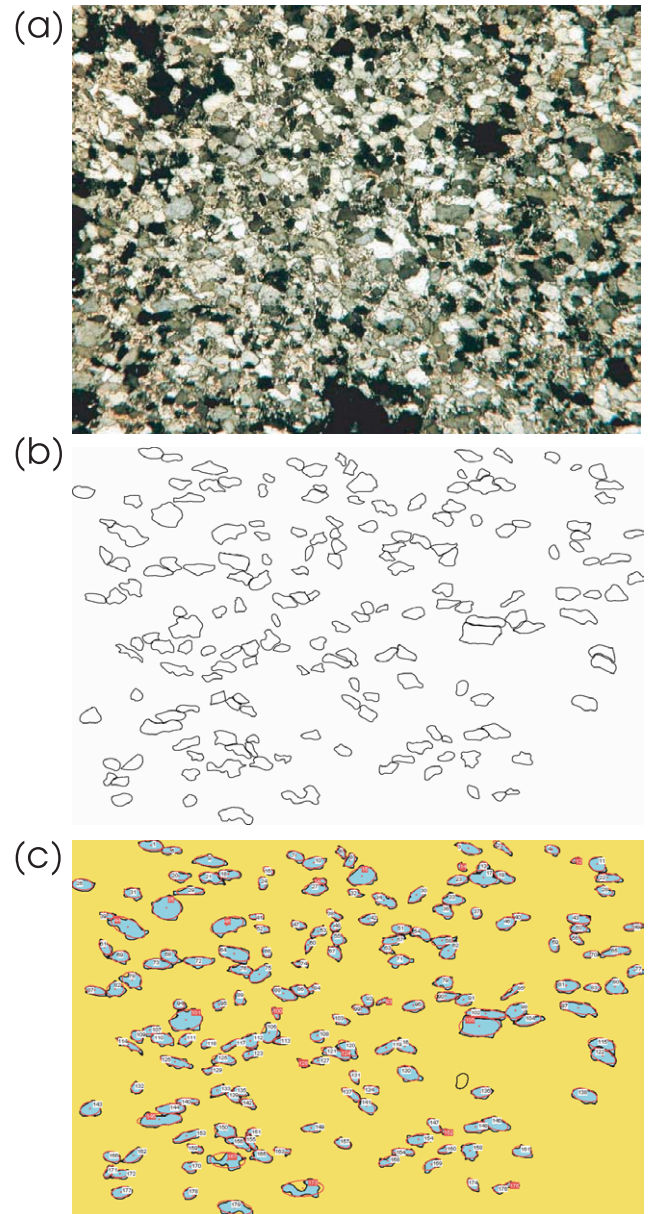


Fig. 5. The sandstone example used for validation. (a) The original microphotograph, (b) object outlines, the input image, and (c) the identified regions with labels and best fit ellipses.

error R_i and ϕ_i against clast area in pixels squared, demonstrate that for both R_i and ϕ_i there tends to be larger errors for small clasts and smaller errors for larger clasts for each sample. This is to be expected because the quality of fit of an ellipse to a given clast will deteriorate with decreasing clast size due to the practical difficulty in accurately defining long and short axes for smaller clast sizes. The effect of R_i on errors was also investigated (see Fig. 8). As might be intuitively expected, larger R_i values show lower errors for ϕ_i values, presumably due to the greater clarity of clast long axes. However, an opposite, almost counterintuitive effect is seen in the R_i data whereby increasing errors correspond to increasing R_i values. This again may be due to the elongate

