An Automated Visual Event Detection System for Cabled Observatory Video

Cline Danelle, Edgington R. Duane, Mariette Jérôme (dcline, duane, jerome)@mbari.org

The permanent presence of underwater cameras on cabled observatories, such as the Victoria Experimental Network Under the Sea (VENUS) and Eye-In-The-Sea (EITS) on Monterey Accelerated Research System (MARS), will generate precious data that can move forward the boundaries of understanding the underwater world. However, sightings of underwater animal activities are rare, resulting in the recording of many hours of video with relatively few events of interest. Automatic detection of events of interest could significantly improve productivity and efficiency of professional annotators by reducing the time required to analyze videos, and would enable utilization of multiple underwater survey cameras without excessive operator oversight. Towards this end, an Automated Visual Event Detection (AVED) software is in development at the Monterey Bay Aquarium Research Institute (MBARI). The software is based on a neuromorphic-selective attention algorithm, modeled on the human vision system. Frames are decomposed into specific feature maps that are combined into a unique saliency map. This saliency map is then scanned to determine the most salient locations. The candidate salient locations are then segmented from the scene and tracked over multiple frames, resulting in the detection of visual events.

To address the image processing challenges in underwater video like low non-uniform light, marine snow, etc. AVED contains a collection of different segmentation and tracking algorithms to use. Through experiment, algorithms are chosen *a priori* to use in the observatory workflow that improve the number of events detected. In the case of a fixed observatory camera with minimum pan and tilt or zoom movement, a segmentation algorithm that removes the weight of the detected events from the background calculation is shown to improve the segmentation of faint objects. In this segmentation algorithm, an image average from a running image cache is used with a graphcut-based algorithm to extract foreground objects from the video. Only pixels detected as background are included in this image cache, excluding pixels detected as visual events to effectively remove their weight on the background computation. In the case of a moving camera or a fixed camera with pan and tilt and zoom movement, a segmentation algorithm based on an adaptive threshold and Otsu's method is used. This method begins by building a histogram based upon the image, and then the threshold is determined by the value that maximizes the between-class variance of the gray level histogram. To track visual events, a nearest neighbor tracking algorithm is used for still or slow moving cameras or a Kalman filter is used for moving cameras.

The overall workflow for processing observatory video with AVED is computationally and data intensive. In the EITS experiment video is collected and processed 24 hours a day which requires a distributed network of computers to manage data storage, video transcoding, and AVED processing, including an 8-node Beowulf cluster for AVED processing. The software that manages the workflow includes the workflow management and batch job management tools Pegasus and Condor DAGMan respectively. These are used to manage mapping abstract workflows into concrete

workflows and executing them in a distributed environment. To integrate with the VENUS system, a service-oriented architecture is used to integrate with the existing VENUS Data Management and Archive System (DMAS). using the Simple Object Access Protocol (SOAP). When new data coming from the observatory are recorded, DMAS sends a request to start an AVED process. The Condor batch job management tool is then used to queue the request and execute AVED.

A user-friendly graphical user interface is provided to edit the AVED results that are written to an Extensible Markup Language (XML) file. This same graphical interface is used to setup AVED processing options like the segmentation and tracking options. After editing AVED results for false detections or other non-interesting events, results can then stored in a custom database, or in the Video Annotation and Reference System (VARS) that was developed at MBARI and is currently in use at MBARI.

In this paper we describe in more detail the design challenges, algorithms and workflow design to achieve an end-to-end automated visual event detection system for use in the VENUS and MARS observatories. We also aim to describe how the science user uses this system to access the valuable video data.