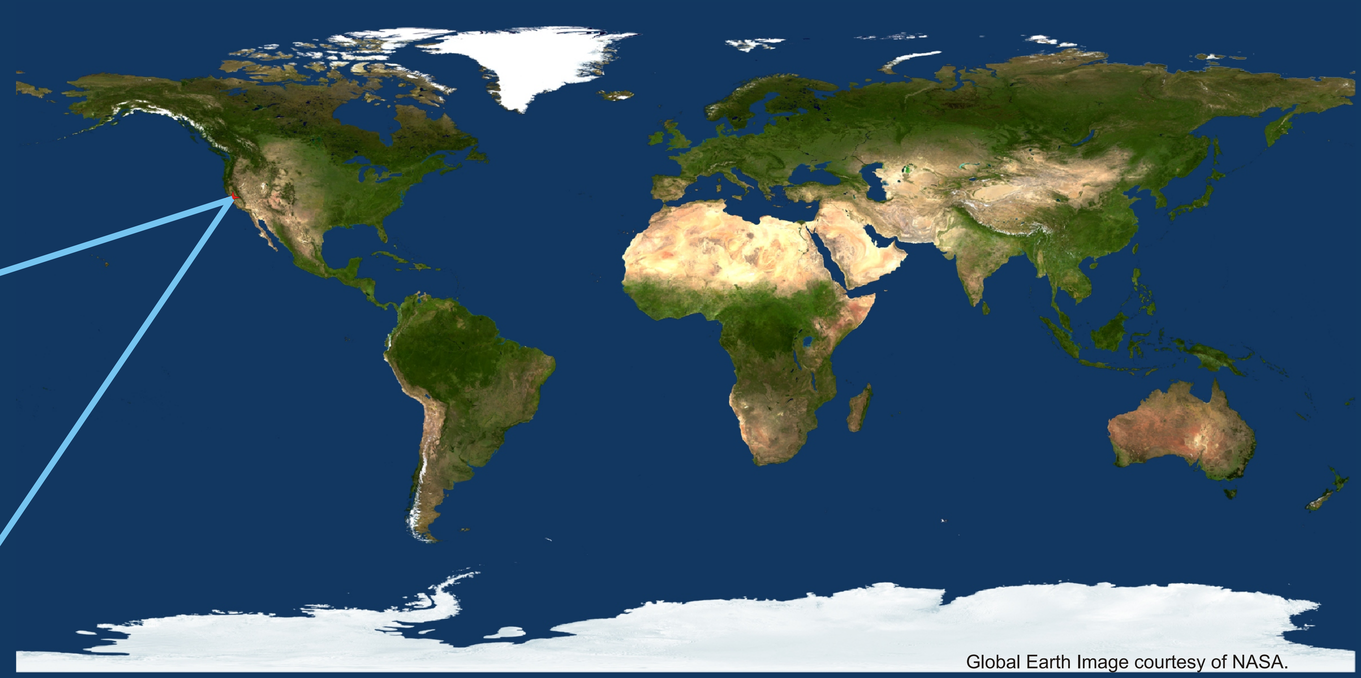
An Automated Visual Event Detection System for Cabled Observatory Video

