

## Imaging-in-Flow: Digital holographic microscopy as a novel tool to detect and classify nanoplanktonic organisms

Eva-Maria Zetsche<sup>1\*</sup>, Ahmed El Mallahi<sup>2</sup>, Frank Dubois<sup>2</sup>, Catherine Yourassowsky<sup>2</sup>, Jacco C. Kromkamp<sup>3</sup>, and Filip J.R. Meysman<sup>1,3</sup>

<sup>1</sup>Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

<sup>2</sup>Microgravity Research Center, Université Libre de Bruxelles, 50 Av. F. Roosevelt, CP 165/62, B-1050 Brussels, Belgium

<sup>3</sup>Royal Netherlands Institute for Sea Research, P.O. Box 140, 4400 AC Yerseke, The Netherlands

### Abstract

Traditional taxonomic identification of planktonic organisms is based on light microscopy, which is both time-consuming and tedious. In response, novel ways of automated (machine) identification, such as flow cytometry, have been investigated over the last two decades. To improve the taxonomic resolution of particle analysis, recent developments have focused on “imaging-in-flow,” i.e., the ability to acquire microscopic images of planktonic cells in a flow-through mode. Imaging-in-flow systems are traditionally based on classical bright-field microscopy and are faced with a number of issues that decrease the classification performance and accuracy (e.g., projection variance of cells, migration of cells out of the focus plane). Here, we demonstrate that a combination of digital holographic microscopy (DHM) with imaging-in-flow can improve the detection and classification of planktonic organisms. In addition to light intensity information, DHM provides quantitative phase information, which generates an additional and independent set of features that can be used in classification algorithms. Moreover, the capability of digitally refocusing greatly increases the depth of field, enables a more accurate focusing of cells, and reduces the effects of position variance. Nanoplanktonic organisms similar in shape were successfully classified from images captured with an off-axis DHM with partial coherence. Textural features based on DHM phase information proved more efficient in separating the three tested phytoplankton species compared with shape-based features or textural features based on light intensity. An overall classification score of 92.4% demonstrates the potential of holographic-based imaging-in-flow for similar looking organisms in the nanoplankton range.

Phytoplankton are a vital component of all aquatic ecosystems, as they play key roles in the biogeochemical cycling of oceans, rivers, and lakes, and sustain the higher trophic levels in the food web of these ecosystems (Culverhouse 2007). The phytoplankton encompass a diverse group of organisms, including diatoms, cyanobacteria, dinoflagellates, and coccolithophorids, which serve as indicators of environmental and climate change, whether natural or human-induced (Lee and Lee 2009; Pomati et al. 2011; Schaap et al. 2012). Quantitative information on the abundance, diversity, and dynamics of these different taxonomic groups within the phytoplankton is therefore essential (Jalba et al. 2004; Malkasian et al. 2011; Pereira and Ebecken

2011). Apart from ecological questions targeting bloom dynamics, species succession, and spatial and temporal patchiness, the monitoring of phytoplankton composition is also crucial for many economically important applications, such as the analysis of ballast water (Ruiz et al. 2000), the development of algal-based biofuels (Singh and Gu 2010), and the monitoring of harmful algal blooms (Schaap et al. 2012).

Techniques for the determination of phytoplankton composition are inherently characterized by a trade-off between processing speed (i.e., the number of cells that can be classified per unit of time) and taxonomic resolution (i.e., the lowest taxonomic level to which cells can be classified). Historically, the determination of phytoplankton composition has been based on detailed microscopic analysis, which is capable of high taxonomic resolution down to cell sizes of approximately 5  $\mu\text{m}$ . Microscopy can be performed in a variety of modes (transmitted light, scattered light, fluorescence, phase contrast, etc.), and each of these “imaging” modes provides distinct and complementary information about the morphology and internal

\*Corresponding author: E-mail: eva-maria.zetsche@nioz.nl  
Current address: Royal Netherlands Institute for Sea Research, P.O. Box 140, 4400 AC Yerseke, The Netherlands

### Acknowledgments

Full text appears at the end of the article.

DOI 10.4319/lom.2014.12.757

subcellular structure of phytoplankton cells. Such detailed information has the advantage that cells can be classified to a low taxonomic level. However, conventional (i.e., not automated) microscopic analysis is extremely labor intensive and time consuming, which strongly limits the throughput of samples. This way, it is not suitable for the high-frequency monitoring of phytoplankton communities as required in modern day field applications (Walker and Kumagai 2000; Embleton et al. 2003; Culverhouse 2007; Pereira and Ebecken 2011).

At the opposite end of the spectrum is flow cytometry, which allows a high throughput of cells in an automated fashion, but collects less detailed information on the structure and identity of the cells. In flow cytometry, cells are characterized in terms of fluorescence signatures and scattering properties (Collier 2000; Schaap et al. 2012). Fluorescence intensity provides information on the fluorescent pigments produced by the cells, whereas the degree of scatter is correlated to the size of the cell (forward scattering) and the cell's granularity or internal complexity (90° light scattering or side scattering). This allows the automated classification of planktonic cells into functional groups, but does not typically allow identification at lower taxonomic levels (Rutten et al. 2005; Pereira and Ebecken 2011). Yet, the strength of flow cytometry is its ability to rapidly and automatically analyze many populations of cells (Rutten et al. 2005). Because of this statistical power, flow cytometry has had a strong impact on phytoplankton composition studies, ever since it was introduced into the aquatic sciences a few decades ago (Yentsch et al. 1983).

More recently, new instruments are being developed that try to link both ends of the throughput-resolution spectrum, that is, enable rapid and automated identification of plankton together with high taxonomic resolution (Sieracki et al. 2010). One line of development involves enhanced flow cytometry techniques, such as Cytobuoy (Dubelaar and Gerritzen 2000) and the laser scanning flow cytometer CLASS (Spizzichino et al. 2011), which collect a more sophisticated optical fingerprint of the cell instead of only simple morphometrics (Malkasian et al. 2011). A separate line of development, represented by systems such as FlowCAM (Sieracki et al. 1998; Zarauz et al. 2009) and Imaging FlowCytobot (Olson and Sosik 2007) take flow analysis one step further by performing actual microscopic imaging. In these so-called 'imaging-in-flow' techniques, a digital image is captured for each particle encountered in flow, from which various features are subsequently extracted by image analysis. This feature information is then processed by a classification algorithm for identification and enumeration of cell types within a sample (Sosik and Olson 2007; Poulton and Martin 2010).

Both the FlowCAM and the Imaging FlowCytobot are based on classical brightfield microscopy. The Imaging FlowCytobot is designed for the study of natural phytoplankton and microzooplankton assemblages in the size range of ~ 10 to 100  $\mu\text{m}$  (Sosik and Olson 2007). The FlowCAM system is able to cover a wider range of organisms from large (~1 mm) to relatively

small (~1  $\mu\text{m}$ ), but is more restrictive, as cells need to have a distinct morphology to be readily identified (Poulton and Martin 2010).

Imaging-in-flow systems aim to combine the statistical power and high throughput of standard flow cytometry with the quantitative identification capabilities of light microscopy. However, the challenges associated with imaging cells in a flow are many, both in terms of data acquisition (e.g., imaging every single cell in a flow stream, producing imagery with high enough quality and spatial resolution, dealing with projection variance and out-of-focus objects) as well as in terms of data analysis (e.g., processing large numbers of images, accurate detection of organism boundaries within an image, automated classification). Yet recent advances in CCD camera technology, optical filtration, and digital computing have resulted in a continuous improvement of the capabilities and reliability of imaging-in-flow techniques, making them promising tools for phytoplankton composition studies (Sosik and Olson 2007).

Here, our goal is to address some of the issues that are currently encountered when using imaging-in-flow for the identification and classification of phytoplankton cells. More specifically, we will examine the use of digital holographic microscopy (DHM) as an imaging-in-flow technique. Where traditional light microscopy records the intensity image of the object, DHM records the interference of a light beam passing through an object (the object beam) with the light of a reference beam (Dubois et al. 1999). The resulting interference pattern is digitally recorded as a hologram, and subsequently, the object image is digitally reconstructed by a computer using a numerical reconstruction algorithm (Zhang and Yamaguchi 1998).

Holography was invented in the late 1940s (Gabor 1948) and was first applied to visualize marine plankton in the 1960s (Knox 1966). However, these early studies were essentially proof-of-concept demonstrations. Only in recent years have digital image sensors and computers become powerful enough to reconstruct images with sufficient speed, resolution, and quality. This has recently led to the development of in situ instruments that use digital in-line holography to investigate particles and organisms in the water column, e.g., the 'holo-sub' by Pfitsch et al. (2005) or a 'point-source digital inline holographic microscope' (Bochdanský et al. 2013). These instruments are based on an in-line holographic configuration, where the interference pattern results from the light that is scattered by the object with the unscattered beam that surrounds the object (i.e., object and reference beams are collinear). In-line holography has the advantage that it requires a simple, inexpensive optical setup, but also has some inherent drawbacks (the presence of a ghost image accompanying any refocused object, less accurate optical phase information, a high noise level inherent to the use of a coherent laser beam). These issues can be resolved by combining an off-axis DHM configuration, where object and reference beams

are separated and recombined on the camera sensor, with partial coherent illumination (Dubois et al. 1999; Dubois et al. 2004). Recently, such an off-axis DHM configuration has been implemented to study particles in a split cell (Dubois et al. 2006a), detect and discriminate *Giardia lamblia* cysts for drinking water applications (El Mallahi et al. 2012), and to record a high throughput of plankton in flow (Yourassowsky and Dubois 2014). The aim here is to investigate how off-axis DHM with partial coherence can improve the detection and classification of phytoplankton, and specifically nanoplanktonic organisms.

With respect to the analysis of phytoplankton composition, DHM has two principle advantages over classical light microscopy: (1) Improved object classification. Compared with classical microscopy, off-axis DHM not only records the light intensity information, but also the phase information of an object. This phase information allows a two-dimensional visualization of the optical thickness of an object, which provides additional detail about the internal structure of the cell. This additional textural information improves the resolution of classification (El Mallahi et al. 2013). (2) Improved object localization. A second unique feature of DHM is the ability to adjust the focus after the image is recorded, since all focus planes are recorded simultaneously by the hologram (Dubois et al. 1999). Hence, DHM has the ability to refocus objects recorded out of focus and thus allows a much greater depth of focus compared to classical microscopy (~70-fold larger depth of focus than a light microscope using a 63× objective lens). This increase in depth of focus allows that larger sample volumes can be analyzed per unit of time (i.e., increase the throughput) (Yourassowsky and Dubois 2014). In this study, we will investigate whether and how these two advantages of DHM can improve the classification of nanoplankton.

## Materials and procedures

Samples of pure nanoplankton cultures were used in so-called supervised classification experiments, i.e., one knows which type of cells are present, and one tests whether the classification method gives the same result. In a classification experiment, cells are characterized in terms of a set of features (i.e., parameters that enable classification), and based on these features, the decision is made to catalogue the cell into a given class. In supervised classification, the identity of the cells in the sample is known a priori. Accordingly, one knows exactly to which class a given cell should be classified, and this way, one can test the performance of the classification method. At its best, the classification method should achieve a classification score of 100% (all cells are correctly classified). In this study, supervised classification experiments were performed with both flow cytometry (the classical reference method in plankton studies) and digital holographic microscopy. Three nanoplankton species were used that are very similar in shape and size and are not readily distinguishable by classical microscopy.

## Nanoplanktonic cell cultures

Three separate single-celled nanoplankton species were used in the supervised classification experiments: two cyanobacterial species, *Cyanothece* sp. and *Stanieria* sp., and the green algae *Chlorella autotrophica* (hereafter referred to as *Chlorella*). These three species were specifically chosen because (1) they show a close similarity in size, shape, and color, and (2) they have a simple spherical shape (~5 µm diameter) without further discriminatory morphological features, which makes classification with classical light microscopy inherently difficult. These three species are found in a range of diverse marine and freshwater habitats (De Philippis et al. 2005; Bandyopadhyay et al. 2011; Mansoor et al. 2011), and can co-occur within the same water body (Descy et al. 2006). Cultures were obtained from the Culture Collection Yerseke (NIOZ-Yerseke). The two cyanobacteria *Cyanothece* (CCY 0408) and *Stanieria* (CCY 0820) were grown at 27°C with 20 µmol photons m<sup>-2</sup> s<sup>-1</sup> light on a 12L:12D light-dark cycle in a T° medium. The green algae *Chlorella* (CCY 9931) was grown at 14°C with 60 µmol photons m<sup>-2</sup> s<sup>-1</sup> light on a 16L:8D cycle in a MDV-Si medium. Note that these culturing conditions are different between species, as they comprise the optimal growth conditions for the individual species. This way we excluded potential effects of nutrient or light stress on cell morphology and classification features. In an additional experiment, we tested the robustness of classification features under different growth conditions, and hence, the impact of growth conditions on classification performance (see below).