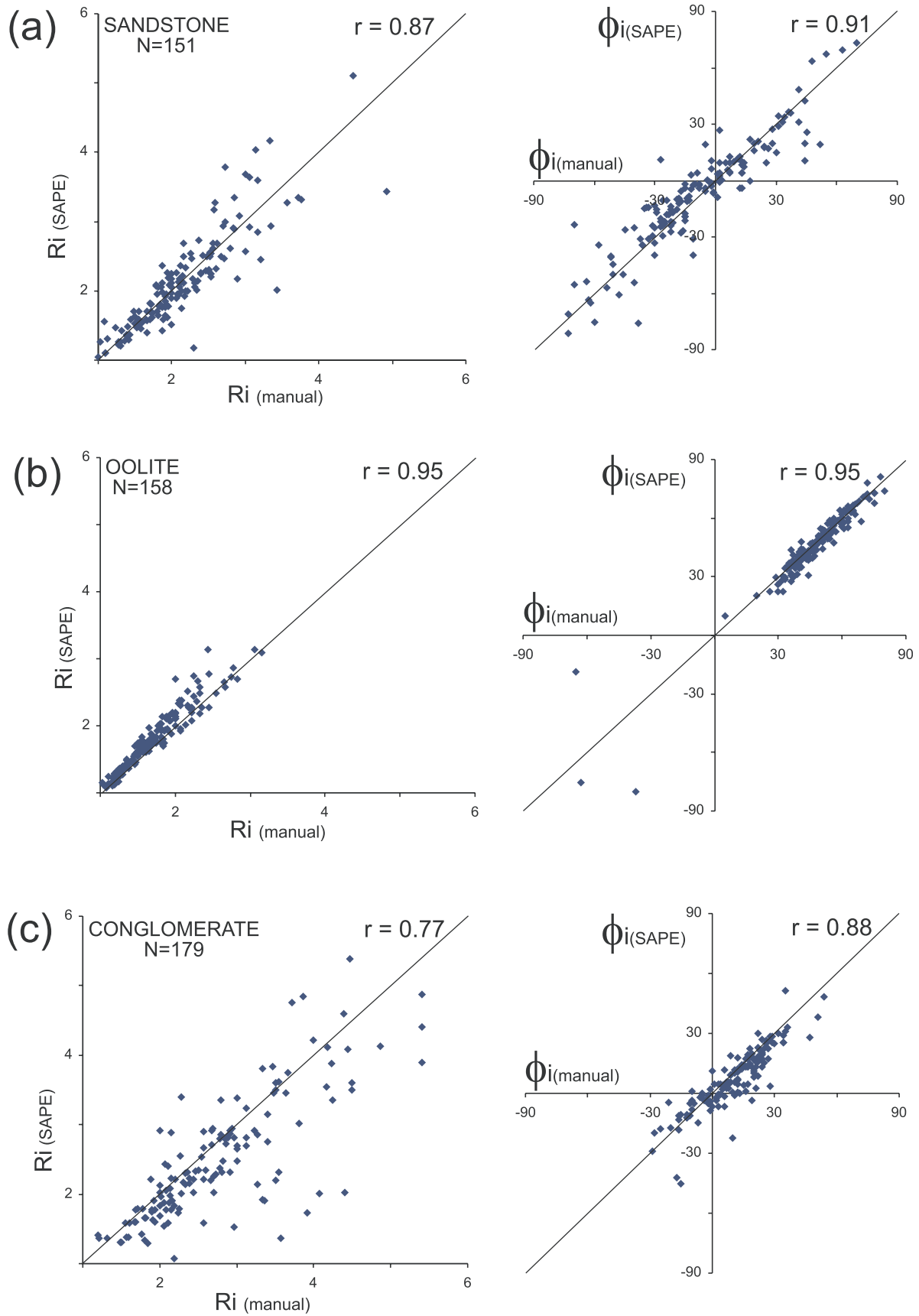


Fig. 6. Bivariate plot for clast aspect ratios (R_i) and long axis orientation (ϕ_i) of SAPE versus manual data for (a) sandstone, (b) oolite, and (c) conglomerate.

shape and narrow terminations of high R_i clasts that are excluded by the SAPE software and result in shorter clast long axis estimates.

5. Non-overlap measure

5.1. Introduction

The analysis of Section 4 demonstrates that measurements made by the manual and SAPE methods tend to agree quite well with each other. However, this analysis fails to discriminate between the methods, i.e. say which method is ‘better’. Of course, to make such a statement, one needs to have a definition of what constitutes being ‘better’. In evaluating measurement methods, it is usual practice to compare a proposed method against some sort of ‘ground truth’ or ‘gold standard’. Although a ‘gold standard’ method of measurement does not exist in our setting, a ‘ground truth’ does indeed exist: they are the grain boundaries of the input image. In this context, one can define a measurement method to be ‘better’ than another if the ellipse generated by these measurements approximates the grain boundary more closely.

The non-overlap statistic is a measure of closeness of two regions. The non-overlap statistic is defined as the area of ‘non-overlap’ between two regions, normalised by the area of the second region. Mathematically, a region can be defined as a *point set*, i.e. a region is defined by the set of points that belongs to it. Using the language of set operations, the non-overlap of two regions A and B , denoted by $nol(A,B)$, can be defined as a ratio of two areas:

$$nol(A,B) = \frac{\text{Area}(A \setminus B) + \text{Area}(B \setminus A)}{\text{Area}(B)}$$

The numerator of this expression is the area of the symmetric difference of regions A and B , i.e. the set of all points that belong to A but not to B and vice versa. Two regions are identical *if and only if* their non-overlap is 0. In Fig. 9a and b, the non-overlap region between the circle (A) and the square (B) is shown in black. The denominator of the expression is the area of B (which is the ‘reference region’) and is used as a scaling factor. It ensures that $nol(A,B)$ remains unchanged no matter what scale is used to make the measurements. In this context, note that in general $nol(A,B) \neq nol(B,A)$. The choice of a symmetric denominator, such as $\text{area}(A \cup B)$, would ensure this. However, it is unsuited to our purpose. The non-overlap measure will be used for comparing regions A and A' (e.g. circles), constructed by two different methods for the same reference region B (e.g. a square). In this context, the choice of $\text{area}(B)$ ensures that the denominator remains unchanged for both $nol(A,B)$ and $nol(A',B)$, thus ensuring comparability, whereas the choice of $\text{area}(A \cup B)$ does not. When interpreting non-overlap values, it should be noted that

non-overlap is an area based (i.e. two-dimensional) measure of goodness of fit. Consequently, typical non-overlap values will be higher than corresponding one-dimensional goodness of fit measures, such as those for curve-fitting.

