Automated Processing of Oceanic Bubble Images for Measuring Bubble Size Distributions underneath Breaking Waves

RAIED S. AL-LASHI[a] AND STEVE R. GUNN

Electronics and Computer Science, Faculty of Physical Sciences and Engineering, University of Southampton, Southampton, United Kingdom

HELEN CZERSKI

Department of Mechanical Engineering, University College London, London, United Kingdom

ABSTRACT

Accurate in situ measurements of oceanic bubble size distributions beneath breaking waves are needed for a better understanding of air–sea gas transfer and aerosol production processes. To achieve this goal, a novel high-resolution optical instrument for imaging oceanic bubbles was designed and built in 2013 for the High Wind Gas Exchange Study (HiWinGS) campaign in the North Atlantic Ocean. The instrument is able to operate autonomously and can continuously capture high-resolution images at 15 frames per second over an 8-h deployment. The large number of images means that it is essential to use an automated processing algorithm to process these images. This paper describes an automated algorithm for processing oceanic images based on a robust feature extraction technique. The main advantages of this robust algorithm are it is significantly less sensitive to the noise and insusceptible to the background changes in illumination, can extract circular bubbles as small as one pixel (approximately 20 μm) in radius accurately, has low computing time (approximately 5 seconds per image), and is simple to implement. The algorithm was successfully used to analyze a large number of images (850 000 images) from deployment in the North Atlantic Ocean as part of the HiWinGS campaign in 2013.

1. Introduction

Bubble plumes entrained by breaking waves in the open ocean have a significant influence on various oceanographic phenomena, including air–sea gas transfer (Farmer et al. 1993; Wanninkhof et al. 2009), marine aerosol production (Fuentes et al. 2010), and scavenging of surfactants (Zhou et al. 1998). Also, they have an important influence on the optical (Frouin et al. 2001; Stramski 1994; Stramski and Tegowski 2001; Zhang et al. 1998) and acoustical (Ainslie 2005; Terrill and Melville 2000) properties in the upper ocean. The most important factor in understanding these processes is the distribution of bubble sizes in the top few meters of ocean (Deane and Stokes 2002). The bubble plumes formed during the first seconds after a wave breaks are characterized by very high void fraction (0.1%–10%), rapid changes in bubble size distributions, and a wide range of bubble radii, from a few microns to a few millimeters (Czerski et al. 2011). Consequently, a detailed understanding of the physics requires in situ and precise measurements of bubble populations with adequate temporal and spatial resolution.

A high-resolution bubble imaging instrument was designed and successfully deployed in the North Atlantic Ocean. The instrument frame rate was 15 frames per

---

[a] Current affiliation: School of Food Science and Nutrition, Faculty of Mathematics and Physical Sciences, University of Leeds, Leeds, United Kingdom.

Corresponding author address: Raied S. Al-Lashi, School of Food Science and Nutrition, Faculty of Mathematics and Physical Sciences, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, United Kingdom.
E-mail: r.al-lashi@leeds.ac.uk

Publisher's Note: This article was revised on 24 October 2016 to include the CCBY license that was added after initial publication.

second and the effective exposure time for each frame was 4 μs. The image resolution was 2048 $\times$ 2048 pixels, and the total recording time for a single deployment was approximately 8.5 h. Extracting bubble sizes accurately from these images is very important for further data analysis. Oceanographic conditions change slowly over many hours, so any meaningful monitoring of bubble plumes must record data over long time periods while maintaining high time resolution. Therefore, an automated and robust algorithm is required for bubble image analysis.

Several algorithms have been developed in the past to analyze bubble images. Some of these algorithms have been used to analyze oceanic images. For instance, Stokes and Deane (1999) developed an optical instrument to study the bubbles within breaking waves. Their image processing algorithm included two preprocessing operations: correcting nonuniform illumination and thresholding. Correcting the nonuniform illumination involved four steps: morphological operation (closing) to estimate the background illumination, convolution with a 15 $\times$ 15 Gaussian kernel to smooth the image, subtraction of the smoothed image from the original image to remove background illumination variation, and scaling to improve the image contrast. Thresholding was then applied to create a new binary image followed by a Hough transform to detect the bubbles in that binary image. Leifer et al. (2003) introduced another instrument for imaging bubbles within breaking waves. They used a thresholding technique to produce binary images. However, this can potentially introduce large errors in the measured bubble size distribution because of the change in the background illumination. Furthermore, the bubbles positioned within the light sheet are much brighter compared to the bubbles outside the sheet. Thomanek et al. (2010) demonstrated an automated gas bubble imaging system to measure bubble sizes at the seafloor. They used a Canny edge detector and the MATLAB function "regionprops" to determine the size and center of each bubble. In comparison to thresholding, edge detection is more accurate in terms of identifying bubbles in inhomogeneous illuminated areas and analyzing images with a gradual decrease in light intensities. However, the main disadvantage of edge detection is the enhancing and shrinking in the bubble boundaries. Zielinski et al. (2010) showed a laboratory setup that consists of frontal illumination and video camera to image bubbles in an aquarium. They used a sequence of optical flow algorithm, thresholding, and region filtering to process the video images. The optical flow analysis was also used by Boelmann and Zielinski (2015) to identify bubbles in images collected by a remotely operated underwater vehicle in west Svalbard. In the optical flow method, the apparent motion of bubbles in the images can be related to each other

as a collection of displacements in the image plane. However, this method is prone to inaccuracies caused by illumination changes, occlusion, and noise (Nixon and Aguado 2002). Wang and Socolofsky (2015) developed a stereoscope imaging system for measurement of natural seep bubble size distributions in the Gulf of Mexico. They applied Canny edge detection for low-density bubble images. For high-density bubble images, they used a Sobel gradient mask to obtain a binary image and a watershed transform to perform image segmentation on the overlapping bubbles. However, the watershed transform is sensitive to noise (Honkanen et al. 2005).

On the other hand, a number of algorithms have been developed to analyze bubble images collected from chemical reactors and bioreactors. For instance, Taboada et al. (2006) presented a semiautomated image analysis algorithm to count bubble sizes and oil droplets in complex dispersions occurring in bioreactors. Their algorithm consists of two stages: preprocessing to obtain a binary image and postprocessing to extract circular bubbles and droplets using the Hough transform. The first preprocessing stage was achieved using commercial imaging software (Image-Pro Plus 5.0, Media Cybernetics) and consists of three filtering operations (median, flatten, and well filter) and two morphological operations (opening and skeleton). Honkanen et al. (2005) described an experimental setup to image a turbulent flow in a pipe and to detect in focus, overlapping, and elliptical bubbles. Three preprocessing steps were carried out according to their recognition algorithm: median filter to remove background noise, image normalization to normalize pixel intensities, and thresholding and grayscale gradients to obtain a binary image. The overlapping elliptical bubbles in this binary image were extracted by examining the perimeter arcs of individual bubbles. This is achieved by calculating the overall perimeter of a segment, finding the connected points at the perimeters of the overlapping objects, grouping the perimeter arcs for the same object, and fitting ellipse to the clustered perimeter arcs of the object. The main disadvantage of their algorithm is it is computationally expensive. Zhong et al. (2016) established an experimental setting to image bubbly water in a gas–liquid reactor. Their image analysis method was based on three preprocessing operations (background subtraction, median filtering, and thresholding), forming a template database of single-bubble images, splitting contours for every overlapped bubble, and reconstructing the segmented bubbles. The main drawbacks in this method are that it cannot run autonomously because single-bubble images are required for every bubble in the original image and it is very sensitive to the background change in illumination.

