

## Stability of Kerogen Classification with Regard to Image Segmentation

J.J. Charles · L.I. Kuncheva · B. Wells · I.S. Lim

Received: 25 June 2007 / Accepted: 11 February 2009 / Published online: 27 March 2009  
© International Association for Mathematical Geosciences 2009

**Abstract** This paper investigates the stability of an automatic system for classifying kerogen material from images of sieved rock samples. The system comprises four stages: image acquisition, background removal, segmentation, and classification of the segmented kerogen pieces as either inertinite or vitrinite. Depending upon a segmentation parameter  $d$ , called “overlap”, touching pieces of kerogen may be split differently. The aim of this study is to establish how robust the classification result is to variations of the segmentation parameter. There are two issues that pose difficulties in carrying out an experiment. First, even a trained professional may be uncertain when distinguishing between isolated pieces of inertinite and vitrinite, extracted from transmitted-light microscope images. Second, because manual labelling of large amount of data for training the system is an arduous task, we acquired the true labels (ground truth) only for the pieces obtained at overlap  $d = 0.5$ . To construct ground truth for various values of  $d$  we propose here label-inheritance trees. With thus estimated ground truth, an experiment was carried out to evaluate the robustness of the system to changes in the segmentation through varying the overlap value  $d$ . The average system accuracy across values of  $d$  spanning the range from 0 to 1 was 86.5%, which is only slightly lower than the accuracy of the system at the design value of  $d = 0.5$  (89.07%).

**Keywords** Classification · Microfossils · Image analysis · Segmentation · Palynomorph

---

J.J. Charles (✉) · L.I. Kuncheva · I.S. Lim  
School of Computer Science, Bangor University, Bangor LL57 1UT, UK  
e-mail: [j.j.charles@bangor.ac.uk](mailto:j.j.charles@bangor.ac.uk)

B. Wells  
Conwy Valley Systems Ltd, Conwy, UK

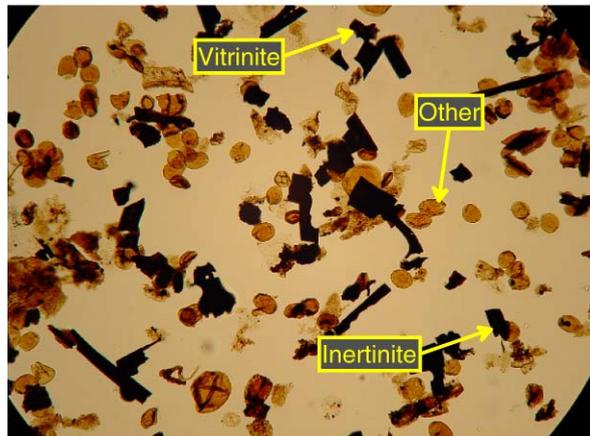
## 1 Introduction

Oil exploration relies on accurate analysis of fossilised material in sieved rock samples. Kerogen is a solid, insoluble organic matter, typically composed of algae and woody plant material, that can yield oil upon heating. Classification of kerogen into subtypes and measuring the volumes thereof can give important clues about the presence of oil. Figure 1 shows a typical image of a dispersed kerogen preparation slide containing microfossils and other organic debris.

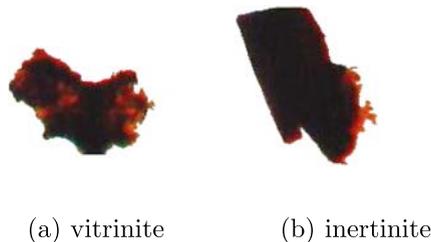
Automated classification of fossilised material has been a topic of interest for decades (Hills 1988; Swaby 1992; Athersuch et al. 1994; Bollmann et al. 2004; Weller et al. 2005). A common theme in all these works is the imperative need for automatic extraction of the objects and their features from the original image. It is sometimes difficult even for the trained eye to determine whether a cluster of dark-coloured forms contains several objects (touching or overlapping) or represents a single piece of kerogen. Thus the extraction of the individual kerogen pieces from an image is a challenge itself. To add to the complexity facing an automated classification system, distinguishing between inertinite and vitrinite is not straightforward. Zoomed examples of typical “vitrinite” and “inertinite” pieces indicated in Fig. 1 are shown in Fig. 2. The differences are subtle, which implies that a combination of features based on colour, size, shape and texture should be used.