## Flow cytometry

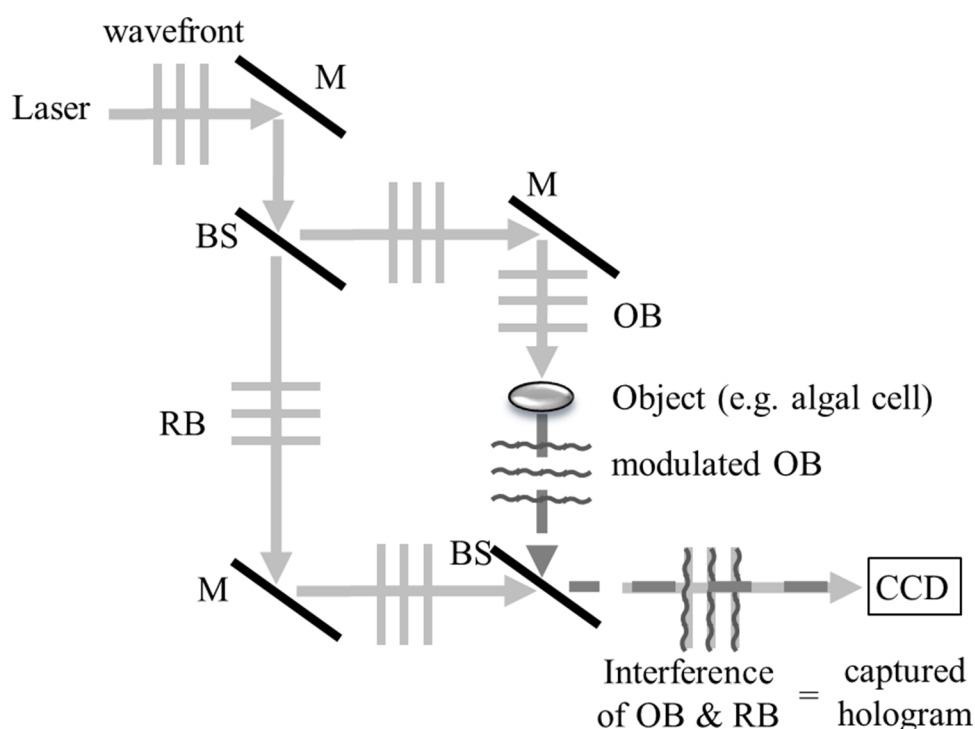
Samples of 50 µL (~900-1000 cells) from each of the three cultures (*Cyanothece* sp., *Stanieria* sp., and *Chlorella autotrophica*) were run on a BD FACSVerse™ flow cytometer with a 3-laser configuration. This allowed the following parameters to be measured: forward scattering (FSC) and side scattering (SSC); fluorescence excited with a blue laser and emission measured in the red (FBR), in the green (FBG) and in the orange (FBO) regions. Flow cytometer results were analyzed using EasyClus v1.18 software (Thomas Rutten Projects, NL).

## Digital holographic microscopy

### Hardware configuration

The principle of digital holographic microscopy (DHM) using an off-axis Mach-Zehnder interferometry configuration is based on a laser beam that is split into two separate components: a so-called reference beam, which travels unhindered from its source to the CCD camera used for the image capture, and a so-called object beam, which travels through the object or sample of interest (Fig. 1). The object beam interferes with the reference beam to create an interference image or hologram, which is recorded by the CCD camera.

The off-axis DHM hardware configuration with partially coherent light used in this study has previously been described in detail (Dubois et al. 2006b; El Mallahi et al. 2013) and is therefore only briefly discussed here. The DHM system uses a coherent light source (red monomode laser diode:  $\lambda =$



**Fig. 1.** Conceptual diagram of the set-up used for off-axis digital holographic microscopy (Mach-Zehnder configuration). A wavefront of light (a laser) is split at a beamsplitter (BS) into an object beam (OB) and a reference beam (RB). The OB passes through the object of interest and is thereby modulated. The RB and OB are then interfered with each other to create a hologram (interference image) on the image sensor (CCD camera). M: Mirror.

635 nm), which is transformed into a partial coherent light source by guiding the incident beam through a rotating ground glass. This use of partially coherent light strongly reduces the speckle noise in the digital holograms, and hence, substantially improves the image quality (Dubois et al. 1999, 2004). In all experiments, we used a JAI CV-M4 camera (with a CCD array of  $1280 \times 1024$  pixels, cropped to a  $1024 \times 1024$  pixel window) and Leica objective lenses of  $63\times$  magnification (numerical aperture [NA] = 0.70). This setup provided a field of view of  $144 \mu\text{m} \times 115 \mu\text{m}$  (subsequently cropped to  $115 \times 115 \mu\text{m}$  during image analysis). All experiments were carried out with a custom-built breadboard version of the microscope configuration displayed in Fig. 1.

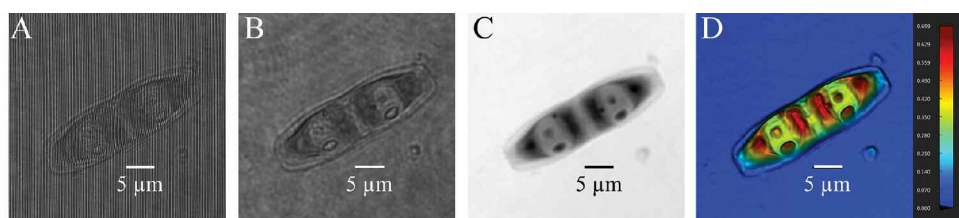
#### Hologram digital processing

To illustrate the procedure of hologram processing, Fig. 2 shows a set of images displaying a benthic diatom and its intracellular structure (imaged in static mode). The actual hologram, as captured by the CCD camera of the DHM setup is shown in Fig. 2A. This hologram shows the characteristic fringe pattern, which results from the interference of the object and reference beams. From this primary hologram, two secondary images are extracted by means of a numerical algorithm as described in detail in Dubois et al. (1999). First, the light intensity (LI) image of the diatom (Fig. 2B) quantifies the amount of light that is transmitted by the various internal organelles of the algal cell and equates to what would be seen

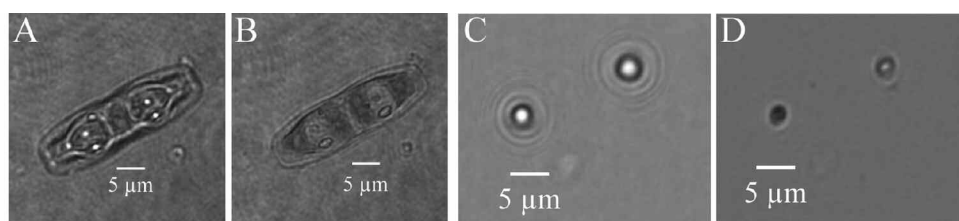
with a classical brightfield light microscope. Second, the phase information (PI) image of the diatom (Fig. 2C) reflects the degree to which light is slowed down upon passing an object, and hence, the PI image quantifies the so-called optical thickness of the internal structures within the diatom. The PI image is composed of a range of gray levels, where each gray level corresponds to a change in optical thickness of 2.5 nm. This high resolution allows the visualization of small differences in optical thickness associated with the internal structure of living cells. Both the LI and PI images can be used for feature extraction in the classification procedure as described below.

#### Hologram digital refocusing

When classical light microscopy is used to image cells in suspension, only the cells that are within the specific focus plane of the microscope, will be captured sharply by the digital camera. As the focus plane is typically narrow ( $\sim 1.4 \mu\text{m}$  for a  $63\times$  objective), this poses challenges when imaging-in-flow, as cells can easily migrate out of the focus plane. Out-of-focus cells do not allow reliable feature extraction, and hence, they cannot be used for classification. As DHM captures both light and phase information, it is possible to bring out-of-focus cells into focus in a digital manner after the image has been acquired using an algorithm based on Kirchhoff–Fresnel propagation equations (Yamaguchi and Zhang 1997; Dubois et al. 1999). This capability of digital refocusing increases the depth of focus of a DHM substantially compared to a classical light microscope ( $\sim 70$ -fold



**Fig. 2.** (A) Hologram of a benthic diatom as obtained with an off-axis digital holographic microscope with partial coherence. This hologram contains both light intensity information (B) as well as phase information (C) representing the optical thickness of the object. The optical thickness of an object is more clearly visualized in the pseudo-3D visualization of the phase image (D) with the color range going from an optical thickness of 0 (blues) up to 0.7  $\mu\text{m}$  (reds) for this diatom.



**Fig. 3.** The procedure of digital refocusing is shown here for the diatom given in Fig. 2 as well as for one of the investigated species—*Chlorella autotrophica* (C-D). A, C: Cells were deliberately captured out of focus from their optimal focus plane (at 10  $\mu\text{m}$  below the correct focus plane). B, D: DHM has allowed the cells to be refocused post-acquisition to their correct focus plane as shown here. In (D), the left *C. autotrophica* cell is now in its correct focus plane, whereas the cell on the right would require further refocusing to achieve its optimal focus plane. All images are light intensity images.

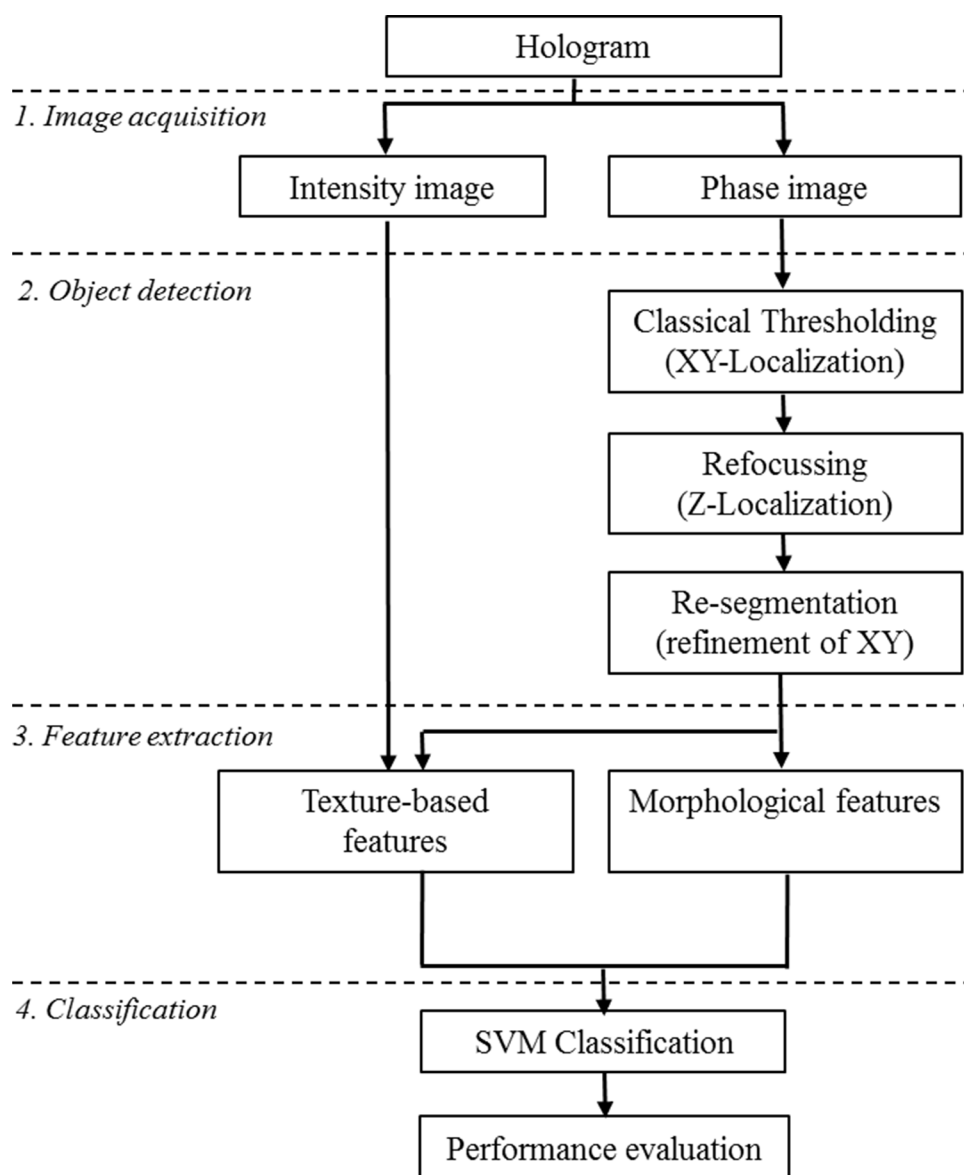
given a 63 $\times$  objective). The procedure of digital refocusing is illustrated in Fig. 3, which shows cells of *Chlorella* as well as a larger diatom cell. In this diatom the internal cell details are more clearly seen. To deliberately put the cells out of focus, the focus plane of the microscope was mechanically offset by  $\sim 10 \mu\text{m}$  in the Z-direction. If captured out of focus, DHM allows the cells to be refocused to their correct focus plane after the hologram has been acquired. This refocusing procedure typically needs to be repeated for different focus planes, as cells are present in different planes. This is exemplified for the *Chlorella* cells in Fig. 3 (C-D) whereby the left lower *Chlorella* cell was originally captured at 10  $\mu\text{m}$  below the correct focus plane, whereas the upper right cell resides in a different plane and remains out of focus. To bring the latter cell into focus, the refocusing procedure has to be repeated with a new refocusing distance. Any feature extraction carried out on a cell that is out of focus (as in Fig. 3A or C) leads to erroneous parameter extraction as described in the Assessment section.