The algorithms discussed above used preprocessing techniques such as filtering to remove noise and low-level

feature extraction approaches such as thresholding and edge detection to extract basic features in the image to speed up the subsequent high-level feature extraction stage (Gonzalez and Woods 2008; Nixon and Aguado 2002). However, there are many limitations in the performance of these techniques. The performance of thresholding techniques is limited by the object (bubble) size, contrast, noise, mean difference between the object and the background intensities, and variances of object and background (Lee and Chung 1990). Noise reduction by filtering operations results in blurred and distorted edges since both noise and edges contain high-frequency contents (Liu and Fang 2015). Traditional edge detector operators such as Sobel, Laplacian, Roberts, and Canny are based on gradient methods (Gonzalez and Woods 2008; Nixon and Aguado 2002). Therefore, these first-order detectors are sensitive to noise (Liu and Fang 2015). Abdou and Pratt (1979) developed a quantitative figure of merit to evaluate the ability of these traditional operators to detect edges as close as possible to the ideal edges. They demonstrated that this figure of merit substantially decreased with reducing the signal-to-noise ratio in the image. Liu and Fang (2015) used the Abdou figure of merit to evaluate applying these traditional edge detector operators on various images with the same noise level. They showed that the performance of these detectors depends significantly on the shape of the objects in these images. Moreover, they found that the detection accuracy decreased to approximately 54% in an image of circular coins. Because of these difficulties, any preprocessing must be applied in such a way as to not remove or distort the underlying signal of interest, and the ideal solution is an algorithm that avoids the preprocessing stage entirely.

This paper describes a robust automated algorithm for analyzing oceanic images based on the Hough transform. The algorithm uses the intensity information in the images directly without any preprocessing stage. The algorithm can extract circular bubbles over a wide radii range from 1 to 25 pixels (approximately 20–406 μm) accurately. The paper is organized as follows: a brief description of the imaging instrument is given in section 2; the deployment and measurement in the ocean is illustrated in section 3; section 4 describes the automated bubble extraction algorithm, its implementation, and evaluation; section 5 presents bubble size distributions obtained from applying the algorithm on a sequence of images; the discussion of the results is in section 6; and we conclude with section 7.

2. Optical instrument for imaging bubbles

Many acoustical and optical techniques have been developed to measure bubble size distributions in the laboratory and open sea. The acoustical techniques (Farmer et al. 1993; Medwin 1970; Medwin 1977) are sensitive to bubble radii from 1 to 500 μm and low void fraction, while optical techniques (Geißler and Jähne 1995; Jähne and Geißler 1995; Leifer et al. 2003; Wang and Monahan 1995) can be used to measure bubble distributions at low and high void fraction and over a wide radii range, from 20 μm to 5 mm.

This section provides a brief description of the bubble imaging instrument that we designed to capture high-quality images and to increase the measured bubble size range. More details about the design considerations and hardware components can be found in Al-Lashi et al. (2016, manuscript submitted to *IEEE J. Oceanic Eng.*) The camera and its control electronics were housed in a waterproof pressure case (see Fig. 1). Strobe lighting illuminated a thin slice of water approximately 4 cm × 4 cm × 5 mm, and this sample volume was positioned a few centimeters in front of the camera housing.

The hardware components of the bubble imaging system can be divided into six modules: the power management board to supply the necessary power to the components; the strobe system to form the light sheet; the machine vision camera; the single-board computer that controls the camera and saves the images on the solid-state drive; and the waterproof enclosure to protect the electronic modules. The principal operation of the instrument is based on the formation of a light sheet in front of an optical Perspex window. Images are formed by focusing the scattered light caused by the bubbles inside the light sheet through a megapixel telecentric lens mounted on a high-resolution charge-coupled device (CCD) camera.

The hardware components were mounted in a waterproof housing that is divided into three chambers separated by two steel disks as shown in Fig. 1. The top chamber contains the single-board assembly with the main electronics, the middle chamber contains the camera assembly with the imaging components, and the bottom chamber contains the strobe assembly with the illumination components. The mirror assembly was fixed outside the housing to form the light sheet in front of camera optical window.

3. Deployments and measurements in the open sea

The bubble imaging instrument in its autonomous configuration was deployed 7 times in the North Atlantic Ocean in 2013 (including buoyancy and instrument configuration trials) and deployment lengths ranged from a few hours to five days. These deployments were part of the High Wind Gas Exchange Study (HiWinGS) campaign to study air–sea gas exchange
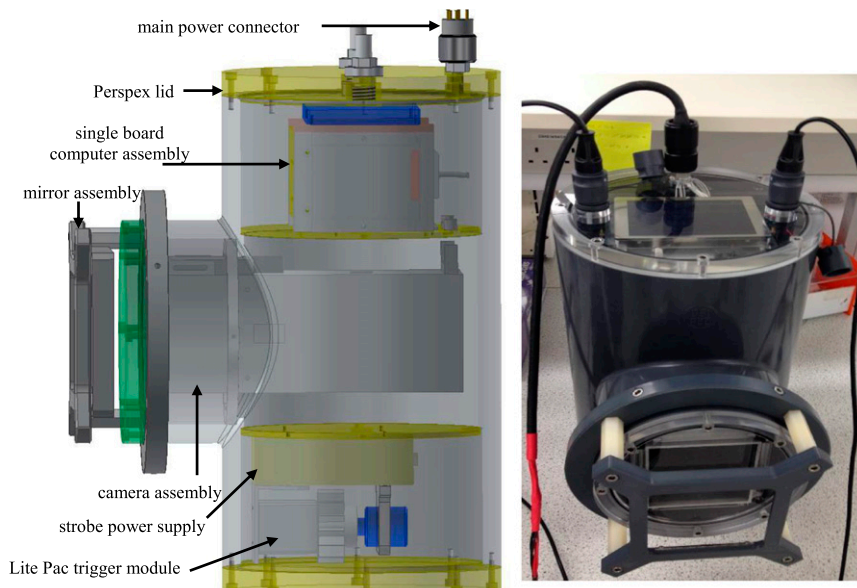
FIG. 1. (left) A cross section through the instrument housing, showing the layout of the hardware components. (right) The bubble optical imaging instrument and its housing. The case diameter is 225 mm × 160 mm, and its length is 407 mm.

during storms. As part of this campaign, a suite of bubble measurement instruments was mounted on a large free-floating buoy. The average hourly wind speed during the deployments ranged from 10 to $30\,\mathrm{m\,s^{-1}}$, and the significant wave height varied from 1 to 10 m. Figure 2 shows the bubble instrument mounted on the free-drifting spar buoy as the buoy is being deployed. The camera was positioned approximately 2 m below the ocean surface when the buoy was floating upright. The design and performance of the spar buoy was well described in Pascal et al. (2011). Whitecaps (the patches of foam left at the surface after the passage of a breaking wave) were imaged by another camera system positioned above the water surface. To synchronize both instruments, the power supplied to them was controlled by two timers that were programmed at preset intervals. The total number of captured images during seven deployments was approximately 850 000.

Depending on the bubble shapes and the activities in the ocean, the captured images can be classified into three main categories: large bubble, small bubble, and complex images. The vast majority of these captured images contain small bubbles. The number of complex and big bubble images in the seven deployments was 5 and 269, respectively. These images are classified automatically by the algorithm as illustrated in the implementation section (section 4a).

