Project Number: 900260

Project Start: Jan 1, 2003 **Project Duration (in years):** 3

Please indicate what type of project is being proposed (mark all that apply).

x	x			
Feasibility/Scoping	**Development**	**Research**	**Infrastructure**	**Education**

Project Manager/PI: Duane Edgington

Project Title: Automated Visual Event Detection (AVED)

Lead Scientist: Jim Barry, Steve Haddock, Bruce Robison, Edie Widder
Lead Engineer: Duane Edgington
Team: Danelle Cline, Dan Davis, Christy Herren, Bill Kirkwood, Tim Meese, Steve Rock, Rob Sherlock, and Nancy Jacobsen Stout

1) ABSTRACT (new and continuing projects)

In this project we develop and apply technology to automatically process images for event detection and for recognition of target biological species. The applications for these technologies include 1) lab analysis of underwater video, 2) real time analysis of video from cabled observatory cameras, 3) analysis of video gathered by ROVs in real time, and 4) systems to be deployed on autonomous underwater robotic vehicles (AUVs). This is a follow-on project to the 2002 Visual Image Event Detection Technology Exploration project.

In this continuing proposal we apply selected technologies to solve real scientific and operational problems. Our initial driving vision was to develop imaging systems to add to AUVs to detect when they pass an organism, and to take an image of the organism. Our dream is to be able to identify organisms autonomously in real time. We have expanded the vision to related problems that have fewer technical constraints than AUVs while still having large impact and importance to MBARI programs, science and discovery.

We apply the powerful biomimic (neuromorphic) saliency model of visual attention that has been implemented in software by Itti et al. as a processing approach inspired by biological vision systems to detect visual events. We augment this model with selected best approaches from machine vision work on object extraction, size estimation from single camera imaging, and efficient implementation in software, including deployment on compute clusters. We couple this system to state of the art object recognition methods to create a complete end-to-end automated image processing system. We extend the system to analyze and classify bioluminescent zooplankton. We explore extension of our system to event response cameras. Our science application focus includes mid-water and benthic animals, and animals that use bioluminescence in the deep sea.

2) PROPOSAL

a) The problem addressed (new projects only)

Statement of the problem

One of several areas where unmanned remotely operated underwater vehicles (ROVs) have proven to be extremely useful tools in oceanography is the capture of video and visual images of biological creatures. The great success of ROVs in collecting high-resolution video during dives encourages us to look for ways to increase the amount of information and knowledge we gather from underwater cameras. We see tremendous potential benefit in automating portions of the analysis of video, and to extend image collection to unmanned autonomous underwater vehicles (AUVs), to other underwater instruments, and to cabled video cameras deployed as part of ocean observatories.

ROVs are natural platforms for collecting broadcast quality high-resolution video, since high-speed telemetry from vehicle to surface and electrical power for camera and lights are available via a cable tether. Some analysis or adaptive sampling is possible during the dive via surface video monitors and annotation systems, and outline or detailed post-dive analysis is routinely done in the video lab using sophisticated tools such as the MBARI Video Annotation and Reference System (VARS). There are two bottlenecks in collecting and analyzing this data that we hope to address in this project:

e) Video annotation is demanding and tedious work; highly trained scientists can work at most four hours a day annotating video; the capacity of the video lab and research staff to post-process dive footage is a historical bottleneck in reaping the benefit of these expensive dives. One of our goals is to develop systems and method to augment the productivity of video annotation analysts who are processing video from ROV dives.

2. The ROVs are expensive, high demand resources. The science staff would like to double or more the length of dives that are taking quantitative species and abundance samples (transects), as well as increase the frequency and sites of such transects, which is impracticable with the ROV resources and budget that MBARI has. It is very attractive to augment ROV dives with AUVs, which promise the capacity to go longer, more frequently, at lower cost than ROVs. However, AUVs are very power limited, since they are not tethered, and must carry their power sources and data storage with them during missions. One of our major long-term goals is to implement technology for enabling AUVs to fulfill the same kind of science missions that ROVs currently provide for species and abundance sampling.

The goal of this project is to implement high-payoff technologies that will enable detection and classification (recognition) of a visual event in the presence of "noise" in the visual field. We have investigated both research work and commercial systems that do visual image processing and image capture. We will apply and implement selected best technologies for our specific problem areas.

This development project will be done with participation and input from

e) Science users (Robison group, Barry group, Haddock group and others),
e) Video lab personnel and science researchers who analyze video (expertise on features of interest, "noise" in the visual field, and illumination requirements),
e) 600034 Steve Rock's group (expertise on automated video processing and controls), outside organizations deploying cameras on AUVs, and the 600034 project team,
d) Research organizations developing neuromorphic vision systems and chips (Kwabena Boahen, [U. Penn], Tobi Delbruck, [ETH Zurich]),
e) External collaborators, with their own funding (Dirk Walther, [Caltech], Christof Koch, [Caltech], Laurent Itti, [USC], Pietro Perona [Caltech], Marc'Aurelio Ranzato [NYU], Andres Castano [NASA JPL]).

