

OvMeter: an automated 3D-integrated opto-electronic system for *Ostreopsis cf. ovata* bloom monitoring

Francesca Sbrana¹ · Ettore Landini¹ · Nikolla Gjerci² · Federica Viti¹ · Ennio Ottaviani² · Massimo Vassalli¹

Received: 8 September 2016 / Revised and accepted: 18 January 2017
© Springer Science+Business Media Dordrecht 2017

Abstract Over the last decade, toxic events along the Mediterranean coast associated with exceptional harmful blooms of the dinoflagellate *Ostreopsis cf. ovata* have increased in frequency and distribution, causing not only the death of marine organisms and human health problems, but also economic loss on the tourism and aquaculture industries. In order to reduce the burden of routine algal counting, an innovative automated, low-cost, opto-electronic system called OvMeter was developed. It is able to speed up the monitoring process and therefore it enables early warning of incipient harmful algal blooms. An ad-hoc software tool provides automated cell recognition, counting and real-time calculation of the final algal concentration. The core of dinoflagellate recognition relies on a localization step which takes advantage of the synergistic exploitation of 2D bright-field and quantitative phase microscopy images, and a classification phase performed by a machine learning algorithm based on Boosted Trees approach. The architectural design of the OvMeter device is presented here, together with a performance evaluation on sea samples.

Keywords *Ostreopsis Cf. ovata* · Dinoflagellate, automated environmental monitoring · Image processing · Pattern recognition

Introduction

Over the last decade, episodes of exceptional harmful algal blooms (HABs) of dinoflagellate *Ostreopsis cf. ovata* have been reported repeatedly around the basin of the Mediterranean sea. *Ostreopsis ovata* is a potentially toxic benthic microalga that, when reaching high concentrations ($\geq 10^4$ cells L⁻¹, (Tichadou et al. 2010)) during its bloom, may cause dangerous side effects to human health, therefore being deleterious for commercial activities, tourism, fishing and aquaculture. *Ostreopsis ovata* grows in shallow and sheltered waters, characterized by low hydrodynamic movements and usually on rocky bottoms as benthic substrate, with the formation of filaments and mucilaginous substances (Totti et al. 2010; Honsell et al. 2011). In the presence of strong wave motion, the microalgae can be resuspended in the water column, giving rise to flocculates and sometimes to surface foam, which can be transported in marine aerosols. Traces of palytoxin, among the most powerful non-peptidic marine toxins known to date (Hilgemann 2003; Ciminiello et al. 2006), and ovatoxins are responsible for the toxicity of *O. ovata*, causing mass mortality of benthic organisms (sea urchins, limpets and barnacles) (Faimali et al. 2012; Giussani et al. 2015). Human intoxication is characterized by irritation of the upper airways, coughing, rhinorrhea, shortness of breath, muscle and joint pain, fever and irritation of the mucous membranes of the eye and skin (Ciminiello et al. 2008, 2012). To date, national and regional monitoring programs of *O. ovata* have being carried out, implementing a counting procedure based on a microscopic inspection of water samples and the manual recognition by taxonomy expert operators. The complexity of this estimation procedure causes long delays between data collection, analysis and information reporting, thus limiting the effectiveness of the monitoring programs.

In the last few years, many automated optical systems have been developed to speed up the recognition task of marine

✉ Francesca Sbrana
francesca.sbrana@ge.ibf.cnr.it

¹ Institute of Biophysics, National Research Council, Via De Marini 6, 16149 Genoa, Italy

² OnAir srl, Via Carlo Barabino 26/4B -, 16129 Genoa, Italy

organisms, such as plankton and microalgae. Verikas et al. (2012), for example, worked on *Prorocentrum minimum* to discriminate organisms in overcrowded images, which represent one of the worst scenarios, they combined phase congruency detection of circular objects, stochastic optimization and image segmentation, achieving 93.25% of sensitivity. The most commonly exploited approaches for rapid identification and counting of microorganisms are imaging systems based on machine learning (Rodenacker et al. 2006; Tao et al. 2010), which are often able to achieve accuracy similar to manual identification performed by trained personnel. In particular, Support Vector Machines (SVMs) (Sosik and Olson 2007) and Artificial Neural Networks (Culverhouse et al. 2003; Embleton et al. 2003; Mosleh et al. 2012; Schulze et al. 2013; Coltelli et al. 2014) are among the most popular classification algorithms. Interestingly, most of the cited methods rely on fluorescent and/or pictorial features. Rodenacker et al. (2006) proposed an automated image acquisition system which, based on a hierarchical tree classification and taking advantage of the chlorophyll auto-fluorescence properties of marine algae, was able to precisely discriminate between phytoplankton cells and non-algal structures. Schulze et al. (2013) proposed an automated analysis system for the identification of phytoplankton, exploiting bright field and fluorescence images. The system relied on segmentation of the organisms out of the background, calculation of a large range of features and classification of imaged organisms into different groups of plankton taxa through a neural network. Sosik and Olson (2007) developed a high-resolution imaging system based on flow cytometry that provided taxonomically resolved estimates of phytoplankton abundance. Exploiting chlorophyll fluorescence and running a SVM on a large set of examples, the system rapidly outputs quantitative measurements of individual organisms. An automatic, real-time, multi-algal recognition tool mounted on a motorized image analysis system was presented by Coltelli et al. (2014), based on pigment signatures of microalgae. The system was designed to recognize algal feature vectors, grouping them into classes by means of an unsupervised neural network and was able to screen, identify and enumerate 23 different microalgal species, representing the major algal phyla, with the finest taxonomic resolution, attaining 98.6% accuracy from a set of 53,869 images. In recent years, with the advances in digital image sensors and the increase of computing power, innovative approaches based on quantitative phase microscopy (QPM) raised a growing attention. This class of instruments provides additional quantitative information by recording the phase shift produced by a specimen at each point within an optical image, which can be employed to identify and classify particles and microorganisms (Mir et al. 2012). In particular, Javidi et al. (2005) demonstrated the potential of QPM-based quasi-3D imaging for marine sample analysis. The authors implemented a complex workflow on QPM images by

chaining object segmentation, shape features extraction (based on Gabor wavelets) and localization through rigid graph matching that was used to recognize two filamentous algae, *Sphaelaria* and *Tribonema aequale*. Moreover El Mallahi et al. (2013) applied a specialized QPM configuration to automatically detect and classify living organisms a fluidic set-up without any staining. Morphological and textural features were extracted from the phase map. A robust SVM classifier was built to unambiguously detect and identify parasitic protozoan *Giardia lamblia* cysts and algal cells such as *Chlorella autotrophica* and *Scenedesmus dimorphus* which are often present in drinking water resources. An enhancement of taxonomic resolution in automated phytoplankton classification provided by additional phase information was also presented by Zetsche et al. (2014) who worked on nano-planktonic organisms. In particular, three single-celled nanoplankton species similar in shape, (*Cyanothece* sp., *Staniera* sp. and *Chlorella autotrophica*) were imaged through QPM and injected in a SVM classification algorithm reaching a classification score of 92.4%. Textural features based on phase information resulted efficient in separating the three tested phytoplankton species.

Unfortunately, the previously mentioned methods are not suitable for a quick quantification of *O. ovata* cell abundance. In fact, *O. ovata* cells have a drop-like shape (very similar to few other microalgae populating the same environment) which helps in distinguishing *O. ovata* from most other marine microorganisms, but which is rarely compatible with the already developed approaches, based on common geometrical shapes. Moreover, existing methods cannot be adapted to the monitoring guidelines, described by the Italian National Institute for Environmental Protection and Research (ISPRA) (ISPRA-ARPA 2012), which represents the reference for environmental monitoring in Italy, and described in the European operational protocols (<http://m3-habs.net/guidelines/>), because Lugol's acidic solution (suggested in the guidelines as a preservation method for maintaining water samples) drastically reduces auto-fluorescence signal, disabling the exploitation of all fluorescence-based classification methods. Therefore, in order to reduce the burden of routine identification and quantification of *O. ovata* HABs, an innovative, automated, low-cost, opto-electronic system exploiting QPM was developed. As a result of the tight integration of a motorized image acquisition system with on-line cell recognition and counting software, based on a supervised machine learning algorithm (a Boosted Tree, BT), the system is able to quickly determine *O. ovata* concentration, ensuring a fast activation of precautionary measures for preserving human health. The tool takes advantage of a synergistic combination of 2D bright field and QPM images, particularly effective for the analysis of critically high-density images, where cells tend to overlap, or dirty samples (with massive presence of debris and fragments) and inverting the approach. The

architectural design of the OvMeter device is presented here, and its performance are evaluated on field samples.