The air bubbles in the light sheet appear as bright circular objects in the images. The bubble shape depends on the surface tension that dominates the shape as a bubble gets smaller. Therefore, small bubbles tend to be spherical, while large bubbles are more likely nonspherical (Leighton 1994). More complex bubble shapes can be described mathematically using spherical harmonics (Leighton 1994). Figure 3 shows a sample of the captured images during deployment in the North Atlantic Ocean. The big bubbles are mainly nonspherical in Fig. 3a, while the small bubbles are circular in Fig. 3b. Moreover, the small bubbles that are not located in the light sheet appear as disks. Some of the small bubble images also contain marine creatures, such as a copepod, as shown in Fig. 3c.
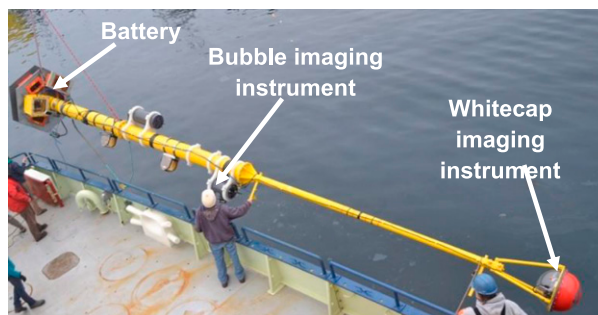

FIG. 2. The bubble imaging instrument discussed here attached to a spar buoy during deployment in the sea. The spar buoy length is 11 m. The automated algorithm described in this paper was used to analyze the images collected by the bubble imaging instrument while it was mounted in this configuration.
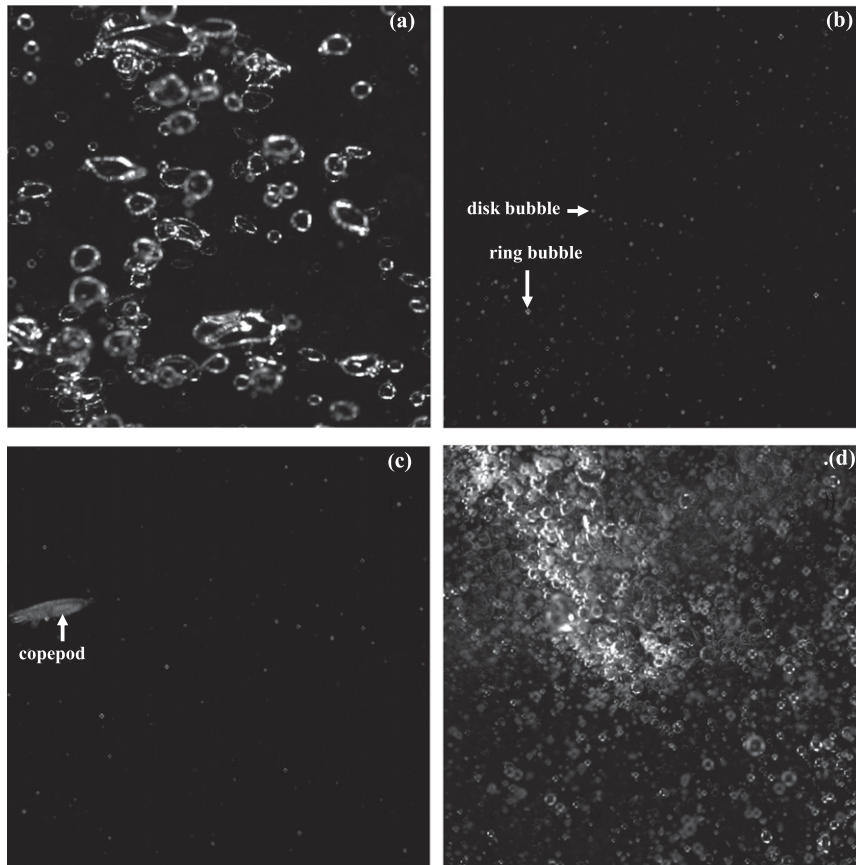
FIG. 3. Sample of the images captured during deployment in the North Atlantic Ocean. The dimensions of these images are 4 cm × 4 cm. (a) Nonspherical bubbles. (b),(c) Two types of circular bubbles: rings and disks. The ring bubble appears as a bright ring (high intensity) around a dark center (low intensity), while the disk bubble appears as a filled circle. (c) A copepod. (d) Complex image that contains a bubble plume.

Complex images contain bubble features that are not extracted by the automated algorithm described in this paper. We chose to implement a simpler algorithm that avoids processing complex images since the number of these images is significantly small (five images only). However, bubbles in these images can still provide valuable information. For instance, the bubble plume shown in Fig. 3d could provide useful information about bubble formation mechanisms (Deane and Stokes 1999).

4. Automated bubbles extraction algorithm

Bubbles in the image appear as high-intensity rings or disks as shown in Fig. 4. The intensities of the bubble wall pixels are higher than the surrounding pixels in the image background and can be used by an algorithm to automatically locate and extract the position and size of the bubbles. There are a number of possible algorithms that could be used for the extraction. These range from model-based approaches to deformable approaches that can accommodate large variations in shape characteristics. Here, we are primarily interested in extracting bubbles in the size range 20–500 μm and these are spherical to a good approximation due to surface tension. Hence, a model-based approach using the Hough transform is chosen for its robustness. The automated algorithm described in this section is not suitable to extract noncircular and overlapped bubbles in complex images.

a. Hough transform for circular shapes

The Hough transform (Hough 1960) is a high-level feature extraction technique based on shape matching (Nixon and Aguado 2002). In particular, it is widely used to extract lines (Duda and Hart 1972), circles (Kimme et al. 1975), and arbitrary shapes (Ballard 1981) from images. The technique is based on an evidence gathering
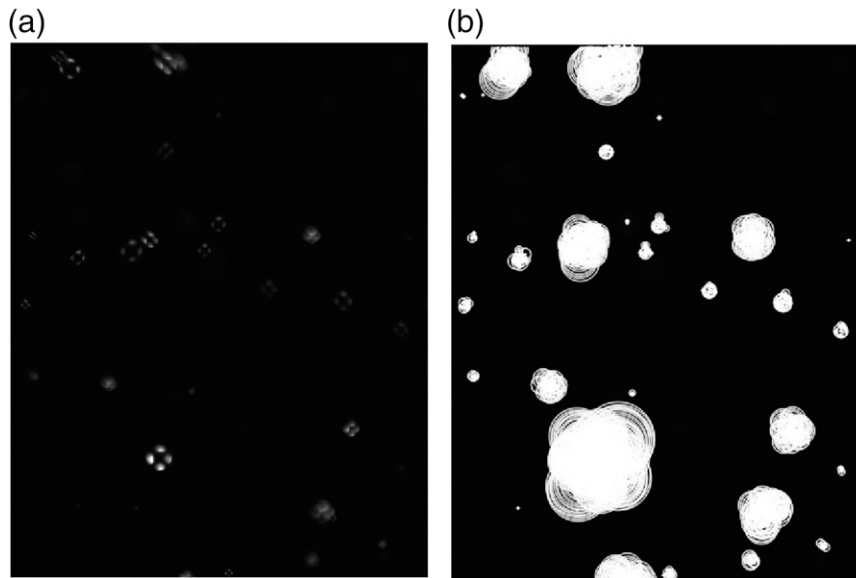
(a)   (b)

FIG. 4. Hough transform feature extraction without peak filtering. (a) A slice in an origin image that contains ring and disk bubbles. (b) The extracted circles are highlighted with white color. There are a significant number of extracted circles that are not related to the oceanic bubbles. The peaks in the accumulator space that correspond to these extra circles are due to bubble wall thickness and bubble wall background.

approach where votes are cast in an accumulator array that is parameterized according to the model of the shape to be extracted, and it can be shown to be an optimal form of template matching.

The Hough transform operates on the principle that candidate edge points have high intensities. In practice this means that an edge detector is often required as a preprocessing stage to transform object boundaries (Ballard 1981; Deane and Stokes 1999; Kimme et al. 1975; Zheng et al. 2004). However, because of the formation of the images from the bubble camera, it is possible to avoid this step as the edges are already in this form. This is advantageous because the gradient-based edge operators can amplify noise and skew peaks off center in the accumulator (Gonzalez and Woods 2008; Nixon and Aguado 2002).

