Automatic screening and multifocus fusion methods for diatom identification

Manuel G. Forero[a], Filip Šroubek[b], Jan Flusser[b], Rafael Redondo[a] and Gabriel Cristóbal[a]

[a] Instituto de Óptica (CSIC), Serrano 121, 28006 Madrid, Spain
[b] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, Czech Rep.

ABSTRACT

The first part of this paper presents a new method for the classification and screening of diatoms in images taken from water samples. The technique can be split into three main stages: segmentation, object feature extraction and classification. The segmentation part consists of two modified thresholding and contour tracing techniques in order to detect the majority of objects present at the sample. From the segmented objects, several features have been extracted and analyzed. For the classification, a diatom training set was considered and the centroids, means and variances of four different classes were found. For the identification process diatoms were classified according with their Mahalanobis distance. The results show the method ability to select at least 80% of usable diatoms from images contaminated with debris. Secondly, full automation of the diatom classification is achieved when multi-focal microscopy is utilized for water sample acquisition. In this case, a necessary preprocessing step is image fusion. A novel wavelet-based fusion method proposed here returns a sharp image that can be directly used for segmentation. For a better understanding of the diatom shape, a 2.5D reconstruction is given.

Keywords: diatom identification, pattern recognition.

1. INTRODUCTION

Diatoms are unicellular algae which have a size between 2 μ-m and 4 mm. They are present in anyplace with water and in general in places with enough degree of humidity. Their importance resides in the fact that they can be used in several research and scientific fields. For instance, they can be used to measure sediment PH, in forensic medicine, as water quality indicators, etc. They have been presented in Earth since Cretaceous age, and also due to their silica wall, the fossils are often well preserved in lake and marine systems. By these reasons they constitute a valuable tool for studies of fossil organisms and climatic history of Earth environment changes. Nowadays, they form large deposits of white chalky material, which is mined for its use in cleansers, paints, filtering agents and abrasives. Diatoms grow in different environments with limited conditions that, depending on species, become an indicator of pollution, water temperature, water depth, nutrient levels or salinity. In addition, they contribute around the 20% of the world's carbon fixation.[1] In a few words, they could determine the Earth evolution and their importance lie in their use as indicators of the ecosystem health.

The recognition and identification of diatoms is a tedious work. Each new image has to be match with those included in the main atlas. This classification process is complicated even for the experts, because there are hundreds of different taxa with many variations in shapes and biological characteristics. The diatom detection from the image processing point of view is even more complex. Beginning by the fact that images are not achieved from a controlled environment. This means that all diatoms present in the image could be rotated, overlapped or partially occluded by other objects.

In this paper, we present the results of a new diatom classification process from debris screening. The technique consists of three main steps. The first stage is based on a segmentation procedure by means of an optimal edge operator and a enhanced contour tracing technique. The second stage consists of the feature extraction based on morphological descriptors of diatoms. After analyzing the numerical representation for all the detected objects in the image, we found that descriptors based on momentum provide a better discrimination ability of diatoms. The last stage consists of the classification of the main diatom classes found where a threshold value must be selected to discriminate them from debris.

The goal of the particle screening process is to remove a substantial number of debris particles that are not required to be analyzed at further stages. However, at medium magnification it is difficult to elucidate which parts of the image are debris and which parts deserve to be analyzed. We explored different screening strategies described elsewhere.[2] The current methods provide an valuable tool for selecting only those candidate objects based on morphological descriptors. This is a very challenging problem because even for an expert in the field it is difficult to take a decision by inspection of such images.

The last part of this paper deals with multifocus fusion and 3D visualization of the surface of diatoms. A multiresolution approach is applied to stack of images, obtaining a decision map which is the central element both for fusion and for 3D visualization. This method provides a very valuable tool in the identification process of biological specimens.

2. MATERIALS

Different diatom samples from fresh water * were photographed on a Cohu 4910 CCD analog camera connected to a LG-3 frame grabber from Scion Co. that includes the NIH image processing shareware. The camera were attached to a Zeiss Axiophot photomicroscope illuminated with a 100W halogen light with 10x, 20x and 40X lenses. Images have been digitized with 8 bit/pixel and 256x256 pixel image format. The microscope slide was moved with a X-Y-Z motorized stage from Prior Instruments, with a step size of 0.1 μm for the X-Y axis and 1 μm for the Z-axis.