Segmentation of the image into adequate kerogen pieces is a prerequisite for accurate classification. If the image is under-segmented, inertinite and vitrinite may

**Fig. 1** Dispersed kerogen preparation slide containing kerogen (dark pieces) with indicated examples of inertinite and vitrinite. Any other microfossil not in this category is classed as ‘other’



**Fig. 2** Zoomed images of typical examples of **a** vitrinite and **b** inertinite as indicated in Fig. 1



be lumped together in a single object. On the other hand, over-segmentation may produce a multitude of small objects that may be too small to be identified. The classification accuracy of the system is likely to be adversely affected by any variability of the segmentation. Our system uses the Centre Supported Segmentation method (CSS) (Charles et al. 2008a). The resolution of the segmentation is controlled by a parameter  $d$ , called overlap and spanning the interval from 0 (no overlap allowed) to 1 (complete overlap allowed). Hence, this paper examines the effect of changes in  $d$  on the classification accuracy.

## 2 System for Kerogen Recognition

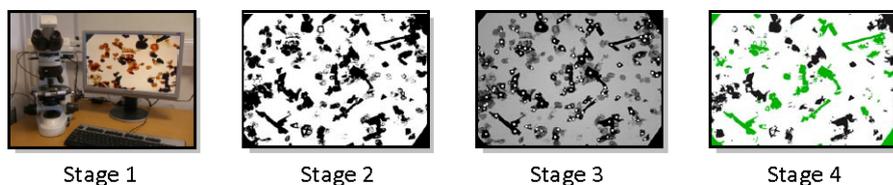
Unlike previous systems for processing palynological images (e.g. Weller et al. 2005), our system is completely automatic. This is why it was deemed important to ensure that the system is sufficiently robust to variations in parameter choices. Fig. 3 displays the four stages of the system: image acquisition, background segmentation, object extraction, and classification into inertinite and vitrinite.

### 2.1 Image Acquisition

High resolution images (approximately 4 megapixels) of dispersed kerogen preparation slides are captured using a digital camera attached to a microscope. The microscope set-up consists of an automatic stepping platform that can be attached to any microscope and is controlled through Petrog software (Wells 2008). A digital FireWire camera captures the image and a control box transfers it to the computer (Stage 1 in Fig. 3). Due to hardware restrictions, images were stored in jpeg format even though it does not guarantee lossless compression. The accuracy of segmentation and classification of kerogen is unlikely to be affected by using jpeg for the following reasons. The compression may cause minor changes to object boundaries but the fine detail in the boundaries of the kerogen pieces is not vital for classification and is mostly ignored by the features we have chosen.

### 2.2 Background Removal

As a pre-processing step the image is converted to greyscale and further segmented into a binary image. The foreground contains the objects of interest (vitrinite or inertinite) which appear in darker colour, while the background contains everything else



**Fig. 3** The four stages of the system for automatic recognition of kerogen in microscope images: image acquisition, background segmentation, object extraction, and classification into inertinite and vitrinite

(non-kerogen fossils, amorphous matter, and blank space). Segmentation by simple thresholding is not suitable because, in transmitted microscopy, the overall light intensity diminishes away from the image centre. A global threshold might wrongly pick the shadowy edges of the image as foreground. A correction of the non-uniform lighting will allow for such a global thresholding to be successful. In a previous study we proposed a method for illumination correction using a collection of quadratic approximations on horizontal and vertical strips of the image (Charles et al. 2008b). The image is first split into horizontal and vertical strips. The background within each strip is approximated through a parabola. The image is ‘normalised’ using the background approximation. A threshold is found next from the grey image histogram. The histogram is typically bimodal; the left peak corresponding to dark-coloured objects (kerogen) and the right peak corresponding to background. After applying the threshold, the black-and-white image is fed to the segmentation algorithm. We note that both the background approximation and the thresholding are done automatically.