It is not difficult to represent the output of both manual and SAPE data acquisition procedures as fitted ellipses. This means that the non-overlap measure can be used to quantify how well the fitted ellipse generated by a data acquisition method fits a grain boundary. However, it is not possible to treat the non-overlap value between the grain boundary and the fitted ellipse as an absolute measure of how well the data acquisition works. A poor value could be due to one of two factors: (a) poor ellipse fitting, or (b) non-elliptical shape of grain. For instance, in Fig. 9a, the reference region (B) is in the shape of a square. The fitted ellipse is the circle. The non-overlap statistic in this case computes to 0.22 (i.e. the non-overlap area is 22% of the area of the square). The number 0.22 in part represents how far a square is from an elliptical shape.

5.2. Non-overlap measure as a bench mark for ellipse fitting

Although the non-overlap measure cannot be used as an absolute measure of how well an ellipse fits across regions, it can be used as a tool for comparing different fits on the same region. For instance, in Fig. 9b, the same square as in Fig. 9a is approximated by a different circle. The non-overlap measure of the circle in Fig. 9b (with respect to the same square) turns out to be 0.80. This means that the circle in Fig. 9a is a better approximation to the square than that in Fig. 9b. It is important to note that this comparison can only be made if the same reference shape is being used in both cases.

In a similar manner, the non-overlap measure can be used to provide a relative comparison of the approximation properties of two different ellipse-fitting techniques for a given set of grains. Given a set of grains, the set of non-overlap values for the ellipses fitted to them by the manual method can be computed. Correspondingly, a second set of non-overlap values for the ellipses fitted to the same grains by the SAPE method can be computed. A statistical comparison of these two sets of non-overlap values will determine how one method performed relative to the other. Comparisons of these two sets of non-overlap values can be done in a number of ways. A graphical method would be to compare their distributions using histograms. Using the oolite sample as a reference, we can see the distribution of non-overlap values for manually fitted and SAPE fitted ellipses in Fig. 10. It is clear that the non-overlap values for SAPE fitted ellipses are in general lower than their manually fitted counterparts.

The distribution of non-overlaps can be summarised by looking at the average (median) non-overlap value for each method (see Table 2). For the oolite example, the average non-overlap for manually fitted ellipses is 0.15 and that for SAPE fitted ellipses is 0.06. As mentioned earlier, the other