A training set consisting of 19 diatoms was used to find the clusters of the classes, and a dataset of 85 objects was used for testing. Only 25 of the test images correspond to diatoms and the remaining 60 correspond to debris (rejected class)

3. SEGMENTATION METHOD

3.1. Image preprocessing

The segmentation process consists of several steps. First, a gradient operator is employed to detect the borders in the image. Diatom segmentation is a very complicated task. The diatoms valves are transparent and present similar gray level intensities than the background. Hence, contrast between diatoms and background is very low, and the typical histogram of a diatom image is unimodal (see Fig. 2). Therefore, in most of the situations, at medium magnification only the edges of the diatoms provides a cue for distinguishing them from the background. In addition, images are intensity-variant and unevenly illuminated. The particle's screening is performed at medium magnification (e.g. 20X) (see Fig. 1).
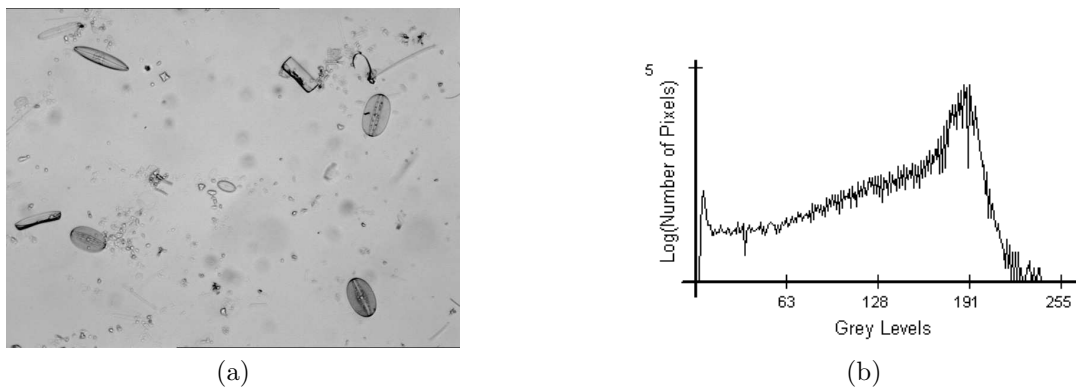

(a) (b)

Figure 1. a. A typical image of a diatom sample at medium magnification (e.g. 20X). Note the different shapes of diatoms and the difficulty to discriminate between diatoms and debris.**b.** Histogram plot corresponding to image **a.**

Several gradient based techniques were tested: the Roberts, Sobel and Prewitt masks and Canny, Deriche, Marr-Hildreth and Shen and Castan optimal filters. Best results were obtained with optimal filters, but detection depends on the image focus and diatom definition. Finally, Canny's operator was selected as the edge detection operator to be used with $\alpha = 1$. This step is followed by a non-maxima suppression and a hysteresis threshold.

Then the located regions were labelled and filtered according to the number of pixels of their contour, rejecting the objects whose contour were too small or too big to be considered as diatom. Because some structures can remain opened, a morphological closing operation was applied.
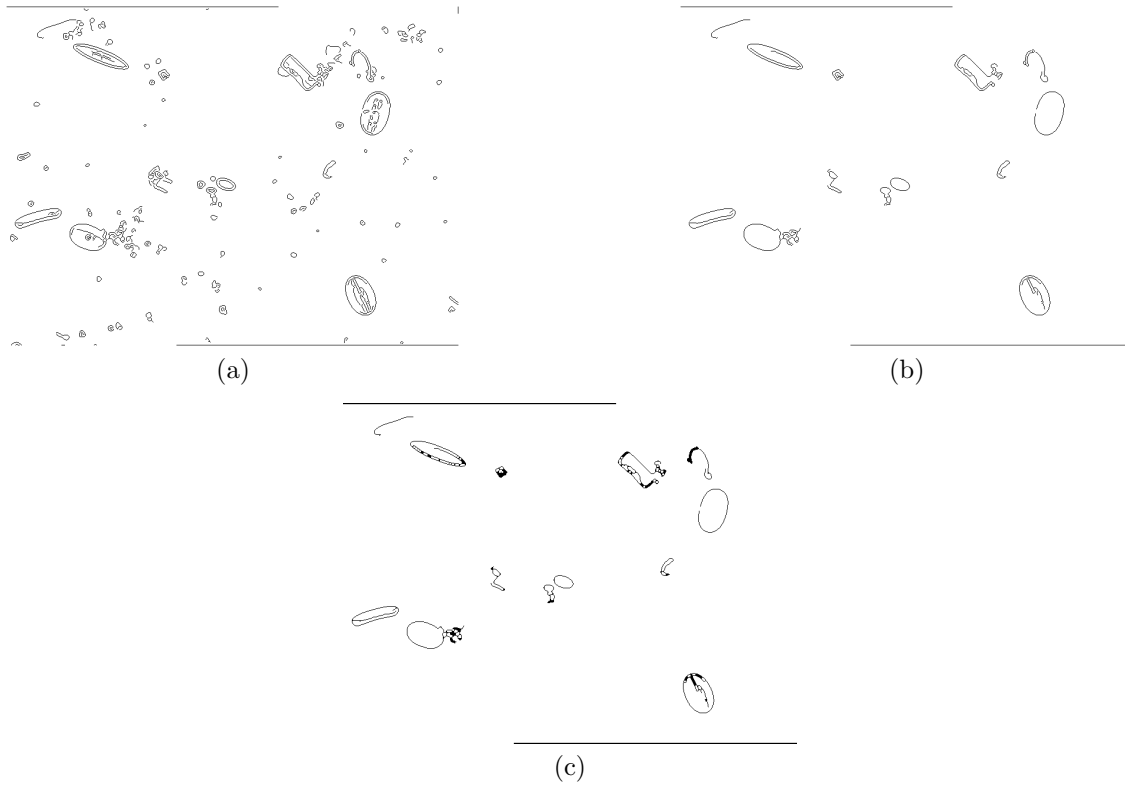

(a)

(b)

(c)

Figure 2. Segmentation.**a.** Canny detection. **b.** Size filter. **c.** Closing.