1) ACCUMULATOR SPACE

The equation of circle is given by

$$(x - x_0)^2 + (y - y_0)^2 = r^2, \tag{1}$$

where the point (x_0, y_0) represents the circle center, points (x, y) correspond to the circle locus, and r represents the circle radius. A bright pixel in the image is a candidate for a point on the locus of a number of possible circles. The algorithm votes by incrementing the accumulator for those values of the parameters (x_0, y_0, r) that satisfy Eq. (1) given the pixel coordinates (x, y).

This corresponds to a cone in the three-dimensional accumulator space parameterized by (x_0, y_0, r). The votes from all sets of edge points of a circle in the image will pass through the same point in the accumulator space. Thus, this point has the maximum vote (peak) in the accumulator space and can be used to extract circle parameters. Since higher intensities indicate greater confidence in edge points, this can be used to weight the vote in the accumulator space,

$$A(x_0, y_0, r) + = g(x, y), \tag{2}$$

where A is the accumulator value at the coordinates (x_0, y_0, r) and g is the pixel intensity at coordinates (x, y). For each value of r, the accumulator coordinates (x_0, y_0) are calculated using the parametric form of Eq. (1),

$$x_0 = x - r\cos\theta, \quad y_0 = y - r\sin\theta, \tag{3}$$

where $\theta \in [0, 2\pi)$. The spatial resolution of the coordinates (x, y, r) in the accumulator space is one pixel, and θ is quantized to 1°.

2) PEAK DETECTION

Promising circles are indicated by large values in the accumulator array. To avoid multiple circles being detected where only one exists in the image, it is necessary to do peak detection. The accumulator array is searched for local maxima by selecting the accumulator cells that have

higher votes than their 3×3 spatial neighborhood. However, there are many additional peaks in the accumulator array that are not relevant to the circular bubbles. For instance, peaks may be created by the surrounding bubble walls, the bubble walls and background, and the marine creatures. These are similar to the bubble peaks. Moreover, the bubble wall can generate many peaks due to its thickness. Figure 4 shows extracted circles in an image that corresponds to these redundant peaks. Thus, it is very important to filter out these unwanted peaks.

The peak filtering can be divided into six stages: radial distribution (RD), suppressing wall thickness peaks, filtering surrounding bubble walls or wall background peaks, filtering edge peaks, and filtering copepod peaks. The radial distribution measures the homogenous distribution of the pixel intensities around the bubble center in the image space. It can be expressed as

$$RD = \frac{\left| \int_{\theta=0}^{2\pi} e^{i\theta} g(x,y) \, d\theta \right|}{\int_{\theta=0}^{2\pi} g(x,y) \, d\theta}. \qquad (4)$$

The pixel coordinates (x, y) are calculated using Eq. (2) for each candidate peak in the accumulator space. To implement this measure, the RD value calculated in Eq. (4) needs to be compared with an absolute value. It is found by evaluating a number of images that the RD values for the bubble peaks are less than 0.4.

The RD is very efficient in removing the peaks between the bubble walls and background. In a highly dense bubble image, approximately 60%–75% of the unwanted peaks are removed by the RD measure. Nevertheless, it is not effective in filtering the wall thickness and the surrounding walls peaks. Before proceeding to the next refining stages, the peaks on the candidate list are sorted in descending order according to their accumulator votes and radii. Consequently, the larger circles are selected first from the list in the next stages. It was decided to follow this hierarchical approach that concentrates first on the larger circles because many false smaller circles exist inside the bubble, on the bubble wall, and between the bubble wall and the background. The wall thickness peaks can be identified by comparing the distance between each candidate peak with all the peaks in the accumulator array. The distance between two circles with peak coordinates (x_1, y_1, r_1) and (x_2, y_2, r_2) can be calculated using the following formula:

$$\delta = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} - (r_1 + r_2), \qquad (5)$$

where δ is positive if the circles do not overlap and it measures the minimum distance between the circle loci.

Equation (5) was used to remove the entire set of overlapping peaks that belong to a particular bubble. This set of overlapping peaks includes the wall thickness peaks as well as the small circle peaks on the bubble wall and between the bubble wall and background. These redundant peaks are removed from the candidate list by following the hierarchical approach mentioned in the previous paragraph. The larger peaks are first selected from the candidate list. Then their distances are compared with the smaller peaks on the list using Eq. (5). It was found by evaluating a number of images that $\delta \leq 1$ is adequate to detect the entire set of smaller peaks that overlap with the bubble peak. Consequently, these overlapping peaks are removed from the list.

The surrounding wall peaks are generated by the walls of adjacent bubbles. It is observed that some parts of the background are covered by the surrounding wall circles in the image space. Therefore, the detection of the surrounding wall peaks is based on counting the number of pixels at the circumference in the image space that have intensities less than double the background intensity.

Even though copepods are not circular, many connected circles are generated inside the copepod bodies due to their high pixel intensities. The copepod and bubble peaks in the accumulator space can be very similar and therefore it is challenging to distinguish between them. Two criteria have been used to identify the copepod peaks. These are the peak value in the accumulator space, and the distance between this peak and its neighbors. The copepod peak is the one with the highest vote and has a close distance with at least three neighboring peaks. If these two conditions are satisfied, then the distance between the connected circles in this region is evaluated using Eq. (5). The algorithm continues evaluating the distance between each new detected copepod circle and its neighbors when satisfying the value of δ in Eq. (5). Two types of images were analyzed to select the right value of δ. The first type of image contains copepods with different shapes and sizes, while the second type contains only bubbles. It was found that $\delta \leq 12$ in Eq. (5) is adequate to detect the entire set of peaks that belong to a particular copepod. Moreover, it would not detect the bubbles in the copepod-free images. These copepod peaks are removed from the candidate list.

The bubbles that are close to the image edges appear as incomplete circles. Therefore, it is not possible to extract them with the same accuracy. The peaks of theses incomplete bubbles are excluded from the candidate list by testing how close their coordinates are from the edge of the accumulator array, and the reduction in effective field of view for large bubbles

can be compensated appropriately in any final histogram calculation.

3) CLASSIFICATION

After refining the unwanted peaks, the remaining peaks in the accumulator space are classified into two categories: rings (bright rings around dark center) and disks (filled circles). The illumination around the ring bubble wall varies because the light source does not illuminate the light sheet evenly from all directions, but it has an increased intensity around the centers of the mirrors creating the light sheet. As a consequence, the ring bubbles appear as four dots for the smallest resolvable bubbles in the image space as shown in Fig. 4.

The coordinates and radii of the ring and disk bubbles are successfully detected by the hierarchical approach of radii extraction that was described in the previous section [section 4a(2)]. However, the bubble rings have a more precise bubble radii distribution than the disks. This is because the circular bubble rings are in the light sheet (in focus) and are very well identified. The ring center intensity is close to the background intensity and is lower than its circumference intensities. In contrast, the filled disks (out of focus) are not in the light sheet and the pixel intensities are irregularly distributed around their centers. Therefore, these out of focus bubbles are excluded from the bubble size distribution calculations.

The classification into rings and disks is achieved by comparing the average intensity of the central bubble region with that of its periphery.