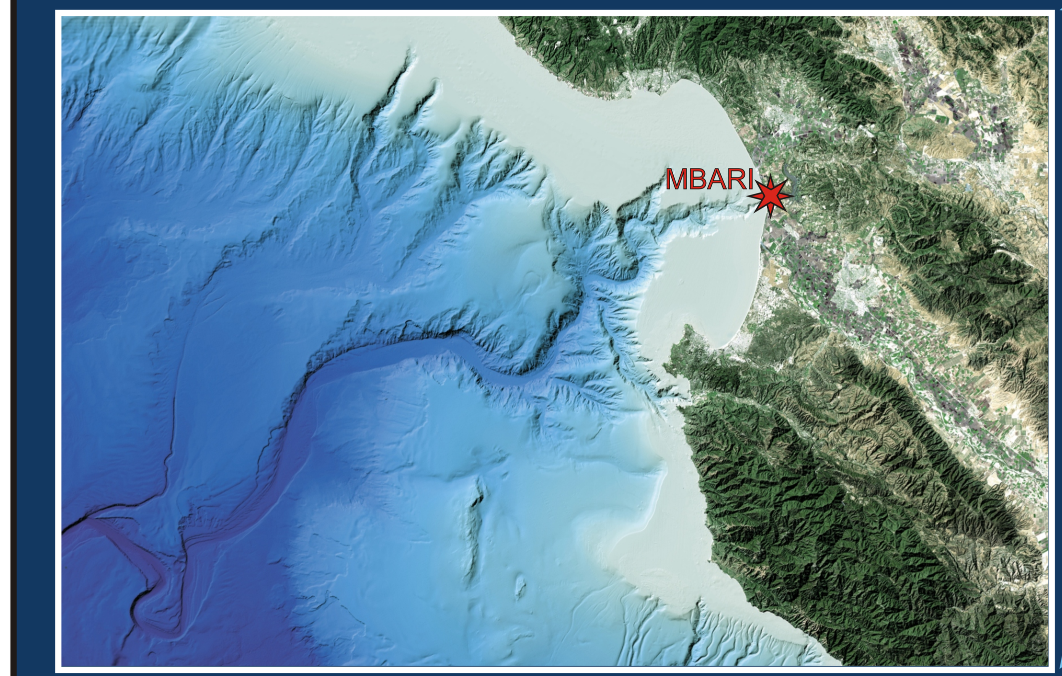
Danelle E. Cline, Duane R. Edgington, and Jerome Mariette *(dcline@mbari.org, duane@mbari.org, jerome.mariette@gmail.com)*

Monterey Bay Aquarium Research Institute, 7700 Sandholdt Rd., Moss Landing, CA 95039

INTRODUCTION

This poster presents an overview of a system for processing video streams from underwater cabled observatory systems called the Automated Visual Event Detection System (AVED) by the Monterey Bay Aquarium Research Institute (MBARI). This system identifies potentially interesting visual events within underwater video captured using a neuromorphic vision algorithm and tracks events frame-by-frame. The events can later be previewed or edited in a graphical user interface for false detections, and subsequently imported into a database, or used in an object classification system.