Materials and methods

TIE-quantitative phase microscopy

TIE-QPM is a non-interferometric phase imaging technique, used to retrieve information about the phase shift of the light radiation passing through a semi-transparent medium computed from a Z-stack acquisition of out-focus bright-field images by solving the Transport of Intensity Equation (TIE). Compared to other widely used phase retrieval techniques, like phase-contrast microscopy (Barty et al. 1998) and interferometer techniques, such as digital holographic microscopy (DH), which entangle the phase and intensity images, TIE-QPM can quantitatively map both. Moreover, since TIE-QPM can be performed with any standard optical microscopy setup, as long as it can move either the sample or the objective along the Z-axis, it drastically reduces the cost of the optical setup.

The relation between phase and intensity images can be modelled through the TIE (Teague 1983):

$$\nabla \cdot [I(\vec{r}, z) \nabla \psi(\vec{r}, z)] = -k \partial_z I(\vec{r}, z) \tag{1}$$

where

- r is the position vector (x, y) along the plane perpendicular to the propagation axis (imaging plane),
- z is the position along the propagation axis,
- $I(r, z)$ is the intensity of the light passing through the specimen, in a given point in space,
- ∇ is the 2D gradient operator acting in the imaging plane,
- $\nabla \cdot$ is the divergence operator acting on the imaging plane,
- $\phi(r, z)$ is the incident light phase at a given point in space,
- k is the wave number ($k = 2\pi/\lambda$, where λ is the light wavelength).

Assuming that a solution exists, this equation shows a relation between the derivative of the intensity calculated along the propagation axis and the phase value. Moreover, assuming an absence of irradiance zeros $I(\vec{r}, z) > 0$, a uniform irradiance and no phase singularities in the optical field, it can be solved for the phase (Nugent et al. 2000):

$$\phi(\vec{r}, z) = -F^{-1} \|\vec{q}\|^{-2} F \left[\nabla \cdot [I(\vec{r}, z)^{-1} \nabla \psi(\vec{r}, z)] \right] \tag{2}$$

where

- $\psi(\vec{r}, z) = F^{-1} \|\vec{q}\|^{-2} F [k \partial_z I(\vec{r}, z)]$,
- $\vec{q} = (q_x, q_y)$ and $\|\vec{q}\|$ is the Euclidean norm of \vec{q} ,

- q_x and q_y are the variable that conjugate to x and y in the frequency domain,
- F and F^{-1} are the Fourier transform and its inverse acting on the imaging plane.

This solution allows calculation of the incident light phase values across the plane of focus by knowing just $\partial_z I(r, z)$ and the intensity of the light. The phase image provides information also about the thickness of the sample, since the phase shift of a light beam passing through a semi-transparent object is proportional to the thickness of the sample, assuming that the object is optically uniform. Consequently, information can be inferred about the 3D structure.

Unfortunately, *O. ovata* is not optically uniform; therefore, it is not possible to retrieve the exact morphology of the sample. On the other hand, since the phase shift in the light passing through an object depends on its refractive index, which is tied to the internal structure of the sample, it may be possible to use this information to get a better localization and segmentation of different species thanks to their optical features.

To obtain an efficient numerical implementation of (2), it is possible (Gureyev and Nugent 1997) to write (2) as follows:

$$\phi(\vec{r}, z) = \phi^{(X)}(\vec{r}, z) + \phi^{(Y)}(\vec{r}, z) \tag{3}$$

$$\phi^{(X)}(\vec{r}, z) = -F^{-1} q_x \|\vec{q}\|^{-2} F \left[I(\vec{r}, z)^{-1} F^{-1} q_x \|\vec{q}\|^{-2} F [k \partial_z I(\vec{r}, z)] \right] \tag{4}$$

$$\phi^{(Y)}(\vec{r}, z) = -F^{-1} q_y \|\vec{q}\|^{-2} F \left[I(\vec{r}, z)^{-1} F^{-1} q_y \|\vec{q}\|^{-2} F [k \partial_z I(\vec{r}, z)] \right] \tag{5}$$

Then, since dividing by the spatial frequency results in a numerical instability for q_x and q_y (that are calculated using the image resolution and the conversion factor micrometer per pixel, provided by the user) approaching zero, two regularization filters were applied (Engl et al. 1996):

$$\Omega_x = \frac{\|\vec{q}\|^2 q_x}{\|\vec{q}\|^4 + \alpha q_x^2}, \quad \Omega_y = \frac{\|\vec{q}\|^2 q_y}{\|\vec{q}\|^4 + \alpha q_y^2} \tag{6}$$

where

$$\alpha = \alpha' \max \left(\|\vec{q}\|^2 \right)$$

α' is a regularization parameter.

By substituting Ω_x and Ω_y in Eqs. (4) and (5), a regularized solution can be obtained:

$$\phi^{(X)}(\vec{r}, z) = -F^{-1} \Omega_x F \left[I(\vec{r}, z)^{-1} F^{-1} \Omega_x F [k \partial_z I(\vec{r}, z)] \right] \tag{7}$$

$$\phi^{(Y)}(\vec{r}, z) = -F^{-1} \Omega_y F \left[I(\vec{r}, z)^{-1} F^{-1} \Omega_y F [k \partial_z I(\vec{r}, z)] \right] \tag{8}$$

The intensity derivative $\partial_z I(r, z)$ is calculated by fitting the values of every pixel with respect to the position along the Z-axis with a polynomial expression and then deriving it analytically.

Supervised machine learning

Supervised machine learning (Bishop 2007) is a branch of machine learning that deals with inferring a function (referred to as the learned function or trained model) from a labelled training set. The training set consists of examples that are made up of pairs of inputs and their expected outputs. A supervised learning algorithm analyses training data and generates the learned function: once the inferring process has ended, it is possible to use this function to make predictions on novel data sets. It is of critical importance to choose a training set that is representative of the general usage of the learned function, so that the predicted classification is generalized correctly. The training set is usually converted into a feature vector that encodes the intrinsic information that describes each object to analyse.

The learning algorithm can be chosen from a wide range of available types such as SVMs (Burgess 1998; Nguyen and De la Torre 2010; Orabona et al. 2010), neural networks (Egmont-Petersen et al. 2002; Bishop 2007), and decision trees (Breiman 2001; Sok et al. 2016). No single algorithm is able to solve every possible machine learning problem (Wolpert and Macready 1997), each one has strengths and weaknesses that should be carefully taken into consideration every time a new learning problem arises. For the development of OvMeter, the classification task was implemented through the machine learning algorithm known as BT (Friedman 2001). This approach is similar to random forests (Breiman 2001; Sok et al. 2016); the difference lies in the fact that in a random forest, every tree (a small classifier) works in parallel with the others, whereas in BT, they act in series, each one trying to correct the classification errors of the previous one.

The BT algorithm combines the performance of many *weak* classifiers to produce a powerful ensemble. A *weak* classifier is only required to be better than chance and thus can be very simple and computationally inexpensive. However, many of them smartly combine results to a *strong* classifier that often outperforms most popular classifiers such as SVM and Neural Networks.

Different variants of boosting are available: Discrete AdaBoost, Real AdaBoost, LogitBoost, and Gentle AdaBoost. All of them are very similar in their overall structure. Therefore, in OvMeter, the attention was focused only on the standard two-class Discrete AdaBoost algorithm, briefly outlined below. Initially, the same weight was assigned to each training sample. Then, a weak classifier (a single tree) was

trained on the weighted training data and its weighted training error was computed. Based on the total error, the classifier got a score defining that its weight was the final ensemble. Then the sample weights were increased for training samples that have been misclassified and all weights were normalized. The process of finding the next weak classifier continued for a fixed number of times. The final classifier was the sign of the weighted sum over the individual weak classifiers.

