An Automated Event Detection and Classification System for Abyssal Time-Series Images of Station M, NE Pacific

Danelle E. Cline, Duane R. Edgington
Ken L. Smith, Michael F. Vardaro, Linda Kuhnz, Jacob A. Ellena
Monterey Bay Aquarium Research Institute
7700 Sandholdt Road
Moss Landing, CA 95039 USA

Abstract—**The time-study data collected at the Station M site off the coast of central California includes high quality still-frame images taken in 1-hour time-lapse increments. The approximately 67,000 time-lapse images collected would take an unfeasible amount of time to fully analyze manually, and therefore would benefit from automated analysis. Towards this end, this work is an aid in the significant effort to analyze megafaunal activity and sedimentation events using an adapted version of the Automated Video Event Detection and Classification System (AVEDac) formerly designed by MBARI to analyze video collected from MBARI's remotely operated underwater vehicles (ROVs) video. This paper describes, in general, the automated system that will aid in the abundance and distribution studies of animals at the Station M site.**

I. Introduction

The extensive study conducted at a site called "Station M" (34deg 50'N, 123deg 00'W; 4,100 m depth) off the central coast of California [1] represents perhaps the most detailed investigation of any abyssal area in the world ocean. Sixteen years of time-lapse photography and measurements of sediment community oxygen consumption were continuously collected and analyzed, from 1989 to 2004 [2, 3].

Although a significant effort has gone into analyzing the megafaunal activity and sedimentation events in the ~67,000 frames gathered from time-lapse cameras at the Station M site, it would take an unfeasible amount of time to fully analyze all images manually. Therefore, to assist in the analysis of the time-lapse images taken 1 hour apart, a software-based system that enables automated annotation of these time-lapse images for quantitative studies of benthic boundary layer ecology is being developed by MBARI.

The software system for this application is based on an adaptation of the Automated Video Event Detection and Classification System (AVEDac) formerly designed by MBARI to analyze video collected from MBARI's remotely operated underwater vehicles (ROVs) video. For this application, minor modifications to AVEDac were made to process still images instead of video, for which it was originally designed.

The AVED software was most recently applied to find interesting events from the Eye-in-the-Sea™ video recordings at a benthic site on the MARS cabled observatory [4].

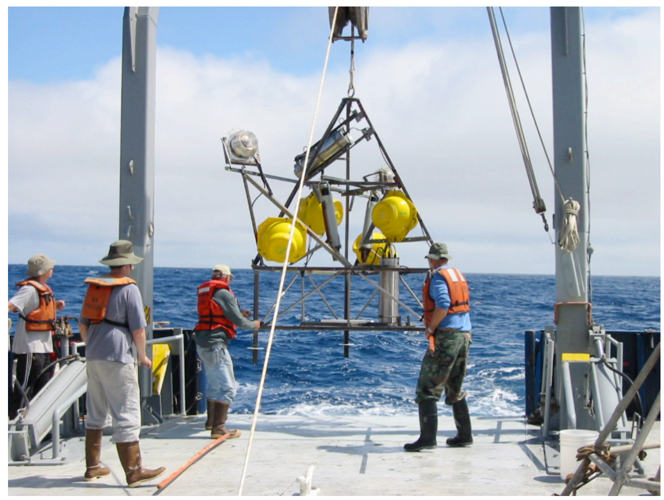


Figure 1 Station M tripod camera deployment

This paper first gives a general overview of a system for automating the detection and subsequent classification of animals found in the abyssal time-series images recorded from Station M. Lastly, future work is discussed.

II. Motivation

Automated classification is a powerful tool in the analysis of abundance and distribution of animals. It allows for software automation to find commonly targeted animals during normal science annotation, leaving the annotator time to focus on other observations. It can also automate the process of finding animals incorrectly identified, or missed entirely on previously analyzed data.

The time-lapse images from Station M are also very well suited to computer classification because of both the nature of the site and the images. The Station M site supports a relatively discrete biological community, making it easier to collect representative images to train the AVEDac classifier. Also, the time-lapse images have nominal variance in lighting, lending themselves to better segmentation of the animals from the background.