It is our distinct intent not to do machine vision research as part of this project. Rather, we focus on applying best technologies and proven robust systems that are being researched and implemented in the top machine vision research labs. We partner with those labs to offer our compelling application and video database as test bed for their implementations. It is our intent to develop a unique compelling system with wide applicability to marine studies that is well engineered and crisply executed.

Why it is important

MBARI has significant investment in ROV video image capture technology and scientific investigation tied to video analysis. This project implements technology that would augment the effectiveness of the staff and equipment to solve scientific questions. In addition, technologies to add autonomous vehicles and autonomous cameras to the scientific arsenal would prove extremely valuable in increasing our ability to sample more thoroughly, more often and in more places economically.

The ability to autonomously detect and identify a creature could open up new lines of scientific enquiry. Analyzing archived video for a key group or species one at a time has emerged as a disruptive technology that could enable many folds amplification of the scientific value and use of this expensive asset.

Technology to aid analysis of video from cabled observatory cameras will be critical for those cameras to be cost effective instruments.

Finally, implementation and ongoing investigation of alternative image processing and detection techniques could suggest follow-on innovative work, such as optical-flow based navigation, feature following, extension to acoustic imaging and analysis and so on.

Present state of art/understanding

Visual image processing is a huge field with a rich history of investigation. Automatic visual feature detection has been a research and engineering topic in a number of disciplines, including medical image analysis, surveillance, robotics (for example, place recognition) and industrial control. There has been major advances, particularly in the last 10 years, that move machine vision from being an esoteric research area to being a field that is delivering robust, mathematically based engineering solutions. Here we present highlights of appropriate technology to motivate the proposed implementation tasks.

There are two classes of approaches to visual feature detection. The first is a parametric approach, where a captured image is processed to detect edges, lines, corners, junctions and so on, and then these parameters are used to match some model of the feature of interest. The images and parameters can be 2D, or 3D images can be constructed and analyzed. Video image processing can add multi-frame images to the analysis, comprehending blurring, rotation, noise, and shifts between frames due to motion. This enormous field of research and engineering has yielded surprisingly few robust general systems, but there are systems of focused application that are successful in feature detection and visual event detection. An example of such as system is the rock detection and recognition system deployed on MARS Exploration Rovers Opportunity and Spirit, developed at the Jet Propulsion Laboratory by Andres Castano.

One active area of research, which has yielded some successful deployed systems, is in the area of surveillance of natural scenes (i.e. outdoors under natural lighting). In modern successful systems Gabor filtering is often used to do noise filtering and feature detection; Gabor filters are a mathematical center excitation-surround inhibition (center-surround) operation that has direct parallels to information processing in mammalian retina and early vision processing in the brain. These are a best practice method we propose to continue implementing and exploring. This approach is a cornerstone of the AVED system implemented to date.

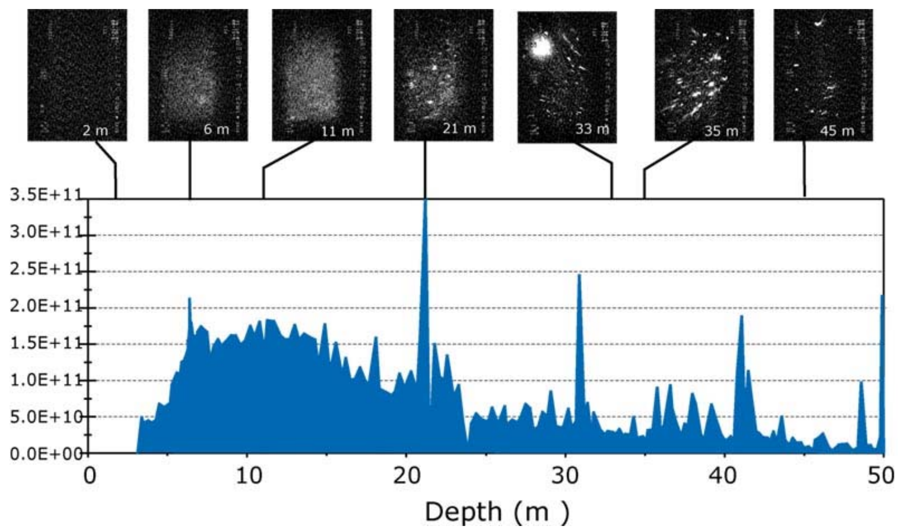
Many vision systems include learning algorithms, to tune the systems to the particulars of the environment. Modern systems use Bayesian learning systems, extensions of the classical artificial neuronal networks (ANNs) that have had extensive research and practical application. Again, comparison of Bayesian learning with neuromorphic models of biological systems shows direct parallels. MBARI is in an unusually favorable position to apply learning algorithms to video image analysis, given our wealth of fully annotated video (some 11,000 hours worth) and some 950,000 annotated events available as a teaching set.

