LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 10, 2012, 1–9

Improving semiautomated zooplankton classification using an internal control and different imaging devices

Eneko Bachiller, Jose Antonio Fernandes and Xabier Irigoien*
Marine Research Division, AZTI Foundation, Herrera Kaia Portualdea, z/g 20110, Pasaia, Spain

Abstract

The rapid development of image-based methods for counting and classifying zooplankton has made it possible to analyze large numbers of samples in a semiautomated way. However, using semiautomated methods to deal with hundreds of samples increases the risk of propagating errors during the procedure. Furthermore, classification methods based on training sets require constant validation to ensure that systematic errors do not affect the results. In this study, we propose using an internal control to check the quality of the procedure for counting and classifying zooplankton. We also evaluate the advantages and disadvantages of two different laboratory imaging devices (scanner and photographic camera) at two resolutions (4800 dpi and 8500 dpi).

Since Victor Hensen took the first plankton samples in 1887 (Benfield et al. 2007), microscopic analysis of preserved samples has been the main tool in plankton research. Nevertheless, manual microscopic analysis requires large numbers of individuals to be subsampled, counted, and sorted into taxonomic groups; a work-intensive task that contrasts with the rapid acquisition and storage of physical oceanographic data.

The influence of the expert (Culverhouse et al. 2003; Benfield et al. 2007) and the limited number of samples that can be accurately processed in a cost-effective way are two disadvantages that have increased the interest in new approaches (Tang et al. 1998; Alcaraz et al. 2003; Grosjean et al. 2004; Boyra et al. 2005; Culverhouse et al. 2006; Irigoien et al. 2006; Benfield et al. 2007; Bell and Hopcroft 2008; Gislason and Silva 2009; MacLeod et al. 2010). Automated identification allows more samples to be processed faster, with much less effort, and increases the spatial and temporal resolution of the studies (Benfield et al. 2007; MacLeod et al. 2010). There is also the possibility of saving plankton sample records in a digital format, which prevents information from being lost due to either deterioration of the preservative (Ortner et al. 1979; Ortner et al. 1981; Leakey et al. 1994; Alcaraz et al. 2003; Zarauz 2007) or sample handling (Benfield et al. 2007).

However, the very same advantage of semiautomatic methods of being able to process many samples in less time also increases the risk of propagating errors. There are two main sources of possible error:

First, because samples are processed in large batches, any errors in the procedure, either in sample preparation (subsampling, staining, etc.) or when data are introduced into the classification system (dilution factors, sampled volumes), are rapidly propagated to a batch of results, which can be as large as a whole survey. Obviously, this risk also exists when samples are counted manually, but in this case, samples for zooplankton counting are usually prepared and analyzed individually. Therefore, the risk of propagating an error is reduced, although the analysis takes more time.

A second source of error can appear at the classification stage. In the automatic classification procedures used nowadays, all the individuals have to be attributed to one of the classes defined in the training set (Grosjean et al. 2004). This error is evaluated and minimized when the training set is set up properly (Grosjean et al. 2004; Fernandes et al. 2009). However, a high number of samples implies high amount of vignettes so that rare items more susceptible to be unnoticeably classified as false positives in robust categories. Hence, when the training set is applied to a large number of samples,

*Corresponding author: E-mail: ebachiller@mail.com

Acknowledgments

We are grateful to the captain and crew of the research vessel R/V *Emma Bardán* and to the onboard scientists and analysts for their support during the sampling. Very special thanks are due to P. Grosjean and K. Denis (University of Mons, Belgium) for their help and advice on image analysis, to D. Conway (Plymouth Marine Laboratory, UK) for his help with taxonomic identification, and to M. Iglesias (scholarship holder in AZTI during summer of 2009) for her efforts in the camera configuration experiments. This research was funded by Ecoanchoa (Department of Agriculture, Fisheries and Food of the Basque Country Government) and Malaspina (Consolider-Ingenio 2010, CSD2008-00077) projects. E. Bachiller is supported by a doctoral fellowship from the Iñaki Goenaga Zentru Teknologikoen Fundazioa (IG-ZTF).

This is contribution no. 552 from AZTI-Tecnalia (Marine Research Division).

DOI 10.4319/lom.2012.10.1

evaluating the consequences in terms of abundance is not always obvious unless some categories are manually counted (i.e., semiautomated classification procedure) to contrast the results (Gorsky et al. 2010). Processing many samples manually before their digitalization allows obtaining information than can be used to detect errors during the image analysis procedure as well as to improve the classifier by adding known vignettes to the training set. However, having to count many samples manually reduces the advantages of the method.