#### Holographic imaging-in-flow and cell classification

A supervised classification was performed to investigate how well DHM could differentiate three nanoplanktonic organisms that are of similar shape and size. For these DHM classification experiments, a sample of a cell culture was channelled through a microfluidic flow cell and every single cell within a given flow volume was imaged. The same total number of cells ( $\sim 200$ ) was analyzed for each of the three nanoplankton species. Since the cell concentrations varied between cell cultures, the sample volume that was analyzed was adjusted accordingly to achieve the same total number.

The whole procedure consisted of four consecutive steps: (1) image acquisition, (2) object detection, (3) feature extraction, and (4) object classification (a diagram of the procedure is shown in Fig. 4). In a first step, primary holograms were captured as video images for each separate culture, and secondary intensity and phase images were extracted for each frame as described above. Then, in the second step, all cells were detected within each image and their correct XYZ-positions were determined, thereby digitally refocusing the cells. This generated a final set of refined LI and PI images, which, in a third step, were used to extract a set of features for each cell. These features are quantitative parameters describing each cell, which were then used in the final and fourth step to classify a given cell. Each of these steps (Fig. 4) is now described in detail below.

**Image acquisition**—Holograms of cell suspensions were captured using the DHM hardware configuration described above. Samples of the cell cultures were injected into Ibidi flow cells (Ibidi  $\mu$ -Slide I with channel dimensions of  $W = 5 \text{ mm}$ ,  $L = 50 \text{ mm}$ ,  $H = 100 \mu\text{m}$ ) attached to a syringe pump (programmable KDS Legato 270P push/pull). The focus plane of the DHM was set to the middle of the flow channel, which allowed the reconstruction of cells over the full height of the flow cell. All cells in a given “frame volume” (i.e., the flow volume captured by a single video image) could therefore be imaged via digital refocusing of a single hologram. Images were taken at a frame rate of 24 frames per second with an exposure time of 200  $\mu\text{s}$ . Dedicated software (Dubois et al. 2006b) was used to extract the LI and PI images from the holograms as described above.



**Fig. 4.** Overview of the four main steps in the image processing and classification procedure. This procedure includes processes to extract the images themselves (1), locating the cells in the dimensions of the flow cell and refocusing them (2), as well as extracting features (3) for use in the classification procedure (4). SVM: Support Vector Machine.

**Object detection**—The second step involved the detection of separate objects within the PI images obtained in Step 1. Each phase image may contain a varying number of cells depending on the cell concentration of the sample. An automated detection procedure was run on each phase image to determine both the XY-plane position as well as the Z-axis position of each cell in the image. First, the detection of an organism in the XY-plane was performed. A classical threshold method (Otsu method) was used to segment the images in order to detect all the cells in each PI image (Sezgin and Sankur 2004). A first set of XY-coordinates for each organism was thus extracted through the centroid of the segmented object. Then, the Z-axis position was determined using the

Robust Refocusing Criterion described in Dubois et al. (2006b) and El Mallahi and Dubois (2011). This Z-axis position was then used to digitally refocus the object. Once the objects were refocused, re-segmenting each object in their newly determined optimal focus plane refined the XY-coordinates. This provided a set of refined intensity and phase images for each object. For each of the three species investigated here, the dataset consisted of ~ 200 segmented objects (i.e., imaged nanoplankton cells) in total.

**Feature extraction**—From the set of refined images, we extracted a set of features for each object. A feature is an image-derived parameter that can be used to perform object classification. Two types of parameters were derived (Gonzalez

**Table 1.** Definitions of features used for the supervised classification. From each cell that has been segmented (segmented region), the detected perimeter (boundary) of the cell is used to extract the following morphometric features: area, perimeter length, major and minor axis, eccentricity, and equivalent circular diameter. For texture-based features, pixel values from within the boundary of the cell are important and allow the calculations of: average intensity and contrast, smoothness, skewness, uniformity and entropy (based on Gonzalez and Woods [2008]).

| Feature                      | Definition   | Equation  |
|------------------------------|--|---|
| Area                         | Nr of pixels within the boundary of the segmented region   |   |
| Perimeter length             | Nr of pixels composing the boundary of the segmented region  |   |
| Major axis                   | Length in pixels of the major axis of an ellipse that has the same normalized second central moments as the segmented region   |   |
| Minor axis                   | Length in pixels of the minor axis of an ellipse with the same normalized second central moments as the segmented region   |   |
| Eccentricity                 | The amount by which the object deviates from a perfect circle (0 = perfect circle, 1 = straight line)  |   |
| Equivalent circular diameter | The diameter of a circle with the equivalent area as that of the object  |   |
| Average intensity            | Mean value of the gray level scale   | $m = \sum_{i=1}^L q_i p(q_i)$                     |
| Average contrast             | Standard deviation of the gray level scale   | $\sigma = \sqrt{\sum_{i=1}^L (q_i - m)^2 p(q_i)}$ |
| Smoothness                   | Measurement of the relative smoothness (R) of the intensity in a region. R equals 0 for a region of constant intensity and increases with the fluctuations in the pixel values (range: 0-1). | $R = 1 - 1/(1 + \sigma^2)$                        |
| Skewness                     | Skewness of the histogram (0 = symmetric histogram, positive value = skew to the right, negative value = skew to the left). Calculated from the third moment.                                | $\mu_3 = \sum_{i=1}^L (q_i - m)^3 p(q_i)$         |
| Uniformity                   | Is at its maximum when all gray levels are equal and decreases from there.   | $U = \sum_{i=1}^L p^2(q_i)$                       |
| Entropy (Randomness)         | Quantifies the randomness of the texture   | $e = \sum_{i=1}^L p(q_i) \log_2 p(q_i)$           |

$q_i$ , random variable indicating intensity

$p(q)$ , histogram of the intensity levels in a region

$L$ , number of possible intensity levels ( $L = 256$  for gray level images)

and Woods 2008): (a) shape or morphometric features, which are based on the external characteristics of the segmented region (the outer shape of the segmented region is used to calculate, e.g., equivalent spherical diameter, major and minor axis, etc.), and (b) textural features, which describe internal characteristics of the segmented region (the gray level of the pixels inside the segmented region are used to calculate e.g., smoothness, coarseness, and regularity). For each object detected, 18 different features were computed: 6 morphometric features, 6 textural features based on the light intensity images, and 6 textural features based on the phase images (see definitions in Table 1). The six morphometric features were area, perimeter length, major and minor axis, eccentricity, and equivalent circular diameter. The six textural features were average gray level, average contrast, smoothness, skewness, uniformity, and entropy. These textural features were calcu-

lated both for the intensity and phase images, adding thus a further 12 values to the classification.

**Classification design and evaluation**—Having completed the first three steps, each cell was tagged with 18 features, which described each of the cells in terms of their internal and external characteristics. The fourth and final step, the classification (or categorization) of the cells, investigated how well this set of descriptors could identify the three different species. The novelty with DHM is that textural features are available both from the intensity and the phase images, thus extending the number of parameters that can be used to classify a given cell.

The classification algorithm used in this study is based on a supervised learning model (machine learning task) where labeled training data are used to predict in which of the known categories new examples fall (see El Mallahi et

al. 2013 for details). Specifically, a support vector machine (SVM) algorithm was used, which is a powerful pattern recognition technique known to give good model performance (Sadeghi et al. 2012). The final outcome of the SVM classification is a so-called confusion matrix as presented in the Assessment section. Such a matrix is a contingency table, where the matrix diagonal stores the number of correctly classified organisms, while off-diagonal matrix elements give the misclassified objects.

Detailed descriptions of SVMs are found elsewhere (Cortes and Vapnik 1995; Van Der Heijden et al. 2004); here only the basic methodology is briefly outlined. The SVM classifier algorithm maps the dataset into a higher-dimensional feature space by using suitable kernel functions (linear or non-linear). The expectation is that the data will be more easily structured and/or separated in this higher-dimensional space. During the SVM mapping, a set of (N-1) hyperplanes (N = number of features; 18 in this study) is built in such a way that the data classes are optimally separated into the N-dimensional feature space. Optimal separation occurs when the distance to the nearest training data points in the hyperplane is maximal for any class (the classification error has been minimized). The multi-class SVM is implemented by training one SVM per class using the one-against-all method (Chang and Lin 2011). The SVM classifier uses the entire set of features and takes also the mutual correlation between the features into account through the different feature spaces.

The performance of the classifier is evaluated with a nested cross-validation method (Devijver and Kittler 1982) composed of an outer and an inner cross-validation. The inner cross-validation selects and optimizes the SVM parameters required for the SVM classifier as described in El Mallahi et al. (2013). The outer cross-validation estimates the misclassification rate and involves a 10-fold cross-validation method (Mosteller and Tukey 1968). This is a good estimate of true accuracy (Davis et al. 2004). The method divided the dataset into 10 randomly chosen subsets. All subsets except one (i.e., in total 9 subsets) were used to build a classifier (training set), against which the tenth subset was tested (testing set). This was reiterated ten times in turn to ensure that each subset was used exactly once as the validation data. The 10 performances of the folds were then averaged to obtain a single robust estimation of the classification performance with its corresponding standard deviation.

#### Statistical analysis of features

Since assumptions of normality and/or homogeneity of variance failed on most of the dataset, non-parametric tests were applied. Given that only some, but not all data distributions lent themselves to data transformations, Kruskal-Wallis one-way analysis of variance by ranks was applied. Post-hoc testing employed Dunn's method. All statistical data analysis was carried out using the statistical analysis packages incorporated into SigmaPlot 12.0 (Systat Software Inc.) with a level of significance of  $P < 0.05$ .

#### Robustness of optical cell properties

An additional experiment was undertaken to examine the effect of growth conditions on features and classification. The goal was to examine whether the optical characteristics of cells are robust enough between different stages and growth conditions so they can be used for classification purposes. To this end, the green algae *Chlorella autotrophica* (CCY 9931) was first grown in nutrient-replete MDV medium (cf. CCY) under normal light conditions ( $30 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) for 1 week before being inoculated into phosphate-deplete MDV medium. Cells were centrifuged twice (2000 rpm; 10 min) and rinsed with P-free MDV medium before they were inoculated in the P-depleted medium. Culturing flasks were set up in triplicate. The cells were maintained in the same light environment and measurements were taken on day 2, 4, 7, and 9 to monitor the cell physiology and perform holographic imaging of cells.