Perona and colleagues at Caltech have developed the state of the art classification system to recognize objects in natural scenes. This system, in its most advanced form recently published at the July 2004 CVPR conference, has been demonstrated to recognize 101 classes of objects, being trained with only 6-8 examples of each object. The performance of the system exceeds 90% correct classification. The input to the system, implemented in Matlab, is objects segmented from the scene. Segmentation using the AVED system has been successfully demonstrated, and integration of segmentation from a system similar to ours with the Perona "Constellation" system was also presented at CVPR 2004 (Rutishauser, Walther, Koch, Perona)

Thus, state of the art understanding in machine vision, along with robust programs implementing these algorithms, presents for the first time the opportunity to integrate robust systems to detect, classify and recognize marine organisms from video taken in their natural environment. The potential impact of this integration on ecological work at MBARI, and ecological work in oceanography and fisheries research, could be enormous.

<u>Automated Bioluminescence Video Detection (ABVD)</u>
Although oceanographers routinely use fluorometers to estimate the abundance of phytoplankton, we currently have few reliable ways to autonomously sample zooplankton communities, other than time-consuming microscopic examination of discrete samples. Existing optical and acoustic methods for assaying zooplankton biomass each have their strengths and weaknesses – the main weakness being a limited ability to resolve species composition or even broad taxonomic groups. Thus, there is no standard for measuring the growth kinetics and distributions of zooplankton; there is essentially a 'missing-link' among oceanographic sensors.

The solution to one part of this dilemma is the use of stimulated bioluminescence. Bioluminescence is purported to be an optical proxy for zooplankton biomass (Piontkovski et al. 1997; Lapota 1998). Indeed, small portable bathyphotometers (MBBPs), such as those used on the MBARI AUVs and other platforms, are now being integrated into standard oceanographic packages. However, datastreams from MBBPs, which provide 'bulk water' light levels – the intensity of stimulated light per depth per volume of water –provide indications, but not conclusive information on plankton identities. We propose that a combination of MBBPs used in conjunction with low-light cameras that record stimulated bioluminescence on a mesh screen (SPLAT cameras, *sensu* E. Widder) can help provide the missing link in automated plankton analysis.



For example, the figure above shows BL intensity as a function of depth in one location in Monterey Bay during MUSE 2000. The bulk light signal is shown in conjunction with video data from a SPLAT camera (E. Widder, HBOI), which was mounted adjacent to the MBBP on the ROV Ventana. The video frames are matched to depths at which they were recorded. Without the video data, the shape of the BL profile over depth and its signal characteristics are the only source of information we have on the identity of the BL organisms responsible for the light signals. With the addition of the video frames to the dataset, we know conclusively that uniformly distributed single-celled dinoflagellates were responsible for the majority of the BL at depths less than 11 m, but brighter and more rare copepods and jellyfish became increasingly important at depths greater than 15 m.

While actual plankton samples are still a vital component to our research, BL video data adds a new spatial dimension that samples or MBBP data alone cannot provide. The resulting data combination can reveal species-identifying flash patterns, concentrations of and distances between key bioluminescent species, and spatial heterogeneity of plankton species. The ability to determine the types of organisms that are producing BL at discrete depths and their absolute concentrations enhances current efforts in the Haddock lab to model the distribution of BL in Monterey Bay. In addition, as the price of commercial, low-light, image-intensified cameras continues to decrease, we believe this software can be made available to the oceanographic community as an essential diagnostic for plankton distributions.

How it relates/contributes to MBARI's mission and strategic plan

This proposal attempts to incorporate leading-edge research and proven robust mathematically based engineering approaches to visual image processing, applied to problems in oceanography. Our application of these technologies is to our knowledge novel and new. The project would offer breakthrough approaches to analyzing video from ROVs, AUVs, manned submersibles and towed cameras for species and abundance surveys that would have wide applicability and interest to oceanography. The project would yield critical technology to make use of real time video from cabled observatory cameras on MARS, ORION and other observatory efforts. This high-reward project would unlikely be undertaken at other oceanographic institutions, but is yielding approaches and products of great interest to those researchers and engineers.

b) Approach to the problem (new and continuing projects)

If the proposed project includes multiple individual projects, please list each of these projects as a subheading in the proposal. (For example, Benthic Biology and Ecology subheadings include Biology of deep-sea CO_2 disposal and Cold seep biology – see example on PCO web.) For each subheading, please address the following:

Status of project

We have completed the tasks attempted in 2003, and are on track to complete the tasks planned for 2004. These tasks include:

- Extending the "vision" program from USC to comprehend the often-sparse data of our underwater video, where many individual frames are "boring" with no visual events to be detected.
- Extending the program to comprehend motion as equivalent to space in video collected from a moving camera, identifying related events in each frame to the object responsible for each event.
- Migrating from proof-of-concept NTSC video copies and frame grab hardware to full resolution frame grab hardware fully integrated with the digital time code-based annotation database.
- Testing the extended program on multi-CPU clusters for real-time video analysis.
- Comparing the performance of the program with expert human annotations (detection of events worthy of annotation) for both midwater and benthic quantitative transects.
- Completion of specific 'annotation assistant' video lab aids

