Phase congruency-based detection of circular objects applied to analysis of phytoplankton images

A. Verikas, A. Gelzinis, M. Bacauskiene, I. Olenina, S. Olenin, E. Vaiciukynas

1. Introduction

Assessment of water quality parameters, studies of long-term changes in aquatic ecosystems, monitoring of toxic algal blooms are some examples, where identification and counting of plankton cells is being used. Much work in this area still remains in the form of conventional microscope analysis and is very time consuming and labor intensive. For example, derivation of quantitative abundance estimates requires recognizing and counting cells in tens or even hundreds of microscopic views. A robust automated image analysis-based system capable of recognizing different plankton species would be of great help and would enable analysis at much larger scales [1]. Benfield et al. present an overview of recent developments and challenges in this area [2].

Different systems are usually designed for automated analysis of zooplankton and phytoplankton images. Objects to be detected in zooplankton images are usually smaller than those found in zooplankton images. Fig. 1 presents two examples of phytoplankton images containing three decision classes: invasive (Prorocentrum minimum) (an example of this class is enclosed in a rectangle), native (enclosed in a circle), and other (enclosed in a pentagon). This work is concerned with automated analysis of such type of phytoplankton images.

There have been a number of successful developments in the area of automated analysis of zooplankton images. Davis et al. [3] and Hu and Davis [4] developed several techniques for automated analysis of plankton images obtained from a video plankton recorder. In [5,3], image analysis procedures applied included in-focus object detection, object feature extraction, and object classification. In total, 237 features were extracted, including shape factor, seven invariant moments, Fourier boundary descriptions, granulometric curves [6], and ratios computed from dimensions of the bounding box. The first 20–30 principal components were then used for object classification based on learning vector quantization (LVQ) [7]. The overall classification accuracy using seven classes was about 61%. Hu and Davis have elaborated upon the system presented in [3]. They used co-occurrence matrices [8,9] computed for distances $d = 1, 4.8$, and 16 pixels to extract features and a support vector machine (SVM) [10] for object classification. The system was verified using 20 000 plankton images and the overall classification accuracy of 71% was achieved using seven categories. Grosjean et al. developed the ZooScan digital imaging system for automated analysis of zooplankton images and...
RF are also implemented in the ZooImage software [16]. When tested in a zooplankton classification task with 63 categories, RF showed the best performance. RF was also the best in the task with 17 categories [17]. Based on their previous experience, Hu and Davis proposed using a sequential classifier consisting of a neural network trained on shape-based features followed by an SVM exploiting texture features [18]. Fernandes et al. [19] studied the issue of selecting the appropriate number of classes (as a trade-off between the number of classes identified and the accuracy) in automated zooplankton classification.

Attempts to develop techniques for classification of binary zooplankton images have also been made. Luo et al. [20] tested several classifiers including SVM, RF, C4.5 trees, and the cascade correlation neural network in the task of categorizing binary zooplankton images into six and seven classes. SVM provided the highest classification accuracy equal to 90% and 75% for the six and seven classes, respectively. Moment invariants of the original image, moment invariants of the contour image after closing, granulometric features, and domain specific features, such as object size, convex ratio, eigenvalue ratio were used. Zhao et al. [21] augmented the aforementioned set of features with additional measurements such as circular projections, boundary smoothness, object density, moment ratios, and some other geometric features. To reduce the dimensionality, principal component analysis was applied. A committee, designed using the random subspace approach, was used to make a decision. The 93% classification accuracy was achieved, when distinguishing between seven classes of objects. A very similar approach was taken by Tang et al. [22].

While automated analysis of zooplankton images is a rather active field of research, developments in the field of automated analysis of phytoplankton images are rather limited. Work by Gorsky et al. is one of the pioneering attempts in this area [23]. Using simple geometric features the authors were able to distinguish between three species of distinct size and shape. Sosik and Olson developed a system for automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry [24]. Embleton et al. [25] trained a multilayer perceptron (MLP) to identify four species in lake water samples. Fourier descriptors, geometrical features, and features characterizing the grey level distribution in a region were used to select a set of suitable features for training the MLP. Cuiping et al. by using a variety of moment features in a support vector machine (SVM) with a radial basis function kernel were able to distinguish between 241 species of marine phytoplankton with 89% accuracy [26]. Blaschko et al. [27] achieved 50–70% classification accuracy in a task of phytoplankton categorization into 12 classes plus an “unknown” class. A large variety of features: shape features, moments, texture features, contour features (780 features in total) were used. Several classifiers, including decision trees, naive Bayes, ridge linear regression, k-NN, SVM, and bagged as well as boosted ensembles were explored. SVM was found to be the best classifier for the task. Culverhouse et al. [28] studied the classification accuracy achieved by the neural network committee-based automated system DiCANN [29] and argued that accuracy of about 72% achieved by the system in a six-class phytoplankton categorization task was similar to the accuracy achieved by the trained personnel. To capture more information for discrimination between five classes of phytoplankton, Rodenacker et al. applied fluorescence imaging in their image acquisition system [30]. Sosik and Olson [24] presented, perhaps, the most elaborated study regarding multi-class phytoplankton categorization using data obtained from Imaging FlowCytobot [31]. A combination of video and flow cytometric technologies is used in the Imaging FlowCytobot. Images of 1380 × 1034 pixels and resolution equal to 3.4 pixels/μm were used in the study. A set of 6600 visually identified and manually inspected images distributed across 22 categories was used in the study. In total, 210 features characterizing geometry, shape, symmetry, texture, and invariant moments of objects were extracted and 131 features were selected and used in an SVM classifier for the categorization. Ten-fold cross validation on the training set was applied to select the hyper-parameters of the classifier. The overall accuracy of 88% was achieved on the test set. The abundance estimates were corrected for classification errors according to the procedure suggested by Solow et al. [32]. While the obtained accuracy is very encouraging, one very important task – object detection – is not addressed in the article.

