Automatic Segmentation of Diatom Images for Classification

ANDREI C. JALBA, MICHAEL H.F. WILKINSON, AND JOS B.T.M. ROERDINK*
Institute for Mathematics and Computing Science, University of Groningen, 9700 AV Groningen, The Netherlands

KEY WORDS segmentation techniques; mathematical morphology; connected operators; watershed from markers; diatom identification; decision trees

ABSTRACT A general framework for automatic segmentation of diatom images is presented. This segmentation is a critical first step in contour-based methods for automatic identification of diatoms by computerized image analysis. We review existing results, adapt popular segmentation methods to this difficult problem, and finally develop a method that substantially improves existing results. This method is based on the watershed segmentation from mathematical morphology, and belongs to the class of hybrid segmentation techniques. The novelty of the method is the use of connected operators for the computation and selection of markers, a critical ingredient in the watershed method to avoid over-segmentation. All methods considered were used to extract binary contours from a large database of diatom images, and the quality of the contours was evaluated both visually and based on identification performance. *Microsc. Res. Tech. 65:72–85, 2004.*

INTRODUCTION

In this report, we consider the problem of automatic segmentation of diatom images. This work grew out of the Automatic Diatom Identification and Classification (ADIAC) project (du Buf and Bayer, 2002), aimed at automating the process of diatom identification by digital image analysis. Diatoms are microscopic, single-celled algae, which build highly ornate silica shells or frustules. They live in aquatic environments (fresh and salt water), and comprise an estimated 200,000 species. Diatoms can be identified based on (1) *contour shape* and (2) pore patterns or "striae" of the shells, called *striation* or *ornamentation* (du Buf and Bayer, 2002, chapter 2, pp. 9–40). The sensitivity of certain diatom species to environmental parameters means that they can be used as indicators of water quality, pollution histories, or climate change. Other applications arise in oil exploration and forensic science (diagnosis of drowning) (du Buf and Bayer, 2002, chapter 3, pp. 41–53). All these applications require counting and identification of different species present in the sample of interest. However, prior to automatic identification, reliable segmentation should be performed. Our goal in this report is to present a framework for automatic segmentation of high-magnification, grey-scale diatom images, which improves initial segmentation results obtained within the ADIAC project (du Buf and Bayer, 2002; Fischer et al., 2002).

There are four main approaches for the segmentation of grey-scale images (Adams and Bischof, 1994; Haralick and Shapiro, 1985; Pal and Pal, 1993): threshold techniques, boundary-based methods, region-based methods, and hybrid techniques, which combine both boundary and region criteria.

Threshold techniques (Sahoo et al., 1988) assume that all pixels whose grey-level values are within a certain range belong to one class. They do not use any spatial information of the image, are sensitive to noise, and do not cope well with blurred edges (Adams and Bischof, 1994). *Boundary-based methods* (Davis, 1975) postulate that changes between regions are abrupt. Examples are local filtering techniques, such as edge detectors (Canny, 1986), or active contour methods (Cohen, 1991, Kass et al., 1987). Because these methods cannot ensure continuous edge-detection, an edge-linking step must be used to produce closed contours. Active contour methods automatically produce closed contours and (usually) provide better edge localization, but they are sensitive to noise and require an initialization step that is hard to automate. *Region-based methods* assume that neighboring pixels within the same region have similar values. Representative methods are region growing (Haralick and Shapiro, 1985), split-and-merge techniques (Horowitz and Pavlidis, 1974; 1976), and clustering methods (Haralick and Shapiro, 1992). The main advantage of region-based methods is that they use and adapt the statistics inside the region, but they generate small holes and irregular boundaries. *Hybrid techniques* combine both boundary and region criteria. Two important representatives of this class are morphological watershed segmentation (Meyer and Beucher, 1990; Roerdink and Meijster, 2000) and seeded region growing (Adams and Bischof, 1994). Advantages of watershed segmentation are that it (generally) leads to closed boundaries of the image regions and it can describe edge junctions (Najman and Schmitt, 1994). In contrast, edge detectors based on zero-crossings of differential operators such as the Laplacian-of-Gaussian methods (Haralick and Shapiro, 1992) do not allow detection of T-junctions (Torre and Poggio, 1986). However, the watershed technique has one major drawback, namely severe over-segmentation. Various methods have been proposed to deal with this issue: image pre-processing by (adaptive) smoothing (Weickert, 2001), region merging as a post-

*Correspondence to: Jos B.T.M. Roerdink, Institute for Mathematics and Computing Science, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands. E-mail: j.b.t.m.roerdink@cs.rug.nl

Received 5 May 2004; accepted in revised form 6 October 2004
DOI 10.1002/jemt.20111
Published online in Wiley InterScience (www.interscience.wiley.com).

Fig. 1. Some examples of diatom shells. Spatial dimensions (left to right): 48 × 15, 40 × 7.5, and 42.5 × 16.2 μm.

processing step (Haris et al., 1998), watershed from markers (Lotufo and Falcao, 2000; Meyer and Beucher, 1990), hierarchical segmentation (waterfall algorithm) (Beucher, 1994; Grimaud, 1992), and multiscale gradient watershed (Jackway, 1995).

In this report, we have adopted a systematic approach. We review the results by Fischer et al. (2002), try to adapt existing methods from each of the four standard approaches, and finally propose a watershed algorithm with a new marker selection scheme, which shows the best overall results. We emphasize that this study is not a review on segmentation methods in general. Instead, it focuses on the segmentation of diatom images for the purpose of classification.

MATERIALS AND METHODS

Automatic Slide Scanning

The input of all segmentation methods considered here consists of grey-scale, high-magnification images of diatom shells obtained by automatic slide scanning (Pech-Pacheco and Cristobal, 2002).

The overall procedure of the automatic slide scanning system consists of three parts. First, image acquisition at low magnification is used to obtain a panoramic view of the whole slide, which allows the extraction of the position and size of all particles. Second, an intermediate resolution particle screening is carried out in order to eliminate non-diatom particles. Third, images are captured at high magnification using autofocusing and multi-focus fusion. Particle screening is used to remove a substantial number of particles (debris, broken valves) that are not required to be analyzed. The output of the system is an annotated gallery of images that can be used for diatom segmentation and identification. Some examples of diatom shells obtained by automatic slide scanning are shown in Figure 1; the typical size of these images ranges from 200 × 400 to 600 × 900 pixels. For further details on slide scanning, we refer to Pech-Pacheco and Cristobal (2002).

Image Acquisition

Diatom samples were analyzed using a Zeiss Axiophot photomicroscope, with a 100-W halogen light source, and with 4×, 10× (low magnification), 20× (medium magnification), and 40× (high magnification) lenses. The ocular magnification was 0.6×. Image acquisition was performed using the LG-3 grey-scale frame grabber from Scion. The LG-3 is connected to a Cohu 4910 series monochrome CCD camera, which is a high-quality, economical choice for grey-scale scientific imaging applications. The frame grabber resolution (CCIR) is 768 × 512 pixels and the pixel depth 8 bits.

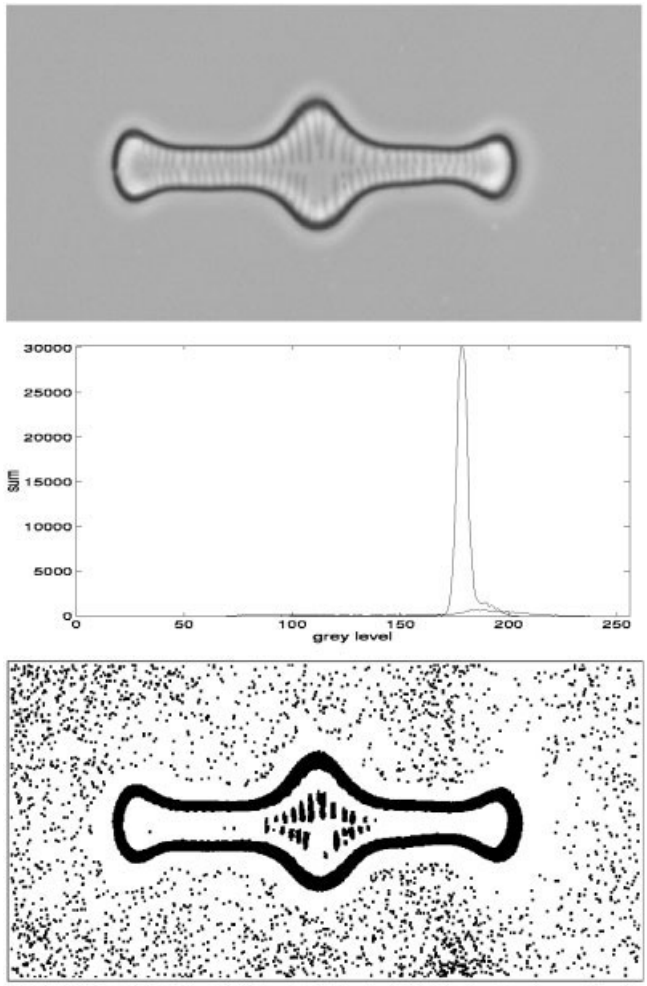

Fig. 2. Heuristic thresholding example. **Top:** Diatom image (28 × 7 μm). **Center:** Background (solid) and diatom (dashed) histograms. **Bottom:** Resulting binary image.

The microscope stage was controlled with a H101 (4″× 3″) controller from Prior Scientific, with a step size of 0.1 μm for the X-Y axes and 1 μm for the Z-axis. The total travel distance of the microscope stage for the X-Y-Z axes is 11 × 7 × 1 cm, respectively.