### 2.3 Centre Supported Segmentation

The CSS algorithm is applied to the black and white image after removing the background from the slide (Charles et al. 2008a). The images and the Matlab code are available on the IAMG server <http://www.iamg.org/CGEditor/index.htm>. The output of CSS is a set  $C$  that contains the locations of all segmented pieces. A complete description of the CSS algorithm can be found in Charles et al. (2008a). We reproduce part of the algorithm here because it is needed for understanding the label-inheritance trees introduced later. Using an iterative approach the CSS algorithm applies thresholds across  $D(p)$  to build up a set of all possible centres  $C$ . A filter is then applied to  $C$  to determine the most likely centres. Using these centres the corresponding individual objects can be extracted. CSS belongs in the class of marker-controlled segmentation algorithms. The watershed algorithm is applied to the negative of the distance image  $D(p)_{\text{neg}} = -D(p)$  modified so that only the regional minima occur at each centre. This technique is known as image imposition (Soille 2003). Let  $D(p)$  be the Euclidean distance from pixel  $p$  to the nearest white pixel in the black and white image (the distance function). We define the centre of an object as the pixel  $p$  with the largest distance  $D(p)$  within the object. For example, the centre of a filled-in circle will be its geometrical centre. However, a doughnut-shaped object will have infinitely many centres, none of which will be its geometrical centre.

Two lists are initialised and subsequently expanded: list  $C$  of centres and list  $V$  of their merging heights.

- Step 1. Initialise  $C = \text{empty set}$ . Find the distance image  $D$  of the black and white image  $B$ .
- Step 2. Sort all distinct values of  $D$  in descending order  $m_1 > m_2 > \dots > m_T$ .
- Step 3. Find the binary image  $B_1$  by thresholding  $D$  at  $m_1$ . This image will contain one black point  $q_1$ ; add this point to  $C$ .
- Step 4. The binary image  $B_2$  is created by thresholding  $D$  at  $m_2$ . This image will be a superset of  $B_1$  and will include points around pixel  $q_1$ . Since the object with centre  $q_1$  is already accounted for we remove this connected component from  $B_2$ .

Step 5. A single pixel from each connected component left in  $B_2$  is chosen arbitrarily and added to  $C$ . Note that these connected components will be single points or clusters of points at the same height, thus any point in this cluster can serve as a centre of the component.

Step 6. The process of thresholding, removing connected components and finding new centres is repeated until the threshold is equal to  $m_T$ . Over-segmentation would result if we were to use all the centres in  $C$  at this point, hence the next step is to apply a filter.

Step 7. A merging height  $v_i$  is attached to each centre  $q_i$ . The merging height of a centre  $q_i$  is the smallest  $m_k$  where  $q_i$  is in a connected component of image  $B_k$  disjoint from any connected components containing the centre  $q_j$  such that  $D(q_j) > D(q_i)$ . Figuratively speaking, this is when an object of smaller size (smaller peak  $D(q_i)$ ) is eclipsed by an object of bigger size ( $D(q_j) > D(q_i)$ ).

A threshold  $s$  is applied to account for the minimum allowable size of an object. All objects with centres  $q$ , such that  $D(q) < s$ , are discarded. If the algorithm is run with  $s = 0$ , it will find all the specks in the image and declare them objects of interest. The value of  $s$  can be estimated subjectively or can be learned from a sample of training images where the objects of interest have been pre-labelled by hand. Weller et al. (2005) propose an empirical threshold of  $14 \mu\text{m}$ . However, applying this threshold alone will only remove the centres of objects considered to be dust specks; we still require a filter to remove centres that over-segment an object. This is done by determining the degree of overlap assigned to each centre. The degree of overlap for centre  $q_i$  is

$$O(q_i) = 1 - \sqrt{1 - (v_i/D(q_i))^2}.$$

By definition,  $D(q_i) > v_i$ , so  $O(q_i) \in [0, 1)$ . A centre  $q_i$  with overlap  $O(q_i) = 0$  means that the object is isolated. As the overlap approaches 1, the object is increasingly covered by a larger item. We can set an upper limit  $d$  on the amount of overlap a centre can have in order for it to remain in the set  $C$ .