Cell concentrations were determined with a Multisizer 3 Coulter Counter (Beckman Coulter) in triplicate. In vivo whole cell absorption was measured spectrophotometrically (Cary 100 Bio UV-visible Spectrophotometer; Agilent Technologies, equipped with a DRA-CA-3300 integrating sphere [Labsphere]) in the visible light range 350-750 nm in triplicate. Pigment content was also measured spectrophotometrically. For this, duplicate samples of 5 mL were taken from each culturing flask and centrifuged (2500 rpm; 10 min). Subsequently 4.5 mL of the supernatant was carefully removed and replaced with a mix of 80% acetone and 20% DMSO. Samples were placed in the fridge overnight ( $4^{\circ}\text{C}$ ) and centrifuged again (2500 rpm; 10 min) before the supernatant was measured spectrophotometrically. Chlorophyll concentrations were calculated according to Ritchie (2006). The optical absorption cross section  $a^*$  ( $\text{m}^2 [\text{mg chlorophyll } a]^{-1}$ ) was obtained by averaging the absorption between 400 and 700 nm (A) and dividing it by the chlorophyll *a* content after a cm to m and log(10) to natural log conversion according to Kromkamp and Limbeek (1993):  $a^* = [A \times 100 \times \ln(10)/\log(10)]/[\text{chl } a]$ .

A pulse amplitude modulated fluorometer (WATER-PAM, Heinz Walz GmbH) was used to measure the photosynthetic activity of the cells (Maxwell and Johnson 2000; Ralph and Gademann 2005). From the rapid light curves (RLC), the relative photosynthetic electron transport rate (rETR) was calculated as the product of the irradiance (E) and the effective photosystem II (PSII) quantum efficiency ( $\Delta F/F_m' = (F_m' - F)/F_m'$ , where  $F_m'$  and  $F$  are the maximum and steady state fluorescence in the light). Cells were dark acclimated for 15 min to measure the minimum ( $F_0$ ) and maximum ( $F_m$ ) fluorescence after which the RLC was started, using 30 s light steps) using the built-in software (wincontrol) supplied with the PAM. The maximum PSII quantum efficiency  $F_v/F_m$  is calculated as  $(F_m - F_0)/F_m$ . The saturating pulse applied to measure  $F_m$  or  $F_m'$  lasted 0.6 s and had an intensity of approximately  $4500 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ , high enough to close all reaction centers. The RLC were fitted according to Silsbe and Kromkamp (2012) to



obtain the following fit parameters: the maximum rate of rETR ( $rETR_{max}$ ), the initial slope of the RLC  $\alpha$  (a proxy for the photosynthetic efficiency at low light) and  $E_k$  ( $= rETR_{max}/\alpha$ ). Assuming that 50% of the light is absorbed by PSII, the absolute rate of PSII electron transport ETR can then be calculated as  $ETR = rETR \times \alpha \times 0.5$  ( $\mu\text{mol electrons [mg chl } a]^{-1} \text{ s}^{-1}$ ) (Kromkamp and Forster 2003).

A holographic module ('qMod' from Ovizio Imaging Systems NV/SA, Belgium) was mounted on a Zeiss Axioplan and used for holographic imaging of the cells on the respective sampling days at 100 $\times$  oil magnification. Samples were placed on glass slides and 50 cells imaged per replicate. Features including morphological and textural features of light intensity images were extracted as well as textural features of the phase information images in the same manner as for the 3-species classification. Values were normalized to adjust for the difference in holographic imaging system.

### Assessment

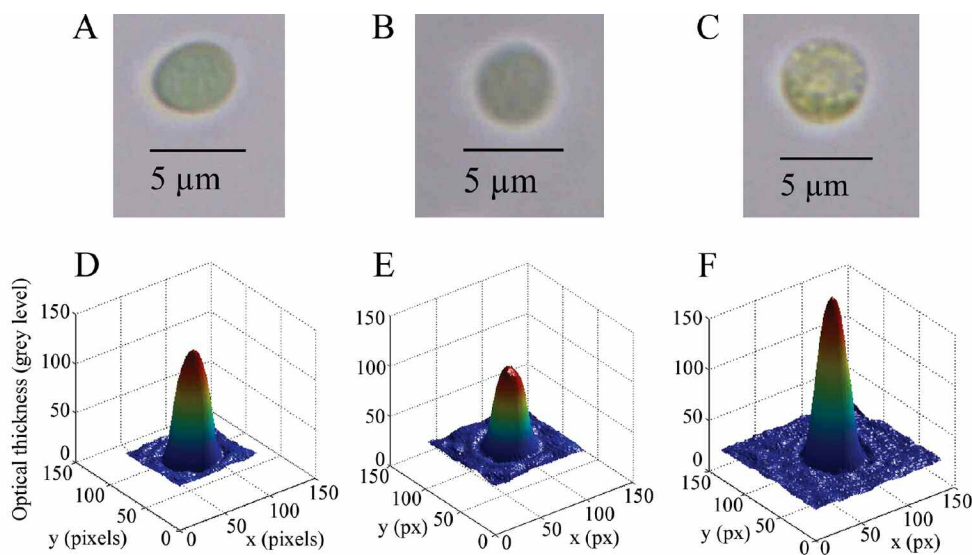
Three nanoplanktonic species were selected in this study for a classification trial. Their similarity in size, shape, and color makes them difficult to discriminate with brightfield microscopy, as can be seen in Fig. 5A-C. All three species appear as similar sized spherical green dots, which make them challenging, if not impossible, to distinguish in traditional phytoplankton composition studies based on visual observation and classical microscopy.

When flow cytometry is employed, there is a large overlap between the three species when forward scattering (FSC) and side scattering (SSC) are used to assess cell morphology (Fig. 6A). A better distinction between species is obtained

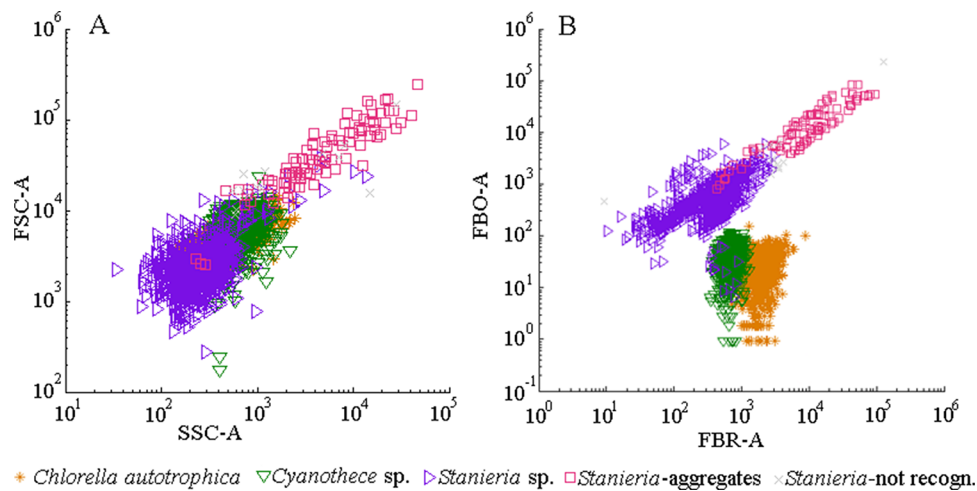
when fluorescence in the orange (FBO), representative for phycoerythrins, is plotted against fluorescence in the red (FBR), representative for chlorophyll *a* (Fig. 6B). Yet even here, the clusters consisting of *Chlorella* and *Cyanothece* are tangent and show overlap. Exactly in this region of overlap, *Stanieria* cells also appear as a small subcluster, in addition to a separate, but broadly dispersed *Stanieria* cluster at higher FBO values. The cause of this broad distribution of *Stanieria* cells is likely explained by the presence of solitary cells alongside cells that aggregate in groups or colonies.

Digital holographic microscopy (DHM) provides additional information about phytoplankton cells compared with classical microscopy and flow cytometry. In addition to light intensity information, DHM provides quantitative phase information, which provides the spatial distribution of the optical height across cells. When comparing the phase images for the three nanoplanktonic cell types, DHM reveals clear differences in the optical thickness distribution (Fig. 5D-F). These phase images reveal that the mean optical thickness of *Chlorella* is twice as high as *Stanieria*, whereas the optical thickness of *Cyanothece* is intermediate. These large differences in optical thickness suggest that the phase images recorded by DHM could provide complementary information to better discriminate these planktonic species.

This idea was confirmed by the statistical analysis of the 18 features quantified for each of the three cell types. Fig. 7 shows the box-whisker plots for the (A) 6 morphology-based features, (B) 6 textural features based on the LI images, and (C) 6 textural features based on the PI images. The morphology features reveal some statistically significant differences between species. For example, *Cyanothece* cells have a larger major axis (Kruskal-



**Fig. 5.** Top Row: Micrographs of the three study organisms, the cyanobacteria (A) *Cyanothece* sp., and (B) *Stanieria* sp., and (C) the green algae *Chlorella autotrophica*, showing their similarity in size and shape. Images were taken with a Zeiss Axioplan light microscope. Bottom row: Respective pseudo-3D visualizations from holograms captured of the three species: *Cyanothece* sp. (D), *Stanieria* sp. (E) and *Chlorella autotrophica* (F). Here the optical thickness (Z-axis) clearly demonstrates differences in the phase information of the organisms that can be exploited for improving the classification using textural features.



**Fig. 6.** Flow cytometry results from three single-species cultures (*Cyanothece* sp., *Stanieria* sp., and *Chlorella autotrophica*). The three resulting datasets are plotted onto a single graph (~900-1000 cells per culture). (A) Forward scattering (FSC) and side scattering (SSC) signals provide an indication of cell size and cell granularity, respectively. The three species are clearly similar in size and morphology and thus hard to distinguish if their fluorescence properties are not considered. (B) Better distinction is achieved with a scatter plot of emission measured in the red (FBR) indicative of chlorophyll *a*, versus the emission in the orange bandwidth (FBO) indicative of phycoerythrins and phycocyanins. Note that the data reveals unidentified clusters (*Stanieria* sp. – aggregates and – not recognized).

Wallis-Test,  $H = 216.22$ ,  $df = 2$ ,  $P < 0.001$ ), a larger area (Kruskal-Wallis-Test,  $H = 107.22$ ,  $df = 2$ ,  $P < 0.001$ ), larger perimeter (Kruskal-Wallis-Test,  $H = 120.65$ ,  $df = 2$ ,  $P < 0.001$ ), and higher equivalent circular diameter (Kruskal-Wallis-Test,  $H = 107.22$ ,  $df = 2$ ,  $P < 0.001$ ) (Fig. 7A). However,  $H$  values typically remain low (~100), and for many intensity-based and morphometric features there are no significant differences between species. For example, no statistically significant differences are revealed by post-hoc testing between *Chlorella* and *Cyanothece* cells for their eccentricity (Dunn's Method,  $Q = 0.178$ ,  $P > 0.05$ ), and for their average contrast calculated from LI images (Dunn's Method,  $Q = 2.251$ ,  $P > 0.05$ ). Uniformity calculated from LI images is also not substantially different between *Cyanothece* and *Stanieria* cells (Dunn's Method,  $Q = 2.278$ ,  $P > 0.05$ ).

One can clearly observe a distinct differentiation between the three species for textural features based on the phase information (PI), where differences are much more pronounced than for morphology-based features or features based on light intensity (LI) (Fig. 7). Except for smoothness, the other five textural features based on PI all have  $H$  values greater than 330, more than three times larger than the  $H$  values for corresponding features in the LI images. Hence the discrepancy between species is much more distinct for the PI-based features as opposed to the LI-based features or morphology-based features.

