

# Robust and automatic cell detection and segmentation from microscopic images of non-setae phytoplankton species

Haiyong Zheng<sup>1</sup>✉, Nan Wang<sup>1</sup>, Zhibin Yu<sup>1</sup>, Zhaorui Gu<sup>1</sup>, Bing Zheng<sup>1</sup>

<sup>1</sup>College of Information Science and Engineering, Ocean University of China, Qingdao 266100, People's Republic of China

✉ E-mail: zhenghaiyong@ouc.edu.cn

**Abstract:** Saliency-based marker-controlled watershed method was proposed to detect and segment phytoplankton cells from microscopic images of non-setae species. This method first improved IG saliency detection method by combining saturation feature with colour and luminance feature to detect cells from microscopic images uniformly and then produced effective internal and external markers by removing various specific noises in microscopic images for efficient performance of watershed segmentation automatically. The authors built the first benchmark dataset for cell detection and segmentation, including 240 microscopic images across multiple phytoplankton species with pixel-wise cell regions labelled by a taxonomist, to evaluate their method. They compared their cell detection method with seven popular saliency detection methods and their cell segmentation method with six commonly used segmentation methods. The quantitative comparison validates that their method performs better on cell detection in terms of robustness and uniformity and cell segmentation in terms of accuracy and completeness. The qualitative results show that their improved saliency detection method can detect and highlight all cells, and the following marker selection scheme can remove the corner noise caused by illumination, the small noise caused by specks, and debris, as well as deal with blurred edges.

## 1 Introduction

Phytoplankton as the major producers of the primary food supply in the sea is vital microscopic organisms. Data on the abundance, distribution and species composition of phytoplankton are significant for the study of marine ecosystems. The traditional method of phytoplankton study relies mainly on well-experienced experts using microscopes in the laboratory, which is time consuming. The development of in situ optical imaging system [1–5] leads to the extensive studies on the automatic analysis of phytoplankton microscopic images for cell identification and classification [2, 6–11].

Reliable cell detection and segmentation play a key role in automatic phytoplankton analysis [12]. Considering the characteristics of setae in microscopic images, we divided phytoplankton into two categories, namely setae species and non-setae species, for efficient cell segmentation. For phytoplankton species with setae, we proposed an image-processing model called greyscale surface direction angle model (GSDAM) and developed an algorithm based on GSDAM to segment setae from microscopic images accurately and completely [13]. Our goal was to detect and segment the phytoplankton cells from microscopic images of non-setae species automatically, which could be used for further phytoplankton analysis, such as identification and classification. The microscopic images used in this study were all captured under laboratory-controlled environment.

Threshold is a common method for image segmentation and is useful for discriminating foreground from the background. Embleton *et al.* [14] separated the regions of interest within the image from background by selecting pixels below a threshold value for automated identification and counting. Davis *et al.* [15] binarised a greyscale image by a certain brightness threshold for in-focus object detection within a video plankton recorder (VPR) system. Rodenacker *et al.* [16] performed threshold-based segmentation on bright field (BF) and fluorescence images for their developed plankton structure analysis system. Bi *et al.* [17] developed an adaptive threshold approach to segment small organisms from images acquired by a zooplankton visualisation system within turbid estuarine waters. The main issue with

threshold techniques is that they often cut an object into different parts such that the object is difficult to segment as a whole.

Edge (contour) detection is a fundamental tool used in most image-processing applications, especially object segmentation, to detect outlines and boundaries between objects and background in the images. Blaschko *et al.* [18] used intensity-based segmentation to capture fine details and snake-based segmentation to produce a smooth closed contour from the images obtained in situ by flow cytometer and microscope. Sosik and Olson [19] applied simple threshold-based edge detection to phase congruity phytoplankton images generated by Imaging FlowCytobot (submersible flow cytometer) for boundary extraction. As a commonly used segmentation method, Canny edge detector (Canny) has been also used for detecting edges and contour of cells from microscopic images [20, 21, 10]. Although edge-based methods, such as Canny, sometimes perform good for detecting edges of objects, they are difficult to be used automatically for optimal results because of the requirement of appropriate initialisation or parameter setting [22].

Phytoplankton cells usually present blurred edges due to surrounding water or inappropriate focus while imaging; thus, region-based methods [23] are valid for phytoplankton cell segmentation. Verikas *et al.* [24] applied fuzzy C-means clustering algorithm to segment the circular objects representing the *Prorocentrum minimum* species in phytoplankton microscopic images. Jalba *et al.* [12] developed a marker-controlled watershed algorithm for automatic segmentation of diatom microscopic images, which improved the initial segmentation results obtained within the automatic diatom identification and classification project [6, 25] and was adopted for further diatom analysis and classification [26].

Although watershed from markers can be useful for automatic non-setae phytoplankton image segmentation [12, 26], designing an appropriate marker selection scheme for low-contrast microscopic images to segment all cells accurately and robustly remains a challenging problem. The selection of markers, containing internal markers (associated with objects of interest) and external markers (associated with the background), is used to control over-segmentation. Markers provide a priori knowledge to bear on the segmentation problem, which may be learned from the human visual system (HVS) segmentation and higher-level tasks.