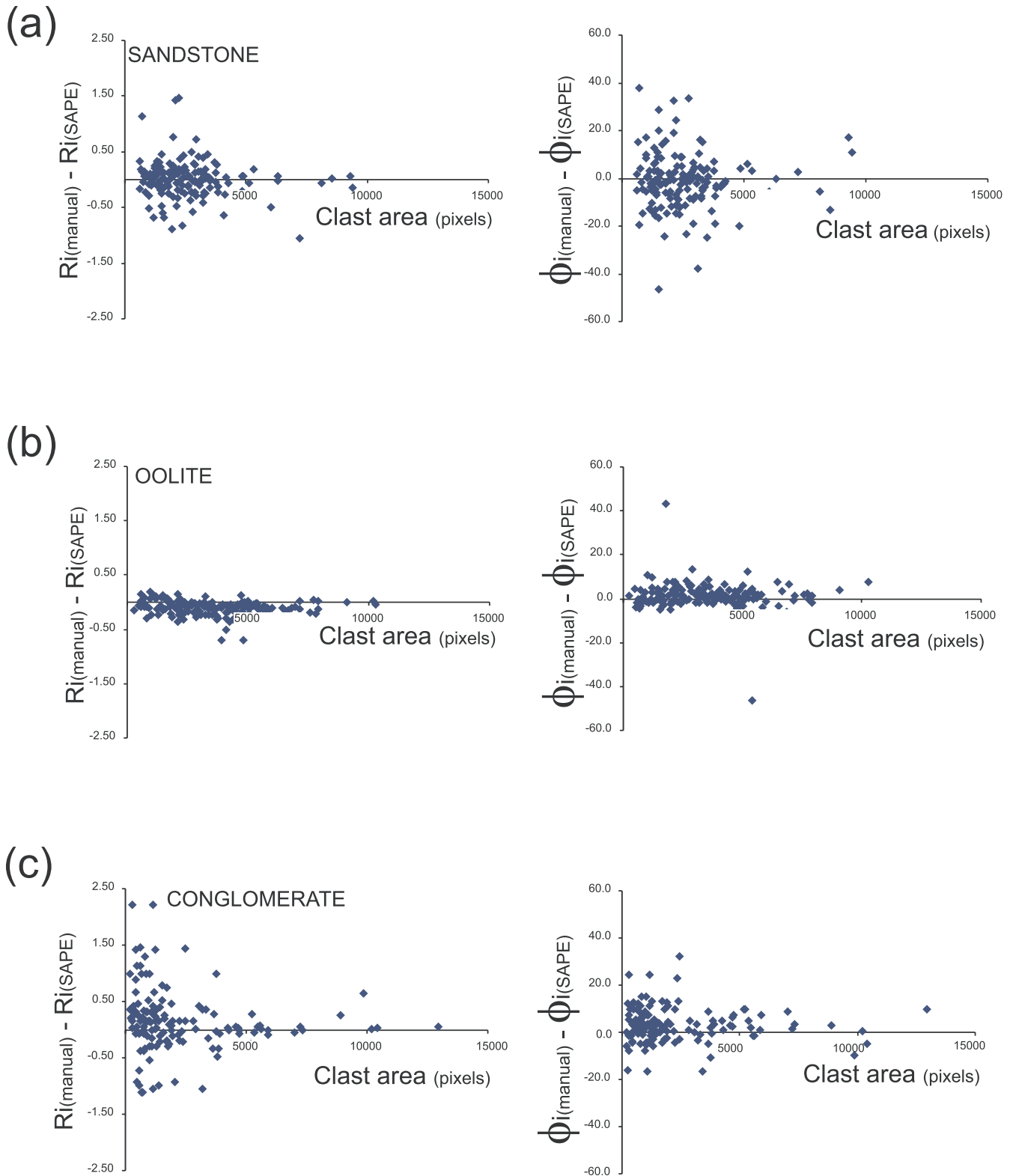


Fig. 7. Bivariate plot of estimated error in aspect ratio (i.e. $[R_i(\text{manual}) - R_i(\text{SAPE})]$) and orientation (i.e. $[\phi_i(\text{manual}) - \phi_i(\text{SAPE})]$) versus manually measured clast aspect ratio for (a) sandstone, (b) oolite, and (c) conglomerate.

aspect of non-overlap values is that they are (in part) a measure of ellipticity of the underlying grains. In this example the underlying grains are ooids, which generally

have a good elliptical shape. This is reflected in the low non-overlap values for both sets of values.

The corresponding values for the dataset (conglomerate)