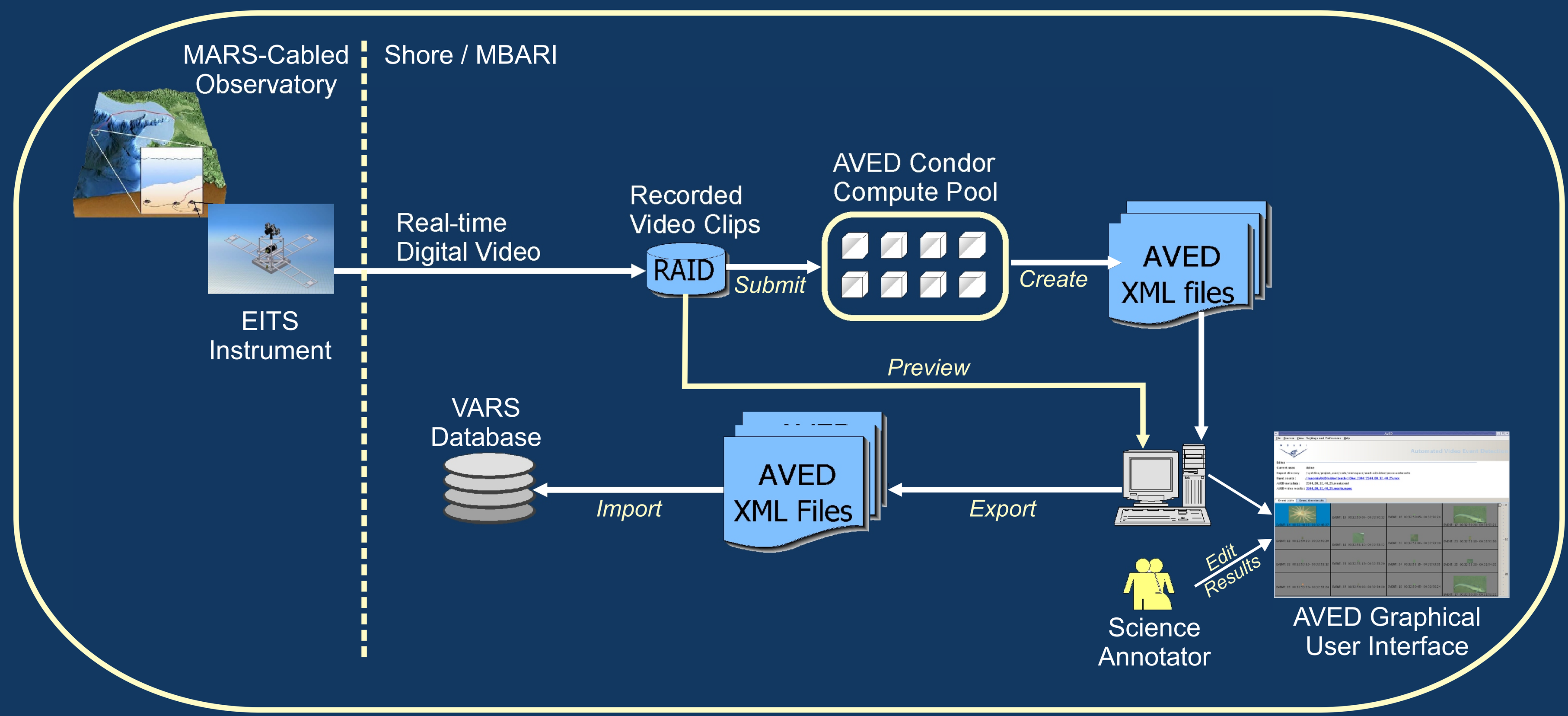
ABOUT MBARI

The Monterey Bay Aquarium Research Institute (MBARI) is a private, non-profit research institute funded by the David and Lucile Packard Foundation located along the Monterey Bay in Central California. We are comprised of approximately 220 scientist, engineers, and operations and administrative staff. Our mission is to advance and maintain a position as a world center for advanced research and education in ocean science and technology, and to do so through the development of better instruments, systems, and methods for scientific research in the deep waters of the ocean.

GOALS OF THE PROJECT

Ocean observatories have the potential to unlock important discoveries with new and existing camera systems, however, these observing systems can have long periods of sparse activity, which can be difficult to visually annotate. The burden of video analysis and management often requires reducing the amount of video recorded and later analyzed which reduces the scientific value of video monitoring. Therefore, the overall goal of automation is to:

- Provide computer-aided assistance in detecting interesting events. This detection is particularly useful during long periods of boring video sequences where human attention strain can result in missed events.
- Design a system to manage the compute resources needed for automatically processing large data sets. Observing system can generate many TeraBytes of video data than need to be processed with many computers which demands some kind of management.
- Provide the tools for science users to quickly preview and remove false detections.

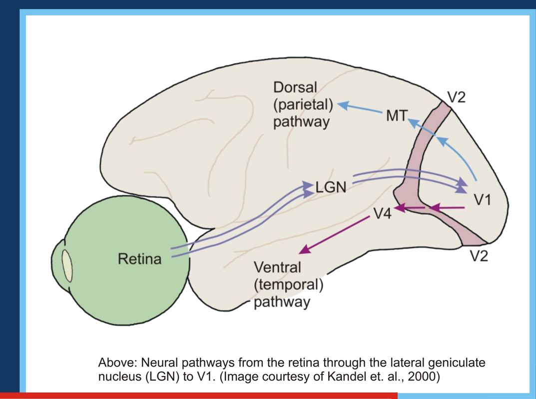


SYSTEM OVERVIEW

The diagram above shows the end-to-end data flow for the EITS system on MARS. The MARS high bandwidth network enables digital video to be transmitted to shore. This digital video stream is then captured on shore into individual clips. To execute and manage this workflow, we use Condor, a specialized workload management system for compute and data intensive jobs developed by the University of Wisconsin Madison. Condor provides scheduling queuing and resource management. Video clips are then submitted for processing in a pool of Condor-enabled compute resources, including an 8-node, 16 CPU Beowulf cluster. The AVED software finds interesting events, saves these events to a metadata XML file. A science annotator then edits events in the AVED user interface for false detections or other non-interesting events. The edited XML metadata are then imported into a database for use with the Video Annotation and Reference System (VARS) that forms a catalogue of the clips as well as the annotations of interesting events by AVED.