3.2. Contour following methods

In order to calculate several shape descriptors, a contour following method must be employed for finding the boundary of each object in the image. In order to obtain more accurate contours a 8-connectivity is selected. Two inner boundary tracing algorithms were tested: the "bug following"[3,4] and the "contour following" methods proposed by Sonka.[5] Because of 4-connectivity used by the bug method, pixels in corners can be skipped. Due to the order used by the Sonka's method to search for the next pixel in the contour, this method do not always track boundary pixels which will lead to errors. In addition, both algorithms are not suitable in this case because diatoms are not filled objects and in some cases the noise, even inside the segmented object, affects the tracking process.

Therefore, it was necessary to develop a new method using the neighborhood of the pixel contours. This algorithm, based in Sonka's contour tracking method (see Fig. 4), is simple and adapts its search direction according to the last contour transition. However the directions are scanned in a different order, so we can be sure that the next pixel always belongs to the border, which allows to obtain good results in all cases. The direction of searching is counter-clockwise sense, starting in the pixel placed in the direction $(dir+3)mod 8$ where dir is the direction of the previous movement along the border (see Fig. 4).
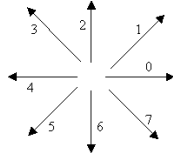
Figure 3. Direction notation.

In the new tracking method presented here, the image is scanned from the top left corner until a pixel P_0 of a new region is reached. Because it is known that there are not pixels above P_0 belonging to the region, the initial direction *dir* is taken in order to look for a new pixel and this is assigned to 0. The searching process stops when the two first pixels are found again. The results obtained through this contour tracking algorithm allows to obtain closed contours even though the shape is noisy and irregular. Fig. 5 shows an example of the of the new contour tracking algorithm. Compared with the Sonka's method, in this algorithm the rule for looking for the following pixel is always the same.
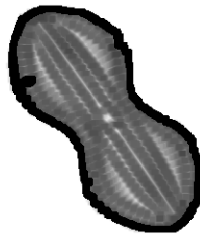
(a)

Figure 4. Example of contour tracing of a diatom image.

4. FEATURE EXTRACTION

Once each region and its boundary have been found, several descriptors were evaluated for the contour characterization: perimeter, area, convex area, compactness (C), major and minor axis length, eccentricity (E), solidity, the seven Hu's moments and the first ten normalized Fourier descriptors.

The abstract representation of the diatom must be preferably invariant against translations, rotations and scale changes in order to group the same species into the same class even if they have been rotated or scaled. In addition, because not all the descriptors present similar discrimination ability, they have been selected according with their mean and variance. A more in depth study will require the use of a Principal Components Analysis for extracting those descriptors that provide maximum variance. The perimeter, area, major and minor axis length and solidity turn out to not provide a good discrimination ability because their variances extend along the classification domain providing classs overlapping due to the diatom's size dependence. The Hu's moments are also invariant descriptors, that provide invariance to rotations, translation and scale. In this problem, it can be observed that the last five Hu's moments are very small, therefore only the first two moments are significant. Other invariant feature descriptors such as the normalized Fourier descriptors are good candidates to be considered but this will be the subject of a future research. In addition, the compactness feature can be taken into account because, although it is not really an invariant descriptor, it provides a good characterization of the shape. Therefore, a simple feature vector is constructed for each taxa composed by the first two Hu moments and the compactness.

5. MULTI-FOCAL FUSION

In multi-focal microscopy images of a 3D sample are acquired at different focal lengths. We assume that the surface of the sample is opaque. The distance of the surface from microscope lenses at a position (x, y) is given by

a function $z(x, y)$. The luminance of the surface ('original image) is denoted by $o(x, y)$. The obtained gray-scale images form a stack $\{g_i(x, y)\}$ and within the stack each image g_i corresponds to a unique focus length d_i and contains different parts of the sample in focus. The relation between the acquired image g_i and the original image o is expressed as

$$g_i(x, y) = H(o(x, y), z(x, y) - d_i) + n(x, y) \tag{1}$$

where H is a degradation operator that depends on the distance $z(x, y) - d_i = \lambda$ and acts on o, and n is white noise of zero mean and variance σ^2. We model the degradation operator as a space-variant convolution with the microscope PSF h that varies with the distance λ, i.e. $H(o, \lambda)(x, y) = \int h_\lambda(x - x', y - y')o(x', y')dx'dy'$. Further we assume that for every position (x, y) there exists at least one image g_i such that $z(x, y) - d_i = 0$, i.e. the location is observed undistorted in that image. This is approximately true if the minimum focal step used in the acquisition is equal to the depth of field of the microscope. The image indices of the undistorted parts reflects the depth map, since $z(x, y) \approx d_i$ in the undistorted section of g_i. The depth map is then approximated be a piece-wise constant function.