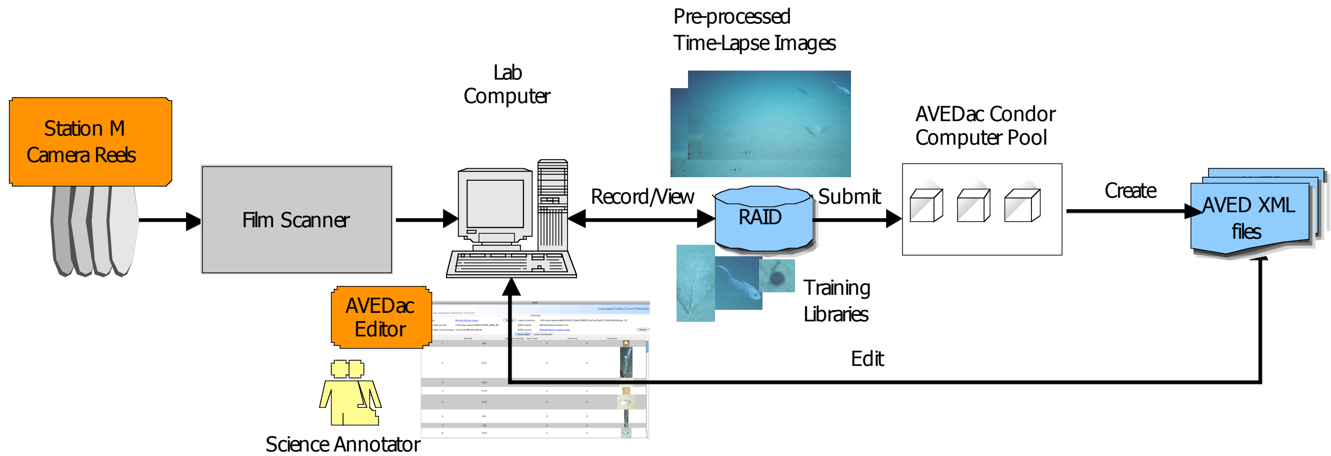


Figure 2 Station M Processing System

III. PROCESSING SYSTEM

Images from Station M from the 1989 to 2004 period are recorded on 35 mm 400-foot reels and because the AVEDac software required images in digital format, reels must be digitized in a format compatible with the AVEDac software. Reels are digitized with a Kodak HR 500 film scanner into 4504x2790 color images at 72 pixels per inch resolution. Once images are in a digital format, pre-processing in the form of a Photoshop batch process of the images removes artifacts left from the scanning process like black bars, a time code overlay, and names files according to an ISO8601 naming convention. These pre-processed images are then stored to a RAID.

After the images are digitized and pre-processed, they are ready for automated processing with the AVEDac software.

Processing video with AVEDac is a compute and data intensive activity that involves a series of processing steps. These steps can be described as a workflow, where each step in the workflow has data input, output and/or control dependencies and each step in this workflow doesn't need to be executed on the same computer.

To execute and manage this workflow, we use a specialized workload management system for compute and data intensive jobs called Condor [5] developed by the University of Wisconsin, Madison.

For this application, a Condor pool of three AVEDac computers was configured and should the need for more processing arise, one can simply install AVEDac on additional computers and add them to the pool.

The AVEDac software analyzes each image searching for interesting events using a neuromorphic software model based on the human vision system. The AVEDac system identifies and tracks potentially interesting visual events spanning multiple image frames based on low-level properties of salient objects. Interesting events are then passed to a Bayesian classifier utilizing a Gaussian mixture model to determine the lowest possible taxonomic category and then analyzed for abundance and distribution. For more information on the AVEDac algorithms see [6].

The output of the AVEDac software is recorded into XML formatted data, with each animal classified into the lowest taxonomic category. These XML results can be edited in the AVED graphical editor for miss-classifications or exported for further analysis in conjunction with ancillary data from the Station M site.

IV. TRAINING LIBRARIES

Classification learning is the process of learning a discrete-valued function. In the case of the AVEDac classifier, it learns from discrete inputs (images in this case). Learning is done through a "training" phase, whereby the training class is defined as a collection of images that represent a particular taxonomic category.

To generate these training classes, a collection of time-lapse images populated with the animals of interest is processed through the AVEDac software. AVEDac is enabled with an option to find all interesting events and extract cropped images of these events. These cropped images are then sorted using a simple graphical interface into representative classes.

The targeted classes for this project include: *Echinocrepis rostrata* (a sea urchin), Macrourids (rattail fish), other fish, *Benthocodon* (a jellyfish) and "Other". The "Other" class includes otherwise uncategorized animals, organic matter, and geological features.

The process of creating training libraries is perhaps the most critical component to the success of the system. Testing to verify the accuracy of the libraries is planned and involves taking images not used to generate the training classes, independently classifying the events in these images manually, then comparing the human versus computer classification.