AVED SOFTWARE OVERVIEW

The AVED software started in collaboration with the California Institute of Technology and has continued development over the past few years. The AVED software is a collection of software written in C++ and Java designed to work on Linux enabled computers. AVED has shown promising results when applied to video from video surveys conducted with video cameras on Remotely Operated Vehicles (Walther, 2003; 2004).

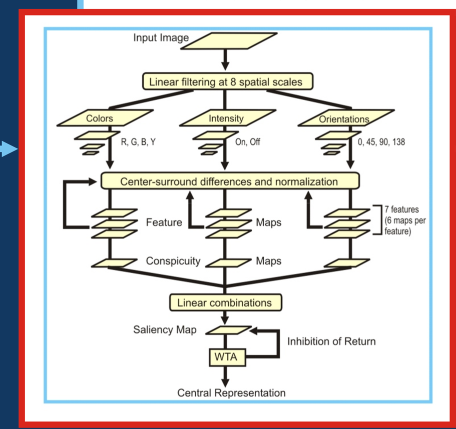


Above: Neural pathways from the retina through the lateral geniculate nucleus (LGN) to V1. (Image courtesy of Kandel et. al., 2000)

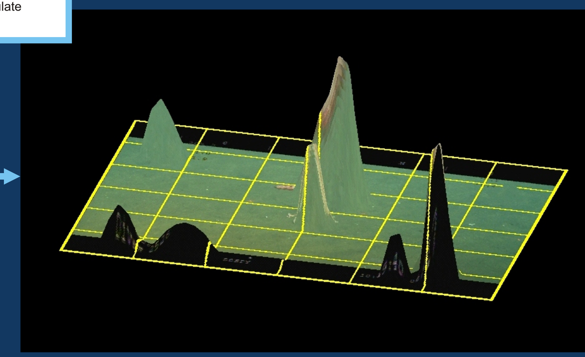
BIOLOGICALLY INSPIRED EVENT DETECTION

Central to the AVED software design is the detection step, where candidate events are identified using a neuromorphic vision algorithm developed by the Itti and Koch (Itti, 1998). This model was developed based on the biology of human perception and vision system. The figure of the brain (shown left) illustrates the pathway of visual processing including the pathway of bottom-up attention modeled in our approach (red arrows). In the saliency model, each input video frame is decomposed into seven channels (intensity, contrast, red/green and blue/yellow double color opponencies, and four canonical, spatial orientations at six spatial scales) yielding 42 feature maps. After iterative spatial competition for saliency within each map, maps are then combined to form a unique saliency map. This saliency map is then scanned for the topmost salient locations using a winner-take-all neural network.

Input Image



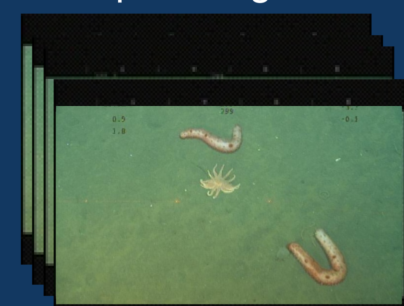
Saliency Attention Model



3D Map of Saliency Output
(Peaks indicate points of high saliency)

