UC San Diego
UC San Diego Previously Published Works

Title

ZOOPS- O2: A broadband echosounder with coordinated stereo optical imaging for observing plankton in situ

Permalink

Authors

Briseño-Avena, Christian
Roberts, Paul LD
Franks, Peter JS
et al.

Publication Date

DOI

Peer reviewed
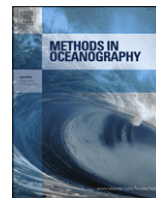
Full length article

ZOOPS-O^2: A broadband echosounder with coordinated stereo optical imaging for observing plankton *in situ*

CrossMark

Christian Briseño-Avena [*], Paul L.D. Roberts, Peter J.S. Franks, Jules S. Jaffe

Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

A B S T R A C T

Here we describe the configuration, calibration, and initial results from the combination of two recently developed underwater instruments that measure acoustic reflectivity and, simultaneously, the location, pose and size of millimeter-sized plankton relative to the sonar beam. The acoustic system, ZOOPS (ZOOPlankton Sonar), uses a broadband chirp signal that operates with a single monostatically configured transducer in the 1.5–2.5 MHz frequency range. We demonstrate that the system can record, with adequate signal-to-noise levels, identifiable reflections from single copepods with lengths as small as 360 μm. To simultaneously identify taxa and measure orientation, a pair of "O-Cam" microscopes were stereoscopically calibrated and geometrically co-registered with the orientation and range-resolved acoustic transmissions of the sonar beam. The system's capability is demonstrated via the *in situ* measurement of acoustic reflectivity as a function of orientation for 224 individual pelagic copepods comprising three orders of free-living taxa. Comparison with a well-known model, the Distorted Wave Born Approximation (DWBA), using a spheroidal formulation, yields both differences and similarities between the *in situ* field data and the model's predictions.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

---

* Corresponding author.
  E-mail address: cbriseno@ucsd.edu (C. Briseño-Avena).

1. Introduction

As acoustic systems continue to be employed routinely in underwater monitoring, there is strong interest in the refinement of methods for *in situ* sensing of various pelagic and benthic features of the environment. This is especially true given the continued and increasing human exploitation – and in many cases depletion – of marine resources. Simultaneously, as the effects of climate change become more apparent there is a need for improved tools to assess the changes.

In the ocean, acoustic methods have many advantages over other potential remote sensing approaches: the low absorption and frequency-dependent properties of sound can be used to acquire information about many important aspects of the submarine world that are invisible to other methodologies. Acoustic methods have been prominent in, for example, the management of fish stocks and the survey of the sea floor for oil and gas exploration or the retrieval of lost objects such as jetliners.

Because of the importance of plankton in underwater ecology, a variety of remote sensing methods to measure abundance as a function of taxonomic category, including optical and acoustic techniques, have been under development for quite some time (e.g., Fernandes et al., 2002; ICES, 2011, 2012, 2013, 2014). To interpret acoustic field data it is important to have either direct empirical observations from known organisms, or a model that can be inverted to obtain the desired results. Ideally, the model would be derived from physical principles based on the organism's morphological and acoustic properties, and applied in concert with *a priori* knowledge of the planktonic (or nektonic) species assemblage known to occur in a given ecosystem.

The end product of a typical field program to verify the sources of acoustic backscatter is a set of observed size-class distributions of dominant taxa, a practice oftentimes referred to as "ground truthing". The methods to collect the data include nets (Sutor et al., 2005; Lavery et al., 2007, 2010; Powell and Ohman, 2012), pumps (Costello et al., 1989; Pieper et al., 1990), and/or optical systems (Benfield et al., 1998; Wiebe et al., 1996; Lavery et al., 2007). Acoustic backscatter data has also been collected in the lab, though this has been limited to a few taxa such as euphausiids, pteropods and physonect (gas-bearing) siphonophores that can be collected without damaging the organisms (Stanton et al., 1996, 1998a,b, 2004).

In addition to using the more traditional volume backscatter systems that were primarily developed for fisheries and bottom surveys, several systems have been formulated specifically for plankton. One such system, developed by Van Holliday and co-workers over a number of decades had a variety of realizations; all were based on the use of a number of narrow-band frequencies spanning a given frequency range. Early versions of the system used as many as 21 frequencies (Holliday and Pieper, 1980; Pieper et al., 1990) while the last versions (TAPS) mostly used four. Building on this work, the BIOMAPER system, as developed by Wiebe and colleagues (Greene et al., 1989; Wiebe et al., 1990, 1996), employed 5 frequencies (43, 120, 200, 420, and 1000 kHz). This system was deployed extensively; however, the complexity of the backscattered waveforms, the complications due to orientation dependence, and the strong dependence on rare but extremely strong scatterers (Stanton et al., 1996; Wiebe et al., 1996), and the resultant ambiguities has restricted the interpretation of the data in many applications.

One system that used concurrent optic and narrowband acoustic technologies to identify individual targets producing a given *in situ* echo was the FishTV multibeam system (Jaffe et al., 1998). Here, individual acoustic signals of relatively large mesozooplankton (gammarid amphipods, euphausiids: 5.3–20 mm) and micronekton (gadid fish: 36–92 mm) were measured. A later version of the system, FishTV Jr. was used in the Red Sea, without optical technology, to test the hypothesis that zooplankton, likely copepods, were holding depth in the face of vertical currents (Genin et al., 2005).

A more recent advance in acoustic echosounders is the use of spectrally continuous, broadband systems (Stanton, 2012; Fornshell and Tesei, 2013). Such systems result in higher range resolution (Chu and Stanton, 1998; Stanton and Chu, 2008; Stanton, 2012), which increases the ability to observe echoes from individual scatterers, allowing echo counting. Furthermore, continuous spectral information can provide an advantage in taxonomic discrimination (Roberts and Jaffe, 2007, 2008). Recent broadband systems of Ross and Lawson (2009) used 85–155 kHz sound and that of Lavery et al. (2010) used frequencies of 50–600 kHz to study copepods, pteropods, amphipods, and euphausiids with sizes between 2 mm and close to 3 cm.

Since the most abundant net mesozooplankton in pelagic habitats are <5 mm (e.g., Gallienne and Robins, 2001; San Martin et al., 2006; Thompson et al., 2013) there is clearly a need to measure small organisms. In pursuit of this goal, and building on work of previous investigators, we invented the ZOOPS-O^2 (ZOOPlankton Sonar with stereo Optical imaging) system. ZOOPS-O^2 is designed to obtain broadband acoustic reflections *in situ* from small (360 μm to <2 cm), individual plankton with concurrent stereoscopic images of the organisms. The instrument combines two stand-alone technologies – acoustic and optic – to unambiguously assign the acoustic echo of individual living and non-living particles to their simultaneous stereoscopic images. The system employs a broadband, ultrasonic transducer (1.5–2.5 MHz) with close to millimeter range resolution and a stereoscopic camera system to image organisms. ZOOPS-O^2 is self-contained, and can be profiled vertically or towed. To date, the system has been used to measure thousands of acoustic reflections with concurrent optical identification. Importantly, this includes delicate, free-living taxa that are abundant, but difficult to assess in lab-based experiments due to their fragility.

Here we detail the configuration, calibration, and at-sea deployment of the ZOOPS-O^2 system. First, the optical and the acoustic systems are described separately. The stereo-camera calibration and its progression to the acousto-optic calibration (mapping acoustic data into the camera coordinate system) are then documented. This includes quantifying the 3-dimensional location and pointing angles of all components. The acousto-optic coincident volume and the frequency response of the acoustic system are then detailed. The use of the geometrically calibrated system to estimate correspondence between individual sonar reflections and the target's 3-dimensional location and pose (orientation) relative to the sonar is then discussed. Finally, the application of the system to estimate the acoustic reflectivity, size, and orientation of 224 individual pelagic copepods from the orders Calanoida, Cyclopoida and Poecilostomatoida is described, with comparison to a well-known model, the Distorted Wave Born Approximation (DWBA) (Chu and Ye, 1999).

2. Material and methods

2.1. O-Cam—a shadowgraph camera system

The O-Cam (Fig. 1) is based on a shadowgraph design (Settles, 2001) that uses a machine vision camera for imaging and a Light Emitting Diode (LED) strobe for illumination. The camera, an AVT GX-1910 (Allied Vision, USA) employs a Kodak 1920 × 1080 CCD with 5.5 μm pixels. It is used with a Rainbow S6X11M-II motorized 2/3″ format CCTV zoom lens (Rainbow, USA). The magnification of 0.2 results in a 5.3 × 3.0 cm field of view, from which a 3.0 cm diameter circular sub-area is used. A 10W Blue LED array (LedEngin, USA) collimated with a condenser lens of focal length equal to 60 mm and diameter of 50.8 mm (Thor labs, USA) is used for illumination with a holographic diffuser (Edmund Optics, USA) placed between the LED array and the condenser, creating a more uniform illumination. The illumination system, in one housing, is aimed so the light beam projects directly into the camera in the second housing. The distance between the two housings is 71 cm, with the sampled volume approximately centered between the two.

