NOAA Technical Memorandum NMFS-F/SPO-121

Report of the National Marine Fisheries Service Automated Image Processing Workshop

September 4-7, 2010
Seattle, Washington

by
Kresimir Williams, Chris Rooper, and John Harms (editors)

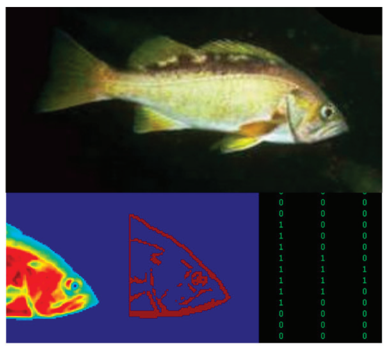
U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service

March 2012

NOAA Technical Memorandum NMFS-F/SPO-121

Report of the National Marine Fisheries Service Automated Image Processing Workshop

September 4-7, 2010
Seattle, Washington



Kresimir Williams[1], Chris Rooper[1], and John Harms[2] (editors)

[1]Resource Assessment and Conservation Engineering Division
Alaska Fisheries Science Center
7600 Sand Point Way N.E.
Seattle, WA 98115
www.afsc.noaa.gov

[2]Fishery Resource Analysis and Monitoring Division
Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle, WA 98112
www.nwfsc.noaa.gov

U.S. DEPARTMENT OF COMMERCE
John Bryson, Secretary
National Oceanic and Atmospheric Administration
Jane Lubchenco, Under Secretary and Administrator
National Marine Fisheries Service
Samuel Rauch III, Acting Assistant Administrator for Fisheries

NOAA Technical Memorandum NMFS

This document should be cited as follows:

Kresimir Williams, Chris Rooper, and John Harms (editors). 2012. Report of the National Marine Fisheries Service Automated Image Processing Workshop. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-121, 48 p.

Reference in this document to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

A copy of this document may be obtained from:

Alaska Fisheries Science Center
NMFS, NOAA
7600 Sand Point Way N.E.
Seattle, WA 98115

online at: http://spo.nmfs.noaa.gov/tm/

Contents

Introduction 1

Part 1. Description of Current and Future Image-Based Sampling Programs at NMFS Fisheries Science Centers that Could Gain from Automation of Image Processing Tasks

Alaska Fisheries Science Center
Chris Rooper, Kresimir Williams, Richard Towler, and Michael Cameron* 4

Northwest Fisheries Science Center
John Harms 8

Southwest Fisheries Science Center
Kevin Stierhoff, George Cutter, David Murfin, and David Demer* . . 13

Pacific Islands Fisheries Science Center
Jacob Asher and John Rooney* 18

Part 2. Presentation Summaries from Invited Experts in the Field of Automated Image Processing and Underwater Video Collection

New Challenges To Deep-Water Stock and Biodiversity Assessments: The Chronobiological Scenario and Automated Video-Imaging Solutions
Jacopo Aguzzi and Juan B. Company* 22

An Automated Visual Event Detection and Classification System Monterey Bay Aquarium Research Institute in Moss Landing, California, USA.
Danelle Cline and Duane Edgington* 25

Automatic and Manual Plankton Identification: a Comparison
Phil Culverhouse 28

Fish Species Recognition for Migration Monitoring and Future Applications
Dah-Jye Lee. 32

Automated Techniques for Detection and Recognition of Fishes Using Computer Vision Algorithms
Janarbek Matai, Ryan Kastner, George R. Cutter Jr., and David A. Demer* 35

StereoMeasure and StereoFeatures: Measuring Fish and Reconstructing Scenes Using Underwater Stereo Photographs
Yuri Rzhanov and George Cutter*. 38

Scantrol DeepVision – Species And Length Of Fish In The Trawl
Darren Hammersland-White. 41

An Automated, Real-Time Identification and Monitoring System for Coral Reef Fish
Joseph Wilder, Gareth J. Russell, and Paul Boyle* 42

* presenting author

Conclusions and Recommendations 45

List of Participants 48

Introduction

Images are increasingly being used in quantitative ways for various National Oceanic and Atmospheric Administration (NOAA), National Marine Fisheries Service (NMFS) activities, from fisheries surveys to ecosystem and behavioral studies. The increased number of camera-based projects recently developed within NOAA reflects evolving trends toward technology-driven methodology that has the potential to redefine established approaches to marine fisheries surveys and ecosystem studies. Image-based sampling generally provides higher spatial and temporal resolution and non-lethal sampling of organisms compared with more traditional extractive methods such as trawls. However, practical implementation of camera systems as routine scientific tools will require further development on several fronts. The primary challenge with these approaches is developing the process for extracting relevant data from the images. Manual processing of images can be a tedious and time-intensive process which is in many cases untenable due to limitations in human resources. Automation of image processing is therefore needed for these methodologies to become practical tools that can be implemented into NOAA's scientific mission.

In most industrial computer vision applications, automated image processing is a vital component of any system. Software and algorithm developments originating in other fields such as the security industry hold promise for reducing the amount of manual analysis required in marine applications. Automating the tasks of marine organism counting, classification, sizing, and movement tracking is a complex undertaking, largely due to the uncontrolled conditions of image capture and challenges of underwater photography. Despite these challenges, considerable progress has already been made in recent years, and new methodology is continually being developed though collaborations between marine scientists and computer vision experts.

This report is a summary of presentations and discussions from a workshop on automated image processing conducted in Seattle, Washington, from 4-7 September, 2010. The objective of the workshop was to examine current and future applications of automated image processing for fisheries and marine ecology research. The workshop provided a platform for representatives from all six NMFS fisheries science centers to present image–based sampling systems that are being used and developed for a wide range of purposes, including essential fish habitat research, target identification for acoustic biomass surveys, and fish behavior studies.

Experts in the field of image processing presented their past and current projects that incorporate automated processing in various stages, showing what can be achieved though automation and where the challenges lie. The majority of the projects presented by computer vision experts dealt with marine ecology or fisheries applications, even though the analytical methodology is general to the field of computer vision. Their examples illustrate the possibilities for future collaborations as automated processing solutions for image-based sampling programs continue to expand. We hope this publication will serve as a networking tool for biologists and computer vision experts and provide concrete examples of automated image processing work, as well as guidance for developing future projects.

Part 1.

Description of Current and Future Image-Based Sampling Programs at NMFS Fisheries Science Centers that Could Gain from Automation of Image Processing Tasks

An Overview of Image-Based Sampling and Research Programs at the Alaska Fisheries Science Center

Chris N. Rooper, K. Williams, R. Towler, and M. Cameron

Alaska Fisheries Science Center, NMFS, NOAA
Seattle, Washington
chris.rooper@noaa.gov

Image-based sampling at the Alaska Fisheries Science Center (AFSC) crosses all programs and divisions from stock assessment to behavioral ecology research. There are common themes in most of the research using imagery that drives the need for common types of data products. If automated image processing can be successfully implemented for some of these tasks, huge benefits may result in terms of taking individual projects from "one-off" projects into "production" mode, where they can provide routine benefit to fisheries and ecosystem management in Alaska.

Many of the studies that use imagery at the AFSC fall into the category of using video and still images for direct assessment of marine mammals, fishes and invertebrates. For example, both beluga whales in Cook Inlet, Alaska, and ice associated seals in the Alaska polar regions are assessed using aerial photo and videography (Rugh et al. 2005, Boveng et al. 2008, Cameron et al. 2009). These and other similar projects using underwater video to examine the distribution of benthic invertebrates (Rooper, unpublished data), directly enumerate individuals to calculate abundance estimates. Another line of assessment research is currently developing methods to provide count, size and species identification from underwater video that can be used with

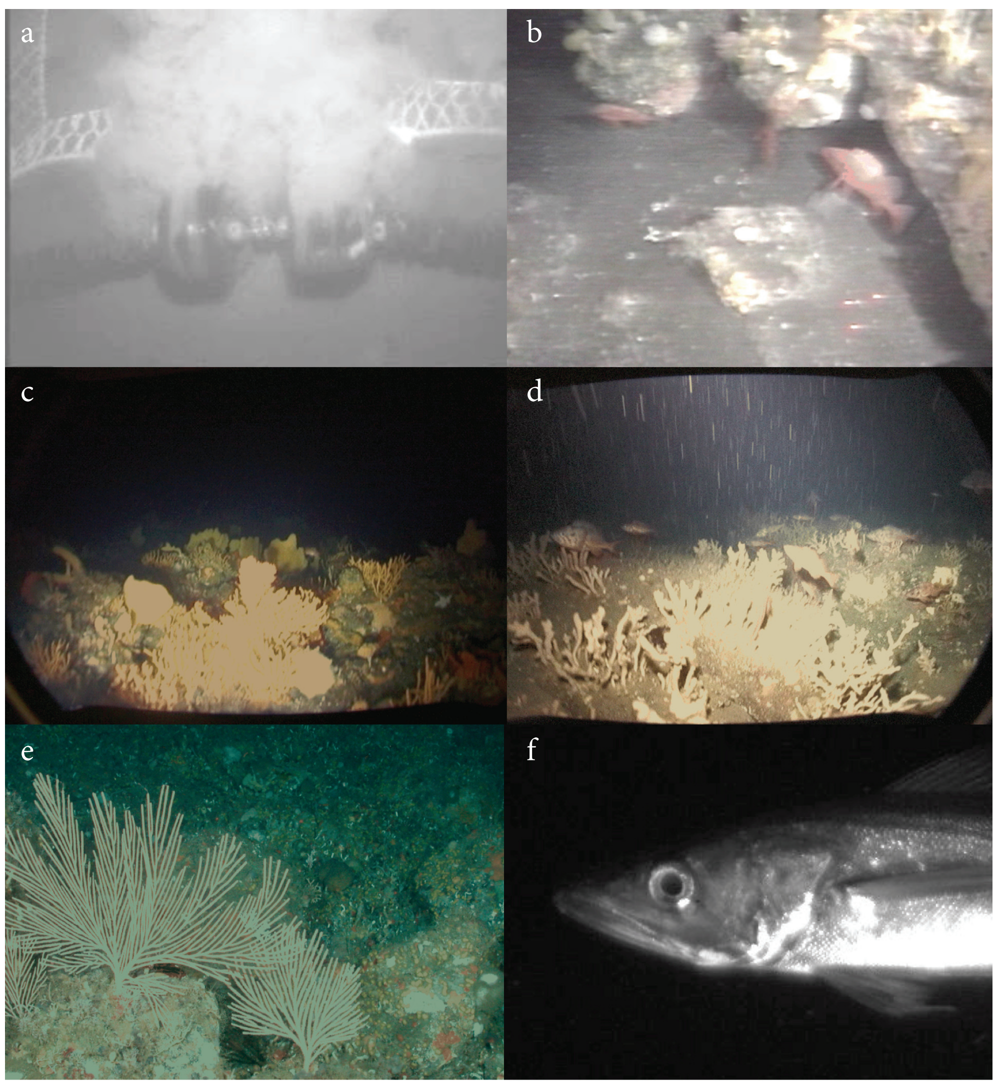


Figure 1. Imagery from research projects at the AFSC: a) still image from analog black-and-white video from a bottom trawl mounted camera showing footrope contact with the seafloor, b) still image from analog color video from study examining habitat use by rockfish, c) still image from digital video collected on study examining distribution of coral and sponge, d) still image from digital video collected on acoustic-optic survey of rockfish, e) digital still image collected on acoustic-optic survey of rockfish, and f) digital still of Pacific hake from trawl-mounted camera shot at 5 m from the fish.

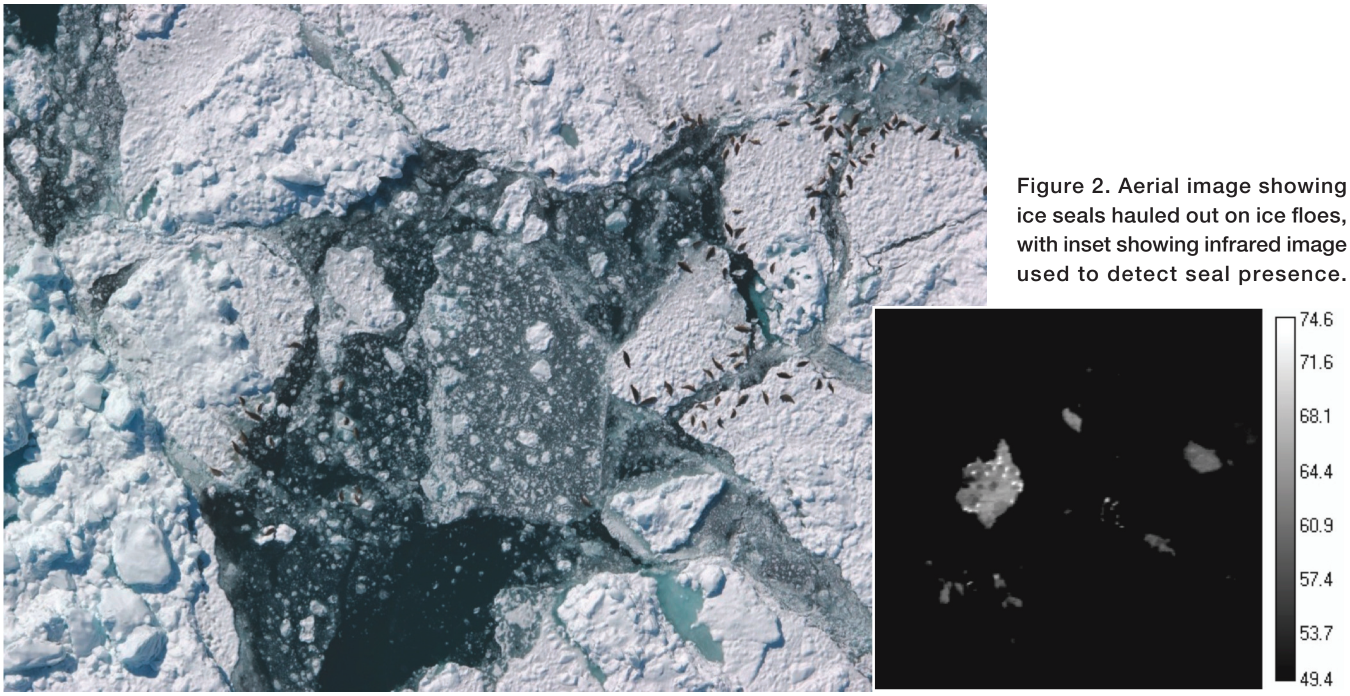


Figure 2. Aerial image showing ice seals hauled out on ice floes, with inset showing infrared image used to detect seal presence.

more traditional fisheries assessment methods such as acoustic surveys (i.e., Williams et al. 2010).

There are substantial amounts of imagery that have been collected during habitat research projects. Typically the foci of these projects have been documenting fish and invertebrate associations with substrate types (i.e., Rooper et al. 2007, Stoner et al. 2007), documenting species distributions (i.e. Woodby et al. 2009), or projects verifying substrate types for benthic mapping applications (i.e. Rooper and Zimmerman 2007, Lomnicky and McConnaughey 2008).

Other sources of imagery at the AFSC are from studies examining both research trawls and commercial fishing gear. These have included measuring the effects of commercial fishing on vulnerable habitats and recovery of these habitats (Freese 2001, Heifetz et al. 2009), estimating catchability of bottom trawl survey gear (Weinberg and Kotwicki 2008) and using imagery to visualize fish behavior in order to reduce bycatch of unwanted species (Rose et al. 2010).

The data products generally needed for all these studies are common. They include species identifications, counts of individual organisms, and size estimates of individual organisms. For some specific applications, orientation of individuals, behavioral observations, and tracking of individual movements are necessary. Finally, for studies examining benthic habitat use, estimates of substrates and percent coverage of important invertebrate and vegetative structures may be important.

The quality and quantity of images collected for AFSC projects have varied substantially, even over recent years (Fig. 1). Both digital and analog imaging systems have been used for collecting video, although most of the analog systems have been retired at this point and still images now are almost always in a digital format. The absence of standardized formats, and requirements for metadata and image storage should be addressed immediately within the AFSC to make the transition to automated image processing easier. Since most of the image collecting projects have been initiated and conducted by individual researchers, there are many different image types and databases of varying complexity associated with these images.

The need for automation is great. In a recent study in the Gulf of Alaska using video and still images from a remotely operated vehicle and a stereo drop camera, we found that even for an expert at video analysis the minimum amount of time spent searching through video to identify and count rockfish was at least 2 hours per hour of video collected. Additionally, to measure rockfish using stereo analysis took an extra 3.2 minutes per fish (including time to find and extract still images from the video source, locate a randomly selected measureable image, perform the measurement and save the data). For comparison, up to 150 rockfish caught in a bottom trawl can be manually measured for length in 15 minutes. Measuring 150 lengths using current stereo video analysis methods would take an estimated 7.5 hours.

To date, very little progress has been made in automating image processing at the AFSC. One of the most important issues to overcome is the volume of images that can be collected in a limited amount of time. In some cases, most of the images may contain little or no relevant information. For example, many of the photographs of ice in harbor seal surveys show no seals present. AFSC researchers are developing an image collecting system where an infrared camera is linked to the digital camera for still image collection. When the infrared camera senses heat above a certain threshold (indicating an ice-associated seal is present), a digital photo is taken (Fig. 2). Afterwards, manual analysis of the digital photo for species identification is undertaken. This is an innovative method to reduce the amount of time spent reviewing imagery with no seal data. Similar attempts to reduce the amount of duplicated information that has to be viewed have been the impetus for seafloor video mosaicing research at the AFSC (Lomnicky and McConnaughey 2008). Pattern recognition software is also being developed for ice-associated seals, to try to automate the process of identifying species from still images (Fig. 3). To date, these attempts encompass the entirety of automating image processing at the AFSC. They are limited to a small number of programs and projects with very specific applications.