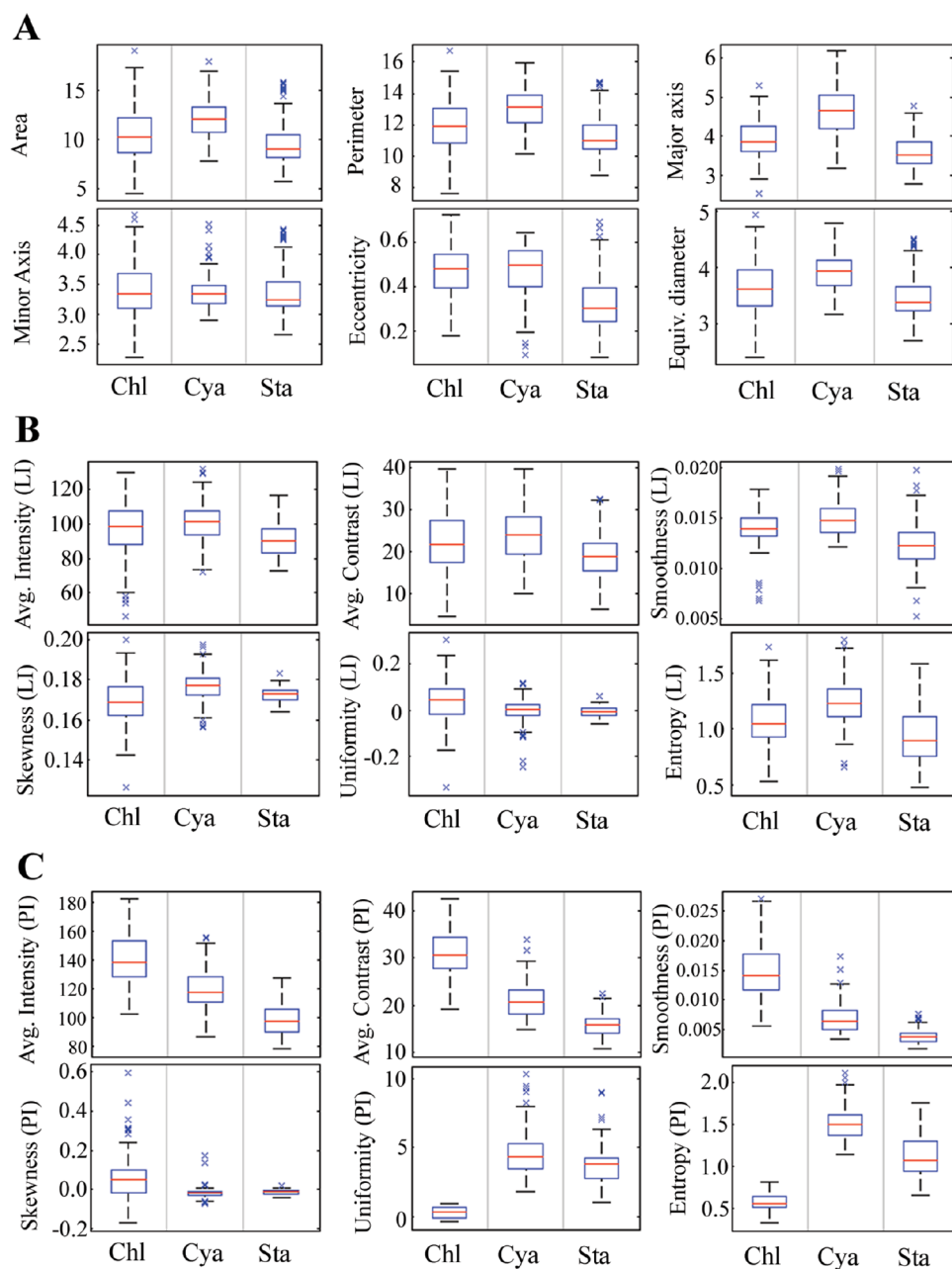
To gain a better understanding of the features and how the species are similar or dissimilar, the overlap between the interquartile ranges (IQR: the interval between the 25th and 75th percentiles) was calculated for each species and each feature combination. Values were averaged to obtain a single IQR overlap value for each feature (Table 2). The IQR represents 50% of the data for each species and therefore no overlap means that < 25% of the data may be overlapping. The %

overlap between the IQR between species revealed an overall average overlap of 48% and 61% for the morphometric features and LI-image-based textural features, respectively. This contrasts markedly with the overlap of only 12% for PI-image-based textural features. The 12% are also the result of only two features (skewness and uniformity) whereas the IQR had 0% overlap between species for the other four PI-image-based textural features (average intensity and contrast, smoothness, and entropy) (Table 2).

The observation that the phase information has a distinct signature for the three species is further exemplified by a representation in feature space as shown in Fig. 8. There is a clear similarity in the cell size and shape between the three species (Fig. 8A), but also textural features calculated from the LI images overlap each other substantially (Fig. 8B). This overlap is markedly reduced when textural features are calculated from the PI images (Fig. 8C). However, the best differentiation of species is obtained when all 18 features of the classification procedure are included, thus combining both information from LI as well as PI images and combining both morphometric as well as textural features (Fig. 8D).

The 18 features derived for each cell object were subsequently used in the automated SVM classification procedure. In automated classification algorithms, some level of error is unavoidable. The level of misclassification should therefore be assessed. The confusion matrix provides such an assessment and quantifies the performance of the classification procedure (Table 3).

The SVM classification procedure is highly accurate and provides an excellent identification level of 92.4% (i.e., percentage of correct predictions) over the three species. Almost all of the *Chlorella* cells were correctly identified (97.8%),



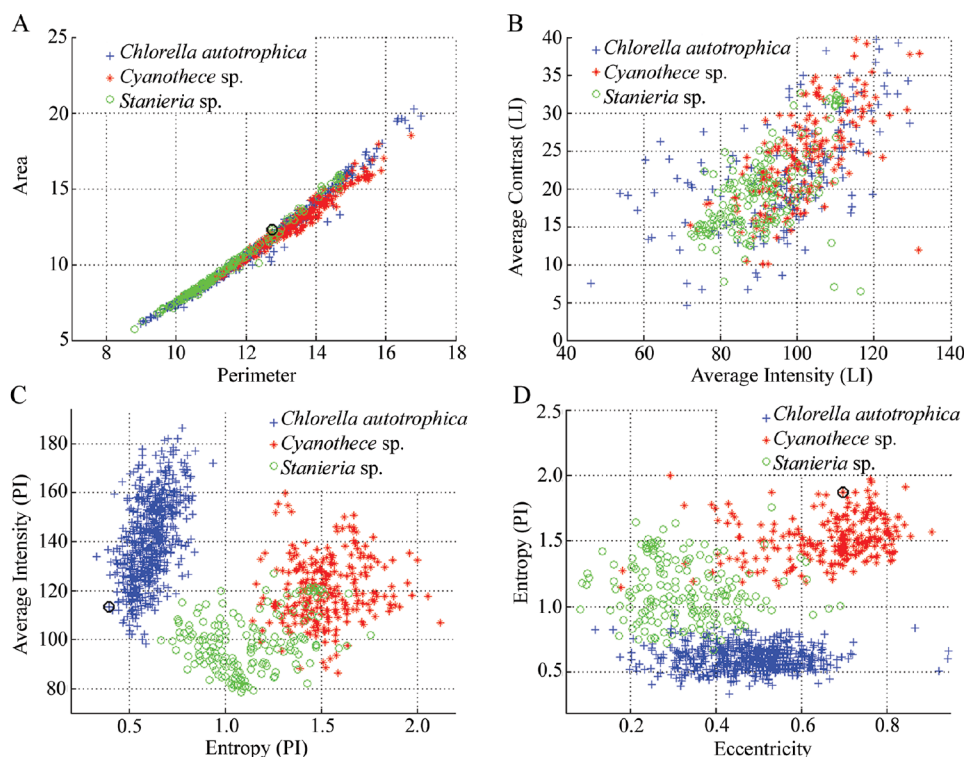
**Fig. 7.** Box-whisker plots for the extracted features for each of the three species *Cyanobacteria* sp. (Cya), *Stanieria* sp. (Sta), and *Chlorella autotrophica* (Chl): (A) morphology-based features, (B) textural features based on the light intensity (LI) information and (C) textural features based on the phase information (PI). Equiv. diameter = equivalent circular diameter, avg. = average. Box-whiskers represent the following: means are shown, with the lower and upper box limits showing the 25th and 75th percentile, respectively. Whiskers show the range, whereas crosses represent outliers.

whereas the classification of *Cyanobacteria* was slightly less accurate at 90.8% (Table 3). *Stanieria* was classified the least well (88.6%); all misclassified *Stanieria* were wrongly labeled as *Cyanobacteria* cells. This can be explained by the greater similarity in optical height between *Cyanobacteria* and *Stanieria* compared to *Chlorella* (Fig. 5D-F). This is clearly seen in Fig. 8C, where *Chlorella* cells distinctly separate from *Cyanobacteria* and *Stanieria*, whereas the latter two show a small degree of overlap.

The textural features obtained from the phase information (PI) particularly contribute to the classification success. When the classification is based on the morphology and LI-based textural features only (12 in total) a classification success rate of only 58.9% is achieved. The confusion matrix for this classification shows *Chlorella* to be identified correctly at 41.1%, *Cyanobacteria* at 74.4%, and *Stanieria* at 61.1%. The PI-based textural features are thus significant contributors to the high classification rate. Features such as average PI intensity and PI

**Table 2.** Average interquartile range (IQR) overlap (in %) calculated from the various species combinations per feature for *Chlorella*, *Cyanothece*, and *Stanieria* species.

| Morphometric features        | % overlap | Textural features | LI-image based<br>% overlap | PI-image based<br>% overlap |
|------------------------------|-----------|-------------------|-----------------------------|-----------------------------|
| Area                         | 39        | Average intensity | 65                          | 0                           |
| Perimeter                    | 36        | Average contrast  | 59                          | 0                           |
| Major Axis                   | 18        | Smoothness        | 31                          | 0                           |
| Minor Axis                   | 118       | Skewness          | 68                          | 45                          |
| Eccentricity                 | 38        | Uniformity        | 80                          | 16                          |
| Equivalent Circular Diameter | 39        | Entropy           | 35                          | 0                           |

**Fig. 8.** Feature space representation of features extracted from both the light intensity (LI) as well as phase information (PI) images. Examples are shown for (A) two morphometric features extracted from the LI, and (B) textural features extracted from the LI images. The extensive overlap between the three species (*Cyanothece* sp., *Stanieria* sp., and *Chlorella autotrophica*) is reduced with the use of extracted textural features from the PI images (C), but maximised when all 18 features are considered for the classification procedure, as exemplified in (D) showing entropy (a PI textural feature) versus eccentricity (a LI morphometric feature).

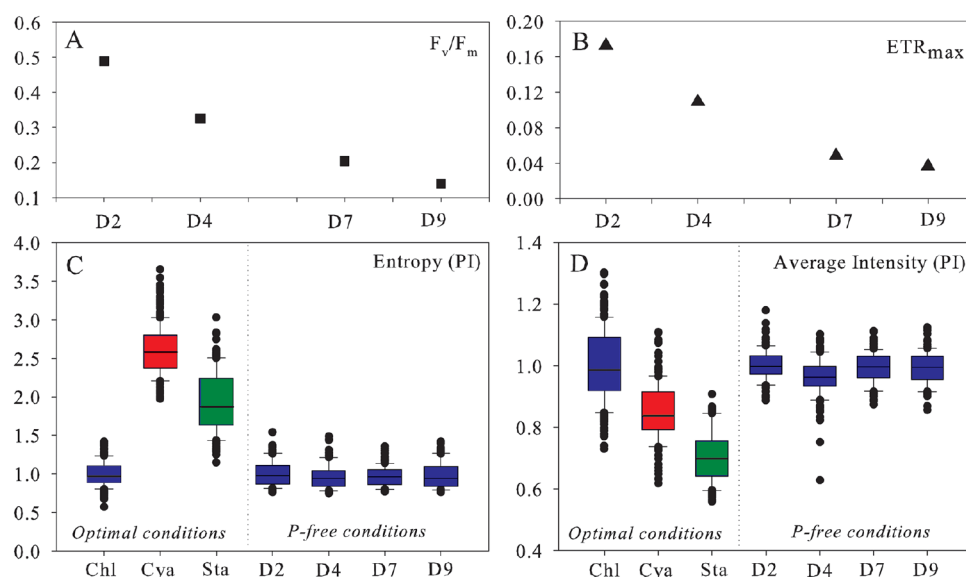
entropy clearly separate the species (Fig. 7C, 8C).