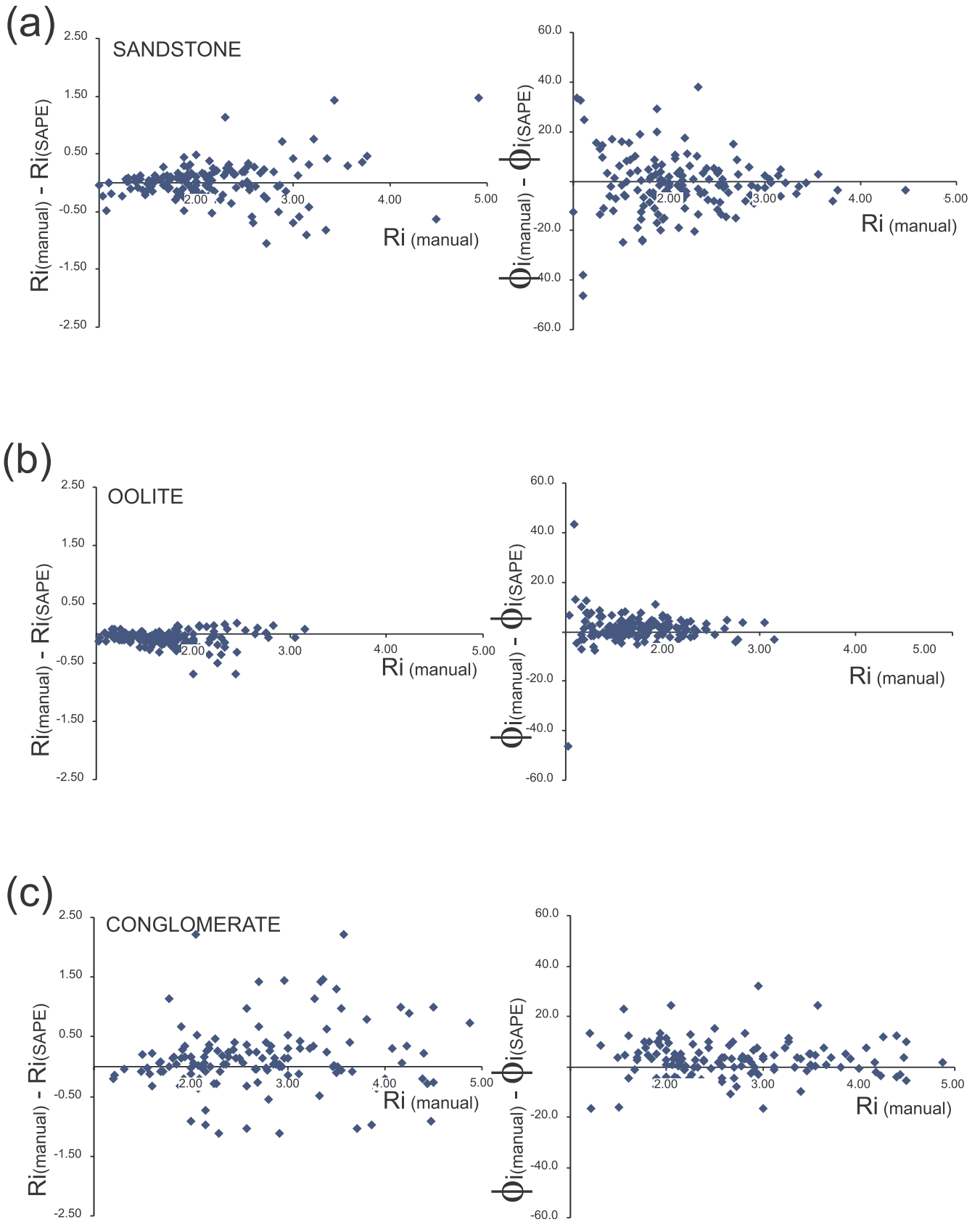


Fig. 8. Bivariate plot of estimated error in aspect ratio (i.e. $[R_i(\text{manual}) - R_i(\text{SAPE})]$) and orientation (i.e. $[\phi_i(\text{manual}) - \phi_i(\text{SAPE})]$) versus clast area in pixels for (a) sandstone, (b) oolite, and (c) conglomerate.

data set are an average of 0.24 for manually fitted ellipses and 0.18 for SAPE fitted ellipses. Again SAPE performs favourably compared with manual measurements, but in this case the magnitude of improvement is not very large. Note also that non-overlap values are larger on average for this dataset, an indication that grain shapes in this example are less elliptical than in the case of ooids. A similar set of results is obtained for the sandstone grains, although the median non-overlap is slightly worse than the conglomerates.

Summarising the performance of the two methods, one can say in terms of the non-overlap statistic that the SAPE method tends to perform at least as well as the manual method. Of course, in addition to providing modest improvement in terms of quality of fit, SAPE also provides a semi-automatic method for extracting data for strain analysis, which is far faster than manual methods.

6. Measurement error analysis

In Section 4.3 it was noted that the errors in measurements (for both SAPE and manual methods) tend to show strong relationships with size and object aspect ratio (R_i). In this section, elementary error analysis is used to provide some theoretical justification for these empirical findings. The analysis is restricted to a consideration of the errors in manual measurements for reasons of simplicity. The result for expected error in measurement derived in this

section can also be used in other applications, such as the construction of confidence intervals for strain estimates or the development of a weighted strain estimate.

The R_i measurement is effected by manually measuring the major and minor axes and then taking the ratio. As with any manual (or indeed mechanical) measurement, the measured axes are subject to measurement error. The cause of this measurement error could be manifold, from things like the accuracy of the measuring instrument, parallax, the resolution of the image to processes like dissolution, etc. Given the heterogeneous (hence unattributable) nature of this error, it is convenient to treat the magnitude of error as a random quantity. Thus the measured major axis can be expressed as $M^m = M^t + \epsilon_1$, where M^t is the true major axis and ϵ_1 is the random error of its measurement. Similarly, the measured minor axis can be expressed as $m^m = m^t + \epsilon_2$, where m^t is the true minor axis and ϵ_2 is the random error of its measurement. How these errors affect the measurement of R_i is now examined (note that R_i^m is the measured value of R_i and R_i^t is the true value of R_i).

$$R_i^m = \frac{M^m}{m^m} = \frac{M^t + \epsilon_1}{m^t + \epsilon_2} = \frac{M^t}{m^t} \left(1 + \frac{\epsilon_1}{M^t}\right) \left(1 + \frac{\epsilon_2}{m^t}\right)^{-1} \quad (11)$$

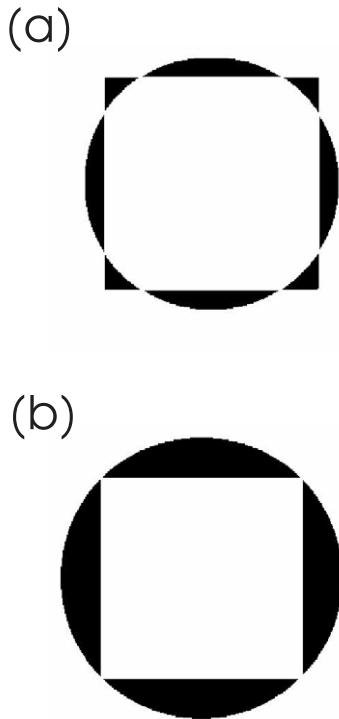


Fig. 9. (a) Region of non-overlap between a square and a circle. (b) Region of non-overlap between a square and circumscribed circle.