There are two main strengths with exploiting boosted tree technique: (1) it performs well on heterogeneous feature vectors with no a priori normalization information, (2) it very seldom suffers from overfitting.

Algal collection and specimen preparation

Marine water samples containing *O. ovata* were collected during the Summer 2015 following *O. ovata* bloom events. Ten fixed stations in protected harbours or bays were chosen along coasts of Spain, Italy and France: Pins Mar (41° 32' 60" N, 2° 29' 24" E) and Blau Mar—Barcelona (41° 32' 20" N, 2° 26' 58" E) (Spain); Chiavari (44° 19' 10" N, 9° 19' 2" E), Fiascherino, (44° 3' 53" N, 9° 55' 16" E) and Genova Quarto (44° 23' 22" N, 8° 59' 30" E) (Italy); Rochambeau Villefrance (43° 41' 40" N, 7° 18' 17" E) and Haliotis—Nice (43° 41' 21" N, 7° 14' 27" E) (France). Samples were collected and treated following the guidelines suggested by ISPRA (ISPRA-ARPA 2012). The macro-alga *Halopteris scoparia* was collected at a depth of 30 cm above pebbles closer to the coast by diving, carefully cutting and placing it in plastic bottles under water. After shaking the macro-algae in the storage bottle vigorously to dislodge epiphytic dinoflagellates, the thalli was also rinsed in filtered sea water. The recovered seawater was filtered again through 200 mm mesh sieves to remove large particles. An aliquot of 1 mL was injected in a Sedgewick-Rafter counting chamber, covered by a coverglass and immediately observed. Column seawater samples were also taken and sampled in shallow waters at 30 cm depth by using plastic jars (250 mL). Column water samples were poured into sedimentation chambers according to the Utermöhl method (Utermöhl 1958) and left to settle for 24 h before observation. A sedimentation column volume of 50 mL was used. All samples, both from macro-algae and column water, were immediately preserved by using a 1% (v/v) Lugol's acidic solution. The exploitation of two different counting chambers (Utermöhl and Sedgewick) to observe cells of *O. ovata* does not influence the analysis methodology proposed in this work.

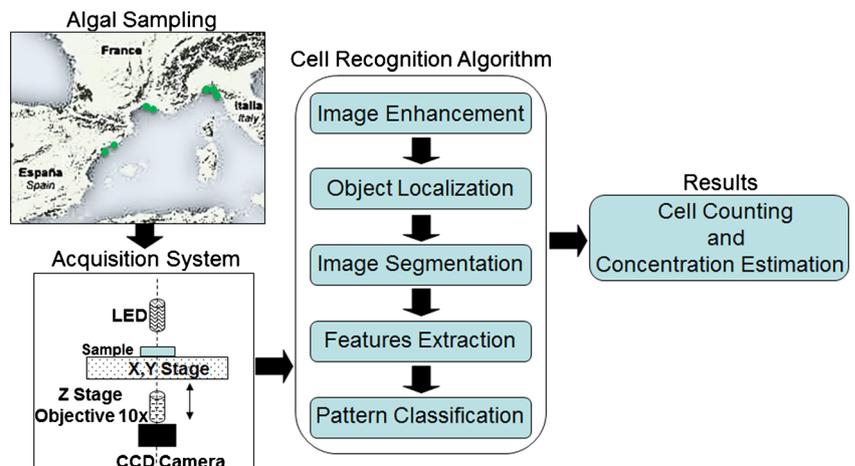
OvMeter acquisition system

OvMeter relies on an opto-electronic platform that integrates standard modular components (Optem FUSION, Qioptiq Photonics GmbH & Co KG) and is equipped with

an Optem 10× high-resolution, long working distance (NA 0.30; WD 34 mm), infinity-corrected objective, mounted on a motorized Z-axis with a 0.01 μm resolution step. The sample is scanned through a motorized X-Y stage with a 0.5 μm resolution step. A modular, compact and freely programmable stepper motor controller (phyMOTION™, Phytron, Germany) drives three motorized axes. A LED lamp is integrated for illumination and a Gig-E DMK 23G274 camera (The Imaging Source, Bremen, Germany) equipped with a CCD (b/w, 1600 × 1200 pixel) is used. In this configuration, the field of view corresponds to 0.5286 × 0.7048 mm², with a spatial resolution of 1 μm. A sample holder was specifically designed to house both the Utermöhl and Sedgewick-Rafter counting chambers with a standard K-frame.

The OvMeter control software was developed in a LabVIEW (National Instruments, USA) environment and automates all microscope functions such as movement of the three axes, acquisition of the 2D bright-field image and acquisition of a Z-stack of images. A scanning procedure is implemented in order to acquire images following a defined pathway chosen by the user. A raster movement is performed with a tuneable step in the X-Y directions. The number of images acquired in the Z-stack and the Z-step size can be both defined by the user. A LabVIEW module was developed for executing the TIE-QPM algorithm, which produces quantitative phase (QP) images. An automated machine learning approach for classification and identification of *O. ovata* cells was implemented in C++ language, operating in Windows environment and wrapped in a dynamic-link library (DLL), housed in the management software. All metadata, including temperature, salinity and pH of sea water, together with sampling date, cell extraction method and total cell counts are summarized and saved in a text file. A simplified scheme of the OvMeter tool is shown in Fig. 1.

Fig. 1 Overview of the processing steps of the OvMeter tool. After the collection of the algae and sample preparation, image acquisition is performed, followed by the presented pipeline of image processing, which culminates in an estimate of the algal concentration



Performance measurements

In order to evaluate the trained model's ability to generalize, an independent test dataset was used. The output of the learned function on this novel set is arranged in a contingency table, where each column represents instances of the predicted class and each row instances of the actual class. In contingency table, if predicted and actual values are both true, true positive results (TP) are noted, while if both of them are false, true negative results (TN) are noted. Vice versa, false positive results (FP) indicate that a given condition has been fulfilled, when it actually has not been fulfilled, and false negative results (FN) indicate that a condition failed, while it actually was successful.

Based on the contingency table entries, different performance indexes can be derived, each of them considering specific aspects of the recognition process (Table 1).

Cell recognition algorithm

The cell recognition algorithm consists of five main processing steps, listed below, operating in cascade on an image in order to perform the *O. ovata* counting task.

(1) A pre-processing module based on the functions of the library OpenCV (<http://www.opencv.org>) implemented an edge detection scheme. Some of the operations that take place in this module are adjustment of the local contrast of the image through the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm (Zuiderveld 1994), image resizing to a standard size suitable for subsequent elaboration and global histogram normalization to enhance edge structures. This step was performed on both bright-field and quantitative phase images.

(2) An automated algorithm (Bernal et al. 2008) for the localization of interesting regions in the images was implemented (i.e. localization of possible *O. ovata*). Behind the localization scheme, the main peculiarity is that *O. ovata* has

Table 1 Description of the main performance measurements

Name	Formula	Note
Sensitivity	$TP/(TP + FN)$	Measures the proportion of positives that are correctly identified.
Specificity	$TN/(TN + FP)$	Measures the proportion of negatives that are correctly identified.
Precision	$TP/(TP + FP)$	Relates to reproducibility and repeatability and refers to the level of measurement that yields consistent results when repeated.
Accuracy	$TP + TN/(TP + TN + FP + FN)$	Level of measurement that yields true (no systematic errors) and consistent (no random errors) results, and represents the degree of closeness of measurement to the true value. An accuracy of 100% means that the measured values are exactly the same as the given values.
Fall-out	$FP/(FP + TN)$	Closely related to specificity and is equal to 1 specificity. It can be looked at as the probability that a non-relevant data is retrieved by the query.
F1-score	$2TP/(2TP + FP + FN)$	Considers both the precision and sensitivity of the test.