In order to proceed with automated image analysis on a larger and more comprehensive scale, we need to improve our ability to automate image processing by setting some standards for image collection and storage, including having adequate metadata and data. We also recommend that emphasis be put on development of technologies that address the challenges that are universal to many research projects in fisheries. The tasks that need to be automated are generally common across all studies and include:

- Target identification
- Target measurement
- Target tracking

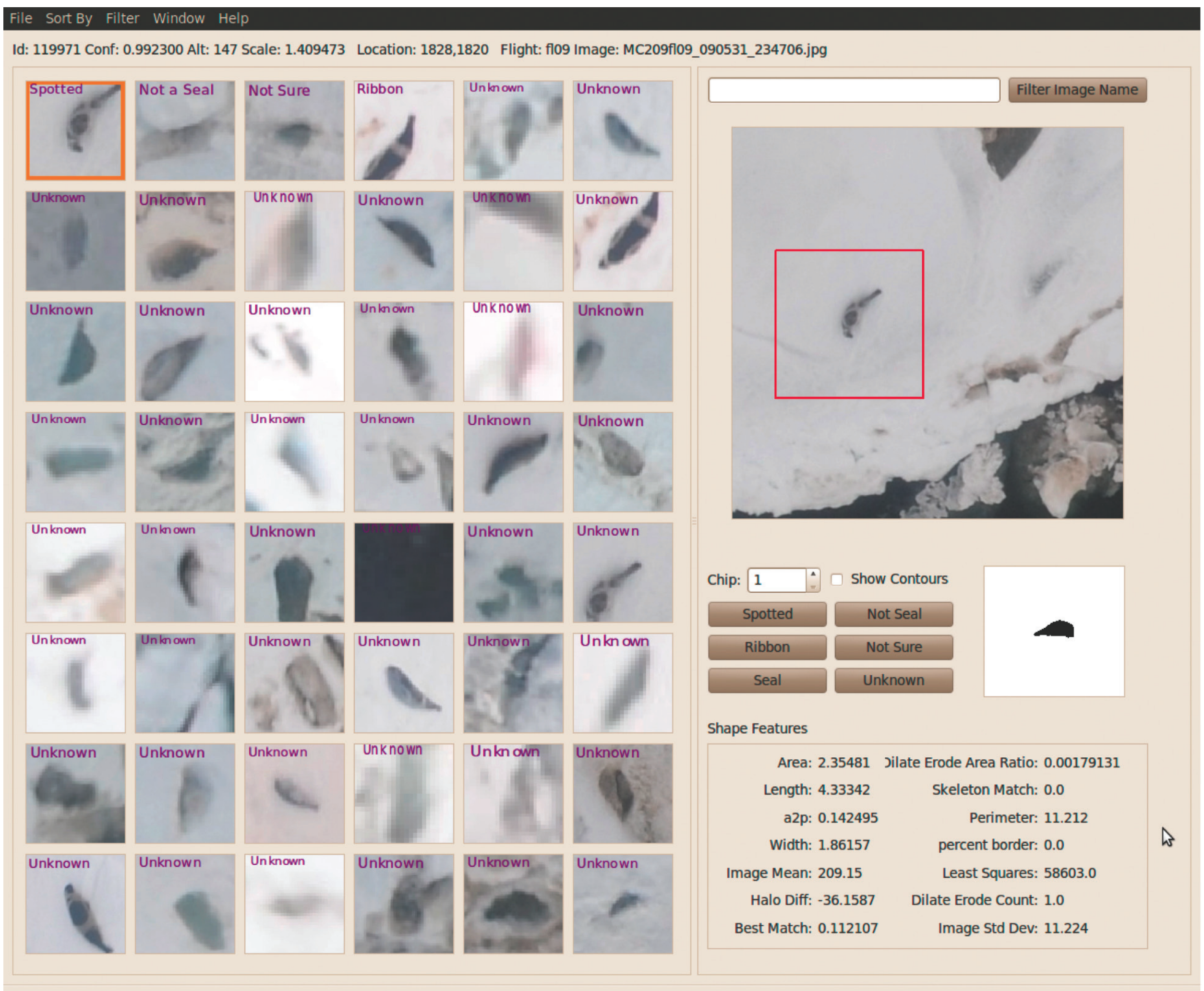


Figure 3. Pattern recognition software interface for identifying species of ice-associated seals from still images.

If we can successfully automate image processing and reduce the amount of time dedicated to extracting useful data from images, we can make photo and video data collection as useful a sampling tool for fisheries and ecosystem managers as traditional methods such as bottom trawling and acoustics. The ability to enhance the collection of data remotely with minimal impact not only benefits those who study theses marine ecosystems, but these unique habitats as well.

Citations

Boveng, P. L., J. L. Bengtson, T. W. Buckley, M. F. Cameron, S. P. Dahle, B. A. Megrey, J. E. Overland, and N. J. Williamson. 2008. Status review of the ribbon seal (*Histriophoca fasciata*). U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-191, 115 p.

Cameron, M.F., E.E. Moreland, and P.L. Boveng. 2009. Testing a ship-based unmanned aerial system (UAS) for surveying the Bering Sea pack ice. AFSC Quarterly Report (April–May–June 2009)

Freese, J. L. 2001. Trawl-induced damage to sponges observed from a research submersible. Mar. Fish. Rev. 63(3):7-13.

Heifetz, J., R. P. Stone, and S. K. Shotwell. 2009. Damage and disturbance to coral and sponge habitat of the Aleutian Archipelago. Mar. Ecol. Prog. Ser. 397:295-303.

Lomnicky, J., and R. A. McConnaughey. 2008. Integrating charting and acoustic habitat research. Sea Technol. 49(8):10-15.

Rooper, C.N., G. R. Hoff and A. De Robertis. 2010. Assessing habitat utilization and rockfish (*Sebastes* spp.) biomass in an isolated rocky ridge using acoustics and stereo image analysis. Can. J. Fish. Aquat. Sci. 67:1658-1670.

Rooper, C. N., J. L. Boldt and M. Zimmermann. 2007. An assessment of juvenile Pacific ocean perch (*Sebastes alutus*) habitat use in a deepwater nursery. Estuar. Coast. Shelf Sci. 75:371-380.

Rooper, C. N. and M. Zimmermann. 2007. A bottom up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification. Cont. Shelf Res. 27:947-957.

Rose, C. S., J. R. Gauvin, and C. F. Hammond. 2010. Effective herding of flatfish by cables with minimal seafloor contact. Fish. Bull., U.S. 108:136-144.

Rugh, D. J., K. E. W. Sheldon, C. L. Sims, B. A . Mahoney, B. K. Smith, L. K. Litzky, and R. C. Hobbs. 2005. Aerial surveys of belugas in Cook Inlet, Alaska, June 2001, 2002, 2003, and 2004. U.S. Dep Commer., NOAA Tech. Memo. NMFS-AFSC-149, 71 p.

Stoner, A. W., M. L. Spencer, and C. H. Ryer. 2007. Flatfish-habitat associations in Alaska nursery grounds: Use of continuous video records for multi-scale spatial analysis. J. Sea Res. 57:137-150.

Weinberg, K., and S. Kotwicki. 2008. Factors influencing net width and sea floor contact of a survey bottom trawl. Fish. Res. 93:265-279.

Williams, K., C.N. Rooper, and R. Towler. 2010. Use of stereo camera systems for assessment of rockfish abundance in untrawlable areas and for recording pollock behavior during midwater trawls. Fish. Bull., U.S. 108:352-362.

Woodby, D., D. Carlile, and L. Hulbert. 2009. Predictive modeling of coral distribution in the central Aleutian Islands, USA. Mar. Ecol. Prog. Ser. 397:227-240.

Image-Based Research at the Northwest Fisheries Science Center

John Harms

Northwest Fisheries Science Center, NMFS, NOAA
Seattle, Washington
john.harms@noaa.gov

Figure 1. Two-part video camera and recording system

Figure 2. Screen capture of video during a study of the performance of the footrope of the net used in the NWFSC's West Coast groundfish bottom trawl survey.

The Northwest Fisheries Science Center (NWFSC) has conducted a variety of image-based research projects in support of monitoring, assessing, and managing the U.S. West Coast's marine resources and habitats. Video and still imagery analyses have been used to examine the performance of the gear used in NWFSC fishery-independent research surveys; develop fishing gear that reduces bycatch of non-target species in commercial fisheries; classify seafloor habitats along the West Coast; identify marine mammals for population monitoring; and differentiate among salmonid species in hatcheries, lakes, and mixed-stock fisheries. NWFSC researchers have collaborated with a diverse group of partners including the Alaska Fisheries Science Center (AFSC), Pacific Islands Fisheries Science Center (PIFSC), Pacific States Marine Fisheries Commission (PSMFC), Woods Hole Oceanographic Institution (WHOI), Oregon State University (OSU), and commercial and sportfishing industries along the West Coast. These partnerships have played an important role in strengthening the quality of the research, leveraging the existing resources of various organizations to control costs, and assuring the credibility of results.

Program Overview

Much of the NWFSC's image-based research has supported the monitoring and assessment of the 90+ species managed under the Pacific Coast Groundfish Fishery Management Plan as required by the Magnuson-Stevens Fishery Conservation and Management Act. A significant portion of this work is conducted using a custom-built video recording system developed in partnership with the AFSC. This system consists of a small sled that can be outfitted with different cameras and light arrays connected to a titanium pressure housing which contains the system's batteries, electronics, and video recorder to capture the footage *in situ* to mini-DV tapes (Fig. 1). Several separate systems based on this portable and adaptable design have been used during the course of many of the NWFSC's image-based research projects.

The system has been mounted on the net and rigging used in the NWFSC's annual trawl survey to evaluate gear performance and identify patterns of fish behavior in response to the net during fishing operations. In one project, the camera system was suspended in front of the net opening with the camera facing aft to collect images of the trawl's footrope and assess the potential for flatfish escapement under or around the footrope (Fig. 2). The system has also been mounted along the net's bridles and door sweeps to capture video to evaluate whether the interaction of the net's

Figure 3. Screen capture of video from a study of potential flatfish “herding” by the trawl survey net’s door sweeps and mud gear.

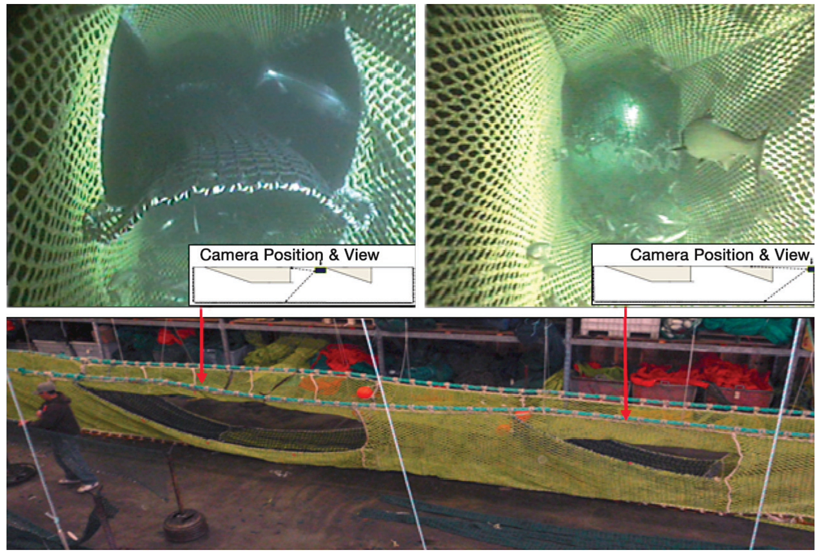


Figure 5. Screen captures of video and a still image of a bycatch reduction device (BRD) developed for use in the west coast commercial hake fishery. The upper left image depicts a Chinook salmon (*Oncorhynchus tshawytscha)* exiting out the starboard side of the forward escape window. The upper right image depicts a Chinook salmon exiting out the starboard side of the aft escape window. The bottom image provides a port side view of the two escape windows.

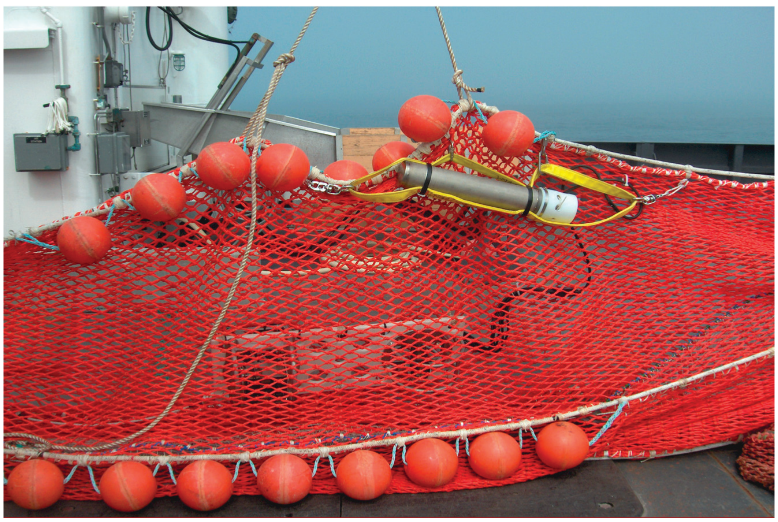


Figure 4. Sled and housing system mounted within a midwater trawl net to visually verify acoustic backscatter collected during the acoustic survey for hake.

mud gear with the seafloor effectively “herds” flatfish between the trawl doors and the net opening (Fig. 3). Both projects generated useful information about net performance and will help reduce uncertainty about the assumptions used in stock assessments that include trawl survey data (Bryan et al., In prep.)

Variations of this camera system were also used for conservation engineering within other NWFSC surveys and within the commercial groundfish industry. The Center’s acoustic survey for hake conducts midwater trawls to provide biological information and a physical groundtruth of the acoustic backscatter data. By mounting a video system inside the survey net (Fig. 4) and leaving the codend open, the camera collects video for a visual, rather than physical groundtruth of fish in the net that can also be used for length frequency analyses. Because the codend is open and no fish are actually captured, survey mortality using this technique is near zero, and larger portions of targeted schools can be sampled than would be feasible with a closed net. Another application of this system included developing less-lethal methods of visually sampling other midwater species such as widow rockfish (*Sebastes entomelas*; Bonacci et al. In prep.)

This system was used for conservation engineering within the commercial hake fishery to reduce bycatch of salmon and rockfish. NWFSC researchers collaborating with the PSMFC and the commercial fishing industry analyzed video from cameras mounted inside commercial nets to help design special panels to serve as bycatch reduction devices (BRDs). BRDs reduce incidental catch of salmon and rockfish in the commercial hake fishery (Fig. 5; Lomeli and Wakefield In press) NWFSC, PSMFC, Oregon Department of Fish and Wildlife, and the commercial shrimp industry have also collaborated in deploying this system to collect video to evaluate footrope performance within commercial shrimp trawls (Mark Lomeli, NWFSC, August, 2010, personal communication).

Another version of this basic system was mounted on a towed aluminum sled and configured to send real-time video to the vessel where it is viewed and recorded onto mini-DV tapes (Fig. 6). This arrangement provides an efficient means of collecting visual observations of the seafloor that can be used to identify and classify specific habitat types at locations of interest. This towed sled system has also been used in partnership with the AFSC and OSU to provide location-explicit groundtruthing for habitat maps developed with multi-beam acoustic data and in partnership with the commercial fishing industry to collect video of schooling midwater rockfish to supplement acoustic surveys (Ressler et al. 2009).

Non-groundfish image-based research projects at the NWFSC include identification of individual killer whales for population monitoring within Puget Sound. Photographs of Puget Sound's southern resident killer whales (SRKWs) are manually reviewed, interpreted, and linked to known individuals within the SRKWs' three pods (Brad Hanson, NWFSC, August 2010, personal communication). Other research includes computer-assisted analysis of photographs of salmonids overlaid by a truss connecting key morphometric landmarks. This approach has been applied to differentiate among juvenile salmonid species at hatcheries and evaluate whether genotypic differences within species may be identified phenotypically (e.g., Winans et al. 2003).

Automated Analysis of Existing Video Collections

The potential for automated image analysis of video generated with the NWFSC's primary underwater camera system may be limited due to the design of the projects, their implementation, and the resolution and overall clarity of the video. Most of these projects were designed to be reviewed and interpreted by human analysts for qualitative observations, and therefore may not be appropriate for analysis via machine vision or other automated means. For example, the lighting or background may not provide sufficient contrast of the items of interest for image segmentation and recognition. In other cases, mud clouds from gear interactions with the seafloor, excessive light reflectance from marine snow, or poor visibility in general may present problems for automated processing. Certainly the potential improvements in efficiency and volume that can be attained through automation warrant some exploration with our existing video, however, it is likely that the best opportunities for successful image processing may reside with some of the NWFSC's newer projects.

AUV

Recent projects have benefited from hardware, software, and research designs that are more explicitly aimed to generate imagery suitable for automated analyses. For example, the NWFSC has partnered with the PIFSC and WHOI to operate an autonomous underwater vehicle (AUV) to conduct various resource and habitat monitoring projects throughout the Pacific Ocean. The AUV is an adaptable, modular platform that can be configured with multiple digital still and video cameras, various light arrays including strobes, and sensors such as a multibeam sonar (Fig. 7). Current applications include providing visual observations to survey mesophotic coral reefs and collecting imagery to help calculate density estimates of fish and

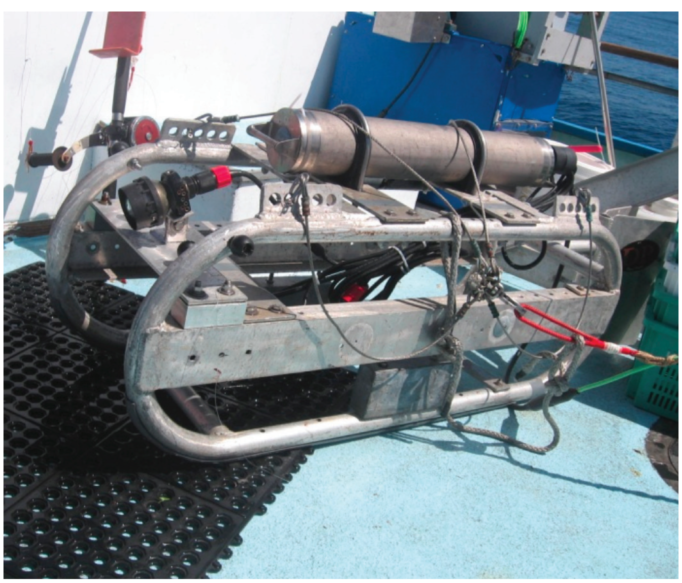


Figure 6. Towed video camera sled used to identify habitats in real-time at sites sampled during NWFSC groundfish surveys.

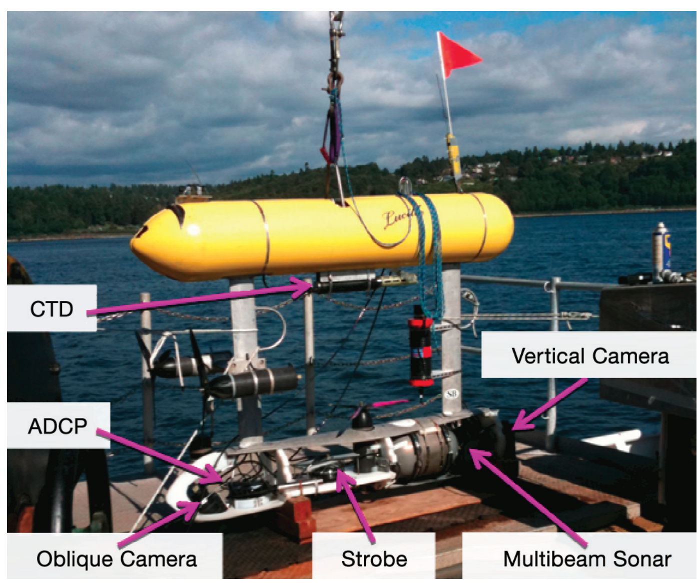


Figure 7. Image of the autonomous underwater vehicle operated jointly by the NWFSC and PIFSC showing some of the various components and sensors that can be mounted on the vehicle.

invertebrates associated with rocky, "untrawlable" habitats. Images may be examined singly or "stitched" together in a mosaic to examine larger areas of the seafloor. The high-resolution digital still cameras, strobe lighting, and the vehicle's ability to maintain constant altitude (e.g., 3 m) off the seafloor enable the collection of high quality images suitable for automated processing.

QUOTAS

Another project with potential for automated image analysis is the quantitative optic trawl analysis system (QUOTAS). Currently in development, this system employs six separate camera, laser-pair, and LED arrays that are oriented in a hexagonal "ring" and mounted inside a research trawl net. Fish captured in the net are illuminated and photographed from the six camera and LED arrays as they pass through the ring (Fig. 8). Each of the six assemblies of camera, laser-pair, and LED arrays is programmed to coordinate with the opposing assemblies to prevent a flash from one array from "washing out" another camera's photograph. The ring's frame consists of panels that provide a dark background to improve contrast with most fish species of interest and reduce excessive light reflectance back into the camera. The QUOTAS system captures images of the same fish from different angles to aid in species identification and length estimates thereby realizing the efficiency of research trawling to sample large areas while reducing mortality to near zero. This system is devised to generate images readily analyzed via automated protocols. Photographs are taken of each fish from multiple angles to increase the likelihood of capturing a fish in an orientation amenable to effective image segmentation and color analysis thus aiding in automated species identification and measurement (Fig. 9). Photographs can also be stitched together to form continuous images of the entire tow from six different angles that will be used to generate density estimates for the area swept by the net. The QUOTAS system can also be combined with other means of image collection in the same trawl net such as downward and horizontal-looking cameras, stereo cameras, and multi-beam sonars providing a large assortment of intriguing potential for automated analysis.