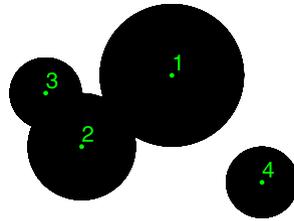
During the CSS algorithm a decreasing sequence of thresholds  $m_i$  are applied to the distance function  $D$ . Threshold  $m_i$  generates a black and white image  $B_i$ . Let  $K_i$  be the set of connected components in  $B_i$ .

**Definition 1** We say that two centres  $p$  and  $q$  merge if there exists a pair of consecutive thresholds  $m_{i-1}$  and  $m_i$  such that

- (1) At threshold  $m_{i-1}$ ,  $p$  and  $q$  are contained within two different connected components in  $K_{i-1}$ .
- (2)  $D(p)$  is the largest distance function value within the connected component in  $K_{i-1}$  containing  $p$ .
- (3)  $D(q)$  is the largest distance function value within the connected component in  $K_{i-1}$  containing  $q$ .
- (4) At threshold  $m_i$ ,  $p$  and  $q$  appear in the same connected component of  $K_i$ .

To illustrate the concept of overlap, consider the image in Fig. 4a. Applying CSS algorithm to the image, centre 1 will be stored first, and will grow a connected component until centre 2 appears. Next centres 3 and 4 will appear as the third and fourth

**Fig. 4** **a** An artificial image with 4 circular objects; **b** label-inheritance trees for the two connected components in **a**



**Fig. 5** Effect of the allowable overlap  $d$  on the image segmentation. Best segmentation (3 centres) is obtained for  $d$  between 0.3 and 0.5

No. Objects = 2  $d = 0.1$	No. Objects = 3  $d = 0.3$	No. Objects = 3  $d = 0.5$
No. Objects = 4  $d = 0.7$	No. Objects = 35  $d = 0.9$	No. Objects = 56  $d = 1$

connected components. Further lowering the distance function threshold  $m_i$ , the connected components of centres 2 and 3 will join into a single connected component, meaning that centres 2 and 3 merge (Definition 1). At that instant centre 3 will receive its degree of overlap,  $d^{(3)} = 0.4453$ . Finally, the connected component containing centres 2 and 3 will join that of centre 1, making centres 1 and 2 merge, and giving centre 2 a degree of overlap  $d^{(2)} = 0.1150$ . Note that centre 3 does not merge with centre 1 because it violates one of conditions 2 or 3 of the definition. For centres 1 and 4,  $d^{(1)} = 0$  and  $d^{(4)} = 0$ . Thus if the overlap threshold is set at 0.3, the image will be segmented into three objects so that objects 2 and 3 will be taken as one.

As a further illustration, a cropped segment of three overlapping pieces of vitrinite and inertinite are shown in Fig. 5. The CSS algorithm is applied and various values of  $d$  are used to remove redundant centres. For automatic segmentation the value of  $d$  must be fixed. In previous studies (Charles et al. 2008a), it was found that setting  $d = 0.5$  produces results that outperform commonly used segmentation methods such as watershed (Vincent and Soille 1991) and extended  $h$ -maxima (Soille 2003). Using these centres found through CSS, the corresponding individual objects are extracted next. CSS belongs to the class of marker-controlled segmentation algorithms. The watershed algorithm is applied to the negative of the distance image  $D(p)_{\text{neg}} = -D(p)$  modified so that only the regional minima occur at each centre. This technique is known as image imposition (Soille 2003). As a result, an object is segmented around each centre. These objects are further labelled as inertinite or vitrinite.

## 2.4 Classification

After segmentation, each piece of kerogen is labelled as either inertinite or vitrinite using a pre-trained classifier. First, 32 features are extracted from each object accounting for colour, shape, size and texture, given in Table 1 with a short comment. Our previous study showed that the logistic classifier suits this type of data well (Kuncheva et al. 2008). The classification accuracy of the system will depend