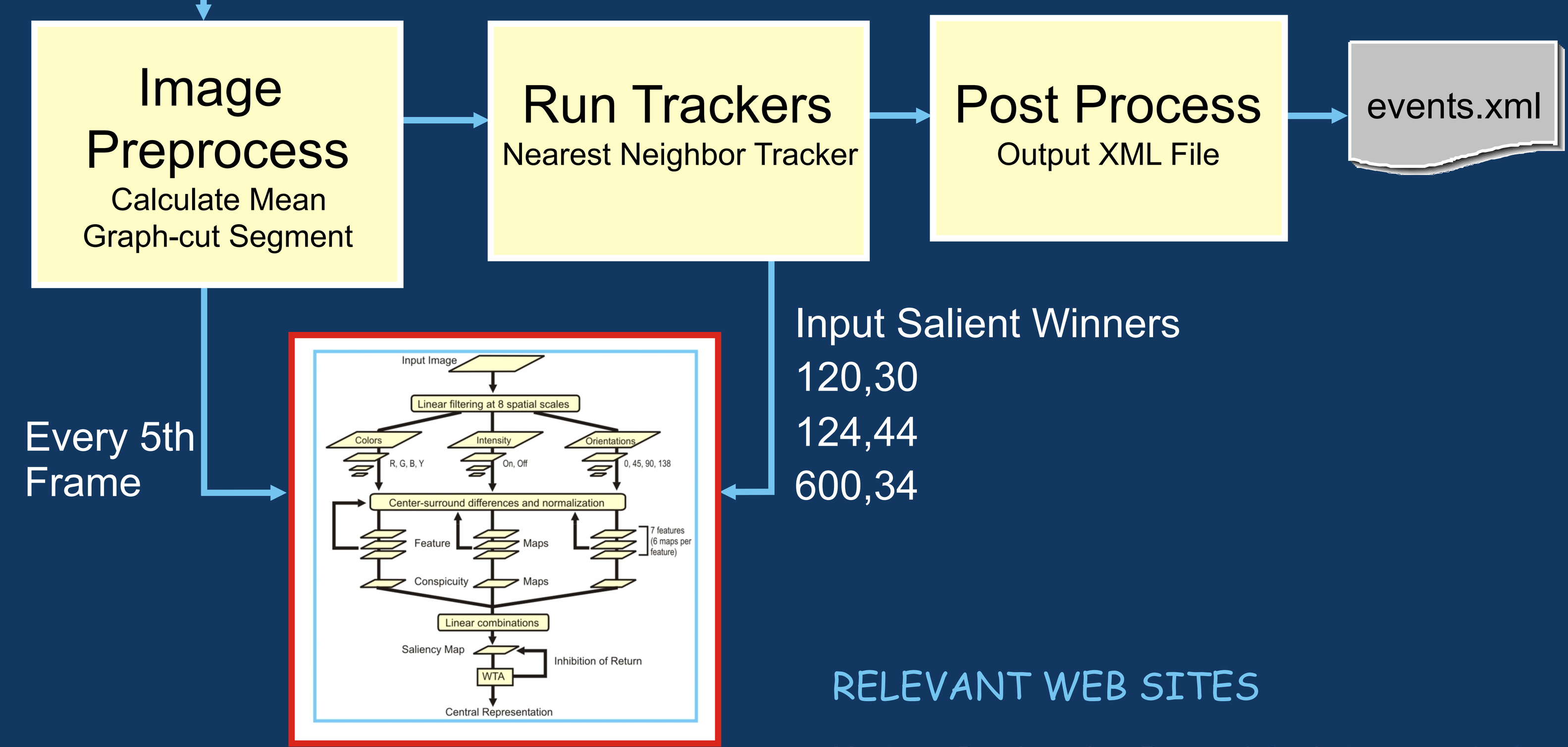
Figure left: Saliency map from the iLab toolkit warped onto a 3-D map. Peaks in the map show points of high visual attention in the center of the image where the Rathbunaster and Leukothele animals are. These points determine where to look for an event and each point is output as a "winner" coordinate, e.g. (120,30), (124,44), (600,34) etc. used for subsequent tracking.

Input Images



AVED ALGORITHM OVERVIEW

Every 5th frame new events are scanned using the detection step. Images are segmented using a Graph-cut algorithm, and a simple Nearest Neighbor tracker tracks the visual events. Results are formatted to XML, which can then be editing in conjunction with the original input video with the AVED user interface.