The concentration of organisms on the plate can also cause problems in the automatic counting method. A small subsample underestimates rare organisms, but large subsamples increase the probability of having touching organisms that the system will identify as a single, large organism. This results in a bias in terms of counts, biomass per individual estimates, and the slope of the size spectra. This can be avoided by separating items manually, but again at the expense of the main advantage of semiautomatic counting and identification: speed.

On the other hand, although image analysis seems to be limited to providing information on coarse taxonomic composition (Davis et al. 2005; Gorsky et al. 2010), there are constant requests for higher resolution images in the belief that if the specialist who creates the training set is able to identify individuals better, the classification procedure will perform better. However, the features used by the classification software are not the same as those used by a taxonomist, and improved resolution may not significantly improve classification.

In this study, we propose using an internal control to obtain a reliable internal quality check of the whole zooplankton semiautomatic analysis procedure (both counting and identification). We also compare the taxonomic accuracy of the classifier in relation to the imaging device and image resolution. We test how improved resolution contributes to increasing the number of classes that can be identified in the training set and whether this leads to an overall improvement in classifier accuracy. We also test whether higher resolution images result in a better classification of items in a training set with the same number of classes.

Materials and procedures

Manual counting

All zooplankton individuals and *Amberlite beads* used in this study were first manually counted and identified under a NIKON SMZ645 stereo microscope.

Sample preparation

Samples used for this study were obtained from the ECOANCHOA survey carried out aboard R/V *Emma Bardán* covering the southeast area of the Bay of Biscay in May 2009. Vertical plankton hauls were taken using a 63 µm PairoVET net (2-CalVET nets [Smith et al. 1985]). Samples were preserved in 4% formaldehyde buffered with sodium tetraborate (Harris et al. 2000) and stored in 250 mL jars.

Control beads (internal control)

A station at 43°33′N and 2°10′W was selected for applying the proposed internal control method, which consists in adding a previously known amount of control beads to the plankton sample bottle in order to detect any anomalies during the process. The control beads should behave similarly to the zooplankton when the sample is stirred to avoid subsampling artefacts (Table 1). As a preliminary step, we tested different materials (metallic and glass micro-marbles [www.blockheadstamps.com] and Amberlite™ XAD-2 Polymeric Adsorbent [www.supelco.com]). The sinking rate of the glass and metallic micro-marbles appeared to be too high, whereas Amberlite beads had a slower sinking rate.

One milliliter wet Amberlite beads (previously filtered through a 500 µm sieve) was manually counted under a microscope. Three replicates were counted manually and independently, resulting in an abundance of 2756 (SD ± 84) Amberlite beads per milliliter. Subsequently, 2 mL of wet Amberlite beads were added to a 250 mL zooplankton sample bottle (whole sample) to obtain a concentration of 22 Amberlite beads per mL (Fig. 1).

Once the Amberlite beads had been added, the sample was divided using a Folsom plankton divider to obtain three replicates for each 1, 2, 4, 8, and 16 mL aliquot subsample.

Collection of zooplankton for training set comparisons

We established 36 taxonomic categories (classes) to create the training sets for comparing the image devices and resolu-

Table 1. Physical properties of Amberlite XAD-2 resin (from *SUPELCO®*).

Appearance: Hard, spherical, opaque beads
Solids: 55%
Porosity: 0.41 mL pore mL^{-1} bead
Surface area (min.): 300 m^2 g^{-1}
Mean pore diameter: 90Å
True wet density: 1.02 g mL^{-1}
Skeletal density: 1.08 g mL^{-1}
Bulk density: 640 g L^{-1}

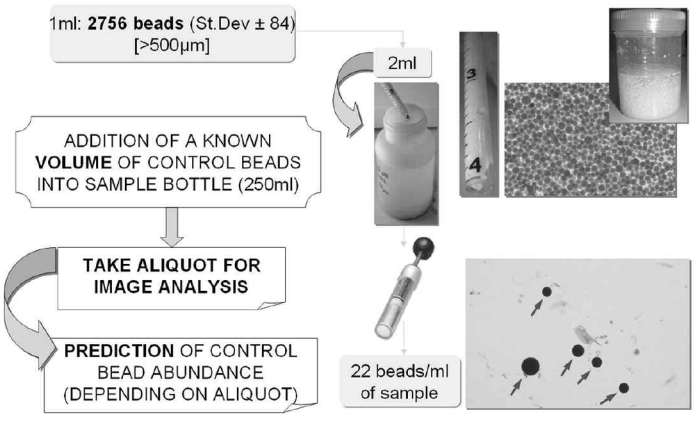

Fig. 1. Schematic diagram of the proposed internal control methodology that consists in adding Amberlite beads to a zooplankton sample bottle.

tions. Forty individuals per class were separated under the microscope from random samples of the same survey.