V. AVEDAC CLASSIFIER IMPROVEMENTS

For this project, improvements were made to the AVEDac computer classification program used to classify benthic species [7]. Although the previous AVEDac classifier performance was good, only grayscale features were employed in its algorithm.

Through a simple addition to the classifier of red, green, and blue color features, significant improvements in the accuracy of several test classes were recently achieved as seen in Tables 1 and 2. We are hopeful that this will improve the image classification but more work is needed to test this improvement against the Station M images.

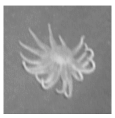
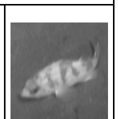
		True Class		
				
		Rathbunaster	Parastichopus leukothele	Other
Estimated Class	Rathbunaster	**91.667**	0.000	0.262
	Parastichopus leukothele	0.000	**68.571**	1.393
	Other	2.833	3.286	**79.024**
	Unknown	5.500	1.857	19.321

Table 1 AVEDac Classifier confusion matrix for Rathbunaster and Parastichopus classes using grayscale features

		True Class		
				
		Rathbunaster	Parastichopus leukothele	Other
Estimated Class	Rathbunaster	**100.00**	0.000	0.286
	Parastichopus leukothele	0.000	**94.857**	0.988
	Other	0.000	3.286	**96.940**
	Unknown	0.000	1.857	1.786

Table 2 AVEDac Classifier confusion matrix for Rathbunaster and Parastichopus classes using color features

VI. SUMMARY AND FUTURE WORK

A system for detecting and classifying visual events in 1-hour time-lapse still images from the Station M site using the AVEDac software is in development and planned for completion by the end of 2009. This system executes on a Condor managed pool of AVEDac enabled computers and results are edited using a graphical user interface for misclassifications or exported for further analysis in conjunction with ancillary data from the Station M site.

Work has been done to modify the AVEDac computer classification program [8] through the addition of color features that resulted in better classification rates for a few benthic classes. Future work includes testing as these changes improve classification of the Station M Images.

While the primary goal of this system is to assist in the analysis of Station M images, future work to package the software system and make it available to the broader community upon completion is planned.

ACKNOWLEDGMENT

We thank the David and Lucille Packard Foundation for their continued generous support. We thank Marco Aurélio Guimarães Moreira for working to add the color features to the AVEDac classifier during the MBARI 2009 intern program. This project originated at the 2002 Workshop for Neuromorphic Engineering in Telluride, Colorado, USA in collaboration with Dirk Walther, California Institute of Technology, Pasadena, California, USA.

REFERENCES

[1] Smith, K. L. Jr., Druffel, E. R. M., Long time-series monitoring of an abyssal site in the NE Pacific: an introduction. *Deep-Sea Research* II 45, 573-586. 1998.

[2] M.F. Vardaro, D. Parmley and K.L. Smith, Jr. A study of possible "reef effects" caused by a long-term time-lapse camera in the deep North Pacific. *Deep-Sea Research*. I. 54, 1231-1240, 2007.

[3] Smith, K.L.Jr., Kaufmann, R. S. and Wakefield, W. W. Mobile megafaunal activity monitored with a time-lapse camera in the abyssal North Pacific. *Deep-Sea Research*. 40, 2307, 1993.

[4] Cline, D.E., Edgington, D.R., Mariette, J., An Automated Visual Event Detection System for Cabled Observatory Video. *MTS/IEEE Oceans 2007 Conference Proceedings*, Vancouver, Canada September 2007. IEEE Press.

[5] Condor High Throughput Computing, The University of Wisconsin, Madison,, viewed 10 August, 2007, <http://www.cs.wisc.edu/condor/ >.

[6] D. Walther, D. R. Edgington, C. Koch, "Detection and Tracking of Objects in Underwater Video", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR),* demonstration, Washington, D.C., 2004.

[7] Edgington, D.R., Cline, D.E., Davis, D., Kerkez, I., and Mariette, J., Detecting, Tracking and Classifying Animals in Underwater Video (060331-207). *MTS/IEEE Oceans 2006 Conference Proceedings*, Boston, MA September 2006. IEEE Press.

[8] Edgington, D.R., Cline, D.E., Davis, D., Kerkez, I., and Mariette, J., Detecting, Tracking and Classifying Animals in Underwater Video (060331-207). *MTS/IEEE Oceans 2006 Conference Proceedings*, Boston, MA September 2006. IEEE Press.