Image fusion consists of comparing the stack images in the image domain or in the wavelet domain, selecting an image in which the pixel (or a block) is depicted undistorted and finally of mosaicing the undistorted parts. To find the undistorted image for the given pixel, a local focus measure is calculated on its neighborhood and the image which maximizes the focus measure is chosen. Various focus measures were proposed for this purpose. Most of them are based on the fact that out-of-focus blurs primarily attenuate high spatial frequencies in an image. Image variance,[6] energy of a Fourier spectrum,[7] norm of image gradient,[6] norm of image Laplacian,[6] inverse image moments[8] and energy of high-pass bands of the wavelet transform[8–10] belong to the most popular focus measures. Numerous experiments published in the literature have demonstrated that the wavelet-based approach performs better and is therefore considered in the sequel.

The 2D wavelet transform is required for images but since the following reasoning holds also for a 1D case and to simplify the notation, the 1D wavelet transform is considered here. The dyadic wavelet decomposition is similar to subband filtering, where the wavelet function ψ at a scale j behaves as a high-pass filter $\hat{\psi}_j(\omega) = \hat{\psi}(2^j\omega)$ and the corresponding scaling function ϕ at the scale j acts as a low-pass filter $\hat{\phi}_j = \hat{\phi}(2^j\omega)$. The decomposition of the signal g at the scale j returns two bands $\phi_j * g(x)$ and $\psi_j * g(x)$. In the Fourier domain the decomposition takes the form of the low pass $\hat{\phi}(2^j\omega)\hat{g}(\omega)$ and the high pass $\hat{\psi}(2^j\omega)\hat{g}(\omega)$. By increasing the scale we get signals representing lower chunks of frequency. Let us assume that the surface $z(x, y)$ is constant on a neighborhood $U(x, y)$ and o^U denotes the original image $o(x, y)$ on this neighborhood. Using model (1), the energy of the low pass and high pass of g_i on U is then given by

$$E_L(g_i(U)) = \|\hat{\phi}_j(\hat{h}_\lambda \hat{o}^U + \hat{n})\|^2 \leq \|\hat{\phi}(2^j\omega)\hat{h}_\lambda \hat{o}^U\|^2 + \sigma^2 \frac{\|\phi\|^2}{2^j}$$

and

$$E_H(g_i(U)) = \|\hat{\psi}_j(\hat{h}_\lambda \hat{o}^U + \hat{n})\|^2 \leq \|\hat{\psi}(2^j\omega)\hat{h}_\lambda \hat{o}^U\|^2 + \sigma^2 \frac{\|\psi\|^2}{2^j} \, ,$$

respectively, where $\lambda = z(U) - d_i$. The second terms of both expressions are noise terms and they depend solely on the noise level and scale. Typically, the amplitude hull of out-of-focus blurs can be expressed as $|\hat{h}_\lambda(\omega)| \leq 1/(|\lambda|\omega)$. Since ψ acts as the high-pass filter, it is the first term in E_H that is primarily reduced by increasing λ and the image with the highest E_H should be the one in focus at U. However, if the underlying image $o(x, y)$ is smooth on U, i.e. its higher frequencies are negligible, this term is small compare to the noise term irrespective of λ and we cannot distinguish between a blurred and sharp image at this particular point. To reduce the effect of noise, a common approach is to perform the decomposition to higher scales and measure E_H at each scale. One drawback of this technique is that different decisions on which image is in focus may occur at different scales and some consistency check should be thus performed. The energy E_L is also reduced as λ increases (probably less than E_H) and since it corresponds to the low pass, it emphasizes low frequencies of the underlying image $o(x, y)$. Therefore in comparison with E_H, E_L is less affected but, on the other hand, consistently affected by the blur size λ. This leads us to a conclusion that a simple but efficient focus measure is a product of both energies:

$$m(g_i(U)) = E_L(g_i(U))E_H(g_i(U)) \, . \tag{2}$$

In two dimensions, the wavelet decomposition brakes into one low band and three high bands and thus we define E_H as the mean of the energies of the high bands. Contrary to the previously proposed methods,[8, 9, 11] not only the high pass but also the low pass is considered in the focus measure and only one level of the wavelet decomposition is required.