Analysis of the literature shows that classification accuracy achieved when solving phytoplankton classification problems varies in a broad range depending on the task and the data. Due to the variety of tasks considered and data sets used, comparison of different approaches is a rather complicated matter. It is worth noting that one very important problem, namely object detection, is almost never addressed in the literature. Robust object detection, however, is a prerequisite step if an accurate tool for automated analysis of images is desired. Especially this is the case when rather simple imaging systems are employed and objects in resulting pictures appear contiguous or overlapped, as exemplified in Fig. 1. Touching or overlapping organisms cause
difficulties in automated categorization and abundance estimation of the species. Irigoien et al. point out that a linear relation between the number of items and the automatic counting holds if the percentage of image area occupied by the items remains below 3%. Above this threshold, automatic counting underestimates abundance due to increased percentage of organisms touching each other [17].

In contrast to previous techniques, a simple imaging system is used to obtain phytoplankton images in this study. A long term goal of this work is an automated system for detection, recognition and abundance estimation of objects representing different phytoplankton species. Such automation can essentially facilitate treatment of phytoplankton samples and can be useful in many practical applications (e.g. rapid ship ballast water analysis, harmful algae bloom monitoring, etc.). This article is limited, however, to analysis of one invasive species, a dinoflagellate P. minimum (Pavillard) Schiller (P. minimum), which is known to cause harmful blooms in many estuarine and coastal environments [33–35]. A new technique for automated analysis of images with high percentage of image area occupied by objects was developed. The technique combines phase congruency-based detection of circular objects, stochastic optimization, image segmentation, and SVM— as well as RF-based classification.

2. Data

Phytoplankton samples used for obtaining images were received from: (1) natural south-eastern Baltic Sea phytoplankton containing P. minimum cells, (2) cultured P. minimum; and, (3) natural phytoplankton mixed with cultured P. minimum. All samples were fixed with acetic Lugol’s solution (in proportion: 0.5 ml of solution for 100 ml of sample). Cells of P. minimum in natural phytoplankton varied in shape from triangular- to oval- and heart-shaped, while cultured cells were mostly oval. The length of P. minimum varied from 14 to 22 μm, while the width ranged from 12 to 18 μm.

2.1. Images

Images for the analysis were obtained from a colour camera of 1280 × 960 pixels attached to an inverted microscope with magnification of 400 ×. Images were recorded in the RGB colour space and then converted to the more homogeneous Lab colour space [36] for further analysis. Presence of pixels in a central part of cells (both invasive and native) with B component values significantly lower than in the background is one characteristic feature of the images. Otherwise, information on colour does not provide much help for the analysis. The background colour in the images may differ significantly depending on the solution applied in the sample preparation process. Such variation can easily be noticed in images of Fig. 1. Three decision classes, namely the invasive (an example of this class is enclosed in a rectangle), native (enclosed in a circle), and other (enclosed in a pentagon) are identified in Fig. 1. Invasive (P. minimum) cells dominate the image shown on right-hand side of Fig. 1. In total, 114 images have been collected for the analysis at several occasions.

3. Methods

3.1. Image preprocessing

A clear 2D shell is a characteristic feature of the invasive cells. To emphasize this feature, we applied preprocessing via the phase congruency-based enhancement of image edges [37]. The phase congruency idea is based on the assumption that features (edges, corners) are perceived in image points, where Fourier components of the signal are maximally in phase. The purpose of such Fourier analysis is to give local frequency information. As suggested by Kovesi [37], we obtain this information by applying a bank of Gabor filters tuned to different spatial frequencies rather than performing direct Fourier analysis. When working with 2D images, local energy is first computed in several orientations θ using 2D Gabor filters [37] and 2D phase congruency. PC(x) is then computed according to the following equations:

\[
PC(x) = \sum_{\theta} \sum_{\omega} W_{\omega}(x) A_{\omega}(x) \Delta \Phi_{\omega}(x) - T_{\theta} \frac{1}{\sum_{\theta} A_{\omega}(x) + \varepsilon}
\]

\[
\Delta \Phi_{\omega}(x) = \cos(\phi_{\omega}(x) - \Phi_{\omega}(x)) - \sin(\phi_{\omega}(x) - \Phi_{\omega}(x))
\]

where the index θ indicates orientation, A_{\omega}(x) and φ_{\omega}(x) is an amplitude and a phase angle, respectively, of the nth frequency component at location x and orientation θ, Φ_{\omega}(x) is the amplitude weighted mean phase angle at orientation θ, and \( |y| \) = y if y > 0 and |y| = 0 otherwise. The appropriate T_\theta values are usually obtained from the responses of the filters applied to the image. Resulting values of the PC(x) measure belong to the interval [0, 1]. If all Fourier components are in phase, PC(x) approaches 1, and if there is no coherence in phase PC(x) approaches 0. Since the measure is independent of the signal magnitude, the measure is invariant to variations in contrast and/or illumination. This is a very important property.