The characterization of the camera system was relatively straightforward. Using test targets, the results indicated that a best resolution of approximately 30 μm is obtained at the focal plane with a field of view of 3 cm. Moving ±1.5 cm away from this center focal plane results in a decrease of resolution to approximately 54 μm. Images collected with the system permit reasonable identification of organisms as small as 360 μm in length (Fig. 2). In practice, the images are relatively immune to the potentially smearing effects of ship heave and water motion: the pulse length employed by the LED was ≤50 μs in all experiments reported here.

2.2. ZOOPlankton Sonar (ZOOPS)

The ZOOPlankton Sonar (ZOOPS) system consists of four single transducers that can be used to simultaneously transmit and receive (T/R) sound. However, here only one transducer is used in the T/R mode. All systems are controlled and synchronized with a National Instruments PXI-8081 embedded
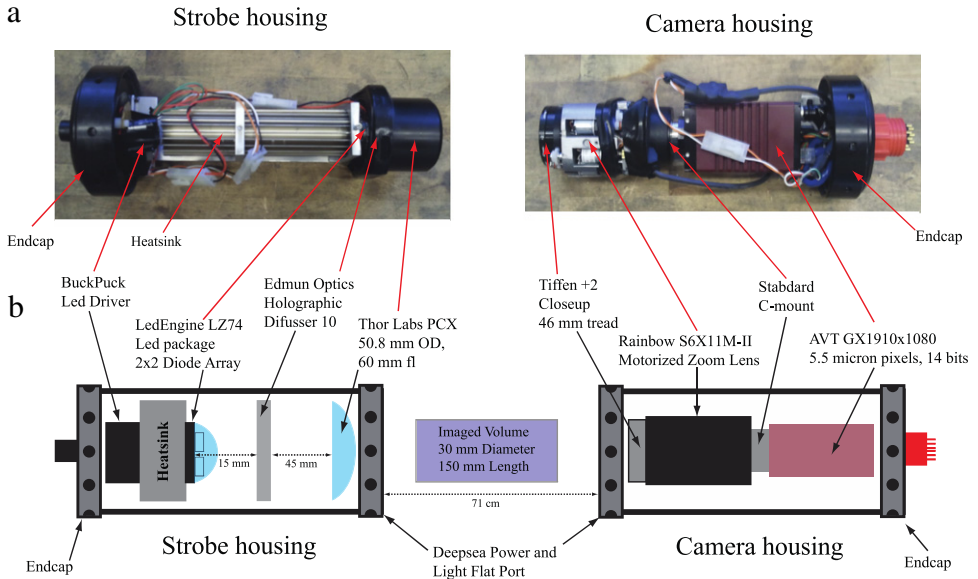
Fig. 1. O-Cam configuration. (a) The strobe and camera components with the housings removed. (b) Schematic of the strobe and camera housings. The imaged volume size representation is exaggerated in this diagram, but noted dimensions are accurate.
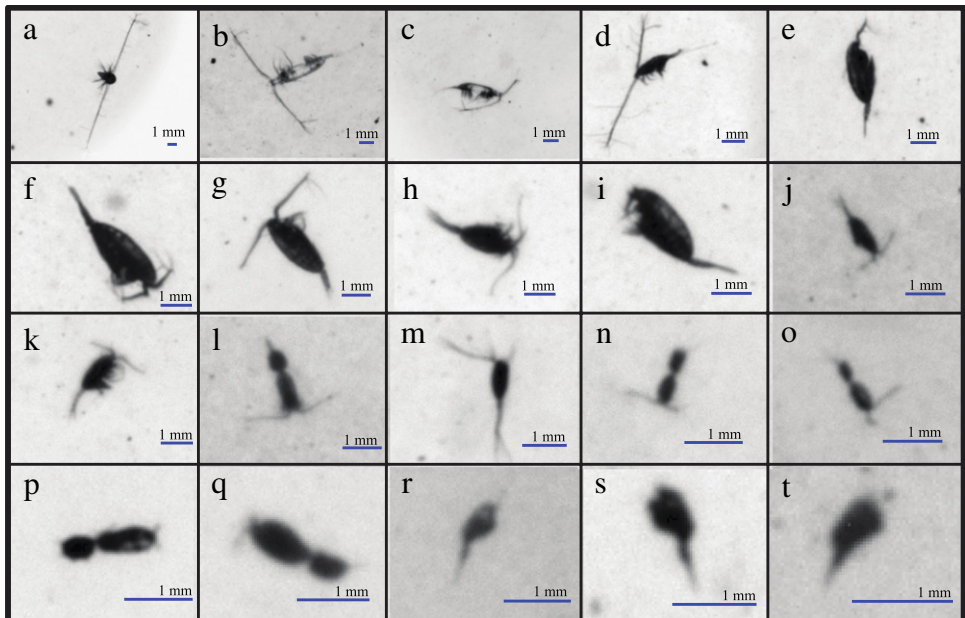

Fig. 2. Examples of O-cam *in situ* images of Calanoida (a–k), Cyclopoida (l–q), and Poecilostomatoida (r–t) copepods. Images are arranged in order of size from large to small (all scale bars were kept to 1 mm) for ease of comparison. Notice the wide range of orientations.

single board computer running Windows XP. In addition, a custom LabVIEW software program logged data from all of the sensors and controlled basic system operations. The signal synthesis and

acquisition boards used a PXI-5412 100 MHz analog output board (National Instruments, USA), and a PXI-6115 10 MHz 12 bit analog to digital converter (National Instruments, USA). For transmit, the low-level analog output signal from the PXI-5412 was amplified with a 250 W power amplifier (AR Worldwide, USA). For receive, the low-level output from the transducer (TC3021, Reson, USA) was amplified with low noise preamplifiers (N.T.S. Ultrasonics, Australia).

A linear frequency modulated (LFM) chirp was used for all experiments:

$$y(t) = \cos\left(2\pi t\left[\left(\frac{f_{max} - f_{min}}{2T}\right)t + f_{min}\right]\right)\sin\left(\frac{t\pi}{T}\right)^{\frac{1}{4}}, \quad \text{for } 0 \le t \le T \tag{1}$$

where $f_{min} = 1.5$ MHz, $f_{max} = 2.5$ MHz and $T = 0.5$ ms. The amplitude weighting via the sine term with fractional power in (1) provided a compromise between bandwidth and side lobes in the transmit signal. A digital version of (1) using a 10 MHz sample rate

$$y_{ref}[n] = \cos\left(\frac{2\pi Tn}{N}\left[\left(\frac{f_{max} - f_{min}}{2T}\right)\frac{Tn}{N} + f_{min}\right]\right)\sin\left(\frac{n\pi}{N}\right)^{\frac{1}{4}}, \quad \text{for } n = 0, 1, \ldots, N \tag{2}$$

was used for signal synthesis and correlation processing, where $N = 5000$ samples.

The received signal was amplified with 70 dB gain and digitized at 10 MHz with 12 bit dynamic range to yield the measured waveform $y_m[n]$. In adjusting the range of the A/D a value was chosen so that as much of the full dynamic range as possible was used from the anticipated targets at the given range. The digitized waveforms were then cross-correlated with the reference signal (2) to yield the pulse-compressed waveform

$$cp_m = y_{ref} \otimes y_m \tag{3}$$

where $\otimes$ denotes cross-correlation and the subscript m denotes the measurement. The compressed pulse has much better range resolution than the transmitted signal and therefore allows identifying reflections from individual targets and surfaces.

The sonar system was calibrated with a 5 mm-diameter tungsten-carbide sphere tethered with a 25 μm nylon monofilament, and using the partial-wave technique of Dragonette et al. (1981) that was extended to broadband waveforms and pulse-compressed processing by Stanton and Chu (2008). This method relies on isolating the first reflection from the calibration target, which is possible in this case due to the bandwidth of the system. To obtain the target strength of the calibration sphere, let the wavenumber of the sonar be k and the radius of the calibration sphere be a. At the center frequency of the sonar (2 MHz), $ka = 21$. It has been shown (Stanton and Chu, 2008) that when $ka \gg 1$, the magnitude of the first reflection from the sphere is approximately frequency-independent and the target strength of the sphere can be approximated as

$$TS_{sphere}(a) = 10\log\left(\frac{a^2}{4}\right). \tag{4}$$

To characterize reflections from sonar targets with a single scalar value, the pulse-compressed echo from the calibration sphere was then used to estimate the magnitude-squared of the first reflection

$$y_{cal} = \max_n\left(\text{Env}\left[cp_{sphere}[n]^2\right]\right) \tag{5}$$

where Env[] denotes the envelope of the waveform computed from the absolute value of its Hilbert Transform (Oppenheim et al., 1999) and the max is taken over a window around the echo. In a similar fashion, the magnitude-squared of the compressed pulse of the echo from a target is computed as

$$y_{bs} = \max_n\left(\text{Env}\left[cp_{bs}[n]^2\right]\right). \tag{6}$$

Next, using the partial-wave TS method described above, the Broadband Target Strength was defined as

$$BTS = 10\log_{10}\left(\frac{y_{bs}}{y_{cal}}\right) - 58.1 \text{ dB}. \tag{7}$$
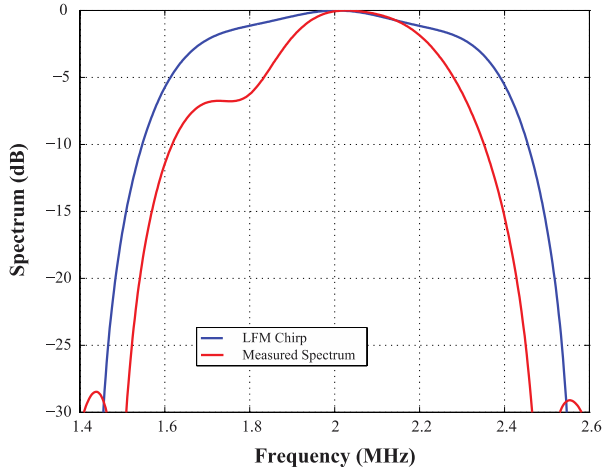
Fig. 3. A comparison of the LFM chirp spectrum (blue) with that of the measured spectrum (red). See text for definitions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Note that the BTS is a weighted superposition of frequencies that are coherently added. In this sense, it can be viewed as a weighted average of the individual Target Strengths as a function of frequency, with suitable weights imposed by the envelope of the transmitted signal over the signal bandwidth.