The proposed fusion of the multifocal stack $\{g_i\}$ is conducted as follows. First, we calculate the one-level wavelet decomposition of each image g_i. Then a decision map $D(x, y)$ is created in accordance with a rule $D(x, y) = \arg\max_i\{m(g_i(U(x, y)))\}$. The decision map is of size of the subband image, i.e., quarter of the original image, and it tells us from which image the wavelet coefficients should be used. The decision map is applied to all four bands and the inverse wavelet transform follows.

The decision map D can also be used for the reconstruction of the depth map z. Indeed, the piece-wise constant approximation of the surface is $\tilde{z}(2x, 2y) = d_{D(x,y)}$. For a better visual representation of the surface, higher-order surface reconstruction is often necessary. A partial 3D representation of the sample, also known as the 2.5D reconstruction, can be achieved by covering the approximated surface with the fused image.

6. RESULTS

6.1. Diatom screening

The feature descriptors above described were used in a classifier to determine the number of centroids, means and variances. Eight classification techniques were tested at this stage. Seven non-supervised clustering algorithms were tested in order to verify which one provides the best performance: adaptive, chain-map, ISODATA, k-means, max-min, sequence and similarity matrix. These approaches have different heuristics and parameters, allowing to compare the results in order to draw conclusions about the clusters and choose the best results. Four different types of shapes were selected for classifying diatoms: circular, elliptical, elongated and squared (see Fig. 6). These are the most frequent shapes present at the image samples, although this study can be extended to other diatom shapes. A set of 25 isolated diatoms was used for training, 8 samples belonging to class 1, 5 to class 2, 5 to class 3 and 7 to class 4. Then the clustering algorithms were used to validate the four classes selected. The results obtained with the algorithms were very similar, validating the correctness of the four types of classes.
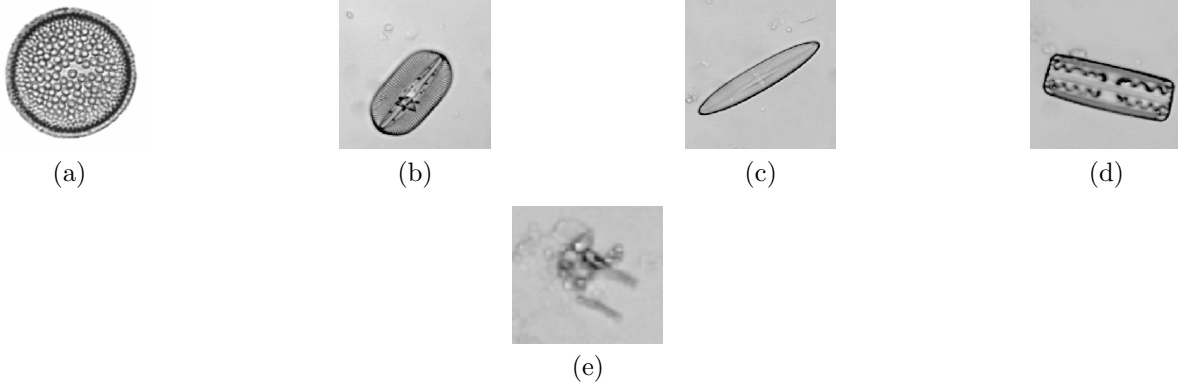

(a)　　　　　　(b)　　　　　　(c)　　　　　　(d)

(e)

Figure 5. Diatom's classes. **a.** Class 1: Circular. **b.** Class 2: Elliptic. **c.** Class 3: Elongated. **d.** Class 4 Square. **e.** Example of debris.

Classification is made by using a nearest-neighbor algorithm applying the Mahalanobis distance as similarity criterion, due to the magnitude of the descriptors and their variance are not similar. The independent probability distribution of the Hus moments and compactness are depicted in Fig. 7. It must be note that debris expands along all the classification domain.