We use a PC image to detect objects in the original phytoplankton image. In addition to the information on phase congruency available from a PC image, we also utilize information of phase congruency variation with orientation. This information is obtained from an angle image Φ (an angle of the principal axis about which the phase congruency moment is minimized [38]) and an image of the magnitude of the maximum moment \( M(\theta) \) (the moment about an axis perpendicular to the principal axis [38]) and the minimum moment m of phase congruency. Values of Φ, M, and m computed in each image pixel x, are given by [38]

\[
\Phi(x) = \frac{1}{2} \arctan \left( \frac{b}{\sqrt{b^2 + (a-c)^2}} \right)
\]

\[
M(x) = \frac{1}{2} \left[ c + a + \sqrt{b^2 + (a-c)^2} \right]
\]

\[
m(x) = \frac{1}{2} \left[ c + a - \sqrt{b^2 + (a-c)^2} \right]
\]

where \( a = \sum_{\theta} |PC_{\omega}(x)\cos(\theta)|^2 \)

\( b = 2 \sum_{\theta} |PC_{\omega}(x)\cos(\theta)||PC_{\omega}(x)\sin(\theta)| \)

\( c = \sum_{\theta} |PC_{\omega}(x)\sin(\theta)|^2 \)

where \( PC_{\omega}(x) \) is the phase congruency value computed at orientation \( \theta \)
3.2. Determining centre points of circular-shaped objects

P. minimum cells are approximately circular- or oval-shaped, see Fig. 1. To facilitate the cell detection step, a technique for determining centre points of circular objects was developed, based on M and Φ images. One can consider image M as an image of edge clarity (certainty) values in each pixel of the original image, while image Φ reflects edge direction in each image pixel. Fig. 2 presents an example of an original image and corresponding M and Φ images. Image C reflecting clarity (certainty) of a centre point in each image point is created first. The size of C is equal to the size of the original image, m x n. We start by setting C = 0, ∀i, j. Image C is then obtained by applying the following algorithm, where r is the supposed radius of an object.

\[
\text{for } i = 1, \ldots, m \\
 \text{for } j = 1, \ldots, n \\
 t^+ = i + r \sin(\Phi_y + 90), \quad j^+ = j + r \cos(\Phi_y + 90) \\
 t^- = i - r \sin(\Phi_y + 90), \quad j^- = j - r \cos(\Phi_y + 90) \\
 \text{if } 0 < t^+ \leq m \text{ and } 0 < j^+ \leq n \\
 C_{ij}^+ = C_{ij}^+ + M_y \\
 \text{end (if)} \\
 \text{if } 0 < t^- \leq m \text{ and } 0 < j^- \leq n \\
 C_{ij}^- = C_{ij}^- + M_y \\
 \text{end (if)} \\
\text{end (for)} \text{ end (for)}
\]

Fig. 3 illustrates the generation process of C based on information available in Φ_y and M_y. The idea is similar to that used in the Hough transform. However, the technique we developed is much more robust to noise. Based on information available in Φ_y, a normal to the edge is determined and the centre certainty values C_i,j and C_i,j^+ are increased by M_y at two possible centre positions, [t, j] and [t^+, j^+]. For the case shown in Fig. 3, many edge pixels contribute to the centre position at [t^+, j^+]. Fig. 4 presents the C image computed from M and Φ images shown in Fig. 2 using r=45. As can be seen from Fig. 4, the approximate centre position manifests itself by high C_i,j values.

Since shape of P. minimum cells deviates from a circle, many pixels in the vicinity of the centre position exhibit high C_i,j values. Moreover, P. minimum cells differ in size. Therefore, C images are filtered by a rotationally symmetric Gaussian low-pass filter with the standard deviation σ = 3. Fig. 5 presents five examples of a filtered C image computed from M and Φ images shown in Fig. 2 for five different values of r: 35, 40, 45, 50, and 55. As can be seen from Fig. 5, the C image is rather insensitive to the r value. When varying r from 40 to 50, rather similar C images, with a clear centre position, are obtained. Thus, using a given r value, centre positions of objects of radius differing from r in a rather broad range, are determined quite accurately. It is worth mentioning that robust contour determination of an object, influences the determination accuracy of a centre position of the object to a greater extent than selection of an exact r value.