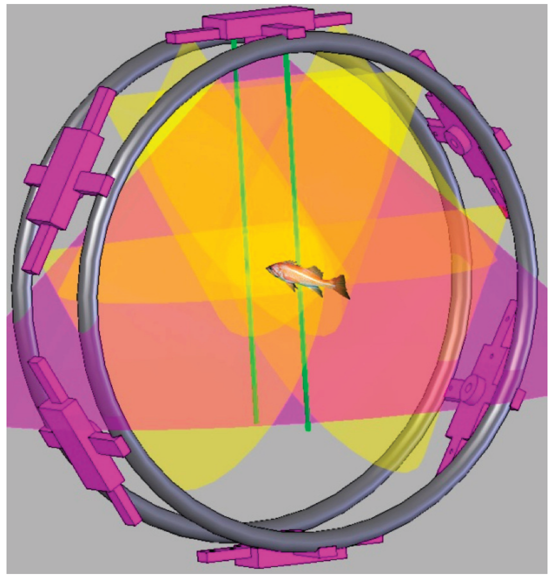


Figure 8. Diagram of the QUOTAS system showing the hexagonal orientation of the six camera/LED array/laser assemblies.

The Future

The primary mission of the six NMFS fisheries science centers of monitoring and assessing the Nation's marine fisheries resources, and habitats has remained generally constant over time. However, technology has improved the tools available to researchers. Direct visual observations of fish, their habitats, and other marine resources is an important component in effective monitoring programs, but often comes with the drawbacks of extensive and sometimes tedious manual review of the collected imagery. As the ability to generate higher resolution images increases and the cost of generating those images decreases, opportunities for automating their review and analysis are expanding within the field of marine science. Further, the technology underlying image processing has exploded in the past decade—driven in part by advances in the security and law enforcement sectors—and the benefits from these advances are spilling over into other fields.

Automated image processing holds considerable potential to reduce the amount of time researchers spend reviewing imagery without the problems inherent with manual review such as individual bias and fatigue. Researchers are then free to spend more time on analysis and interpretation. Clearly, there are significant barriers that constrain an organization's ability to implement automated image processing schemes on a large scale: high initial investment in gear, equipment, and software; potentially steep learning curves associated with the hardware and software; trial and error in collecting imagery of sufficient quality for automated analysis; and developing appropriate QA/QC protocols among many others. However, because visual observations are an adaptable,

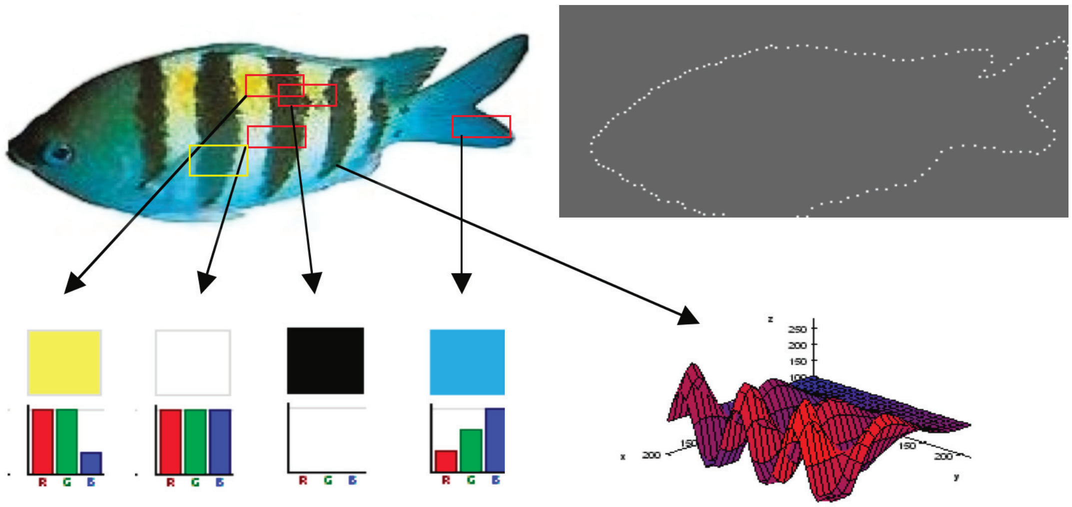


Figure 9. Diagram of potential image segmentation and color spectrum analysis of imagery collected by the QUOTAS system to aid in automated species identification.

widely-used, and a non-lethal means of data collection and likely to remain an essential component of marine science, researchers stand to benefit by tailoring their image-based projects to generate output appropriate for automation. These investments of time and resources into forward-looking technologies can yield improvements not only in terms of volume and efficiency, but also in accuracy of results.

Part of the process of making tools such as automated image processing more affordable and accessible is building partnerships between the scientists who can benefit from these new technologies and the universities, businesses, and organizations that develop them. Whether through structured workshops such as this one, or informal relationships built on mutual interests, it is essential to maintain ongoing communication about research needs, funding opportunities, and scientific challenges to help ensure that the best available tools continue to be utilized in our research.

Citations

Bonacci, L.A., W.W. Wakefield, and C.R. McGourty. In prep. Developing a non-extractive survey method for widow rockfish (*Sebastes entomelas*): a pilot study off the central Oregon coast. Available from NWFSC.

Bryan, D.R., K.L. Bosley, W.W. Wakefield, A.C. Hicks, and M.A. Haltuch. In prep. Video evidence of flatfish herding by NWFSC groundfish survey trawl warps. Available from NWFSC.

Lomeli, M.J.M., and W.W. Wakefield. In press. Efforts to reduce Chinook salmon (*Oncorhynchus tshawytscha*) and rockfish (*Sebastes* spp.) bycatch in the U.S. west coast Pacific hake (*Merluccius productus*) fishery Fish. Res.

Ressler, P.H., G.W. Fleischer, G.W., V.G. Wespestad, and J.H. Harms. 2009. Developing a commercial vessel-based stock assessment survey methodology for monitoring the U.S. west coast widow rockfish (*Sebastes entomelas*) stock. Fish. Res. 99:63-73.

Wakefield, W.W., K.M. Matteson, R.W. Hannah, M.J. Lomeli, and S.J. Parker. 2010. Observations of fish behavior in the vicinity of a selective flatfish trawl footrope using DIDSON imaging sonar. Poster presentation, Western Groundfish Conference. Juneau, AK.

Winans, G.A., S. Pollard, and D.R. Kuligowski. 2003. Two reproductive life history types of kokanee, *Onchorynchus nerka*, exhibit multivariate morphometric and protein genetic differentiation. Env. Biol. Fish:67:87-100

An Overview of Visual Survey Research and Imaging Technology Development at the Southwest Fisheries Science Center

Kevin Stierhoff, George Cutter,
David Murfin, and David Demer

Southwest Fisheries Science Center, NMFS, NOAA
La Jolla, California
kevin.stierhoff@noaa.gov

Overview of Visual Survey Research and Technology at the SWFSC

The Fisheries Resources Division (FRD) at the Southwest Fisheries Science Center (SWFSC) in La Jolla, California has been conducting visual surveys of benthic fishes and invertebrates since 2001. Initially, a remotely operated vehicle (ROV) program was developed by the Benthic Resources Group to monitor populations of a newly-listed endangered species, white abalone (*Haliotis sorenseni*). However, the research focus of this program soon expanded to include surveys of market squid (*Doryteuthis opalescens*) spawning habitat (Zeidberg et al. 2011), distribution and abundance of groundfishes (primarily rockfishes in the genus *Sebastes*), and benthic habitats throughout southern California (CA). These ROV surveys have generated a vast collection of images, including nearly 900 hours of video and 37,000 high resolution photographs.

The Advanced Survey Technologies (AST) Group at SWFSC has developed camera systems, and is collaborating to develop algorithms for automated detection, measurement, and identification of fish in underwater video, still, or stereo images (Matai et al., this volume; Rzhanov and Cutter, this volume). The AST group has developed single- and stereo-camera systems that are deployed by divers or from vessels as tethered systems, attached to an ROV, or placed on moorings or landers (Fig. 1). The AST group has also developed a towed, undulating, optical and environmental sampling system (FasTowCam)

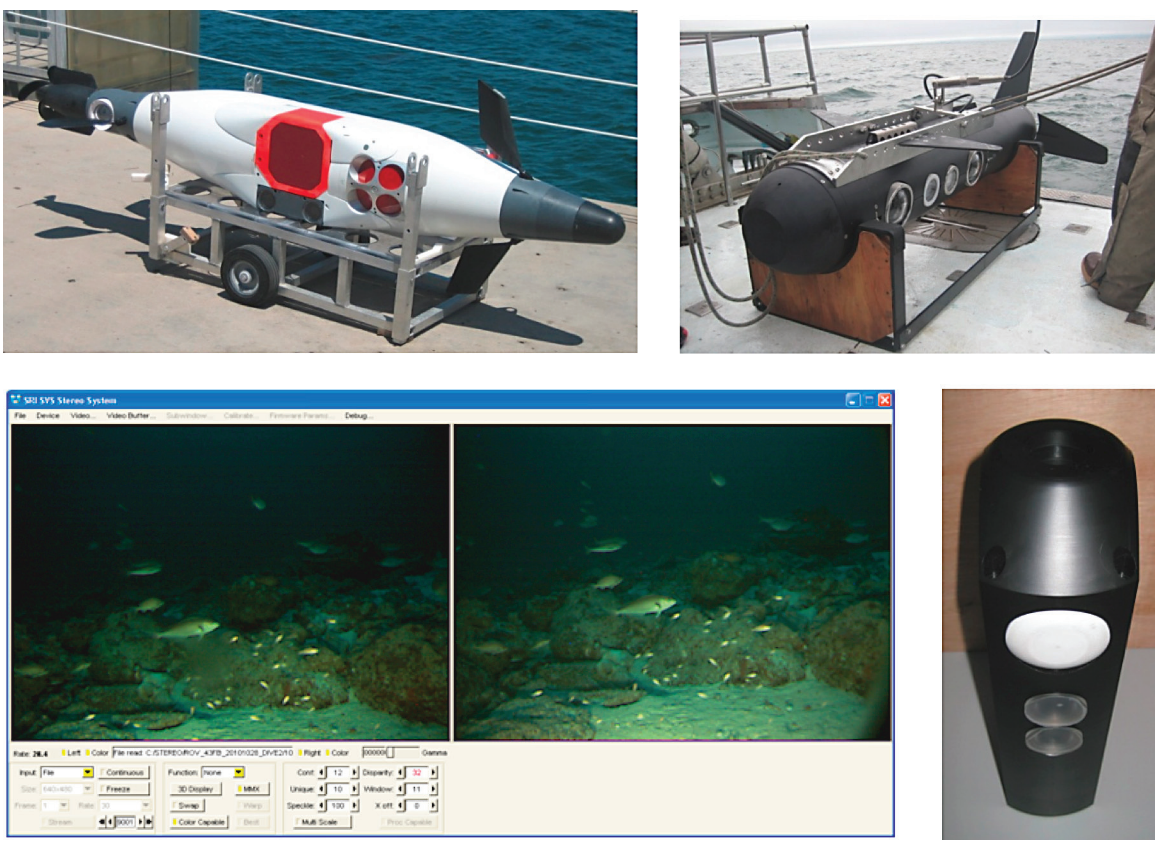


Figure 1. Examples of visual survey technologies developed by the Advanced Survey Technology Group. Clockwise from top left: NMFS autonomous underwater vehicle (AUV); the towed, undulating, optical and environmental sampling system (FasTowCam) with integrated stereo camera, strobes and CTD; the self-contained micro-echosounder system with adaptive-sampling camera (Acoustic-Optical Sampler; AOS); and images of ocean whitefish (*Caulolatilus princeps*) from the self-contained stereo camera deployed on the SWFSC ROV at 43-Fathom Bank.

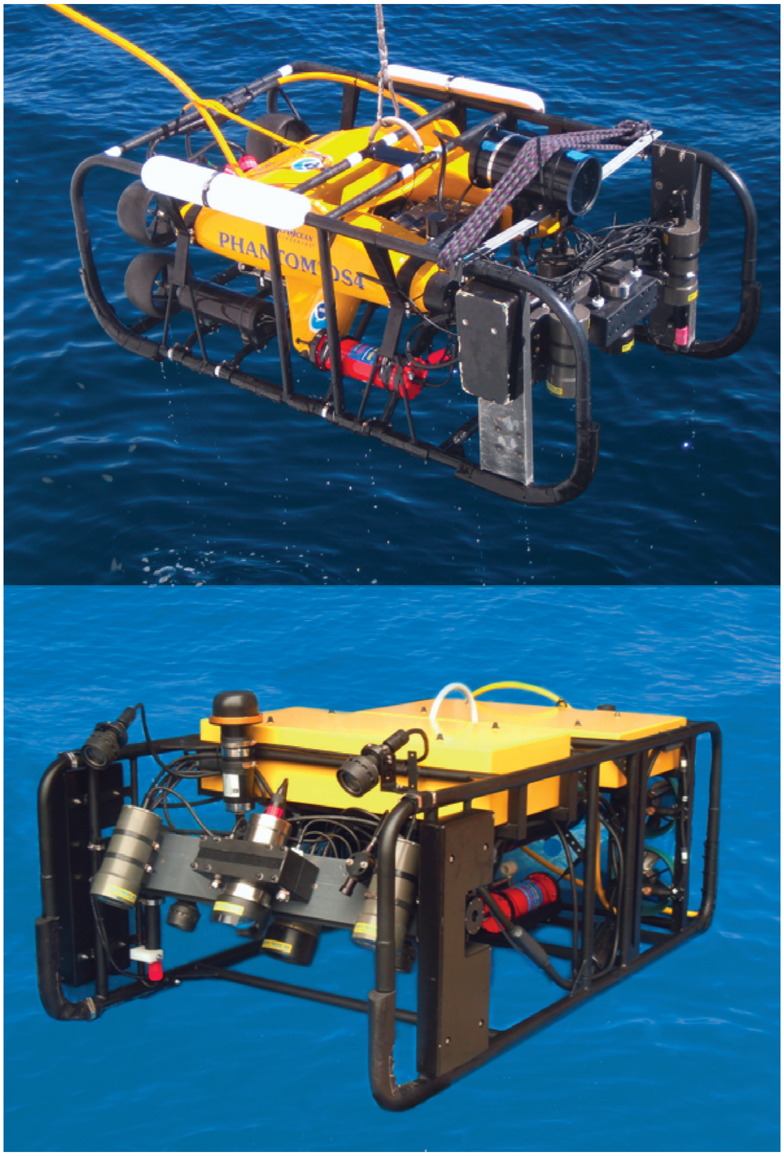


Figure 2. The SWFSC Phantom remotely operated vehicle (ROV, top) with the Videre stereo camera system attached above the camera tilt tray, and the new, Mini Zeus high-voltage, high definition (HDHV) equiped ROV (bottom).

with integrated stereo camera, strobes and connectivity, temperature and depth sensor (CTD) to augment acoustic surveys of coastal pelagic species (CPS; e.g., sardines, anchovies, and jack mackerel). Additionally, AST develops and operates the NMFS autonomous underwater vehicle (AUV, http://swfsc.noaa.gov/AUV/) that includes a stereo camera system adjacent to a scientific echosounder to collect images of acoustic targets for identification and measurement. Similarly, AST has developed a self-contained micro-echosounder system with adaptive-sampling camera (Acoustic-Optical Sampler; AOS) which was designed for deployment on large marine animals (e.g., elephant seals), and is opportunistically deployed on multiple buoys.

The need to streamline processing of stereo images collected during all of these surveys has led to the development of the semi-automated stereo-image measuring software StereoMeasure; algorithms for automated detection and recognition of fish; and the StereoFeatures application that combines the recognition algorithms and three-dimensional reconstruction for measurement and identification (Matai et al., this volume; Rzhanov and Cutter, this volume). Each of these systems will be detailed below.

SWFSC Remotely Operated Vehicle (ROV) System

Since 2001, the Benthic Resources Group at the SWFSC has been using a modified Deep Ocean Engineering Phantom ROV to conduct visual surveys of benthic fishes and invertebrates (Fig. 2). In its present configuration, the Phantom ROV is equipped with a forward-looking color-video camera (Sony FCB-IX47C with 468x720 lines of resolution and an 18x optical zoom) and a high-resolution-still camera with 4x zoom (Insite Pacific, Inc. Scorpio with Nikon Coolpix 995). The ROV is tracked in real-time using a combination of a differential GPS mounted on the ship, an ultra-short baseline (USBL) acoustic tracking system (ORE Offshore TrackPoint II-Plus), and a Doppler velocity log (Workhorse Navigator, Teledyne RDI). Additional sensors include a CTD (Citadel 2" Micro-CTD, Teledyne RDI), oxygen optode (Model 3975, Aanderaa Instruments), scanning sonar (MS1000, Kongsberg Mesotech), and laser caliper system for measure objects and calculating field of view. All navigation and oceanographic data are synchronized and logged using integrated navigation software (WinFrog, Fugro Pelagos, Inc.).

The Benthic Resources Group recently completed the development of a custom ROV to replace the Phantom system (Fig. 2; http://swfsc.noaa.gov/HDHV-ROV/). Improvements include the replacement of the standard definition color video camera with an Insite Pacific high-definition Mini Zeus video camera, quieter and more powerful brushless DC-powered thrusters (Technadyne), and a tether with three optical fibers that significantly improves the bandwidth for transmission of video and data collected from onboard instrumentation. This system should greatly improve our ability to detect and identify cryptic organisms and minimize the impact of the ROV on fish behavior.

The Use of ROV Surveys to Quantify Fishes and Invertebrates

The primary focus of the ROV program has been to monitor populations of the endangered white abalone in southern California since being listed in 2001. Surveys were conducted at several locations where white abalone were once abundant or present. A combination of multibeam sonar mapping and ROV strip transects were used to comprehensively map white abalone habitat and estimate the densities, abundances, size distributions, and group sizes for their sub populations (Butler et al. 2006). In 2008 and 2010, additional ROV surveys at one site indicated that the white abalone population has continued to decline sharply (Butler et al. In prep.). In both studies, white abalone were identified *in situ* and counted from the video, with the aid of higher-resolution photographs. Due to the cryptic appearance of these abalone, which closely resemble the algae-encrusted rocks and macroalgae on which they reside, automated detection and classification of abalone in images is exceptionally challenging.

Other work has included ROV surveys to quantify the populations of cowcod (*Sebastes levis*) and several species of severely depleted rockfishes and other groundfishes that inhabit deep, rocky, offshore banks (Fig. 3). A substantial subset of those surveys involved a collaborative, optically assisted acoustic survey technique (COAST), wherein active acoustic surveys were used to provide estimates of fish biomass and seabed type over large areas (Demer et al. 2009), and subsequent ROV surveys provided species composition, size distribution, habitat associations, and seabed classification. The data from the ROV surveys are used to apportion acoustic backscatter to various species groups and size classes. In some cases, substantially more post-processing time and effort is required to identify and quantify the abundance, species composition, and size distribution of observed fishes compared to the abalone surveys. Consequently, this is an area of research that would greatly benefit from the ability to automatically detect, measure, and classify fish targets by species.

One area where automated image processing techniques has greatly improved our existing visual survey techniques has been the development of a 3Beam© quantitative measurement system that is used to more accurately quantify the area searched during strip transects with the ROV (Fig. 4, Pinkard et al. 2005). In brief, the 3Beam software detects the location of parallel lasers in compressed video frames, and uses the pitch, roll, altitude, and camera viewing angle to compute the width of the field of view at a user-defined time or distance interval. The software also allows the analyst to review and correct erroneous laser detections from the automated algorithm. In combination with high-resolution and highly accurate distance measurements from the DVL (±1% over 1,500m, Stierhoff et al. In prep.), more accurate estimates of search area and species density can be calculated.