The size range of the particles analyzed was between 20 and 260 μm, which can be considered part of the microplankton. Smaller diatoms (nanoplankton) cannot be observed using the available system, because they require oil-immersion lenses. The size ranges of

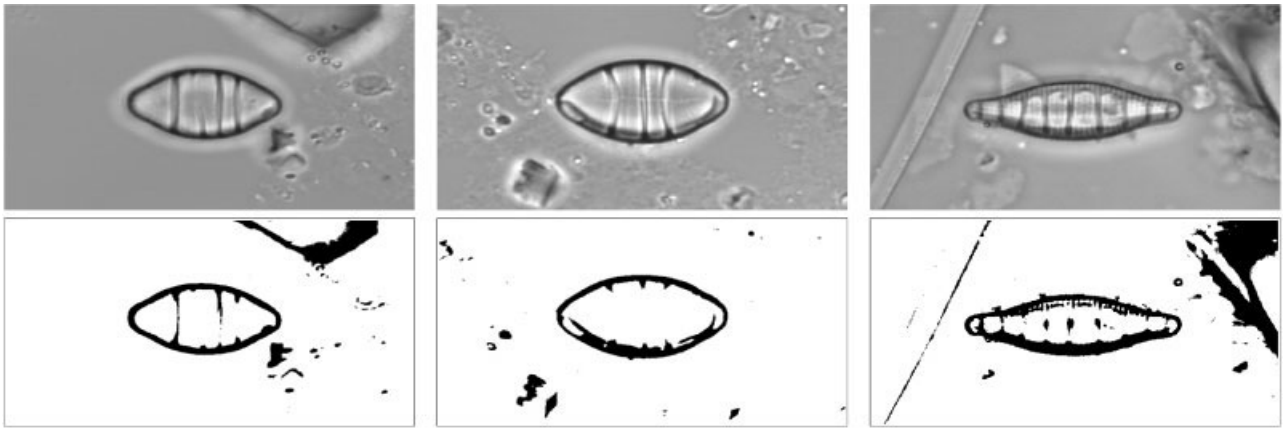

Fig. 3. Iterative threshold selection. Examples. **Top:** Input images (left to right: 19 × 16, 20 × 17, and 19 × 8 μm). **Bottom:** Resulting binary images.

the particles that can be observed with the lenses are as follows: 10×: range = 131–260 μm; 20×: 61–130 μm; 40×: 20–60 μm. Subsequently, we will indicate the spatial dimensions of the bounding box of the central diatom shell present in each image.

Problematic Images

Ideally, each image should contain a single diatom shell. However, there are some serious obstacles to be overcome by all segmentation methods. We mention a number of these problems.

First, diatoms may overlap or be very close to each other. Also, they can be fragmented or may not be in proper focus. Dust specks and debris may be present in the images. If the illumination around the diatom is not uniform, most global thresholding methods fail to find the outlines. In addition, diatoms in microscopic images exhibit the same grey levels as the background, so that the histogram is unimodal (Fischer et al., 2002). This fact obstructs most threshold selection methods that assume multimodal histograms. Moreover, if the diatom is not in proper focus, the edges are blurred and can only be partly detected by most edge detection techniques. Although region-based methods are more flexible, pixels belonging to debris or fragmented diatoms can often not be uniquely assigned to either diatom or background regions.

RESULTS
Segmentation Methods

In this section, we describe each image segmentation technique (grouped according to the classification given above), give example results, and, if appropriate, describe specific pre-processing steps. Postprocessing and boundary-extraction steps, which are common to all segmentation methods, are described later in this section. The output of each segmentation method is assumed to be a binary image, in which candidate diatom contours are depicted in black over a white background.

In each technique, parameter values have to be determined that give the best results for the diatom images. This involves an iterative process of guessing suitable initial parameter values, evaluating the results, and refining the values. Two remarks should be made: (1) the values reported here are not the result of an exhaustive search of the parameter space, because such a search would require an impractically long time; (2) the parameter values were adjusted in an attempt to give best *average performance* across all the images.

Thresholding

Since most diatom images exhibit a unimodal histogram, threshold selection methods based on statistics of the grey-level distribution fail. Thus, we shall use only those threshold techniques that make no assumption about the distribution of grey-level values.

Heuristic Thresholding. Fischer et al. (2002) proposed a simple heuristic procedure for global threshold selection, which was successfully used to extract diatom contours. Starting from the maximum value of the histogram of the image, a search proceeds towards the left tail of the histogram until the first entry is located that has a frequency not higher than 15% of that of the mean value. The grey level that corresponds to this entry is chosen as the threshold. This heuristic is motivated by the fact that the histogram of the background is strongly peaked, while the histogram of the diatom itself is just a flat curve. In addition, diatoms have grey levels similar to those of the background, and only small ranges to the left and right of the main peak of the histogram characterize the dark borders or the light diffraction halos typically occurring around diatoms in optical microscopy images. Typically, such grey levels do not occur in the background, except for very dark regions related to debris such as ash or sand grains. An example is shown in Figure 2, where the mean value is 177 and, by applying the criterion of Fischer et al. (2002), a threshold value of 173 is found. All pixels with grey level above the threshold were marked in white and all other pixels in black. Notice that by following the black border around the diatom, the correct contour can be found.

Iterative Threshold Selection. Ridler and Calvard (1978) have suggested a simple but efficient iterative method for threshold selection. Assume that the input

Fig. 4. RATS method. Examples. **Left:** Input images (19 × 8 μm). **Middle:** Threshold maps. **Right:** Resulting binary images. **Top:** No preprocessing. **Bottom:** After bottom-hat filtering.

image contains only two classes of pixels: (1) pixels within diatom regions, and (2) pixels in the region surrounding the diatoms (background). The segmentation threshold is selected through the following iterative procedure.

1. Initialize the threshold T^0 with the smallest grey level present in the input image.
2. Let T^i be the threshold at step i. Let μ_d and μ_b be the mean grey-level of the current set of diatom and background pixels respectively, after segmentation with threshold T^i. Choose the new threshold for step $i + 1$ as

$$T^{i+1} = \frac{\mu_d + \mu_b}{2}.$$

3. Repeat until stability, i.e., until $T^{k+1} = T^k$; the value of the final threshold is given by T^k.

Some results of this method are shown in Figure 3. Although the input images are difficult to segment because of debris and fragments of other diatoms, the contours of the central diatoms are correctly found and can be traced.

Robust Automatic Threshold Selection (RATS). Kittler et al. (1985) introduced the Robust Automatic Threshold Selection (RATS) algorithm for segmentation. The RATS method computes thresholds either locally or globally using a weighted average of the grey levels within arbitrary areas of an image. A variant of the RATS algorithm (Wilkinson, 1998) uses a quadtree representation of the input image, and the weight assigned to each pixel is the squared response of a Sobel edge detector at that pixel. Within a given region R_j indexed by j, the threshold T_j is computed using the gradient-magnitude image E and the original grey-level image I by

$$T_j = \frac{\sum_{i \in R_j} E[i] \cdot I[i]}{\sum_{i \in R_j} E[i]}$$

The computed threshold values are assigned to the centers of the smallest regions, and then interpolated across the entire image space. For more details about this method we refer to Wilkinson (1998). The method has three parameters, the number of levels L in the quadtree, the noise level σ and a parameter λ, which is used during threshold computation to omit the pixels with gradients below $\lambda \cdot \sigma$. In our experiments, the best results (where "best" is defined in terms of visual quality or identification performance) were obtained for $L = 5$, $\sigma = 4.0$, and $\lambda = 3.0$. All these values are identical to those indicated in Wilkinson (1998), except that of the noise level σ, which was increased from 2.0 to 4.0. Because of the high content of debris present in most diatom images, some pre-processing was necessary before using this method. As mentioned at the end of the second section, most diatoms show a prominent dark outline caused by the inwardly curved cell wall. Therefore, a morphological bottom-hat filter (Haralick and Shapiro, 1992) with a structuring element of size 50 × 50 was used to extract the dark border, prior to thresholding.

An example is shown in Figure 4. The top and bottom rows refer to the case of no preprocessing, and preprocessing by the bottom-hat filter, respectively. Notice that without preprocessing, the diatom outline is connected to some circular debris at its right end, so that a correct contour cannot be traced. When the bottom-hat filter is applied, a much better result is obtained and the correct contour can be found.