ISSN 1751-9659

Received on 13th February 2017

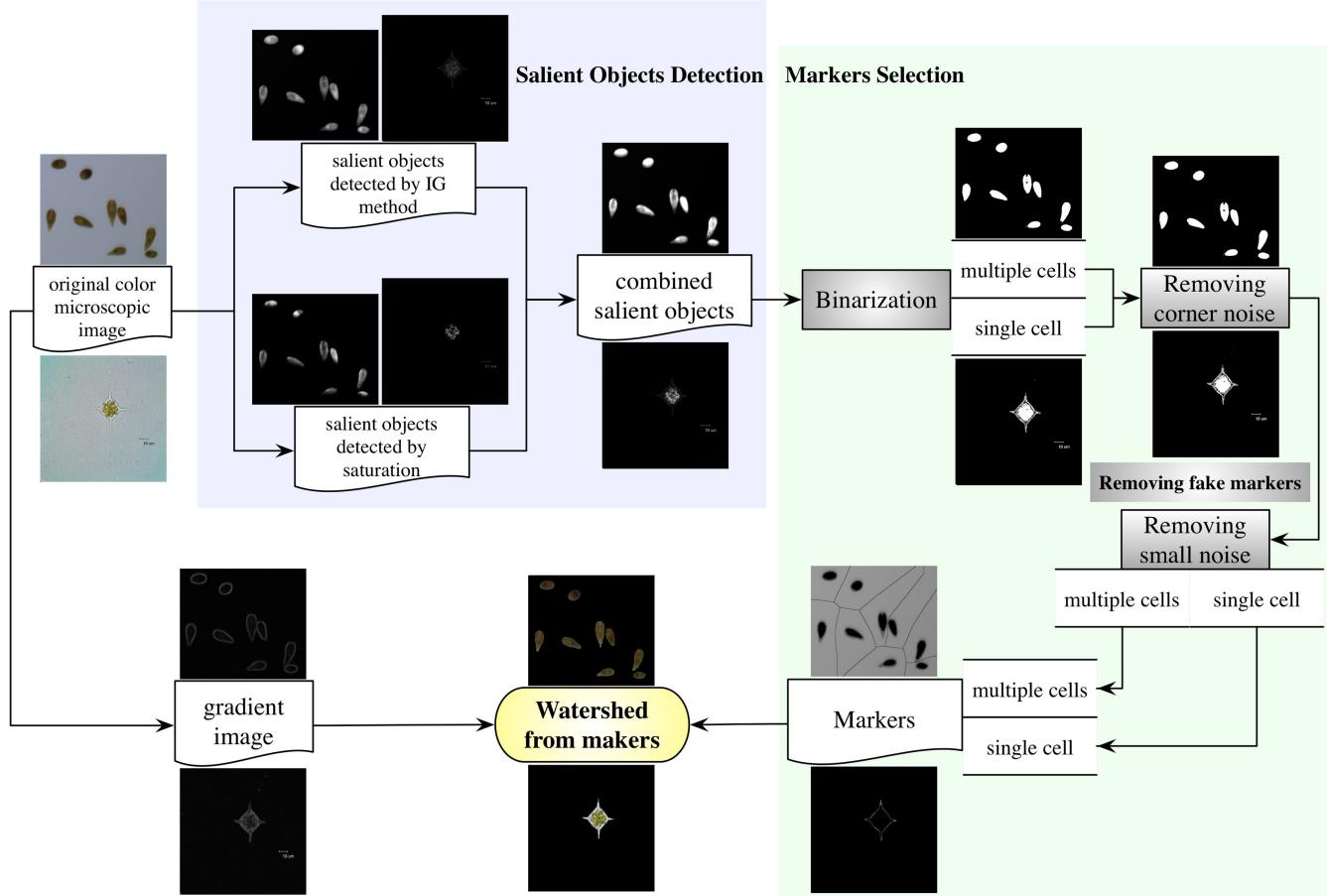
Revised 28th June 2017

Accepted on 20th August 2017

E-First on 25th September 2017

doi: 10.1049/iet-ipr.2017.0127

www.ietdl.org



**Fig. 1** Our proposed method for detecting and segmenting the phytoplankton cells from microscopic images of non-setae species

The capability to perceive objects before segmenting or identifying them is closely related to bottom-up visual attention (*saliency*), which refers to the capability to rapidly locate the most important information (also known as foreground) in a scene. Microscopic images are usually captured by involving the visual attention mechanism of the HVS to highlight the phytoplankton cells as the foreground objects. Therefore, the image saliency can be used to determine the foreground cells and can be the appropriate prior knowledge as the markers for automatic watershed segmentation.

Saliency-related research can be broadly classified into two categories: fixation prediction and salient object detection [27]. Fixation prediction aims at computing a ‘saliency map’ that simulates human eye movements to highlight edges and corners rather than the entire objects, which are unsuitable as markers for watershed; salient object detection tries to detect the most salient and attention-grabbing objects in a scene that can be used to select markers for watershed algorithm. This study presented a novel marker selection method based on salient object detection and according to the characteristics of phytoplankton microscopic images to make watershed segmentation accurate and robust against noises, such as specks and debris, uneven illumination and blurred edges. As shown in Fig. 1, we first improved the simple and computationally efficient IG (the authors used ‘IG’ to represent their salient region detection method in [28]) saliency detection method by combining the saturation feature with color and luminance feature to detect the salient objects (regions) in the microscopic images uniformly. We then removed the possible corner and small noise from the binary salient objects to avoid fake markers and thus produce effective internal and external markers for efficient performance of watershed segmentation. The proposed method was applied to automatic detection and segmentation of phytoplankton cells from non-setae species on our benchmark dataset containing 240 microscopic images. We compared our cell detection method with seven popular saliency detection methods and our cell segmentation method with six commonly used segmentation methods by evaluating precision recall (PR) curve,  $F$ -

measure and multiple-cell counting quantitatively, and the visual image results. The quantitative and qualitative experimental results validate that our method performs more robustly than the popular saliency detection methods on cell detection and more accurately than the commonly used methods on cell segmentation.

The main contribution of the study is threefold: (i) a simple and efficient salient object detection algorithm was developed, (ii) an effective noise removal method was designed on the basis of the observation of microscopic image characteristics and (iii) a novel marker selection scheme for watershed segmentation was proposed. We also provided the first benchmark dataset to study phytoplankton automatic cell detection and segmentation from microscopic images built with the knowledge of a phytoplankton taxonomist.

## 2 Automatic cell detection and segmentation

### 2.1 Image characteristics

From our observation and analysis, the characteristics of phytoplankton microscopic images can be summarised as follows. Specks and debris, such as ash or sand particles, may be present as small noise in the images. The uneven illumination on the water of samples in different depths can result in low contrast, namely bright in the centre, whereas dark in the surroundings, especially the corners, because the microscopic images are always captured under the illumination deployed on the centre of the human visual field. Cells in the microscopic images usually present blurred edges due to surrounding water or inappropriate focus. The saturation of the cells is different from that of background because the water is nearly transparent in the microscopic images.

As for the microscopic image segmentation of non-setae phytoplankton species, threshold methods may fail because of the uneven illumination and strong noise of specks and debris, and contour-based methods may not be valid due to the low contrast and blurred cell edges. Although region-based methods are flexible, pixels belonging to specks or debris cannot often be

uniquely assigned to either cell or background regions. The discriminative saturation information has also never been used in previous methods.

## 2.2 Salient object detection

The phytoplankton microscopic images are usually snapped manually, which means that the imaging procedure is processed by the visual attention mechanism of the HVS, and its computational model (also known as saliency detection) can be used to find the salient regions (objects). The salient objects in this study are those regions of an image that are visually conspicuous by contrast with respect to surrounding regions; these objects are usually cells and strong noise in the phytoplankton microscopic image.

**2.2.1 Gaussian smoothing:** The high-frequency components of an image correspond to pixel values that change rapidly across the image (e.g. edge and noise), and strong low-frequency components correspond to large-scale features in the image (e.g. a single, homogeneous object that dominates the image). Very small particles in water are high-frequency noise in the phytoplankton microscopic image that should be suppressed before saliency computation. Therefore, we first used Gaussian low-pass filter on the original microscopic image to retain the low-frequency components while attenuating the high-frequency components, that is

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2 + y^2/2\sigma^2)}, \quad (1)$$

with a cutoff value (standard deviation) of  $\sigma = 0.5$  and a window size of  $3 \times 3$  in our experiments.

**2.2.2 Improved IG method:** The IG method outputs full-resolution saliency maps with well-defined boundaries of salient objects [28]. We improved this method for the phytoplankton microscopic images by combining saturation feature to determine salient objects uniformly, given that the background is less saturated than cells in the microscopic image. The IG method uses *Lab* colour space with dimension  $L$  for luminance, which closely matches human perception of lightness and  $a$  and  $b$  for the colour-opponent dimensions because it is designed to approximate human vision. The background (water) in the phytoplankton microscopic image is nearly transparent and makes the difference in saturation between cells and background conspicuous. Thus, we combined the features of luminance and colour in *Lab* colour space with the saturation feature in hue, saturation and intensity (*HSI*) colour space, which represents colours similarly to human eye colour senses, to compute the saliency of the phytoplankton microscopic image. The IG method of finding the saliency map  $S_{Lab}$  for an image  $I$  of height  $H$  and width  $W$  can be formulated as [28]

$$S_{Lab}(x, y) = \| \mathbf{I}_{Lab} - \mathbf{I}_{Lab}(x, y) \|, \quad (2)$$

where  $\mathbf{I}_{Lab}$  is the mean image feature vector  $([L, a, b]^T)$ ,  $\mathbf{I}_{Lab}(x, y)$  is the corresponding image pixel vector value in  $I$  (the Gaussian blurred version of the original microscopic image), and  $\| \cdot \|$  is the  $L_2$  norm. The *Lab* colour space is used, wherein each pixel location is an  $[L, a, b]^T$  vector, and the  $L_2$  norm is the Euclidean distance. However, the saturation feature in *HSI* colour space cannot be regarded as an element of the above-mentioned image feature vector directly because it is not linearly independent with the three dimensions of *Lab* colour space. Thus, we computed the saliency of saturation  $S_S$  similarly according to (2) shown as

$$S_S(x, y) = \| I_S - I_S(x, y) \|, \quad (3)$$

where  $I_S$  is the mean image saturation feature value in *HSI* colour space, and  $I_S(x, y)$  is the corresponding image pixel value in the Gaussian blurred version  $I$  of the original microscopic image. We

then combined  $S_{Lab}$  and  $S_S$  linearly to generate the final saliency map  $S$  with all the detected salient objects by

$$S = S_{Lab} + S_S. \quad (4)$$

## 2.3 Marker selection

The saliency map  $S$  with all the detected salient objects, generated by combining features of luminance, colour and saturation, can be used to extract internal and external markers to avoid over-segmentation of the watershed algorithm. Effective markers for watershed can be obtained by binarising  $S$  and then removing the fake markers caused by the corner noise and small noise on the basis of the analysis of image characteristics. For the different purposes of segmenting single and multiple cells from the microscopic image, marker selection should be different for efficient segmentation.