In addition to the BTS, an estimate of the spectrum, $B(f)$, from an individual target was computed. Let the Fourier transform of the windowed and zero-padded, first reflection from the calibration sphere be $CP_{sphere}(f)$, and the Fourier transform of windowed and zero-padded echo from a target be $CP_{bs}(f)$, where f denotes frequency. Then

$$B(f) = 10 \log_{10} \left| \frac{CP_{bs}(f)}{CP_{sphere}(f)} \right|^2 - 58.1 \text{ dB}. \tag{8}$$

A comparison of the spectrum of the digitally synthesized signal with that of the impulse response from the pulse-compressed first reflection is shown in Fig. 3. As can be seen, there is some reduction in bandwidth and a bit of shaping, primarily a result of the frequency response of the transducer and downstream electronics.

As part of the calibration, a value for the system noise level was established. This was determined by examining the average magnitude of the compressed-pulse output when no target was present, over a small range-resolved volume about 0.85 m from the transducer, where the sonar and optical volumes intersect (see Section 2.3.2). The BTS value here was found to be close to -130 dB. A target was judged to be present when the BTS exceeded this level by 10 dB resulting in a minimum target detection level of BTS $= -120$ dB and an SNR of 10 dB.

2.3. ZOOPS-O^2 system configuration and components

The two stand-alone systems (O-Cam and ZOOPS) were assembled into a platform named ZOOPS-O^2, the "O^2" indicating that two O-Cams were configured together to build a stereoscopic system. ZOOPS-O^2 is a system that simultaneously acquires both optical images and the broadband backscatter from particles within the common sampling volume of the transducer and cameras. In this configuration, it is possible to identify and extract the echo produced by a target while optically imaging the same target that produced the echo. The two cameras allow calculation of the 3D locations of the imaged particles. The projected optical images can also be used to infer the 3D locations of parts of the target by identifying common points in the images. This facilitates the measurement of particle length, width, and the orientation of the longest axis, relative to the incident sonar beam. The resultant
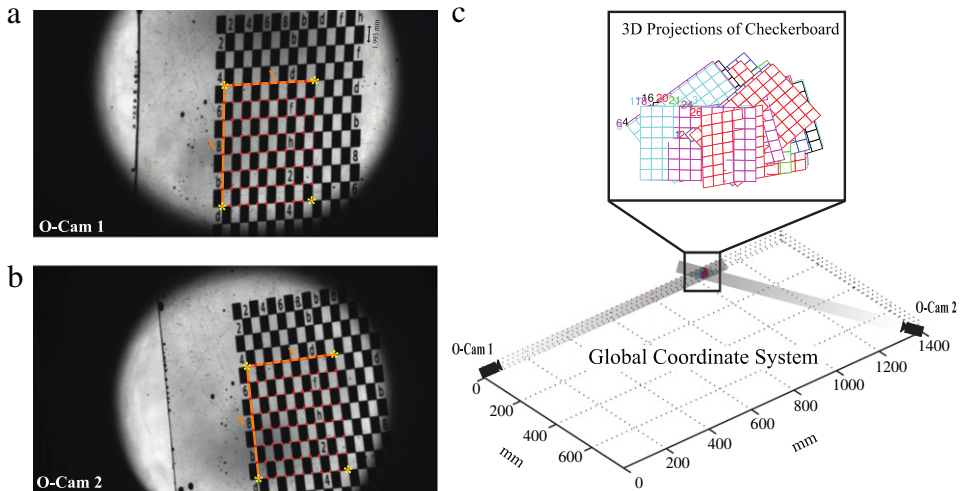
Fig. 4. Example of the stereo camera calibration. (a–b) Images of the checkerboard target simultaneously imaged by O-cam 1 and O-Cam 2. The yellow asterisks are the corners of the effective area used during calibration. The Orange lines indicate the *x* and *y* axes with respect to the upper left corner. (c) The Global Coordinate System (GCS) showing the relative position of O-Cam 2 with respect to O-Cam 1. The intersected imaged volume (gray shaded paths) is enlarged to show the checkerboard projections in the stereoscopic volume. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sampled volume in this system (estimated at 20.03 mL) is constrained by the stereoscopic system and can be approximated by the volume of two intersecting cylinders (Hubbell, 1965).

2.3.1. Stereoscopic system

The O-Cams were configured to obtain a stereoscopic view of a common volume, subsequently referred to as the "stereoscopic volume". The cameras were positioned and aimed so that the center of the common volume and the center of the two camera images were coincident. In addition, the camera volumes were configured so that the volume from each camera intersected the center of the other camera's depth of field. This resulted in a stereoscopic volume that was approximately one fifth as large as the total imaged volume. However, it allowed accurate triangulation of targets that appeared in both of the cameras.

2.3.1.1. Stereoscopic system calibration. The cameras were calibrated using the stereo calibration toolbox for MATLAB (Jean-Yves Boguet, http://www.vision.caltech.edu/bouguetj/calib_doc/, Camera Calibration Toolbox for MATLAB). Although this software was used for convenience, the calibration of stereo imaging systems is well known; methods generally rely on imaging a common set of targets to establish the intrinsic and extrinsic camera parameters with the extrinsic parameters for each camera referenced to a global coordinate system (Jain et al., 1995). Given the small imaged volume and the geometry of the cameras, calibration was performed using a checkerboard target with 2 mm $\pm$ 20 μm squares printed on a transparency (Fig. 4(a), (b)). The stereoscopic calibration maps points in the 3D space defined by the intersection of two cameras' fields of view onto pixels in each camera image. Fig. 4(c), for example, shows all the back-projections of the calibration checkerboard images in 3D space and with respect to the relative position of the O-cams. Because only a small area of the checkerboard was visible simultaneously in both cameras (area bounded by yellow asterisks and orange lines in Fig. 4(a), (b)), the target was marked with letters and numbers to uniquely identify the squares used during the calibration process. All calibration exercises were done in a large tank filled with fresh water. This did not likely introduced significant bias into our calibration measurements.

The result of the calibration provided the input to the algorithm that computed the 3D positions of objects simultaneously imaged by the two cameras. All of the objects, including cameras, their
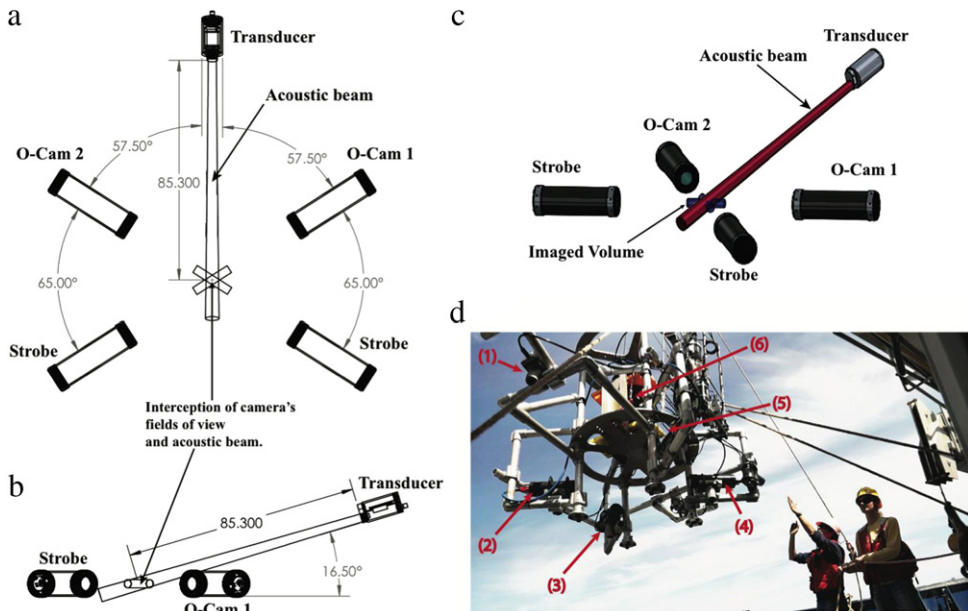
Fig. 5. Line diagrams (a–c) and picture (d) of ZOOPS-O^2 showing different views. (a) Top-down view of the configuration of the O-Cams and the transducer. The acoustic beam and imaged volumes from the cameras are also drawn. (b) Side view (minus one O-cam for visual clarity) to show the inclination of the transducer with respect to the cameras. (c) 3D rendering showing the two O-cams and transducer as well as the acoustic beam intercepting the stereoscopic volume. (d) The system being deployed at sea from the R/V New Horizon. (1) Transducer; (2) O-Cam 1; (3 and 4) Strobe housings (4) O-Cam 2; (5) temperature–pressure sensor.

pointing angles, and the subsequent location of the transducer, were defined in a common coordinate system, referred to as the Global Coordinate System (GCS) (Fig. 4(c)).