Boundary-Based Techniques

We discuss one of the representative methods of this class, i.e., the *Canny edge detector* (Canny, 1986). This method uses a discrete gradient operator in a local neighborhood of each pixel. High responses of this operator provide candidate edge pixels. Hysteresis thresholding is used to decide whether weak responses correspond to edges or not. For diatom segmentation, the Canny edge detector yielded good results (Fischer et al., 2002), and it successfully complemented the threshold-based methods. Some examples are shown in Figure 5. In spite of the presence of fragmented dia-

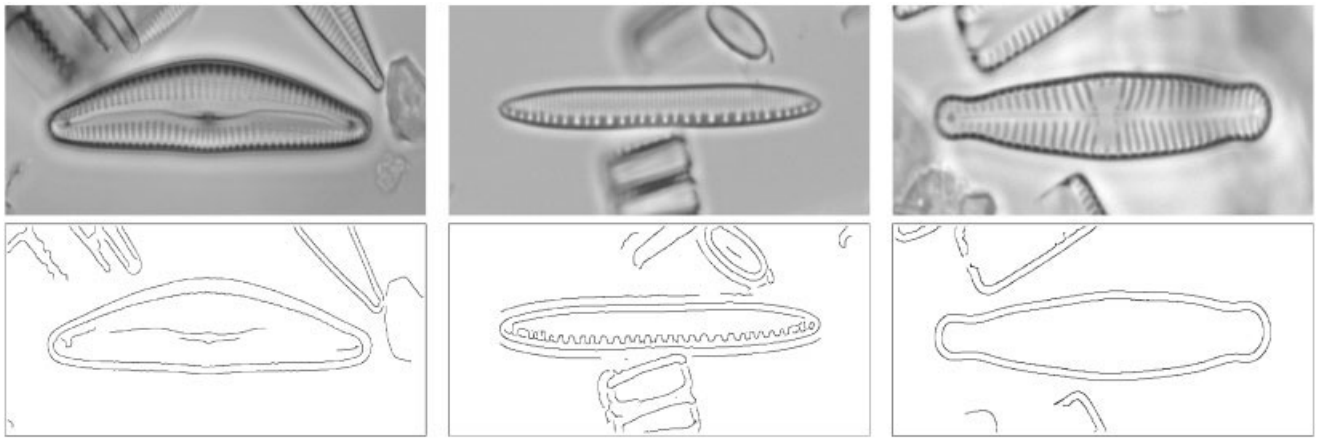

Fig. 5. Canny edge detector. Examples. **Top:** Input images (left to right: 65 × 13.3, 41.2 × 4, 37.5 × 8 μm). **Bottom:** Resulting binary images.

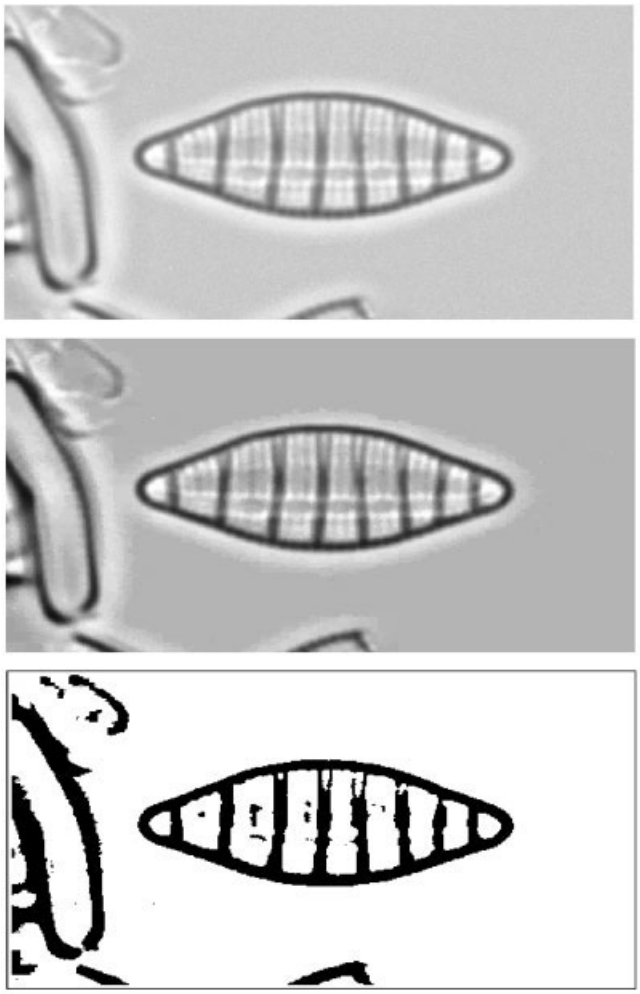

Fig. 6. Split-and-merge method. Example. **Top:** diatom image (18.7 × 6.2 μm). **Center:** Output of the split-and-merge method. **Bottom:** resulting binary image.

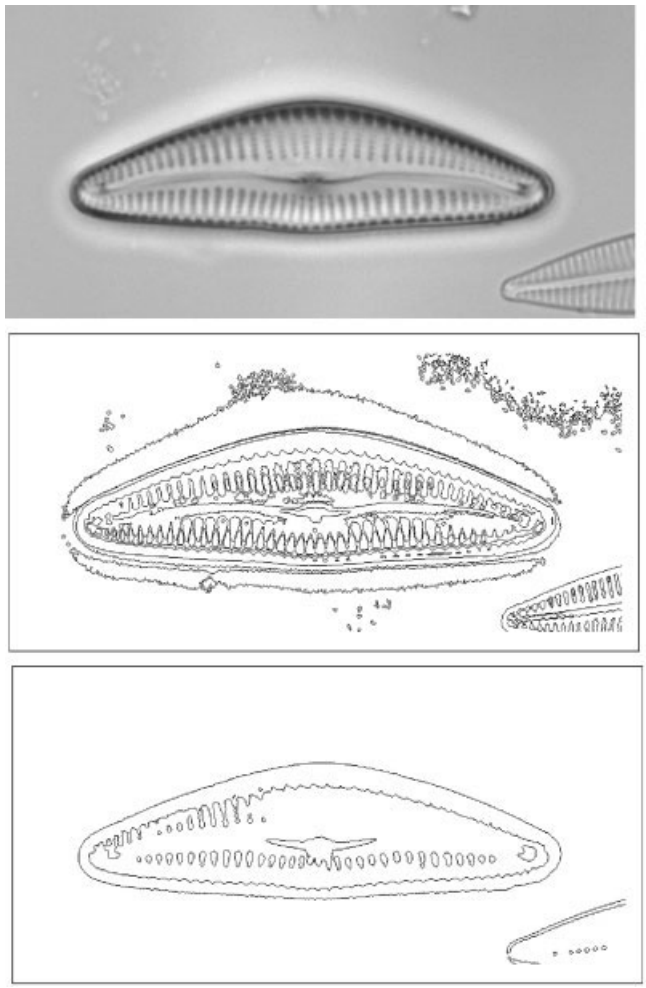

Fig. 7. Clustering. Example. **Top:** Diatom image (63.3 × 10.5 μm). **Center:** Over-segmentation, $r = 0.4\sigma$. **Bottom:** Undersegmentation, $r = 1.2\sigma$.

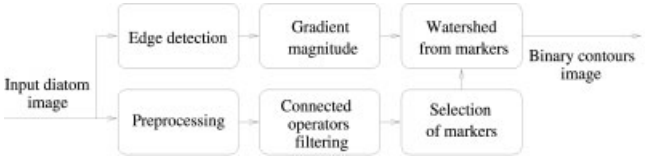

Fig. 8. Hybrid segmentation method.

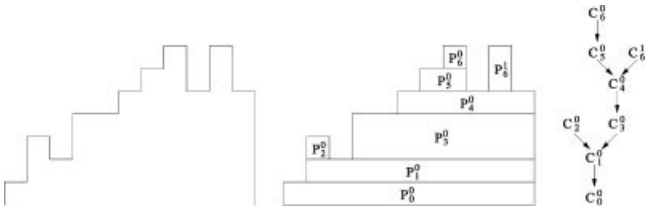

Fig. 9. The max-tree structure. **Left:** A 1-D signal. **Center:** Peak components P_h^k of the signal. **Right:** Its corresponding max-tree.

toms, the contours of the central diatoms are correctly found.

The three parameters of the method, the width σ of the Gaussian smoothing filter and the low and high values τ_l and τ_h of the hysteresis-based thresholding, were tuned for best performance. The values obtained were: $\sigma = 4.0$, $\tau_l = 0.5$, and $\tau_h = 0.8$.

Region-Based Techniques

Region-based methods consider segmentation as a process that partitions the entire image space R into n subregions, $R_1, R_2, \ldots, R_n$, such that (i) $\bigcup_{i=1}^{n} R_i = R$, and (ii) $R_i \cap R_j = \phi, \forall\, i, j$ with $i \neq j$. The first condition requires that every pixel is assigned to a region, while the second says that the regions must be disjoint.

Split-and-Merge Technique. The split-and-merge technique supplements the above requirements by two criteria: (iii) $P(R_i) = TRUE$, $\forall\, i$, and (iv) $P(R_i \cup R_j) = FALSE$, for $i \neq j$ and R_i adjacent to R_j. Here $P(R_i)$ is a logical predicate defined over the points in set R_i. The first condition represents the properties that must be satisfied by the pixels in a segmented region, while the second condition indicates that any two adjacent regions must be different in the sense of predicate P. The general procedure (Horowitz and Pavlidis, 1976) is to subdivide an image into a set of disjoint regions that are subsequently merged and/or split in an attempt to satisfy the above conditions. In the segmented image, the mean intensity value of each identified region is used as the grey value of that region.

For diatom segmentation, good results are obtained using the variance of the grey-level distribution as the measure of homogeneity. The predicate $P(R_i)$ is defined to be true $\sigma^2_{R_i} < T$, where $\sigma^2_{R_i}$ is the variance of intensity values inside R_i and T is a threshold. However, using this predicate, it is not possible to obtain a partitioning of the input image only into diatom regions and background. Instead, we can partition the image into background and other "inside-diatom" (non-background) regions.