Table 1 shows the experimental results where two types of classification were evaluated: the first one using only the diatom images (shown in column 2) and the second one using all the set of images (diatoms and debris)
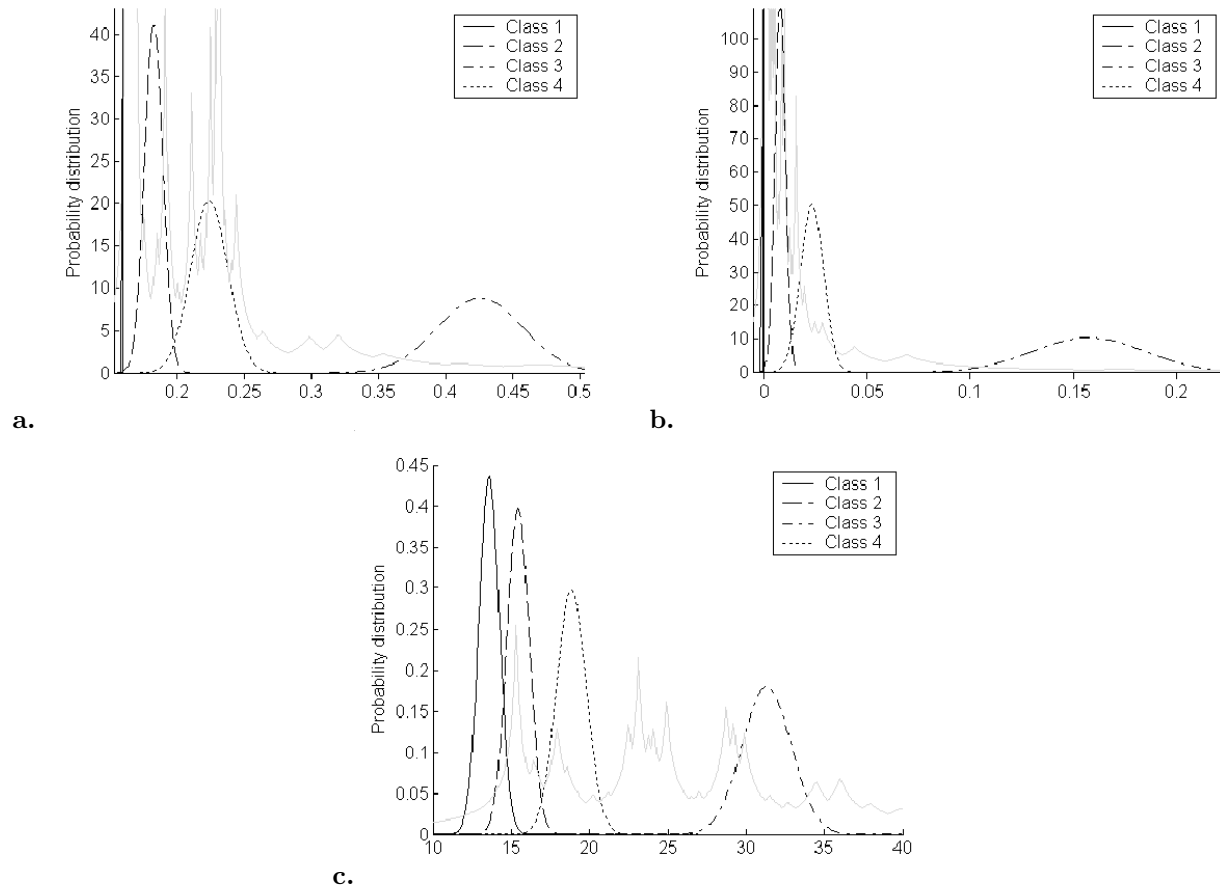
Figure 6. Independent probability distribution of descriptors. **a.** Hu 1. **b.** Hu 2. **c.** Compactness.

(column 3). In the last case, a fixed threshold value must be chosen to discriminate between diatoms and debris. Two different vectors of descriptors were employed for testing as can be noted in the table. The incorrect identification is due to the presence of noise and structures in the images that provides characteristic vectors that are near the clusters centroids of the class.

Table 1. Diatom from debris screening results

Mahalanobis	Diatoms	Diatoms+Debris
Hu1, Hu2	$96\% \pm 7.6$	$77.6\% \pm 8.8$
Hu1, Hu2, Compactness	$80\% \pm 15.6$	$82.35\% \pm 8.1$

6.2. Visualization of the diatom 3D surface

The proposed image fusion method was tested on multi-focal images of a diatom. The stack contained 70 images acquired with the focal step 0.1μm. Fig.7 depicts few images from the stack. The wavelet type and the size of the neighborhood U are the only two parameters that the method requires. We have conducted several experiments and found that the size of decomposition filters used in the discrete wavelet transform and the size of neighborhoods have a negative impact on the precision of the decision map. On the other hand, extremely