3.2.1. Comparison to the Hough transform

We used the Hough transform to compare the circular objects detection robustness. To assess the detection robustness of a circular object centre position, we computed a ratio, ρ_m and ρ_a, between the maximum/average C image intensity value at the centre position and the maximum/average intensity value at locations do not representing circular objects

\[
\rho_m = \frac{\max_{(i,j)\in K} C_{ij}}{\max_{(i,j)\notin K} \frac{1}{N} \sum_{(i,j)\in K} C_{ij}}
\]

\[
\rho_a = \frac{\max_{(i,j)\in K} C_{ij}}{\max_{(i,j)\notin K} \frac{1}{N} \sum_{(i,j)\in K} C_{ij}}
\]

where K_o and K_n stand for neighbourhood around an object centre position and neighbourhood around a non-centre position, respectively, and N is the neighbourhood size. We used circular neighbourhoods of 8 pixels radius in the evaluation. The same measures were applied for accumulator values of the Hough transform. Before computing the measures, values of both C images and accumulators of the Hough transform were normalized to the [0, 1] interval. Fig. 6 (top) presents a phytoplankton image example containing four P. minimum cells that are to be detected. Fig. 6 (middle) displays the normalized C image obtained
from the technique proposed, while Fig. 6 (bottom) shows the normalized accumulator image obtained from the Hough transform. In this experiment, we set $r=45$ for the phase congruency-based technique, while the optimal interval of $r$ values for the Hough transform was found experimentally and was equal to $[42, 48]$.

Centre positions of the four $P. minimum$ cells manifest themselves by higher intensity values inside the rectangles 1, 2, 3, and 4. The maximum normalized intensity value found in each rectangle is written above the corresponding rectangle. In Fig. 6 (middle) and (bottom) are also shown three rectangles, 5, 6, and 7, enclosing the highest intensity values that do not represent centre positions of the $P. minimum$ cells. The shown maximum intensity values reveal that all the four centre positions can be correctly detected by simple thresholding using the $C$ image, while the Hough transform accumulator image allows detecting only two centre positions. The average $\overline{I}_y$ value computed for the four centre positions of the cells in the $C$ image is $\overline{I}_y^C = 2.74$, while the average $\overline{I}_y$ value for the Hough transform accumulator image is $\overline{I}_y^H = 1.73$. Thus, higher centre position detection robustness is obtained when using the phase congruency-based technique, if compared to the Hough transform. A similar level of difference between the measures has also been obtained when analyzing other phytoplankton images for detection of $P. minimum$ cells as well as images from other domains. For example, values of $\overline{I}_y^C = 3.79$ and $\overline{I}_y^H = 2.36$ were obtained when analyzing the widely used image of coins, shown in Fig. 7.

3.3. Determining contour of circular objects

Having a centre position of a circular-shaped object, a square of size $(2r + \beta) \times (2r + \beta)$, where $\beta$ is a parameter, is clipped out around the centre together with identical rectangle area of $M$ image for further analysis, see Fig. 8. The analysis is based on the assumption that an approximate object radius $r$ is known and that an approximate object contour is given by the brightest line in the corresponding part of $M$ image transformed to the polar coordinate system, using the determined centre position. Fig. 9 presents the $M$ image shown in Fig. 8 transformed to the polar coordinate system, where the angle $\alpha$ is computed clockwise starting at “3 o’clock” and bottom corresponds to the centre position in the original image. To emphasize image elements making horizontal lines and to suppress isolated image elements, $M$ image in the polar coordinate system is filtered by a Gaussian filter with a much larger standard deviation in the angle direction than in the $r$ direction.

An iterative algorithm is then used to find a contour line $c$. The algorithm aims at placing a contour line in high intensity points of a filtered $M$ image $E$. A contour line $c$ is given by a set radius values – one value each angle $\alpha$ from a set of given angle values – and determines a row index in each column of $E$. The optimization
Values of in the neighborhood of a c at the best step size given by the largest r object in the polar coordinate system (top denoted by ns for G where version of this article.)

For each a, the initial contour position c^0, a = 0,...,360°, is given by the largest r value, such that E_{a,r} > \text{mean}(E_{a,r})

c^0 = \max(\arg(E_{a,r} > \text{mean}(E_{a,r})))

(12)

where E_{a,r} is an image E element—intensity at angle a and radius r. At each iteration, the contour line c is updated by Δc

\[ c^{i+1} = c^i + \Delta c, \quad \Delta c = B_{\text{step}} G \]

(13)

where i denotes an iteration index, B_{\text{step}} stands for a vector of best step sizes at different a and the convolution operation, denoted by * s, with a Gaussian filter G is applied to make adjustments at different a dependent on neighbours. At i th iteration, the best step size B_{\text{step}} at the randomly selected angle a' and using the Gaussian window of randomly selected width w' is given by

\[ B_{\text{step}} = \arg \max_{\text{step}} \sum_{j = w'}^{j - w'} E_f c'_{j} \cdot \text{step} G(j) \]

(14)

where G is a Gaussian window with maximum at G(0) and \( f = a' + j \). If B_{\text{step}} ≠ 0, contour adaptation takes place at a' and in the neighbourhood of a', given by \( a' \pm w' \)

\[ c_{j'}^{i+1} = c_{j'}^i + B_{\text{step}} \cdot G(f - a'), \quad f = a' - w', \ldots, a' + w' \]

(15)

The width of the Gaussian window w' is randomly selected from the interval \( w_{\min} = 20 \) and \( w_{\max} = 100 \). The search continues for \( n_s \) successful iterations, B_{\text{step}} ≠ 0, or for \( n_s \) iterations in total. Values of \( n_s = 40 \) and \( n_s = 1000 \), found experimentally, worked well in all the tests. Fig. 10 presents an example of a filtered object in the polar coordinate system (top), (middle): the initial contour line (yellow) along with the contour line determined by the algorithm (green-smooth), and (bottom): the contour line after 5, 7, 15, and 25 iterations. As can be seen from Fig. 10, the learning process converges rather fast and after a small number of iterations the determined contour line conforms to the perceived one quite well.