4) IMPLEMENTATION

To implement Eqs. (1)–(3), it is necessary to specify the size and resolution of the accumulator. To speed up the image analysis, the maximum bubble radius needs to be known before running the algorithm. To find this radius, the average intensity of each single image in two major deployments (8.5 h each) was calculated. It was found that the maximum bubble radii in high-intensity images did not exceed 20 pixels (approximately $400\,\mu$m). Therefore, the radii range used in the algorithm is from 1 to 25 pixels (approximately 20–406 μm). The accumulator coordinates (x_0, y_0) have the same range as the image space (2048×2048) and the resolution for the coordinates (x_0, y_0) and r is one pixel. The implementation of the algorithm can be summarized as follows.

1) Calculate the image histogram to assess the image type (small bubbles, big bubbles, or complex image). This is accomplished by examining the number of pixels at a particular gray level in the image histogram. The algorithm continues running the next steps (2–11) only if the image contains small bubbles.

2) Build the three-dimensional accumulator array $A(x_0, y_0, r)$ using the voting scheme in Eq. (2). For each value of r, there are 2048 rows and 2048 columns of the accumulator coordinates (x_0, y_0). This array can be built by varying the coordinates of the circle locus (x, y) in Eq. (3) from 0 to 2047 for a particular value of r. The value of θ in Eq. (3) is varied from 0 to 2π for a given circle locus (x, y). The vote of the accumulator array $A(x_0, y_0, r)$ in Eq. (2) is incremented only for the coordinates (x_0, y_0) that lie within the image dimension.

3) Find the peaks in the accumulator array by selecting the accumulator cells that have higher votes than their 3×3 spatial neighborhood. This can be accomplished by comparing the vote of each accumulator cell $A(x_0, y_0, r)$ with its neighbors. A list of peaks is created in this step. This list contains the accumulator votes and coordinates that have larger votes than its neighborhood.

4) Refine the peaks list using the radial distribution measure in Eq. (4). The circle loci for each accumulator coordinate (x_0, y_0, r) on the list of peaks are calculated by varying θ in Eq. (3) from 0 to 2π. The intensity values of these locus coordinates in the image space are summed to calculate the RD in Eq. (4). The accumulator vote and coordinate for a given peak are removed from the list if the RD value is larger than 0.4.

5) Sort the peaks in descending order according to their radii and votes so that priority is given to the larger bubbles as shown in section 4a(2). This can be done by calculating the product of the accumulator vote and the radius of each peak and comparing its values with all the peaks on the list.

6) Refine the list by removing the surrounding walls and wall background peaks. The circle loci for each accumulator coordinate (x_0, y_0, r) on the list of peaks are calculated by varying θ in Eq. (3) from 0 to 2π. A counter is created and incremented by 1 if the intensity value for a given locus coordinate in the image space is less than double the background intensity. The peak is discarded from the list if the counter value is larger than 180.

7) Refine the list by removing the wall thickness peaks using Eq. (5). The distance between peak coordinates (x_1, y_1, r_1) and (x_2, y_2, r_2) on the list are calculated and the peaks of smaller circles (lower votes) are removed from the list if $(x_1 - x_2)^2 + (y_1 - y_2)^2 \leq (r_1 + r_2 + 1)^2$.

8) Refine the list by removing the peaks at the edges. For a given accumulator coordinate (x_0, y_0, r), its

TABLE 1. Comparison between the errors in the coordinate $\epsilon_{x,y}$ and radius ϵ_r using different spd values (blurring) in Eq. (6).

Spread (pixels)	Max $\epsilon_{x,y}$	Max ϵ_r	$\epsilon_{x,y}$ percentage in 10 images (%)	ϵ_r percentage in 10 images (%)
1	0	0	0	0
2	0	0	0	0
3	2	1	0.3	15
4	8	7	1.17	16.1
5	8	8	2.52	14.53
6	9	9	4.99	25.48

peak is discarded if $(x_0 - r) < 0$ or $(x_0 + r) > 2048$ or $(y_0 - r) < 0$ or $(y_0 + r) > 2048$.

9) Refine the list by removing the peaks caused by copepods. The distance between the coordinate with the highest vote peak (x_1, y_1, r_1) and other peak coordinates (x_2, y_2, r_2) on the list are calculated using Eq. (5). A counter is created and incremented by 1 if $(x_1 - x_2)^2 + (y_1 - y_2)^2 \leq (r_1 + r_2 + 12)^2$. The highest vote peak and its neighbors will be regarded as copepod peaks if the counter value is equal or larger than 3. In this case, the algorithm continues evaluating the distance between theses detected copepod peaks and its neighbors. The neighborhood peaks will also be discarded from the list if $(x_1 - x_2)^2 + (y_1 - y_2)^2 \leq (r_1 + r_2 + 12)^2$.

10) Classify the peaks as rings or disks. For a given accumulator coordinate (x_0, y_0) on the list, the locus coordinates of two circles with radii r and $0.1r$ are calculated by varying θ in Eq. (3) from 0 to 2π. The average intensities of these two circle loci in the image space are calculated and compared with the circle center intensity.

11) Calculate the histogram of the bubble rings and disks, and save them in two different files.

12) Load the next image in the directory and repeat steps 1–11.

Steps 2–10 are explained in more detail in section 4a(2).

b. Evaluation

To demonstrate the performance, the Hough transform was evaluated using synthetic and real images. To simulate the bubble images, the following model was used to generate bubbles:

$$\text{psf} = 255\left(1 - \frac{\text{dist}}{\text{spd}}\right)\left[\frac{1}{2} + \frac{r}{2\,\text{rmax}}\cos\left(4\tan^{-1}\frac{y}{x}\right)\right], \quad (6)$$

where psf is the point spread function that describes the blurring around a particular circle locus (x, y), spd represents the blurring of the bubble wall, dist represents the distance between the circular bubble locus (x, y) and its center, r represents the bubble radius, rmax represents the maximum radius in the synthetic image, and the cosine expression simulates the variation in illumination around the bubble due to the nonuniformity of the four light sources forming the light sheet; the inverse tangent is computed using atan2 to retain sign information. The spd value in Eq. (6) was gradually increased from 1 to 6 to expand the blurring of the bubble walls and to reduce the distance between them. Moreover, the bubble radii were in the range 1–50 pixels and randomly positioned in the images. The resolution of these images was 2048×2048 pixels and the number of bubbles was between 400 and 700. The total number of synthetic images was 60, where 10 images were tested for each spread value. The absolute errors in the coordinate $\epsilon_{x,y}$ and radius ϵ_r are calculated as follows:

$$\epsilon_{x,y} = |D_{x,y} - T_{x,y}|, \quad (7)$$
$$\epsilon_r = |D_r - T_r|,$$

where D and T stand for the detected and true circles, respectively.

Table 1 shows the maximum and the average errors in the coordinates and radii obtained from increasing the spd value from 1 to 6. There is a gradual increase in the maximum and average errors when this value exceeds 2. In addition, the average errors in the radii are much greater than the coordinates. It has been found that the majority of these radii errors result from bubbles with radii smaller than 3 pixels and that the absolute error in the radius is ± 1 pixel. This is because the shapes of these small bubbles change from rings to disks in response to the increase in the spd value. Figure 5 shows the effect of increasing the spd value from 1 to 6 on the bubble shapes and distance between them. It is clear that the pixel intensities forming the bubble walls vary substantially and that a significant number of the bubbles are overlapping when the spread is 6. However, the algorithm correctly extracts these overlapped bubbles as shown in Fig. 5d. In addition, Fig. 5d shows that the algorithm does not detect the incomplete bubbles at the image edge.