TP true positive results, *TN* true negative results, *FP* false positive results, *FN* false negative results

a distinctive drop-like shape. A representation of a typical *O. ovata* cell was elaborated by calculating the average shape among a large set of individual cells (Fig. 2a). This average specimen was used as a template for generating a set of representative cells of various rotations, eccentricities and scales. The number of different representative cells has been set according to a trade-off between computational complexity and detection accuracy of the localization algorithm. After having tested several combinations, it was found that a total of 64 cells (8 rotations, 2 eccentricities and 4 scales) was the best choice for the OvMeter system. This set is called template set (Fig. 2b). Every cell in the template set was pre-processed by means of convolution with the Laplacian of Gaussian wavelet. Ultimately, every template in the template set was convoluted with the preprocessed image of the first step of the algorithm

generating one convolution map for every template. Imposing a threshold across all convolution maps and selecting the extreme points suggests that where possible, *O. ovata* cells could lie on the image (Fig. 2c). This step was first performed on bright-field images, then on the QP images, which enhances the localization scheme performance on samples that contain large amounts of debris or particularly difficult overlaps.

(3) The segmentation (i.e. partitioning of a digital image into multiple sets) of the regions of interest selected in the previous step was performed: this was accomplished by selecting a sufficiently large area around each center and computing the binary map of this local area allowing the best separation of the cell from the external background. Given that the center point is inside the cell, the size of the local area should be large enough to include the whole contour of the

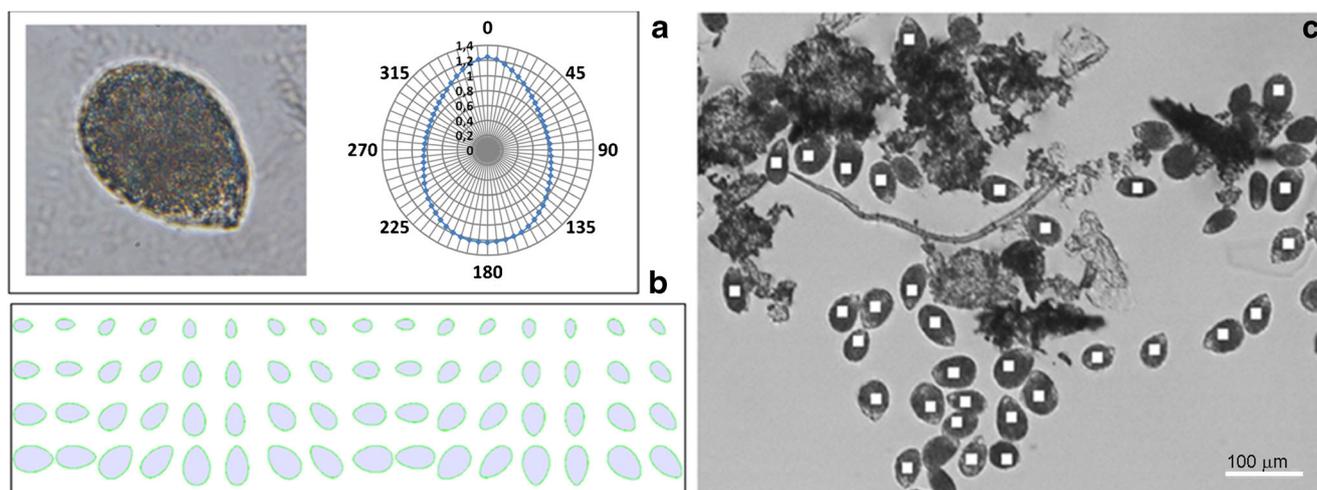


Fig. 2 Basis for localization scheme. **a** Optical image of an *O. ovata* cell (left) and a pictorial representation of the averaged pattern shape of a *O. ovata* cell (right). **b** Example of different values of scale, rotation, and eccentricity chosen as template to localize as many *O. ovata* individuals as possible (considering the inevitable presence of a

percentage of false positive results that will be removed during the classification step). **c** Example of an optical image output from the localization step: the method is able to find almost all *O. ovata* individuals within the image

located cell. The representation of the local image area as a graph allows the interpretation of the image segmentation task as a version of the well-known maximum flow problem on graphs. To deal with it, the Boykov-Kolmogorov algorithm (Boykov and Kolmogorov 2004) for solving mincut/max flow problems on graphs was used, as implemented by Computer Vision Research Group (<http://vision.csd.uwo.ca/code/>). This step was performed on bright-field images if the center of segmentation was found in the bright-field image; otherwise, this step was performed on the QP images. All of the above steps are thoroughly discussed by Gjerci (2015).

(4) A feature set was identified: it is a mix of shape and texture descriptors based on previous scientific works on *O. ovata* (Besada et al. 1982; Accoroni et al. 2012; Honsell et al. 2013; Escalera et al. 2014; Kreshchenovskaya and Orlova 2014). A set of 18 features of which 11 features were extracted from the shape of the region and 7 features were extracted from the texture are reported in Table 2.

In both types of images (2D images and QP images), feature extraction was performed on regions of interest, identified during segmentation step. It is worth to note that only bright-field images carry texture information, useful to discriminate between objects of similar shape but different internal content. Texture of QP images is not meaningful from the point of view of classification, due to the highly variable transparency properties of *O. ovata* cells and the unavoidable presence of artefacts. Nevertheless, QP images are used as a complementary way to locate and segment cells (in particular in cluttered samples) and to compute shape descriptors.

(5) Lastly, the classification was implemented by using the BT learning algorithm.

Results

OvMeter performance

Performance evaluation relied on the comparison of the automated counter with the human inspection. To quantify the performance of the system, an expert user was asked to manually evaluate the same images to which OvMeter applied the classification procedure. The exploited test dataset, which was completely independent from the training dataset, consisted of around 23,000 objects (including individuals and clutter noise), where 5000 items were *O. ovata*, thus reproducing the real proportion of true and false objects that can be commonly found in marine field samples ($10^5 \div 10^4$ cells L^{-1}).

The total number of *O. ovata* cells counted by the system, operating in 2D mode on seven different experiments (all samples collected in Genova Quarto, site), was 4642 (Table 3), compared to 4742 counted by the operator, thus providing a high degree of correlation between the two techniques (Fig. 3), a rather small dispersion (R^2 coefficient > 0.9) and an error of the system of 2.1%. The choice of a test set completely independent from the training set excluded the risk of overfitting, providing statistically reliable performance.

Object recognition performance was evaluated by analysing the contingency table (Table 3), which compares the actual values with those estimated, providing the degree of reliability of the classification based on different performance measurements. Considered indexes are reported in Table 4, ranged between 0 and 1. Sensitivity, specificity, precision, accuracy and F1 score showed values > 0.9 , thus highlighting an excellent reliability of the system in counting cells, compared to manual results. In particular, the obtained

Table 2 Description of shape and texture features

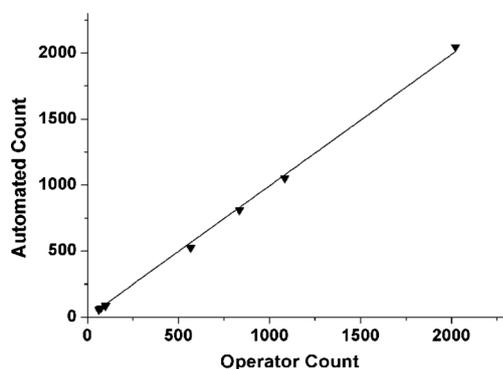
	Shape features	Description
1–7	Invariant Hu moments	A set of shape descriptors robust to scale and rotations (Hu 1962), 7 values up to the third order
8	Axis ratio	Minor axis/major axis (1 for almost circular objects and < 1 for elongated ones)
9	Convexity	The ratio between the area of the object and its convex hull (1 for convex objects)
10	Symmetry	The area of intersection between the object and its version flipped with respect to the major symmetry axis, normalized by the object area (1 for perfect symmetry)
11	Shape entropy	Defined as the entropy of the radius histogram (low values for nearly circular objects)
	Texture features	Description
1	Average grey level	Mean value of the grey level scale (low values for darker objects)
2	Global contrast	Difference of average grey value inside and outside the object, normalized to the sum of the same values
3	Entropy	First order statistics of grey values inside the object
4	Angular second moment	A measure of smoothness of the grey level spatial distribution, computed from the grey level co-occurrence matrix
5	Local contrast	A measure of local grey-level variations, based on the co-occurrence matrix
6	Inverse difference moment	Further second order statistics of grey level local variations
7	Second order entropy	Defined over the grey level co-occurrence matrix instead of the grey level histogram

Table 3 Contingency table of automated identification of *Ostreopsis ovata* respect to counts obtained from a visual inspection. *Ostreopsis ovata* cells counted with the OvMeter numbered 4642 (TP + FP) compared to 4742 (TP + FN) counted by the operator

	Actual value: true	Actual value: false
Predicted value: true	4335 (TP)	307 (FP)
Predicted value: false	407 (FN)	17,992 (TN)

values indicated the ability of the OvMeter tool to correctly identify the target cells and discriminated them from similar objects (sensitivity and specificity), to provide consistent results when the experiment is repeated (precision, F1 score) as well as to reach a good degree of closeness to the true value (accuracy). In accordance with previous values, the fall-out index was near 0, since it represents the probability that a non-relevant data is retrieved as positive during classification.