**2.3.1 Binarisation:** The binarisation of the saliency map  $S$  aims to locate the salient objects as markers for watershed segmentation, which does not require completeness of separation in this stage. For simple computation, we used the threshold method for binarisation but selected the clustering-based Otsu's method for multiple cells and the expectation plus ten empirically for single cells because the salient level of multiple objects usually seems higher than that of single object in the saliency map, that is

$$T_{\text{multiple}} = T_{\text{otsu}}, \quad (5)$$

$$T_{\text{single}} = \frac{1}{H \times W} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} S(x, y) + 10, \quad (6)$$

where  $T_{\text{otsu}}$  represents the threshold computed by Otsu's method on grey-level saliency map  $S$ , and  $S(x, y)$  denotes the pixel value of coordinate  $(x, y)$  in saliency map  $S$  with height  $H$  and width  $W$ . The binary map  $B$  can be obtained using the computed threshold  $T$  on the saliency map  $S$  as

$$B(x, y) = \begin{cases} 1 & \text{if } S(x, y) \geq T, \text{ foreground} \\ 0 & \text{if } S(x, y) < T, \text{ background} \end{cases} \quad (7)$$

**2.3.2 Removing corner noise:** The phytoplankton microscopic images are usually captured under uneven illumination that concentrates on the centre of our visual field, thereby yielding more saturation in the surrounding and corners than in the centre of the images. The four corners of the image may consequently feature the same saturation as the objects, which can also appear as the salient objects in the binary map. Considering the priori knowledge that the cells are nearly impossible to be put in the corners while imaging, we set the four connected components related to the corresponding four vertex pixels of the binary map  $B$  as *background* to remove the possible corner noise, which may appear as fake markers in the subsequent procedure. We yielded the corner-denoised binary map  $B_{\text{dn}}$  as

$$B_{\text{dn}} = B |_{C_{(0,0)} = 0, C_{(H-1,0)} = 0, C_{(0,W-1)} = 0 \text{ and } C_{(H-1,W-1)} = 0}, \quad (8)$$

where  $C_{(0,0)}$ ,  $C_{(H-1,0)}$ ,  $C_{(0,W-1)}$  and  $C_{(H-1,W-1)}$  represent the connected components of the corresponding four vertex pixels  $(0,0)$ ,  $(H-1,0)$ ,  $(0,W-1)$  and  $(H-1,W-1)$ , respectively.

**2.3.3 Removing small noise:** The relatively small noises, such as specks and debris, can then be removed from the corner-denoised binary map  $B_{\text{dn}}$  according to the size of the binary salient objects by the following steps:

- (i) The connected component labelling on  $B_{\text{dn}}$  is used to group all pixels of the corner-denoised binary map into components based on pixel connectivity, and the holes in the components are filled to

obtain the connected regions ( $\{O_1, O_2, \dots, O_N\}$ ,  $N$  denotes the number) representing the salient objects;

(ii) The number of pixels belonging to each connected region is calculated as the size of this salient object as follows:

$$P_{O_k} = N\{(x_i, y_j) | (x_i, y_j) \in O_k, i = 0, 1, \dots, H-1; j = 0, 1, \dots, W-1\}; k = 1, 2, \dots, N \quad (9)$$

where  $O_k$  denotes the  $k$ th salient object, and  $P_{O_k}$  is the size of  $O_k$ .

(iii) The salient objects are classified into cells and noise by size. The salient object with the maximum size is selected as the cell for single cells, and the Otsu's thresholding method is adopted for multiple cells, that is

$$O_{\text{single}} = \arg \max_O P_{O_k}, k = 1, 2, \dots, N \quad (10)$$

$$O_{\text{multiple}} = \{O | P_{O_k} \geq P_{O_{\text{otsu}}}, k = 1, 2, \dots, N\} \quad (11)$$

where  $P_{O_{\text{otsu}}}$  represents the threshold computed by Otsu's method on the set of  $\{P_{O_1}, P_{O_2}, \dots, P_{O_N}\}$ .

We accordingly obtained the binary salient cell map denoted as  $B_{\text{sc}}$ .

**2.3.4 Internal and external markers:** Although the extracted salient objects representing the cells in  $B_{\text{sc}}$  are usually smaller and less accurate than the real cells in the original microscopic image due to their blurred edges, they are still suitable to be used to produce the internal markers  $M_{\text{internal}}$  and external markers  $M_{\text{external}}$  for the watershed algorithm for complete and accurate segmentation of the cells.

(i) For segmenting single cells, the binary salient cell map  $B_{\text{sc}}$  was eroded twice to produce the internal markers, and  $B_{\text{sc}}$  was dilated twice and then reversed to produce the external markers, that is

$$M_{\text{internal\_single}} = (B_{\text{sc}} \ominus E_{3 \times 3}) \ominus E_{3 \times 3} \quad (12)$$

$$M_{\text{external\_single}} = \neg((B_{\text{sc}} \oplus E_{3 \times 3}) \oplus E_{3 \times 3}) \quad (13)$$

where  $E_{3 \times 3}$  is a  $3 \times 3$  matrix of ones as the structuring element.

(ii) For segmenting multiple cells, the internal markers were produced in the same way, whereas the external markers were produced by finding the watershed ridge lines from the watershed transform of the distance transform of the binary salient cell map  $B_{\text{sc}}$ , that is

$$M_{\text{internal\_multiple}} = (B_{\text{sc}} \ominus E_{3 \times 3}) \ominus E_{3 \times 3} \quad (14)$$

$$M_{\text{external\_multiple}} = (\mathcal{W}(\mathcal{D}(B_{\text{sc}})) = 0) \quad (15)$$

where  $E_{3 \times 3}$  is a  $3 \times 3$  matrix of ones as the structuring element,  $\mathcal{D}(\cdot)$  means the Euclidean distance transform, and  $\mathcal{W}(\cdot)$  means the watershed transform.

#### 2.4 Watershed from markers

The marker-controlled watershed was applied for the final cell segmentation, that is the grey-level phytoplankton microscopic image converted from the original RGB image was used to compute the morphological gradient image  $G_D$  by subtracting the erosion of the grey-level image from the dilation of the grey-level image as follows:

$$G_D = (G \oplus E_{3 \times 3}) - (G \ominus E_{3 \times 3}) \quad (16)$$

where  $G$  is the grey-level image of the original microscopic colour image, and  $E_{3 \times 3}$  is a  $3 \times 3$  matrix of ones as the structuring

element. The extracted internal and external markers ( $M_{\text{internal}}$  and  $M_{\text{external}}$ ) were used to modify the gradient image  $G_D$  by use of watershed for the result of the final cell segmentation  $R$  as

$$R = \mathcal{W}(G_D, M_{\text{internal}}, M_{\text{external}}) \quad (17)$$

### 3 Results and discussion

#### 3.1 Image acquisition

The phytoplankton samples we used in this study were collected from China coastal waters, such as South China Sea, Yellow Sea and East China Sea. Most species were collected by the 'Dong Fang Hong 2' marine survey vessel of the Ocean University of China and cultured by the Key Laboratory of Marine Environment and Ecology, Ministry of Education in Ocean University of China. A small number of phytoplankton species were separated from the high-frequency harmful algal bloom HAB occurrence areas of South China Sea and cultured by the Algal Collection of Research Center for Harmful Algae and Aquatic Environment in Jinan University. All the samples were collected for marine survey in China coastal waters, and no specific permissions were required. We confirmed that the field studies did not involve endangered or protected species.