small neighborhoods do not provide enough data for the focus measure and short wavelets are badly localized in the frequency domain and therefore perform poorly as high-pass filters. The best results, according to our visual assessment, were obtained for the biorthogonal spline wavelets 2.2 (the decomposition filters were 5 pixels long) and the neighborhood of 9×9. The calculated decision map is in Fig. 8a and the corresponding fused image is in Fig. 7e. Fig. 8b shows the mask used for extracting the diatom's region. Fig. 8c shows the 2.5D reconstruction of the diatom with the surface shape obtained from the decision map and the fused image as the texture.
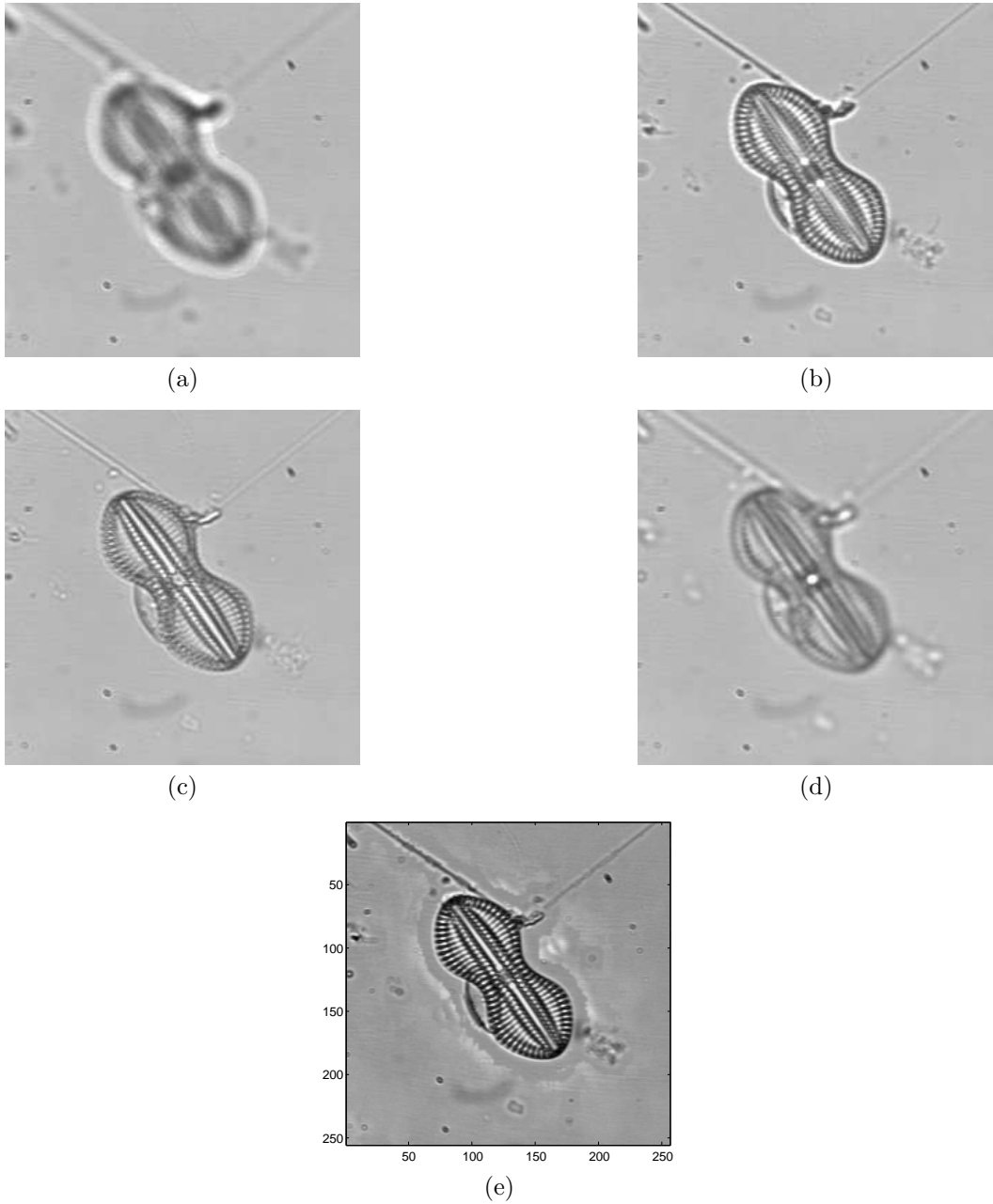

(a)

(b)

(c)

(d)

(e)