3.4. Segmenting original images

To detect objects, not only phase congruency preprocessed images, but also original colour images were used. Since the RGB colour space is rather nonhomogeneous, original RGB images were first transformed to the Lab colour space and then used in the segmentation process, based on the Fuzzy C-Means clustering (FCM) algorithm [39]. In fact, segmentation into two clusters, namely “objects” and “background” is sufficient in this study. However, aiming to obtain more accurate object boundaries, a larger number of clusters was set for the FCM algorithm. Then, in the next step, clusters were merged and too large regions, representing the background, were eliminated. Fig. 11 presents an original image (left), the segmentation result obtained using five clusters (middle), and the result obtained using two clusters (right). It is easy to see that the middle image reflects objects seen in the original image better than the one shown on the right.

3.5. Combining analysis results for object detection

It is well known that fusing results obtained from different analysis techniques is an efficient way to improve the accuracy of the analysis [40,41]. Therefore, the techniques developed to find centres of circular objects, to determine contour lines of such objects, and to segment an image into objects and background...
were combined into one algorithm. The algorithm can be summarized by the following steps:

i. create an image of objects $O$ by segmenting an $Lab$ image;

ii. determine centres of circular objects;

iii. determine contour lines of the circular objects;

iv. eliminate pixels belonging to the circular objects from $O$; this step ends phase I of the algorithm.

v. eliminate pixels belonging to very small and very large objects from $O$; objects much smaller than $P. \text{minimum}$ cells are considered as being very small;

vi. if the image $O$ contains at least one object, take an object from $O$ and perform the following steps (phase II):
(a) determine a centre point of the object; if the object is larger than twice the average $P. \text{minimum}$ cell size, the most distant point from the object edge is assumed to be the centre point;
(b) determine a contour line of the object;
(c) eliminate pixels of the object from $O$;
(d) apply the morphological opening operation on the image $O$;
(e) eliminate very small objects from $O$;
(f) if $O$ contains at least one object goto Step (vi); otherwise stop.

Observe that we do not consider very small and very large objects. Objects with centre points too close to image boundaries are also left unprocessed.

### 4. Feature extraction

In total, 65 features characterizing object geometry and texture were used. The feature list includes: (1) object area; (2) eccentricity; (3) perimeter; (4–10) the Hu set of invariant moments (seven moments) [42]; (11) standard deviation (std) of object grey levels; (12) entropy of object grey levels; (13 and 14) mean and std of the local entropy of object grey levels (the entropy value in the $9 \times 9$ neighbourhood of each pixel of the input grey scale image is computed); (15 and 16) mean and std of the local std (the $3 \times 3$ window is used to compute the local std); (17–30) Haralick’s coefficients [8] computed from the averaged (over four directions) co-occurrence matrices estimated using the distance parameter $d=5$; (31–34) four features representing the mean intensity and $35–38$ four features representing the std of intensity of the input image filtered using Gabor filters of four different scales: $2.5, 5, 10,$ and $20$ (six orientations $30^{\circ}, 60^{\circ}, 90^{\circ}, 120^{\circ}, 150^{\circ},$ and $180^{\circ}$ are used and the results are averaged); (39 and 40) mean and std of $M$ image pixels corresponding to the object; (41 and 42) mean and std of $m$ image pixels corresponding to the object; (43 and 44) mean and std of $C$ image pixels corresponding to the object, (45) mean of binary object image obtained from the FCM clustering; (46–55) 10 features computed from the local std image to characterize intensity variation when going from object’s exterior to the centre; (56–65) as 46–55, except that $M$ image is used instead of local std image.

### 5. Classification

A committee made of a Gaussian kernel SVM and a random forest (RF) classifier [11,12] is used to make a decision. Aggregation...
of SVM and RF is done by applying weighted averaging to the posteriori probabilities obtained from these two classifiers. The posteriori probability from a trained RF is estimated as

$$p((t_1, \ldots, t_L), r, q) = \frac{\sum_{j=1}^{L} f(t_j, r, q)}{L}$$

(16)

where $t_1, \ldots, t_L$ are the trees of the random forest, $r$ is the object being classified, $q$ is a class label, and $f(t_j, r, q)$ stands for the $q$th class frequency in the leaf node where $r$ falls in the $i$th tree $t_i$ of the forest

$$f(t_i, r, q) = \frac{n(t_i, r, q)}{\sum_j n(t_i, r, q_j)}$$

(17)

where $n(t_i, r, q)$ is the number of training data coming from class $q$ and falling into the same leaf node of $t_i$ as $r$.