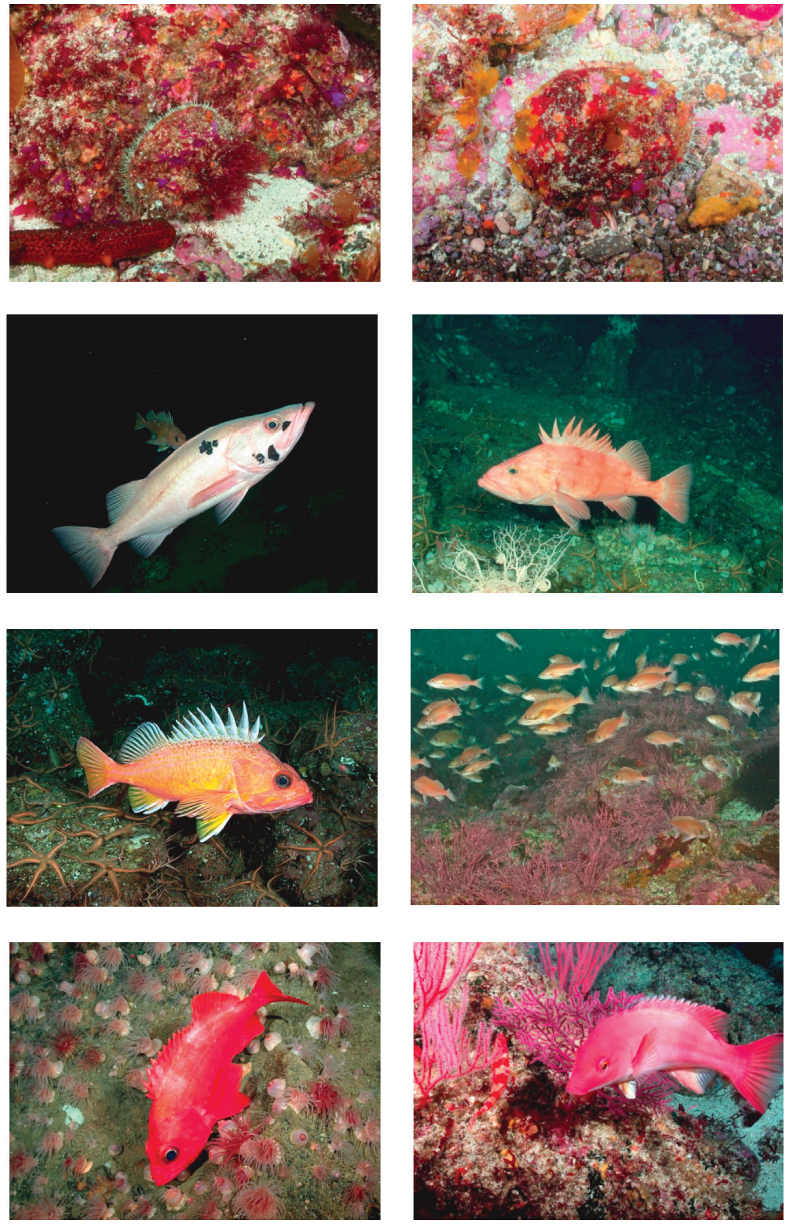


Figure 3. An example of still images collected using the SWFSC remotely operated vehicle (ROV). Clockwise from top left: white abalone (*Haliotis sorenseni*); white abalone; cowcod (*Sebastes levis*); a school of squarespot rockfish (*S. hopkinsi*); a California sheephead (*Semicossyphus pulcher*); splitnose rockfish (*S. diploproa*), greenspotted rockfish (*S. chlorostictus*); and a bocaccio rockfish (*S. paucispinis*).

Areas for Improvement on Traditional Visual Survey Techniques

Two of the greatest challenges to this research have been the time required to post-process video observations and to provide sufficient length measurements to construct size distributions for fishes of interest. For a comprehensive community study, where all encountered species are identified and enumerated, it can take as long as 5 hours to annotate 1 hour of video when fish aggregations are dense, diverse, or both; when species are cryptic; or when habitats are highly complex. This analysis time can be significantly reduced, however, when the analysis is restricted to only a few key species of interest, by the use of on-the-fly annotation hardware such as programmable keyboards, or both. Perhaps more problematic is the difficulty in measuring fishes using the laser-caliper system when fish are smaller than the laser spacing, when the lasers cannot be placed on the fish target, or when fish are not perpendicular to the camera and laser system. These constraints on the visual-survey methods has motivated the development of automated methods for detecting and classifying fishes and also the development of stereo-camera systems and image analysis software for detecting features and measuring targets of interest in stereo-images. Given the present and growing number of visual-survey platforms utilizing single- and stereo-camera systems, the development and improvement of automated image-processing methods could provide considerable savings in time and resources.

Toward Automation of Detection, Measurement and Classification of Fish

The limitation of single-camera laser-caliper systems for measurement of fishes has led the AST group to acquire and develop stereo-camera systems. Images from calibrated stereo or multi-view camera systems enable estimation of three-dimensional coordinates of any point imaged by multiple cameras. This feature allows measurement of distances in three dimensions using two or more images, for targets of any orientation within the field of view. Resulting measurements of fish sizes are critical to scattering models for interpretation of acoustic data and accurate estimation of fish biomass, and for characterization of populations from visual-survey data.

More generally, the AST group has procured and developed imaging systems designed to enable the identification and measurement of organisms during various types of surveys covering a variety of habitats from rocky banks to open-ocean pelagic systems. The large numbers of images produced by these systems and the ROV need to be efficiently and consistently analyzed to detect, identify, and measure organisms. The highly effective algorithms for detection and recognition of faces, for example, or other targets in air, motivate development of similar methods for automatically detecting and recognizing fish and other organisms from these underwater systems. However, the seawater medium poses challenges not encountered in air, and complicates direct

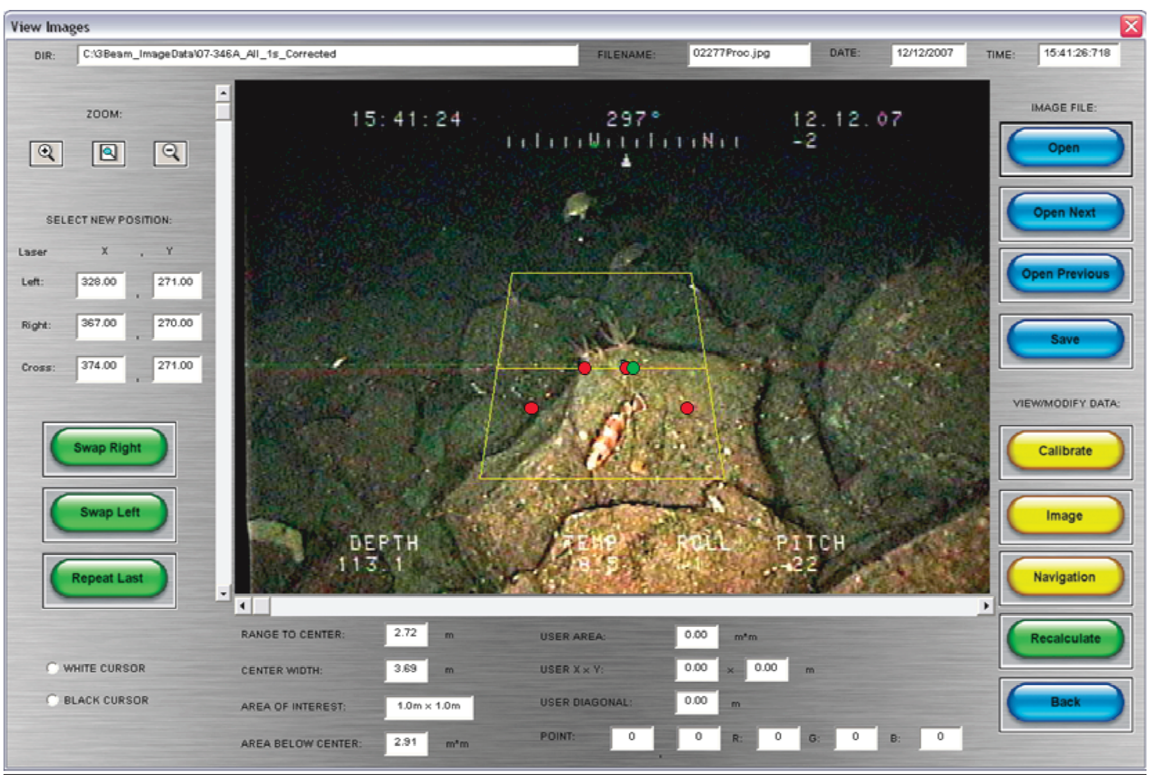


Figure 4. A screen grab from the 3Beam quantitative measurement software, enhanced to illustrate the location of the lasers on the bottom (red and green circles).

implementation of existing algorithms, leading to adaptations of common algorithms. Medium properties can be affected by various conditions that affect the scattering, absorption, and color of light; for example the presence of plankton, suspended particles, or both, and also the types and sizes of such particles.

Toward these goals the AST group is developing methods to automatically process images from single or multiple camera systems to 1) assist image analysts with more automated measurements of fishes (e.g., StereoMeasure; Rzhanov and Cutter, this volume); 2) automatically detect, and identify organisms by adapting algorithms (e.g., Viola-Jones, and principal components analysis) developed for face recognition (Matai et al., this volume); and 3) identification and reconstruction of the entire three-dimensional (3-D) scene including fish targets using algorithms adapted from SIFT or SURF (Rzhanov and Cutter, this volume) for measurement and classification of fish using the combination of shape and pattern by adaptation of recognition algorithms (from stage 2) to stereo images. Such methods will reduce the burden on analysts, increase the rate of analysis, and enable adaptive behaviors of autonomous vehicles.

Citations

Butler J.L., M. Neuman, D. Pinkard, R. Kvitek, and G.R. Cochrane. 2006. The use of multibeam sonar mapping techniques to refine population estimates of the endangered white abalone (*Haliotis sorenseni*). Fish. Bull., U.S. 104:521-532.

Butler J.L., K.L. Stierhoff, and M. Neuman. In prep. On the road to extinction: Population declines of the endangered white abalone, *Haliotis sorenseni*. Available from SWFSC, 3333 North Torrey Pines Court La Jolla, CA 92037-1023.

Demer D.A., G.R. Cutter, J.S. Renfree, and J.L. Butler. 2009. A statistical-spectral method for echo classification. ICES J. Mar. Sci. 66:1081-1090.

Matai J., R. Kastner, G.R. Cutter, and D.A. Demer. This volume. Automated techniques for detection and recognition of fishes using computer vision algorithms.

Pinkard D., D.M. Kocak, and J.L. Butler. 2005. Use of a video and laser system to quantify transect area for remotely operated vehicle (ROV) rockfish and abalone surveys. MTS/IEEE OCEANS ‘05 3:2824-2829.

Rzhanov Y., G.R. Cutter. This volume. StereoMeasure: measuring fish underwater using stereo photographs.

Stierhoff K.L., J.L. Butler, D.M. Kocak, D. Pinkard-Meier, D.W. Murfin. In prep. Improved methods for quantitative underwater visual surveys. Available from SWFSC, 3333 North Torrey Pines Court La Jolla, CA 92037-1023.

Zeidberg L.D., J.L. Butler, D. Ramon, A. Cossio, K.L. Stierhoff, A. Henry. 2011. Estimation of spawning habits of market squid (*Doryteuthis opalescens*) from field surveys of eggs off Central and Southern California. Mar. Ecol. 2011:1–11

Photographic and Videographic Imagery Collection and Analysis Activities at the Coral Reef Ecosystem Division of the NOAA Pacific Islands Fisheries Science Center

Jacob Asher and John Rooney

Pacific Islands Fisheries Science Center
Honolulu, Hawaii
jacob.asher@noaa.gov

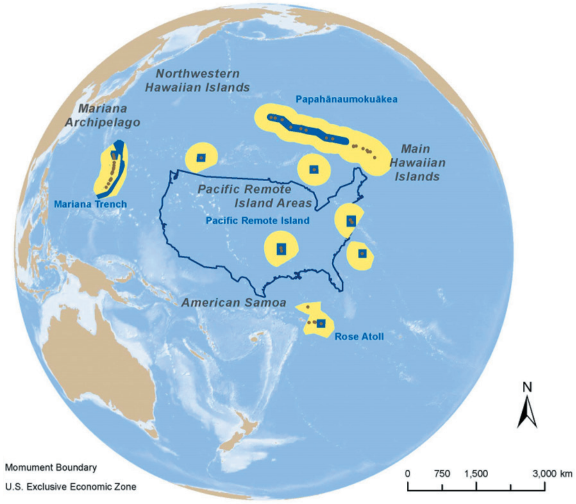


Figure 1. Area of coral reef ecosystem research conducted by CRED. Figure by Tomoko Acoba, Joint Institute for Marine and Atmosperic Research, University of Hawaii, Honolulu.

The NOAA Pacific Islands Fisheries Science Center (PIFSC) Coral Reef Ecosystem Division (CRED) has been historically engaged with a number of different research activities involving the collection and processing of still and videographic imagery (Fig. 1). These activities include but are not limited to towed-diver surveys, towed-optical-assessment-device (TOAD) surveys, autonomous-underwater-vehicle (AUV) surveys, and Rapid Ecological Assessment (REA) benthic and fish surveys during CRED Pacific Reef Assessment and Monitoring Program (Pacific RAMP) cruises around the Pacific Basin (Fig. 2).

Each respective methodology is described below:

- ◇ Towed-diver surveys provide assessments of relatively large areas of reef habitat (~ 2–3 km/survey), which incorporate benthic (Kenyon et al. 2006) and fish components (Richards et al. 2011).
- ◇ The TOAD is used in surveys to provide optical validation data to be correlated against bathymetry and acoustic backscatter imagery (Bare et al. 2010). More information can be found at http://www.soest.hawaii.edu/pibhmc/.
- ◇ The SeaBED AUV was designed by Hanumant Singh at the Woods Hole Oceanographic Institution (WHOI) and is jointly owned and operated by the Fishery Resource Analysis and Monitoring Division (FRAM) of the Northwest Fisheries Science Center (NWFSC) and CRED. Its primary mission is to collect fisheries-independent optical, bathymetric, and oceanographic data below typical diver depths (between 50 m to 2,000 m) to document fish–benthos relationships. More information can be found at http://www.soest.hawaii.edu/pibhmc/pibhmc_auv.htm.
- ◇ Photographic imagery data are collected in conjunction with stationary-point-count surveys at REA fish sites selected using a stratified random sampling design and with transect surveys at REA benthic sites (historically as part of photoquadrat surveys and more recently with line-point-intercept surveys). The REA surveys are used to obtain high benthic taxonomic resolution (compared with the broadscale towed-diver surveys, which target functional groups).

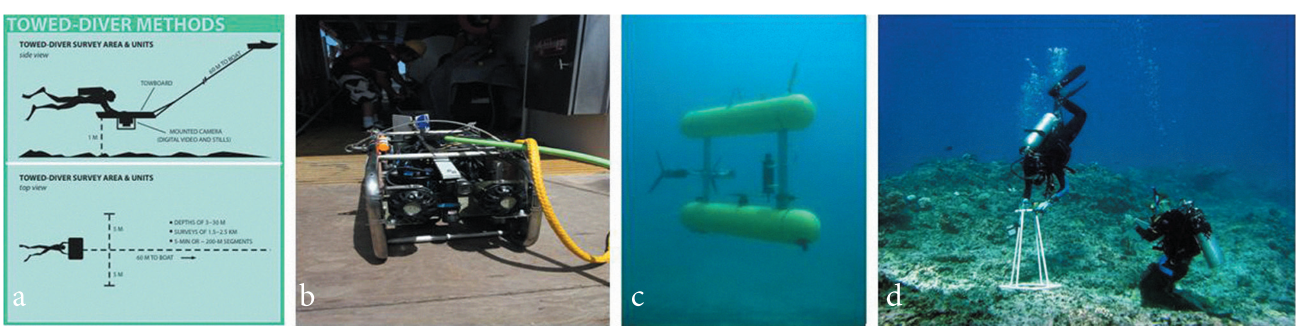


Figure 2. Benthic imagery data are or have been collected through a number of survey methods, including a) towed diver (Figure by Amanda Toperoff, Joint Institute for Marine and Atmospheric Research, University of Hawaii, Honolulu.), b) TOAD c) AUV, or d) REA photoquadrat.

CRED's approach to categorizing, analyzing and processing, and storing photographic and videographic data has continued to evolve after the advanced sampling technology working group (ASTWG) workshop in Seattle in 2010, with a subsequent Image Analysis Workshop on 30 November–3 December, 2010, involving members of CRED and Scripps Institution of Oceanography (SIO) of the University of California San Diego (UCSD). The main driver of this second workshop was to build consensus among the variety of groups within CRED about how best to analyze CRED's large and diverse imagery archive.

Following the workshop, CRED developed several internal goals regarding the direction of benthic image analysis. This mission plan included but was not limited to the following goals:

- ◇ Establish a repeatable, consistent, and statistically robust image analysis protocol to monitor benthic change at the functional group level, with expansion capability to include improved taxonomic resolution
- ◇ Determine the level of temporal change (interannual and decadal cycles) of benthic functional groups (e.g., hard coral and macroalgae cover) at island or atoll scales
- ◇ Conduct detailed comparisons of photographic vs. direct-diver-observation data, where applicable

Highlights of CRED image analysis work since the ASTWG meeting and subsequent image analysis workshop include the following activities:

- ◇ Design and implementation of an updated benthic classification hierarchy and standard operating procedures (Fig. 3). The common classification tiers, categories, and definitions were originally developed during the December 2010 CRED Image Analysis Workshop and then refined during subsequent months. (https://www.st.nmfs.noaa.gov/confluence/display/CRED/Classification+Tiers%2C+Categories%2C+and+Definitions).
- ◇ Development of an in-depth analyst training protocol and region-specific pretests for multiple image analysts covering multiple geographic regions.
- ◇ Standardization of data vocabularies.
- ◇ Broadscale analysis effort, involving a large pool of image analysts, to process American Samoa towed-diver photoarchival data for the 2002–2010 period. A separate effort to process line-point-intercept photo data from 2010 is also underway.

Figure 3. Example of the updated benthic classification tiers used by CRED.

The possibility of automating image analysis of coral reef benthic habitats remains a future challenge with a staggering number of obstacles. A combination of factors would be required to make it successful, including improvements in edge-detection software and increasing the ability of analysts to identify and discriminate between benthic functional groups using spectral signatures. In addition, these advances must account for highly variable factors such as sea and visibility states and habitats (e.g., spur and groove vs. basaltic reef). Until those challenges are addressed, CRED will continue to use a long-standing pool of scientists and analysts to manually classify benthic imagery.

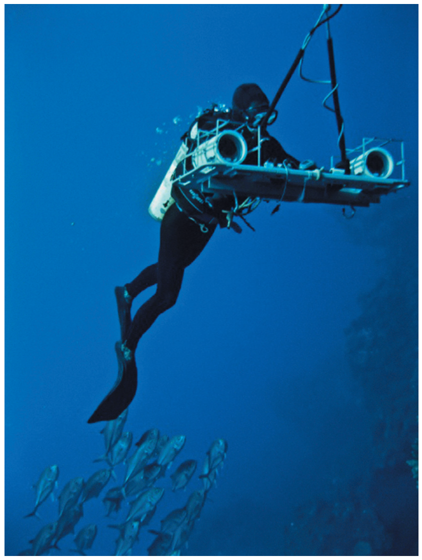


Figure 4. A modified towed-diver platform designed to capture stereoscopic images of coral reef fishes.