The threshold T is estimated as follows. The image domain R is divided into $B \times B$ square-blocks, and the variance inside each block is computed. Then, variance samples are drawn uniformly. Let A_B be the area of the background, and A_D be the area of all "inside-diatom" regions. If we denote by m the number of samples drawn and by k the number of samples that belong to "inside-diatom" regions, then $A_D / A_B \approx k / m$. Assuming that $A_D \leq A_B / 2$, and that the variance of a background sample is smaller than that of an inside-diatom region sample, we obtain an estimate of the threshold T as the median of the smallest $m - k$ values (with $k \leq m / 2$). In our implementation, we use $B = 10$ and $m = 300$. Finally, assuming again that the area of the background is larger than that of the inside-diatom regions, the grey level of the background region is chosen as the maximum value of the histogram of the output image, and a threshold of 10 grey levels below this maximum value is used to produce the final binary image. The value of the threshold was determined experimentally based on the observations made when discussing threshold-based methods. Most diatom images were successfully segmented by this method. An example is shown in Figure 6. The value of the threshold T found

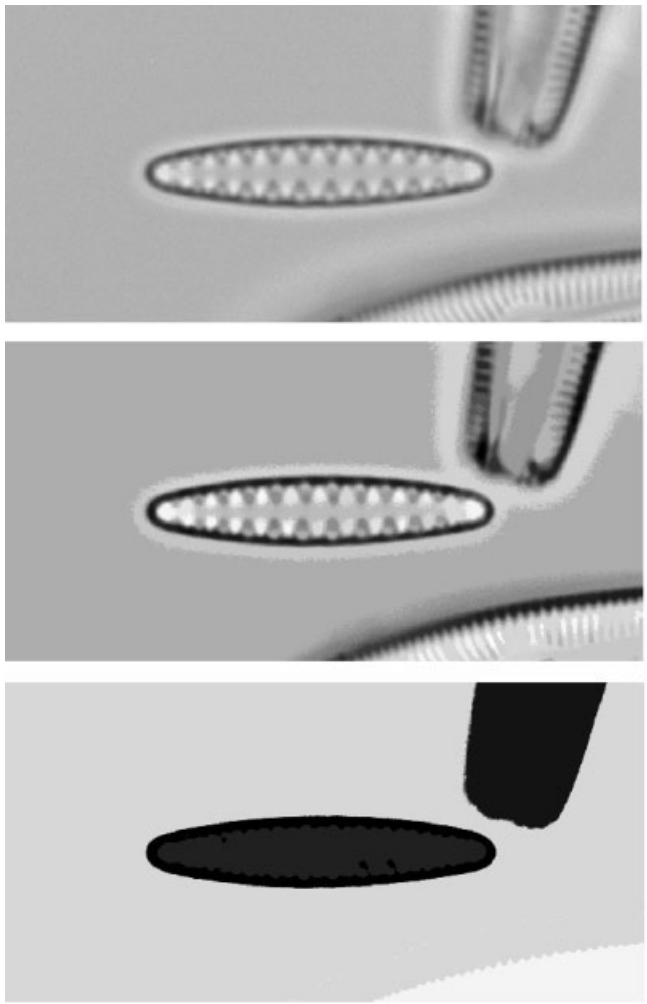

Fig. 10. Connected operator filtering. Example. **Top:** Diatom image (20.5 × 3.7 μm). **Center:** λ = 5%. **Bottom:** λ = 1%.

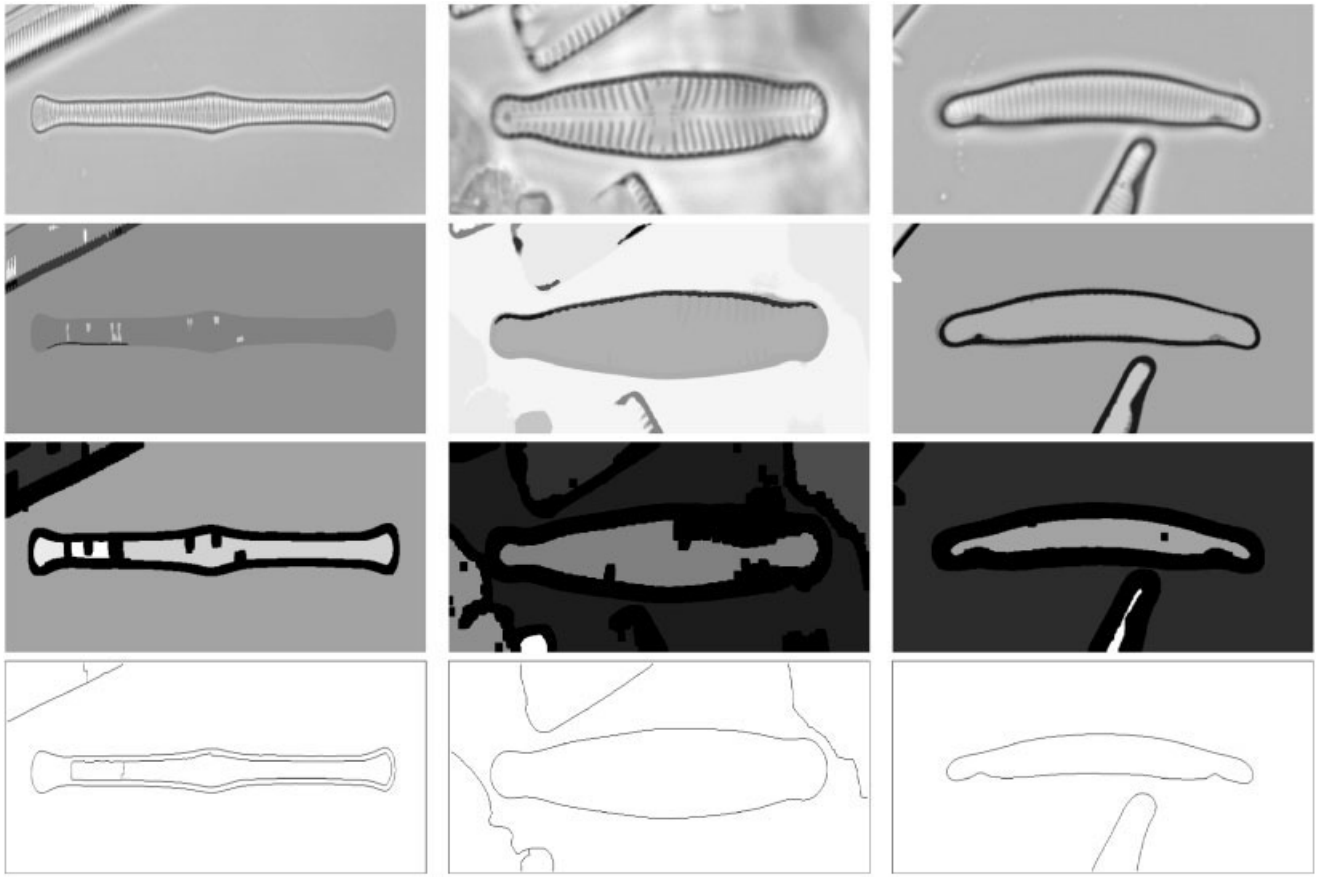

Fig. 11. Hybrid segmentation technique. Examples. **Top row:** Diatom images (81 × 5.3, 37.5 × 8, 40 × 5 μm). **Second row:** Filtered images. **Third row:** Label images. **Fourth row:** Binary images resulting from the watershed transform.

is 9.6, and the grey level of the background region is 206. The resulting binary image, after thresholding at grey level 196, is shown in Figure 6 (bottom).

Clustering. Pattern vectors are extracted from local search neighborhoods of an input image. In Comaniciu and Meer (1997), a general auto-clustering technique was proposed for recovery of significant image features. An advantage of this method is that the number of clusters does not need to be known a priori. The outline of this simple feature space analysis is as follows (Comaniciu and Meer, 1997).

1. Map the image domain into the feature space;
2. Define an adequate number of search windows at random locations in the space;
3. Find the high-density region centers by applying the mean shift algorithm (Fukunaga and Hostetler, 1975) to each window;
4. Validate the extracted centers with image domain constraints to provide the *feature palette* (see the explanations below);
5. Allocate all the feature vectors to the feature palette, using image domain information.

This procedure can be easily cast in the form of an image segmentation algorithm by mapping the mean of a small neighborhood around each pixel to the feature space (Comaniciu and Meer, 1997). Grey-scale images are handled as color images with only the brightness coordinate defined.

Three parameters define the segmentation resolution: (1) r, the radius of the search window, taken to be proportional to the square root σ of the trace of the global covariance matrix; (2) *Nmin*, the smallest number of elements required for a significant color, and (3) *Ncon*, the smallest number of connected pixels required for a significant image region. The initial feature palette is given by significant colors, i.e., colors represented by at least *Nmin* vectors in feature space, and by minimally one connected component with at least *Nmin* pixels in the image domain. The final palette is obtained by mapping the mean values of the feature vectors to the same colors of the initial palette. Finally, in a post-processing step, small, connected components of size less than *Ncon* are removed.

In our experiments, the best results using this method were obtained with the following values: $r = 1.2\sigma$, $Ncon = 10$, and $Nmin = 70 \times 70$. According to the classification in Comaniciu and Meer (1997), this parameter setting corresponds to a severe under-segmentation. Therefore, most diatom images are clustered in two classes, one corresponding to the dark border re-

gions around diatoms and to some dark striae inside diatoms, and the other containing the remaining pixels. An example is shown in Figure 7. If an increased segmentation resolution is selected ($r = 0.4\sigma$), the light diffraction halos around diatoms form new classes whose corresponding regions are usually merged with the diatom regions, resulting in ragged and poorly located boundaries (see Fig. 7, center image).