Performance analysis can be enriched by evaluating the distribution of real and predicted counts for each sample. This allows estimation of the degree of correlation among single counts, therefore inferring if accuracy of the prediction can be related to the sample's content and verifying if error percentage is correlated to the cell density. This study was performed on seven samples listed in Table 5 showing various cell densities, thus providing diverse levels of *O. ovata* abundance. Numerical values obtained for each sample are reported in Table 5. Samples containing few *O. ovata* cells are prone to false negative results, when automatically counted (compared to the manual count), thus increasing the relative error rate. As sample cell density increases, the number of false positive results decrease leading, first, to a reduced final error rate and, then, to invert error trend. In general, the absolute error remains almost constant, being dependent mainly on the number of processed images and not on cell density.

**Fig. 3** Comparison between automated and manual counts for seven independent and differently dense samples collected in Genova Quarto site, highlighting linear trend among automated and manual results (eq. $y = 0.9945x$, $R^2 = 0.9987$)**Table 4** Performance measurements obtained for samples collected in Genova Quarto site

Index	Value
Sensitivity	0.914
Specificity	0.983
Precision	0.934
Fallout	0.017
Accuracy	0.969
F1 score	0.924

Subsampling strategy

Evaluations were carried out in order to assess the minimum number of images to be captured to provide a reliable estimate of the number of cells contained in the counting chambers. In fact, the complete scanning of a biological sample requires the acquisition and processing of more than 1000 images. Additionally, the use of phase reconstructed images required the acquisition of more than four images in each position along the Z-axis, thus leading to a total scanning time for a whole sample of >2 h. Therefore, in order to save processing time, it becomes crucial to assess the minimum number of images necessary for obtaining a reliable statistical counting of cells. Obviously, this approximation generates an additional error, due to inferring data from areas which are not directly inspected. The produced error can be evaluated by assuming a uniform distribution of the cells in the sample. To pursue this aim, statistical simulations of under-sampling of a Sedgewick-Rafter counting chamber have been performed. Figure 4 shows a 2D representation of the distribution of an automated count performed on a selected region of a sedimentation chamber, corresponding to 882 fields of view of the OvMeter optical system. Colours (from blue to red) provide information about the number of *O. ovata* cells identified in each field of view. Considering the standard deviation value as the under-sampling error and assuming uniform cell distribution, the relations to be verified in order to estimate the

Table 5 Numerical values obtained by counting *Ostreopsis ovata* in samples of different cell densities. The samples were collected in Genova Quarto site

Sample	Real count	Automated count	Relative error (%)	Absolute error
1	62	57	-8.0	-5
2	74	69	-6.7	-5
3	99	86	-13.1	-13
4	567	523	-7.7	-44
5	835	810	-3.0	-25
6	1083	1052	-2.8	-31
7	2022	2045	+1.1	+23

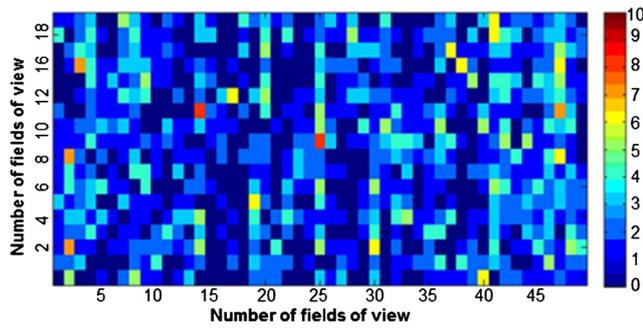


Fig. 4 Two-dimensional representation of the distribution of an automated count on a selected region of the Sedgewick-Rafter counting chamber. The region was 12.7 mm × 26.4 mm sized. Each pixel represents a field of view of the optical system. The count of *O. ovata* cells identified in each field of view is represented with a chromatic scale corresponding to a count range between 0 (blue—lowest) and 10 (red—highest) of *O. ovata* cells identified there

minimum number of images needed to ensure a reliable estimation are

$$\sigma_K < E_A \Rightarrow K > \frac{N \cdot C_0}{E_A^2} \quad \varepsilon_K < E_R \Rightarrow K > \frac{N}{E_R^2 \cdot C_0} \quad (9)$$

where σ_K and ε_K are, respectively, the absolute and the relative errors, E_A and E_R indicated the expected absolute and relative errors, N is the total number of images, K is the number of sampled images, and C_0 is the expected concentration of cells (obtained by screening K_{min} images). The relations reported in (9) determine K by imposing specific absolute and/or relative errors. Table 6 reports hypothetical values of K obtained by simulating sedimentation in a Sedgewick-Rafter counting chamber of samples presenting diverse and typical *O. ovata* concentrations, assuming $E_A = 80$ and $E_R = 5\%$. Graph in Fig. 5 shows the error variation caused by image under-sampling correlated with the average number of cells present in each image, considering sets containing 100, 200, 300 and 400 images.

As reported, the statistical uncertainty introduced by the under-sampling is high when samples are modestly populated

Table 6 Hypothetical values of K obtained by simulating sample sedimentation in a Sedgewick-Rafter counting chamber. E_A indicates the number of images obtained by assuming a threshold on the absolute error value. E_R indicates the number of images obtained by assuming a threshold on the relative error value

Total count	Counts per image	Total no. of images (E_A)	Total images no. (E_R)	Optimal images no. (K)
1328	1.5	183	266	183
2656	3.0	234	133	133
13,280	15.1	1171	27	27

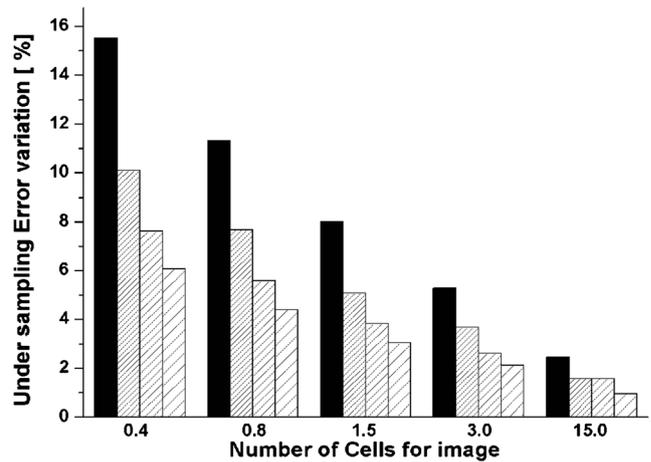


Fig. 5 Plot of under sampling error variation in percentage respect to number of cell per image. Each bar represents a set of 100 (full black), 200 (oblique dense stripes), 300 (oblique medium stripes) and 400 (oblique sparse stripes) images

(first stages of a HAB), while it reduces considerably for high density samples (corresponding to the actual peak of HAB), which are in fact the cases of interest for risk assessment.