Results to date

Progress to date:

1. We have developed a process flow to enable automated comparison of automated event detection with human annotation. We have demonstrated algorithms to separate "interesting" events from "boring" events detected by the saliency algorithm and compared results with human annotation. We have enhanced the core saliency algorithms to improve detection of long string-like animals such as siphonophores, and have implemented tracking of detected events.
2. We have developed an approach to stream video to our system and capture the results continuously in near real-time. We have demonstrated the software system running on our compute cluster, with the goal of achieving near real time performance with a parallel computer implementation of the software system by the end of 2004.
3. We have developed detailed requirements and use cases from the MBARI video lab to drive development of tools for video annotators to use.

4. We have developed detailed requirements and use cases from MBARI scientists for specific missions that drive publication quality research.
5. We have published, presented and demonstrated our work at the IEEE Computer Vision and Pattern Recognition Conference (CVPR), presented to the MBARI board of directors at MBARI Day of Science and Technology, presented a poster at the MTS/IEEE Oceans Conference, presented a poster at ASLO/TOS and ASLO summer meetings to the marine ecology community, and presented a poster at the Fall AGU meeting session on artificial intelligence applications in the Geosciences. Highpoints of our presentations and publications to date include 1) our 2004 CVPR paper, which was double blind peer reviewed by machine vision experts and selected as one of only 54 out of over 1000 submissions (5% acceptance rate) for oral presentation at CVPR in plenary presentation to an audience of over 500 machine vision and pattern recognition specialists, 2) our concurrent day-long demonstration of the system at CVPR which was seen by over 100 individuals, and 3) Rob Sherlock's poster at ASLO summer meeting which was very well attended and received.
6. While continuing analyzing midwater quantitative video transects, we have begun analyzing benthic transects.
7. We have quantified the performance of our detection system: over 80% detection of annotated events in midwater transects, and over 90% detection in benthic transects at Smooth Ridge.

   The system has proven to be robust and consistent, requiring no tuning to adapt to changes in lighting, camera settings, turbidity, or other natural conditions.

Development goals for 2005:

1) Implement algorithms for "large object" detection and segmentation based on characteristic features.

The current system detects events in the video, and segments the image by flooding events that exceed an intensity threshold. The system displays the events outlined in contrasting color, or surrounded by a bounding box. To drive an automatic classification system, we need to insure that the object is segmented in its entirety, and not clipped by this simple segmentation scheme. Rutishauser, Walther, Koch, and Perona have demonstrated a segmentation method that retains the entire detected object and is suitable to drive classification. This task is to integrate their method into our MBARI system. We will also evaluate the rock identification and image segmentation algorithm implemented by Andes Castano at JPL, based on active contours (see 4 below).

2) Object classification and recognition.

Detecting a visual event, and extracting the object underlying the event, is a powerful and important capability to segment the scene. Currently humans need to look at each detected event, and identify what the object (animal) actually is. This task explores automatic autonomous object recognition, a Holy Grail of artificial intelligence research for many decades. Active research in the Perona laboratory at Caltech, focused on doing object recognition, is widely recognized as state of the art in object recognition and classification in natural scenes. In this task we propose to continue to collaborate with that group as appropriate to apply object recognition approaches to our application domain. MBARI has a tremendous database in its fully annotated library of dive video that makes us a valuable collaborator in object recognition research and development. We hope to exploit that valuable resource through these external collaborations.

Towards this end we have brought to MBARI a recognition system implemented by Perona student Marc'Aurelio Ranzato. This program, implemented in Matlab, recognizes biological "particles" in single images. Like our midwater images, the "particles" to be classified are relatively small (some 50x50 pixels) in a frame with other particles. In this task we will collect a corpus of images of selected organisms, for example, nanomia, radiolarian, and chaetognatha from midwater dives, and train the system to recognize these common species, and to classify them. We have already demonstrated that AVED detects these organisms over 90% of the time when present in midwater transects.

Similarly, we will collect a corpus of images from benthic transects, for example, Rathbunaster californicus, Parastichopus leukothele, and sebastes, and train the system to recognize these common species, and to classify them. Summer Intern Ishbel Kerkez demonstrated that AVED detects these organisms over 90% of the time when present in benthic transects.

3) Motion detection and classification.

Implement extension of *vision* program for motion detection. This task involves testing the implementation of motion detection as feature map that was recently implemented at Caltech, and extending as necessary.

The task includes removing camera motion from the overall motion seen in each scene. We are computing camera motion as part of the Kalman filter based tracking of objects that we have implemented. After removing camera motion, residual motion is biological motion plus currents (assuming no turbulence). We will work to robustly extract the motion due only to biological motion of creatures in the scene.