2.3.2. Combining optics and acoustics

In order to use the stereo optics to identify the target that generated the observed acoustic reflection, the two systems, ZOOPS and the two O-Cams, now O^2, were assembled using a metal frame (Fig. 5(d)). For this study, the two O-Cams were mounted on the frame 25.5 cm below a circular aluminum plate with a post-calibrated angle of 115° between them (Fig. 5(a)). The transducer was aimed downward with a post-calibrated angle of 16.5° with respect to the O-Cams (Fig. 5(b)) and, as measured, approximately 85.3 cm from the center of the stereoscopic volume (Fig. 5(a), (b)). This is where the acoustic beam crosses the stereoscopic volume. This geometry resulted in the transducer beam pointing toward the stereoscopic volume (Fig. 5(c)) while avoiding strong reflections from the aluminum frame.

The next task was to establish object positions in the GCS and measured range in the acoustic system, given the physical location of all devices. Here, the tungsten-carbide sphere tethered with a 25 μm nylon monofilament used for acoustic calibration was imaged by the stereoscopic camera system, while simultaneously recording acoustic reflections (pings). The acoustic reflection was characterized by its range (r) in meters (subtracting 2.5 mm to compute the center of the sphere) and the BTS in dB. For each acoustic measurement, the position of the tungsten-carbide sphere in the GCS was recorded. The total data set consisted of 1342 image pairs and "pings".

To co-register the optic and acoustic systems, the 3D origin of the ping in the coordinate system of the stereo camera is needed. To accomplish this a simple minimization algorithm was used. Given the set of 1342 3D locations of the sphere, where (x_i, y_i, z_i) designates the "ith" location, and the corresponding, measured, acoustic range as $r_{measured_i}$ we seek the acoustic origin of the transducer (h, k, l) so that $r^2_{measured_i} = (x_i - h)^2 + (y_i - k)^2 + (z_i - l)^2$ over the entire set of data. The equation
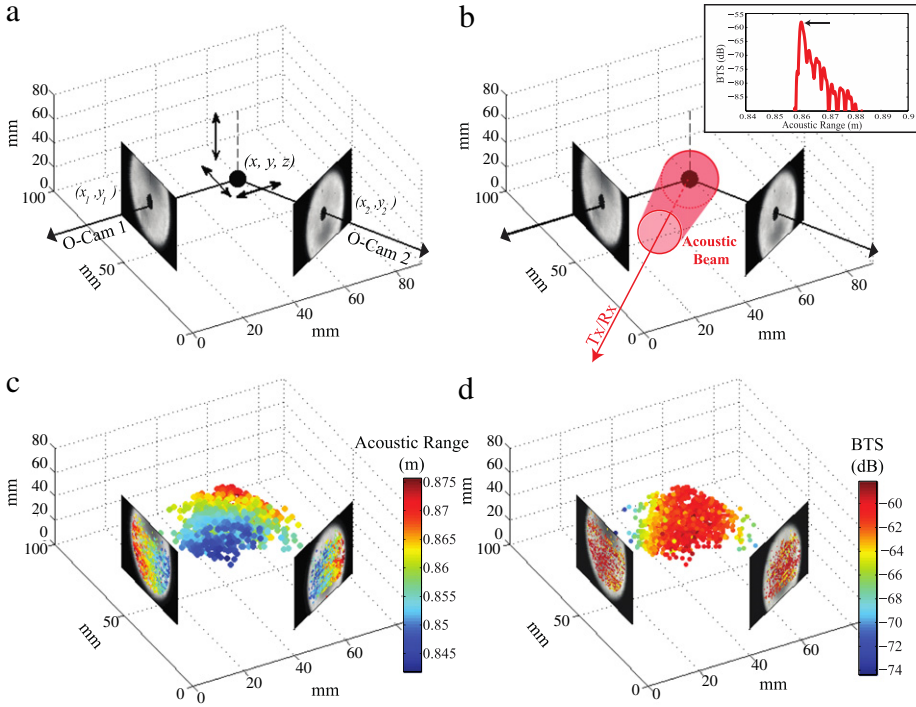
Fig. 6. The optic and acoustic data in the 3D global coordinate system (GCS). (a, b) The 5 mm tethered tungsten-carbide sphere being simultaneously imaged and ensonified. The inset panel in (b) shows the acoustic record with the peak of the first return indicated with an arrow. The red truncated cone represents a portion of the idealized acoustic beam and the general direction and relative location of the transducer (Tx/Rx). (c, d) The acoustic data (acoustic range, broadband target strength: BTS) mapped in the GCS with the corresponding camera viewpoints, as re-projected into the coordinate system of the cameras. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ensures that the distance from the origin of the acoustic sound to the measured range is equal to the measured range to the sphere as computed from the 3D optical data. The position of the acoustic origin is estimated by minimizing

$$\min \left\{ \sum_{i=1}^{N} abs \left[(x_i - h)^2 + (y_i - k)^2 + (z_i - l)^2 - r^2_{measured_i} \right] \right\}. \tag{9}$$

The MATLAB function *fminsearch* was used to find (h, k, l) that satisfied Eq. (9), though a variety of optimization algorithms could have been used to solve this equation as there are only three unknown parameters with search bounds given by positions measured on the aluminum frame (Fig. 5). Finally, given (h, k, l) and (x_i, y_i, z_i), the acoustic ranges ($r_{predicted_i}$) were predicted using the expression:

$$r^2_{predicted_i} = (x_i - h)^2 + (y_i - k)^2 + (z_i - l)^2.$$

The optimized values of (h, k, l) taken over all of the 1342 tungsten-carbide sphere locations resulted in a residual between $r_{measured_i}$ and $r_{predicted_i}$ with a mean of 0.063 mm and standard deviation of ±0.2217 mm. This acousto-optic calibration produced a mapping between the 3D, optically imaged volume, computed from the pair of stereo images, and the corresponding acoustic range of the transducer. This then permitted the localization of a particular target's echo in range given the target location in the GCS computed from the stereo image pair.

Data from the acousto-optic calibration is displayed in Fig. 6 where the basic setup that includes the pair of images and the 3D location of the sphere (Fig. 6(a)), and the acoustic beam (Fig. 6(b)) is shown.

Fig. 6(c) shows the set of corresponding acoustic ranges computed from the derived knowledge of 3D position and the range of positions of the sphere that was moved around inside the illustrated volume. As expected, a single range describes an arc of positions, centered on the transducer origin. Fig. 6(d) shows the estimated set of BTS values, as computed using Eq. (7) (Section 2.2). We note that, over the range of target positions, the recorded BTS of the sphere (assuming its value of −58.1 dB is at the center) does not vary by more than ±3 dB. It was therefore not necessary to obtain an estimate for the center of the acoustic beam and then correct for target location relative to the beam axis. Likewise, corrections for spreading and attenuation were not implemented, for the same reason. As a further confirmation, a simulation of the broadband beam pattern for the identical diameter transducer after implementing the correlation receiver indicated that the beam width was in agreement with that measured via the use of the calibration sphere.

In summary, the stereo calibration resulted in the ability to compute a 3D location in the GCS given the set of labeled points from O-Cam 1 and O-Cam 2. In addition, the acoustic and optical calibration permitted the prediction of the range of the target observed by the sonar from the location in the GCS. As observed, the small variance in the set of calibrated records indicate the accuracy of this procedure to at least a quarter of the system's range resolution of close to 1.5 mm and, in almost all cases, when an echo greater than −120 dB was recorded, the 3D position of targets predicted an echo that was observed at the corresponding acoustic range.

2.3.3. Ancillary components

Additional components, not listed above and shared by both systems were: (1) A set of 8, 95 W-hour Lithium-Ion battery packs (BA95HC-FL, Ocean Server) allowing up to 8 h of remote deployment of the entire system, (2) a large pressure housing with depth rating of 500 m, and (3) a temperature–pressure (TP) sensor (SBE 39, Seabird, USA) mounted at the same level as the cameras.

2.4. Acquisition and processing of in situ data

ZOOPS-O^2 was deployed in the Southern California Bight from the R/V New Horizon March 28–29, 2013. The system was vertically profiled at speed of 15–25 m min^{-1} in a self-recording mode to its maximum rated depth (500 m) or close to the bottom when shallower. Data were processed according to the methodology above and then automatically inspected to identify echoes from the correlator output (Eq. (6)) that exceeded the given BTS signal to noise ratio threshold of −120 dB at ranges between 0.84 m and 0.88 m. Corresponding image pairs were then visually inspected and images that were judged to contain the same zooplankter or particle were selected for further processing.