Segmentation comparison

QP reconstructed images affect the pipeline in multiple steps, supporting object recognition and performing a pre-screening of items used to feed the classifier. The use of TIE-QPM proved to be particularly advantageous for complex images, such as those with a high overlap of microalgae and debris, the microalgae aggregation or the presence in the sample of microalgae with a shape very similar to *O. ovata*, such as the dinoflagellates *Prorocentrum lima* and *Prorocentrum micans*. In order to show the complementarity of information coming from the two considered types of images (2D and QP images), a test was carried out on a sample and results of the segmentation step (Fig. 6) are reported by mapping them on both images separately. The algorithm working on 2D images (Fig. 6a) selects *P. lima* (Fig. 6a, item 1) confusing it with *O. ovata*, thus resulting in over-segmentation (although this result does not affect system output since discrimination is performed by BT in the classification step, which follows the ‘segmentation’ phase). Similarly, it happens for the microalga *Coolia monotis* (Fig. 6a, item 2). The exploitation of QP reconstructed images (Fig. 6b) emphasizes *O. ovata* features, avoiding segmenting objects different from the target (Fig. 6b, items 1 and 2). Boxes highlighted in Fig. 6a, b show another typical case where integration of 2D and QP images can retrieve more complete information: *O. ovata* cells hidden by debris. 2D image segmentation correctly identifies *O. ovata* cell in Fig. 6a—item 3, but omits the identification of *O. ovata* cell in Fig. 6a—item 4. Segmentation of QP image was able to highlight both of them (Fig. 6b—items 3 and 4),

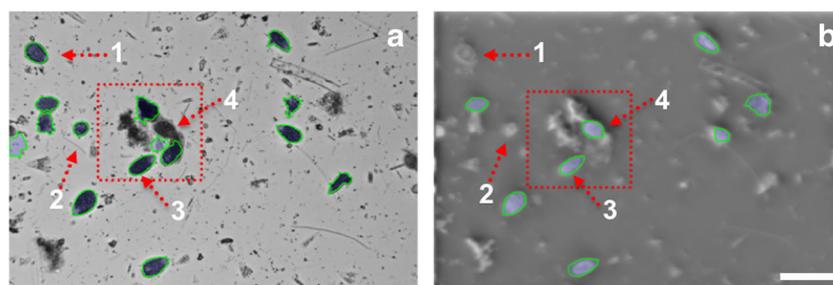


Fig. 6 Marine field sample analysed by 2D and TIE-QPM strategies. Results of the localization/segmentation steps on **a** a 2D image and **b** the corresponding quantitative phase reconstructed image are shown. Identified cells are marked through *green outline*. *Red arrows* and

boxes indicate objects of interest, differently identified by 2D image segmentation and QP image segmentation. (1) *Prorocentrum lima*, (2) *Coolia monotis*, (3 and 4) *O. ovata* cells. *Scale bar* = 100 μ m

assessing the power of 2D and QP reconstructed images synergic analysis to obtain a comprehensive object identification.

TIE-QPM effectiveness analysis

Results obtained from 2D image analysis and from 2D and TIE-QPM integration assess the benefit of using a double 2D/TIE-QPM segmentation step, in order to better define the cell shape and the related geometrical descriptors even in extreme samples condition, such as when the background is very complex and the cell borders are not clearly identifiable in the bright-field image (Fig. 7c). In order to evaluate performance related to TIE-QPM exploitation together with 2D image analysis, comparisons have been run on samples from different sampling sites in Italy, Spain and France.

Analysing data reported in Table 7, two clusters can be identified: the first (cases 1, 3, 5, 6, 7, 8) including samples where similar results come from analysis of 2D images and analysis taking advantage of TIE-QPM, and the other (cases 2, 4, 9, 10) grouping samples where contribution of QP acquisition affects analysis. This discrimination reflects sample composition. Samples presenting *O. ovata* cells with well-defined edges (Fig. 7a, b) gain little improvement from TIE-QPM approach exploitation,

and the algorithm working on 2D images is able to retrieve results analogous to those obtained by TIE-QPM integration. Samples composed by a high number of different and overlapped objects (Fig. 7c) take advantage from TIE-QPM acquisition. The utility of TIE-QPM integration is not related to the number of *O. ovata* cell populating the sample, but to the wide presence of macroalgae fragments or other microorganisms or debris in the sample, partially hiding the dinoflagellates. In order to provide evidence about TIE-QPM advantage, Pins Mar, Chiavari and Fiascherino cases were analysed, where QPM power is largely exploited to dramatically refine automated counting. Manual count represents the reference for results goodness, keeping in mind that operator variability can drastically affect output (Culverhouse et al. 2003). As shown in Table 7, manually inspected images showed the presence of 4713, 2630 and of 839 *O. ovata* cells respectively for samples of Pins Mar, Chiavari and Fiascherino. When considering just 2D image analysis, the system retrieved an automated count of 4276, 2331 and 646 *O. ovata* cells for Pins Mar, Chiavari and Fiascherino respectively, largely down-estimating the real quantity of *O. ovata* cells in the samples estimated by manual counting (absolute errors 437, 299 and 193,

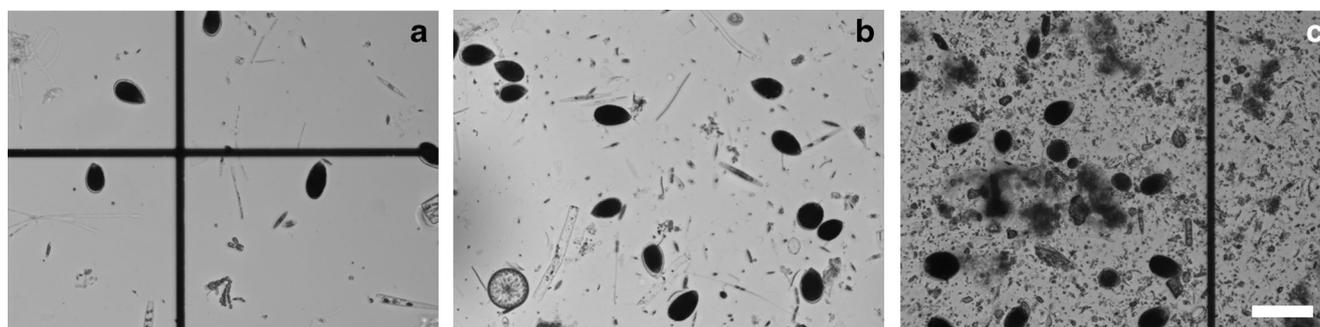


Fig. 7 Representative 2D images showing different composition of samples. **a** Sample (from Haliotis, France) including rare and well-defined *O. ovata* cells, without debris, sand nor other algal fragments. **b** Sample (Rochambeau, France) showing large presence of *O. ovata* cells

in a quite uniform and clean substrate. **c** Sample (from Chiavari, Italy) presenting a mix of elements surrounding *O. ovata* cells which diminish the possibility of dinoflagellate detection. *Scale bar* = 100 μ m

Table 7 Cell counts from 2D automated counting, 2D/TIE-QPM integrated automated counting and a manual inspection of samples from different sampling sites in Spain, France and Italy

Site		2D AC	2D / TIE - QPM IAC	MI	N
1	Blau Mar-Spain	1922	1924	1745	508
2	Pins Mar-Spain	4276	4354	4713	509
3	Rochambeau-France (1)	713	717	704	323
4	Rochambeau-France (2)	4246	4283	4175	509
5	Rochambeau-France (3)	96	96	12	608
6	Haliotis-France (1)	35	35	25	504
7	Haliotis-France (2)	882	886	864	494
8	Genova Quarto-Italy	437	440	378	509
9	Chiavari-Italy	2331	2465	2630	507
10	Fiascherino-Italy	646	744	839	608

AC automated counting, IAC integrated automated counting, MI manual inspection, N number of images analysed

respectively for the three sites; relative errors 9, 11 and 23%, respectively, for the three sites).

Exploiting information from phase reconstructed images, the system counts 4354, 2465 and 744 *O. ovata* for samples collected in Pins Mar, Chiavari and in Fiascherino, with a decrease of both absolute (359, 165 and 95 elements, respectively) and relative (7.6, 6.2 and 11%, respectively) errors. As shown in these cases, a loss of information can be recovered by exploiting TIE-QPM data integration.