The microscopic images were mainly acquired using QImaging Retiga 4000R FAST 1394 charge-coupled device (CCD) digital camera equipped on the Olympus BX61 microscope illuminated by a 12 V 100 W halogen light source in BF mode and snapped and stored digitally using Image Pro Plus software with a pixel depth of 24 bits (RGB). A few microscopic images were captured using the single-panel colour CCD equipped on the Olympus FSX100 microscope illuminated by a white light-emitting diode light source in BF mode with a pixel depth of 24 bits (RGB).

#### 3.2 Dataset

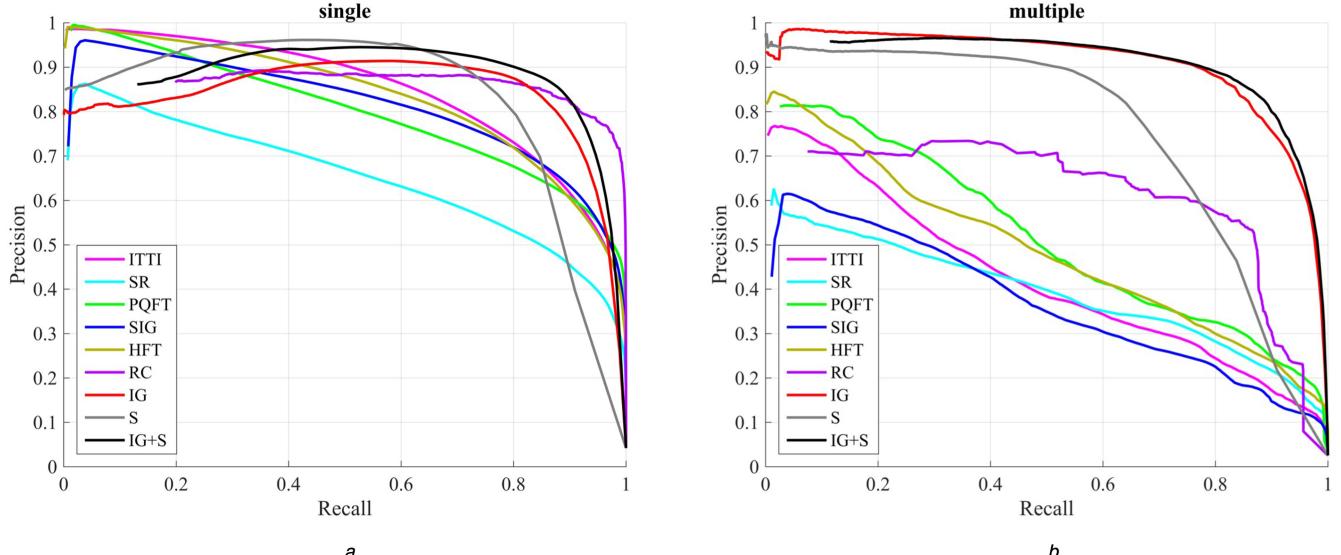
We built the first benchmark dataset for phytoplankton cell detection and segmentation consisting of 240 phytoplankton microscopic images, with 225 single cells and 15 multiple cells, and human-labelled ground-truth (GT) cell regions, to evaluate and compare the performance of our method for cell detection and segmentation quantitatively and qualitatively. These images were acquired in different sizes (256 × 256 to 4080 × 3072), and the different species of non-setae phytoplankton were selected with the help of phytoplankton experts. The pixel-wise GT masks were produced as guided by a phytoplankton taxonomist to retain the biomorphic characteristics of the cells. These data provide useful resource to study the phytoplankton for automatic cell detection and segmentation and are available at <http://vision.ouc.edu.cn/zhenghaiyong/research/acds/>.

#### 3.3 Experimental settings

In this study, the salient object detection methods and the commonly used segmentation methods were collected and implemented for comparison of cell detection and segmentation, respectively. Most methods were shared online with complete code available, and the others were provided by the authors or written by us according to the corresponding published paper; these methods are indicated below as footnotes. The comparative study was implemented by running all the codes of these methods (including ours) with MATLAB R2013a under Mac OS X Yosemite operating system on a computer with 2.6 GHz dual-core Intel Core i5 CPU and 8 GB 1600 MHz DDR3 RAM. The MATLAB implementation of our proposed method and the evaluation metrics are available at <https://github.com/zhenlab/ACDS>.

#### 3.4 Quantitative comparison

**3.4.1 Cell detection:** We compared our proposed method for salient cell detection with seven popular saliency detection methods: ITTI ('ITTI' comes from the family name of the first author from [29], and this implementation comes in the *simpson* package from <http://www.vision.caltech.edu/~harel/share/>



**Fig. 2** Quantitative comparison of cell detection by different methods on our proposed dataset

(a) 225 microscopic images with single cells. The yellow  $\star$  representing the actual threshold generated by our method for binarisation shows the nearly best segmentation on these PR curves, (b) 15 microscopic images with multiple cells

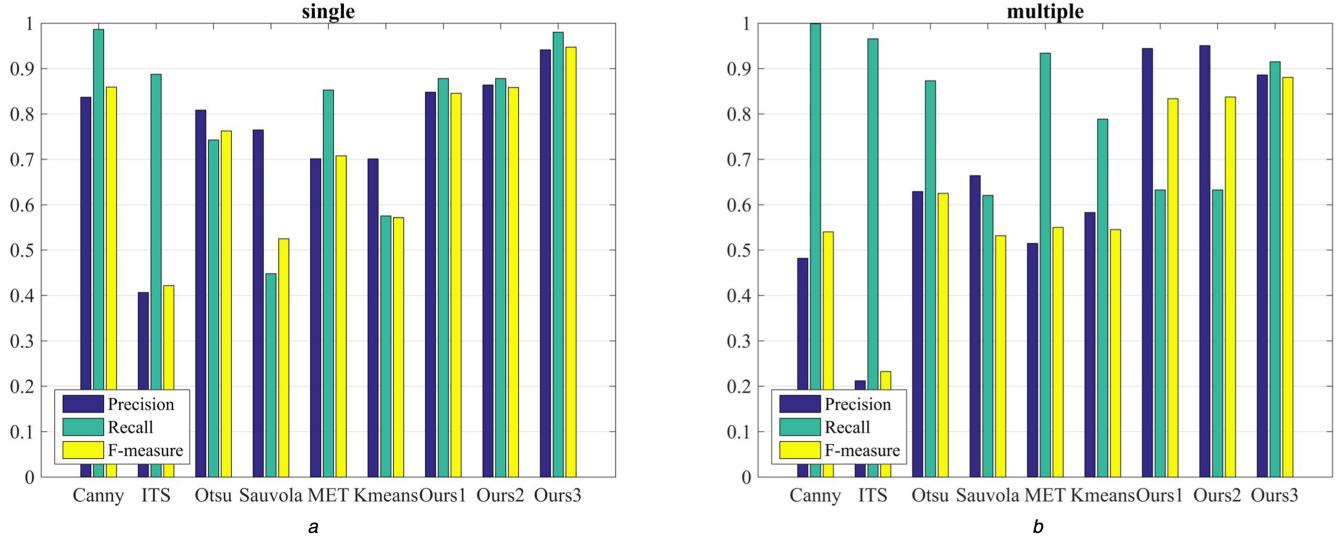
gbvs.php) [29], SR ('SR' means spectral residual from [30], and the code is available at <http://www.its.caltech.edu/~xhou/>) [30], PQFT ('PQFT' stands for the phase spectrum of quaternion Fourier transform from [31], and the code was provided by the authors) [31], SIG ('SIG' stands for SIGnature from [32], and the code is available at <http://www.its.caltech.edu/~xhou/>) [32], HFT ('HFT' stands for hypercomplex Fourier transform from [33], and the code was downloaded from <http://www.cim.mcgill.ca/~lijian>) [33], RC ('RC' means region contrast from [34], and the code is available at <http://mmcheng.net/code-data/>) [34], and IG (the code is available at [http://ivrgwww.epfl.ch/supplementary\\_material/RK\\_CVPR09/](http://ivrgwww.epfl.ch/supplementary_material/RK_CVPR09/)) [28]. The choice of these algorithms was motivated by the following reasons: ITTI was one of the first computational models for bottom-up saliency detection inspired by the spatial attention derived from the biologically plausible architecture, and the other methods were selected due to their computational efficiency for possible real-time analysis, among which SR was the first proposed spectral saliency model based on Fourier transform and then extended to PQFT and HFT via quaternion and hypercomplex Fourier transform, SIG was introduced as a simple powerful image descriptor to highlight sparse salient regions, and RC was proposed recently as a simple efficient multi-scale algorithm to produce full-resolution high-quality saliency maps. Similar to [28, 34], we evaluated the saliency maps from all the methods including only the saturation feature from HSI colour space on our proposed dataset by precision and recall. The precision value corresponded to the ratio of salient pixels correctly assigned to all the pixels of extracted regions, and the recall value was defined as the percentage of detected salient pixels in relation to the GT number. Similar to prior work, the PR curves were obtained by binarising the saliency map using thresholds in the range of 0–255. The resulting precision versus recall curves for images with single cells and multiple cells are shown in Figs. 2a and b, respectively. These curves provide a reliable comparison of how well various saliency maps highlight salient cell regions in images. The ITTI, SR, PQFT, SIG and HFT methods show high accuracy for a low recall (high threshold) and then the accuracy drops steeply. The salient pixels from these methods are within salient cell regions and present near uniform values but do not cover the entire salient cells. Although the RC method consistently outperforms the existing state-of-the-art salient object detection methods in terms of precision and recall [34], it is unsuitable for cell detection from phytoplankton microscopic images, except that it has a higher precision value than IG-based methods at nearly maximum recall for single cell detection (Fig. 2a), which only happens at a very low threshold close to zero. As shown in Fig. 2a, the saturation (S) performs better at low recall and worse at high recall for single cell