All samples were stained for 24 h with 1 mL Eosin (5 gr L^{-1}). This stains the cell cytoplasm and muscle protein and thus improves organism identification and reduces the number of artefacts (inorganic particles) to be counted and classified by the system. The samples were then poured onto polystyrene plates (126 × 84 mm) for later analysis with different methods.

Hardware

Scanner system

Previously prepared zooplankton sample plates were scanned at 4800 dpi resolution using an EPSON V750 PRO scanner (with VueScan Professional Edition 8.5.02 software).

Digital camera system

The camera system consisted in a copy stand & tilting arm (model *Kaiser RS1 5511*) and a micrometric sliding plate (model *Manfrotto 454*) with a Canon EOS 450D digital camera controlled by the computer. In addition, the optic consisted in a macro lens (model *Tamron SP 90mm F/2.8 Di EOS*) and extension tubes (model *KENKO 3 Ring DG P/Canon EOS*, 12 + 20 + 36 mm). A uniform background light was provided by a white LED backlight (model *BIBL-w130/110*) (Fig. 2).

This configuration allows different resolutions to be obtained depending on the focus distance (Table 2). A vignette matrix was created from images taken with the digital camera at different resolutions to obtain a visual guide that defines the clarity of the vignettes for when the training set is developed (Fig. 3). The amount of photographs that is needed to obtain the whole sample plate area is different depending on the configuration of the camera system. Excluding borders, the plate we used had an area of 10168 mm^2. The plates were marked with squares of different sizes (depending on the resolution) to avoid overlaps. Table 2 shows the number of photographs that were taken to cover as much plate area as possible at different resolutions, as well as the real percentage of the plate area photographed. A correction factor was applied in all analyses to standardize the plate area covered by the photos. All configurations were checked with a calibration graticule to measure the real resolution obtained. The 4800 dpi resolution configuration was selected to compare the camera and scanner results, and 8500 dpi was chosen to evaluate possible advantages at the highest resolution obtained with the camera.

Images from both the scanner and camera at selected resolutions were analyzed and processed with Zooimage according to Grosjean and Denis (2007).

Training sets

The training set to evaluate the effectiveness of the internal control was built from organisms identified in the images obtained with the different imaging devices.

To compare the effect of the image resolution on accuracy a training set was built using organisms identified under the microscope before taking the images with the different devices. Therefore, the three training set categories contained exactly the same organisms (30 individuals per class) imaged with different systems.

Fig. 2. Digital camera system: (1) Copy stand & tilting arm (model Kaiser RS1 5511); (2) Micrometric sliding plate (model Manfrotto 454); (3) Camera (model CANON EOS 450D); (4) Macro lens (model Tamron SP 90mm F/2.8 Di EOS); (5) Extension tubes (model KENKO 3 Ring DG P/Canon EOS, 12 + 20 + 36mm); (6) Uniform white LED backlight (model BIBL-w130/110); (7) Connected computer system.

Assessment

Subsample volume effect

If the subsample is small, rare organisms may be underestimated, but if the subsample is large, there is a higher probability of having touching particles that cannot be distinguished by the system. Using an internal control allows us to evaluate the magnitude of these two risks. In the conditions of our example, we observed an overestimation of particle abundance in almost all cases, and the 4 mL aliquot showed the smallest difference between the measured and expected (theoretical) Amberlite bead concentrations (Fig. 4). The scanned images resulted in underestimated bead abundances in aliquots larger than 4 mL, whereas the rest of the methodologies and resolutions still led to overestimation, including manual counting. An ANOVA analysis showed significant differences comparing the different methodologies ($P < 0.01$). However, Dunnett's paired comparison of means with manual

Table 2. Possible configurations of the camera system to obtain different image resolutions. In all cases, the camera was configured at ISO1600, f/22 at 1/80sec. 4800 dpi and 8500 dpi are highlighted on the table as they were the resolutions used in this study.