Toward automation of the analyses of imagery collected around coral reef habitats, several successful steps have been taken, including the following achievements:

- WHOI engineers have developed software algorithms to automate photomosaicing of SeaBED AUV imagery. While manual adjustments are still often required, this capability may generate additional automation possibilities for AUV imagery processing, mapping, and analysis.
- Continuation of collaborative efforts with Oscar Beijbom of UCSD to develop automated reef analyses using computer vision. A combination of texture analyses, color analyses, and testing of a computer-driven identification and classification engine on a suite of Caribbean coral reef images resulted in ~ 70% correct classifications on a functional group level, with a run time of ~ 1,000 images analyzed/24 hours (O. Beijbom, PIFSC, unpublished data). More information can be found at http://vision.ucsd.edu/project/computer-vision-methods-coral-reef-assessment and http://cvce.ucsd.edu/index.php.

Finally, in a separate, stand-alone effort, CRED in 2011 initiated the incorporation of SeaGIS stereoscopic camera systems (Fig. 4) (http://www.seagis.com.au/hardware.html) as an experimental effort to obtain accurate size-class data for fish species encountered during towed-diver surveys. Data collected during the Mariana Archipelago Reef Assessment and Monitoring Program (MARAMP) cruise in the spring of 2011 will be analyzed using PhotoMeasure software and methods similar to those used to analyze video from baited remote underwater video stations (BRUVS) and stereoscopic diver-swim transects (Harvey et al. 2003).

Citations

Bare, A.Y., K.L. Grimshaw, J.J. Rooney, M.G. Sabater, D. Fenver, and B. Carroll. 2010. Mesophotic communities of the insular shelf at Tutuila, American Samoa. Coral Reefs 29(2):369-377.

Harvey, E., M. Cappo, M.R. Shortic, J. Robson, J. Buchanan, and P. Speare. 2003. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (*Thunnus maccoyii*) with a stereo-video camera system. Fish. Res. 63(3):315-326.

Kenyon, J., R. Brainard, R. Hoeke, F. Parrish, C. Wilkinson. 2006. Towed-diver surveys, a method for mesoscale spatial assessment of benthic reef habitat: a case study at Midway Atoll in the Hawaiian Archipelago. Coast. Manage. 34(3):339-349.

Richards, B.L., I.D. Williams, M.O. Nadon, and B.J. Zqliczynski. 2011. A towed-diver survey method for mesoscale fishery-independent assessment of large-bodied reef fishes. Bull. Mari. Sci. 87(1):55-74.

Part 2.

Presentation Summaries from Invited Experts in the Field of Automated Image Processing and Underwater Video Collection

New Challenges to Deep-water Stock and Biodiversity Assessments: the Chronobiological Scenario and Automated Video-imaging Solutions

Jacopo Aguzzi and Juan Batista Company

Instituto de Ciencias del Mar (ICM-CSIC), Barcelona (Spain)
jaguzzi@cmima.csic.es

Corrado Costa and Paolo Menesatti

AgritechLab-Agricultural Engineering Research Unit of the Agriculture Research Council (CRA-ING), Monterotondo Scalo-Rome (Italy)

Sampling methodologies and their applicability represent a major bottleneck in our understanding of the biology of species inhabiting the deep-sea. Trawling is still one of the most effective and economically feasible methods of sampling. Trawling studies are broadly conducted over large seabed areas to analyze the distribution and demography of populations (i.e., stock assessment), as well as for overall biodiversity evaluation. One of the primary limitations in improving the reliability of data proceeding from trawl surveys is their overall variability. Sampling at random times produces unpredictable differences in the species composition of catches related to the rhythmic behavior of individuals forming local targeted populations.

Broad diel (i.e., 24-hour based) variation in demersal community composition occurs as the rhythmic presence and absence of populations from haul samples. Some categories of rhythmic displacement can be categorized as follows: benthopelagic species are those generally located on or just above the seabed; nektobenthic species migrate within the benthic boundary layer, encompassing continental shelves and slopes; and finally, endobenthic burrowing and burying species hide in the substrate during periods of behavioral passivity (Fig. 1).

Technological limitations in direct observation capabilities are at the root of the scarce modelling available on temporal biases in stock and biodiversity assessment. Improvement in this field requires a new remote, continuous, and especially long-lasting observational technology to monitor community changes in relation to contextual habitat metadata. Cabled multiparametric observatories

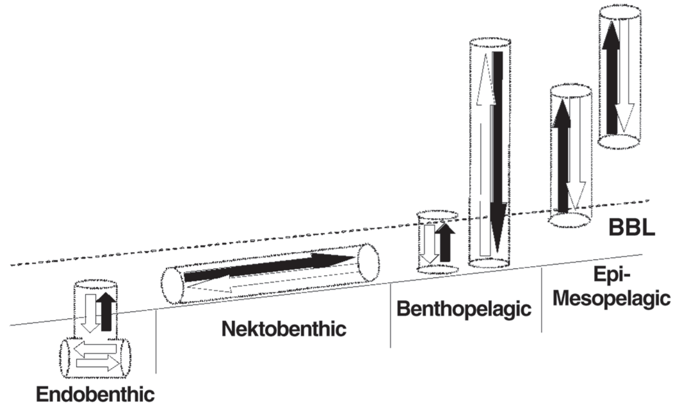


Figure 1. Schematic representation of activity rhythms in relation to the seabed, the water column, and the benthic boundary layer (BBL). Movements within corridors (the dashed cylinders) are indicated by thick black (night) and white (day) arrows.

represent a substantial innovation in this respect. One of the difficulties marine biology faces is the lack of sensors directly measuring biological activity. In contrast, geological and oceanographic sensors are more abundant and able to directly measure the properties and processes of interest. Cabled multiparametric observatories often house video cameras and these could be used as an efficient sensor at population and species level, provided that video-image analysis is sufficiently developed to automatically classify tracked individuals within different species as categories (classification) and to count individuals over time, as a proxy of behavioral rhythms.

Implementing analyses for tracking and classification of animals is crucial in order to extract relevant biological data using automated video-image analysis. Tracking is the process of identifying the same animal within a set of temporally consecutive frames. Classification is the grouping of each animal within a pre-established category, usually a species. While tracking is of critical value for the characterization of behavioral rhythms (i.e., the counting of individuals over time), classification enables the characterization of communities at local scale.

Re-counting of the same individuals is a problem in all studies that aim to estimate local population sizes by video imaging. One way of avoiding this is by tracking individuals and subsequently eliminating all initially counted individuals within the same frame set. Computing of trajectories can be implemented by using Kalman filters.

The shape of a given animal recorded by a static camera can be automatically classified and assigned to a species by Fourier methods. These methods allow the recognition of an organism

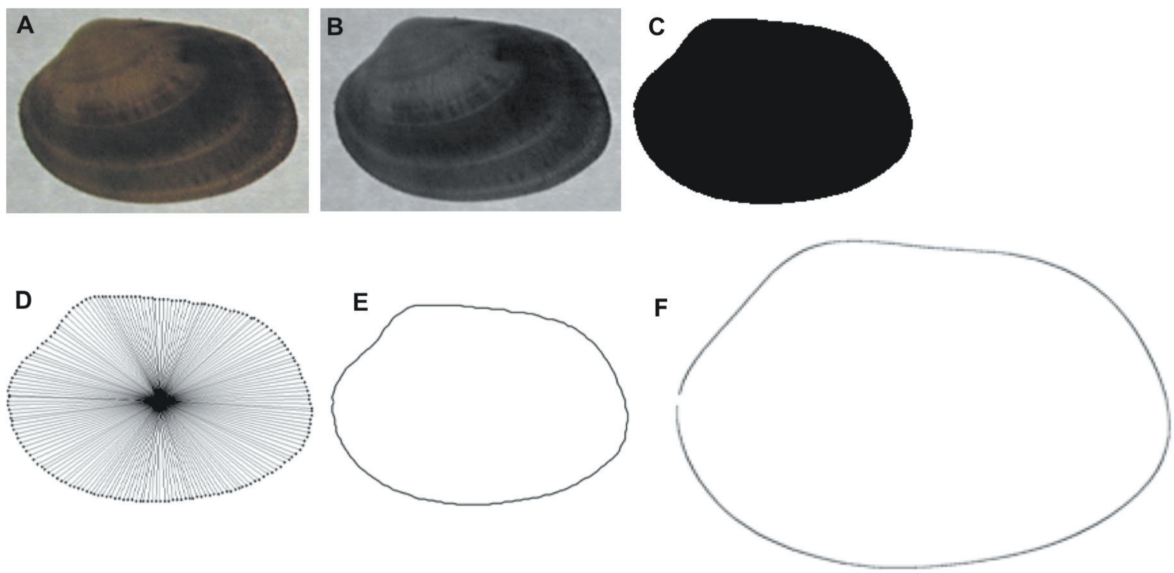


Figure 2. Eliptic Fourier analysis (EFA) carried out for the automatic classification of clams' shells within two species, Tapes decussatus and T. philippinarum. Digital images (A) are acquired in color and subsequently transformed in grey-scale (B) prior image binarization (C). Points are equidistantly placed (D) and joined as an outline (E). Supervised or non-supervised multivariate analyses can be carried out in order to screen the sample for any clustering of shapes. A mean shape outline (F) can be then obtained for each clustering.

profile through the fitting of harmonic functions. In the elliptic Fourier analysis (EFA), the animal's shape can be studied by the automatic positioning of angular or equidistant points along the profile; the transformation of the contour into an incremental harmonic function (*x,y* coordinates are computed for each point); and finally, the fitting of that function with an increasing number of ellipses in order to approximate its variation with the highest precision (Fig. 2). At the end of this fitting process, each biological sample is represented by a set of ellipses, each of them with their four coefficients. As a result, a matrix of all individuals and their respective ellipses' coefficients can be obtained. This matrix is the input required for two types of multivariate statistical analysis: non-supervised, (e.g., principal component analysis-PCA) when we screen the sample for any potential clustering of shapes without prior ecological hypotheses, and supervised when the clustering is attempted according to ecological hypotheses (through the addition to the matrix of indexed complex ecological variables).

Fourier descriptors (FD) are also employed for the automated recognition of tracked animals. They can be utilized to describe the shape outline in terms of its frequency, by the fitting of a set of circular harmonic functions each with its own coefficients (the FDs) onto the outline. In combination with shape analysis, the RGB content (i.e., the average color content coordinates) of organisms can be added to increase recognition efficiency.

Morphological recognition is carried out in a semi-automated fashion (Fig. 3). A library of manually supervised and classified images is required for each target species. Animal images from different angles can be saved in a binary format. The classification of every newly tracked animal is carried out on the basis of shape and color content descriptors by resemblance to an average model extracted from the training set of images. Recognition can be efficiently carried out with partial least squares discriminant analysis PLSDA or Supervised Standard K-Nearest Neighbor (KNN) analyses.

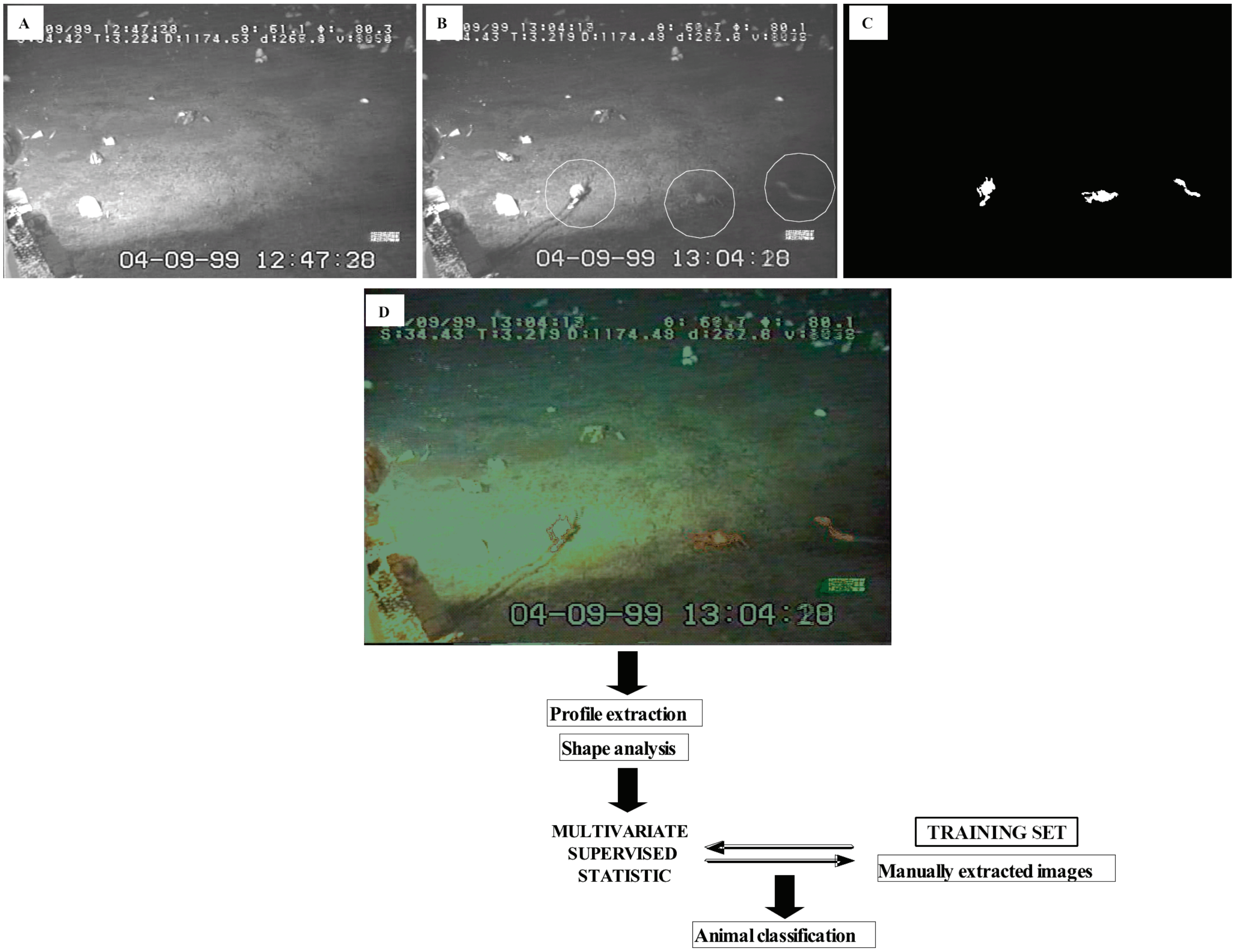


Figure 3. The automated tracking and classification of moving animals (an eelpout, a red crab and a snail are highlighted by the circles) in digital videos of Sagami deep-sea station (1,100 m depth, central Sea of Japan). The identification occurs at frame subtraction (A-B) and after image binarization and area filtering (C).

An Automated Visual Event Detection and Classification System

D. E. Cline and D. R. Edgington

Monterey Bay Aquarium Research Institute
Moss Landing, California
dcline@mbari.org, duane@mbari.org

Introduction

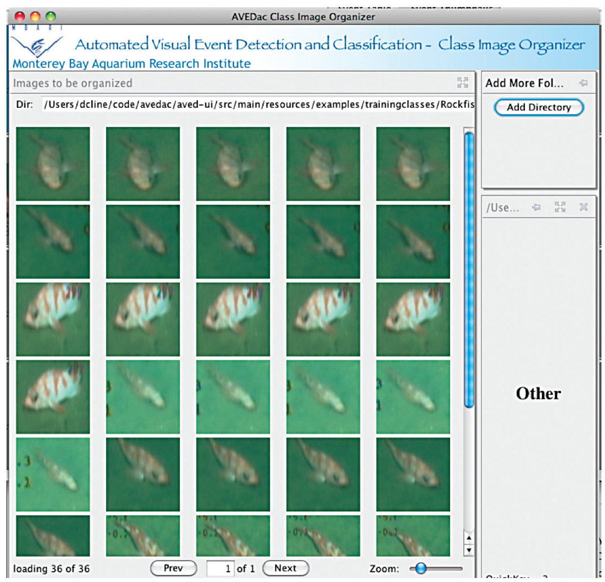
Ocean observatories and underwater video surveys have the potential to unlock important discoveries, yet the burden of video management and analysis often requires reducing the amount of video recorded and later analyzed. To help address this problem, the Automated Visual Event Detection and Classification (AVEDac) software has been under development at the Monterey Bay Aquarium Research Institute (MBARI) since 2002 to help analyze video and still images.

About MBARI

Monterey Bay Aquarium Research Institute is a private, non-profit research company funded by the David and Lucile Packard Foundation. Our staff includes approximately 220 scientists, engineers, and operations and administrative personnel. Our mission is to achieve 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.

Project Motivation

The motivation for this work extends across several diverse applications including: autonomous underwater video surveys in marine protected areas (MPAs); deep-water observatories such as the MBARI Monterey Accelerated Research System (MARS) observatory test bed where we can record video 24-hours a day on shore; time-lapse cameras placed on the seafloor to record high-quality still images from the Abyssal Time-Series Images of Station (Cline et al. 2009). More recently, work is underway to help University of California Davis (UCD) and the U.S. Army Corps of Engineers monitor the passage of lamprey eel and salmon through low-head dams. The studies in each of these applications are primarily for abundance estimation and distribution or behavioral studies. AVEDac currently does not estimate sizes of targets, although it is a desired feature for MPA studies.