detection, thereby making the proposed IG + S method best with the recall in the range of [0.65, 0.9] (closest to the point [1, 1]). The yellow  $\star$  on Fig. 2a represents the actual threshold generated by our method for binarisation, which shows nearly the best segmentation on these PR curves. Overall, the proposed method (IG + S) can outperform all the selected methods on salient cell detection, especially multiple cell detection. The PR curves also indicate that the current detected saliency maps are insufficient (all far away from the ideally best result at [1, 1]) for accurate cell segmentation and need to be processed further.

**3.4.2 Cell segmentation:** We then compared our proposed method for cell segmentation with six commonly used segmentation methods: Canny (the implementation used the MATLAB function 'edge' with canny parameter) [35], iterative threshold selection (ITS) (<http://www.mathworks.com/matlabcentral/fileexchange/44255-ridler-calvard-image-thresholding>) [36], Otsu's thresholding (Otsu) (the implementation used the MATLAB function 'graythresh') [37], Sauvola adaptive thresholding (Sauvola) (<http://www.mathworks.com/matlabcentral/fileexchange/40266-sauvola-local-image-thresholding>) [38], minimum error thresholding (MET) (the code is available at <https://code.google.com/p/binalgorithm/>) [39] and K-means clustering (K-means) (we implemented this algorithm according to [40], and the code is available at <https://github.com/zhenglab/ACDS>) [40]. We tested the necessity and effectiveness of our proposed processing for the binarisation and denoising shown in Fig. 1. The results after binarisation and removing corner noise, as well as the final segmentation results and the selected six methods, were evaluated on our proposed dataset by *F*-measure, which is the overall performance measurement computed by the weighted harmonic of precision and recall, that is

$$F_\beta = \frac{(1 + \beta^2)\text{precision} \times \text{recall}}{\beta^2\text{precision} + \text{recall}}, \quad (18)$$

where we set  $\beta^2 = 0.3$  to emphasise the precision [28]. The comparisons for images with single cells and multiple cells are shown in Figs. 3a and b, respectively. Although several methods, such as Canny, ITS and MET, show high recall, especially for multiple cell segmentation, they perform worse than our proposed method because the recall rate is not as important as precision for segmentation. For example, a 100% recall rate can be achieved by simply selecting the entire image. The results after binarisation (Ours1) and denoising (Ours2) of our proposed method outperform the segmentation results of the selected six methods, and the final segmentation results (Ours3) of our proposed method exhibit the



**Fig. 3** Quantitative comparison of cell segmentation by different methods on our proposed dataset

(a) 225 microscopic images with single cells, (b) 15 microscopic images with multiple cells

**Table 1** Cell counting comparison on 15 microscopic images with multiple cells. The numbers in italics mean the truth (GT), and the numbers in bold indicate the best results

No.	Canny	ITS	Otsu	Sauvola	MET	K-means	Ours1	Ours2	Ours3	GT
#1	10	55	11	51	<b>12</b>	<b>12</b>	16	14	<b>12</b>	12
#2	8	7	7	57	7	8	10	10	<b>9</b>	9
#3	158	839	745	203	177	455	267	198	<b>125</b>	141
#4	207	3390	12	11	46	9	9	5	<b>3</b>	3
#5	12	5323	6	46,292	44	14,232	8	<b>2</b>	<b>2</b>	2
#6	35	107	6	5220	49	3014	17	<b>2</b>	<b>2</b>	2
#7	10	228	10	14	27	10	11	9	<b>5</b>	5
#8	8	11	5	4	11	3	21	3	<b>2</b>	2
#9	35	345	338	53	32	34	39	35	<b>31</b>	31
#10	6	614	9	7	9	4	3	<b>2</b>	<b>2</b>	2
#11	18	627	336	90	58	114	32	14	<b>12</b>	11
#12	35	224	37	42	121	33	<b>30</b>	29	29	31
#13	8	390	8	10	24	7	13	10	<b>5</b>	5
#14	<b>11</b>	<b>11</b>	<b>11</b>	26	<b>11</b>	<b>11</b>	12	<b>11</b>	10	<b>11</b>
#15	12	6934	<b>4</b>	41,825	114	15,283	47	21	5	2

#12, #14 and #15 are exceptions for Ours3 to count cells the most accurate but still very close to the truth (GT).

highest precision, recall and *F*-measure values, thus indicating the best performance.

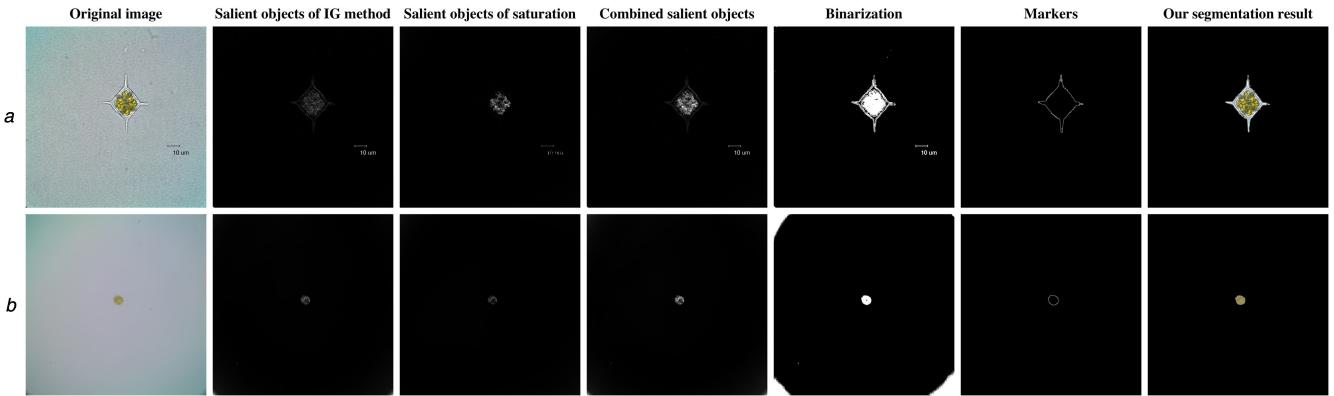
Biomorphic similarity, such as shape, is important for the further phytoplankton identification or classification. Therefore, we also used the modified Hausdorff distance [41] to measure the shape similarity of the GT to the above-mentioned segmentation results of different methods. For 225 microscopic images with a single cell (Supplemental Table S1), the contour (shape)-matching number most similar to the GT is 12 by Canny, 2 by Ours1/Ours2 and 211 by our final segmentation (Ours3). For 15 microscopic images with multiple cells (Supplemental Table S2), the corresponding matching number is 2 by Ours2 and 13 by Ours3, indicating the better performance of our proposed method on shape similarity of phytoplankton cells.