Resolution (dpi)	Camera position in stand (cm)	Macro lens mm (extension tube)	Magnification factor	Pixels per mm (±SD)	Each photograph area's width (mm)	Each photosgraph area's height (mm)	Nr photos inside plate (NO overlap)	Covered plate area with photos (%)
800	73.8	12	9x	33.7 ± 0.8	126.9	84.6	1	105.6
1200	60.3	20	12x	47.2 ± 1.2	90.6	60.4	1	53.8
1600	50.3	20	17x	65.8 ± 0.4	64.9	43.3	2	55.2
2000	47	36	22x	83.7 ± 0.8	51.1	34.0	4	68.4
2400	44.6	36	24x	94 ± 0.9	45.4	30.3	4	52.4
2800	43.7	48	28x	108.3 ± 1.5	39.4	26.3	8	81.6
3200	42	56	33x	126.8 ± 0.8	33.7	22.5	9	66.9
3600	40.5	56	37x	145.2 ± 3.4	29.4	19.6	10	56.8
4000	41	68	40x	150.2 ± 4.2	28.4	19.0	12	63.7
4400	39.6	68	46x	177 ± 1.7	24.1	16.1	16	61.1
4800	**38.9**	**68**	**50x**	**193.7 ± 1.4**	**22.1**	**14.7**	**25**	**79.8**
5200	38.5	68	54x	206.5 ± 1.5	20.7	13.8	25	70.2
5600	38.1	68	57x	220.8 ± 1.2	19.3	12.9	25	61.3
6000	37.8	68	61x	235.3 ± 0.8	18.2	12.1	32	69.1
6200	37.6	68	64x	248.3 ± 1.2	17.2	11.5	32	62.1
6400	37.5	68	67x	255.8 ± 2	16.7	11.1	36	65.8
6800	37.4	68	69x	269.4 ± 0.9	15.9	10.6	36	59.3
7000	37.2	68	71x	277.3 ± 0.9	15.4	10.3	42	65.3
7200	37.3	68	74x	285.2 ± 1	15.0	10.0	42	61.8
7600	37.3	68	77x	301 ± 1	14.2	9.5	60	79.2
8500	**37.2**	**68**	**80x**	**337.6 ± 1.2**	**12.7**	**8.5**	**78**	**82.4**

counts as reference showed a significant difference for the theoretically predicted values as well as for the scanner, whereas results obtained with the camera showed no significant difference. The *P* values of the multiple comparisons are presented in Table 3.

The error in the theoretically predicted abundance estimation of Amberlite beads was plotted against the percentage of plate area covered by organisms (pixels with organisms/total pixels). This plot showed an inflexion point when around 1% of the plate was covered with organisms for all the imaging methods used (Fig. 5). The scanner was the imaging device that obtained the results closest to the predicted values. High resolution camera images (i.e., 8500 dpi) also obtained estimations that were relatively close to the theoretical estimations at 2% plate coverage. A statistically significant difference (ANOVA $P = 0.01$) was found when the percentage of error in abundance (from the theoretically predicted value) was compared between methods.

Taxonomic classification with different methods

Fig. 3 shows the effect of increased resolution in the visual aspect of individual organisms in the vignettes. Several groups could only be detected at the higher resolutions obtained with the camera, i.e., *Oithona* spp., juvenile copepods, and copepod naupli (Table 4). Therefore, resolution appears as a potential limiting factor when the training set is set up directly from images. However, the highest total accuracy of training set obtained for all groups with the high resolution images (72.6% for 44 categories) was not significantly different (ANOVA $P = 0.65$) from that obtained with the 4800 dpi resolution (69.4% for 41 categories in the case of the scanner, and 69.1% for 41 categories for the camera).

To check whether improved resolution had an effect on the classification itself (after setting up the training set), we compared three training sets with identical categories (28) and numbers of individuals per category (30) imaged after classification under the microscope. The classifier showed slightly higher sensitivity (i.e., true positives divided by true positives plus false negatives) when high resolution camera images were used (Table 5), but differences between classifier sensitivities obtained with different resolutions were not significant (ANOVA $P = 0.62$).

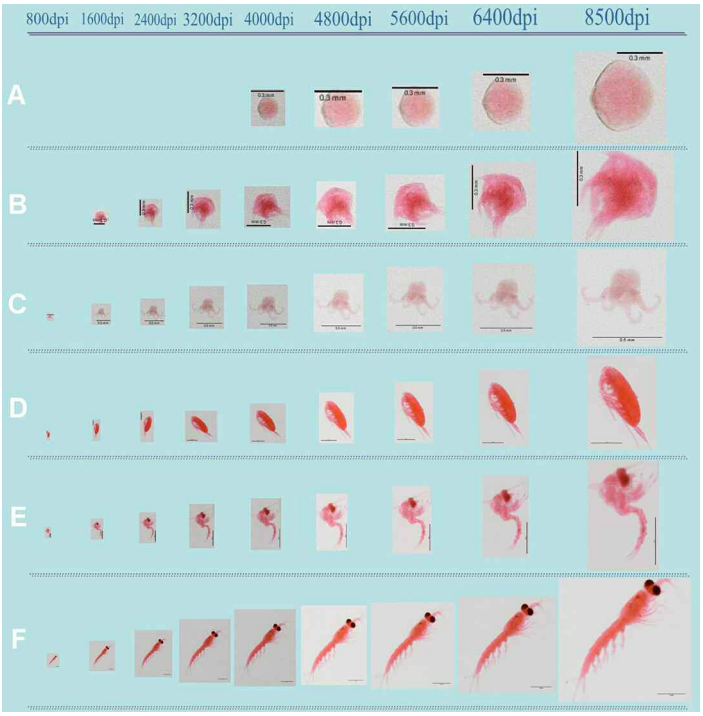

Fig. 3. The vignettes show different clarities depending on the resolution of the images they were extracted from. Therefore, obtaining the training set is more or less difficult depending on the size and abundance of vignettes and the required classification accuracy. All these vignettes were extracted from images taken with a digital camera. (A) Bivalve veliger; (B) Cirriped nauplius; (C) Cephalopoda larva; (D) *Calanus sp.*; (E) ZOEA larva; (F) Euphausiid.