As noted earlier, the cell cultures used in the above supervised classification were grown at optimal growth conditions. However, growth conditions in natural samples cannot be controlled, and may strongly deviate from optimal growth conditions. If discriminant features used for classification (such as the textural phase information features) are highly sensitive to growth conditions and cells are plastic in their acclimation to changing growth conditions, classification performance could potentially be degraded. To examine the effect of growth conditions on features, a *Chlorella* culture was first

grown at optimal conditions, and subsequently exposed to phosphate starvation. Over a period of 9 days in P-free medium, the cell physiology of *Chlorella* became significantly affected, as seen in the response of multiple photosynthetic parameters (Fig. 9A,B). There is a strong reduction in the maximum quantum efficiency of PSII as indicated by the decrease in the ratio of variable to maximum fluorescence ( $F_v/F_m$ ) (Fig. 9A). Furthermore, the maximum rate of photosynthetic electron transport ETR declines during phosphate-limiting conditions (Fig. 9B) and is positively correlated to reductions in the initial slope of the RLC  $\alpha$  ( $r = 0.997$ ,  $n = 4$ ,  $P = 0.003$ , Pearson

**Table 3.** Confusion matrix of the mean performances of the classification procedure (in %  $\pm$  SD) for the three study organisms as computed by a 10-fold cross validation method (cross-validation by rotation). This type of contingency table easily identifies where predicted classes (columns) are not matching their actual classes (rows).

| True labels                   | Predicted classes |                   |                   | Totals (%) |
|-------------------------------|-------------------|-------------------|-------------------|------------|
|                               | Chlorella         | Cyanothece        | Stanieria         |            |
| <i>Chlorella autotrophica</i> | 97.8 ( $\pm$ 2.6) | 0.8 ( $\pm$ 0.2)  | 1.4 ( $\pm$ 0.9)  | 100        |
| <i>Cyanothece</i> sp.         | 0.0 ( $\pm$ 0)    | 90.8 ( $\pm$ 2.6) | 9.2 ( $\pm$ 0.7)  | 100        |
| <i>Stanieria</i> sp.          | 0 ( $\pm$ 0)      | 11.4 ( $\pm$ 0.9) | 88.6 ( $\pm$ 2.3) | 100        |



**Fig. 9.** (A) Changes in maximum quantum efficiency of photosystem II ( $F_v/F_m$ ) and (B) electron transport rates ( $ETR_{max}$ ,  $\mu\text{mol electrons [mg Chl } a]^{-1}\text{s}^{-1}$ ) over the duration of the 9-d period (samples taken on day (D) 2, 4, 7, and 9) clearly indicate the effects on cell physiology over time as *Chlorella autotrophica* cells become starved of phosphate and by the end of the experiment cease to remain photosynthetically active. (C) Average intensity (PI) and (D) entropy (PI) values for the three different species *C. autotrophica* (Chl), *Cyanothece* sp. (Cya), and *Stanieria* sp. (Sta) (as given in Fig. 7C) normalized to *C. autotrophica* grown under optimal conditions, compared with values extracted for *C. autotrophica* grown under P-free conditions on D2, 4, 7, and 9 after inoculation and normalized to D2.

Product Moment Correlation). Finally,  $E_k$  drops after a week in the phosphate-limited conditions to only 36% (not shown). By the end of the experiment (i.e., after 9 days) cells more or less ceased functioning.

Nutrient depletion, however, does not strongly affect the textural features based on phase information. Fig. 9 compares the relative values of two textural PI features, the ‘average intensity’ and ‘entropy,’ before and after phosphate-depletion (Fig. 9C and 9D, respectively). These two features were previously found to be strongly discriminating between species (Fig. 7C, 8C). These two textural features remain constant over time during phosphate-deplete conditions and any variations induced by nutrient depletion within a single species are small compared with the large differences found for these two features between the three phytoplankton species (Fig. 9C,D). This suggests that although physiological properties will change over time for cells grown under nutrient-deplete conditions, textural features based on phase information remain largely insensitive

to growth conditions, thus providing a robust basis for classification under various natural conditions.

In addition to correct identification, imaging-in-flow techniques should have a sufficient recovery (a recovery of 100% implies that all the cells in the sample are detected and imaged). DHM with partial coherence allows an accurate localization of the cells (XYZ-positioning) in the flow channels due to the digital refocusing ability of the technique post-acquisition (Dubois et al. 2006a). Our DHM analysis showed a high recovery; for example, a detection rate of  $80.4 \pm 5.2\%$  was achieved for *Chlorella*. A lower detection rate was achieved for *Cyanothece* and *Stanieria* sp., with  $75.2 \pm 4.8\%$  and  $62.4 \pm 6.1\%$ , respectively. These two species tend to have more variation in their optical phase information, which then introduces uncertainty in the calculation of the reconstruction distance. The contrast between the object phase and the background is more often insufficient to achieve a good detection rate for these two species compared to *Chlorella* (El Mallahi 2013). However, this



could be improved in the future with the use of acoustic or hydrodynamic focusing techniques, or alternatively, by using fluidic devices with a smaller channel depth, which in both cases would reduce the refocusing distance that is needed.

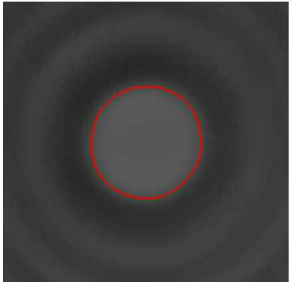
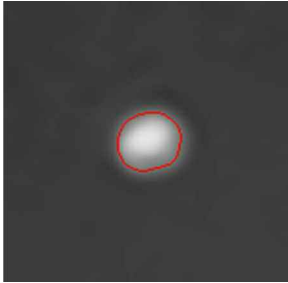
The extraction of feature information from the image of a given object is critically dependent on the object being in focus. Table 4 illustrates how features (both morphometric as well as textural) vary between out-of-focus cells compared to in-focus cells. A *Chlorella* cell was (deliberately) captured out of focus at  $-22\ \mu\text{m}$  relative to the focus plane and was subsequently digitally refocused in the focus plane. Morphological features substantially differ for the out-of-focus object versus the in-focus object: the object area differs 3-fold, whereas a 2-fold difference is noted in other shape features (perimeter, major and minor axis, eccentricity, and equivalent circular diameter). For textural features, the differences are even higher, amounting to orders of magnitude (Table 4). This sensitivity experiment stresses the importance of using in-focus cells for feature extraction and classification, and illustrates the advantage of DHM, which has the ability to digitally refocus objects post-hoc that were recorded out of focus.

## Discussion

### Holographic configuration

Holography is not entirely new to plankton research (Katz and Sheng 2010). Previous systems developed for plankton enumeration and classification were in situ systems with an in-line rather than an off-axis optical configuration (Pfirsch et al. 2005; Graham and Nimmo-Smith 2010; Bochdansky et al. 2013). An in-line configuration provides a large depth of field at high resolution with the ability to reconstruct the position of objects in three-dimensional space (Bochdansky et al. 2013). However, in-line holography does not allow to extract accurate quantitative phase information of small objects. Off-axis holography, as used here, does provide such phase information, which—as we show here—strongly improves the classification of nanoplanktonic organisms. The introduction of partial coherence (Dubois et al. 1999, 2004) has been an important development in off-axis DHM, as it substantially improves the quality of the resulting holograms. This then allows the digital refocusing of small particles with high accuracy, which we here demonstrate for nanoplanktonic organisms.

**Table 4.** Segmentation of the cell is carried out on the compensated phase image (the cells are phase objects making segmentation on light intensity images not possible). In segmentation, a label is assigned to every pixel in an image; pixels that share certain visual characteristics share the same label, creating a detectable object. Morphometric features are computed using the extracted perimeter (in red), whereas textural features for both the compensated phase image as well as intensity image are computed on the given information inside the perimeter of the segmented cell.

| Features based on compensated phase images     | Out-of-focus object<br>(at recorded plane)   | In-focus object<br>(refocused at $-22\ \mu\text{m}$<br>from the recorded plane)       |
|--|--|---|
|  |  |  |
| Morphological features                         |  |   |
| Area ( $\mu\text{m} \times \mu\text{m}$ )      | 39.79  | 11.73   |
| Perimeter ( $\mu\text{m}$ )                    | 23.25  | 12.5  |
| Major axis ( $\mu\text{m}$ )                   | 7.22   | 4.10  |
| Minor axis ( $\mu\text{m}$ )                   | 7.03   | 3.65  |
| Eccentricity                                   | 0.23   | 0.46  |
| Equivalent circular diameter ( $\mu\text{m}$ ) | 7.118  | 3.865   |
| Textural features                              |  |   |
| Average intensity                              | 85.98  | 164.67  |
| Average contrast                               | 4.65   | 36.3  |
| Smoothness                                     | 0.000332   | 0.019   |
| Skewness                                       | -0.002   | 0.215   |
| Uniformity                                     | 0.001  | 0.00000841  |
| Randomness                                     | 0.69   | 0.347   |

### Classification performance

Traditionally, phytoplankton classification has been strongly dependent on expertise of dedicated taxonomic experts performing tedious microscopy work. Yet taxonomic experts are not faultless, and it has been estimated that the average classification performance of human experts (i.e., the percentage of cells identified correctly) is about ~ 70% (Culverhouse et al. (2003). Human 'errors' affecting classification performance include fatigue, boredom, recency effects, bias from prior expectations, and short-term memory (Culverhouse et al. 2006; Culverhouse 2007). Automation tries to counter these problems, and thus, the impulse to develop improved automated classification systems continues (Macleod et al. 2010). The state of the art suggests classification accuracies between 70% to 80% for 10-20 class problems (Benfield et al. 2007; Sieracki et al. 2010). However, it is difficult to compare classification performance, as studies differ in the number of classes that are distinguished as well as in the sizes and types of organisms that are investigated (Schulze et al. 2013). Similar high classification accuracies as found in this study (>90% for three species) have been obtained for six plankton taxa analyzed from the Video Plankton Recorder (Tang et al. 1998) and six cyanobacteria species imaged with a high-resolution microscope (Walker and Kumagai 2000).

Improved automated taxonomic categorisation, as seen in the last decade, has been the result of two different lines of development: (1) novel hardware instrumentation, like 3D imaging, allows that more information is gathered on plankton cells, thus enlarging the portfolio of features that can be used in automated classification (Embleton et al. 2003; Culverhouse et al. 2006; Boistel et al. 2011), and (2) new developments in software have led to improved image processing and classification algorithms (Benfield et al. 2007; Schulze et al. 2013).

Here we have demonstrated that the additional phase information gathered with off-axis DHM can strongly increase the taxonomic resolution and automated classification performance, thus supporting the recent results of El Mallahi et al. (2013). Using a relatively limited number of features (18 in total), a very good classification performance (>90% overall) was achieved for highly similar and hence "difficult to classify" nanoplanktonic organisms. Traditionally, classification algorithms have been based on morphometric parameters. However, nanoplanktonic cells are generally spherical or ellipsoid in shape, lack distinct morphological features, and therefore these species are not well distinguished based solely on shape-based features (Malkassian et al. 2011). This hence calls for the use of texture-based features in classification rather than shape-based features. In recent years, this has been an active field of research, and a range of computational methods and techniques have been proposed to derive textural features for use in cell classification (Fourier descriptors, invariant moments, granulometric curves, etc.) (Hu and Davis 2006; Sosik and Olson 2007). As computing capabilities have increased, so have the numbers of features that are used in

plankton classification, with some studies using over 200 features (Blaschko et al. 2005; Hu and Davis 2006). The inclusion of non-morphometric feature types such as texture and invariant moments, can lead to strongly improved classification performance (e.g., 88% overall classification accuracy for 22 phytoplankton categories in Sosik and Olson [2007]). Our study proposes a new type of textural features based on holographic phase information, which adds to the portfolio of features that can be used in automated classification.