**Table 1** Groups of features

Type	Feature	Notation	Explanation
Colour	mean red		
	mean blue		
	mean green		
	mean grey		
Size	inner radius	$r_i$	Radius of the largest circle contained entirely within the object
	outer radius	$r_o$	Radius of the smallest circle which contains the entire object
	diameter	$d$	The maximum distance between 2 contour pixels
	perimeter	$p$	Number of pixels on the border between object and background
	circle difference	$r_o - r_i$	Difference between the outer and inner radii
	area	$a$	Total number of pixels comprising the object
	distance	$\delta$	Mean distance from centre of gravity to all contour pixels
Texture	entropy		Entropy of the grey-level histogram taken as a pdf
	anisotropy		Symmetry of the grey-level histogram about its median
	correlation		Correlation between grey level intensity of neighbouring pixels
	homogeneity		Homogeneity of neighbouring pixels in the grey level image
	contrast		Contrast of neighbouring pixels in the grey level image
	energy		Energy of neighbouring pixels in the grey level image
	rim variability		Variance of the grey level intensity in a “rim” of width $r_i/5$
Shape	anisometry	$e_+/e_-$	Ratio of the lengths of the major and minor elliptic semi-axes
	eccentricity	$d_-/d$	Ratio of the length of the minor axis of the object to $d$
	rectangularity	$a/a_b$	Ratio of object area to the area of smallest bounding rectangle
	bulkiness	$\frac{\pi(e_+)(e_-)}{a}$	Ratio of the areas of a corresponding ellipse and the object
	convexity	$a/a_c$	Ratio of the object’s area to its convex area
	variance $x$		Variance across $x$ -axis with respect to centre of gravity
	variance $y$		Variance across $y$ -axis with respect to centre of gravity
	covariance		
	compactness	$4\pi a/p^2$	Ratio of the area to that of a circle with the same perimeter
	sigma	$\sigma$	Standard deviation of distances from centre of gravity to contour
	roundness	$1 - \sigma/\delta$	
	sides	$1.41(\frac{\delta}{\sigma})^{0.4724}$	Number of pieces of a regular polygon
	equant/lath	$r_i/d$	Equant/lath ratio
structure factor	$\frac{\pi(e_+)^2}{a} - 1$	anisometry $\times$ bulkiness $- 1$	

on the quality of the segmentation at stage 3. Because this is governed by the overlap threshold  $d$ , this study investigates the effect on variations in  $d$  on the classification accuracy.

### 3 Label Inheritance Trees

Varying the overlap  $d$  will change the segmentation; the objects may get broken down into smaller objects (higher  $d$ ) or merged (lower  $d$ ). To assess the classification accuracy at various  $d$ , all the objects (broken down or merged) must have labels that are to be considered as ground truth. The bottom right subplot in Fig. 5 suggests that the number of small objects might be prohibitive for high  $d$ , hence seeking expert opinion on the labels will become impractical. Therefore, the labels for the objects extracted at  $d \neq 0.5$  have to be estimated from the labels at  $d = 0.5$  that were available to us. In order to determine which larger object breaks into which smaller pieces, a tree is constructed for each centre that has degree of overlap zero. The purpose of the tree is to define how the class labels of the pieces are inherited for different values of the overlap  $d$ . The intuition behind the label-inheritance trees is that a piece with class label  $X$  should break down into pieces of the same class. On the other hand, when pieces of different classes happen to be merged, the largest class label is taken as the class of the new piece. For the example in Fig. 4, there will be two trees, one for centre 1 and one for centre 4. First, all centres are sorted with respect to their degree of overlap  $d$ . The sequence of sorted overlap values are denoted by  $d_0, d_1, \dots, d_K$ . The centres with  $d_0 = 0$  are taken to be roots of trees. All trees will be built to the same height,  $d_K = \max_i d_i$ . Let  $C_i$  be the set of centres at height  $d_i$ , also serving as nodes of the trees. We note that  $C_i \subseteq C_{i+1} \subseteq C$ , where  $C$  is the set of all centres (found before applying the overlap threshold).

**Definition 2** Node  $p \in C_{i-1}$  is a parent of node  $q \in C_i$  if either

- (1)  $p = q$ .
- (2)  $p$  and  $q$  merge and  $q \notin C_{i-1}$ .