The algorithm was evaluated using 80 real images. The selected images contain a large number of ring and disk bubbles and copepods as shown in Fig. 6. The figure shows that the algorithm extracts the majority of the ring
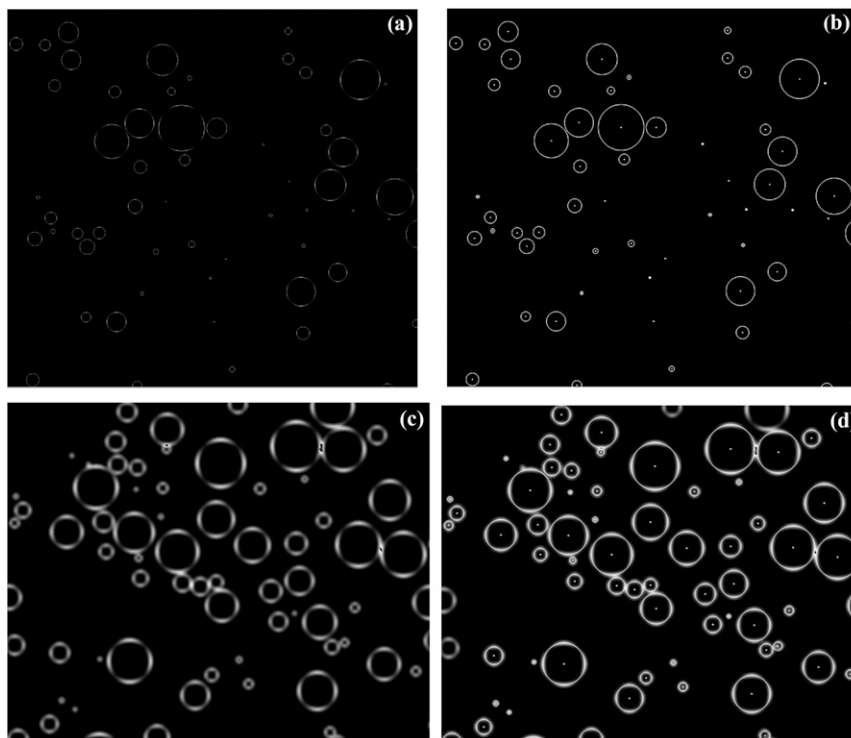
FIG. 5. Subregions of the synthetic images to illustrate the change in spread values in Table 1. (a) The spread values are equal to 1. (b) The extracted bubbles in (a). The extracted bubbles are highlighted with a white dot in the center and a white circumference line. (c) The spread value is equal to 6. (d) The detected bubbles in (c). This region is taken from the top of an image and therefore the incomplete bubbles at the edges are not extracted.

and disk bubbles in these images without counting the copepod.

5. Bubble size distributions

The bubble extraction algorithm was applied to the entire sequence of images collected in two main deployments of the bubble imaging instrument. The bubble imaging instrument was configured to capture images for 40 min every 3 h. The total number of processed images was approximately 850 000. Figure 7 shows a sample of the data collected during deployment in the North Atlantic Ocean on 25 October 2013. The wind speed was $26 \, \mathrm{m \, s^{-1}}$. The void fraction versus time of recording is shown in Fig. 7a. The 1-s averages were calculated based on the 15 sequential frames taken in that 1 s, and this average was calculated for every second throughout the 40-min measurement period. A subsection identified by the vertical dashed line in Fig. 7a is plotted in more detail in Fig. 7b. The circle (16:18:33), diamond (16:19:45), and triangle (16:19:51) in Fig. 7b are the markers for the three bubble size distributions plotted in detail in Fig. 7c.

Each distribution was scaled by the measured volume of water, which was approximately $4.0 \times 4.0 \times 0.5 \, \mathrm{cm^3}$. The bubble numbers quoted here are the bubble number per micron radius increment per unit volume, which is the conventional unit used in the oceanography literature. Each size distribution shown is also a 1-s average, and the total number of bubbles counted to calculate each 1-s size distribution was 28 (circles), 542 (diamonds), and 1225 (triangles).

6. Discussion

The focus of this paper was the automated extraction of bubble images using the Hough transform as the basis. Most of the methods based on the Hough transform for circular shape detection use gradient information that is obtained from applying a first-order edge detector to the original image. Therefore, the success of these methods significantly depends on accurate estimates of the edge information. In our approach, the pixel intensities are directly used to build the accumulator array for the voting scheme. This was followed by several stages of filtering and a hierarchical radial extraction
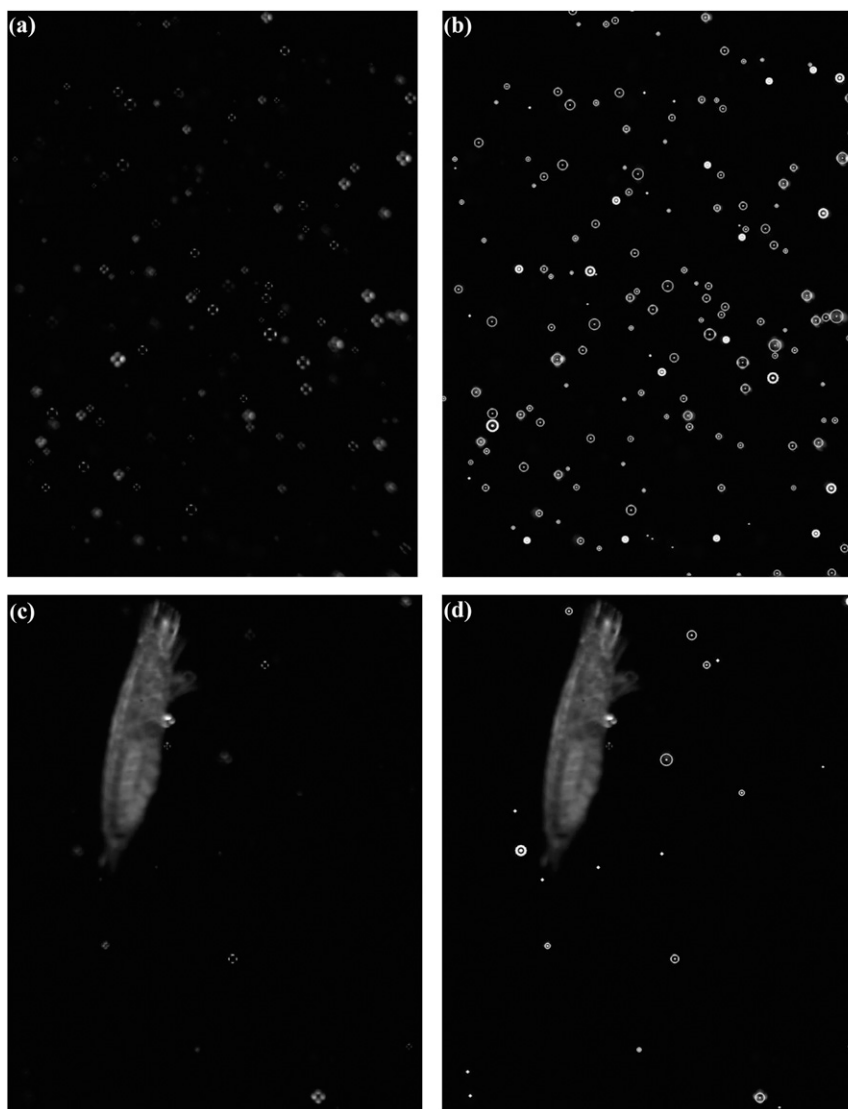
FIG. 6. Subregions of the real images illustrating the algorithm extraction accuracy. (a) Ring and disk bubbles. (b) The extracted ring and disk bubbles in (a). The extracted bubbles are highlighted with a white dot in the center and a white circumference line. The brighter white color was used to identify disk bubbles. (c) A copepod , and disk and ring bubbles. (d) The extracted bubbles in (c), successfully avoiding false detection associated with the copepod.

approach to remove unwanted peaks. This improves the detection accuracy of the Hough transform by approximately 50% (Liu and Fang 2015).