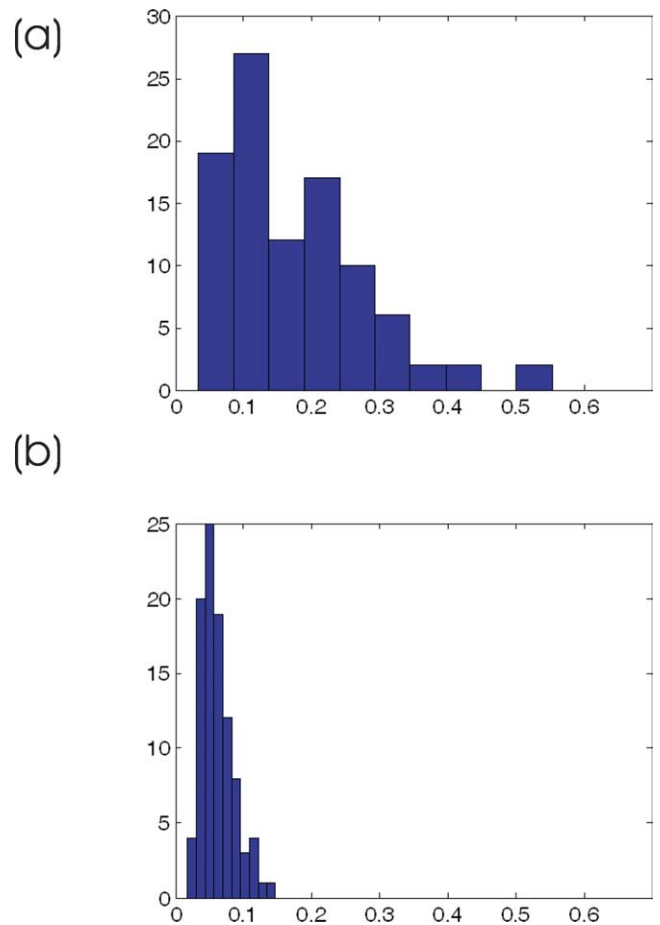


Fig. 10. Distribution of non-overlap values for (a) manually fitted ellipses and (b) SAPE ellipses for the oolite example.

Table 2
Non-overlap results for manual and SAPE processing of samples

Sample	Manual	SAPE
Sandstone	0.27	0.22
Oolite	0.15	0.06
Conglomerate	0.27	0.22

Eq. (11) can be simplified by expanding the last term as a Taylor series and ignoring higher order terms (which is justified provided the magnitude of the error ε_2 , is small compared to the true minor axis, m^t). Therefore:

$$R_i^m = \frac{M^t}{m^t} \left(1 + \frac{\varepsilon_1}{M^t} - \frac{\varepsilon_2}{m^t} \right) \quad (12)$$

The error in the measurement of the axial ratio (ε_R^m) is thus the difference between R_i^m and $R_i^t = \frac{M^t}{m^t}$, the true axial ratio. Thus using Eq. (11):

$$\varepsilon_R^m = R_i^m - R_i^t = \frac{M^t}{m^t} \left(\frac{\varepsilon_1}{M^t} - \frac{\varepsilon_2}{m^t} \right) = R_i^t \left(\frac{\varepsilon_1}{M^t} - \frac{\varepsilon_2}{m^t} \right) \quad (13)$$

Eq. (13) expresses the error in terms of random quantities. It can thus be small or large depending on the particular circumstance. However, an expression for the mean squared error (MSE) or variance ($V(\varepsilon_R^m)$) of this error can be obtained. This expression indicates how the square of the error behaves on average:

$$V(\varepsilon_R^m) = (R_i^t)^2 \left(\frac{\sigma_1^2}{(M^t)^2} + \frac{\sigma_2^2}{(m^t)^2} \right) \quad (14)$$

In Eq. (14), the quantities σ_1^2 and σ_2^2 indicate the MSE of measuring the major and minor axes, respectively. In deriving Eq. (14) it has been assumed that these errors are independent of each other. Although this is not a crucial assumption in what follows, it makes the expression simpler. Eq. (14) immediately indicates that the relationship between measurement error of R_i and its actual value is an increasing one.

To be more precise than this, the relationship between area (A_i) and the major and minor axes must be specified. Intuitively it may be asserted that A_i is roughly proportional to the magnitude of the major and minor axes. Consider first though, the simplest case of an elliptical grain, in which case:

$$A_i = \frac{\pi}{4} M^t m^t \quad (15)$$

Noting that $R_i^t = \frac{M^t}{m^t}$ then:

$$(M^t)^2 = \frac{4A_i R_i^t}{\pi} \quad (16)$$

$$(m^t)^2 = \frac{4A_i}{\pi R_i^t} \quad (17)$$

Substituting into Eq. (14):

$$\begin{aligned} V(\varepsilon_R^m) &= (R_i^t)^2 \left(\frac{\pi \sigma_1^2}{4A_i R_i^t} + \frac{\pi R_i^t \sigma_2^2}{4A_i} \right) \\ &= \frac{\pi}{4A_i} (R_i^t \sigma_1^2 + (R_i^t)^3 \sigma_2^2) \end{aligned} \quad (18)$$

Eq. (18) clearly gives the relationship between the MSE of measuring the axial ratio R_i by the manual method and the size and axial ratio of the clast, in the case of an elliptical clast. In detail, the MSE is inversely proportional to the size of the clast, but directly proportional to the cube of the axial ratio. This is in broad agreement with the empirical results presented earlier as well as those of McNaught (2002) and Meere and Mulchrone (2003). Extending this result to include other more general shapes is analytically difficult and has not yet been attempted. However, it is conjectured that a similar relationship would be found.