Discussion

Monitoring of aquatic microscopic species is a necessary step for controlling HABs or population trends of endangered species. Generally, monitoring is mainly run by visual inspection, which is a repetitive, boring and tiring job and highly prone to errors. In the last decade, many tools have been proposed for automating the process. In this paper, the OvMeter tool is presented: it performs automated counting of *O. ovata* cells and returns algal cell concentration estimation, in order to help monitoring HAB and to prevent risks to human health. The development of a new device was necessary since the already existing tools for other marine species rely on features not suitable for quickly determining the concentration of *O. ovata*. In particular, for *O. ovata*, image-based classification fluorescence detection, used as a marker in many recognition methods for marine microorganisms (Rodenaeker et al. 2006; Sosik and Olson 2007; Schulze et al., 2013), cannot be exploited since Lugol's acidic solution, suggested by international guidelines for preserving *O. ovata*, drastically reduces the strength of the fluorescence signal. Moreover, the drop-like shape of *O. ovata* cells is peculiar and rarely compatible with already developed segmentation approaches. In order to

reduce the burden of routine identification and quantification of *O. ovata* HAB, OvMeter, a low-cost, automated, embedded opto-electronic system, was developed and here presented.

OvMeter tool allows the acquisition of automated, timely, reliable and reproducible results, and includes valuable features that make it suitable for *O. ovata* monitoring. OvMeter aims to substitute human cell counting, maintaining sample treatment procedures totally analogous to guidelines suggested by ISPRA-ARPA (2012) and by European operational (<http://m3-habs.net/guidelines/>) protocols. This allowed a reasonable comparison among OvMeter results and human inspection output. OvMeter performance was measured (Table 3). An excellent reliability of OvMeter was highlighted in (1) correctly identifying the target cells and discriminating them from similar objects (sensitivity and specificity), (2) providing consistent results when the experiment is repeated (precision, F1 score), (3) reaching a good degree of closeness to the true value (accuracy), and (4) avoiding the positive labelling of non-relevant data during classification (fall-out). Samples showing higher concentrations of *O. ovata*, more critical when assessing intoxication risks, are subjected to small errors (1–3%), indicating reliability of the system in monitoring critical conditions. Moreover, OvMeter is able to retrieve repeatable and unbiased results, and could be easily run even by non-expert operators.

The high level of performance shown by OvMeter is ascribable majorly to the exploitation of TIE-QPM technique, which provides quantitative phase reconstructed images computed from a Z-stack acquisition of out-focus bright-field images solving TIE, in synergistic cooperation with 2D images analysis. This represents one of the most distinctive aspects of OvMeter tool and is added value. TIE-QPM is a non-interferometric technique belonging to the emerging QPI methodologies, such as single-exposure online digital holography (Javidi et al. 2005) and off-axis digital holography microscopy (El Mallahi et al. 2013; Zetsche et al. 2014), which gain quantitative information on physical and chemical properties of semi-transparent samples by optical phase shift of the light radiation passing through. Commonly, QPM images add information related to thickness and refractive index fluctuations of the examined organism, which provide knowledge about the presence of organelles or other internal structures (Mir et al. 2012, Zetsche et al. 2014). In the described procedure, dinoflagellates' internal structures are obscured by the exploitation of Lugol's acidic solution. Nevertheless TIE-QPM remains a valid strategy, because it improves cells' edges definition (increasing contrast with background) and acting in cooperation with 2D bright images; it is able to resolve objects even at critical sample conditions, thus improving counting reliability. In particular, the contribution of TIE-QPM acquisition is carried out in the localization and segmentation steps and in the computation of the shape descriptors in the pipeline scheme of the cell recognition

algorithm. Although phase images do not have a meaningful texture to be exploited, the borders extracted from the phase images can be imposed over the bright-field image, defined over the same pixel grid and used as a new possible segmentation, while the texture information is easily recovered the bright-field images. Thus, the classifier can operate on data coming from both 2D and QP images in the same way, without the need of any specific 2D–QPM training. TIE-QPM integration shows major advantage when working on samples including other microorganisms and debris hiding the dinoflagellates. In these cases, the detection and identification of *O. ovata* appeared to be remarkably improved, diminishing distance among automatic results and manual count (taken as reference for results goodness). It is worth to note that no cases showed identical values of manual and automated results: this can be ascribed to operator variability which can drastically impact results. Culverhouse et al. (2003) reported that the performance of human operators in identifying and counting correctly cells is affected by several psychological factors such as fatigue and boredom, bias from expectations and short-term memory limit, and it is estimated to be about 70%. The benefit of using a double 2D/TIE-QPM segmentation step, in order to define cell localization and the related geometrical descriptors was showed in the comparative analysis running on samples from different sampling sites in Italy, Spain and France collected during the summer 2015. In particular, samples collected in Pin Mar, Chiavari and Fiascherino sites were composed by a mix of elements including debris and various microorganisms other than *O. ovata* hiding the dinoflagellate cells: these samples represent good case studies where TIE-QPM power exploitation dramatically refine automated count.

Finally, it is worth noting that OvMeter works as a unique, embedded, built-in system, where software and hardware components cooperate to perform a well-defined task. This brings a set of advantages. It enhances performance in terms of time consumption, since image acquisition and processing steps can be run in parallel (i.e. it can acquire the next image while the current one is being processed). This approach speeds up the counting results, thus permitting to monitor coasts more frequently and widely, and enabling a fast response to the general population in case of intoxication risks. Moreover, the automatic policy of the system can decide to stop image acquisition and cell counting when reaching a sufficiently reliable value (i.e. number of *O. ovata* cells is either very low or very stable, and the number of processed images reaches a predefined threshold to avoid under-sampling) before the analysis of the entire counting would have been completed.

In conclusion, in this paper, the OvMeter tool is presented as an integrated optical system devoted to monitoring HABs of *O. ovata*. This tool is able to measure the concentration of *O. ovata* present in marine field samples, automatically and

independently from the skills of the operator. The main advantage of the system is that it fully adheres to international guidelines concerning algal sampling and analyses and works in the complete absence of a human operator. The developed algal recognition tool relies on a machine learning approach (i.e. BT) and in order to enhance the counting performance, the implemented algorithm was fed with 2D and QP images. The integration of these two data types improves object identification and facilitates target recognition, assessing performance similar to those achieved by human inspection and shortening time. Finally, it is worth noting that the developed counting algorithm was finely tuned to *O. ovata* features, thus exploiting the shape of the species to be searched. However, the system can be adjusted to recognize other microorganisms, if a novel template set and new training and test datasets for classifications are produced.

Acknowledgements This work was supported by the ENPI Project: “Risk Monitoring, Modelling and Mitigation of Benthic Harmful Algal Blooms along Mediterranean coasts (M3-HABs)—II-B/2.1/0096”; by Regional program FAS 2007-2013 Progetto 4 “Programma triennale per la ricerca e l’innovazione: progetti integrati ad alta tecnologia”; and by PO-CRO fondo sociale europeo regione Liguria 2007-2013.

The authors wish to thank Sitem srl Italy, for the support in software development, and VacuumFAB srl Italy for the opto-mechanical design.

References

- Accoroni S, Romagnoli T, Pichierri S, Colombo F, Totti C (2012) Morphometric analysis of *Ostreopsis cf. ovata* cells in relation to environmental conditions and bloom phases. *Harmful Algae* 19:15–22
- Barty A, Nugent KA, Paganin D, Roberts A (1998) Quantitative optical phase microscopy. *Opt Lett* 23:817–819
- Bernal AJ, Ferrando SE, Bernal LJ (2008) Cell recognition using wavelet templates. *Proc. Canadian Conference on Electrical and Computer Engineering* 1219–1222
- Besada EG, Loeblich LA, Loeblich AR (1982) Observations on tropical, benthic dinoflagellates from ciguatera-endemic areas: *Coolia*, *Gambierdiscus*, and *Ostreopsis*. *Bull Mar Sci* 32: 723–735
- Bishop C (2007) *Pattern recognition and machine learning*. Springer, Heidelberg
- Boykov YY, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans Pattern Anal Mach Intell* 26:1124–1137
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Burges CJC (1998) A tutorial on support vector machines for pattern recognition. *Data Min Knowl Disc* 2:121–167
- Ciminiello P, Dell’Aversano C, Fattorusso E, Forino M, Magno GS, Tartaglione L, Melchiorre N (2006) The Genoa 2005 outbreak. Determination of putative palytoxin in Mediterranean *Ostreopsis ovata* by a new liquid chromatography tandem mass spectrometry method. *Analyt Chem* 78:6153–6159
- Ciminiello P, Dell’Aversano C, Fattorusso E, Forino M, Tartaglione L, Grillo C, Melchiorre N (2008) Putative palytoxin and its new analogue, ovatoxin-a, in *Ostreopsis ovata* collected along the Ligurian coasts during the 2006 toxic outbreak. *J Am Soc Mass Spectrom* 19:111–120

