

Edgington, D. R., D. Walther, D. E. Cline, R. Sherlock, K. A. Salamy, A. Wilson, and C. Koch, Detection of Visual Events in Underwater Video Using a Neuromorphic Saliency-based Attention System, *Eos Trans. AGU*, 84(46), Fall Meet. Suppl., Abstract H11F-0912, 2003

AVED AGU Abstract:

**Detection of Visual Events in Underwater Video Using a Neuromorphic Saliency-based Attention System**

Duane R. Edgington (duane@mbari.org), Danelle E. Cline, R. Sherlock, Karen A.

Salamy, Alexis Wilson

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

Dirk Walther, Christof Koch

California Institute of Technology, Computation and Neural Systems Program, Pasadena, CA 91125, United States

The Monterey Bay Aquarium Research Institute (MBARI) uses high-resolution video equipment on remotely operated vehicles (ROV) to obtain quantitative data on the distribution and abundance of oceanic animals. High-quality video data supplants the traditional approach of assessing the kinds and numbers of animals in the oceanic water column through towing collection nets behind ships. Tow nets are limited in spatial resolution, and often destroy abundant gelatinous animals resulting in species undersampling. Video camera-based quantitative video transects (QVT) are taken through the ocean midwater, from 50m to 4000m, and provide high-resolution data at the scale of the individual animals and their natural aggregation patterns. However, the current manual method of analyzing QVT video by trained scientists is labor intensive and poses a serious limitation to the amount of information that can be analyzed from ROV dives.

Presented here is an automated system for detecting marine animals (events) visible in the videos. Automated detection is difficult due to the low contrast of many translucent animals and due to debris (“marine snow”) cluttering the scene. Video frames are processed with an artificial intelligence attention selection algorithm that has proven a robust means of target detection in a variety of natural terrestrial scenes. The candidate locations identified by the attention selection module are tracked across video frames using linear Kalman filters. Typically, the occurrence of visible animals in the video footage is sparse in space and time. A notion of “boring” video frames is developed by detecting whether or not there is an interesting candidate object for an animal present in a particular sequence of underwater video – video frames that do not contain any “interesting” events. If objects can be tracked successfully over several frames, they are stored as potentially “interesting” events. Based on low-level properties, interesting events are identified and marked in the video frames. Presented here is performance data that compare the automated detection method with that of human annotators. The system enhances the productivity of human video annotators and/or cues a subsequent object classification module by omitting “boring” frames and marking candidate objects.

URL: <http://www.mbari.org/AVED>