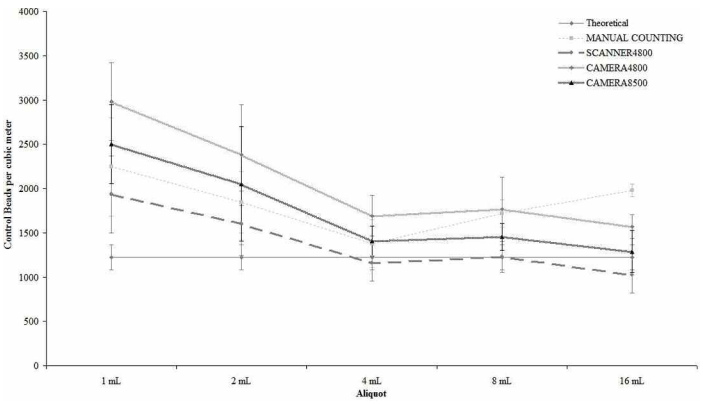

Fig. 4. Abundance of Amberlite beads per m^3 with corresponding standard deviations obtained in different aliquot processing images with different methodologies and resolutions. The theoretically predicted value did not change because the same abundance per m^3 should be observed if the subsampling is carried out correctly.

Discussion

Internal control as a contribution to zooplankton studies

Semiautomated procedures allow processing more samples in less time, but this also implies less exhaustive control of results so that they are more susceptible to the propagation of errors in an unnoticeable manner for the expert. Furthermore, replicate random variability (Barnes and Marshall 1951; Downing et al. 1987) and subsampling can also affect the results. Amberlite beads are a reasonable internal control for semiautomated counting and automatic classification procedures. In addition to procedure error control, they provide a tool for evaluating the appropriateness of the sample treatment procedures according to the objectives because they have a very similar density and behavior to the zooplankton collected in bottles to obtain the same level of susceptibility to be sampled. They can be used to evaluate whether the subsample volumes are sufficient for rare organisms, whether the percentage of touching particles is reasonable and whether the training set is equally effective at all stations.

Table 3. *P* values of Dunnett's Multiple Comparison of means.

Control group = Manual	**α = 0.05**
Level	***P***
Theoretical	0.002
Scanner 4800 dpi	0.04
Camera 4800 dpi	0.4
Camera 8500 dpi	0.9

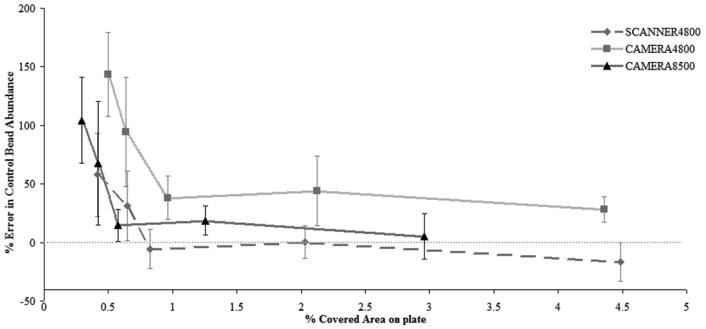

Fig. 5. Percentage of error in Amberlite bead abundance (from theoretically predicted values) plotted against the percentage of area covered by particles on the sample plate. Standard deviations have also been plotted.

Defining the appropriate resolution

Potentially, higher image resolutions can improve classifiers by producing better defined training sets, which has three advantages: 1) the expert can identify more detailed categories, 2) the "shape" parameters measured at a higher resolution reveal differences not detected at lower resolutions, and 3) the higher resolution allows more "shape" parameters to be measured.

High resolutions allow the expert to consider more groups in the training set that would not have been detected at lower resolutions. This is especially true for small plankton groups (Table 4). The essential identification features are reflected well in high resolution vignettes (Fig. 3), which makes manual vignette classification much easier and quicker, even for small

Table 4. Number of recognized vignettes (and percentages in relation to the manually counted vignettes) when the training set was obtained with different methodologies. Artefact vignettes were not presented.