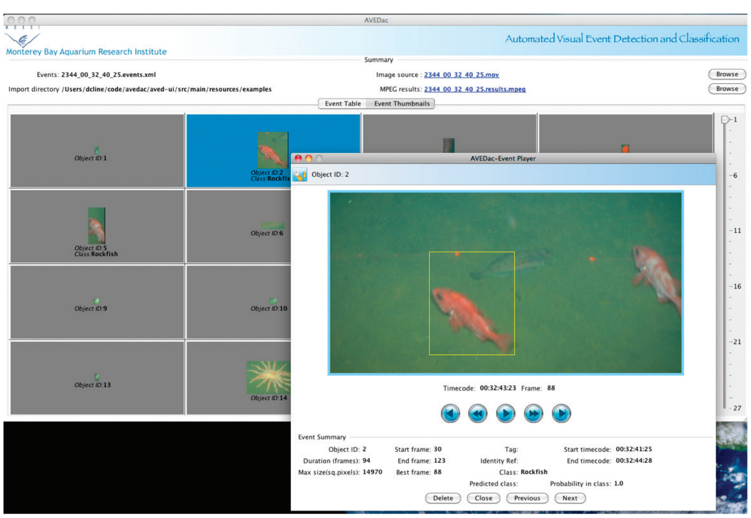


Figure 1. Example views of AVEDac user interface

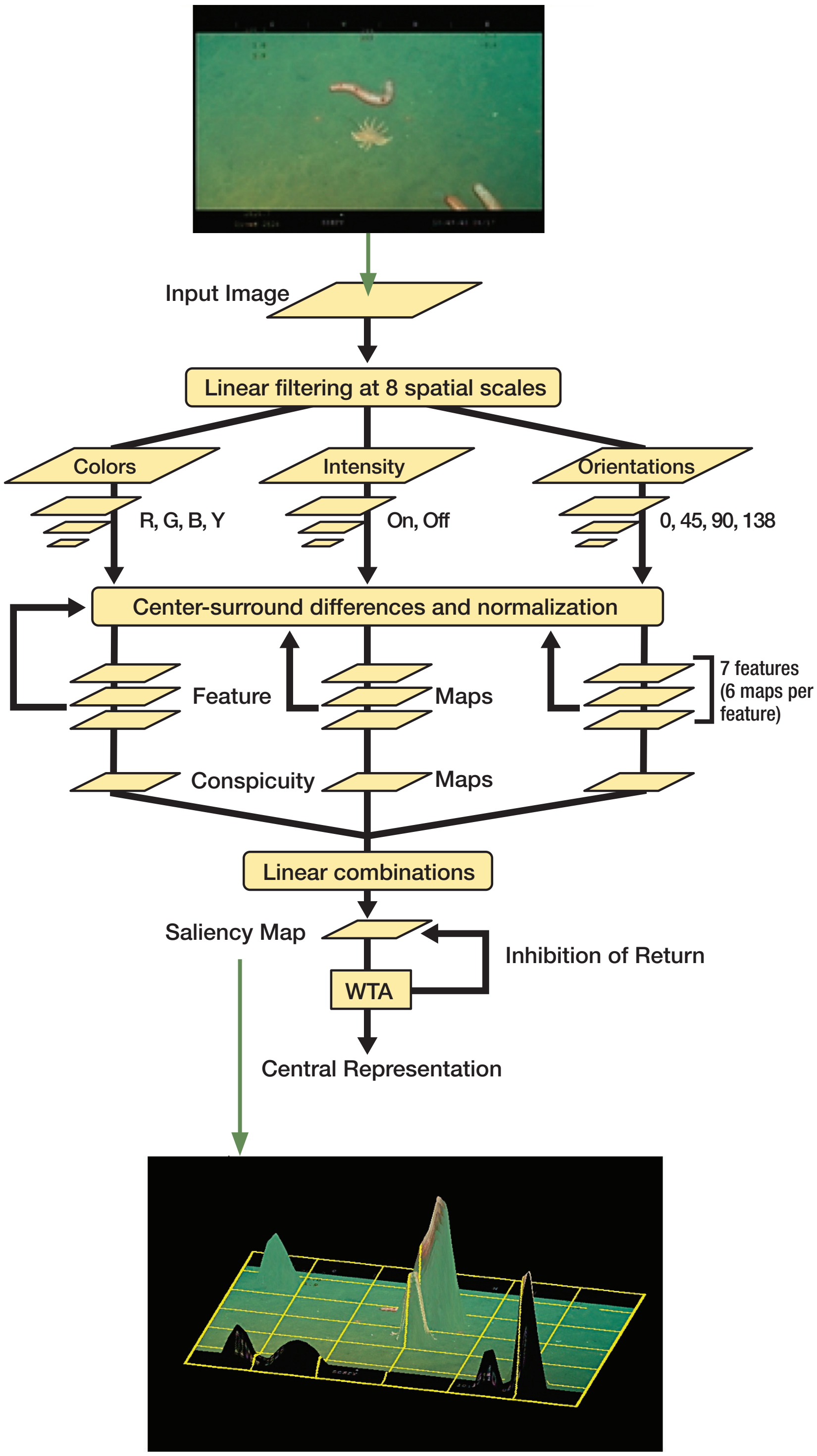


Figure 2. Saliency map from the iLab toolkit warped onto a 3-D map. Peaks in the map show points of high visual attention where the Rathbunaster and Leukothele are in the center image.

Software Overview

The AVEDac software analyzes each image in search of interesting events using a neuromorphic software model based on the human vision system (Itti et al. 1998). Potentially interesting visual events spanning multiple images or frames based on low-level properties of salient objects are tracked (Walther 2003, 2004). 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. To generate these training classes representing each taxonomic category required for the AVEDac classifier, a collection of video and still-images populated with the items of interest is processed through the AVEDac software. The user then sorts all potentially interesting events into representative classes through the aid of a graphical interface (Fig. 1).

Processing imagery with AVEDac is both data and computationally intensive and 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 may not necessarily be executed on the same computer (Fig. 2). To execute and manage applications with particularly large data sets or complex workflows, a specialized workload management system for computer and data intensive jobs called Condor, developed by the University of Wisconsin–Madison, is used.

Results and Future Work

The system has shown promising results when applied to imagery from video surveys conducted by remotely operated vehicles (Walther 2003, 2004). We continue to make improvements to the classification algorithms, as this has proven to be the most difficult task.

Work is underway to explore alternative classification algorithms to help UCD and the U.S. Army Corps of Engineers identify and distinguish lamprey eels from salmon during passage through low-head dams (weirs). Work is also ongoing to add new features to the graphical interface in preparation for use with MPA studies. We are also currently using AVEDac to analyze video collected from the Eye-in-the-Sea™ camera located in Monterey Bay (Widder et al. 2005).

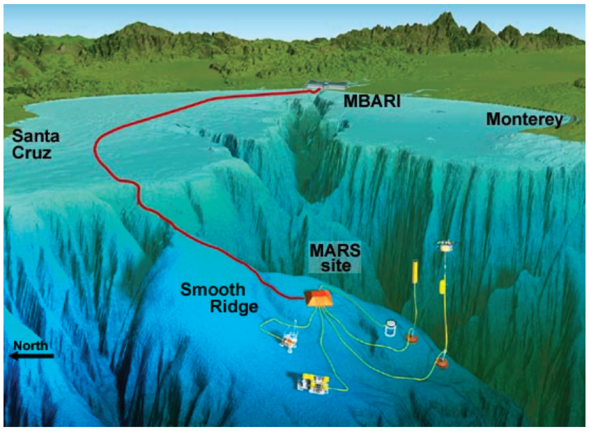


Figure 3. The MARS cabled-to-shore-observatory site 891 meters (2,923 feet) below the surface of Monterey Bay. Power and 100 megabit communications is provided to a "science node". The Eye-in-the-Sea™ connected to the MARS "science node" for six months, recording video 24 hours a day in 2008-2009.

Acknowledgments

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 AVEDac to help them analyze the EITS video.

Citations

Cline, D.E., D.R. Edgington, K.L Smith, M.F. Vardaro, and L. Kuhnz. 2009. An Automated Event Detection and Classification System for Abyssal Time-Series Images of Station M, NE Pacific. MTS/IEEE Oceans 2009 Conference Proceedings, October, 2009. Biloxi, Mississippi 2009.

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): 1254-1259.

Walther, D, D.R. Edgington, and C. Koch. 2004. Detection and tracking of Objects in Underwater Video, 2004, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Washington, D.C. Available from www.mbari.org.

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. Available from www.mbari.org.

Widder, E.A., B.H.Robison, K.R.Reisenbichler, and S.H.D.Haddock. 2005. Using red light for *in situ* observations of deep-sea fishes, Deep-Sea Res. I 52:2077-2085.

Relevant Web Sites

Monterey Bay Aquarium Research Institute: www.mbari.org

Automated Visual Event Detection (AVED) Project:
www.mbari.org/aved avedac.googlecode.com

Video Annotation and Reference System (VARS) Project: www.mbari.org/vars

iLab at University of Southern California: ilab.usc.edu

Condor Project: www.cs.wisc.edu/condor

Automatic and Manual Plankton Identification: a Comparison

Phil Culverhouse

Centre for Robotics and Intelligent Systems,
University of Plymouth, Plymouth, UK. PL4 8AA

Humans cannot efficiently process the volume of images generated during large-scale plankton surveys for proper identification of field samples and also cannot operate with *in situ* image collection. There is a need to automate. The one-dimensional case illustrated in Figure 1 shows how characteristics are extracted from an image (in this case, by multi-channel laser absorption spectra) and used to create discriminating features of cells. These features can then be harnessed to analyze water samples (Fig. 2). Extending this to two dimensions through the use of digital cameras, flatbed scanners can be employed to evaluate digitized images. There are several commercial products (e.g., FlowCAM, ZooSCAN, software toolsets for microscopy) and several free software toolsets (e.g., Zoo/PhytoImage, ZooProcess, Weka and Tanagra for statistical analysis) available to assist in these analyses.

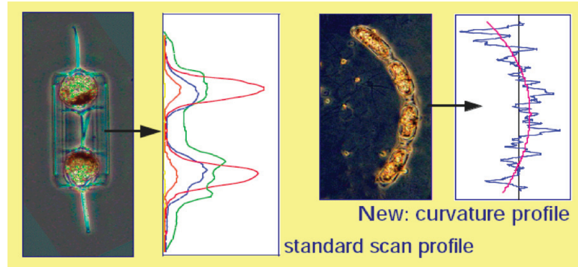
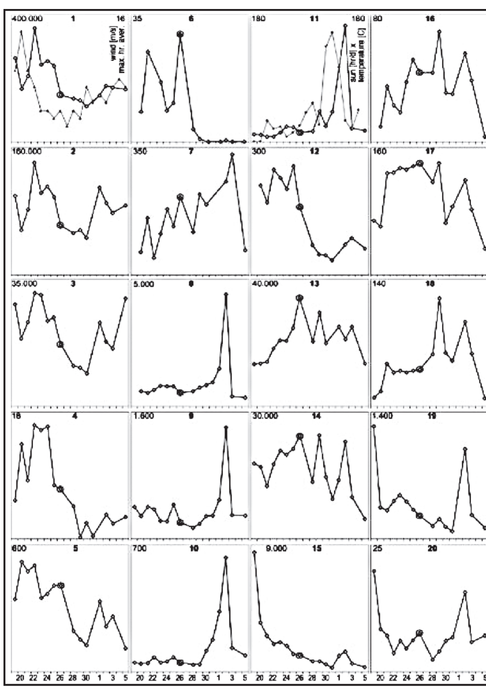


Figure 1. Cytosense image in flow examples of two phytoplankton and their accompanying laser fluorescence traces (source: Dubelaar et al. 2004).

The least expensive option is to use a flatbed scanner (Fig. 3) which can examine objects as small as 0.5 mm in length. These images can then be analyzed with appropriate software (in this case, Zoo/PhytoImage) to extract vignettes (i.e., regions of interest or ROIs), obtain measurements, and then train a classifier on the selected training set specimens. Figure 4 provides an example of a flatbed scan of a water sample (cf. Fig. 3). The sample has been stained with Bengal rose to highlight specimens and reduce the contribution of detritus to the analysis. A subsample of this scan is shown magnified in Figures 5 and 6 show a commercial variant of the flatbed scanner, the Zooscan. Zooscan has been designed for faster sample processing and reduced calibration effort



1 Non fluorescent particles of 1–3 μm (sediment-debris).
2 Non fluorescent particles of 3–10 μm (sediment-debris).
3 Non fluorescent particles of 10–50 μm (sediment-debris).
4 80–140 μm organisms with small local (12–18 μm) orange fluorescent section.
5 30–100 μm particles with localized weak to moderate chlorophyll fluorescence.
6 Filaments > 50 μm, cyano type.
7 Picoplankton cyano type.
8 *ca.* 8 μm cells, moderate accessory pigment.
9 20–35 μm Cryptomonas type cell.
10 50–70 μm Cryptomonas type cell.
11 Synura colonies (50–150 μm diam.)
12 *ca.* 7 μm cells, with accessory pigment fluorescence.
13 Picoplankton, green.
14 Green cells, 3–15 μm.
15 *ca.* 8 μm cells, high chlorophyll fluorescence.
16 Closterium-like, symmetrical cells, 50–80 μm.
17 Closterium-like, symmetrical cells, 70–100 μm, very thin.
18 Spindle shaped, 70–90 μm cell with double chloroplast (30–40 μm).
19 15–40 μm cells, high chlorophyll fluorescence.
20 Dinobryon colonies, 100–200 μm length.

(No taxonomic analyses were available for further identification of these groups.)

Figure 2. Distributions of 20 groups of particles in a Cytosense-analyzed water sample (a), with group labels assigned to sample groups (b). Vertical scale is individuals per milliliter. (Dubelaar et al. 2004)

Figure 4. CEH.01-07-02.p7+B2. Flatbed scan of mixed zooplankton (Di Mauro et al. 2009)

Figure 5. CEH.01-07-02.p7+B2. Magnified scan of a subsample of the mixed zooplankton shown in Figure 5. A showing a small sample of copepods drawn from Figure 5

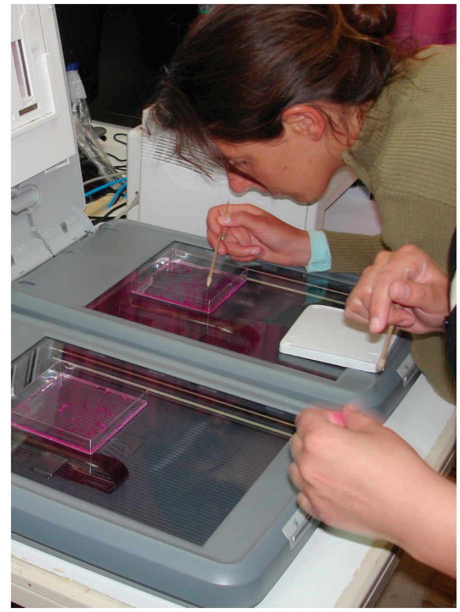


Figure 3. Using a flatbed scanner to image Bengal rose-stained zooplankton (source: Culverhouse, P.F.C., Austral Summer Institute, USC, Chile 2007).

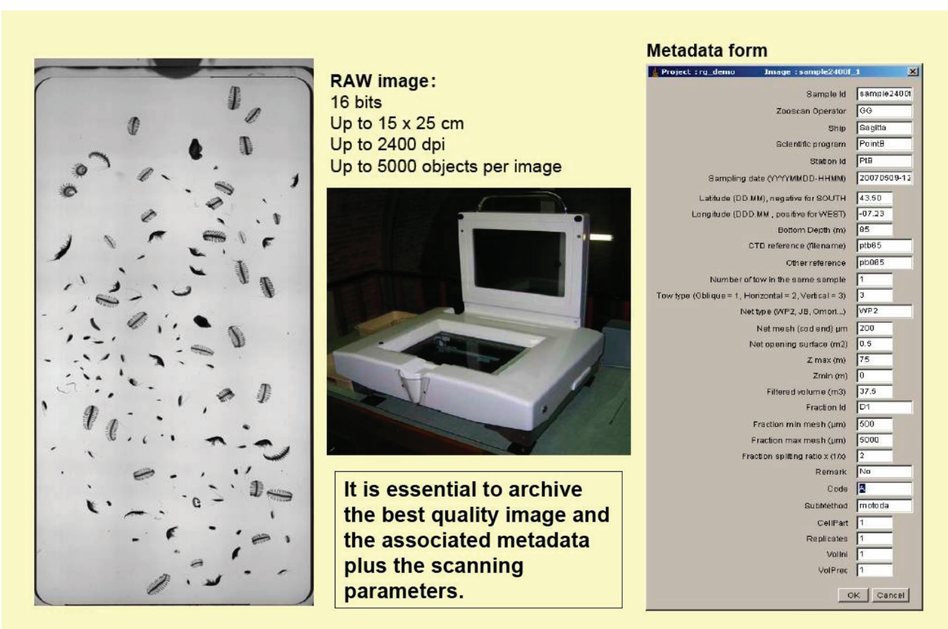


Figure 6. The Zooscan instrument with associated metadata form (Gorsky et al. 2010).

Table 1. A summary of study results (source: Culverhouse et al. 2007).

Categorization task	Self consistency within panel individuals	Panel consistency across individuals performance	Machine	Specimen Data size
Ceratium longipes and *C. arcticum*	expert: 94-99% 'book' expert 67-83% -8 experts	95% to 43%	99%	60 epidiascope tracings
5 spp. Tintinnidae	N/A	91 % - 6 experts	87%	100 printed photomicrographs
3 spp. Fish larvae	N/A	N/A	70%	1562 computer analysed photomicrographs
5-9 spp. Dinoflagellates	N/A	91-95% - 6 experts	67%	30 computer displayed photomicrographs
23 spp. Dinoflagellates	N/A	83-86% - 6 experts	83%	310 computer displayed photomicrographs
10 taxa Zooplankton	46 to 86% - 6 experts	N/A	80%	>400 specimens in suspended in IMS with detritus

ID	major axis	Minor axis	Area	Centroid X	Centroid Y
50	49.847467	22.949748	773	820.274256	98.421734
51	38.66348	12.328536	359	405.041783	94.415042
52	36.453996	11.220497	303	212.277228	146.082508
55	54.676793	20.648836	816	616.53799	293.973039
58	42.046903	15.77498	501	1313.52495	309.854291
61	46.289936	14.698126	438	1215.51598	349.0570
62	34.615556	12.756711	320	62.78125	419.856
63	28.20686	7.952657	163	1108.36196	440.0184
65	39.420003	14.781337	446	250.210762	504.948
66	85.896223	21.420149	625	226.5472	557.14
67	50.898697	21.809834	746	1167.30027	541.2828
70	87.031542	23.041159	821	120.952497	608.6699
75	34.293377	14.11345	358	1197.28492	630.3212

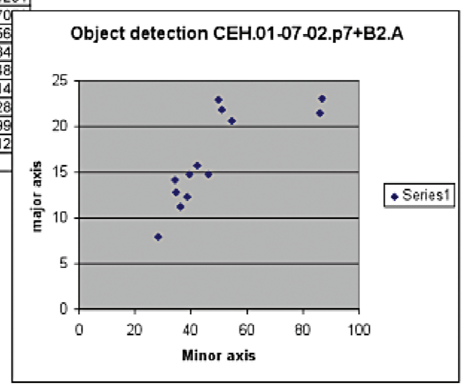


Figure 7. Morphological data extracted automatically from CEH.01-07-02.p7+B2.A (from Fig. 4).

as compared to a normal flatbed scanner. In conjunction with these analyses, metadata containing accurate and relevant cruise-specific information is essential for subsequent ecological study and application of the image analysis results.

We can try and extract discriminating measurements from each specimen's image, such as the fit to an ellipse (major axis, minor axis) and the specimen's pixel area. Figure 7 shows the tabulations for the specimens in Figure 4. After manual inspection of the data, we can suggest there are three or four clusters, based upon these three measurements. This is the basis of machine categorisation - but instead of three parameters, often 60 or more parameters are extracted from each image vignette, resulting in a 60-dimensional clustering problem. Training a classifier algorithm such as a random forest with such data is straightforward. The limiting step is the need to first label by hand each specimen in the training set. The value of this type of analysis is that net-caught plankton samples can be processed quickly, allowing for the analysis of over 10 net samples per day for up to 30 functional groups.

Let us now look at how expert analysts perform at this type of inspection. Review Figure 8 yourself and decide how many of the organisms in the key on the right hand side of the figure appear in the image. This type of manual review is not only difficult, it can be monotonous and boring with fatigue setting in quickly, perhaps within 15 minutes of starting the review. Human performance in identifying and sorting organisms is affected by several psychological factors including:

- Human short-term memory limit of five to nine items
- Fatigue and boredom: severe loss of categorisation performance (> 50% error)
- Recency effects where a new classification is biased toward those in the set of most recently labels
- Positivity bias, where specimen identification is biased by one's expectations of the species likely to be present in the sample

Context and other prior cues as to category speed up recognition significantly. These human factors give rise to biases in tally counting, as can be seen from the consistency within individuals and between experts (Table 1).

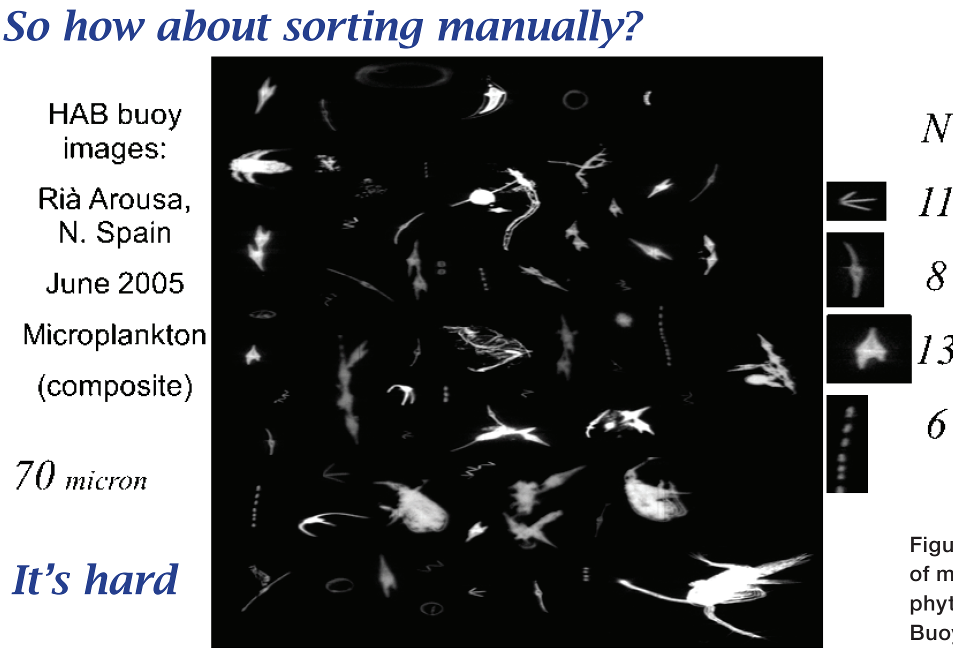


Figure 8. A composite image of microzooplankton and phytoplankton taken with HAB Buoy (Culverhouse et al. 2007).