6. Feature selection

It is well known that some of the features, that can be extracted in many pattern recognition applications, may be redundant or even irrelevant. In such cases, better performance can be achieved by discarding such features. Feature selection in general is a difficult problem, especially in the case of classification committees, since an accurate committee is obtained by combining members, which are not only accurate but also diverse. Committees targeted feature selection can help creating diverse members of a committee. In this work, however, we perform feature selection separately for SVM and RF.

Support vectors-based feature selection [42] is a good choice to select features for SVM. However, aiming to integrate feature selection with the process of determining the optimal values of SVM hyper-parameters – the width of the Gaussian kernel and the regularization constant – we use a genetic algorithm (GA) [44] to solve this task. The fitness function optimized by the genetic algorithm is given by the classification accuracy of the validation set data.

Feature selection for RF is based on feature importance evaluations available from the RF software and it’s relation to the classification accuracy is indirect. Each RF tree is grown on a bootstrap sample. For an RF tree grown on a bootstrap sample, the out-of-bag (OOB) data are used to estimate feature importance. The importance measure $\overline{D}_j$ for feature $x_j$ is given by [45,12]

$$\overline{D}_j = \frac{1}{B} \sum_{b=1}^{B} (K_{OB}^{b} - K_{OB}^{b})$$

(18)

where $B$ is the number of bootstrap samples (trees in the forest), $K_{OB}$ is the number of correct classifications of the OOB data by the tree $T_b$ and $K_{OB}$ is the number of correct classifications of the OOB data when the values of $x_j$ in the OOB set were randomly permuted. It is expected that for an important feature large decrease in the number of correct classifications will be observed when the feature values are randomly permuted. Sequential backward feature elimination based on feature importance evaluations given by Eq. (18) was applied to select features for RF.

7. Experimental investigations

7.1. Parameters of the algorithms

There are several parameters governing behaviour of the algorithms. The appropriate values of the parameters are chosen experimentally. The average radius of $P. minimum$ cells computed using cell area data was equal to $r = 45.8$ pixels and the standard deviation of the radius was $3.7$ pixels. Based on this information, we set $r = 45$ and this value allowed accurate enough estimation of centre positions of $P. minimum$ cells of various sizes encountered in the images. The $r$ value should be adjusted if the magnification factor of the microscope is changed. Three Gaussian filters are used by the algorithms. A $C$ image is filtered by a rotationally symmetric Gaussian low-pass filter with the standard deviation $\sigma$. The parameter $\sigma$ was set to $\sigma = 3$ and worked well in all the tests. An $M$ image in the polar coordinate system is filtered by a Gaussian filter with a much larger standard deviation in the angle direction than in the $r$ direction. The standard deviation equal to 8 and 1 in the angle and $r$ directions, respectively, was a good choice. The width of the Gaussian window $w'$ used by the stochastic contour detection algorithm is randomly selected from the interval $w_{\text{min}} = 20$ and $w_{\text{max}} = 100$. The number of successful iterations $n_t$ and the total number of iterations $n_\text{max}$ used by the algorithm were set to $n_t = 40$ and $n_\text{max} = 1000$. Objects larger than five times the average cell size are considered as being very large. The optimal number of randomly selected features used to split a tree node in RF was equal to 11 and the number trees in the random forest was set to 1000. It is worth mentioning that the algorithms are rather insensitive to the choice of the parameters. Parameters of the algorithms are listed in Table 2, for convenience.

We used Breiman’s implementation of random forests [45], Kovesi’s functions to compute images $PC$, $\Phi$, $M$ and $r$ related to phase congruency analysis [46], and the LIBSVM software [47]. All other algorithms used in this work were implemented by us. About 2 h of computation time was needed to select features and SVM hyper-parameters by GA when using 1.6 GHz Intel Q720 processor. Observe that to evaluate fitness of each chromosome during the genetic search process, the SVM was trained 20 times using different random splits of the data set into training and validation subsets and results were averaged. The computation time can be reduced by reducing the number of repetitions. Processing of one phytoplankton image takes about 30 s, where about 15 s go to computations related to phase congruency and image segmentation and 15 s to object and feature extraction.

7.2. Results

Two criteria are used to assess the performance of the algorithms: the percentage of detected $P. minimum$ cells (if compared to all $P. minimum$ cells present in the images) and the test data set classification accuracy when classifying detected objects into $P. minimum$ cells and other objects (since objects not belonging to the class of $P. minimum$ cells are also detected). The accuracy reported here is the average accuracy obtained from 20 trials using random split of the data set into learning and test subsets. In total, 114 images were processed and the detection results were verified by manual inspection. The manual inspection has shown that there are 2088 $P. minimum$ cells in these images in total. The algorithms found 2412 objects. Among these objects,
1947 were detected as *P. minimum* cells. Thus, 93.2% of *P. minimum* cells were detected.

Fig. 12 presents an example of object detection results for a full-size image. As can be seen from Fig. 12, majority of small “uninteresting” objects were eliminated and all *P. minimum* cells, except those with centre points too close to the image boundaries, were detected. All objects detected in phase II do not belong to the class of *P. minimum* cells. Figs. 13 and 14 present two more examples of object detection results. Objects labelled by a yellow cross were left aside from the analysis, as being too close to the image boundaries with respect to the chosen parameter $r$ value. The experiments have shown that most of *P. minimum* cells were detected in the 1st analysis phase. This demonstrates the importance of information on phase congruency.