**3.4.3 Cell counting:** We also evaluated the preceding segmentation results by counting the number of connected regions automatically for cells counting on 15 images with multiple cells of our proposed dataset. Table 1 shows the results compared with the counting number from GT images representing the truth, indicating that our proposed method (Ours3) provides accurate counting results against the other methods (#12, #14 and #15 are exceptions but still very close to the truth). The inaccurate counting results of Ours1 and Ours2 also prove the necessity and effectiveness of the proposed processing for binarisation and denoising shown in Fig. 1.

### 3.5 Qualitative comparison

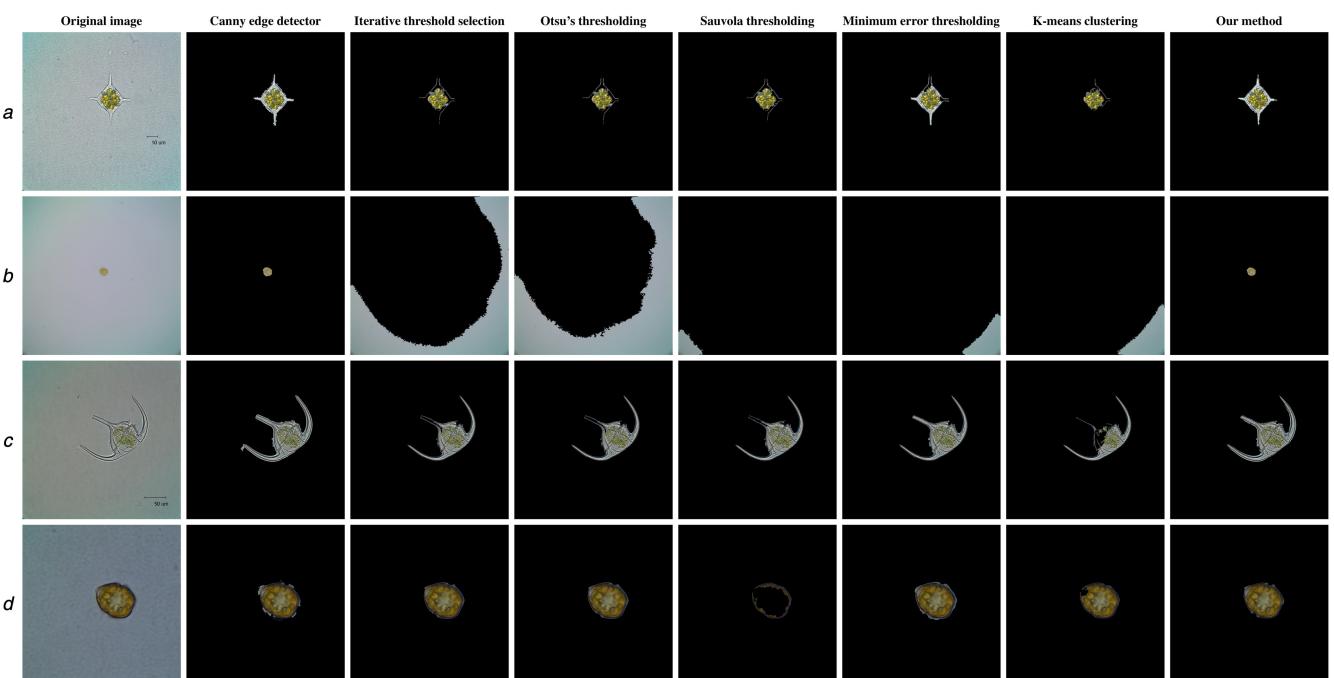
**3.5.1 Single-cell detection and segmentation:** Fig. 4 shows two examples of our method to segment the single cells from the microscopic images of non-setae phytoplankton species. The saliency maps generated by the IG method (column 2) preserve the boundaries, whereas the saliency maps generated by saturation (column 3) highlight the objects, which make the combined salient objects (column 4) detected and enhanced uniformly. Although the binary maps (column 5) still contain some noise, such as the ruler label (10  $\mu\text{m}$ ) in the lower right of the image shown in the column 5 of Fig. 4a, as well as the corner noise shown in the column 5 of Fig. 4b, the markers (column 6) are produced without noise by removing fake markers in our method. The segmentation results (column 7) indicate the feasibility and effectiveness of our proposed method for single cell detection and segmentation.

**3.5.2 Single-cell segmentation comparison:** Fig. 5 provides the visual single-cell segmentation comparison of our method and the six commonly used methods: Canny [35], ITS [36], Otsu [37], Sauvola [38], MET [39] and K-means [40] (the first column shows the original RGB microscopic images, and the remaining columns present the methods mentioned and our proposed method). We extracted the region of maximum size as the cell to remove the relatively big chunks of noise for the six methods to find exactly the cell object that always occupies the largest area compared with



**Fig. 4** Experimental results of our proposed method for single-cell detection and segmentation. The first column shows the original RGB microscopic images of the following non-setae species in each row

(a) *Dictyocha fibula*, (b) *Chattonella marina*. Columns 2–6 present the image results of salient objects detected by the IG method, the salient objects detected by saturation, our combined salient objects, the binarisation of combined salient objects, and the markers containing internal (black regions in the objects) and external (black regions outside the objects) markers imposed on the grey-level microscopic images, respectively. The last column shows the final segmentation results of our proposed method



**Fig. 5** Visual comparison with the commonly used segmentation methods on single-cell segmentation. The first column shows the original RGB microscopic images of the following non-setae species in each row

(a) *Dictyocha fibula*, (b) *Chattonella marina*, (c) *Ceratium tripos*, (d) *Scrippsiella trochoidea*. For comparison, the remaining columns present the results obtained by the following segmentation methods consecutively: Canny, ITS, Otsu, Sauvola, MET, K-means and our proposed method