	MANUAL COUNTING	SCANNER 4800 dpi	CAMERA 4800 dpi	CAMERA 8500 dpi
	Total nº individuals	Vignettes (%)	Vignettes (%)	Vignettes (%)
Noctiluca scintillans (Phytopl.)	50	0 (0)	0 (0)	1 (2)
Anthomedusae	56	21 (37.5)	11 (19.6)	18 (32.1)
Diphyidae family (Siphonophorae)	45	8 (17.8)	4 (8.9)	9 (20)
Gastropod veliger (Prosobranchia)	48	29 (60.4)	37 (77.1)	46 (95.8)
Bivalve veliger larvae	42	17 (40.5)	18 (42.9)	20 (47.6)
Podon sp. (Cladocera)	46	30 (65.2)	34 (73.9)	35 (76.1)
Evadne spp. (Cladocera)	46	13 (28.3)	7 (15.2)	11 (23.9)
Acartia spp.	48	31 (64.6)	29 (60.4)	36 (75)
Candacia armata	41	35 (85.4)	35 (85.4)	40 (97.6)
Centropages spp.	51	32 (62.8)	44 (86.3)	50 (98)
Temora longicornis	45	35 (77.8)	33 (73.3)	41 (91.1)
Calanidae family	43	40 (93)	40 (93)	42 (97.7)
Small Calanoid	48	38 (79.2)	43 (89.6)	43 (89.6)
Oithona spp.	45	0 (0)	3 (6.7)	8 (17.8)
Oncaea spp.	48	33 (68.8)	39 (81.3)	40 (83.3)
Corycaeus anglicus	45	26 (57.8)	26 (57.8)	33 (73.3)
Euterpina acutifrons	44	27 (61.4)	34 (77.3)	36 (81.8)
Microsetella spp.	43	28 (65.1)	31 (72.1)	34 (79.1)
Juvenile copepod	47	0 (0)	6 (12.8)	13 (27.7)
Copepod naupli	76	0 (0)	4 (5.3)	11 (14.5)
Cirripedia Naupli	43	31 (72.1)	31 (72.1)	37 (86.1)
Cirripedia	72	58 (80.6)	66 (91.7)	70 (97.2)
Mysida	48	43 (89.6)	48 (100)	48 (100)
Amphipoda	48	41 (85.4)	47 (97.9)	48 (100)
Calyptopis larvae (Euphausiacea)	52	43 (82.7)	35 (67.3)	52 (100)
Furcilia larvae (Euphausiacea)	80	74 (92.5)	77 (96.3)	80 (100)
Decapoda	40	36 (90)	34 (85)	38 (95)
ZOEA larvae (Decapoda)	38	34 (89.5)	37 (97.4)	38 (100)
Megalopa larvae (Decapoda)	7	7 (100)	7 (100)	7 (100)
Polychaeta larvae	51	45 (88.2)	38 (74.5)	46 (90.2)
Sagitta setosa (Chaetognatha)	54	46 (85.2)	30 (55.6)	53 (98.2)
Fritillaria sp.	42	11 (26.2)	6 (14.3)	20 (47.6)
Oikopleura spp.	43	8 (18.6)	0 (0)	10 (23.3)
Doliolida (Thaliacea)	62	51 (82.3)	38 (61.3)	58 (93.6)
Lancet larvae (Cephalochordata)	43	38 (88.4)	31 (72.1)	40 (93)
Engraulis encrasicolus eggs	48	47 (97.9)	36 (75)	47 (97.9)
Unidentified fish eggs	64	59 (92.2)	64 (100)	64 (100)
Clupeid larvae	83	73 (87.9)	24 (28.9)	83 (100)
Unidentified fish larvae	30	27 (90)	19 (63.3)	27 (90)

zooplankton groups. Therefore, a high resolution allows the expert to reduce error when the training set is build up from vignettes taken directly in the sample since more details could be seen on vignettes. However, the extent to which having more and better defined categories contributes to the overall accuracy of the classifier depends on the complexity and characteristics of the plankton community. More classes do not necessarily improve the overall accuracy (Fernandes et al. 2009) and the effect will depend on whether the higher resolution improves the identification of broad categories that could tend to include organisms from many other categories. On the other hand, higher resolution implies longer digitizing

Table 5. Classifier sensitivity percentages for the same training set at different resolutions. All groups were limited to 30 vignettes to avoid influences other than resolution on the results.