Figure 4b shows the two trees for the images in Fig. 4a. These trees are called label-inheritance trees because the labels of broken up connected components are evaluated on their basis. At threshold  $d = 0$ , there are only two centres representing the two connected components. By construction (CSS algorithm), these centres are 1 and 4. Next is centre 2 at  $d_1 = 0.1150$ . As centre 1 is a parent of centre 2 according to Definition 2, they are joined at  $d_1$ . The number of objects will not change when  $d$  is increased until the next value,  $d_2 = 0.4453$ , is reached. At this value, centres 2 and 3 merge, and centre 2 is a parent of centre 3. For overlap threshold  $0.4453 \leq d \leq 1$ , four distinct objects will be detected. The same algorithm was used to build the trees of all images for the experiment. The range for  $d$  was discretized into  $d_0 = 0, d_1 = 0.1, \dots, d_{10} = 1.0$ , and the respective subsets of centres,  $C_i$ , were determined. All objects in our experiment were extracted at  $d = 0.5$ . These objects were labelled by the expert palynologist. To carry out a classification experiment for different values of  $d$  we require a mapping  $f_i : C_i \rightarrow L$ , where  $L$  is the set of possible class labels.

The only available mapping is  $f_k$ , at layer  $k$  with height  $d = 0.5$ . We estimate the mapping  $f_j$  for all nodes on layer  $C_j$  as follows:

if  $j > k$  Let  $q \in C_k$  be a node connected to  $p \in C_j$ . Then  $f_j(p) = f_k(q)$ .

if  $j < k$  We form a set  $Q = \{q \in C_k \mid q \text{ is connected to } p \in C_j\}$ . The nodes in  $Q$  represent objects. We calculate the areas of these objects and attach them to the corresponding nodes in  $Q$ . Then  $f_j(p) = \text{weighted vote of set } \{f_k(q) \mid q \in Q\}$ , where the areas of the objects are used as weights.

For example, suppose that an object labelled as vitrinite at  $d = 0.5$  is split up into smaller pieces by increasing the overlap threshold. According to the adopted procedure, each small piece will be also labelled as vitrinite. On the other hand, consider two objects: a large piece of inertinite and a small piece of vitrinite with a high overlap degree. At some value of the overlap parameter  $d < 0.5$ , these two objects will be merged into one. Since the inertinite piece occupies more area than the vitrinite piece, the weighted vote will label the combined object as inertinite.

## 4 The Stability Experiment

The purpose of this experiment is to study the accuracy of a classifier with regard to changes in the segmentation parameter  $d$ . A classifier will take as input the features of an object and output the suggested class label. In our case we will use a 2-class classifier for the vitrinite and inertinite pieces.

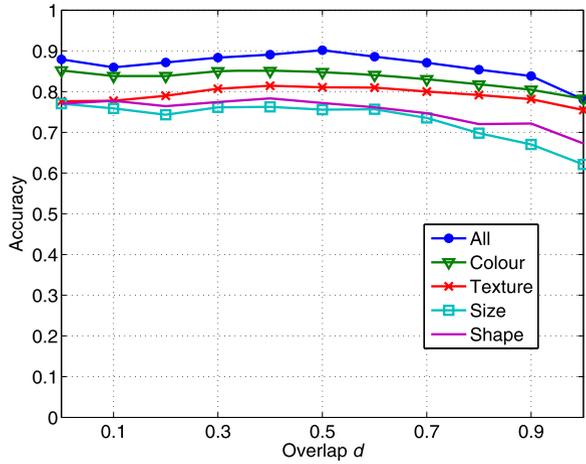
### 4.1 Classification at $d = 0.5$

The training set of 609 objects extracted from 7 images using CSS with overlap  $d = 0.5$ , was subsequently labelled by an expert palynologist. The eight objects that were labelled by the expert as “other” were excluded from the experiments, leaving a 2-class data set of 601 objects. In a previous study we compared 10 state-of-the-art classifiers using Weka (Witten and Frank 2005), with 10 times 10-fold cross-validation. The classifiers used were Naive Bayes, Decision tree, Logistic, Nearest Neighbour, Multilayer Perceptron (MLP), Support Vector Machine (SVM), AdaBoost, Bagging, LogitBoost, and Random Forest. The logistic classifier produced the highest accuracy (Kuncheva et al. 2008).