Motion is a strong signal of saliency in midwater video; details of the motion are cues to classification by human annotators. We will investigate classifying the motions we detect, to use as a feature for classification, based on state of the art egomotion research.

4) Extension to bioluminescent zooplankton (Automated Bioluminescence Video Detection)

Recently, there has been increasing interest in characterizing the populations of bioluminescent organisms at varying depths in the water column. Devices that are sensitive to low light environments have been employed in the past and modern CCD and CMOS-based imaging arrays provide additional means of gathering the characteristic bioluminescent signatures of these organisms.

At shallower depths, bioluminescent organisms such as phytoplankton are spatially nondescript and can be measured quantitatively with flourometric methods targeted at a specific wavelength. At deeper depths, where zooplankton such as dinoflagellates and copepods are encountered, populations are less numerous, but the bioluminescent signature of each organism is somewhat more distinct. This signature can be identified by three characteristics:

- Shape, both before and after mechanical excitation
- Duration ('flash kinetics') after mechanical excitation
- Wavelength of bioluminescence emission

In this task we build on AVED. We apply machine vision methods for segmentation and identification of discrete objects in an image field. Active contours, or “snakes”, are a computer graphics construct where a boundary surrounding arbitrary objects of interest in an image is generated and is made to converge and define the visible boundaries of these objects in images. Active contours are therefore used to quantify objects of interest in images, such as a 'dark' region that might represent a tumor in a medical CT scan. The same method was originally employed utilizing a B-Spline implementation by Widder, et al at HBOI for characterizing the bioluminescent organisms encountered by the 'Splat CAM' device. In this task we would integrate active contours into the AVED system. In addition, new methods outlined below would be applied to the video stream to enhance the potential for characterization of the images.

Implementation Plan
We propose to extend the work of Edie Widder's group:

- Use the Intel OpenCV Computer Vision Library (http://www.intel.com/research/mrl/research/opencv/) to do initial development of a standalone proof of concept application. The OpenCV library is a comprehensive image processing library that can be used to develop applications under Windows or Linux.

- Evaluate enhancements to the Kass algorithm such as Gradient Vector Flow, an optimization that improves the convergence of the active contour to the boundary of the object of interest.

- Perform contrast enhancement on each frame of video and, utilizing a combination of AVED and newly developed code identify, characterize, and track each object which can be perceived to be a bioluminescent organism of interest. Evaluate the effect of newer noise reduction techniques, such as application of Haas Wavelet Transforms, in improving the efficacy of the object detection algorithms.

- Use the object tracking mechanisms in the AVED software (Kalman filter, Optical Flow) to track each identified bioluminescent object in multiple frames. This will provide another possible mechanism of identification as the organism may experience morphological change over the period of time it is actively bioluminescent.

- Develop a Graphical User Interface displaying relevant data to the operator.

5) Event response Close-up Video Camera scoping effort

This task is to explore feasibility of automatically controlling a paired pan/tilt/zoom close-up camera to focus on a potential event of interest during fixed velocity video transects. Midwater and benthic ROV quantitative video transects are taken with the science video camera fixed at a relatively wide field of view (2 meters for midwater transects, 1 meter for benthic transects); the ROV is controlled to transect at fixed velocity. This allows a reasonable number of events and amount of space to be covered during the limited ROV time available. However, this means that the minimum size of an object that can be reliably recognized is restricted by the wide field of view – there just aren’t enough pixels in the image to resolve organisms less than a few cm in total length, and it is often difficult or impossible to positively identify the smaller creatures. At transect speed it is impractical for a human to carefully monitor the video stream and control a second camera to zoom in to small objects to get the close-up image needed for identification. In future applications (such as cabled observatory cameras or AUV transect systems) there may not be a human in the loop at all.

We propose to build on the automated video event detection (AVED) project, to explore controlling a second paired camera to zoom in on selected events in real time. We propose to develop algorithms for camera control in the lab, using a commercial video teleconferencing pan/tilt/zoom camera that is already integrated into the Saliency program on which the AVED system is built. We propose to determine the technical requirements for event response cameras, to evaluate if commercial video cameras are fast enough to respond in the various use cases of ROV video transecting, AUV or towed sled video transecting, and fixed observatory cameras. We will research the state of the art and state of industry in various pan/tilt/zoom technologies, including steered mirror systems. Our products will include a laboratory demonstration system, trade study and a technology whitepaper. This work is primarily a controls exercise, and thus will be done in collaboration with Prof. Steve Rock and any students he designates to work on the problem from Stanford.

6) Refine and iterate proof of concept system to develop system that is useful to the video lab staff for their mission.

One goal of the AVED project is to develop an 'annotation assistant' tool that uses the AVED technology to facilitate the identification of objects in video. The current prototype technology can be used to detect several objects in priority order of interest, as defined by the saliency-based algorithm. Our focus has been on an application that identifies frames, and objects within frames, of interest for subsequent review and annotation by the VARS system. Such an application could also be used to process video recorded from benthic camera deployments, or in background mode process video from cable-based video. A specific task of the requirements analysis has been to determine how this application will interface with the VARS annotation system. The prototype will be able to create a stream of data that can be assimilated by the VARS system to aid post-cruise annotation. A long-term goal will be to integrate the application for real-time use with the VARS system during cruise operations.