Although the findings of Eq. (18) are a broad confirmation of the results of Section 4, it should be noted that Eq. (18) is not immediately applicable to the errors discussed in Section 4. In particular, Section 4 looks at the difference between two measurement methods. Of these, only manual measurement has been considered in the theoretical analysis above. Moreover, strictly speaking, Eq. (18) only holds for objects that are perfectly elliptical, a condition rarely encountered in nature.

7. Discussion and conclusions

SAPE is a program for extracting information relevant to strain analysis from input digital images that consist of white regions enclosed by a black outline on a white background. Digital images of natural rocks or microphotographs never conform to this specification and the creation of the input image therefore involves either a complex computer program or human input. Writing a computer program that can handle the complexities of identifying regions of interest in any given material is not practical right now. Therefore, the logic of SAPE is to allow the human user to perform this computationally onerous task. Once the input image is prepared, SAPE rapidly extracts the required data. As such, SAPE is an intermediate step along the way to the ideal solution.

SAPE works by using a simple region-growing algorithm to identify regions of interest from the input image. Subsequently, the second moments are calculated for each region enabling the common strain analysis parameters (i.e. R_i and ϕ_i) to be readily computed (Mulchrone and Roy Choudhury, 2004). The program was used to extract information from microphotographs of deformed sandstone and oolite as well as a meter scale photograph of deformed conglomerate. SAPE compared well with a manual analysis of the same samples. The non-overlap statistic was introduced as a measure by which the performance of two

different methods for data extraction could be judged. In all three examples SAPE performed at least as well as the manual method.

Previously, manual recording of strain data would entail for each object (1) identification of a suitable object, (2) identification of long and short axes (a totally subjective procedure), (3) measurement of length and orientation of long axis and length of short axis, (4) recording of results to paper (optional), and (5) transfer of results to computer. SAPE allows for rapid extraction of data from images because, from a human perspective, the above five steps are reduced to one step, i.e. trace the outlines of the objects of interest. In our experience, the time required to manually measure 150 objects is of the order of 1.5 h, whereas tracing outlines of the same takes only 15 min. Additionally, tracing is found to be less tedious and error prone and the task could be performed for long periods of time without tiredness. Combining the time advantage with the objectivity, robustness and ease of the approach, it is felt that use of SAPE will allow for easier capture of larger datasets and therefore more reliable strain data in structural studies.

Error analysis of measurement methods is important for a variety of reasons. Apart from validating the accuracy of the method, knowledge of the distribution of error and how it behaves with respect to other variables, such as clast size and shape, can be useful in strain analysis calculations. In particular, it can be used for the calculation of error bounds for the strain estimate and may also be incorporated into strategies such as weighted strain estimation.

Empirical analysis of the discrepancy between SAPE and manual measurements done in Section 4 demonstrates certain trends in relation to clast size and aspect ratio. In particular, we observed that the discrepancy decreased with increasing clast size while it increased with aspect ratios. However, the degree of relationship between these variables appears to differ between lithologies, with conglomerate displaying the most marked trends and oolites the least marked. A theoretical analysis of errors for the manual measurement method also confirmed these relationships between the measurement error versus clast size and aspect ratio. However, it was noted that the theoretical analysis was only valid under a restricted setting. It is hoped that subsequent analysis, under less restrictive conditions, will demonstrate how errors behave under a variety of settings. It is also hoped that future work will investigate errors in the SAPE measurement method in more detail.

Acknowledgements

Constructive, detailed and critical reviews by W.H. Owens and Dr M. Mukul helped to greatly improve the final draft of this manuscript. This work forms part of a project funded by Enterprise Ireland (SC/2000/104/Y).