A Hybrid Technique Based on the Watershed From Markers

In this subsection, a method based on the morphological watershed (Meyer and Beucher, 1990) with a new marker selection scheme is proposed. The main steps of our method are shown in Figure 8. The processing branches in two paths according to the desired output. One of the paths ends with the selection of markers, which produces a label image, while the other ends with the computation of a gradient-magnitude image. Both these images are then used in the watershed-segmentation step.

The novelty of the proposed technique is the computation and selection of markers using the concept of connected operators. A new connected operator is used to simplify the input image and to produce candidate marker regions. A further selection step is carried out to produce the final markers as a label image. Although this label image can be readily used to trace the diatom contours, better contour localization is obtained if a watershed process initiated from the label image is applied to the gradient of the input image.

Preprocessing. A non-linear method for contrast enhancement (Fairfield, 1990) is applied to the input image. The basic idea of the method is as follows. First, the gradient of the image is computed using the Sobel operator (Haralick and Shapiro, 1992). Then, for each pixel of the image, a sliding downhill is performed on the gradient-squared surface until a local minimum is found. All pixels along the path followed during the latter sliding downhill are given the grey-level value of the local minimum. This contrast enhancement is used for marker extraction, but not in the step that leads to the gradient-magnitude image. The reason is that this technique performs a rough quantization of the grey levels, which may result in false edges hampering the evolution of the watershed. Regions associated with false edges can be eliminated either by the subsequent filtering step, or during the selection of marker regions. If some regions still survive, they can be neglected when the contours are extracted, due to the property of the watershed to allow for T-junctions.

Connected Operator Filtering. Connected operators preserve contours of an image and only alter the grey values of its level components (connected components, in the binary case) (Heijmans, 1999; Salembier and Serra, 1995; Serra and Salembier, 1993). To implement anti-extensive connected set operators, we use the *Max-tree* representation introduced by Salembier et al. (1998).

Before describing this representation, we introduce some notation. Let R be the domain of a greyscale image f. A flat zone L_h at level h of grey-scale image f is a connected component of the level set $X_h(f)=\{p \in R : f(p) = h\}$. A *peak component* P_h at level h is a connected component of the threshold set $T_h(f) = \{p \in R : f(p) \geq h\}$. At each level h there may exist several such components, indexed as L_h^i, P_h^j, with i, j from two index sets.

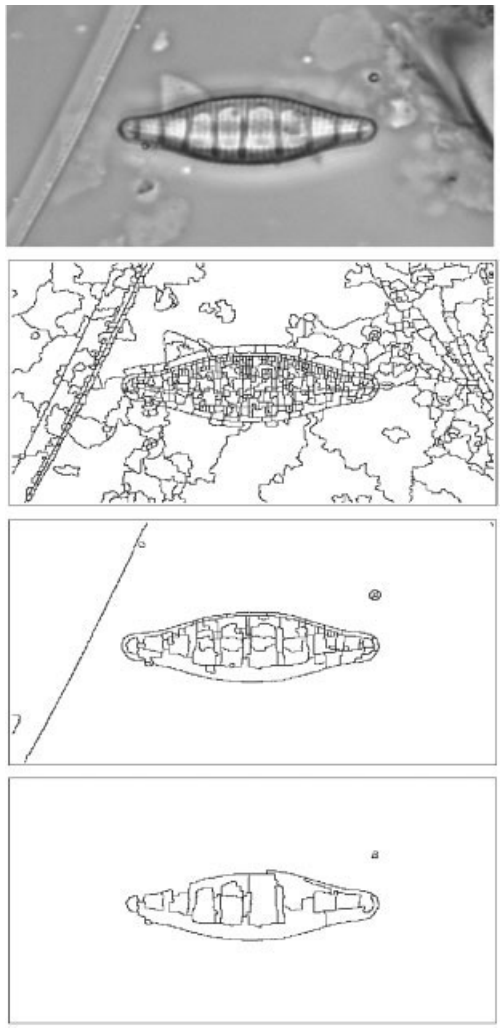

Fig. 12. Hierarchical watershed. Examples. **Top-to-bottom:** Input image (19×8 μm); binary images resulting from the watershed transform with dynamics thresholds set to 5, 13, 20, respectively.

A max-tree is a rooted tree, in which each of the nodes C_h^k at grey-level h corresponds to a peak component P_h^k. However, C_h^k contains only those pixels in P_h^k which have grey level h. In other words, it is the union of all flat zones $L_h^j \subseteq P_h^k$. An example of a 1-D signal, its peak components, and its max-tree representation is shown in Figure 9. For processing the input image, we have a three-step process: (1) construction of the max-tree, (2) deciding which nodes to keep or remove according to some criterion, and (3) image restitution, which transforms the filtered max-tree into an output image.

Let us assume that components with large areas compared with the area of their parent component are to be preserved. Let $A(P_m^n)$ be the area of the peak component P_m^n associated with the node C_m^n. As a measure, we use the *percentage difference* between the area

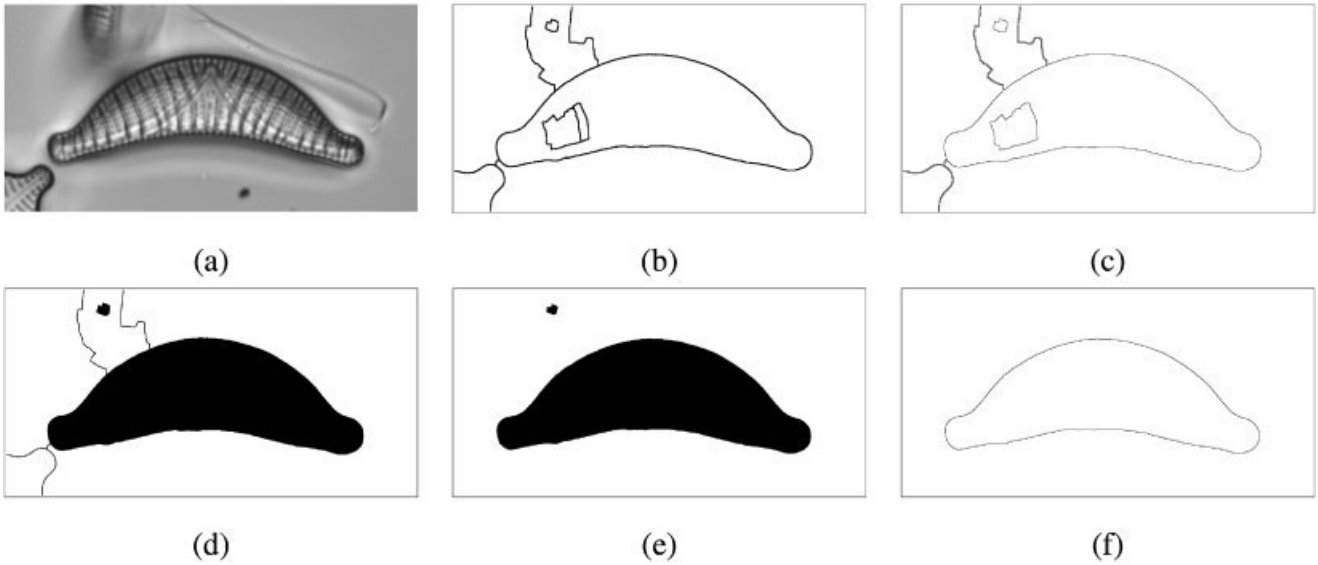

Fig. 13. Contour extraction. Example. **a:** Initial image (48.7 × 15 μm). **b:** Binary image with watershed lines. **c:** Extracted contours. **d:** Regions obtained after flood-filling the contours. **e:** Closed image. **f:** The final contour.

of a component and the sum of areas of its child components (corresponding to child nodes $C_{h_i}^k$), i.e.,

$$\delta_A = 100 \times \frac{A(P_m^n) - \sum_i A\,(P_{h_i}^k)}{A(P_m^n)} = 100 \times \frac{A(C_m^n)}{A(P_m^n)}\,(\%),$$

with $h_i > m$, and k, n from two index sets. Starting with the root node, δ_A is recursively computed according to the above equation. If for a given node C_h^k this value is larger than a threshold λ, all its *direct* child nodes are marked as deleted. In a subsequent step, all marked nodes are merged with their nearest preserved ancestors.

The measure δ_A is suitable for simplification of diatom images. The reason is that the shells of most diatoms present striae patterns (i.e., alternating dark and light stripes), and the regions corresponding to the light stripes will be removed during filtering because the sum of their areas is small compared to the area of their parent component.

By the duality $f \leftrightarrow -f$, one can construct a min-tree, as explained in Salembier et al. (1998). The effect of both operators is a simplification of the input image, controlled by the parameter λ. An example is shown in Figure 10. Notice that for other segmentation problems, these operators can be augmented by using some information related to the grey level of the components or to the variance of the grey-level distribution of the child components. However, here this filtering step is used only to provide candidate marker regions, and such extensions are not necessary.

Selection of Markers. Marker selection is the most critical part of the watershed method. As the number of markers does not change during the watershed evolution, a marker region lost during marker selection cannot be recovered later. Therefore, special care must be taken during the marker selection process, for which we propose the following procedure.

- Compute the morphological gradient with a structuring element of size 13 × 13, and label with zero all pixel positions where its value is greater than zero;
- Do a connected component labeling (Haralick and Shapiro, 1992) (with labels greater than zero) of all regions which are not assigned a value of zero;
- Regions with areas smaller than a threshold of 100 pixels are not considered marker regions, i.e., they are marked with a zero label.

At the end of this procedure, all marker regions have a unique label greater than zero, and all other regions are marked with zero. Next, basins are grown from marker regions by the watershed transform under the control of the magnitudes of edges.