Saliency Attention Model

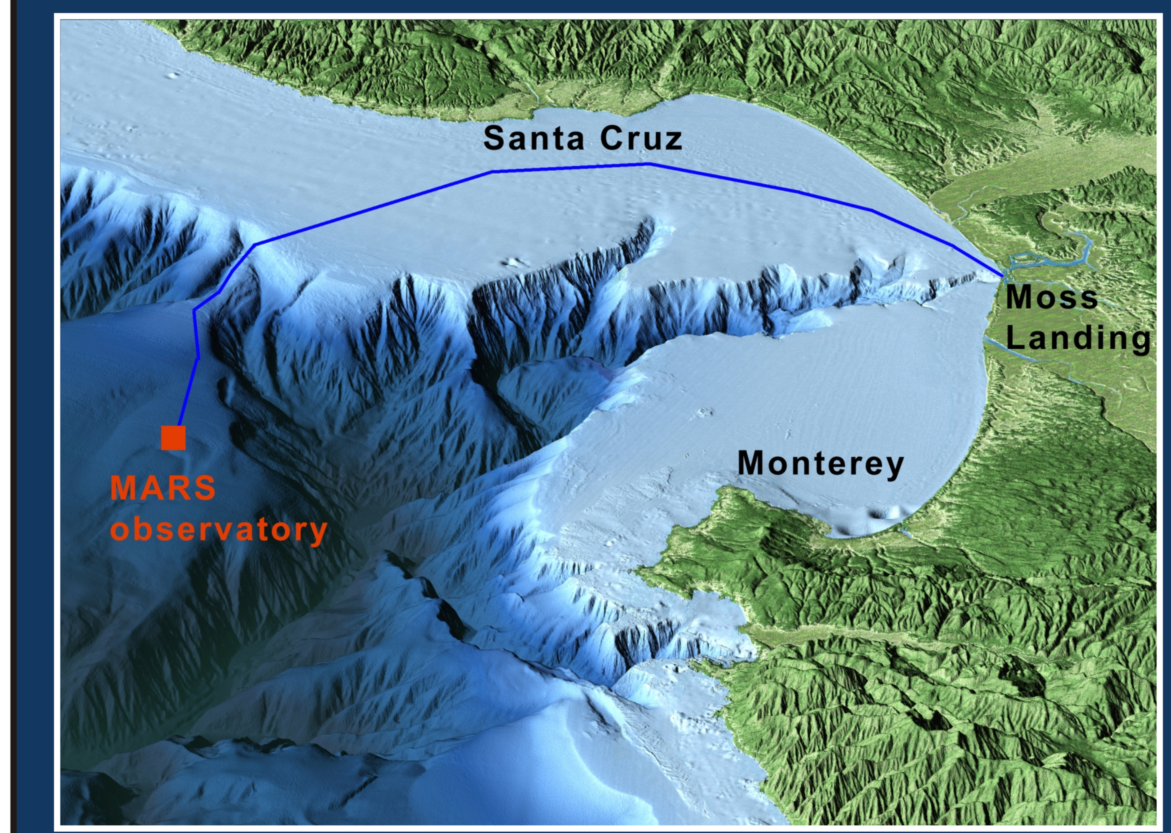


Figure left: The MARS cabled-to-shore-observatory site 891 meters (2,923 feet) below the surface of Monterey Bay. Power and 100 megabit communications is provided to a "science node".

MARS

The Monterey Accelerated Research System (MARS), currently under construction, will provide electrical power and data connections for new research instruments in the deep-sea. Most oceanographic instruments on the seafloor have no connections with the surface, so they have to run on batteries and store their own data. A cabled observatory like MARS removes those restrictions, allowing scientists to design new types of oceanographic equipment and study the ocean in new ways.

Eye-in-the-Sea (EITS)

The Eye-in-the-Sea, developed by the Ocean Research & Conservation Center Association (ORCA), will connect to the MARS "science node" for three to six months at a time, with its camera running 24 hours a day. It can be thought of as a deep-water surveillance system with its low-light camera and far-red illumination. The far-red LED light source illuminates the scene in a color most animals can't see. This allows the scientist to observe animals without blinding them with bright light. The MARS data link allows scientists to watch, trigger experiments with the electronic jellyfish, grab video, or still images, as soon as they see something extraordinary.