Until now, imaging-in-flow has been based on brightfield microscopy, and this allows textural features to be derived from light intensity images alone. As we show here, the interferometric information recorded in the holograms can strongly increase the classification capabilities, as off-axis DHM provides access to a new type of textural features, derived from the phase information image. Our results show that textural features based on phase information have much more discriminative power for classifying nanoplanktonic organisms than textural features based on light intensity (Fig. 7). As shown for a diatom (Fig. 2), the phase information image provides a distinctive fingerprint of the cell morphology, and consequently, the phase information appears to reveal better the variability in internal structure between different cell types. A light intensity image (produced by a classical microscope or a DHM) reflects the degree to which light is absorbed by the various cell organelles. In contrast, the phase intensity image produced by a DHM is based on a very different physical process, as it quantifies the degree to which the light phase is modulated by the internal cell structures, thus revealing subtle differences in optical thickness between various parts of a cell. DHM technology is able to quantify these differences in optical thickness at very high resolution (detecting changes ~ 2.5 nm), which hence leads to a novel set of features providing stronger discriminative power for classifying nanoplankton.

A preliminary experiment, where *Chlorella* cells were subjected to growth under nutrient-depleted conditions, suggests that textural features based on phase information are not strongly dependent on the growth conditions, despite clear changes in cell physiology. These textural features vary much more between species than within a single species growing at different conditions (Fig. 9). Although more studies should be undertaken, which should examine the response of different algal species subject to a range of growth conditions (e.g., light stress), our results suggest that the optical thickness distribution of a cell is not strongly dependent on the cells' physiological state, and hence, PI-based textural features may provide a robust basis for nanoplankton classification under a range of growth conditions.

### Improved cell detection

Another important capability of DHM is the capability to refocus a particle post acquisition, i.e., to find its optimal focus after recording the image, and thereby locating its 3D position accurately in the flow channel also reduces errors associated

with another set of problems typical for automated sample analysis, namely, the problems of position variance and focus. These problems, addressed for example by Walker and Kumagai (2000) and Schaap et al. (2012), typically force researchers to simply disregard out-of-focus cells, as any defocus adversely influences calculations of statistical properties for each cell. In addition, as cells or particles move along the flow channel, they rotate and present themselves differently to the detection sensor. In contrast to texture-based features, shape-based feature measurements can suffer greatly from projection variance (differences in focus and changes in orientation). Depending on the position in the field of view, the same cell may have a different segmented shape, reducing the accuracy of the classification algorithms (Hu and Davis 2005; Jakobsen and Carstensen 2011). El Mallahi and co-authors (2013) also showed that, depending on the cell's orientation, the calculated morphometrics of a cell (*Giardia lamblia* cyst) showed a wide range of values as it rotated along the flow trajectory.

Projection variance is one of the important problems that exists in the field of pattern recognition today (Tang et al. 1998). The refocusing capability of digital holography removes problems associated with the position in the field of view (Table 4), by being able to refocus all particles and applying further image processing procedures only on in-focus cells, which in our case, is all cells that are encountered. Occlusion on the other hand, remains an important problem for pattern recognition systems, including DHM. Occlusion may occur where one cell is blocked by another cell, cells may be touching or nonlinear illumination leads to parts of an organism being obscured (Hu and Davis 2005). However, some progress has been made on extrapolating properties for overlapped cells in order to separate them (El Mallahi and Dubois 2013) and may therefore be less of a problem for the simple morphology of nanoplanktonic organisms.

### Future outlook and recommendations

Digital holographic microscopy has already been successfully implemented in cell biology studies (Dubois et al. 2006c; Kemper and Von Bally 2008; Pavillon et al. 2010), and here we have assessed its capability for phytoplankton classification. As the need for intensified environmental monitoring is increasing (Schaap et al. 2012), there is a clear potential for DHM in aquatic applications, as an additional tool in the imaging-in-flow toolbox for plankton studies. Our results suggest that one promising application could be the automated monitoring of cyanobacteria in the context of harmful algal blooms. Novel EU regulations were adopted in 2006 (EU Directive 2006/7/EC) that require that the ecological status and water quality of lakes is monitored employing cyanobacterial and algal markers (Schaap et al. 2012). As harmful algal blooms become more prominent, there is a continued need to better understand bloom dynamics, rapidly identify the cause of the bloom, and monitor the bloom progress (Anderson et al. 2012). Online monitoring of cyanobacterial species is also

important for drinking water applications (El Mallahi et al. 2013). In these applications, one has to deal with small nanoplanktonic organisms that so far have not drawn much attention in the field of automation, and where manual sample analysis by trained experts still prevails (Walker and Kumagai 2000). Studies on the automation of cell counting and identification of cyanobacteria are presently limited (e.g., Walker and Kumagai 2000; Thiel and Wiltshire 1995; Mansoor et al. 2011). Moreover, these studies have focused on the differentiation of cyanobacteria types that are very distinct in shape, such as *Microcystis*, *Oscillatoria*, *Chroococcus*, and *Anabaena*. This makes classification relatively easy and results in classification performances of more than 90% accuracy. Most cyanobacteria, however, are cell-like units of frequently 5–10  $\mu\text{m}$  diameter and more or less spherical shape (Kremer 2006), and species such as *Cyanothece* and *Stanieria* are certainly very similar in appearance (Fig. 5). The genera *Cyanothece* also contains toxic species (Schaap et al. 2012), highlighting the need for this size range to be investigated more thoroughly in the coming years, especially in freshwater ecosystems. Our study shows that the classification of cyanobacteria can greatly benefit from the use of DHM with (1) improved object classification and (2) improved object localization as most important advantages.

DHM has the potential to improve the reliability, processing speed and taxonomic resolution of automated phytoplankton classification, compared with existing imaging-in-flow techniques based on light microscopy. Our approach using off-axis DHM with partial coherence is able to successfully detect nanoplanktonic cells in a flow with high accuracy due to its interferometric capability of refocusing post-acquisition. This reduces errors such as position and focus variance associated with some of the current automated systems. Standard flow cells were used in this experiment, but the technique may well benefit from using an approach of hydrodynamic focusing, similar to what is currently used in flow cytometers and systems such as CytoBuoy, or acoustic focusing (e.g., Leckey and Hinders 2012). When such focusing techniques are implemented, DHM could potentially provide superior measurements of size and other shape-based features. We were further able to demonstrate that the full interferometric information obtained with DHM was able to improve the classification performance for three similar looking organisms. Further studies are required to examine how well our results extend to other planktonic organisms, both in laboratory conditions as well as in natural samples, and using supervised as well as unsupervised classification routines. The true test of combining DHM with imaging-in-flow will be its performance during the analysis of natural plankton samples. If DHM is equally performant in the analysis of field samples as in the controlled experiments performed here, it provides a promising new tool for the detection, enumeration, and classification of natural phytoplankton populations, with the prospect of higher recovery and less classification errors.



## Acknowledgments

We thank Michele Grego (Culture Collection Yerseke, Royal Netherlands Institute for Sea Research) for his algal culturing work. We also thank Roland Langelaar (BD - Becton, Dickinson and Company) for running samples on the flow cytometer, and Thomas Rutten (Thomas Rutten Projects) for his analysis of the results. This study was supported by The Brussels Institute for Research and Innovation (INNOVIRIS) in the framework of the HoloFlow Impulse-Environment Project. F.J.R.M. was further supported by an Odysseus grant (FWO, Belgium), VIDI grant 864.08.004 (NWO, The Netherlands), and ERC Starting Grant 2012-306933.

## References

- Anderson, D. M., A. D. Cembella, and G. M. Hallegraeff. 2012. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annu. Rev. Mar. Sci.* 4:143-176 [doi:10.1146/annurev-marine-120308-081121].
- Bandyopadhyay, A., and others. 2011. Novel metabolic attributes of the genus cyanothecae, comprising a group of unicellular nitrogen-fixing cyanobacteria. *mBio* 2:1-10.
- Benfield, M. C., and others. 2007. RAPID—Research on automated plankton identification. *Oceanography* 20:172-187 [doi:10.5670/oceanog.2007.63].
- Blaschko, M. B., and others. 2005. Automatic in situ identification of plankton., p. 79-86. *In* Proceedings of the Seventh IEEE Workshops on Application of Computer Vision (WACV/MOTION'05). IEEE Computer Society Washington, DC, USA ©2005.
- Bochdansky, A. B., M. H. Jericho, and G. J. Herndl. 2013. Development and deployment of a point-source digital inline holographic microscope for the study of plankton and particles to a depth of 6000 m. *Limnol. Oceanogr. Methods* 11:28-40 [doi:10.4319/lom.2013.11.28].
- Boistel, R., J. Swoger, U. Kržič, V. Fernandez, B. Gillet, and E. G. Reynaud. 2011. The future of three-dimensional microscopic imaging in marine biology. *Mar. Ecol.* 32:438-452 [doi:10.1111/j.1439-0485.2011.00442.x].
- Chang, C.-C., and C.-J. Lin. 2011. LIBSVM: a library for support vector machines. *ACM TIST* 2:27.
- Collier, J. L. 2000. Flow cytometry and the single cell in phyecology. *J. Phycol.* 36:628-644 [doi:10.1046/j.1529-8817.2000.99215.x].
- Cortes, C., and V. Vapnik. 1995. Support-vector networks. *Mach. Learn.* 20:273-297 [doi:10.1007/BF00994018].
- Culverhouse, P. F. 2007. Human and machine factors in algae monitoring performance. *Ecol. Inform.* 2:361-366 [doi:10.1016/j.ecoinf.2007.07.001].
- , R. Williams, B. Reguera, V. Herry, and S. González-Gil. 2003. Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Mar. Ecol. Prog. Ser.* 247:17-25 [doi:10.3354/meps247017].
- , and others. 2006. Automatic image analysis of plankton: future perspectives. *Mar. Ecol. Prog. Ser.* 312:297-309 [doi:10.3354/meps312297].
- Davis, C. S., Q. Hu, S. M. Gallager, X. Tang, and C. J. Ashjian. 2004. Real-time observation of taxa-specific plankton distributions: an optical sampling method. *Mar. Ecol. Prog. Ser.* 284:77-96 [doi:10.3354/meps284077].
- De Philippis, R., C. Faraloni, C. Sili, and M. Vincenzini. 2005. Populations of exopolysaccharide-producing cyanobacteria and diatoms in the mucilaginous benthic aggregates of the Tyrrhenian Sea (Tuscan Archipelago). *Sci. Total Environ.* 353:360-368 [doi:10.1016/j.scitotenv.2005.09.078].
- Descy, J., and others. 2006. Climate variability as recorded in Lake Tanganyika (CLIMLAKE): Final report. Brussels: Belgian science policy.
- Devijver, P. A., and J. Kittler. 1982. Pattern recognition: A statistical approach. Prentice/Hall International Englewood Cliffs, NJ.
- Dubelaar, G. B. J., and P. L. Gerritzen. 2000. CytoBuoy: a step forward towards using flow cytometry in operational oceanography. *Sci. Mar.* 64:255-265.
- Dubois, F., L. Joannes, and J.-C. Legros. 1999. Improved three-dimensional imaging with a digital holography microscope with a source of partial spatial coherence. *Appl. Opt.* 38:7085-7094 [doi:10.1364/AO.38.007085].
- , M.-L. Novella Requena, C. Minetti, O. Monnom, and E. Istasse. 2004. Partial spatial coherence effects in digital holographic microscopy with a laser source. *Appl. Opt.* 43:1131-1139 [doi:10.1364/AO.43.001131].
- , N. Callens, C. Yourassowsky, M. Hoyos, P. Kurowski, and O. Monnom. 2006a. Digital holographic microscopy with reduced spatial coherence for three-dimensional particle flow analysis. *Appl. Opt.* 45:864-871 [doi:10.1364/AO.45.000864].
- , C. Schockaert, N. Callens, and C. Yourassowsky. 2006b. Focus plane detection criteria in digital holography microscopy by amplitude analysis. *Opt. Express* 14:5895-5908 [doi:10.1364/OE.14.005895].
- , and others. 2006c. Digital holographic microscopy for the three-dimensional dynamic analysis of in vitro cancer cell migration. *J. Biomed. Opt.* 11:054032-054035.
- El Mallahi, A. 2013. Automated 3D object analysis by digital holographic microscopy. PhD. Université Libre de Bruxelles.
- , and F. Dubois. 2011. Dependency and precision of the refocusing criterion based on amplitude analysis in digital holographic microscopy. *Opt. Express* 19:6684-6698 [doi:10.1364/OE.19.006684].
- , A. Detavernier, C. Yourassowsky, and F. Dubois. 2012. Automated 3D detection and classification of Giardia lamblia cysts using digital holographic microscopy with partially coherent source, p. 84291D-84291. *In* Proc. SPIE 8429, Optical Modelling and Design II. SPIE—The International Society for Optical Engineering.