The VARS system is in the final process of development and will, shortly, replace the current VICKI annotation system. The current development is constrained by the existing database of annotations and the model on which it is based, although it is likely that this model will be modified to meet new requirements. Since the development of an annotation assistant is necessarily dependent on the model used for 'annotations', its development will need to be coordinated with the VARS development.

The use cases identified in cooperation with the Video Lab staff for an 'annotation assistant' are:

- Identify timecodes of 'interesting' frames in ROV video stream prior to full annotation by annotators. The results should be integrated/editable with VARS. Frame grabs should be taken of the interesting frames. Comments: The AVED system has already been shown to be capable of achieving this. The work that needs to be done is to establish a data format and interface for these results to be used in VARS.

- Indicate the position and extent of an annotated object within a frame. Comments: AVED already has this capability. A data format and interface to VARS will need to be established.

- Detect the beginning and end of a sequence containing an object. Comments: AVED already has this capability. A data format and interface to VARS will need to be established.

- Identify segments of video with technical problems (e.g. timecode jumps, video glitches/drop outs). Comments: The AVED system may be helpful with achieving some of this depending on the problem. What constitutes a 'timecode jump' depends on the timecode setting policy. AVED can detect pixel dropouts in most cases.

- Identify objects in the video based on sets of typical images of objects. Use this on any segment of video to id any and all objects, or use it to look for only certain objects in a video stream. Frame grab should be taken for each. A report of objects that could not be identified should be automated. Comments: This is proposal task (2) above.

It is our intent to deliver these identified 'annotation assistant' capabilities. Through our interactions with potential users we have discovered that AVED technology is disruptive – some potential users see no need for assistance, and have no interest in testing a system that would disrupt their current workflow, even if it might aid their productivity. Others, however, are very interested in assistance, particularly for tasks that are currently tedious, or would be become tedious with increased video data flowing in from new data streams. Thus, the team intends to deploy 'annotation assistant' capability with users who have specific needs that motivate them to go through the learning curve of adopting a new system. Specific science applications for AVED are enumerated below:

Science Applications:

1) A key science application of our system that has emerged is the quantitative analysis of a large set of videos for single species of interest. In each case, being able to fully analyze the large video archive that MBARI has collected for a species or group would lead to multiple publications of impact. For example:

A. Radiolarians have recently emerged as a species of interest to Dr. Haddock, with a hypothesis that they could be major contributors to carbon cycling in the deep-sea. The AVED system with recognition capability for radiolarians would be a way to quantify their impact. The alternative is to re-annotate much of the 11,000 hours of video MBARI has collected, since radiolarians have only recently been routinely annotated, and have not been in the "top 10" list of creatures that the midwater annotators have focused their annotations on. The radiolarians are in the video, but re-annotation, by machine or by humans, will be required to develop the scientific story.

   To quote and paraphrase from Dr. Haddock's 2004 proposal, radiolarians are an example where many potential breakthroughs can come from answering basic questions and from quantifying organism occurrences in midwater video. As with most jellies, these protists can only be collected using specialized sampling methods, so that live radiolarians have received almost no scientific attention. This lack of effort is in stark contrast is their importance in many ecosystems: as was learned during blue-water dives on the Hawaii transit, polycystines can be dominant members of the ecosystem in oligotrophic waters. These numbers, when considered along with their siliceous skeletons, suggest that they may replace diatoms as the major factors in silica cycling outside of coastal upwelling zones.

In the deep sea, Haddock and colleagues have discovered that, contrary to a century of literature, all of the tuscaroid radiolaria examined are colonial. This morphology is key to their method of feeding, allowing them to act as "marine Velcro", accumulating many times their own biomass of detritus from the water column. Because of their ability of accumulate and repackage marine snow, they will play a role in carbon transport in deep sea. Quantitative annotation of their distribution will enable calculation of contribution to carbon cycling in the deep ocean. That contribution could be significant.

B. Poebius has emerged as an important midwater species. In early studies midwater ecologists did not appreciate how widely spread they are in the water column. Poebius is now thought to be important in the ecology of eastern North Pacific. The distribution and the effect of el nino events on their abundance is unknown. A particular question of interest in the Robison group is whether they have a bimodal distribution, or is the distribution uniform. Ironically, they are hard to annotate because they are so abundant. They have not been annotated quantitatively in transect data. The task to study these creatures is to carefully annotate midwater transects for Poebius, and to correlate the annotations with depth and environmental conditions. An automated poebius classifier with predictable performance would directly yield scientific results.

C. Jane Lee, a Ph.D. student in adjunct Bill Hamner's lab, is studying the giant siphonophores praya and Apolemina. She is not using the video database for her work because these animals are not quantitatively annotated. Again, an automated re-annotation of video transects for these large animals would directly enable scientific results.

