Zoo/PhytoImage, a software for automatic analysis of plankton samples based on R and ImageJ

Kevin Denis[1,*], Xabier Irigoien[2], Romain François[3], Véronique Rousseau[4], Jean-Yves Parent[4], Christiane Lancelot[4] and Philippe Grosjean[1]

1. Numerical Ecology of Aquatic Systems, Mons University, 8 avenue du Champ de Mars, 7000 Mons, Belgium
2. AZTI …, Spain
3. Independent R consultant romainfrancois@free.fr, France
4. Ecology of Aquatic Systems, Université Libre de Bruxelles, Bd du Triomphe, 1050 Bruxelles, Belgium

* Contact author: Kevin.Denis@umh.ac.be

Keywords: Plankton, Image analysis, Machine learning, Ecology, Oceanography

Zoo/PhytoImage (ZooImage and PhytoImage, depending on users) is a complete solution to analyze so-called, numerically fixed plankton samples, that is, samples that were digitized usually by imaging system, either in cultures or at sea. Zoo/PhytoImage (http://www.sciviews.org/zooimage) can import and analyze images obtained with digital cameras (micro- or macrophotographies), with flatbed scanners, with the FlowCAM (http://www.fluidimaging.com), or even other digitizing devices like the Zooscan (http://www.zooscan.com), or underwater cameras, provided an adequate importation plugin is available.

Zoo/PhytoImage segments the images, localize particles and measure them (more than 30 measurements on each particle: size, shape, moments, transparencies, texture, etc.). It also extracts 'vignettes', that are little pictures of each original particle. Taxonomists can then classify manually a subset of these vignettes in different taxonomic groups (with hierarchy between these groups if relevant).

This manually classified subset constitutes the training set used to build a classification algorithm for similar samples, using machine learning algorithms. Zoo/PhytoImage can then process a series of plankton samples in batch. It counts, measures and classifies particles found in the samples and calculates ecologically meaningful variables, like relative or absolute abundances, size spectra and biomasses for each taxon.

The software also provides tools to visualize and assess the performances of the classifiers. It proposes a format to store compressed data on disk, and uses dedicated S3 objects in R for it. Metadata, including series, station, cruise, sampling method and information, digitizing technique, … are also handled by Zoo/PhytoImage.

A GUI eases the process from image importation to export of final results for those who are not familiar with R. For the others, Zoo/PhytoImage is more a toolkit of functions that~~y~~ can assemble in scripts, or in their own R code for complex processing of plankton images. Zoo/PhytoImage has also proven useful in other applications, like bugs counting and classification, or in bacteriology.

Zoo/PhytoImage is partly developed for the AMORE III project (Advanced Modelling and Research on Eutrophication, http://www.ulb.ac.be/assoc/esa/AMORE/objectives.htm) funded by the Belgian Science Policy. It is also a contribution to the SCOR WG130: Automatic visual plankton identification (http://www.scor-wg130.net, Benfield et al. 2007).

References

Benfield, M.C., Ph. Grosjean, Ph. Culverhouse, X. Irigoien, M.E. Sieracki, A. Lopez-Urrutia, H.G. Dam, Q. Hu, C.S. Davis, A. Hansen, C.H. Pilskaln, E. Riseman, H. Schultz, P.E. Utgoff & G. Gorsky (2007). RAPID : Research on Automated Plankton Identification. *Oceanography*, 20(2):12-26.