A graphical user interface algorithm (echo-locator) based on the results of the calibration exercises in Sections 2.3.2 and 2.3.1.1 was used to manually identify the echo of the particle imaged *in situ*. The echo-locator algorithm works in the following way: selected imaged pairs are displayed side by side. The user selects the particle present in pictures from both cameras by clicking first on the image from O-cam 1, then on the image from O-cam 2, trying to select the same part of the zooplankton or particle being displayed (red circles on Fig. 7(a)–(b) and (d)–(e)). By clicking on the images, the points (x_1, y_1) for O-cam 1 and (x_2, y_2) for O-cam 2 are recorded in pixel units. The algorithm then uses the *stereo_triangulation* function from the MATLAB camera calibration toolbox to triangulate the pixel coordinates, which locates the particle in the GCS. To ensure that the same object was selected in both cameras, the point from O-cam 1 is re-projected back onto O-cam 2. If the re-projected point (yellow circles on Fig. 7(b), (e)) lies on top of the selected point (red circles on Fig. 7(b), (e)) from camera 2 then the particle selected on both images is inferred to be identical. We call this the re-projection condition.

Once the re-projection condition was satisfied, the particle location was used to estimate its distance to the transducer ($r_{predicted}$, see Section 2.3.2). The echo found at that location was then interpreted as being produced from the *in situ* optically identified target. The echo was then extracted for further processing that included determining the BTS, and spectrum, $B(f)$.
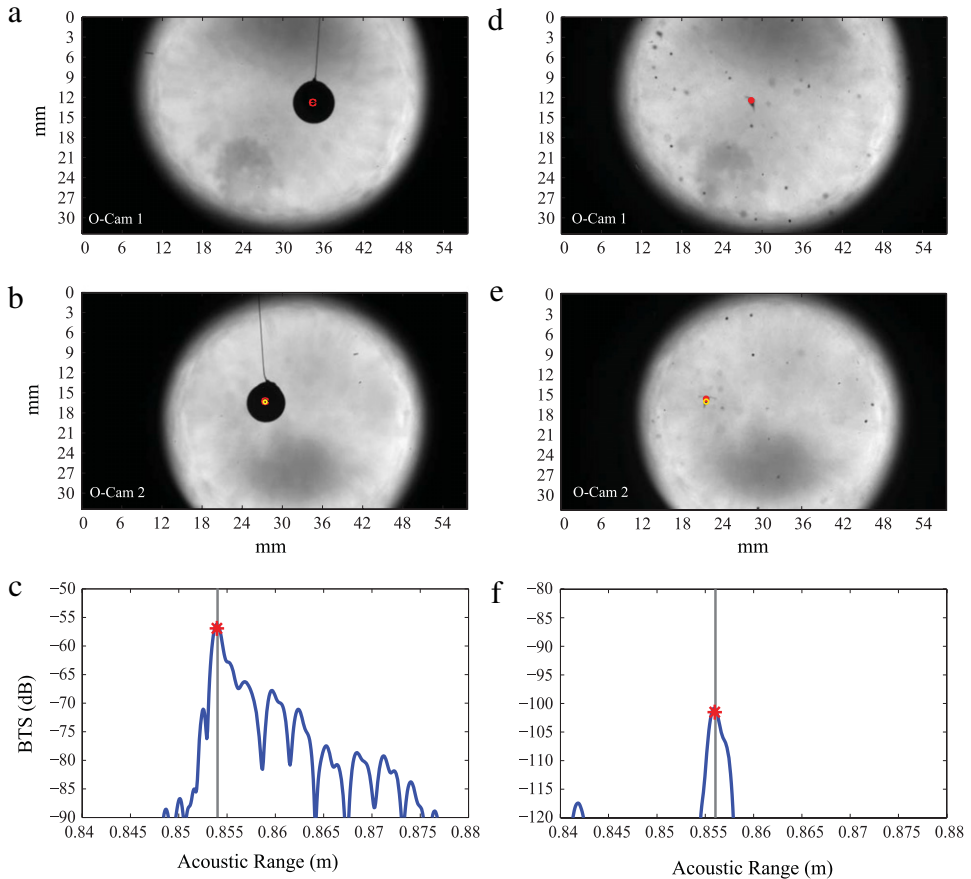
Fig. 7. Results of the echo-locator algorithm for calibration sphere (a, b, and c) and an *in situ* recorded calanoid copepod (d, e, and f). Red circles in a, b, d, and e indicate manually selected points on individual images. The yellow circles are the points on images from camera 1 (O-Cam 1) re-projected onto images from camera 2 (O-Cam 2). Red asterisks in c and f indicate the peak location of the extracted echo (blue line) and the gray vertical lines represent the predicted peak location in the acoustic range. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. Computing BTS and modeled spectra, B(f), for 224 copepods

Given that thousands of stereo images and their concurrent acoustic data were processed, there were many options for illustrating the performance of the system. Here, the system's performance is demonstrated via the analysis of a subset of targets: 224 individual pelagic copepods from the orders Calanoida, Cyclopoida, and Poecilostomatoida. Data from these 224 organisms were used in a statistical analysis of the reflected energy as a function of target size and orientation. The spectra from the individual copepods were compared to predictions of the Distorted Wave Born Approximation (DWBA) (Chu and Ye, 1999; Lavery et al., 2002) implemented for a homogeneous, prolate spheroid. This comparison will allow comparison of the actual (data) and predicted (DWBA model) effects of size and orientation on the reflected acoustic signal.

2.5.1. Estimating geometric properties from stereoscopic image data

Taking advantage of the stereoscopic calibration (Section 2.3.1.1) each of the 224 copepods was modeled as a prolate spheroid, and its major axis **a**, and minor axis **b** were estimated from image data. To accomplish this, four points were selected on the copepod shown on the O-Cam 1 image in
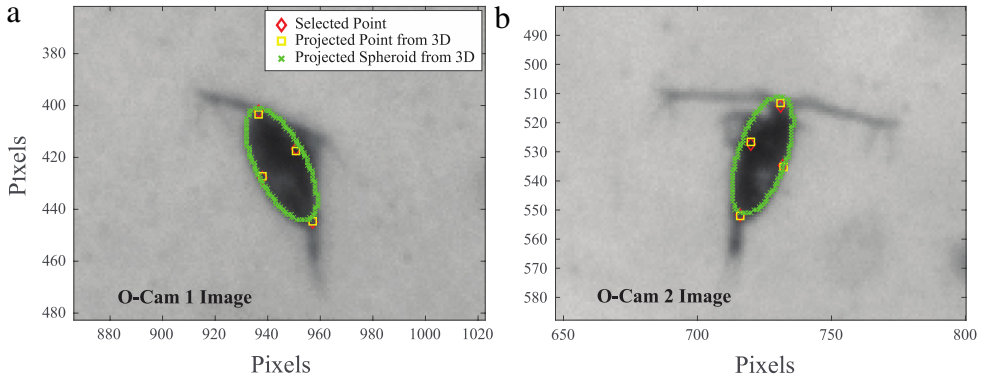
Fig. 8. Example of the geometric parameter estimation from stereo image pairs. Panels a and b show the O-Cam 1 and O-Cam 2 images of the same copepod, respectively. The red diamonds show the four points selected manually in each image. The yellow squares show the re-projection of the four 3D points estimated from the *stereo_triangulation* back into each image. The green xs demarcate the boundary points of the estimated spheroid projected back into each image. This boundary encloses the copepod's prosome. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

this specific order: tip of the prosome (P1), end of the prosome (P2), left side of prosome (P3) and right side of prosome (P4). The same was repeated for the corresponding copepod on the O-Cam 2 image. The prosome is the main part of the copepod's body (see Fig. 8). Because these points come from the O-cam images, they are in two dimensions. Hence, these two-dimensional points were translated into the GCS using the *stereo_triangulation* algorithm to yield a set of four 3D points p_1, p_2, p_3, and p_4, respectively. The major axis length was then estimated as $a = \|p_1 - p_2\|_2$, the minor axis length as $b = \|p_3 - p_4\|_2$, and θ, the orientation of the major axis **a** with respect to the acoustic beam was estimated using

$$\theta = \cos^{-1}\left\langle \frac{p_1 - p_2}{\|p_1 - p_2\|}, \hat{\mathbf{w}} \right\rangle \tag{10}$$

where the unit vector $\hat{\mathbf{w}} = \mathbf{w}/\|\mathbf{w}\|$ is the center orientation of the beam in the GCS and $\langle\,\rangle$ denotes the inner product. Finally, the transducer vector **w** (o_1, o_2) was defined by the 3D point o_1 (origin of the transducer (h, k, l) in the GCS system estimated in Section 2.3.2) and a point o_2. Point o_2 fell on the middle of major axis vector **a** (p_1, p_2) and was estimated using the mid-point formula: $o_2 = [\frac{1}{2}(x_1 + x_2), \frac{1}{2}(y_1 + y_2), \frac{1}{2}(z_1 + z_2)]$. Note that **w** is estimated dynamically depending on the location of vector b in the GCS. This way, **w** intersected the major axis of the copepod in question. Because the beam spread angle was 2.2°, this method did not introduce any significant bias to the angle estimates.