Conclusions

People are not perfect identification machines, although experts can be highly self-consistent. In contrast, novices often lack even self-consistency, and forming a consensus of opinion between people can be difficult (e.g., two people sorting the same sample will very often come up with different tally counts for the selected categories).

Automated identification of plankton is feasible when applied in constrained circumstances (i.e., it is not a detailed taxonomic analysis of specimen images). It can be effective for generic discriminations with results comparing favourably with human performance, given that human error rate increases with sample size and machines offer consistent results regardless of sample size.

Future challenges in the field of automated image analysis include:

- Scaling up to > 100 and > 1,000 categories
- Handling three-dimensional specimens for *in situ* work
- Validating training data using scarce human expert resources
- Funding research in this area.
- Cross-disciplinary interactions with ecosystems modellers
- Difficult for referees to review, so there is a need to promote wider awareness of problems and solutions
- Greater support from the wider community.

Citations

Culverhouse, P.F. 2007. Natural object categorisation: man versus machine, *In* Automated Taxon Identification in Systematics: Theory, Approaches, and Applications. MacLeod, N. (ed.). Systematics Association. CRC Press. Boca Raton, LA, US. 24–45

Culverhouse, P.F., R. Williams, B. Reguera, V. Herry, and S. González-Gil. 2003. Do experts make mistakes? Mar. Ecol. Prog. Ser. 247: 17-25.

Dubelaar, G.B.J., P.J.F. Geerders, and R.R. Jonker. 2004. High frequency monitoring reveals phytoplankton dynamics. J. Env. Monitor. 6:946-952.

Gorsky, G., M.D. Ohman, M. Picheral, S. Gasparini, L. Stemmann, J.-B. Romagnan, A. Cawood, S. Pesant, C. Garcia-Comas, and F. Prejger, 2010. Digital zooplankton image analysis using the ZooScan integrated system. J. Plankton Res. 32:285–303.

Fish Species Recognition for Migration Monitoring and Future Applications

Dah-Jye Lee

Department of Electrical and Computer Engineering
Brigham Young University, Provo, Utah
djlee@byu.edu

The quantification of abundance, distribution, and movement of fish is critical to viable fishery management and ecological and environmental studies of fish communities. This includes assessing the use of fish bypass facilities, estimating entrainment at canals and dams, studying behavioral responses of fish to environmental and weather changes, and collecting counts and sizes of fish at multiple points along migration routes. Hydroelectric facilities, diversion dams, and reservoirs on most western rivers block fish migration routes. Passageways are often installed to allow fish movement through these structures. Monitoring fish movement at these passageways can provide important ecological and management information about fish migration patterns. Other biological research projects not performed at these passageways often require capturing and releasing of fish.

On many western rivers, there can be multiple diversion dams built along each migration route. Monitoring fish migration at these diversion dams or at narrow passages along the rivers can provide extensive information about the numbers, seasonal timing, periodicity of movement, and passage survival of fish. From these, projections of the strength of runs can be made, long-term trends in populations can be compared, and even the relative success of various mitigation measures can be appraised. Condition factors can be used to determine the influence of abiotic factors on fish growth. The timing of migrations can be related to regional sea surface temperatures, and changes in relative species composition can be investigated relative to upstream habitat changes. Even mortality rates can be estimated as fish move progressively upstream. Without monitoring, this type of management is not possible. However, monitoring requires labor, either on-site observers to identify and count species passing through the station, or some mechanism for automated recording of the fish.

Figure 1. Prosser Dam on the Yakima River in Prosser, Washington (top), and its fish ladder (bottom).

The Columbia River basin covers 259,000 square miles and includes the Snake, Deschutes, Okanagan, Wenatchee, Spokane, and Yakima rivers. It is also home to more than 400 dams used for hydroelectric power, irrigation, transportation, flood control, and recreation. At present, fish are counted and monitored at 50 percent or more of the U.S. Bureau of Reclamation (USBR) and U.S. Army Corps of Engineers facilities. Almost all monitoring is done by human observers, requiring significant time and financial costs. It is also subject to human error and labor constraints. To reduce the cost of manual fish tracking, videotaping has been implemented. However, the resulting large volume of video recordings must be reviewed manually in order to collect this critical information. Typically, the image quality obtained from video recordings is poor, making it difficult for biologists to identify species further complicating the data acquisition process. Maintenance of video recordings has also been a challenging task as analog video recording technology becomes obsolete. According to biologists at the USBR, state governments, universities, and Native American tribes, an automated fish recognition and monitoring system is urgently needed. The USBR and U.S. Department of Agriculture (USDA) supported this research effort.

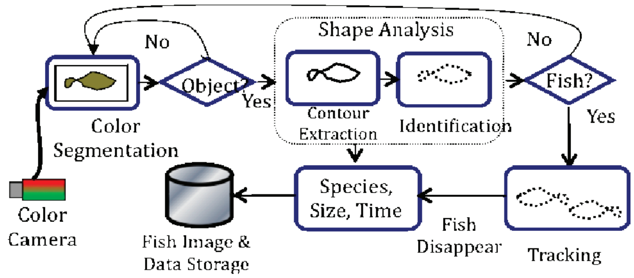


Figure 2. FishID system diagram.

Figure 2 illustrates the overall design of one such automated project: the FishID system. It outlines the need for each of the five processing steps. The first and most important step is image acquisition. Object detection algorithms detect the presence of an object. Object contour extraction and identification help to identify whether the object is indeed a fish and perform species recognition. Shape-based recognition can be performed to obtain size and species information. Tracking of fish movement determines the location of the fish and identifies the best image frame for recognition. In addition to fish species, timestamp and fish image (if necessary) can be stored for further analysis by the user.

Pictures in Figure 3 show the fish viewing window at the test site on the right and the functional graphical user interface on the left. The viewing window is built in a vault approximately 30 feet below the top surface of the dam. The software was developed with a friendly graphical user interface for testing. Calibration and parameter adjustments were done with simple mouse clicks. The viewing window was 4×4 and the distance allowed for mounting the camera was also 4 feet.

Figure 3. Image of fish ladder window and screen grab of FishID system interface.

We convert the chosen image into a binary image using the data calculated in the object detection stage. From this we determine the fish outline, or contour of the fish. This raw data are simply a list of points with the left-most fish pixel being the starting point and the contour traced in a clockwise direction. We reduce these points to a useful number (between 30 and 50 points appeared to give the best results with our data) by taking points equally spaced along the contour of the fish. Using equally spaced points implicitly makes our shape representation scale-invariant. The reduced-point contour of the fish forms a closed polygon. We calculate the turn angles of this polygon, again starting from the left-most point and working clockwise. This list of turn angles contains no information concerning the position of the fish and thus makes our shape description translation-invariant. Also, as long as the left-most point on the fish is the mouth of the fish (which it generally is in this case) then the turn angles are rotation-invariant as well. Figure 4 illustrates the whole process of generating a fish model.

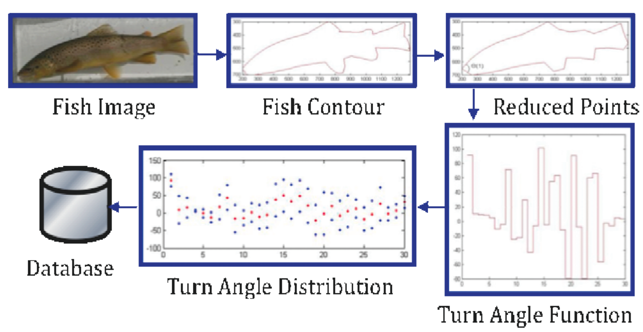


Figure 4. Turn Angle Distribution Analysis (TADA).

At Prosser Dam and most other facilities there are only a small number of species the biologists are concerned about monitoring. Furthermore, species that confuse the recognition algorithm may not present at the same facility, as most of the species monitored are larger than most of the other species in the river. The accuracy of our TADA recognition algorithm tested on a few selected species is shown in Table 1. The accuracy decreased when more species were added for testing. Testing using only two species lead to accuracies as high as 97%. It maintained 97.5% accuracy when four species were used. Accuracy was negatively affected by the addition of speckled dace (*Rhinichthys osculus*) and brown trout (*Salmo trutta*), both similar in shape to salmon. Our classification algorithm — coupled with a size filter to ignore the smaller species — would perform very well in this case. Adding color features could also further improve the accuracy for certain species.

Table 1. Fish species recognition results using TADA.

Species	2 Species	3 Species	4 Species	5 Species	6 Species
Speckled dace	97%	93%	93%	77%	77%
Whitefish	97%	97%	97%	93%	93%
Cottid	-	100%	100%	97%	97%
Utah sucker	-	-	100%	80%	77%
Salmon	-	-	-	90%	83%
Brown trout	-	-	-	-	13%
Average %	97%	96.7%	97.5%	87.4%	73.3%

New Developments

One of the most important computer vision research areas is the detection and recognition of specific objects. The difficulty of detecting and labeling objects in images is due in part to issues such as lighting conditions, object orientation, distortions, and naturally varying parameters of the objects. Overcoming these obstacles to achieve robust object recognition would be beneficial for many scientific fields and applications such as automatic target identification for the military, surveillance for security applications, medical imaging, and abundance estimation or invasive species detection for biological and environmental studies.

Object recognition is a well-studied but extremely challenging field. Most current approaches rely on human experts to construct features for object recognition. We have developed a novel approach in the Robotic Vision Lab at BYU that addresses a few common drawbacks of current practices. Our approach aims to perform automatic feature construction for object detection called Evolution-COnstructed Features (ECO features). ECO features are automatically constructed by uniquely employing a standard genetic algorithm to identify series of transforms that are highly discriminative. Using ECO features provides several advantages over other object detection algorithms including:

1. Effective features can be discovered without the use of a human expert.
2. Non-intuitive features can be constructed that are not normally considered by human experts.
3. ECO features are not limited to certain image sources including data originating from complementary metal-oxide semiconductor (CMOS) sensors, synthetic aperture radar (SAR), infrared (IR), and potentially others such as magnetic resonance imaging (MRI), computed tomography (CT), X-ray, etc.
4. ECO features can be learned offline for any object type. In other systems the human expert creates features that are good for one class of objects but may do poorly on other objects types.
5. ECO features can be global or local feature types.

Potential Applications

As mentioned previously, many applications are in need of a robust object detection and recognition method. Our ECO features approach is able to address many critical issues in those areas. Potential applications of this novel object recognition includes invasive fish species detection and coral reef fish abundance evaluation for biological and environmental studies. There are also other potential applications in the military, homeland security, and medical applications.

Automated Techniques for Detection and Recognition of Fishes using Computer Vision Algorithms

J. Matai[1], R. Kastner[1], G. R. Cutter, Jr. [2], and D. A. Demer[2]

[1]University of California, San Diego
[2]Southwest Fisheries Science Center, NMFS, NOAA
jmatai@es.ucsd.edu

Introduction

Automated recognition and classification of fish and other organisms is beneficial to efforts of counting fish for population assessments, for describing associations between fish and habitats, or monitoring ecosystems. In this work, we summarize current efforts to automate the process of fish detection and recognition from a video or still camera source using computer vision algorithms. In order to recognize a fish from video source, there are two steps involved. First is the fish detection process, in which the fish is detected and separated from background. The detected fish image from previous stage is then passed to a recognition algorithm to identify the species of the fish. The latter is known as the recognition or identification stage.

Fish Detection Methodologies

The detection process consists of identifying fish locations in an image frame (i.e., its x,y pixel coordinates), fish extent (width, height), followed by a clear segmentation of fish from background. The outcome is an image that only contains fish targets, with the background masked out, and individual non-overlapping fish targets separately labeled. The Viola and Jones (VJ) object detection algorithm based on haar-like features(Viola and Jones 2004) was evaluated for identifying fish. First, a training image set was assembled consisting of positive (with fish) and negative images (without fish). Then this training set was used to identify test sets of images to determine the effectiveness of the method. The detection of two fish species, the Scythe butterfly fish (*Prognathodes falcifer*) and flag rockfish (*Sebastes rubrivinctus*) from images was tested using this approach. Images of butterfly fish in an aquarium collected by Benson et al. (2009) and rockfish images collected *in situ* by an ROV were provided by J. Butler, NOAA SWFSC Benthic Resources Group (Fig. 1).

Fish Identification Methodologies

The recognition of fish is the process of identifying fish targets to species based on similarity to images of representative specimens (testing sets of images of know species). Following is a brief description of PCA (principal component analysis) and SIFT (shift invariant feature transform) algorithms used for the recognition process.

- ◊ **PCA (Principal Component Analysis)** Turk and Pentland (1991) introduced an algorithm for face recognition based on PCA. It is the simplest and most widely used face recognition algorithm, and is quite effective. The PCA recognition algorithm has two stages. As in the fish identification stage the first step consisted of assembling the test sets, and in the second stage this test set was compared to unknown fish targets.
- ◊ **SIFT (Scale Invariant Feature Transform** Introduced by Lowe (2004), the scale invariant feature transform (SIFT) can be used for matching images or for object recognition. The main objective of SIFT is to find important key points in two images and match those points against other images. The main focus of SIFT is to find these points by dimensionality reduction. The SIFT approach is robust to variations in scale, rotation, and illumination in test set images. We used the VLFeat software tool for training the SIFT process and for validation of the results. For further information see http://www.vlfeat.org/.

Fish Detection Results

An example of the application of the VJ algorithm to identify fish targets is presented below. Table 1 summarizes results for six different test cases in detecting butterfly fish. The first three test cases use 1,000 positive images and 3,000 negative images as the training set and the second three test cases use 2,689 positive images and 3,000 negative images.

Table 1. VJ fish detection algorithm results. P/N indicates positive and negative image ratio in the training set. TS indicates test set size. Hits indicate the number of and percentage of correctly detected fishes. Missed indicates the number missed fishes. The “False” column indicates false positives.

#Test	P/N	TS	Hits	Missed	False
1	1000/3000	112 R	94 (83%)	18	23
2		112 L	68 (60%)	44	48
3		224 LR	162 (72%)	62	71
4	2689/3000	112 R	101 (90%)	11	16
5		112 L	91 (81%)	21	19
6		224 LR	192 (85%)	32	35

S. constellatus — *S. levis* — *S. miniatus* — *S. rubrivinctus (flag rockfish)* — *Scythe butterflyfish*

Figure 1. Training set for principal components analysis (PCA).

Table 2. Flag rockfish detection with Viola and Jones (2004) algorithm.

#Test	P/N	TS	Hits	Missed	False	#Stages	#Weak Classifiers
1	3100/3000	1272 LR	245 (19%)	916	138	3	3
2	3800/3000	1272 LR	615 (49%)	546	196	6	11

Using test image set of known fish targets for validation consisting of 112 right side, 112 left side, 224 left and right side fish images, we got 83%, 68%, 72% hit rates for first three test cases and 90%, 81%, 85% hit rates for the second three test cases. The results show that larger training image sets result in higher hit rates.

This analysis was repeated on images of flag rockfish. The first test consisted of a training set of 3,100 positive images and 3,000 negative images (Table 2.). Flag rockfishes were less successfully detected with a 19% hit rate for a test set which contains 1,272 left and right side images. In the second test the positive images were increased to 3800, improving the hit rate to 49%.

Fish Recognition Using PCA

PCA approach was used with four species of rockfish (genus *Sebastes*) and one species of butterfly fish. The images used in this experiment are shown in Figure 1. Seven images of each species were used as a training set. In order to produce high quality training data, training images were normalized for position and had similar illumination. The result of applying the PCA resulted in 100% successful clustering for every case. This result may be unrealistic, as it was limited by the number of high quality training images, and should be further evaluated with larger image sets, and with fish in different positions and varying illumination. However, as a preliminary assessment, the PCA shows promising results.

There are also modular PCA (MPCA) and weighted modular PCA (WMPCA) which are reported to be more robust than normal PCA (Gottumukkal and Asari 2004) and could further improve performance over the PCA approach.

SIFT Results

The SIFT approach was applied using the VLFEAT tool for four different test cases (Table 3). Using five positive images of butterfly fish and flag rockfish resulted in a 50% recognition rate. With an increase in the number of positive images to 10, a 100% hit rate (#Test = 2) was achieved. Performance seem to have decreased when more potential classes were added to the analysis. As with the PCA, these results are limited by the number of training images. As a result, the SIFT approach will be further evaluated with more images in the future. Current studies showed that SIFT works well when images vary in scale, illumination and pose. Therefore, we think SIFT may be more suitable than PCA for underwater fish recognition.

Conclusions and Future Direction

We tested different detection and recognition algorithms in this project. Our main conclusion is that with a larger training set, we obtain better results. In order to evaluate existing classical object detection and recognition algorithms, we need more robust training data set. In the future, first we will move towards preparing a standardized training and testing database, which will allows us to 1) make a direct comparison between different algorithms for fish detection and identification, 2) identify the most promising fish classification/detection algorithms, 3) assess the state of the art algorithms for fish detection/recognition, 4) to identify future directions of research for fish detection/identification, and 5) advance the state of the art in fish detection and identification. In addition, we plan to test emerging object detection and recognition algorithms with standardized data set. For example, we will test combining computer vision with human effort for fish recognition following a method introduced by Branson et al. (2010).

Table 3. Scale invariant feature transform (SIFT) results.

<table>
<tr><th>#Test</th><th>Used Images</th><th>P / Test Set</th><th>Hits</th></tr>
<tr><td>1</td><td rowspan="2">P. falcifer (butterfly fish) and S. rubrivinctus (flag rockfish)</td><td>5/5</td><td>50%</td></tr>
<tr><td>2</td><td>10/10</td><td>100%</td></tr>
<tr><td>3</td><td>S. miniatus, S. constellatus, and S. levis</td><td>4/4</td><td>33%</td></tr>
<tr><td>4</td><td>P. falcifer and S. rubrivinctus

S. miniatus, S. constellatus, and S. levis</td><td>10/10

4/4</td><td>40%</td></tr>
</table>

Citations

Benson B., J. Cho, D. Gosorn, and R. Kastner. 2009. Field Programmable Gate Array (FPGA) based fish detection using haar classifiers. American Academy of Underwater Sciences, Atlanta, Georgia, USA, March 10-14, 2009.

Branson, S., C. Wah, B. Babenko, F. Schroff, P. Welinder, P. Perona, and S. Belongie. 2010. Visual recognition with humans in the loop. European Conference on Computer Vision (ECCV), Heraklion, Crete, Sept., 2010.

Freund, Y., and R.E. Schapire. 1997. A decision-theoretic generaliztion of on-line learning and an application to boosting. J. Comp. Sys. Sci. 55:119-139

Gottumukkal. R., and V.K. Asari. 2004. An improved face recognition technique based on modular PCA approach. Pat. Rec. Let. 25:29-436.