The proposed Hough transform has been shown to be successful in extracting bubbles in synthetic and real images as mentioned in section 4b. The test images contain a large number of ring and disk bubbles that overlap in some cases. In addition, the size of the bubble radii was between 1 and 50 pixels (approximately 20–1000 μm). The unwanted peaks caused by the overlapping bubbles have been effectively removed by the hierarchical radial extraction approach that first emphasizes the larger particles. Consequently, this hierarchical approach gives a bias to detect larger bubbles. However, this bias can be corrected and included as error bars on the bubble size distributions. The accuracy of extracting ring bubbles is much higher than the corresponding disk bubbles. This is because the bubbles in the light sheet are in focus and are seen as a white ring surrounding dark centers. Nevertheless, the unfocused bubbles are observed as disks with an irregular distribution of pixel intensities around the center. Therefore, the histograms of the disk bubbles were saved in another file to separate them from the ring bubbles.
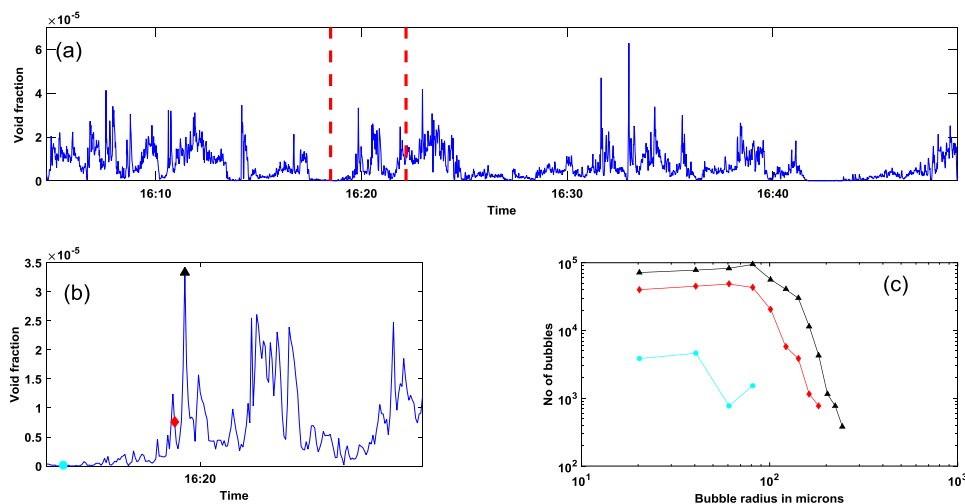
FIG. 7. Sample of the data processed by the bubble extraction algorithm. (a) The void fraction with time for one 45-min measurement period. (b) An enlarged section of the void fraction plot. (c) The detailed bubble size distributions at the three marked times.

The entire image sequence for two major deployments was analyzed, and a sample of the bubble size distributions measured during an 18-s period is shown in Fig. 7. There is significant change in the bubble numbers during these short periods. This is likely due to a combination of highly inhomogeneous bubble plumes and the advection of the bubbles past the optical instrument. In general, the number of bubbles is significantly lower when their size exceeds $100\,\mu$m. Moreover, the smallest extracted bubbles were approximately $20\,\mu$m in radius, which is equivalent to the Rayleigh resolution limit for such types of oceanic bubble imaging instruments (Deane and Stokes 1999; Walsh and Mulhearn 1987).

The required computation time and memory for the Hough transform depends on the total number and size of the bubbles in the image, the discrete resolution of the accumulator parameters (radii and centers), and the possible range of these parameters. The proposed algorithm was coded in C++ and used the OpenCv library for computational efficiency (Bradski and Kaehler 2008). The C++ program requires approximately 5 s to extract one frame on a 2.5-GHz Core i5 Mac laptop with 4-GB system memory. It is mainly steps 2 and 3 of implementation (feature extraction stage) that are time consuming.

Although the proposed Hough transform is very robust to extract and count the bubbles, the algorithm's accuracy significantly decreases in particular images. For instance, the bubbles can be large and nonspherical, or tightly packed and highly illuminated as shown in Figs. 3a and 3d, respectively. Bubbles with radii greater than approximately 1 mm are likely to show a significant distortion from a spherical shape. An advanced

approach based on a combination of machine learning and Hough transform may be required to process such complex images.

The proposed algorithm in this paper can be extended to extract elliptical bubbles. This can be done by defining five parameters that represent an ellipse, instead of three parameters that represent a circle. As a result, the algorithm becomes more complicated and slower since it requires significantly greater computational recourses. The complex and elliptical big bubble images were detected and not processed by inspecting their image histograms as illustrated in the implementation section [section 4a(4)]. The percentage of these images was less than approximately 0.04%.

7. Conclusions

We have presented an automated algorithm for bubble extraction based on the Hough transform. The algorithm effectively resolves bubbles with a radius of one pixel and discriminates between bubbles and copepods. It was applied to analyze approximately 850 000 captured images from the ocean. The main limitation of this algorithm is that it processes circular bubbles only and excludes noncircular and significantly overlapped bubbles in complex images.

The main novelty of this automated algorithm is that it extracts the bubbles from the original image by using the pixel intensities directly without applying any preprocessing operations. Therefore, the algorithm is less sensitive to noise because it does not use any first-order edge operators. In addition, it is not susceptible to the background changes in illumination and effectively extracts bubble sizes as small as one pixel in radius. This is

because it avoids using any prefiltering and thresholding operations. The implementation of this automated algorithm is significantly simpler than the ones published in literature and it has been applied to process 850 000 real ocean images. In addition, we have found that this algorithm performs well for realistic ocean bubble distributions and removal of copepods. Real data from the ocean were used to refine the algorithm, in contrast to methods that are calibrated solely using laboratory data collected in controlled conditions.

We believe that the proposed algorithm can be used to extract circular bubbles in many other engineering, medical, and chemical applications.

REFERENCES

Abdou, I. E., and W. K. Pratt, 1979: Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proc. IEEE*, **67**, 753–763, doi:10.1109/PROC.1979.11325.

Ainslie, M. A., 2005: Effect of wind-generated bubbles on fixed range acoustic attenuation in shallow water at 1–4 kHz. *J. Acoust. Soc. Amer.*, **118**, 3513–3523, doi:10.1121/1.2114527.

Ballard, D. H., 1981: Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognit.*, **13**, 111–122, doi:10.1016/0031-3203(81)90009-1.

Boelmann, J., and O. Zielinski, 2015: Automated characterization and quantification of hydrocarbon seeps based on frontal illuminated video observations. *J. Eur. Opt. Soc. Rapid Publ.*, **10**, doi:10.2971/jeos.2015.15018.

Bradski, G., and A. Kaehler, 2008: *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, Inc., 580 pp.

Czerski, H., M. Twardowski, X. Zhang, and S. Vagle, 2011: Resolving size distributions of bubbles with radii less than 30 μm with optical and acoustical methods. *J. Geophys. Res.*, **116**, C00H11, doi:10.1029/2011JC007177.

Deane, G. B., and M. D. Stokes, 1999: Air entrainment processes and bubble size distributions in the surf zone. *J. Phys. Oceanogr.*, **29**, 1393–1403, doi:10.1175/1520-0485(1999)029<1393:AEPABS>2.0.CO;2.

——, and ——, 2002: Scale dependence of bubble creation mechanisms in breaking waves. *Nature*, **418**, 839–844, doi:10.1038/nature00967.

Duda, R. O., and E. Hart, 1972: Use of the Hough transform to detect lines and curves in pictures. *Commun. ACM*, **15**, 11–15, doi:10.1145/361237.361242.