An example of this estimation process is shown in Fig. 8, where the estimated spheroid from the point measurements is projected back into the images to confirm that it is consistent with the image data.

To confirm the quality of the estimates of the geometric parameters from the image data, a MATLAB simulation was performed using the stereoscopic calibration data. In this simulation, a spheroidal copepod was projected into each camera view to represent an imaged copepod and the procedure above was used to estimate the best-fit spheroid model. The spheroidal copepods were given "antennae" and a "urosome" as position clues in order that the points on the projected images could be selected in the same order as in the *in situ* data (e.g., Fig. 8). Because the axis parameters of the spheroidal copepod are known, this tested the validity of the estimates of a and b from the *in situ* imaged copepods by the human operator. Out of a set of 10 simulation tests with randomly sized and oriented spheroidal copepods, excellent agreement was found between the simulated and predicted major axis length, and orientation (data not shown).

2.5.2. Applying the Distorted Wave Born Approximation (DWBA) to a spheroid

The DWBA for backscattering from a homogeneous prolate spheroid of major axis a, and minor axis b, and major-axis orientation relative to the sonar beam θ is given by Chu and Ye (1999):

$$f_{bs}(a, b, \theta, k) = k_1^2 ab^2(\gamma_k - \gamma_\rho)\frac{j_1(2k_2\Phi(a, b, \theta))}{2k_2\Phi(a, b, \theta)} \tag{11}$$

where j_1 denotes the spherical Bessel function of order 1, and

$$\Phi(a, b, \theta) = \sqrt{a^2 \cos^2(\theta) + b^2 \sin^2(\theta)} \tag{12}$$

and $\gamma_k = (1 - gh^2)/gh^2$, $\gamma_\rho = (\rho_2 - \rho_1)/\rho_2 = (1 - g)/g$ and $g = \rho_2/\rho_1$, and $h = c_2/c_1$, are the density and sound speed contrasts between the body and the surrounding medium. To predict values of the BTS and spectrum that would be measured from targets with geometric values of measured (a, b, θ) the equation was implemented in MATLAB using values of $g = 1.05$ and $h = 1.05$ (Stanton et al., 1998a,b).

3. Results

3.1. Discerning echo peaks for multiple targets within the stereoscopic volume

To explore the system's ability to discern echoes from closely spaced targets, image pairs containing multiple targets were evaluated. Fig. 9 illustrates one such image pair containing three copepods of different sizes within the stereoscopic volume (Fig. 9(a), (b)), along with their predicted echoes (Fig. 9(c)). Note that one copepod in each image (marked with black arrows in Fig. 9(a), (b)) did not appear in the corresponding image. Thus they were separate individuals, located outside of the acoustic beam and the stereoscopic volume. This conclusion is further validated by the absence of additional acoustic reflections inside the shared acousto-optic volume.

3.2. Measured versus Modeled (DWBA) BTS and spectra for copepods

A comparison between the modeled (DWBA) and measured BTS for 224 individual copepods measured *in situ* is summarized in Fig. 10. The graph shows the BTS, one for each animal, as a function of θ (the angle between the major axis a and the transducer vector **w**, Section 2.5.1). The copepods ranged in length (a) from 360 μm to 4.47 mm and in thickness (b) from 16.4 μm to 1.232 mm. As illustrated, the best agreement between the BTS and DWBA occurs when the copepods are broadside to the acoustic beam. Interestingly, as the organism becomes more aligned with the axis of the sonar beam, the measured BTS decreases more slowly than predicted by the DWBA for the same spheroid. Note that a set of systematic adjustments of the values of the g and h parameters within a reasonable range ($1.01 \leq g, h \leq 1.05$) did not change the BTS values enough to reconcile the two data sets. We also note that the potential error in the angular estimate increases as θ approaches either 0 or π. Accordingly, the variance in the data increases at these more oblique angles.

Spectra, $B(f)$, were also computed for each of the organisms over the usable bandwidth of the *in situ* system between 1.6 and 2.4 MHz. A comparison of the DWBA spectra and those estimated from the acquired data is shown in Fig. 11.

4. Discussion

We have described the combination of a new broadband system (ZOOPS) and a set of underwater microscopes (O-Cam) for viewing plankton. Results indicate that the combined systems have the capability to acoustically observe copepods as small as 360 μm while, at the same time, optically identifying the organism that reflected the sound. While the system performed quite well and the combination of optics and acoustics was successful, a number of instrumentation and interpretation issues remain.
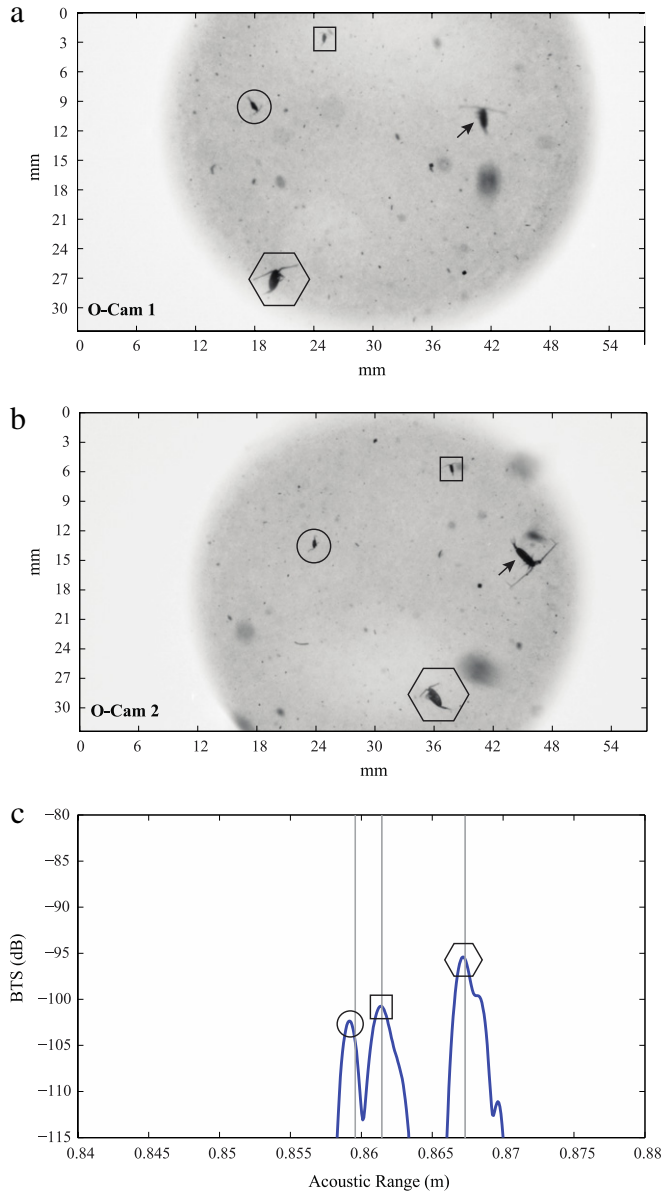
Fig. 9. Panels a and b are two temporally synchronized images. The images were processed as described to yield a 3D position, and hence a prediction of acoustic range, for each target. Three copepods yielded acoustic reflections whose peaks were well predicted, as indicated by the thin vertical lines in panel c. Geometrical shapes denote the corresponding echo to each copepod in a and b. An additional two copepods, shown by the two arrows in panels a and b were judged to be different and, as expected, the targets are outside of the coincident 3d volume and hence, there are no measured acoustic reflections from the organisms in the acoustic record.

It is important to discern what level of confidence exists in the determination of the BTS values. As explained, the tungsten sphere's BTS was predicted, based on the size and the isolation of the first reflection. While this value is undoubtedly reliable, it is possible that the extrapolation of this value to the reflectivity of small animals that reflect four orders of magnitude less sound could produce
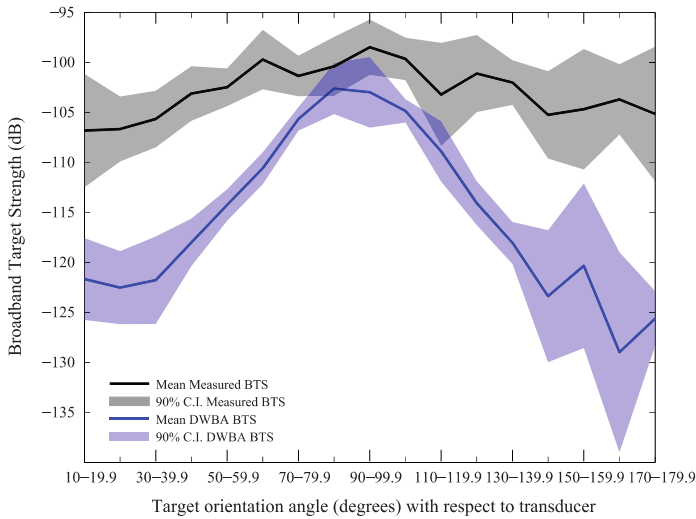
Fig. 10. Mean (±90% Confidence Intervals, shaded areas) estimates for measured (black) and DWBA (blue) BTS for 224 copepods. Data was binned every 10° (target orientation with respect to transducer) to estimate the presented statistics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
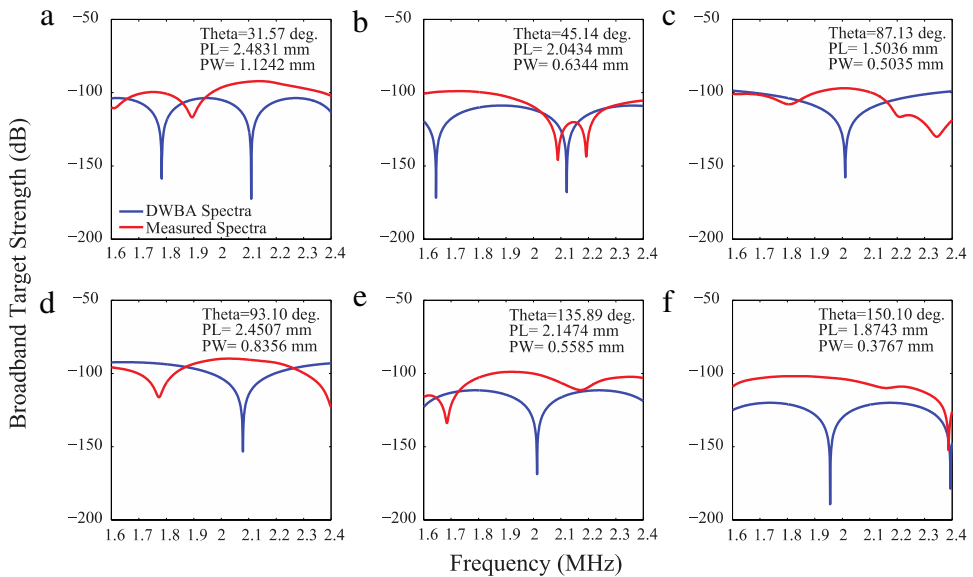