	TRAINING SET	SCANNER 4800 dpi	CAMERA 4800 dpi	CAMERA 8500 dpi
	Total nº vignettes	Sensitivity %	Sensitivity %	Sensitivity %
Gastropod veliger (Prosobranchia)	30	90	93.3	86.7
Podon sp. (Cladocera)	30	73.3	73.3	80
Acartia spp.	30	80	80	93.3
Candacia armata	30	70	80	83.3
Centropages spp.	30	80	76.7	73.3
Temora longicornis	30	76.7	56.7	83.3
Calanidae family	30	86.7	90	96.7
Small Calanoid	30	46.7	66.7	70
Oncaea spp.	30	80	86.7	90
Corycaeus anglicus	30	76.7	53.3	63.3
Euterpina acutifrons	30	90	76.7	73.3
Microsetella spp.	30	86.7	80	83.3
Cirripedia Naupli	30	83.3	80	90
Cirripedia	30	83.3	86.7	83.3
Mysida	30	70	73.3	70
Amphipoda	30	20	46.7	36.7
Calyptopis larvae (Euphausiacea)	30	90	83.3	80
Furcilia larvae (Euphausiacea)	30	83.3	76.7	83.3
Decapoda	30	16.7	16.7	36.7
ZOEA larvae (Decapoda)	30	53.3	73.3	53.3
Polychaeta larvae	30	40	16.7	36.7
Sagitta setosa (Chaetognatha)	30	90	66.7	90
Doliolida	30	90	73.3	76.7
Lancet larvae (Cephalochordata)	30	100	90	96.7
Engraulis encrasicolus eggs	30	100	96.7	100
Unidentified fish eggs	30	93.3	96.7	93.3
Clupeid larvae	30	90	46.7	86.7
Unidentified fish larvae	30	50	73.3	66.7
MEAN SENSITIVITY %		**74.6**	**71.8**	**77**

time (i.e., more photographs), more space on hardware, and slower data processing. Hence the aim of each study should consider all good and bad sides of using very high resolutions. For studies targeting a particular taxonomic category, increased resolution will help to improve accuracy by allowing much better training sets to be developed, however it is not obvious it to be the case for general studies.

In our study, once the training set had been built, the resolution of the images did not significantly modify the accuracy of the classifier. Although it is sometimes difficult to understand, the way the classification algorithms use particle measurements to make classifications is not the same as the algorithms our brain uses. The shape measuring parameters and the way the distribution probability changed as a function of these values did not differ significantly with resolution. However, high resolution allowed obtaining training sets with more categories as well as higher quality of vignettes, especially favorable for experts when performing the machine learning. Another field to be explored is the additional features that could be measured in higher-resolution images that would improve classification, e.g., antennas (and even main antennulas) included in the perimeter of the particles, an improved shape recognition of gelatinous plankton, etc. Nevertheless, the Zooimage filter used already measures 20 different parameters such as the major and minor axes, ECD, area, and the gray scale of the pixels (Fernandes et al. 2009).

Comments and recommendations

Because the proposed control bead methodology allows an easy, cheap, systematic, and reliable quality check of the whole zooplankton analyzing procedure—moreover without altering the sample behavior—including such an internal control for further planktonic research should be considered.

On the other hand, the results obtained from images taken with a digital camera are at least as good as those obtained with a scanner, both in terms of taxonomic accuracy and cost effectiveness. Nevertheless, the digital camera offers new possibilities, such as easiness, speed, and the possibility of working with particularly high resolutions if necessary depending on the aims of the project. Using digital cameras allows a wider range of resolutions, which can be useful for analyzing the abundance of small organisms not detected with the limited resolution of the scanner as well as for obtaining better defined training sets. In addition, some zooplankton groups that can only be detected under the microscope are also represented in higher-resolution images. Furthermore, the image acquisition speed allows digital cameras to be used onboard the research vessel.

However, a digital camera set up is less robust than a scanner-based system because the parameters and configurations (focus, focus distance, number of photographs to take) can be unintentionally altered, and therefore needs to be handled more carefully. As we have seen, higher resolutions imply longer processing time and effort as well as higher computing power requirement, and they do not necessarily improve the classifiers. On the other hand, more taxonomic groups can be automatically identified and the machine-learning process can be better made with high quality vignettes suggesting long-term cost-effectiveness. The aims of the project should define the adequate methodology.

References

Alcaraz, M., E. Saiz, A. Calbet, I. Trepat, and E. Broglio. 2003. Estimating zooplankton biomass through image analysis. Mar. Biol. 143:307-315 [doi:10.1007/s00227-003-1094-8].

Barnes, H., and S. M. Marshall. 1951. On the variability of replicate plankton samples and some applications of 'contagious' series to the statistical distribution of catches over restricted periods. J. Mar. Biol. Assoc. U.K. 30:233-263 [doi:10.1017/S002531540001273X].

Bell, J. L., and R. R. Hopcroft. 2008. Assessment of ZooImage as a tool for the classification of zooplankton. J. Plank. Res. 30:1351-1367 [doi:10.1093/plankt/fbn092].