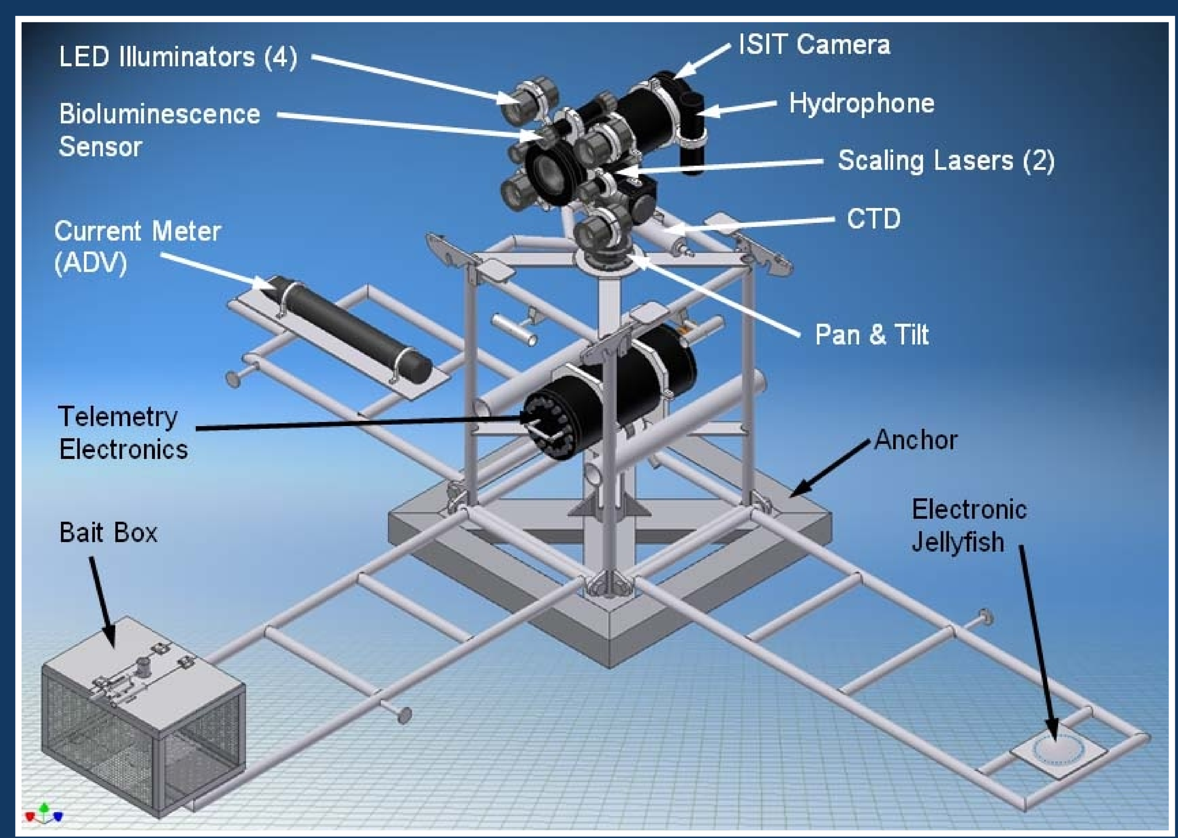


Figure left: The Eye-In-The-Sea Instrument opened on the sea floor in deployed configuration.

Figure above: The EITS instrument open in test tank area at MBARI.

RELEVANT WEB SITES

Monterey Bay Aquarium Research Institute: www.mbari.org
Automated Visual Event Detection (AVED) Project: www.mbari.org/aved
Video Annotation and Reference System (VARS) Project: www.mbari.org/vars
iLab at University of Southern California: ilab.usc.edu
Condor Project: http://www.cs.wisc.edu/condor/
Ocean Research & Conservation Association (ORCA):
http://www.oceanrecon.org/research.htm

ACKNOWLEDGEMENTS

We thank the David and Lucile Packard Foundation for their continued generous support. 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. We thank Edith Widder, Erika Raymond, and Lee Frey for their support and interest in using AVED to help them analyze the EITS video.

LITERATURE CITED

Itti, L., C. Koch, and E. Niebur, 1998. 'A model of saliency-based event visual attention for rapid scene analyses. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(22)*: p 1254-1259.

Walther, D., D.R. Edgington, K.A. Salamy, M. Risi, R.E. Sherlock, and Christof Koch, 2003, 'Automated Video Analysis for Oceanographic Research', *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, demonstration, Madison, WI.

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

Kandel, E. R., J.H. Schwartz, and T.M. Jessell, 2000. Principles of Neural Science, 4th Edition, McGraw-Hill Companies, 1414 pages.

Compiled by Karen A. Salamy for MBARI, 2008