- , and F. Dubois. 2013. Separation of overlapped particles in digital holographic microscopy. *Opt. Express* 21:6466 [doi:10.1364/OE.21.006466].
- , C. Minetti, and F. Dubois. 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 [doi:10.1364/AO.52.000A68].
- Embleton, K., C. Gibson, and S. Heaney. 2003. Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method. *J. Plankton Res.* 25:669-681 [doi:10.1093/plankt/25.6.669].
- Gabor, D. 1948. A new microscopic principle. *Nature* 161:777-778 [doi:10.1038/161777a0].
- Gonzalez, R. C., and R. E. Woods. 2008. Digital image processing, 3rd ed. Pearson International Edition.
- Graham, G. W., and W. a. M. Nimmo-Smith. 2010. The application of holography to the analysis of size and settling velocity of suspended cohesive sediments. *Limnol. Oceanogr. Methods* 8:1-15 [doi:10.4319/lom.2010.8.1].
- Hu, Q., and C. Davis. 2005. Automatic plankton image recognition with co-occurrence matrices and support vector machine. *Mar. Ecol. Prog. Ser.* 295:21-31 [doi:10.3354/meps295021].
- , and ———. 2006. Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction. *Mar. Ecol. Prog. Ser.* 306:51-61 [doi:10.3354/meps306051].
- Jakobsen, H. H., and J. Carstensen. 2011. FlowCAM: Sizing cells and understanding the impact of size distributions on biovolume of planktonic community structure. *Aquat. Microb. Ecol.* 65:75-87 [doi:10.3354/ame01539].
- Jalba, A. C., M. H. F. Wilkinson, and J.B.T.M. Roerdink. 2004. Automatic segmentation of diatom images for classification. *Microsc. Res. Tech.* 65:72-85 [doi:10.1002/jemt.20111].
- Katz, J., and J. Sheng. 2010. Applications of holography in fluid mechanics and particle dynamics. *Annu. Rev. Fluid Mech.* 42:531-555 [doi:10.1146/annurev-fluid-121108-145508].
- Kemper, B., and G. Von Bally. 2008. Digital holographic microscopy for live cell applications and technical inspection. *Appl. Opt.* 47:A52-A61 [doi:10.1364/AO.47.000A52].
- Knox, C. 1966. Holographic microscopy as a technique for recording dynamic microscopic subjects. *Science* 153:989-990 [doi:10.1126/science.153.3739.989].
- Kremer, B. 2006. Mat-forming coccoid cyanobacteria from early Silurian marine deposits of Sudetes, Poland. *Acta Palaeontol. Pol.* 51:143.
- Kromkamp, J., and M. Limbeek. 1993. Effect of short-term variation in irradiance on light harvesting and photosynthesis of the marine diatom *Skeletonema costatum*: a laboratory study simulating vertical mixing. *J. Gen. Microbiol.* 139:2277-2284 [doi:10.1099/00221287-139-9-2277].
- , and R. M. Forster. 2003. The use of variable fluorescence measurements in aquatic ecosystems: differences between multiple and single turnover measuring protocols and suggested terminology. *Eur. J. Phycol.* 38:103-112 [doi:10.1080/0967026031000094094].
- Leckey, C. A., and M. K. Hinders. 2012. Viscous effects in the acoustic manipulation of algae for biofuel production. *J. Appl. Phycol.* 24:145-156 [doi:10.1007/s10811-011-9662-7].
- Lee, C. L., and Y. Lee. 2009. Cyanobacterial bio-indicator survey for two main rivers in Taitung Taiwan. *Environ. Ecol. Proceedings* 2:1-26.
- Macleod, N., M. Benfield, and P. Culverhouse. 2010. Time to automate identification. *Nature* 467:154-155 [doi:10.1038/467154a].
- Malkasian, A., D. Nerini, M. A. Van Dijk, M. Thyssen, C. Mante, and G. Gregori. 2011. Functional analysis and classification of phytoplankton based on data from an automated flow cytometer. *Cytom. Part A* 79A:263-275 [doi:10.1002/cyto.a.21035].
- Mansoor, H., M. Sorayya, S. Aishah, and M. a. A. Mosleh. 2011. Automatic recognition system for some cyanobacteria using image processing techniques and ANN approach, p. 73-78. *In* 2011 International Conference on Environmental and Computer Science. IPCBEE. IACSIT Press.
- Maxwell, K., and G. N. Johnson. 2000. Chlorophyll fluorescence—a practical guide. *J. Exp. Bot.* 51:659-668 [doi:10.1093/jexbot/51.345.659].
- Mosteller, F., and J. Tukey. 1968. Data analysis, including statistics. *In* G. Lindzey and E. Aronson [eds.], Revised handbook of social psychology. Addison-Wesley.
- Olson, R. J., and H. M. Sosik. 2007. A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. *Limnol. Oceanogr. Methods* 5:195-203 [doi:10.4319/lom.2007.5.195].
- Pavillon, N., and others. 2010. Cell morphology and intracellular ionic homeostasis explored with a multimodal approach combining epifluorescence and digital holographic microscopy. *J. Biophotonics* 3:432-436 [doi:10.1002/jbio.201000018].
- Pereira, G., and N. Ebecken. 2011. Combining in situ flow cytometry and artificial neural networks for aquatic systems monitoring. *Expert Syst. Appl.* 38:9626-9632 [doi:10.1016/j.eswa.2011.01.140].
- Pfitsch, D. W., E. Malkiel, Y. Ronzhes, S. R. King, J. Sheng, and J. Katz. 2005. Development of a free-drifting submersible digital holographic imaging system, p. 690-696. *In* OCEANS, 2005. Proceedings of MTS/IEEE. IEEE. Washington DC.
- Pomati, F., J. Jokela, M. Simona, M. Veronesi, and B. W. Ibelings. 2011. An automated platform for phytoplankton ecology and aquatic ecosystem monitoring. *Environ. Sci. Technol.* 45:9658-9665 [doi:10.1021/es201934n].

- Poulton, N. J., and J. L. Martin. 2010. 8 Imaging flow cytometry for quantitative phytoplankton analysis—FlowCAM, p. 47-54. *In* B. Karlson, C. Cusack, and E. Bresnan [eds.], Microscopic and molecular methods for quantitative phytoplankton analysis. Intergovernmental Oceanographic Commission of UNESCO.
- Ralph, P. J., and R. Gademann. 2005. Rapid light curves: a powerful tool to assess photosynthetic activity. *Aquat. Bot.* 82:222-237 [doi:10.1016/j.aquabot.2005.02.006].
- Ritchie, R. J. 2006. Consistent sets of spectrophotometric chlorophyll equations for acetone, methanol and ethanol solvents. *Photosynth. Res.* 89:27-41 [doi:10.1007/s11120-006-9065-9].
- Ruiz, G. M., T. K. Rawlings, F. C. Dobbs, L. A. Drake, T. Mul-lady, A. Huq, and R. R. Colwell. 2000. Global spread of microorganisms by ships. *Nature* 408:49-50 [doi:10.1038/35040695].
- Rutten, T. P. A., B. Sandee, and A. R. T. Hofman. 2005. Phyto-plankton monitoring by high performance flow cytometry: A successful approach? *Cytom. Part A* 64A:16-26 [doi:10.1002/cyto.a.20106].
- Sadeghi, R., R. Zarkami, K. Sabetraftar, and P. Van Damme. 2012. Use of support vector machines (SVMs) to predict dis-tribution of an invasive water fern *Azolla filiculoides* (Lam.) in Anzali wetland, southern Caspian Sea, Iran. *Ecol. Mod-ell.* 244:117-126 [doi:10.1016/j.ecolmodel.2012.06.029].
- Schaap, A., T. Rohrlack, and Y. Bellouard. 2012. Optical classi-fication of algae species with a glass lab-on-a-chip. *Lab Chip* 12:1527-1532 [doi:10.1039/c2lc21091f].
- Schulze, K., U. M. Tillich, T. Dandekar, and M. Frohme. 2013. PlanktoVision—an automated analysis system for the iden-tification of phytoplankton. *BMC Bioinformatics* 14:115 [doi:10.1186/1471-2105-14-115].
- Sezgin, M., and B. Sankur. 2004. Survey over image threshold-ing techniques and quantitative performance evaluation. *J. Electron. Imaging* 13:146-168 [doi:10.1117/1.1631315].
- Sieracki, C. K., M. E. Sieracki, and C. S. Yentsch. 1998. An imaging-in-flow system for automated analysis of marine microplankton. *Mar. Ecol. Prog. Ser.* 168:285-296 [doi:10.3354/meps168285].
- , and others. 2010. Optical plankton imaging and analysis systems for ocean observation, p. 21-25. *In* J. Hall, D. E. Harrison, and D. Stammer [eds.], Proceedings of OceanObs'09: Sustained Ocean Observations and Informa-tion for Society. ESA Publication WPP-306.
- Silsbe, G. M., and J. C. Kromkamp. 2012. Modeling the irradi-ance dependency of the quantum efficiency of photosyn-thesis. *Limnol. Oceanogr. Methods* 10:645-652 [doi:10.4319/lom.2012.10.645].
- Singh, J., and S. Gu. 2010. Commercialization potential of microalgae for biofuels production. *Renew. Sust. Energ. Rev.* 14:2596-2610 [doi:10.1016/j.rser.2010.06.014].
- Sosik, H. M., and R. J. Olson. 2007. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* 5:204-216 [doi:10.4319/lom.2007.5.204].
- Spizzichino, V., L. Fiorani, A. Lai, A. Palucci, K. Semyanov, E. Viaggiu, and P. Albertano. 2011. First studies of pico-and nanoplankton populations by a laser scanning flow cytometer. *J. Quant. Spectrosc. Radiat. Transfer* 112:876-882 [doi:10.1016/j.jqsrt.2010.10.008].
- Tang, X., W. K. Stewart, H. Huang, S. M. Gallager, C. S. Davis, L. Vincent, and M. Marra. 1998. Automatic plankton image recognition. *Artif. Intell. Rev.* 12:177-199 [doi:10.1023/A:1006517211724].
- Thiel, S. U. and R. J. Wiltshire. 1995. The automated detection of cyanobacteria using digital image processing techniques. *Environ. Int.* 21:233-236 [doi:10.1016/0160-4120(95)00013-5].
- Van Der Heijden, F., R. P. W. Duin, D. De Ridder, and D. Tax. 2004. Classification, parameter estimation and state esti-mation. Wiley Online Library [doi:10.1002/0470090154].
- Walker, R. F., and M. Kumagai. 2000. Image analysis as a tool for quantitative phycology: a computational approach to cyanobacterial taxa identification. *Limnology* (Tokyo) 1:107-115 [doi:10.1007/s102010070016].
- Yamaguchi, I., and T. Zhang. 1997. Phase-shifting digital holography. *Opt. Lett.* 22:1268-1270 [doi:10.1364/OL.22.001268].
- Yentsch, C. M., and others. 1983. Flow cytometry and cell sorting: A technique for analysis and sorting of aquatic par-ticles. *Limnol. Oceanogr.* 28:1275-1280 [doi:10.4319/lo.1983.28.6.1275].
- Yourassowsky, C., and F. Dubois. 2014. High throughput holo-graphic imaging-in-flow for the analysis of a wide plankton size range. *Opt. Express* 22:6661-6673 [doi:10.1364/OE.22.006661].
- Zarauz, L., X. Irigoien, and J. A. Fernandes. 2009. Changes in plankton size structure and composition, during the gener-ation of a phytoplankton bloom, in the central Cantabrian Sea. *J. Plankton Res.* 31:193-207 [doi:10.1093/plankt/fbn107].
- Zhang, T., and I. Yamaguchi. 1998. Three-dimensional microscopY with phase-shifting digital holography. *Opt. Lett.* 23:1221-1223 [doi:10.1364/OL.23.001221].

Submitted 16 July 2014

Revised 6 October 2014

Accepted 14 October 2014