Fig. 11. Measured (red lines) and modeled (DWBA, blue lines) spectra $B(f)$ from five copepods at different orientations with respect to the sonar. Theta values are given in degrees. PL = Prosome Length (major axis a); PW = Prosome Width (minor axis b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

errors. To accommodate this large dynamic range the dynamic range of the receivers was adjusted via changes in the preamplifier gains and the full-scale extent of the digitizer. However, there is the potential of introducing error if the changes in amplification and/or the voltage range were not completely linear. We thus regard the observation that the backscatter strength showed less angular dependence than predicted via the DWBA to be a more reliable conclusion than the absolute values of the BTS.

Another important concern is whether the 25 μm tether (visible in Figs. 7(a), (b) and 6(a), (b)) on the 5 mm sphere contributed to the observed BTS during calibration tests. To explore this, we performed experiments where only the monofilament was in the imaged 3D volume. The measured BTS (−120 to −130 dB) was several orders of magnitude lower than the calibration sphere (−58.1; see Section 2.3.2). In addition, no knots were used in wrapping the monofilament around the sphere. Rather, the tip of the monofilament was dipped into waterproof glue and immediately attached to the sphere and let dry overnight. The droplet was mostly flattened at the junction of the line and the sphere's surface, likely causing no effect on the overall reflectivity of the sphere. Though it is virtually impossible to judge the overall effect of the filament, in our estimation it was minimal compared to that of the sphere.

One of the major advantages of combining the optics and acoustics is to measure acoustic echoes with taxonomic identification of the corresponding, visually identified targets. The opto-acoustic calibrations shown in Section 2.3.2 demonstrate that the images and echoes from individual targets can be simultaneously and accurately located in both the stereoscopic volume and in the acoustic record. Furthermore, the minimal difference between the predicted and observed distances of the calibration sphere with respect to the transducer indicates that *in situ* targets can be accurately located in the acoustic record. Fig. 9, for example, shows that the system can resolve the acoustic peaks of three copepods contained within the stereoscopic volume. The figure also shows how the presence of an animal in one camera but not the other implies that the animal is outside of the stereoscopic volume.

A current goal in ground-truthing acoustic systems is to reduce the uncertainties in interpreting field data. Stanton (2012) identified several issues regarding acoustic methods, especially at the high-frequency (MHz) range. One is the influence of orientation (tilt angle) of elongated organisms. The stereoscopic nature of ZOOPS-O^2 provides an ideal tool for studying the dependence of BTS on *in situ* orientation, without the confounding influence of tethers, artifacts associated with laboratory settings, or the physiological state of the organisms due to stress or other post-capture factors (e.g., Greenlaw, 1977; Johnson, 1977). As shown in Section 2.5 the major and minor axes (a and b, respectively), and the orientation of the major axis a with respect to the acoustic beam were estimated for 224 copepods from the *in situ* data using the stereoscopic calibrations. Assuming the copepod prosome resembled a spheroid, the expected BTS inferred using the well-known Distorted Wave Born Approximation (DWBA) model was compared with the measured BTS.

An interesting result from our comparison of *in situ* measurements of 224 copepods with predictions from a theoretical model was that the spheroidal DWBA under-predicts the BTS compared to the measured data (Fig. 10). As is widely appreciated, copepods are not simply spheroidal bags of homogeneous substance but have exoskeletons with antennae, and feeding and swimming appendages. The internal organs of potentially different composition, and oil sacs (spherical or oblong) present in some species of Calanoid copepods could contribute to the observed differences between the measured and DWBA. Furthermore, copepod body shape can be diverse (Fig. 2), and female copepods can often carry spermatophores (sacs containing the male's sperm) or egg sacs attached to their urosome ("tail") (Fig. 2(l), (n)–(q)). The spheroidal DWBA does not consider the effects of these various features; the nulls that are evident in the spectra in Fig. 11 may be caused by these differences and, moreover, the potential increase in backscatter contrast due to these features are not incorporated into the model.

Given that a perpetual confounding influence in interpreting sonar backscatter has been the organism orientation, it is certainly welcome to see that the *in situ* data are less orientation-dependent than predicted from the elongate structure based on the dimensions of the prosome. Although evaluation of the angle between the animal's major axis and that of the sonar beam is more error prone as the animal becomes less broadside, there is a much less decrease in the BTS value than predicted by the spheroidal model approximation. We therefore have confidence that the departures of the data from the model are real. From this observation we conclude that it is likely zooplankton surveys are less susceptible to copepod orientation than previously thought.

There is a wide range of other applications of the ZOOPS-O^2 system, with many potential benefits. For one, the system could be mounted on the sea floor. This would be useful in a shallow water environment to study plankton abundance and the changes that occur due to predation,

reproduction, and dispersion due to flow over local bottom features. In this case, with slow currents, the hydrodynamic signature of the system and the potential for its detection by the organisms would be small. An additional strategy would be to measure changes in abundance, and possibly rates of predation, along current trajectories by employing multiple spatially separated transducers.

It is natural to think that future systems focused on different targets could use the same methodology. For example, there is a dire need for understanding the relationship between backscatter changes as a function of tilt angle for fish. One approach would be to use a stereo camera system with a sonar to both image the fish and measure its backscatter. The requirement that the organism be in the far field of the sonar imposes a minimal distance for this to work; however, we anticipate that in the future, coupled acoustic and optic systems will become ever more popular, perhaps following the optical design of the Cam Trawl (Williams et al., 2010) and coupling that to acoustics, in a similar way to the work of Ryan et al. (2009) and Macaulay et al. (2013), where they studied the *in situ* target strength of orange roughy (*Hoplostethus atlanticus*).

With the advent of broadband technologies such as the Simrad EK80 system, an interesting option would be to attempt to classify targets using features of the echo. Such features could include echo envelope width, shape or other statistics. These methods are currently being used to classify images (e.g., Sosik and Olson, 2007), and could be fruitfully extended to sonar echoes. This would be especially valuable in turbid conditions, where the optical system would have substantially less range capability than the acoustics. The joint use of acoustics and optics would also be synergistic in the case of object classification from multiple views, where the different modalities can be fused together to mitigate the orientation-dependent appearance of elongated plankton in image and acoustic data (Roberts et al., 2009).

Lastly, an important area of inquiry would be to use the ZOOPS-O^2 system to study the acoustic properties of fragile, gelatinous taxa (e.g., doliolids, siphonophores, hydromedusae, and ctenophores) as well as chaetognaths and appendicularians, which have important ecological roles. Such organisms have been encountered during sampling with the system, and efforts to characterize the BTS of a variety of pelagic entities, including marine snow, are currently underway.

5. Summary

The ZOOPS-O^2 system combines a stereoscopic camera system with broadband sonar to estimate acoustic reflectivity of individual plankton with concurrent measurement of their size and orientation. Measurements of the system's performance show that copepods as small as 360 μm give a reflected broadband target strength (BTS) that is at least 10 dB above the system's noise level. The 1.5 mm range resolution of ZOOPS, along with the visual identification capabilities of the O-Cam (designated O^2 in stereoscopic mode) permits extraction of echo properties of individual copepod targets in their natural environment with concurrent identification. As demonstrated, ZOOPS-O^2 can be used to study the acoustic properties of this ecologically important plankton group *in situ*.