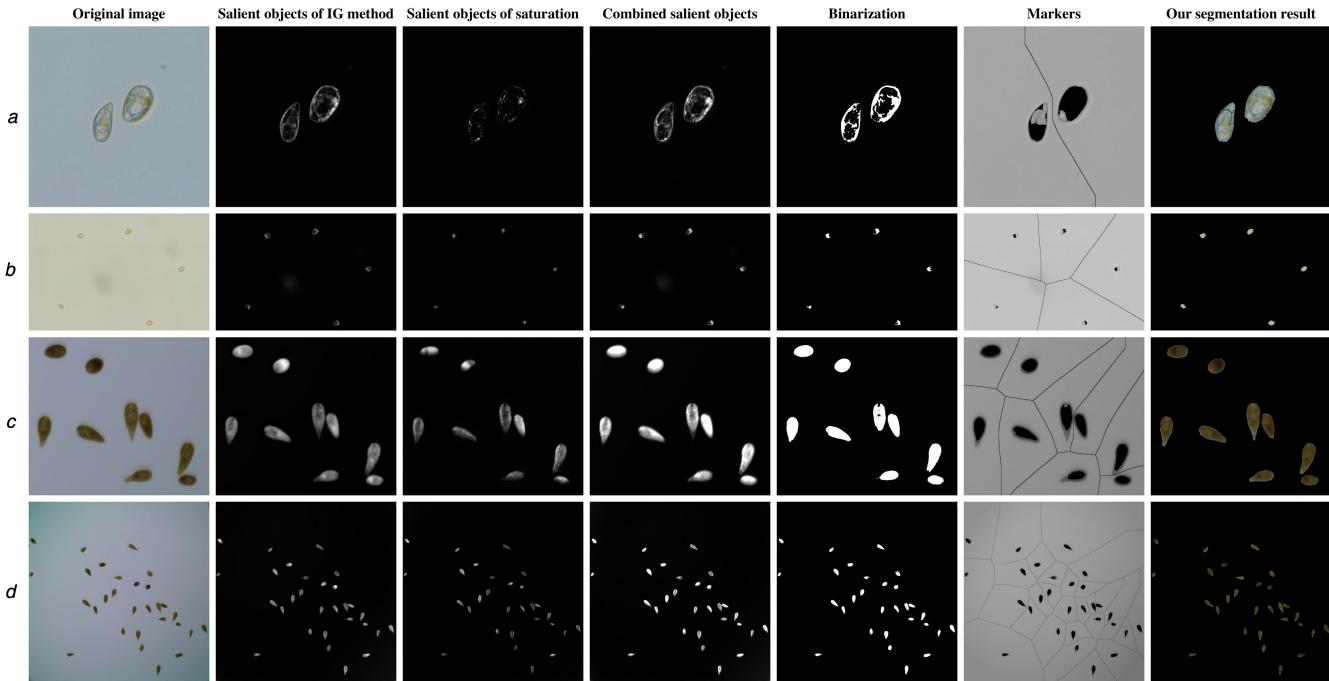
the noise in the microscopic image. Figs. 5a and b depict that the thresholding (columns 3–6) and clustering (column 7) methods are nearly ineffective to segment the entire cell because of the low-contrast blurred edges and uneven illumination. Although the Canny (column 2) can segment the cell better than other commonly used methods, it still fails to keep the edge in details, especially the biomorphic characteristics, such as cingulum (column 2 of Figs. 5c and d) compared with our method (column 8 of Fig. 5c and d).

**3.5.3 Multiple cell detection and segmentation:** The experiments of segmenting multiple cells from non-setae phytoplankton microscopic images using our method are shown in Fig. 6. The improved salient object detection (column 4) can detect and highlight all the salient objects uniformly while disregarding the inconspicuous blurred noise (the upper right small noise in the column 1 of Fig. 6a and the centre large blurred noise in the column 1 of Fig. 6b do not appear in column 4). Although the cells shown in the binary maps (column 5) and the markers (column 6) are incomplete, the segmentation results shown in column 7 indicate that our proposed method can extract all cells in details completely. The selected external markers (the lines shown in

column 6) cause all the cells to segment separately, thereby making the automatic counting easy and efficient (e.g. Fig. 6d). However, our method will fail to segment and separate the cells from the overlapping objects, which will be our future work.

#### 4 Conclusions

The in situ imaging system is a useful tool to monitor and understand how environmental processes affect plankton (especially HABs), which in turn can contribute to understand the impact of climate change. Therefore, numerous plankton-imaging systems are developed and tested in situ (i.e. underwater) to overcome the shortcomings of traditional plankton laboratory analysis that requires expensive laboratory settings and time-consuming analysis, which also makes the automated plankton image analysis attractive. The plankton image analysis mainly focuses on the identification (or classification) and counting of special species to answer the three fundamental questions about planktonic organisms proposed by Victor Hensen [8]. Although cell detection and segmentation are the key steps that cannot be bypassed at the beginning of plankton image analysis (the same



**Fig. 6** Experimental results of our proposed method for multiple cell detection and segmentation. The first column shows the original RGB microscopic images of the following non-setae species in each row

(a) *Prochlorococcus triestinum*, (b) *Amphidinium carterae*, (c, d) *Chattonella marina*. Columns 2–6 present the image results of the salient objects detected by the IG method, the salient objects detected by saturation, our combined salient objects, the binarisation of combined salient objects, and the markers containing internal (black regions on the objects) and external (black lines between the objects) markers imposed on the grey-level microscopic images, respectively. The last column shows the final segmentation results of our proposed method

importance as the object detection and segmentation in computer vision), only a few studies have contributed to them in recent years [12, 13, 24]. Following the state-of-the-art development from computer vision, we investigated the properties of phytoplankton cells in microscopic images and proposed the saliency-based marker-controlled watershed method for automatic cell detection and segmentation.

The capability to perceive objects before identifying them is closely related to bottom-up visual attention (saliency), which refers to the capability to select important visual information for further processing. Thus, accurate and reliable saliency detection can benefit from numerous tasks ranging from low-level segmentation to high-level recognition. Real-time analysis for on-site and in situ systems requires simple implementation and computational efficiency. Therefore, successful salient object detection can also be used to facilitate automated cell detection of in situ plankton-imaging systems, such as VPR [42] and Scripps Plankton Camera [43].

The detected salient objects usually contain some non-cell noise due to the characteristics of illumination and noise in microscopic images. Obtaining the real cells, especially some specific species from all objects, can be considered the class-specific object detection (such as face detection) in computer vision [44]. We classified the objects into cells and noise by dividing the segmentation purpose into single and multiple cells and removed the noise according to the microscopic image characteristics of different purposes using different methods, thereby yielding efficient performance on cell or cell extraction. For the cases wherein whether the image contains one single cell or multiple cells is unknown, the size of the detected salient objects may be helpful to determine the processing method. The completely automatic processing without purpose partition can also be implemented by providing some prior knowledge, such as shape feature (e.g. circle), which belongs to the learning-related vision problems.

The identification and classification of phytoplankton cells rely mainly on the characteristics of biological morphology, which require the accurate and complete segmentation of cells from microscopic images. The analysis of microscopic image characteristics presents that our proposed saliency-based marker-

controlled watershed method can overcome the incorrect and incomplete segmentation caused by the specks and debris, uneven illumination and blurred edges. The quantitative and qualitative experimental results also show that our method not only outperforms the commonly used segmentation methods on the single cell segmentation more accurately and completely but also extracts all the cells separately and completely from the microscopic images containing multiple cells. Therefore, our method presents potential future applications for automatic phytoplankton analysis, such as identification and classification in the on-site and in situ real-time systems.

Although our proposed method can be used to ‘find’ and ‘extract’ the cells from microscopic images automatically, phytoplankton is still difficult to identify and classify due to their diverse shapes from different species in different growth stages and different viewpoints in different orientations under the microscope, even by humans. However, the segmented cells can be used for feature extraction to represent their biomorphic characteristics for further training and learning, and then identification and classification; this task is also the ultimate goal of our research.

To encourage future work, we make the source code, dataset and all results available in the project page: <http://vision.ouc.edu.cn/zhenhaiyong/research/acds/>, and GitHub repository: <https://github.com/zhenlab/ACDS>.

## 5 Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant nos. 61771440, 41776113 and 61301240, and Qingdao Municipal Science and Technology Program under grant no. 17-1-1-5-jch.

## 6 References

- [1] Olson, R.J., Sosik, H.M.: ‘A submersible imaging-in-flow instrument to analyze nano-and-microplankton: imaging flowCytobot’, *Limnol. Oceanogr.: Methods*, 2007, **5**, pp. 195–203
- [2] Erickson, J.S., Hashemi, N., Sullivan, J.M., et al.: ‘In situ phytoplankton analysis: theres plenty of room at the bottom’, *Anal. Chem.*, 2011, **84**, (2), pp. 839–850
- [3] Reynaud, E.G.: ‘Imaging marine life: macrophotography and microscopy approaches for marine Biology’ (Wiley-Blackwell, 2013)