D. Dr. Kevin Raskoff at CSUMB is working with the MBARI midwater group to study aegina, an extremely abundant midwater jelly. Appreciation of their wide distribution is relatively recent. An automated re-annotation of video transects for these abundant animals would directly enable scientific results.

E. Sergestes is another group that has been inconsistently annotated. Their distinctive color, shape and motion make them another candidate for automated re-annotation of video transects of these abundant animals that would directly enable scientific results.

F. Dr. Barry has told us without equivocation that his scientific investigations would make heavy use of a system that could predictably recognize even a few specific benthic species or groups. An example is Ophiuroidea, a large group of echinoderms that includes the brittle stars (Ophiurida) and basket stars (Euryalida). While we have found that benthic transect video is very amenable to automated event detection and classification, it is nonetheless very tedious and time-consuming for humans to annotated quantitatively. The applications for benthic ecology include habitat identification and classification. Areas for us to explore include classifying other features in the video besides the animals, including the floral and the geological features, to enable habitat ID.

   Real-time classification would be very powerful for benthic transects, to enable zooming in of a slaved camera to enable measurement of size, reproductive state, substructure (e.g. larval forms to be released), and parasites (such as those found on Rathbunaster).

2. Automated Bioluminescence Video Detection

Key science deliverables for this proposal task are:

- Develop and evaluate a method for automated quantifying of zooplankton abundance
- Design and utilization of a BL-detecting AUV with a SPLAT screen for time-series transects in the Monterey Bay.
- Analysis of SPLAT video by:
  - Develop algorithms to identify key BL species of plankton using their distinctive flash patterns over time and space (BL kinetics)
  - Quantify concentrations of key bioluminescent species.
  - Quantify distances between key bioluminescent species.

We have discussed this proposed development with Dr. Widder, and her former postdoc Sonke Johnsen (currently a faculty member at Duke University and collaborator with Steve Haddock) who worked on the current splat screen software program, and have their enthusiastic support and interest in this proposal.

Project team: Duane Edgington, Danelle Cline, Dan Davis, Bill Kirkwood, Tim Meese, Steve Rock, Rob Sherlock, Nancy Jacobsen Stout, Bruce Robison, Jim Barry, Steve Haddock, Christy Herren.

Publications (from the last 5 years)

Walther, Dirk, Duane R. Edgington, Christof Koch, Detection and Tracking of Objects in Underwater Video. *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition (CVPR),* Oral Session 2B, Washington, D.C., 2004. (peer reviewed).

Edgington, Duane R., Dirk Walther, Danelle E. Cline, R. E. Sherlock, and Christof Koch, Detecting and Tracking Animals in Underwater Video. *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, demonstration, Washington, D.C., 2004. (peer reviewed).

Edgington, D. R., D. Walther, D. E. Cline, R. Sherlock, and C. Koch. Detecting and tracking animals in underwater video using a neuromorphic saliency-based attention system. American Society of Limnology and Oceanography (ASLO) Summer Meeting, Savannah, Georgia, USA, June 13-18, 2004.

Duane R. Edgington, Dirk Walther, Danelle E. Cline, Rob Sherlock, Karen A. Salamy, Alexis Wilson, and Christof Koch. Detecting and tracking animals in underwater video using a neuromorphic saliency-based attention system. American Society of Limnology and Oceanography (ASLO)/The Oceanography Society (TOS) Ocean Research Conference, Honolulu, Hawaii, USA, February 15-20, 2004

Walther, D., Edgington, D. The art of seeing jellies. *The Neuromorphic Engineer* (2004) **1**, 6-6

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

Walther, D., Edgington, D., Salamy, K., Risi, M., Sherlock, R.E., Koch, C. (2003). Automated video analysis for Oceanographic Research. IEEE Computer Vision and Pattern Recognition Conference (CVPR) Proceedings, Madison Wisconsin, June 16-22, 2003. (peer reviewed).

Edgington, D., Walther, D., Salamy, K., Risi, M., Sherlock, R.E., Koch, C. (2003). Automated Event Detection in Underwater Video. *MTS/IEEE Oceans 2003 Conference Proceedings*, San Diego, CA, September 2003. IEEE Press.

Edgington, D. R. and D. Walther (2002). Detection of visual events in underwater video using a neuromorphic saliency-based attention system. *Report on the 2002 Workshop on Neuromorphic Engineering*, pages 25 – 29, Telluride, CO, July 2002.

Shih Chii Liu, Jorg Kramer, Duane Edgington, Tobi Delbruck. Making direction selective cells using AER retina and spiking neurons. *Report on the 2001 Workshop on Neuromorphic Engineering*, pages 94 – 96, Telluride, CO, July 2001.

Plans for 2005

Engineering/technology development

See Development Goals for 2005, above.

Field programs

None planned.