Gradient Magnitude Computation. The gradient is obtained by convolving the initial image with a derivative Gaussian filter (Haralick and Shapiro, 1992). In our implementation, the width of the kernel σ was set to 4.0.

Watershed From Markers. The classical watershed from markers (Lotufo and Falcao, 2000; Meyer and Beucher, 1990; Roerdink and Meijster, 2000; Vincent and Soille, 1990) grows *catchment basins* from marker regions, which are the regional minima of the (gradient-magnitude) image. Unfortunately, this approach leads to over-segmentation. The traditional way to overcome over-segmentation in the watershed segmentation is to apply a minima imposition operator, which changes the homotopy of the image in such a way that the desired markers are the only regional minima of the image (Meyer and Beucher, 1990). Here we use an alternative approach, by using a particular case of the Image Foresting Transform (IFT), which

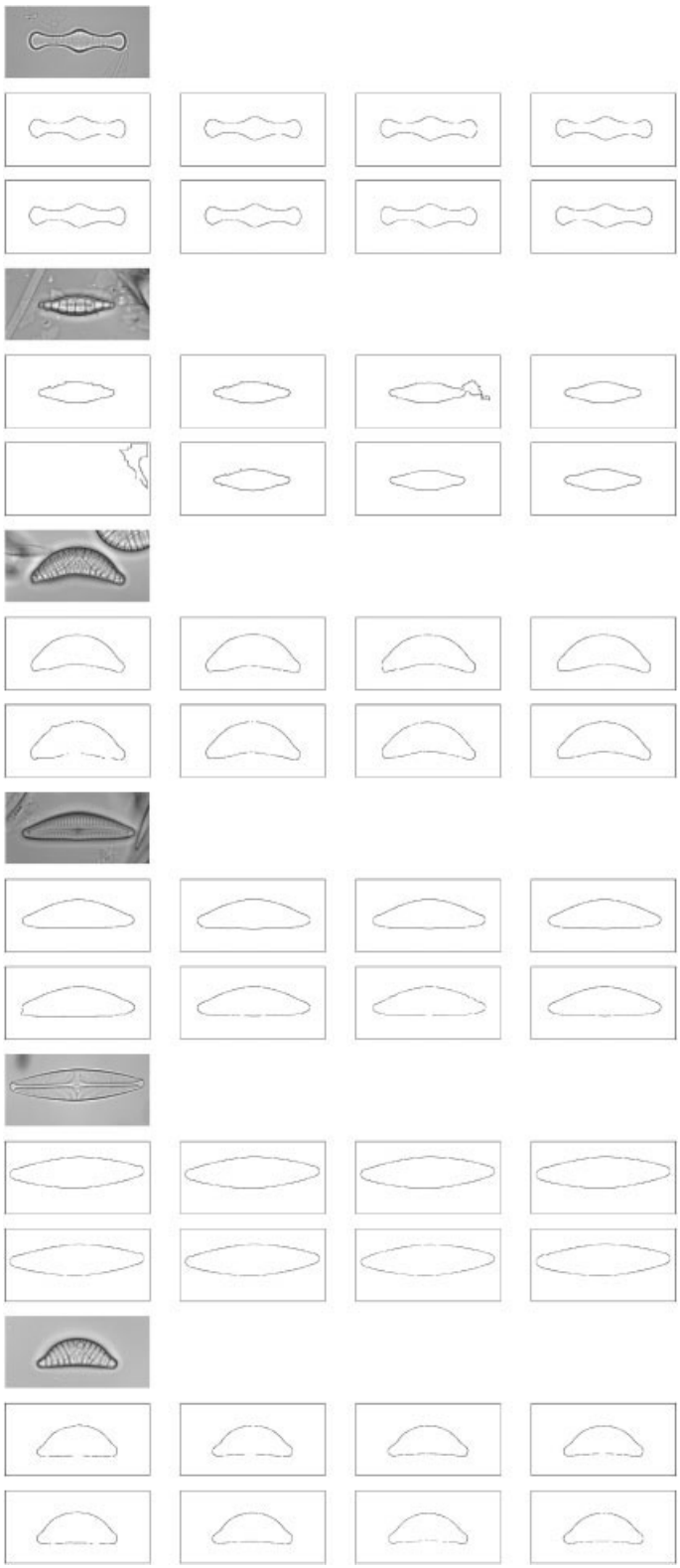

Fig. 14. Contour extraction. For each input image (28 × 7, 19 × 8, 48.5 × 14, 59 × 12, 61 × 13, and 28.7 × 11.2 μm), from left-to-right, top-to-bottom, binary contours extracted using: heuristic thresholding, iterative thresholding, RATS, Canny edge detection, split-and-merge, clustering, hierarchical watershed, hybrid.

TABLE 1. Segmentation results

Segmentation method	Errors	Performance (%)
Heuristic thresholding	176	78.2
Iterative thresholding	132	83.6
RATS	185	77.1
Canny edge detection	125	84.5
Split-and-merge	120	85.1
Clustering	176	78.2
Hierarchical watershed	108	86.6
Hybrid	16	98.0

has the advantage over the classical watershed that it guarantees the optimality of the solution, as long as the cost path is a non-decreasing function of the arc weights (Lotufo and Falcao, 2000). Also, similarly to the watershed from markers using an ordered queue, the IFT does not need a change of homotopy.

All regions in the label image whose pixel values are greater than zero provide the marker regions from which the watershed segmentation is initiated. The watershed propagation is done on the gradient-magnitude image. As a final step, the watershed lines are drawn in order to provide a binary image. Some results using the proposed hybrid segmentation technique are shown in Figure 11. Although the large region corresponding to the central diatom in the first image is split by the marker selection procedure, this is not a problem since all extracted contours are flood-filled in the post-processing step. Some T-junctions produced by the watershed can be observed in the first two cases.

Hierarchical Watersheds

Other techniques to suppress over-segmentation do exist in the literature, such as waterfalls (Beucher, 1994) and hierarchical watersheds (Meyer and Beucher, 1990; Beucher, 1994) using dynamics (Grimaud, 1992) or extinction values (Vachier and Meyer, 1995) as criteria for selecting markers from the set of catchment basins. Here we will use the dynamics concept dynamics (Grimaud, 1992) as the basic notion for hierarchical watersheds, which represents a powerful tool for selecting significant extrema with respect to a contrast criterion. The *dynamics* of a regional minimum is the minimum height a pixel has to climb in a walk to reach another regional minimum with a higher dynamics. After the dynamics of each regional minimum is computed, one uses as markers for the watershed only minima above a given dynamics. Then, by varying this threshold one can build a hierarchy of nested segmentations.

Some examples obtained using this method for different values of the dynamics threshold are shown in Figure 12. For small values of the threshold, the over-segmentation problem still remains (Fig. 12, the second image), whereas larger values result in under-segmentation (Fig. 12, last image). The best overall results using this method were obtained when the dynamics threshold was set to 13. We also experimented with area and volume extinction values (Vachier and Meyer, 1995), but the results were worse than those obtained using the dynamics criterion (results not shown).

Post-Processing and Contour Extraction

The purpose of this last step is to extract the contours of the diatoms present in the binary images, in which diatom borders are depicted in black on a white background. All borders in the binary image are traced using a standard contour tracing algorithm (Haralick and Shapiro, 1992). All extracted contours, which are necessarily closed, are filled at grey-level zero by a flood-fill algorithm, and all obtained regions are drawn in the same image. In a further post-processing step, an opening with a structuring element of size 3×3 is performed, in order to prune thin structures that may still be connected to diatom regions, due to debris or fragments of other diatoms. In this way, the union of all diatom and inner-diatom regions is obtained and all diatom contours can be found by tracing only a single contour per region. Finally, an area closing of $\lambda = 4{,}900$ pixels is performed in order to remove small regions that are not considered diatom regions.

The whole tracing process is illustrated in Figure 13. Notice that inner-diatom regions that were not removed by filtering or by the marker selection procedure are now merged into one large diatom region.

EXPERIMENTAL RESULTS ON A DIATOM DATABASE

All segmentation methods described in the previous section were applied to a database of 808 diatom images comprising 37 different diatom taxa, each taxon (class in the pattern recognition sense) having at least 20 representatives (du Buf and Bayer, 2002). The quality of the extracted contours was evaluated: (1) visually, and (2) based on identification performance.

Visual Quality

Visual estimation of the quality of the contours was guided by the following criteria: (1) contours should be smooth, (2) they should correspond well with the perceived diatom outlines, and (3) they should not enclose debris or diatom fragments.

Some binary contours for typical samples of the diatom database are shown in Figure 14, for all segmentation methods considered above. All methods (with one exception, for the second input image) succeed in extracting the contours of the central diatoms, although fragments of other diatoms or debris may be close to them (see the first, third and the fourth input images). Good quality contours of the central diatom in the second image are extracted only by the edge-detection, hybrid, and hierarchical watershed methods. For this image, all other methods failed, producing either jagged contours (thresholding methods), or no contour at all (split-and-merge method).

The overall quality of each segmentation method is shown in Table 1. All contours that did not fulfill the above requirements and all images for which no appropriate contours could be extracted were considered as errors, and are shown in the second column of Table 1. The third column in Table 1 shows the performance of each method as a percentage of the number of good quality contours. The hybrid technique showed the best results, producing up to 98% correctly extracted contours, while the performance of all other methods is smaller by at least 10%. All 16 images where the hybrid method failed are shown in Figure 15.

An interesting question is how many errors (from these 16) can be corrected by using all other methods.