### 4.2 Classification Stability

The CSS algorithm was applied to the seven microscope slides at 11 overlap values  $d = 0, 0.2, \dots, 1$ . Using the method described above, labellings of the extracted objects were inferred from those at  $d = 0.5$ . A dataset for each overlap value was formed consisting of inertinite and vitrinite. The logistic classifier fared better than the other 9 classifiers and so was trained on all objects extracted at  $d = 0.5$ . This classifier was then tested on the 11 new datasets produced at different  $d$ . The accuracy of the classifier was measured against the labels assigned by the label-inheritance trees as a ground truth. Objects with other ground truth were counted as an error in the

**Fig. 6** Stability of the logistic classifier on various feature sets



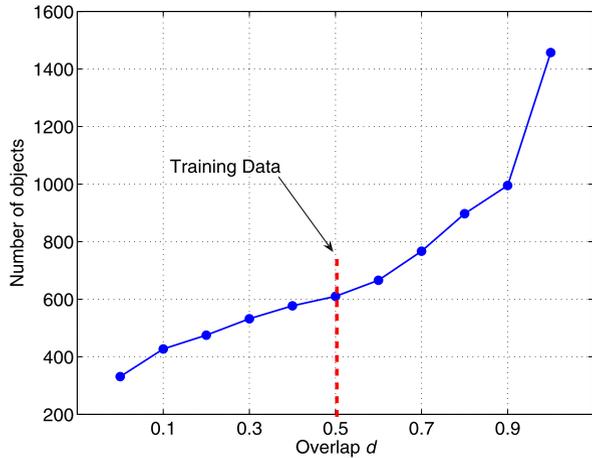
**Table 2** Mean and standard deviation of the classification accuracy

Feature set	Mean	Standard deviation
All	86.5%	3.33%
Colour	83.2%	2.22%
Texture	79.2%	1.85%
Size	73.0%	4.75%
Shape	75.13%	3.37%

classification. The desired result would be a high accuracy across all overlap values  $d$ . A classifier exhibiting this type of behaviour would be deemed robust against the output of the CSS algorithm. Figure 6 shows the accuracy of the logistic classifier as a function of the overlap threshold  $d$ . We also ran experiments with the following groups of features: colour, size, texture, and shape. The accuracy-overlap graphs for these groups are plotted in Fig. 6. The means and standard deviations across the values of  $d$  are shown in Table 2. The standard deviation is used as a measure of stability, the lower this value the more stable the classifier is to changes in  $d$ . Here we notice that all sets of features behave fairly similarly, indicating high stability of the system across the overlap threshold. This is an encouraging finding that gives us ground and confidence to develop the system into a commercial product.

The resubstitution accuracy of the logistic classifier was found to be 90.2%. This is known to be an optimistically biased estimate, hence we check the system’s accuracy against the 10 times 10-fold cross-validation accuracy found during system’s construction. That accuracy was 89.7% with standard deviation 3.44 (Kuncheva et al. 2008). The observed average accuracy of 86.5% is slightly lower but still acceptable for the purposes of kerogen classification. The consistently high accuracy of the classifier is a reflection of the fact that some objects are “stable”, i.e. they will be present in the majority of the values of  $d$ , and will acquire their own labels as in the training set. In other words, the testing sets contain a large proportion of the training objects. This signifies good quality of the CSS algorithm as well as the

**Fig. 7** Data set sizes at different levels of the overlap threshold  $d$



system because the recognition rate will not suffer a significant drop even for moderate changes in the overlap threshold  $d$ . Figure 7 shows the number of objects in each of the 11 data sets. Viewed together with Fig. 6, it indicates that the recognition rate deteriorates for a larger number of objects but is levelled as the objects are aggregated into bigger ones (smaller  $d$ ).

Table 2 shows that colour features strike the best balance between accuracy and stability compared to the other three groups (texture, size, and shape). There are two reasons for this. First, colour features were found to be the best predictors of the class labels in previous studies (Weller et al. 2006; Kuncheva et al. 2008), so we can expect that their overall performance will be better than that of the other groups of features. Second, colour features are not going to change a lot if an object is broken into smaller pieces, whereas shape and size features would. This explains the decline of accuracy for the shape and size features toward the end of the graph in Fig. 6 and the relatively unchanged accuracy of colour and texture features. In fact, texture features exhibit the best stability across different values of  $d$ . However, it is illustrated in Fig. 6 that, regardless of  $d$ , using all the features ensures better accuracy than any of the feature groups alone, hence we recommend that choice.