Figure 7. **a-d.** Examples of four images from the multifocus stack **e.** Fused image.
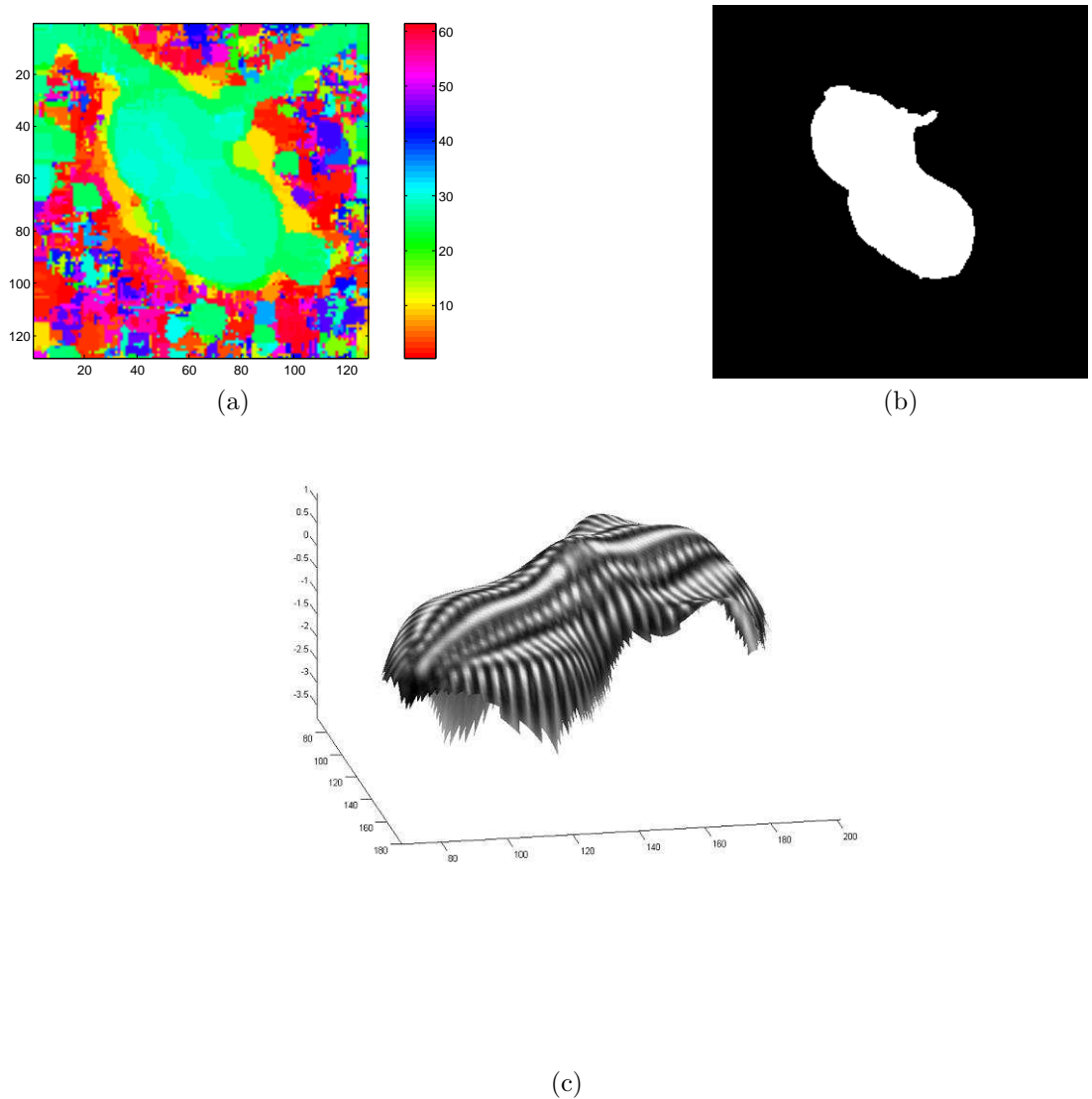
(a)

(b)

(c)

Figure 8. **a.** Decision map (D). **b.** Mask filter **c.** Visualization of the 3D surface of the diatom

7. CONCLUSIONS

A new technique was developed for screening diatom from debris. Based in this study the automatic detection of diatoms in water samples appear to be feasible and practical.It could enhance diatom screening techniques. The method allows to recognize and classify diatoms from debris by using features extracted from the objects. The method was tested for detecting diatoms belonging to four different taxa characterized by four different shapes. Future research will include the use of learning techniques for producing new centroids of new classes when new diatom shapes come through the system. Also a new technique for diatom 3D reconstruction was presented. It's based in a multiscale decomposition of a multifocus stack of images followed by the calculation of a decision map. The decision map plays a central role both in the subsequent fusion process of the multifocus images and also because it can be interpreted as a depth map. This allows to render the 3D surface of the diatom, facilitating a better visualization of the structural details of the organism and therefore its identification.

8. ACKNOWLEDGEMENTS

REFERENCES

1. C. Field, M. Behrenfeld, J. Randerson, and P. Falkowski, "Primary production of the biosphere: integrating terrestrial and oceanic components," *Science* **281**, pp. 237–240, 1998.
2. J. L. Pech and G. Cristóbal, *Automatic Diatom Identification*, ch. Automatic Slide Scanning, pp. 259–288. World Scientific, 2002.
3. W. Pratt, *Digital image processing*, John Wiley and sons, second ed., 1991.
4. D. Paulus and J. Hornegger, *Applied pattern recognition: A practical introduction to image and speech processing in C++*, Morgan Kaufmann, 1998.
5. M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*, PWS publishing, 2nd ed., 1999.
6. T. C. M. Subbarao and A. Nikzad, "Focusing techniques," *Opt. Eng.* **32**, pp. 2824–2836, 1993.
7. M. Subbarao and J. Tyan, "Selecting the optimal focus measure for autofocusing and depth-from-focus," *IEEE Trans. Pattern Analysis and Machine Intelligence* **20**, pp. 864–870, 1998.
8. Z. Zhang and R. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE* **87**, pp. 1315–1326, 1999.
9. B. M. H. Li and S. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Model and Image Processing* **57**, pp. 235–245, 1995.
10. B. Z. J. Kautsky, J. Flusser and S. Šimberová, "A new wavelet-based measure of image focus," *Pattern Recognition Letters* **23**, pp. 1785–1794, 2002.
11. Y. Chibani and A. Houacine, "Redundant versus orthogonal wavelet decomposition for multisensor image fusion," *Pattern Recognition* **36**, pp. 1785–1794, 2003.