Data reduction and analysis

The team members will do data reduction and analysis as part of the project. Additional analysis of dive video (for example, evaluating the output of the visual event detection system, and special annotations of selected videos marking frames at the coordinates of events (animals or rocks)) will be requested of our video lab project support members or MBARI science department members.

Significant changes from previous project proposal (continuing projects only)

Our focus in 2005 will be to deliver products useful to science groups, with less emphasis on short term deliverables to the video lab group.

c) Specific deliverables (new and continuing projects)

- **Products from research and development (instruments, tools, documentation, publications, presentations, etc.):**

1. The main product from this project in 2005 will be a working implementation of the visual image event detection and classification system processing digital BetaCam archived videos on shore at full resolution integrated with the VARS derived time-coded annotations, essential to individual

scientists to produce publication quality results on the occurrence of specific organisms in analyzed video.

2. A secondary product will be a demonstrated capability to scale to real-time on-shore video processing from cabled observatory-class video cameras.
3. Collaborations and explorations will be documented in whitepapers placed on the Visual Image Event Detection website and project presentations.
4. Results of the research and development will be presented at selected conferences and workshops, including the MTS/IEEE Oceans Conference, ASLO, IEEE/ACM Computer Vision and Pattern Recognition Conference. Results from the development will be published in selected peer reviewed IEEE publications as appropriate.

- **<u>How success will be measured</u>**

a) By working implementation of shore-side visual event detection and classification system, useful to science staff.
b) Demonstrated capability to scale to real-time shore-side video processing.
c) Documentation of explorations and collaborations.
d) Accepted presentations and publications at appropriate venues.
e) Successful external and/or internal follow-on project proposals.

d) Project team (new and continuing projects)

Project Manager: Duane Edgington
Lead Scientist: Jim Barry, Steve Haddock, Bruce Robison, Edie Widder
Lead Engineer: Duane Edgington
Team: Danelle Cline, Dan Davis, Duane Edgington, Christy Herren, Bill Kirkwood, Tim Meese, Steve Rock, Rob Sherlock, and Nancy Jacobsen Stout
External Collaborators, with their own funding: Dirk Walther, [Caltech], Christof Koch, [Caltech], Laurent Itti, [USC], Pietro Perona [Caltech], Marc’Aurelio Ranzato [NYU], Andres Castano [NASA JPL]
:

e) Project lab space (new and continuing projects)

We require lab space of about 24 feet by 15 feet, with typing-height surfaces all around, for our video-lab test setup (computer, video capture card, video monitor, tape deck, computer monitor), 3 workstations, and notebook computers. Our current project lab space in A244 is adequate for the project as currently scoped, with 3-4 people working on the project at any one time.

References

Kocak, D. M., N. D. Lobo, et al. (1999). "Computer vision techniques for quantifying, tracking, and identifying bioluminescent plankton." *IEEE Journal of Oceanic Engineering* **24** (1): 81-95.

Lapota, D. (1998). Long-term and seasonal changes in dinoflagellate bioluminescence in the Southern California Bight, University Of California, Santa Barbara.

Piontkovski, S. A., R. Williams, et al. (1997). "Spatial heterogeneity of the planktonic fields in the upper mixed layer of the open ocean." *Marine Ecology Progress Series* **148** (1-3): 145-154.

Widder, E. A., S. Johnsen, et al. (1999). "Thin layers of bioluminescent copepods found at density discontinuities in the water column." *Marine Biology (Berlin)* **134** (3): 429-437.

3) BUDGET JUSTIFICATION (new and continuing projects)

1.	Project computer disks and upgrades.	$3,000
2.	Spare processing node for Beowulf cluster	$2,000
3.	Matlab simulink video processing toolkit	$1,000
4.	Presentation, collaboration and participation in Neuromorphic Engineering workshop	$3,500
5.	Lab supplies and stock	$2,000
6.	Team travel to collaborators (Caltech, USC, NYU, JPL)	$5,000
7.	Travel to SPLAT collaborators (Harbor Branch, Duke)	$5,000
8.	Travel to present work (ASLO, AGU, CVPR)	$7,000
9.	Misc software tools	$3,000
10.	Compression software tools	$2,000
11.	Materials and supplies for summer intern	$1,000
12.	Matlab training	$2300
13.	Project reference books	$600
14.	RTI Control Shell	$10,000
15.	Matlab code generator	$5,000

4) MILESTONES (new and continuing projects)

	Milestone 1	Milestone 2	Milestone 3	Milestone 4	Milestone 5	Milestone 6	Total Resources[1] for each resource category for all milestones listed here
Milestone Date	Q2 2005	Q4 2005	Q3 2005	Q4 2005	Q4 2005	Q4 2005	
Milestone Description	"Large object" detection and segmentation	Object classification and recognition	Motion detection and classification	Extension to bioluminescent zooplankton	Event response camera	Annotation assistant	
EE							
SE	25	45	35	80	20	20	225
ME							
Machine Shop							
ITD Video lab		20	10			10	40

[1] This number should be the same number you entered on the resource requests for this particular resource.