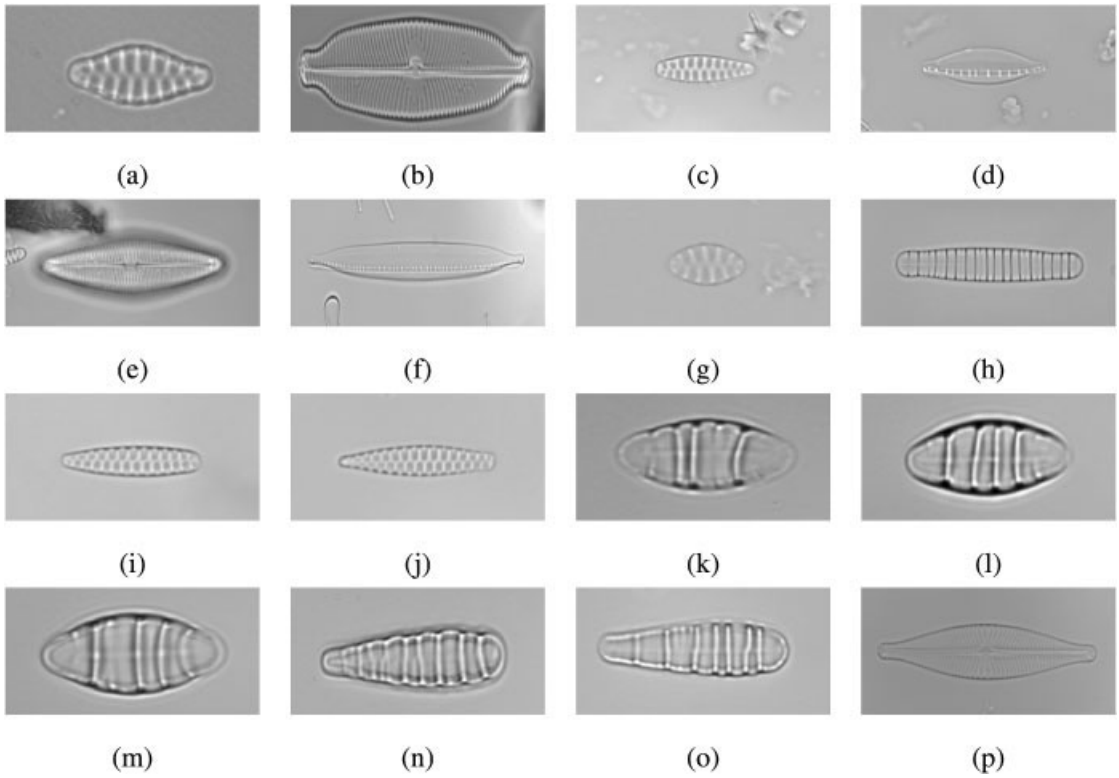

Fig. 15. Hybrid technique. Problematic images (**a–p:** 19 × 13, 61.2 × 10.5, 19 × 9.8, 20 × 7.6, 39 × 11, 81 × 8.2, 19 × 13.5, 55 × 8, 20 × 5, 20.5 × 5, 20.2 × 11, 20 × 11.2, 20.5 × 11.7, 25 × 5, 25 × 5, 73.7 × 13 μm).

TABLE 2. Corrected errors; (g), (h), (i), (p) refer to Fig. 15

Segmentation method	Corrected errors
Heuristic thresholding	(g), (h), (i)
Iterative thresholding	(g), (h)
RATS	(h), (p)
Canny edge detection	(g), (h), (i), (p)
Split-and-merge	(h), (i)
Clustering	—
Hierarchical watershed	(h), (i), (p)

The results of this experiment are shown in Table 2, referring to the images in Figure 15.

As can be seen from Table 2, only 4 errors can be corrected when the results of all other methods are combined, and all these are corrected only by the edge detection technique. The hybrid technique failed in these 4 cases because of the prominent ornamentation of the central diatoms, which led to wrong contours. None of the methods can extract good-quality contours from all images shown in Figure 15a–f. All these images have very low contrast, at least at one ending of their central diatoms (see Fig. 15b and f). The same observation holds also for the diatom shown in Figure 15j. The images shown in Figure 15k–m have both poor contrast at the endings of their diatoms and salient internal structures, and therefore all methods failed. Another poor contrast example is the image in Figure 15p.

This experiment shows that the combined results of all other methods cannot be much better than the result of the proposed hybrid segmentation technique alone. We conclude that the hybrid technique is indeed reliable and can be used for automatic segmentation of microscopic diatom images.

Identification Results

Next, we compared identification performance obtained using contours extracted by the methods in Fischer et al. (2002) and contours extracted by the hybrid technique proposed here. In Fischer et al. (2002), the results of edge detection were combined with those of heuristic thresholding, keeping the best contours.

The taxon (class) of each central diatom present in the input images is known. Both sets of extracted contours were used as input for a feature extractor based on morphological curvature scale spaces (Wilkinson et al., 2002). In our identification experiments, we have used the C4.5 decision tree classifier (Quinlan, 1993), and the identification performance was evaluated using 10-fold cross-validation (Kohavi, 1995). The mean performance (in terms of correctly classified samples) obtained using the hybrid method was 70.8 ± 1.5%. This compares favourably with the performance obtained using the contours extracted by the methods in Fischer et al. (2002) which was 64.5 ± 1.8, and is very close to the performance of 73 ± 1.6 obtained using manually adjusted contours (du Buf and Bayer, 2002).

DISCUSSION

In this report, we have proposed a framework for reliable automatic segmentation of microscopic diatom images, which improves initial segmentation results obtained within the ADIAC project (du Buf and Bayer, 2002; Fischer et al., 2002). Six segmentation techniques were analyzed covering all standard types, i.e., threshold techniques, boundary-based methods, region-based methods, and hybrid techniques. The best results were obtained by a new hybrid method based on the watershed segmentation, which substantially improves existing results. The novelty of this method is the computation and the selection of markers. Our marker selection uses two procedures, one that computes candidate marker regions based on connected operator filtering, and another one that selects the final markers based on the area of each candidate region, after some morphological post-processing is performed.

All segmentation methods were applied on a large database of diatom images, and the results were evaluated in two ways: (1) by visually inspecting the quality of the contours; (2) by comparing the identification performances using the C4.5 decision tree classifier. Contours extracted by the methods in (Fischer et al., 2002) as well as contours extracted by our new hybrid technique were used as input for a feature extractor based on morphological curvature scale spaces (Wilkinson et al., 2002). In the visual inspection, all techniques yielded more than 75% correctly extracted contours, while the best result was that of the hybrid technique, 98%. The mean identification performance using the hybrid method was 70.8 $\pm$ 1.5%, as compared to 64.5 $\pm$ 1.8% when using the contours extracted by the methods in Fischer et al. (2002), and 73 $\pm$ 1.6% when using manually adjusted contours project (du Buf and Bayer, 2002).

Finally, it should be mentioned that the proposed hybrid segmentation technique is fast. The CPU time on a Pentium III at 670 MHz to obtain the binary contour image (including post-processing) is around 3 seconds for typical image sizes of 800 $\times$ 300 pixels. Although thresholding methods are faster (around 2 seconds), the most computationally expensive part is the post-processing and contour extraction step. Therefore, our conclusion is that with the proposed hybrid method, the segmentation of diatom images can be performed automatically, efficiently, and with very good results. The hybrid method may be extended to segmentation of phase-contrast images of cells such as leukocytes, although more research on this subject is needed.

APPENDIX

A. Binary Mathematical Morphology

A binary image A in the N-dimensional Euclidean space E^n (where $E^n = \mathbb{R}^n$ or $E^n = \mathbb{Z}^n$) can be considered as a subset of E^n, i.e., $A \subset E^n$. The *translation* of A by a vector $x \in E^n$ is denoted by $A + x$ or A_x and is defined as

$$A + x = \{c \in E^n : c = a + x \text{ for a} \in A\}.$$

The *reflection* of A, denoted by $\hat{A}$ is defined as

$$\hat{A} = \{c \in E^n : c = -a \text{ for a} \in \text{A}\}.$$

The *erosion* of an input image A by a structuring element $B \in E^n$ is

$$A \ominus B = \{x : B + x \subseteq A\}.$$

Its dual, *dilation* is given by

$$A \oplus B = \{x : \hat{B}_x \cap A \neq \phi\}.$$

Opening of an input image A by a structuring element B is defined as

$$A \circ B = (A \ominus B) \oplus B,$$

while its dual *closing* is

$$A \bullet B = (A \oplus B) \ominus B.$$

B. Morphology for Grey-Value Images

All basic binary transforms can be defined as well for grey-scale images using the concept of *umbrae* (Serra, 1982; Heijmans, 1994). Then, the dilation becomes a maximum transformation over a neighborhood of a given radius, while erosions is the minimum transform over the neighborhood.

The *hat transforms* represent an important class of morphological transforms used for detail extraction from signals or images. Assume a grey-scale image f and a 2-D structuring element K. The residue of the opening compared to the original signal, i.e., $f - (f \circ K)$ represents the *top-hat* transformation. Its dual, the *bottom-hat* transform, is defined as the residue of a closing $f \bullet K$ compared to the original signal f.

The *morphological gradient* is defined as the difference between the dilated and the eroded images.

C. Connected Operators

Connected operators (Salembier and Serra, 1995) are characterized by the powerful property of preserving contours, and they only transform an image by selectively altering the grey values of connected sets of pixels. There are several ways of defining the notions of connectivity and connected operators. Here we shall follow the definitions of (Heijmans, 1999; Salembier and Serra, 1995).