The genetic algorithm, used to design the SVM classifier, enabled selection of both, a feature subset and hyper-parameters (width of the Gaussian kernel and the regularization constant). The optimal feature subset for the SVM classifier was found to consist of 25 features. In the case of RF, 31 features were selected. Fig. 15 presents feature importance values computed for RF designed using all 65 features. Observe that when feature elimination process starts and features are eliminated one by one from RF, importance of the remaining features change, comparing to the values shown in Fig. 15. This is why some features exhibiting relatively high importance values in Fig. 15 were eliminated, while some features with relatively low importance values were included into the final feature subset for the RF-based classification. Features included into the final feature subsets selected for the RF and SVM classifiers are labelled by a triangle and a circle, respectively, in Fig. 15. On the top of Fig. 15, types of features used are shown for the sake of convenience. As can be seen from Fig. 15, several features computed from $M$, $m$ and, $C$ images created by performing analysis on phase congruency were found to be salient for both SVM and RF classifiers.

The test data set classification accuracy obtained from the RF, SVM, and the committee was equal to 90.9%, 94.4%, and 94.9%, respectively. On average, SVM was more accurate than RF. Fig. 16 explores behaviour of the committee when changing the weight value, where “Oracle” stands for a classifier making a correct decision if at least one of the two members, RF or SVM, makes a correct decision. Since in many cases both SVM and RF erroneously classified the same samples, the committee was not so effective. However, as can be seen from Fig. 16, the committee outperformed the SVM for a broad range of weight $w$ values.

Figs. 17 and 18 present ROC curves computed for the SVM, RF and the committee. As can be seen from the ROC curves, even if SVM outperformed RF on average, RF is superior to SVM for small false positive rate values.

Deeper insights into similarity of objects coming from the two classes can be obtained by exploring the data proximity matrix available from RF. By applying some data mapping technique, the matrix can be mapped on the two-dimensional (2D) space. Fig. 19 presents such mapping obtained by using multidimensional scaling. *P. minimum* cells are denoted by rectangles and other objects by triangles in the figure. It is worth mentioning that, due to the simplicity of the imaging system used, there were many images containing rather unfocused regions. *P. minimum* cells extracted from such regions are labelled by filled-in markers in Fig. 19. There were 25% of such *P. minimum* cells in total, while the overall error rate is 5%. Thus, most of *P. minimum* cells extracted from unfocused images were recognized correctly. Moreover, Fig. 19 shows that *P. minimum* cells extracted from unfocused regions do not make a separate cluster and overlap, to
a great extent, with other *P. minimum* cells. Thus, the feature system used is rather robust to such type of distortions. Bearing in mind the simplicity of the imaging system used the results we obtained are rather encouraging.

![Figure 15](image-15.png)  
**Fig. 15.** Variable importance values for all 65 variables.

![Figure 16](image-16.png)  
**Fig. 16.** Committee performance.

![Figure 17](image-17.png)  
**Fig. 17.** ROC curves computed for SVM and RF classifiers.

![Figure 18](image-18.png)  
**Fig. 18.** ROC curves computed for SVM, RF, and the committee.

![Figure 19](image-19.png)  
**Fig. 19.** The data proximity matrix mapped onto the 2D space by multidimensional scaling.
## 8. Discussion and conclusions

A system for automated detection, recognition, and derivation of abundance estimates of different phytoplankton species using phytoplankton images generated by a simple imaging system, is a long term goal of this work. This article is mainly concerned with detection and recognition of objects representing one invasive species, *P. minimum*, which is known to cause harmful blooms in many estuarine and coastal environments. Special emphasis is made on the object detection and feature extraction problems in this work. Nonetheless the robust detection of objects is a prerequisite to create an accurate system for automated analysis of plankton images, this aspect of the problem is very seldom addressed in the literature.

A new technique, combining phase congruency-based detection of circular objects, stochastic optimization, and image segmentation was developed for solving the object detection task. A committee consisting of a support vector machine and a random forest is used to make a decision. A committee decision is obtained by averaging the a posteriori class probabilities obtained from these two classifiers. Experimental tests have shown that even in images, where image area occupied by objects was much larger than 3%. On average 93.25% of objects representing *P. minimum* cells were detected. On average, the SVM showed superior performance to that of the RF, although these two classifiers agreed on most of the erroneously classified data. Therefore, only moderate improvement in accuracy was obtained from the committee. The test data set classification accuracy of 94.9% was obtained from the committee. It is worth mentioning that this accuracy was reached using objects extracted from relatively unfocused images. After examining results obtained using blurred and out-of-focus images it is clear that the features selected for the classification task are robust enough to such type of distortions. Barring in mind the simplicity of the imaging system used and the high percentage of image area occupied by objects in phytoplankton images, the obtained results are rather encouraging and will be used to develop an automated system for obtaining abundance estimates of the species.

### Acknowledgement

This research was funded by a grant (No. LEK-05/2010) from the Research Council of Lithuania. The authors are grateful to Dr. Adrianna Ianora and Dr. Francesco Esposito (Stazione Zoologica, Naples, Italy) for providing the living culture of *P. minimum*, Naples, Italy) for providing the living culture of *M. brevipes*, Naples, Italy. A video showing the automatic detection of objects was present during the conference.