Acknowledgments

References

Benfield, M.C., Wiebe, P.H., Stanton, T.K., Davis, C.S., Gallager, S.M., Greene, C.H., 1998. Estimating the spatial distribution of zooplankton biomass by combining Video Plankton Recorder and single-frequency acoustic data. Deep-Sea Res. II 45, 1175–1199.
Chu, D., Stanton, T.K., 1998. Application of pulse compression techniques to broadband acoustic scattering by live individual zooplankton. J. Acoust. Soc. Am. 104 (1), 39–55.
Chu, D., Ye, Z., 1999. A phase-compensated distorted wave Born approximation representation of the bistatic scattering by weakly scattering objects: Application to zooplankton. J. Acoust. Soc. Am. 106 (4), 1732–1743.

Costello, J.H., Pieper, R.E., Holliday, D.V., 1989. Comparison of acoustic and pump sampling techniques for the analysis of zooplankton distributions. J. Plankton Res. 11 (4), 703–709.

Dragonette, L.R., Numrich, S.K., Frank, L.J., 1981. Calibration technique for acoustic scattering measurements. J. Acoust. Soc. Am. 69 (4), 1186–1189.

Fernandes, P.G., Gerlotto, F., Holliday, D.V., Nakken, O., Simmonds, E.J., 2002. Acoustic applications in fisheries science: the ICES contribution. ICES Mar. Sci Symp. 215, 483–492.

Fornshell, J.A., Tesei, A., 2013. The development of SONAR as a tool in marine biological research in the twentieth century. Int. J. Oceanogr. 2013, 1–9.

Gallienne, C.P., Robins, D.B., 2001. Is *Oithona* the most important copepod in the world's oceans? J. Plankton Res. 23 (12), 1421–1432.

Genin, A., Jaffe, J.S., Reef, R., Richter, C., Franks, P.J.S., 2005. Swimming against the flow: A mechanism of zooplankton aggregation. Science 308, 860–862.

Greene, C.H., Wiebe, P.H., Burczynski, J., 1989. Analyzing zooplankton distributions using high-frequency sound. Limnol. Oceanogr. 34 (1), 129–139.

Greenlaw, C.F., 1977. Backscattering spectra of preserved zooplankton. J. Acoust. Soc. Am. 62 (1), 44–52.

Holliday, D.V., Pieper, R.E., 1980. Volume scattering strengths and zooplankton distributions at acoustic frequencies between 0.5 and 3 MHz. J. Acoust. Soc. Am. 67 (1), 135–146.

Hubbell, J.H., 1965. Common volume of two intersecting cylinders. J. Res. Natl. Bur. Stand. C 69C (2), 139–143.

ICES. 2011. Report of the Working Group on Fisheries Acoustics Science and Technology (WGFAST), 10–13 May 2011, Reykjavík, Iceland. ICES CM 2011/SSGESST:12. 37 pp.

ICES. 2012. Report of the Working Group on Fisheries Acoustics, Science and Technology (WGFAST), 23–27 April 2012, Brest, France. ICES CM 2012/SSGESST:09. 38 pp.

ICES. 2013. Report of the Working Group on Fisheries Acoustics Science and Technology (WGFAST), 16–19 April 2013, San Sebastian, Spain. ICES CM 2013/SSGESST:09. 42 pp.

ICES. 2014. Report of the Working Group on Fisheries Acoustics, Science and Tech-nology (WGFAST), 6–9 May 2014, New Bedford, Massachusetts, USA. ICES CM 2014/SSGESST:07. 41 pp.

Jaffe, J.S., Ohman, M.D., De Robertis, A., 1998. OASIS in the sea: measurement of the acoustic reflectivity of zooplankton with concurrent optical imaging. Deep-Sea Res. II 45, 1239–1253.

Jain, R., Kasturi, R., Schunck, B.G., 1995. Machine Vision. Vol. 5. McGraw-Hill, New York, p. 549.

Johnson, R.K., 1977. Sound scattering from a fluid sphere revisited. J. Acoust. Soc. Am. 61 (2), 375–377.

Lavery, A.C., Chu, D., Moum, J.N., 2010. Measurements of acoustic scattering from zooplankton and oceanic microstructure using broadband echosounder. ICES J. Mar. Sci. 67, 379–394.

Lavery, A.C., Stanton, T.K., McGehee, D.E., Chu, D., 2002. Three-dimensional modeling of acoustic backscattering from fluid-like zooplankton. J. Acoust. Soc. Am. 111, 3304–3326.

Lavery, A.C., Wiebe, P.H., Stanton, T.K., Lawson, G.L., Benfield, M.C., Copley, N., 2007. Determining dominant scatterers of sound in mixed zooplankton populations. J. Acoust. Soc. Am. 122 (6), 3304–3326.

Macaulay, G.J., Kloser, R.J., Ryan, T.E., 2013. In situ target strength estimates of visually verified orange roughy. ICES J. Mar. Sci. 70 (1), 215–222.

Oppenheim, A.V., Schafer, R.W., Buck, J.R., 1999. Discrete-Time Signal Processing, second ed. Prentice-Hall, Inc., Upper Saddle River, NJ, p. 796.

Pieper, R.E., Holliday, D.V., Kleppel, G.S., 1990. Quantitative zooplankton distributions from multifrequency acoustics. J. Plankton Res. 12 (2), 433–441.

Powell, J.R., Ohman, M.D., 2012. Use of glider-class acoustic Doppler profilers for estimating zooplankton biomass. J. Plankton Res. 34 (6), 563–568.

Roberts, P.L.D., Jaffe, J.S., 2007. Multiple angle acoustic classification of zooplankton. J. Acoust. Soc. Am. 121 (4), 2060–2070.

Roberts, P.L.D., Jaffe, J.S., 2008. Classification of live, untethered zooplankton from observations of multiple-angle acoustic scatter. J. Acoust. Soc. Am. 124 (2), 796–802.

Roberts, P.L.D., Jaffe, J.S., Trivedi, M.M., 2009. A multiview, multimodal fusion framework for classifying small marine animals with an opto-acoustic imaging system. In: Applications of Workshop on Computer Vision, WACV, 7–8 December 2009, p. 1:6.

Ross, T., Lawson, G., 2009. Long-term broadband acoustic observations of zooplankton scattering layers in Saanich Inlet, British Columbia. J. Acoust. Soc. Am. 125, 2551.

Ryan, T.E., Kloser, R.J., Macaulay, G.J., 2009. Measurement and visual verification of fish target strength using an acoustical-optical system attached to a trawlnet. ICES J. Mar. Sci. 66, 1238–1244.

San Martin, E., Harris, R.P., Irigoien, X., 2006. Latitudinal variation in plankton size spectra in the Atlantic Ocean. Deep-Sea Res. II 53, 1560–1572.

Settles, G.S., 2001. Schlieren and Shadowgraph Techniques: Visualizing Phenomena in Transparent Media. Springer, Germany, p. 376.

Sosik, H.M., Olson, R.J., 2007. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. Limnol. Oceanogr. Methods 5, 204–216.

Stanton, T.K., 2012. 30 years of advances in active bioacoustics: *A personal perspective*. Methods Oceanogr. 1–2, 49–77.

Stanton, T.K., Chu, D., 2008. Calibration of broadband active acoustic systems using a single standard spherical target. J. Acoust. Soc. Am. 124 (1), 128–136.

Stanton, T.K., Chu, D., Reeder, D.B., 2004. Non-Rayleigh acoustic scattering characteristics of individual fish and zooplankton. IEEE J. Ocean. Eng. 29 (2), 260–268.

Stanton, T.K., Chu, D., Wiebe, P.H., 1996. Acoustic scattering characteristics of several zooplankton groups. ICES J. Mar. Sci. 53, 289–295.

Stanton, T.K., Chu, D., Wiebe, P.H., 1998a. Sound scattering by several zooplankton groups. II. Scattering models. J. Acoust. Soc. Am. 103 (1), 236–253.

Stanton, T.K., Wiebe, P.H., Chu, D., 1998b. Differences between sound scattering by weakly scattering spheres and finite-length cylinders with applications to sound scattering by zooplankton. J. Acoust. Soc. Am. 103 (1), 254–264.

Sutor, M., Cowles, T.J., Peterson, W.T., Lamb, J., 2005. Comparison of acoustic and net sampling systems to determine patterns in zooplankton distribution. J. Geophys. Res. 110, C10S16.

Thompson, G.A., Dinofrio, E.O., Alder, V.A., 2013. Structure, abundance and biomass size spectra of copepods and other zooplankton communities in upper waters of the Southern Atlantic Ocean during summer. J. Plankton Res. 0 (0), 1–20.

Wiebe, P.H., Greene, C.H., Stanton, T.K., Burczynski, J., 1990. Sound scattering by live zooplankton and micronekton: Empirical studies with a dual-beam acoustical system. J. Acoust. Soc. Am. 88 (5), 2346–2360.

Wiebe, P.H., Mountain, D.G., Stanton, T.K., Greene, C.H., Lough, G., Kaartvedt, S., Dawson, J., Copley, N., 1996. Acoustical study of plankton on Georges Bank and the relationship between volume backscattering strength and the taxonomic composition of the plankton. Deep-Sea Res. II 43 (7–8), 1971–2001.

Williams, K., Towler, R., Wilson, C., 2010. Cam-trawl: a combination trawl and stereo-camera system. Sea Technol. 51 (12), 45–50.