- Ciminiello P, Dell'Aversano C, Dello Iacovo E, Fattorusso E, Forino M, Grauso L, Vanucci S (2012) Isolation and structure elucidation of ovatoxin-a, the major toxin produced by *Ostreopsis ovata*. *J Am Chem Soc* 134:1869–1875
- Coltelli P, Barsanti L, Evangelista V, Frassanito AM, Gualtieri P (2014) Water monitoring: automated and real time identification and classification of algae using digital microscopy. *Environ Sci Process Impacts* 16:2656–2665
- Culverhouse PF, Williams R, Reguera B, Herry V, González-Gil S (2003) Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Mar Ecol Prog Ser* 247:17–25
- Egmont-Petersen M, de Ridder D, Handels H (2002) Image processing with neural networks—a review. *Pattern Recogn* 35:2279–2301
- El Mallahi A, Minetti C, Dubois F (2013) Automated three-dimensional detection and classification of living organisms using digital holographic microscopy with partial spatial coherent source: application to the monitoring of drinking water resources. *Appl Opt* 52:A68–A80
- Embleton KV, Gibson CE, Heaney SI (2003) Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method. *J Plankton Res* 25:669–681
- Engl HW, Hanke M, Neubauer A (1996) Regularization of inverse problems. Kluwer Academic Publishers, Dordrecht
- Escalera L, Benvenuto G, Scalco E, Zingone A, Montresor M (2014) Ultrastructural features of the benthic dinoflagellate *Ostreopsis cf. ovata* (Dinophyceae). *Protist* 165:260–274
- Faimali M, Giussani V, Piazza V, Garaventa F, Corrà C, Asnaghi V, Privitera D, Gallus L, Cattaneo-Vietti R, Mangialajo L, Chiantore M (2012) Toxic effects of harmful benthic dinoflagellate *Ostreopsis ovata* on invertebrate and vertebrate marine organisms. *Mar Environ Res* 6:97–107
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Statist* 29:1189–1232
- Giussani V, Sbrana F, Asnaghi V, Vassalli M, Faimali M, Casabianca S, Penna A, Ciminiello P, Dell'Aversano C, Tartaglione L, Mazzeo A, Chiantore M (2015) Active role of the mucilage in the toxicity mechanism of the harmful benthic dinoflagellate *Ostreopsis cf. ovata*. *Harmful Algae* 44:46–53
- Gjeci N (2015) Master thesis, University of Genova
- Gureyev TE, Nugent KA (1997) Rapid quantitative phase imaging using the transport of intensity equation. *Opt Commun* 133:339–346
- Hilgemann DW (2003) From a pump to a pore: how palytoxin opens the gates. *Proc Natl Acad Sci* 100:386–388
- Honsell G, De Bortoli M, Boscolo S, Dell'Aversano C, Battocchi C, Fontanive G, Penna A, Berti F, Sosa S, Yasumoto T, Ciminiello P, Poli M, Tubaro A (2011) Harmful dinoflagellate *Ostreopsis cf. ovata* Fukuyo: detection of ovatoxins in field samples and cell immunolocalization using antipalytoxin antibodies. *Env Sci Technol* 45:7051–7059
- Honsell G, Bonifacio A, De Bortoli M, Penna A, Battocchi C, Ciminiello P, Dell'Aversano C, Fattorusso E, Sosa S, Yasumoto T, Tubaro A (2013) New insights on cytological and metabolic features of *Ostreopsis cf. ovata* Fukuyo (Dinophyceae): a multidisciplinary approach. *PLoS One* 8:e57291
- Hu MK (1962) Visual pattern recognition by moment invariants. *IRE Trans Information Theory* 8:179–187
- ISPRA-ARPA (2012) Monitoraggio di *Ostreopsis ovata* e *Ostreopsis* spp.: Protocolli operativi. Quad – R MAR 5/2012. <http://www.isprambiente.gov.it/publicazioni/quaderni/ricerca-marina/monitoraggio-di-ostreopsis-ovata-e-ostreopsis-spp.-protocolli-operativi>
- Javidi B, Moon I, Yeom S, Carapezza E (2005) Three dimensional imaging and recognition of microorganism using single exposure on-line (SEOL) digital holography. *Opt Express* 13:4492–4506
- Kreshchenovskaya MA, Orlova TY (2014) The ultrastructure of the dinoflagellate *Ostreopsis cf. ovata* Fukuyo, 1981 (Dinophyceae) from the sea of Japan. *Russ J Mar Biol* 40:273–278
- Mir M, Bhaduri B, Wang R, Zhu R, Popescu G (2012) Quantitative phase imaging. *Prog Optics* 57:133–217
- Mosleh MAA, Manssor H, Malek S, Milow P, Salleh A (2012) A preliminary study on automated freshwater algae recognition and classification system. *BMC Bioinformatics* 13:S17–S25
- Nguyen MH, De la Torre F (2010) Optimal feature selection for support vector machines. *Pattern Recogn* 43:584–591
- Nugent K, Paganin D, Barty A (2000) Phase determination of a radiation wave field. Patent: CA 2348912:A1
- Orabona F, Castellini C, Caputo B, Jie L, Sandini G (2010) On-line independent support vector machines. *Pattern Recogn* 43:1402–1412
- Rodenacker K, Hense B, Jütting U, Gais P (2006) Automatic analysis of aqueous specimens for phytoplankton structure recognition and population estimation. *Microsc Res Tech* 69:708–720
- Schulze K, Tillich UM, Dandekar T, Frohme M (2013) PlanktoVision—an automated analysis system for the identification of phytoplankton. *BMC Bioinformatics* 14:115
- Sok HK, Ooi MPL, Kuang YC (2016) Multivariate alternating decision trees. *Pattern Recogn* 50:195–209
- Sosik HM, Olson RJ (2007) Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol Oceanogr Methods* 5:204–216
- Tao J, Cheng W, Boliang W, Jiezheng X, Nianzhi J, Tingwei L (2010) Real-time red tide algae recognition using SVM and SVDD. Intelligent computing and intelligent systems (ICIS). *IEEE International Conference* 1:602–606
- Teague MR (1983) Deterministic phase retrieval: a Green's function solution. *J Opt Soc Am* 73:1434–1441
- Tichadou L, Glaizal M, Armengaud A, Gressel H, Lemée R, Kantin R, Lasalle JL, Drouet G, Rambaud L, Malfait P, de Haro L (2010) Health impact of unicellular algae of the *Ostreopsis* genus blooms in the Mediterranean Sea: experience of the French Mediterranean coast surveillance network from 2006 to 2009. *Clin Toxicol* 48:839–844
- Totti C, Accoroni S, Cerino F, Cucchiari E, Romagnoli T (2010) *Ostreopsis ovata* bloom along the Conero Riviera (northern Adriatic Sea): relationships with environmental conditions and substrata. *Harmful Algae* 9:233–239
- Utermöhl H (1958) Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitt Int Ver Theor Angew Limnol* 9:1–38
- Verikas A, Gelzinis A, Bacauskiene M, Olenina I, Olenin S, Vaiciukynas E (2012) Automated image analysis- and soft computing-based detection of the invasive dinoflagellate *Prorocentrum minimum* (Pavillard) Schiller. *Expert Syst Appl* 39:6069–6077
- Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. *IEEE Trans. Evol Comput* 1:67–82
- Zuiderveld K (1994) Contrast limited adaptive histogram equalization. In: *Graphics Gems IV*, San Diego: Academic Press
- Zetsche EM, El Mallahi A, Dubois F, Yourassowsky C, Kromkamp JC, Meysman FJR (2014) Imaging-in-flow: digital holographic microscopy as a novel tool to detect and classify nanoplanktonic organisms. *Limnol Oceanogr* 12:757–775Methods