## References


[2] M.C. Benfield, P. Grojean, P.F. Culverhouse, X. Irgoien, M.E. Sieracki, A. Lopez-Urrutia, H.G. Dam, Q. Hu, C.S. Davis, A. Hansen, C.H. Pilskaln, E.M. Riseman, H. Schultz, H. Utgoff, G. Gorsky, RAPID: object detection is possible even in images, where image area occupied by objects was much larger than 3%. On average 93.25% of objects representing *P. minimum* cells were detected. On average, the SVM showed superior performance to that of the RF, although these two classifiers agreed on most of the erroneously classified data. Therefore, only moderate improvement in accuracy was obtained from the committee. The test data set classification accuracy of 94.9% was obtained from the committee. It is worth mentioning that this accuracy was reached using objects extracted from relatively unfocused images. After examining results obtained using blurred and out-of-focus images it is clear that the features selected for the classification task are robust enough to such type of distortions. Barring in mind the simplicity of the imaging system used and the high percentage of image area occupied by objects in phytoplankton images, the obtained results are rather encouraging and will be used to develop an automated system for obtaining abundance estimates of the species.

### Acknowledgement

This research was funded by a grant (No. LEK-05/2010) from the Research Council of Lithuania. The authors are grateful to Dr. Adrianna Ianora and Dr. Francesco Esposito (Stazione Zoologica, Naples, Italy) for providing the living culture of *P. minimum*. A video showing the automatic detection of objects was present during the conference.

## References


[2] M.C. Benfield, P. Grojean, P.F. Culverhouse, X. Irgoien, M.E. Sieracki, A. Lopez-Urrutia, H.G. Dam, Q. Hu, C.S. Davis, A. Hansen, C.H. Pilskaln, E.M. Riseman, H. Schultz, H. Utgoff, G. Gorsky, RAPID: object detection is possible even in images, where image area occupied by objects was much larger than 3%. On average 93.25% of objects representing *P. minimum* cells were detected. On average, the SVM showed superior performance to that of the RF, although these two classifiers agreed on most of the erroneously classified data. Therefore, only moderate improvement in accuracy was obtained from the committee. The test data set classification accuracy of 94.9% was obtained from the committee. It is worth mentioning that this accuracy was reached using objects extracted from relatively unfocused images. After examining results obtained using blurred and out-of-focus images it is clear that the features selected for the classification task are robust enough to such type of distortions. Barring in mind the simplicity of the imaging system used and the high percentage of image area occupied by objects in phytoplankton images, the obtained results are rather encouraging and will be used to develop an automated system for obtaining abundance estimates of the species.
Antanas Verikas is currently holding a professor position at both Halmstad University Sweden and Kaunas University of Technology, Lithuania. His research interests include image processing, pattern recognition, artificial neural networks, fuzzy logic, and visual media technology. He is a member of the International Pattern Recognition Society, European Neural Network Society, International Association of Science and Technology for Development, Swedish Society of Learning Systems, and a member of the IEEE.

Adas Gelzinis received the M.S. degree in electrical engineering from Kaunas University of Technology, Lithuania, in 1995. He received the Ph.D. degree in computer science from the same university, in 2000. He is a senior researcher in the Department of Electrical & Control Equipment at Kaunas University of Technology. His research interests include artificial neural networks, kernel methods, pattern recognition, and image processing, texture classification.

Marija Bacauskiene is a senior researcher in Department of Electrical & Control Equipment at Kaunas University of Technology, Lithuania. Her research interests include artificial neural networks, image processing, pattern recognition, and fuzzy logic. She participated in various research projects and published numerous papers in these areas.

Irina Olenina is working for several years as senior specialist at the Department of Marine Research of the Environmental Protection Agency in Klaipeda, Lithuania being responsible for analysis of freshwater and marine phytoplankton. Her main research interest includes phytoplankton biodiversity, long-term changes in phytoplankton community, ecology of invasive phytoplankton species. She is a member of the HELCOM Phytoplankton Experts Group, and a member of the Baltic Marine Biologists Working Group “Intercalibration and distribution of diatom species in the Baltic Sea”.

Sergej Olenin is a Professor in Hydrobiology at Klaipeda University, Lithuania. His research field is aquatic invasion ecology with particular reference to the ecosystem change analysis and vectors of introductions. He participated in several recent EU funded research projects aimed at biological invasions (DAISIE, ALARM, IMPASSE, MEECE) being responsible for analysis of marine bioinvasion impacts. He is a co-editor of the book “Invasive aquatic species of Europe—distribution, impact and management” (Kluwer, 2002), and a member of the Editorial Board of the journal Biological Invasions.

Evaldas Vaicinkynas received the M.S. degree in informatics engineering from Kaunas University of Technology, Lithuania, in 2006. He is currently a Ph.D. candidate in computer science and also a junior research assistant in the Department of Electrical & Control Equipment at Kaunas University of Technology. His research interests include computational intelligence, statistical learning, kernel methods, pattern recognition, and signal processing.