Let E be the domain of a grey-scale image f. A *flat zone* L_h at level h of grey-scale image f is a connected component of the level set $X_h(f) = \{p \in E : f(p) = h\}$. A *peak component* P_h at level h is a connected component of the threshold set $T_h(f) = \{p \in E : f(p) \geq h\}$.

A flexible way of defining connected operators for functions is via partitions (Heijmans, 1999). A function $P : E \rightarrow \mathcal{P}$ is called a *partition* of E if (i) $x \in P(x)$, $x \in E$, and (ii) $P(x) = P(y)$ or $P(x) \cap P(y) = \phi$, for $x,y \in E$. In words, a partition is a subdivision of the underlying space into disjoint zones. Let P and P' be two partitions of E. Partition P is said to be *coarser* than P' (or P' is *finer* than P) if $P'(x) \subseteq P(x)$ for every $x \in E$.

Grey-level connected operators can be introduced if we define a partition associated to a grey-level function f. It can be shown (Salembier and Serra, 1995) that the set of flat zones of a function constitutes a partition of the domain of f. In the following, this partition will be called the *partition of flat zones* of f, and will be denoted by $C(f)$.

Definition 1: An operator γ acting on a grey-level function f is said to be *connected* if $C(\gamma(f))$, the partition of flat zones of $\gamma(f)$, is coarser than $C(f)$.

Thus, the only operations a connected operator can do are merging flat zones, and modifying their grey levels.

Definition 2: The *connected opening* $\Gamma_x(X)$ of a set of X at a point x is the connected component of X containing x if $x \in X$, and ϕ otherwise.

The binary area opening can now be defined as:

Definition 3: Let $X \subseteq E$ and $\lambda \geq 0$. The *binary area opening* of X with parameter λ is given by

$$\Gamma^a_\lambda(X) = \{x \in X: \mathbf{Area}(\Gamma_x(X)) \geq \lambda\}.$$

Definition 4: The *area opening* of a grey-scale image f is given by

$$(\gamma^a_\lambda(f))(x) = \mathbf{sup}\{h: x \in \Gamma^a_\lambda(T_h(f))\}.$$

Thus, the area opening of an image assigns each point the highest threshold at which it still belongs to a connected foreground component of area λ or higher, and it represents one of the simplest connected operators.

Attribute openings (Breen and Jones, 1996) are a general class of openings that allow the use of size criteria other than, for example, area (as used in area openings). In the binary case, attribute openings can be obtained by computing some increasing attribute (such as diagonal of the minimum enclosing rectangle, the area of the largest circle that can fit inside a region, etc.) of each foreground connected component, and removing the components for which the attribute is smaller than a given threshold. *Attribute thinnings* are a generalization of attribute openings to non-increasing criteria.

The reader is referred to Heijmans (1999), Salembier and Serra (1995), and Serra and Salembier (1993) for a complete background on connected operators, Breen and Jones (1996) for attribute morphology and Jones (1999), Meijster and Wilkinson (2002), and Salembier et al. (1998) for efficient algorithms to implement various connected operators.

REFERENCES

Adams R, Bischof L. 1994. Seeded region growing. IEEE Trans. Pattern Anal Mach Intell 16:641–647.

Beucher S. 1994. Watershed, hierarchical segmentation and waterfall algorithm. In: Serra J, Soille P, editors. Mathematical morphology and its application to image and signal processing. Dordrecht: Kluwer Academic Publishers. p 69–76.

Breen EJ, Jones R. 1996. Attribute openings, thinnings and granulometries. Comput Vision Image Understand 64:377–389.

Canny JF. 1986. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8:769–798.

Cohen LD. 1991. On active contour models and balloons. CVGIP. Image Understand 53:211–218.

Comaniciu D, Meer P. 1997. Robust analysis of feature spaces: color image segmentation. Proc IEEE Conf Comput Vision and Pattern Recognition, Los Alamitos, CA. p 750–755.

Davis LS. 1975. A survey of edge detection techniques. Comput Graph Image Process 4:248–270.

du Buf H, Bayer MM. 2002. Automatic diatom identification. Singapore: World Scientific Publishing.

Fairfield J. 1990. Toboggan contrast enhancement for contrast segmentation. In: Proc. 10th ICPR, Los Alamitos, CA. p 712–716.

Fischer S, Bunke H, Shahbazkia HR. 2002. Contour extraction. In: du Buf H, Bayer MM, editors. Automatic diatom identification. Singapore: World Scientific Publishing. p 93–107.

Fukunaga K, Hostetler LD. 1975. Estimation of the gradient of a density function with applications in pattern recognition. IEEE Trans Inf Th 21:32–40.

Grimaud M. 1992. New measure of contrast: the dynamics. In: Gader PD, Dougherty ER and Serra JC, editors. Proc SPIE, Bellingham, WA.

Haralick RM, Shapiro LG. 1985. Image segmentation techniques. Comput Vis Graph Image Process 29:100–132.

Haralick RM, Shapiro LG. 1992. Computer and robot vision. New York: Addison-Wesley.

Haris K, Efstratiadis SN, Maglaveras N, Katsaggelos AK. 1998. Hybrid image segmentation using watersheds and fast region merging. IEEE Trans Image Process 7:1684–1699.

Heijmans HJAM. 1994. Morphological image operators. Boston: Academic Press.

Heijmans HJAM. 1999. Connected morphological operators for binary images. Comput Vis Image Understand 73:99–120.

Horowitz SL, Pavlidis T. 1974. Picture segmentation by a directed split-and-merge procedure. In: Proc. 2nd Int Joint Conf Pattern Recognition, Los Alamitos, CA. p 424–433.

Horowitz SL, Pavlidis T. 1976. Picture segmentation by a tree traversal algorithm. J ACM 23:368–388.

Jackway P. 1995. Morphological multiscale gradient watershed image analysis. In: Proc 9th SCIA-IAPR, Singapore. p 87–94.

Jones R. 1999. Connected filtering and segmentation using component trees. Comput Vis Image Understand 75:215–228.

Kass M, Witkin A, Terzopoulos D. 1987. Snakes: active contour models. Int Jnl Comput Vision 1:321–331.

Kittler J, Illigworth J, Foglein J. 1985. Threshold selection based on a simple image statistic. Comp Vision Graph Image Proc 30:125–147.

Kohavi R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: International Joint Conference on Artificial Intelligence, San Francisco, CA. p 1137–1145.

Lotufo R, Falcao A. 2000. The ordered queue and the optimality of the watershed approaches. In: Goutsias J, Vincent L and Bloomberg D, editors. Mathematical morphology and its application to image and signal processing. Dordrecht: Academic Publishers. p 341–350.

Meijster A, Wilkinson MHF. 2002. A comparison of algorithms for connected set openings and closings. IEEE Trans Pattern Anal Mach Intell 24:484–494.

Meyer F, Beucher S. 1990. Morphological segmentation. J Vis Commun Image Repres 1:21–46.

Najman L, Schmitt M. 1994. Watershed of a continuous function. Signal Process 38:99–112.

Pal N, Pal S. 1993. A review of image segmentation techniques. Pattern Recog 26:1277–1294.

Pech-Pacheco L, Cristobal G. 2002. Automatic slide scanning. In: du Buf H, Bayer MM, editors. Automatic diatom identification. Singapore: World Scientific Publishing. p 259–288.

Quinlan JR. 1993. C4.5: Programs for machine learning. Morgan Kaufmann Publishers.

Ridler TW, Calvard S. 1978. Picture thresholding using an iterative selection method. IEEE Trans Syst Man Cybernet 8:630–632.

Roerdink JBTM, Meijster A. 2000. The watershed transform: Definitions, algorithms and parallelization strategies. Fund Informat 41: 187–228.

Sahoo PK, Soltani S, Wong AKC. 1988. A survey of thresholding techniques. Comput Vis Graph Image Process 41:233–260.

Salembier P, Serra J. 1995. Flat zones filtering, connected operators, and filters by reconstruction. IEEE Trans Image Process 4:1153–1160.

Salembier P, Oliveras A, Garrido L. 1998. Anti-extensive connected operators for image and sequence processing. IEEE Trans Image Process 7:555–570.

Serra J. 1982. Image analysis and mathematical morphology. New York: Academic Press.

Serra J, Salembier P. 1993. Connected operators and pyramids. SPIE Image Algebra Morphol Image Process IV 2030:65–76.

Torre V, Poggio TA. 1986. On edge detection. IEEE Trans Pattern Anal Mach Intell 8:148–163.

Vachier C, Meyer F. 1995. Extinction value: a new measurement of persistence. In: Proc of IEEE Workshop on Non Linear Signal and Image Processing, Piscataway, NJ. p 254–257.

Vincent L, Soille P. 1990. Watersheds in digital spaces: an efficient algorithm based on immersion simulation. IEEE Trans Pattern Anal Mach Intell 13:583–598.

Weickert J. 2001. Efficient image segmentation using partial differential equations and morphology. Pattern Recog 34:1813–1824.

Wilkinson MHF. 1998. Optimizing edge detectors for robust automatic threshold selection: coping with edge curvature and noise. Graph Models Image Proc 60:385–401.

Wilkinson MHF, Jalba AC, Urbach ER, Roerdink JBTM. 2002. Identification by mathematical morphology. In: du Buf JMH, Bayer MM, editors. Automatic eiatom identification. Singapore: World Scientific Publishing Co. p 221–244.