References

- Adams, R., Bischof, L., 1994. Seeded region growing. *IEEE Transactions on Pattern Analysis Machine Intelligence* 16, 641–647.
- Ailleres, L., Champenois, M., Macaudiere, J., Bertrand, J.M., 1995. Use of image analysis in the measurement of finite strain by the normalized Fry method: geological implications for the 'Zone Houillère' (Briançonnais zone, French Alps). *Mineralogical Magazine* 59, 179–187.
- Borradaile, G.J., 1976. A study of a granite/granite-gneiss transition and accompanying schistosity formation in SE Spain. *Journal of the Geological Society, London* 132, 417–428.
- Dunnet, D., 1969. A technique of finite strain analysis using elliptical particles. *Tectonophysics* 7, 117–136.
- Dunnet, D., Siddans, A.W.B., 1971. Non-random sedimentary fabrics and their modification by strain. *Tectonophysics* 12, 307–325.
- Elliott, D., 1970. Determination of finite strain and initial shape from deformed elliptical objects. *Geological Society of America Bulletin* 81, 2221–2236.
- Erslev, E.A., 1988. Normalized center-to-center strain analysis of packed aggregates. *Journal of Structural Geology* 10, 201–209.
- Erslev, E.A., Ge, H., 1990. Least squares center-to-center and mean object ellipse fabric analysis. *Journal of Structural Geology* 8, 1047–1059.
- Fry, N., 1979. Random point distributions and strain measurement in rocks. *Tectonophysics* 60, 806–807.
- Heilbronner, R., 2000. Automatic grain boundary detection and grain size analysis using polarization micrographs or orientation images. *Journal of Structural Geology* 22, 969–981.
- Lisle, R.J., 1977a. Estimation of the tectonic strain ratio from the mean shape of deformed elliptical markers. *Geologie en Mijnbouw* 56, 140–144.
- Lisle, R.J., 1977b. Clastic grain shape and orientation in relation to cleavage from the Aberystwyth Grits, Wales. *Tectonophysics* 39, 381–395.
- Lisle, R.J., 1985. *Geological Strain Analysis: a Manual for the R_f/ϕ Technique*. Pergamon Press, Oxford.
- Masuda, T., Koike, T., Yuko, T., Morikawa, T., 1991. Discontinuous growth of quartz in metacherts: the influence of mica on a microstructural transition. *Journal of Structural Geology* 9, 389–402.
- Matthews, P.E., Bond, R.A.B., Van Den Berg, J.J., 1974. An algebraic method of strain analysis using elliptical markers. *Tectonophysics* 24, 31–67.
- McNaught, M., 1994. Modifying the normalized Fry method for aggregates of non-elliptical grains. *Journal of Structural Geology* 16, 493–503.
- McNaught, M., 2002. Estimating uncertainty in normalized Fry plots using a bootstrap approach. *Journal of Structural Geology* 24, 311–322.
- Meere, P.A., Mulchrone, K.F., 2003. The effect of sample size on geological strain estimation from passively deformed clastic sedimentary rocks. *Journal of Structural Geology* 25, 1587–1595.
- Mukul, M., 1998. A spatial statistics approach to the quantification of finite strain variation in penetratively deformed thrust sheets: an example from the Sheeprock Thrust Sheet, Sevier Fold-and-Thrust belt, Utah. *Journal of Structural Geology* 20, 371–384.
- Mukul, M., Mitra, G., 1998. Finite strain and strain variation analysis in the Sheeprock Thrust Sheet: an internal thrust sheet in the Provo salient of the Sevier Fold-and-Thrust belt, Central Utah. *Journal of Structural Geology* 20, 385–405.
- Mulchrone, K.F., 2003. Application of Delaunay triangulation to the nearest neighbour method of strain analysis. *Journal of Structural Geology* 25, 689–702.
- Mulchrone, K.F., Meere, P.A., 2001. A Windows program for the analysis of tectonic strain using deformed elliptical markers. *Computers and Geosciences* 27, 1253–1257.
- Mulchrone, K.F., Roy Choudhury, K., 2004. Fitting an ellipse to an arbitrary shape: implications for strain analysis. *Journal of Structural Geology* 26, 143–153.

- Mulchrone, K.F., O'Sullivan, F., Meere, P.A., 2003. Finite strain estimation using the mean radial length of elliptical objects with bootstrap confidence intervals. *Journal of Structural Geology* 25, 529–539.
- Panozzo, R., 1984. Two-dimensional strain from the orientation of lines in a plane. *Journal of Structural Geology* 6, 215–221.
- Panozzo Heilbronner, R., Pauli, C., 1993. Integrated spatial and orientation analysis of quartz c-axes by computer-aided microscopy. *Journal of Structural Geology* 15, 369–382.
- Peach, C.J., Lisle, R.J., 1979. A Fortran IV program for the analysis of tectonic strain using deformed elliptical markers. *Computers and Geosciences* 5, 325–334.
- Ramsay, J.G., 1967. *Folding and Fracturing of Rocks*. McGraw-Hill, New York.
- Ramsay, J.G., Huber, M.I., 1983. *The Techniques of Modern Structural Geology*, Volume 1. Strain Analysis. Academic Press, London.
- Robin, P.F., 1977. Determination of geologic strain using randomly oriented strain markers of any shape. *Tectonophysics* 42, T7–T16.
- Seul, M., O'Gorman, L., Sammon, M.J., 2000. *Practical Algorithms for Image Analysis: Description, Examples and Code*. Cambridge University Press, Cambridge, UK.
- Shimamoto, T., Ikeda, Y., 1976. A simple algebraic method for strain estimation from deformed ellipsoidal objects. 1. Basic theory. *Tectonophysics* 36, 315–337.
- Yu, H., Zheng, Y., 1984. A statistical analysis applied to the R_p/ϕ method. *Tectonophysics* 110 (151), 155.