Benfield, M. C., and others. 2007. RAPID: research on automated plankton identification. Oceanography 20:172-187 [doi:10.5670/oceanog.2007.63].

Boyra, G., X. Irigoien, A. Aristegieta, and I. Arregi. 2005. Plankton Visual Analyser, GLOBEC International Newsletter 11: 9-10.

Culverhouse, P. F., R. Williams, B. Reguera, V. Herry, and S. González-Gil. 2003. Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. Mar. Ecol. Progr. Ser. 247:17-25 [doi:10.3354/meps247017].

———, and others. 2006. Automatic image analysis of plankton: future perspectives. Mar. Ecol. Progr. Ser. 312:297-309 [doi:10.3354/meps312297].

Davis, C. S., F. T. Thwaites, S. M. Gallager, and Q. Hu. 2005. A three-axis fast-tow digital Video Plankton Recorder for rapid surveys of plankton taxa and hydrography. Limnol. Oceanogr. Meth. 3:59-74 [doi:10.4319/lom.2005.3.59].

Downing, J. A., M. Pérusse, and Y. Frenette. 1987. Effect of interreplicate variance on zooplankton sampling design and data analysis. Limnol. Oceanogr. 32:673-680 [doi:10.4319/lo.1987.32.3.0673].

Fernandes, J. A., X. Irigoien, G. Boyra, J.A. Lozano, and I. Inza. 2009. Optimizing the number of classes in automated zooplankton classification. J. Plank. Res. 31:19-29 [doi:10.1093/plankt/fbn098].

Gislason, A., and T. Silva. 2009. Comparison between automated analysis of zooplankton using ZooImage and traditional methodology. J. Plank. Res. 31:1505-1516 [doi:10.1093/plankt/fbp094].

Gorsky, G., and others. 2010. Digital zooplankton image analysis using the ZooScan integrated system. J. Plank. Res. 32:285-303 [doi:10.1093/plankt/fbp124].

Grosjean, P., M. Picheral, C. Warembourg, and G. Gorsky. 2004. Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. ICES J. Mar. Sci. 61:518-525 [doi:10.1016/j.icesjms.2004.03.012].

———, and K. Denis. 2007. Zoo/PhytoImage version 1.2-0. User's manual. http://www.sciviews.org/zooimage, last update on September 9, 2011.

Harris, R. P., P. H. Wiebe, J. Lenz, H. R. Skjoldal, and M. Huntley. 2000. Zooplankton methodology manual. Academic Press.

Irigoien, X., P. Grosjean, and A. Lopez Urrutia. 2006. Image Analysis to count and identify zooplankton. In Report of a GLOBEC/SPACC workshop on image analysis to count and identify zooplankton, November 2005, San Sebastian, Spain. 21 pp. http://www.globec.org/structure/regional/spacc/image_wksp/contents.htm, last access on March 26, 2005.

Leakey, R. J. G., P. H. Burkill, and M. A. Sleigh. 1994. A comparison of fixatives for the estimation of abundance and biovolume of marine planktonic ciliate populations. J. Plank. Res. 16:375 [doi:10.1093/plankt/16.4.375].

MacLeod, N., M. Benfield, and P. Culverhouse. 2010. Time to automate identification. Nature 467:154-155 [doi:10.1038/467154a].

Ortner, P. B., S. R. Cummings, R. P. Aftring, and H. E. Edgerton. 1979. Silhouette photography of oceanic zooplankton. Nature 277:50-51 [doi:10.1038/277050a0].

———, L. C. Hill, and H. E. Edgerton. 1981. In-situ silhouette photography of Gulf Stream zooplankton. Deep Sea Res. A. 28:1569-1576 [doi:10.1016/0198-0149(81)90098-4].

Smith, P. E., W. Flerx, and R. H. Hewitt. 1985. The CalCOFI Vertical Egg Tow (CalVET) Net. NOAA technical report NMFS 36. *In* R. Lasker (Ed.), An egg production method for estimating spawning biomass of pelagic fish: Application to the northern anchovy, Engraulis mordax. US Dep Commer.

Tang, X., W. K. Stewart, H. Huang, S. M. Gallager, C. S. Davis, L. Vincent, and M. Marra. 1998. Automatic plankton image recognition. Artific. Intell. Rev. 12:177-199 [doi:10.1023/A:1006517211724].

Zarauz, L. 2007. Description and modelling of plankton biomass distribution in the Bay of Biscay by means of image analysis-based methods. PhD Thesis, Department of Vegetal Biology and Ecology, University of the Basque Country (EHU - UPV).

Submitted 4 March 2011
Revised 13 September 2011
Accepted 17 November 2011