## 5 Conclusion

We describe and examine an automatic image processing system that classifies kerogen in images of sieved rock samples as either inertinite or vitrinite. CSS is applied to segment out kerogen pieces in the image. The algorithm depends upon an overlap parameter  $d$  that needs to be set by the user. The purpose of this study was to evaluate the stability of the classification accuracy of the system for variations in  $d$ . The ground truth labels—inertinite or vitrinite—were provided by an expert-palynologist for each of the 609 pieces extracted from our images at  $d = 0.5$ . However, evaluating the classification accuracy for a different value of  $d$  required bespoke labels (to be used as ground truth) for the objects extracted at that value. Hence, a label-inheritance

tree algorithm was proposed here. By propagating the labels we constructed 10 additional labelled data sets, one for each value of  $d$  from  $\{0.0, 0.1, 0.2, \dots, 1.0\}$ ,  $d \neq 0.5$ . The logistic classifier trained at  $d = 0.5$  was tested on each new data set. The results demonstrate that the system is not inadequately sensitive even if dramatic changes of  $d$  are applied (Fig. 6). Having examined the classification accuracy and the sensitivity of the system to its only parameter (the overlap threshold), further aspects of its quality should be considered before deploying a prototype for research or commercial use. These include transparency of the decision, computational complexity, operation speed, scalability and, transferability. Future plans involve also obtaining a second expert opinion on the labels of the kerogen objects and evaluating the match with the system prediction and with the original labels given by the first expert. A class label “other” will be reinstated, and the system will be given a chance to abstain from making a classification. Similar type of object extraction and classification can be developed for the non-kerogen objects in the image, i.e. for the remaining fossils after all kerogen pieces are removed.

**Acknowledgements** The EPSRC CASE grant Number CASE/CNA/05/18 is acknowledged with gratitude.

## References

- Athersuch J, Banner FT, Higgins AC, Howarth RJ, Swaby PA (1994) The application of expert systems to the identification and use of microfossils in the petroleum industry. *Math Geol* 26(4):483–489
- Bollmann J, Quinn P, Vela M, Brabec B, Brechner S, Cortés M, Hilbrecht H, Schmidt DN, Schiebel R, Thierstein HR (2004) Automated particle analysis: Calcareous microfossils. In: Francus P (ed) *Image analysis, sediments and paleoenvironments*. Dordrecht, Kluwer Academic, pp 229–252
- Charles JJ, Kuncheva LI, Wells B, Lim IS (2008a) Object segmentation within microscope images of palynofacies. *Comput Geosci* 34:688–698
- Charles JJ, Kuncheva LI, Wells B, Lim IS (2008b) Background segmentation in microscopy images. In: *Proc 3rd international conference on computer vision theory and applications VISAPP08, Madeira, Portugal*, pp 283–294
- Hills S (1988) Outline extraction of microfossils in reflected light images. *Comput Geosci* 14(4):481–488
- Kuncheva LI, Charles JJ, Miles N, Collins A, Wells B, Lim IS (2008) Automated kerogen classification in microscope images of dispersed kerogen preparation. *Math Geosci* 40(6):639–652
- Soille P (2003) *Morphological image analysis: principles and applications*. Springer, New York
- Swaby PA (1992) VIDES: An expert system for visually identifying microfossils. *IEEE Expert: Intell Syst Appl* 7(2):36–42
- Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans Pattern Anal Mach Intell* 13(6):583–598
- Weller AF, Corcoran J, Harris AJ, Ware JA (2005) The semi-automated classification of sedimentary organic matter in palynological preparations. *Comput Geosci* 31(10):1213–1223
- Weller A, Harris A, Ware J, Jarvis P (2006) Determining the saliency of feature measurements obtained from images of sedimentary organic matter for use in its classification. *Comput Geosci* 32(9):1357–1367
- Wells B (2008) Digital petrography data collection software. <http://www.stepsstage.com>
- Witten H, Frank E (2005) *Data mining: practical machine learning tools and techniques*, 2nd edn. Morgan Kaufmann, San Mateo