- [4] Watson, J.E., Zielinski, O.: 'Subsea optics and imaging' (Elsevier, Singapore, 2013)
- [5] Jaffe, J.S.: 'Underwater optical imaging: the past, the present, and the prospects', *IEEE J. Oceanic Eng.*, 2015, **40**, (3), pp. 683–700
- [6] du Buf, H., Bayer, M.M.: 'Automatic diatom identification' (World Scientific, 2002)
- [7] Culverhouse, P.F., Williams, R., Benfield, M., *et al.*: 'Automatic image analysis of plankton: future perspectives', *Mar. Ecol. Prog. Ser.*, 2006, **312**, pp. 297–309
- [8] Benfield, M.C., Grosjean, P., Culverhouse, P.F., *et al.*: 'RAPID: research on automated plankton identification', *Oceanography*, 2007, **20**, (2), pp. 172–187
- [9] MacLeod, N., Benfield, M., Culverhouse, P.: 'Time to automate identification', *Nature*, 2010, **467**, (7312), pp. 154–155
- [10] Santhi, N., Pradeepa, C., Subashini, P., *et al.*: 'Automatic identification of algal community from microscopic images', *Bioinf. Biol. Insights*, 2013, **7**, pp. 327–334
- [11] Verikas, A., Gelzinis, A., Bacauskiene, M., *et al.*: 'An integrated approach to analysis of phytoplankton images', *IEEE J. Oceanic Eng.*, 2015, **40**(2), pp. 315–326
- [12] Jalba, A.C., Wilkinson, M.H.F., Roerdink, J.B.T.M.: 'Automatic segmentation of diatom images for classification', *Microsc. Res. Tech.*, 2004, **65**, pp. 72–85
- [13] Zheng, H., Zhao, H., Sun, X., *et al.*: 'Automatic setae segmentation from Chaetoceros microscopic images', *Microsc. Res. Techniq.*, 2014, **77**, (9), pp. 684–690
- [14] Embleton, K.V., Gibson, C.E., Heaney, S.I.: 'Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method', *J. Plankton. Res.*, 2003, **25**, (6), pp. 669–681
- [15] Davis, C.S., Hu, Q., Gallager, S.M., *et al.*: 'Real-time observation of taxonomic plankton distributions: an optical sampling method', *Mar. Ecol. Prog. Ser.*, 2004, **284**, pp. 77–96
- [16] Rodenacker, K., Hense, B., Jüttig, U., *et al.*: 'Automatic analysis of aqueous specimens for phytoplankton structure recognition and population estimation', *Microsc. Res. Techniq.*, 2006, **69**, (9), pp. 708–720
- [17] Bi, H., Guo, Z., Benfield, M.C., *et al.*: 'A semi-automated image analysis procedure for in situ plankton imaging systems', *PLoS One*, 2015, **10**, (5), p. e0127121
- [18] Blaschko, M.B., Holness, G., Mattar, M.A., *et al.*: 'Automatic in situ identification of plankton', Proc. Seventh IEEE Workshops on Application of Computer Vision IEEE, Breckenridge, CO, USA, 5–7 January 2005, vol. 1, pp. 79–86
- [19] Sosik, H.M., Olson, R.J.: 'Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry', *Limnol. Oceanogr. Methods*, 2007, **5**, pp. 204–216
- [20] Luo, Q., Gao, Y., Luo, J., *et al.*: 'Automatic identification of diatoms with circular shape using texture analysis', *J. Softw.*, 2011, **6**, (3), pp. 428–435
- [21] Mosleh, M.A.A., Manssor, H., Malek, S., *et al.*: 'A preliminary study on automated freshwater algae recognition and classification system', *BMC Bioinform.*, 2012, **13**, (Suppl. 17), p. S25
- [22] Kloster, M., Kauer, G., Beszteri, B.: 'SHERPA: an image segmentation and outline feature extraction tool for diatoms and other objects', *BMC Bioinformatics*, 2014, **15**, (1), p. 218
- [23] Verma, S., Khare, D., Gupta, R., *et al.*: 'Analysis of image segmentation algorithms using MATLAB'. Proc. Third Int. Conf. Trends in Information, Telecommunication and Computing, Bangalore, India, 3–4 August 2012, 2013, pp. 163–172
- [24] Verikas, A., Gelzinis, A., Bacauskiene, M., *et al.*: 'Phase congruency-based detection of circular objects applied to analysis of phytoplankton images', *Pattern Recogn.*, 2012, **45**, (4), pp. 1659–1670
- [25] Fischer, S., Shahbazkia, H.R., Bunke, H.: 'Contour extraction', In du Buf, H., Bayer, M.M. (Eds.): 'Automatic diatom identification' (World Scientific, Singapore, 2002), pp. 93–107
- [26] Dimitrovski, I., Kocev, D., Loskovska, S., *et al.*: 'Hierarchical classification of diatom images using ensembles of predictive clustering trees', *Ecol. Inform.*, 2012, **7**, (1), pp. 19–29
- [27] Li, Y., Hou, X., Koch, C., *et al.*: 'The secrets of salient object segmentation', Proc. IEEE Conf. Computer Vision and Pattern Recognition IEEE, Columbus, OH, USA, 24–27 June 2014, pp. 280–287
- [28] Achanta, R., Hemami, S., Estrada, F., *et al.*: 'Frequency-tuned salient region detection', Proc. IEEE Conf. Computer Vision and Pattern Recognition IEEE, Miami, FL, USA, 20–25 June 2009, pp. 1597–1604
- [29] Itti, L., Koch, C., Niebur, E.: 'A model of saliency-based visual attention for rapid scene analysis', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1998, **20**, (11), pp. 1254–1259
- [30] Hou, X., Zhang, L.: 'Saliency detection: a spectral residual approach', Proc. of IEEE Conf. on Computer Vision and Pattern Recognition IEEE, Minneapolis, Minnesota, USA, 18–23 June 2007, pp. 1–8
- [31] Guo, C., Ma, Q., Zhang, L.: 'Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform', Proc. IEEE Conf. Computer Vision and Pattern Recognition IEEE, Anchorage, AK, USA, 24–26 June 2008, pp. 1–8
- [32] Hou, X., Harel, J., Koch, C.: 'Image signature: highlighting sparse salient regions', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, **34**, (1), pp. 194–201
- [33] Li, J., Levine, M.D., An, X., *et al.*: 'Visual saliency based on scale-space analysis in the frequency domain', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (4), pp. 996–1010
- [34] Cheng, M.-M., Mitra, N.J., Huang, X., *et al.*: 'Global contrast based salient region detection', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, **37**, (3), pp. 569–582
- [35] Cannby, J.: 'A computational approach to edge detection', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1986, **8**, (6), pp. 679–698
- [36] Ridler, T.W., Calvard, S.: 'Picture thresholding using an iterative selection method', *IEEE Trans. Syst. Man Cybern.*, 1978, **8**, (8), pp. 630–632
- [37] Otsu, N.: 'A threshold selection method from gray-level histograms', *IEEE Trans. Syst. Man Cybern.*, 1979, **9**, (1), pp. 62–66
- [38] Sauvola, J., Pietikäinen, M.: 'Adaptive document image binarization', *Pattern Recogn.*, 2000, **33**, (2), pp. 225–236
- [39] Kittler, J., Illingworth, J.: 'Minimum error thresholding', *Pattern Recogn.*, 1986, **19**, (1), pp. 41–47
- [40] Hartigan, J.A., Wong, M.A.: 'Algorithm AS 136: a k-means clustering algorithm', *J. R. Stat. Soc. Ser. C Appl. Stat.*, 1979, **28**, (1), pp. 100–108
- [41] Dubuisson, M.-P., Jain, A.K.: 'A modified Hausdorff distance for object matching', Proc. 12th IAPR Int. Conf. Pattern Recognition IEEE, Jerusalem, Israel, 9–13 October 1994, pp. 566–568
- [42] Davis, C.S., Thwaites, F.T., Gallager, S.M., *et al.*: 'A three-axis fast-tow digital video plankton recorder for rapid surveys of plankton taxa and hydrography', *Limnol. Oceanogr. Methods*, 2005, **3**, pp. 59–74
- [43] Jaffe, J.S., Roberts, P.L.D., Ratelle, D., *et al.*: 'Scripps plankton camera system', (2015). Available at <http://spc.ucsd.edu/>
- [44] Alexe, B., Deselaers, T., Ferrari, V.: 'What is an object?'. Proc. IEEE Conf. Computer Vision and Pattern Recognition: IEEE, San Francisco, USA, 13–18 June 2010, pp. 73–80