Farmer, D. M., C. L. McNeil, and B. D. Johnson, 1993: Evidence for the importance of bubbles in increasing air–sea gas flux. *Nature*, **361**, 620–623, doi:10.1038/361620a0.

Frouin, R., S. Iacobellis, and P.-Y. Deschamps, 2001: Influence of oceanic whitecaps on the global radiation budget. *Geophys. Res. Lett.*, **28**, 1523–1526, doi:10.1029/2000GL012657.

Fuentes, E., H. Coe, D. Green, G. de Leeuw, and G. McFiggans, 2010: Laboratory-generated primary marine aerosol via bubble-bursting and atomization. *Atmos. Meas. Tech.*, **3**, 141–162, doi:10.5194/amt-3-141-2010.

Geißler, P., and B. Jähne, 1995: Measurements of bubble size distributions with an optical technique based on depth from focus. *Air-Water Gas Transfer: Selected Papers from the Third International Symposium on Air-Water Gas Transfer*, B. Jähne and E. C. Monahan, Eds., AEON Verlag and Studio, 351–362.

Gonzalez, R. C., and R. E. Woods, 2008: *Digital Image Processing*. 3rd ed. Prentice Hall, 954 pp.

Honkanen, M., P. Saarenrinne, T. Stoor, and J. Niinimäki, 2005: Recognition of highly overlapping ellipse-like bubble images. *Meas. Sci. Technol.*, **16**, 1760, doi:10.1088/0957-0233/16/9/007.

Hough, P. V. C., 1960: Method and means for recognizing complex patterns. U.S. Patent 3,069,654, filed 25 March 1960, and issued 18 December 1962.

Jähne, B., and P. Geißler, 1995: An imaging optical technique for bubble measurements. *Sea Surface Sound '94: Proceedings of the III International Meeting on Natural Physical Processes Related to Sea Surface Sound*, M. J. Buckingham and J. R. Potter, Eds., World Scientific, 290–303.

Kimme, C., D. Ballard, and J. Sklansky, 1975: Finding circles by an array of accumulators. *Commun. ACM*, **18**, 120–122, doi:10.1145/360666.360677.

Lee, S. U., and S. Y. Chung, 1990: A comparative performance study of several global thresholding techniques for segmentation. *Comput. Vision Graphics Image Process.*, **52**, 171–190, doi:10.1016/0734-189X(90)90053-X.

Leifer, I., G. D. Leeuw, and L. H. Cohen, 2003: Optical measurement of bubbles: system design and application. *J. Atmos. Oceanic Technol.*, **20**, 1317–1332, doi:10.1175/1520-0426(2003)020<1317:OMOBSD>2.0.CO;2.

Leighton, T. G., 1994: *The Acoustic Bubble*. Academic Press Limited, 613 pp.

Liu, X., and S. Fang, 2015: A convenient and robust edge detection method based on ant colony optimization. *Opt. Commun.*, **353**, 147–157, doi:10.1016/j.optcom.2015.05.019.

Medwin, H., 1970: In situ acoustic measurements of bubble populations in coastal ocean waters. *J. Geophys. Res.*, **75**, 599–611, doi:10.1029/JC075i003p00599.

——, 1977: In situ acoustic measurements of microbubbles at sea. *J. Geophys. Res.*, **82**, 971–976, doi:10.1029/JC082i006p00971.

Nixon, M. S., and A. S. Aguado, 2002: *Feature Extraction and Image Processing for Computer Vision*. Newnes, 350 pp.

Pascal, R. W., and Coauthors, 2011: A spar buoy for high-frequency wave measurements and detection of wave breaking in the open ocean. *J. Atmos. Oceanic Technol.*, **28**, 590–605, doi:10.1175/2010JTECHO764.1.

Stokes, M. D., and G. B. Deane, 1999: A new optical instrument for the study of bubbles at high void fractions within breaking waves. *IEEE J. Oceanic Eng.*, **24**, 300–311, doi:10.1109/48.775292.

Stramski, D., 1994: Gas microbubbles: An assessment of their significance to light scattering in quiescent seas. *Ocean Optics XII*, J. S. Jaffe, Ed., International Society for Optical Engineering (SPIE Proceedings, Vol. 2258), 704–710, doi:10.1117/12.190117.

——, and J. Tegowski, 2001: Effects of intermittent entrainment of air bubbles by breaking wind waves on ocean reflectance and underwater light field. *J. Geophys. Res.*, **106**, 31 345–31 360, doi:10.1029/2000JC000461.

Taboada, B., L. Vega-Alvarado, M. S. Córdova-Aguilar, E. Galindo, and G. Corkidi, 2006: Semi-automatic image analysis methodology for the segmentation of bubbles and drops in complex dispersions occurring in bioreactors. *Exp. Fluids*, **41**, 383–392, doi:10.1007/s00348-006-0159-0.

Terrill, E. J., and W. K. Melville, 2000: A broadband acoustic technique for measuring bubble size distributions: Laboratory and shallow water measurements. *J. Atmos. Oceanic Technol.*, **17**, 220–239, doi:10.1175/1520-0426(2000)017<0220: ABATFM>2.0.CO;2.

Thomanek, K., O. Zielinski, H. Sahling, and G. Bohrmann, 2010: Automated gas bubble imaging at sea floor—A new method of in situ gas flux quantification. *Ocean Sci.*, **6**, 549–562, doi:10.5194/os-6-549-2010.

Walsh, A. L., and P. J. Mulhearn, 1987: Photographic measurements of bubble populations from breaking wind waves at sea. *J. Geophys. Res.*, **92**, 14 533–14 565, doi:10.1029/JC092iC13p14553.

Wang, B., and S. A. Socolofsky, 2015: A deep-sea, high-speed, stereoscopic imaging system for in situ measurement of natural seep bubble and droplet characteristics. *Deep-Sea Res. I*, **104**, 134–148, doi:10.1016/j.dsr.2015.08.001.

Wang, Q., and E. C. Monahan, 1995: The influence of salinity on the spectra of bubbles formed in breaking wave simulations. *Sea Surface Sound '94: Proceedings of the III International Meeting on Natural Physical Processes Related to Sea Surface Sound*, M. J. Buckingham and J. R. Potter, Eds., World Scientific, 312–319.

Wanninkhof, R., W. E. Asher, D. T. Ho, C. Sweeney, and W. R. McGillis, 2009: Advances in quantifying air-sea gas exchange and environmental forcing. *Annu. Rev. Mar. Sci.*, **1**, 213–244, doi:10.1146/annurev.marine.010908.163742.

Zhang, X., M. Lewis, and B. Johnson, 1998: Influence of bubbles on the scattering of light in the ocean. *Appl. Opt.*, **37**, 6525–6536, doi:10.1364/AO.37.006525.

Zheng, Y., M. S. Nixon, and R. Allen, 2004: Automated segmentation of lumbar vertebrae in digital videofluoroscopic images. *IEEE Trans. Med. Imaging*, **23**, 45–52, doi:10.1109/TMI.2003.819927.

Zhong, S., X. Zou, Z. Zhang, and H. Tian, 2016: A flexible image analysis method for measuring bubble parameters. *Chem. Eng. Sci.*, **141**, 143–153, doi:10.1016/j.ces.2015.10.033.

Zhou, J., K. Mopper, and U. Passow, 1998: The role of surface-active carbohydrates in the formation of transparent exopolymer particles by bubble adsorption of seawater. *Limnol. Oceanogr.*, **8**, 1860–1871, doi:10.4319/lo.1998.43.8.1860.

Zielinski, O., B. Saworski, and J. Schulz, 2010: Marine bubble detection using optical-flow techniques. *J. Eur. Opt. Soc. Rapid Publ.*, **5**, doi:10.2971/jeos.2010.10016s.