Lowe, D.G. 2004. Distinctive image features from scale invariant features. Int. Jo. Comp. Vision 60:91-110.

Turk, M., and A. Pentland. 1991. Eigenfaces for Recognition, J. Cog. Neurosci. 3 1:71-86.

Viola, P., and M. Jones 2004. Robust real-time object detection. Int. J. Comp. Vis., 57:137-154.

StereoMeasure and StereoFeatures: Measuring Fish and Reconstructing Scenes Using Underwater Stereo Photographs

Yuri Rzhanov[1] and George Cutter[2]

[1]Center for Coastal and Ocean Mapping, University of New Hampshire, Durham, New Hampshire
[2]Southwest Fisheries Science Center, NMFS, NOAA La Jolla, California
yuri.rzhanov@unh.edu

Abstract

Photogrammetric techniques can and should be used for reconstruction of underwater scenes, in particular for estimation of fish sizes and species' recognition. However, specifics of the medium render most of the algorithms developed for in-air conditions inapplicable without major alterations. The authors describe results of their experimentation with collected imagery and work in progress.

Introduction

Remote sensing and non-lethal approaches for assessments of underwater habitat and biological resources increasingly has relied on underwater images. NMFS routinely deploys cameras on ROVs, AUVs, and other platforms for underwater surveys of fish and habitats. Some of these images are visually inspected and annotated. Relatively few are being processed and used for semi-quantitative analysis of fish abundance, estimation of coverage by certain species, size measurements, etc., particularly when stereo imaging is considered (e.g., Rooper, et al. 2010). The sheer amount of collected data makes obvious the necessity of development of automated processing techniques allowing for substantial reduction of involvement of human operators. In this paper the authors report the ongoing work on automatic reconstruction of 3-dimensional scenes including fish targets ("3D reconstruction") from underwater stereo imagery, which is an initial step for fish detection in complex environments and subsequent species recognition.

Methods

Multi-view imaging proved to be a reliable tool for reconstruction of 3-dimensional scenes: geo-registered urban scenes, reverse CAD engineering, etc. (Frahm et al. 2010). Since the foundational work on multiple view geometry in the 1970s, the associated mathematics have been well understood and formulated.

Almost all of the significant results in this area have been achieved working with the imagery acquired in the air under ambient lighting. It appears that direct application of the same techniques to underwater imagery does not meet performance expectations. The reasons for this include 1) distance- and wavelength-dependent absorption of light by the medium; 2) particles suspended even in a very clear water increase noise in acquired imagery; and 3) the effects of artificial illumination that is required underwater beyond several meters deep.

These reasons also complicate construction of mosaics from underwater images, compared to those from images taken in-air. Many 3-D reconstruction techniques utilize a brightness constancy constraint. This assumption rarely holds even for in-air imagery – it expects only Lambertian light scattering and careful photometric calibration of all cameras (Pons et al. 2007). Although these conditions are not encountered in the real world, this assumption is still being used due to its advantages – it allows working with a single-pixel resolution and to use computationally efficient global optimization algorithms (e.g., Kolmogorov and Zabih 2002). All other algorithms require rectification of image pairs prior to searching for conjugate points (points in different images corresponding to the same feature in the scene). Due to the noisiness common in underwater images, the algorithm of choice must be noise-tolerant (Leclercq and Morris 2003). It also must be local, that is, window-based, as line-based algorithms (Birchfield and Tomasi 1996) are known to suffer from streaky artifacts. The most robust algorithm is one that utilizes normalized cross-correlation (NCC) as a measure of matching quality, and it was chosen for our implementation of "StereoMeasure" software.

Careful calibration of each camera and the stereo rig as a whole allowed for correction of lens distortions and rectification of each stereo pair of images. Rectification performs such transformations on the images that any point feature (visible in both frames) appears in these images on the same row. In other words, epipolar lines become horizontal, and vertical disparity becomes zero. (The experiments have shown later that features located near the frame edges still often have non-zero vertical disparity, which indicates lack of accuracy in the calibration procedure, or inability

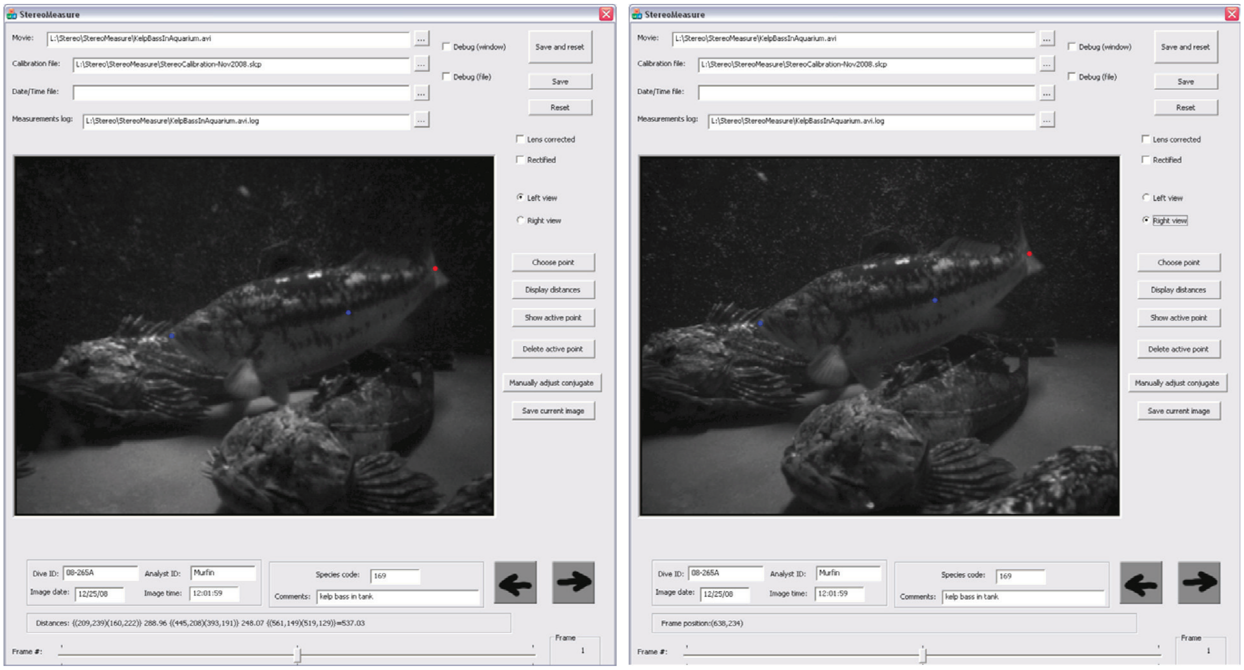


Figure 1. ***StereoMeasure* user interface showing three measurement points and their conjugate points used to measure a fish in images from the left and right camera.**

of the used calibration model to capture the complexity of the optical system.) Measurement of distances between points in 3-dimensional space was performed by the user choosing (clicking in the GUI) these points in one image from a stereo pair. Points were assumed to belong to some opaque surface (e.g., fish skin, rock in the background, etc.). "StereoMeasure" software automatically converts integer pixel locations to a corresponding location in a rectified image (represented by floating point coordinates), searches for a conjugate point in the complimentary rectified image (utilizing epipolar constraint and NCC score), and converts its location back to the space of the original complimentary frame. All points – manually chosen (MC) and automatically found (AF) – are displayed in the GUI for visual verification (Fig. 1). Occasionally, automatic procedure finds conjugate points incorrectly. The primary reasons for this include: 1) choice of initial point in a textureless area; and 2) choice of the point in the area with the repeatable pattern (like fish scales). In these rare cases (less than 2% during our experimentation), the user can open a dialog box with an upsampled version in the vicinity of either an MC or AF point and shift it in any direction. Upon closing the dialog box, all the calculations are refreshed. Changing the location of a MC point would find another AF point (which is then reviewed for accuracy); changing an AF point overrides the automatically found result and accepts manually corrected result as final. Once MC and AF pairs are finalized, "StereoMeasure" performs triangulation in 3D space and calculates distances between sequential triangulated points. The results can be saved in an ASCII file along with metadata provided by the user (e.g., operator's name, number and date of the mission, etc.) More than two points can be identified by the user to enable measurement of length along a curved surface such as the side of a bent fish.

Discussion

"StereoMeasure" proved to be useful and reliable tool for underwater distance measurements, and the extension of the work is underway as the "StereoFeatures" project. The current task is to build a dense disparity map for the image pair which allows 3D reconstruction of a scene—estimation of the spatial locations of all targets and seabed or background elements. The scene resolution is limited only by the resolution of input images. Rectified images then undergo extraction and pair-wise matching of salient point features. We have experimented with SIFT (Lowe 2004) and SURF (Bay 2008) keypoints and descriptors, and the results proved to be very similar.

On smooth surfaces, disparity also changes smoothly, so starting with the extracted "seed" points (which are all assumed to be correctly matched), matching is continued in all directions. The search for a conjugate point is conducted within a square window of predetermined size, and the NCC score is aggregated over a square window of a different size. Obviously, the procedure stops working near occlusions (NCC score decreases), textureless areas (spatial variability of NCC score decreases), and around pixels which, despite being conjugate, demonstrate highly

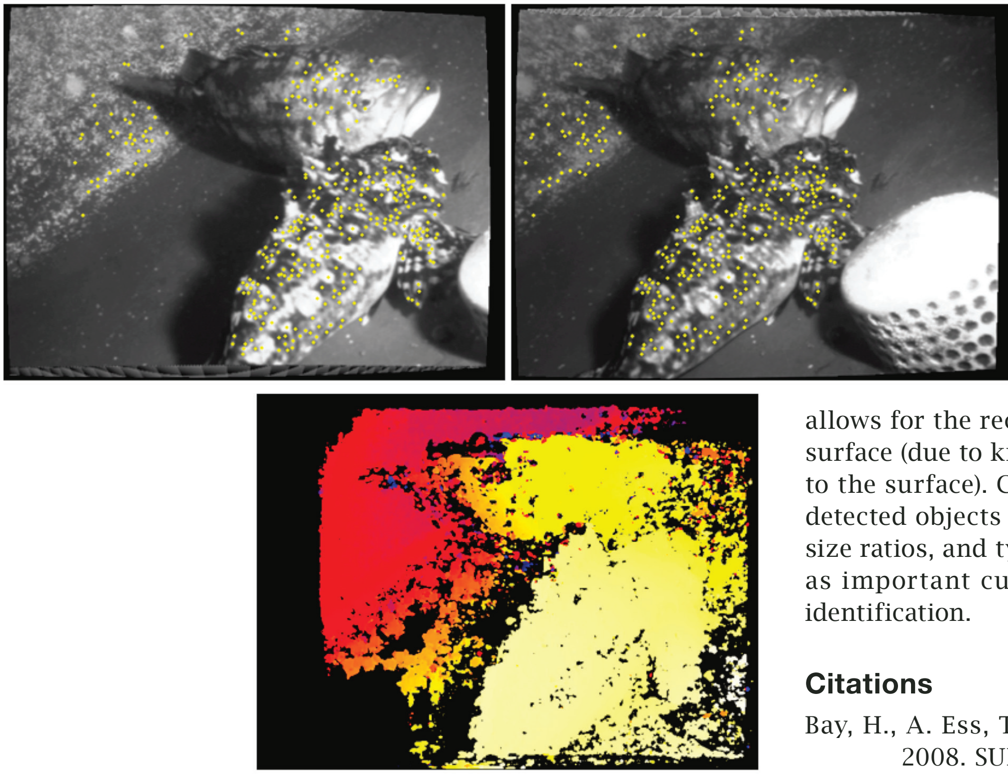


Figure 2. Matched SIFT keypoints for fish and part of the background in left and right stereo camera images acquired from a tank, and (bottom) a rendition of the reconstructed 3-D scene, where colors (yellow to red = near to far) represent distances from the origin, referenced to one of the cameras.

different brightnesses (due to specular reflection, or other reasons). These considerations allow for the formulation of stopping criteria: propagation in a certain direction stops if its NCC score or "textural richness" drops below some threshold. The latter has been defined as the variance of deviations of brightness values from planar surface optimally fit to the current window.

During the first stage, the NCC score is calculated using only integer pixel values. All matches with a score exceeding the threshold are ranked according to this score, and the *n* top-scoring matches are saved as potential candidates (*n* usually ranges from 5-10). Once the entire image is processed, all potential matches are considered in an attempt to minimize the disparity difference for neighboring pixels. The selection process works on "belief propagation" principles. Chosen matches are improved by calculation of subpixel locations of conjugate points. Finally, successful matches are triangulated, creating a cloud of points in the system of coordinates where the left camera is at the origin (Fig. 2). Points calculated from the neighboring matches are immediately linked together creating triangular facets. Those with gaps remain unlinked.

Further processing steps will include detection of 3D objects that potentially represent bodies of fish, separation of these forms from their background, and automated estimation of size and shape. Measurement of water properties (specifically, the dependence of absorption coefficient on light wavelength) allows for the reconstruction of true color of its surface (due to known distance from the camera to the surface). Characteristic measurements of detected objects (sizes in all directions, various size ratios, and typical color patterns) may serve as important cues in species recognition and identification.

Citations

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. 2008. SURF: Speeded-up robust features. Computer Vision and Image Understanding (CVIU). 110:346-359.

Birchfield, S., and C. Tomasi. 1996. Depth discontinuities be pixel-to-pixel stereo. Technical report CS-TR-96-1573. Department of Computer Science, Stanford University.

Frahm, J.-M., M. Pollefeys, S. Lazebnik, B. Clipp, D. Gallup, R. Raguram, and C. Wu. 2010. Fast robust reconstruction of large scale environments. Conference on Information Sciences and Systems (CISS). Princeton, NJ.

Kolmogorov, V., and R. Zabih. 2002. Multi-camera scene reconstruction via graph cuts. Proc. ECCV. 3:82-96.

Leclercq, P., and J. Morris. 2003. Robustness to noise of stereo matching. Proceedings of ICIAP'03. Mantova, Italy.

Lowe, D.G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. 60:91-110.

Pons, J.-P., R. Keriven, and O. Faureras. 2007. Multi-view stereo reconstruction and scene flow estimation wit a global image-based matching score. Int. J. Comp. Vision 72:179–193.

Rooper, C.N., G.R. Hoff, and A. De Robertis. 2010. Assessing habitat utilization and rockfish (*Sebastes* spp.) biomass on an isolated rocky ridge using acoustics and stereo image analysis. Can. J. Fish. Aquat. Sci. 67:1658-1670.

Scantrol DeepVision – Species and Length of Fish in the Trawl

Darren Hammersland-White

Scantrol AS, Bergen, Norway
darren@scantrol.no

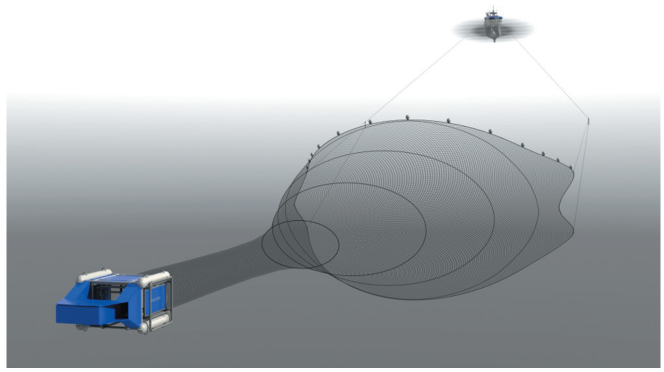


Figure 1. DeepVision system concept.

The collection of biological data on fish species composition and individual length has always been an important requirement for fisheries research, stock assessment and management or policy-making. The actual quantity and quality of data acquired in many areas may not be sufficient to meet this requirement. Recent advances in technology have facilitated the development of certain methods for data acquisition in fisheries, such as video and satellite vessel monitoring systems (VMS) and machine vision systems. One such machine vision system is the CatchMeter (White et al. 2006), a conveyor-based machine vision system for automatic species recognition and length measurement. In this contribution we present an underwater machine vision system, named DeepVision, capable of automatic species recognition, length measurement and sorting of fish in the trawl. The DeepVision system is attached in the codend of a trawl (Fig. 1) and is essentially a frame within which there is a controlled illumination setup. Fish are guided into a transparent channel and as they pass a camera takes side-on, high quality color images (Fig. 2), which are analysed in real-time to give species and length data (Fig. 3). This information is then passed to a sorting mechanism located after the transparent channel, which can either guide fish into a sack or release them again into the sea. The DeepVision can run for 8 hours on batteries and can run automatically or manually via an Ethernet connection over standard netsounder cable to the bridge. So far the system has been trained to recognise four species of fish with a species classification accuracy of 97% for well positioned fish. More fish species will be trained as data becomes available. Initial results show that the length measurements correlate well with manual measurements. With the DeepVision in operation the user has the exact time and depth that individual fish enter the cod end, offering the possibility to acquire species and size data over both time and depth. Furthermore, the user can decide whether to simply release all fish that enter the system or take a subsample for closer examination onboard. The new system will be of interest for use with both research and commercial vessels. The prototype system has been successfully tested on the Norwegian research vessels *G.O. Sars* and *Johan Hjort*. The pilot version of the commercial system is due to be ready autumn 2011.

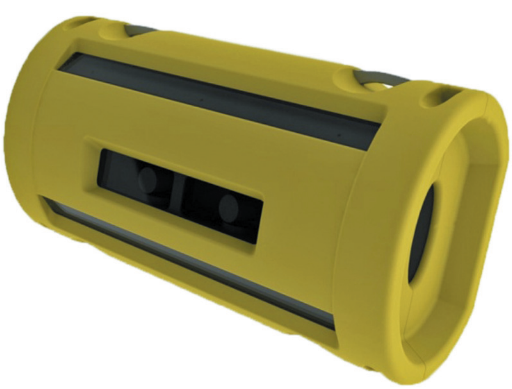


Figure 2. DeepVision subsea unit.

Citation

White, D.J., C. Svellinger, and N.J.C. Strachan. 2006. Automated measurement of species and length of fish by computer vision. Fish. Res. 80:203-210.

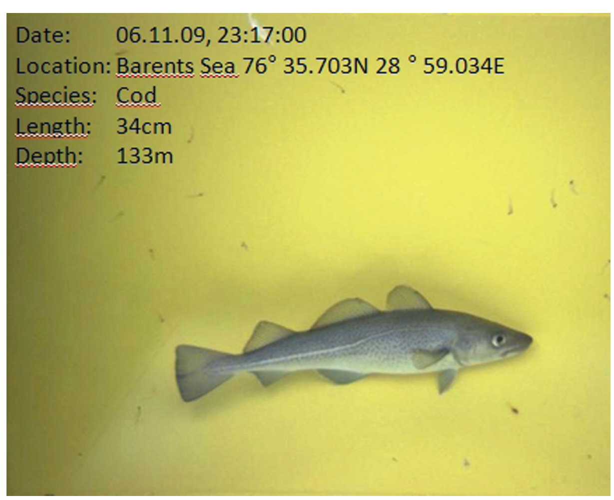


Figure 3. DeepVision image of cod.

An Automated, Real-Time Identification and Monitoring System for Coral Reef Fish

Joseph Wilder[1], Gareth J. Russell[2], and Paul Boyle[3]

[1]CAIP Center, Rutgers University
[2]Department of Biological Sciences, New Jersey Institute of Technology
[3]New York Aquarium, Wildlife Conservation Society
wilder@caip.rutgers.edu

Introduction

The long-term goal of this project is to develop an underwater sensor network to automatically census coral reef fish populations and thereby provide timely warning about possible environmental impacts on the reef. Each sensor may include multiple video cameras and illumination sources. It will identify and count various fish species found near the reef and convey that information to a central data collection node. The data and associated imagery will be remotely accessible for further analysis and for educational activities. The initial phase of this project involves development and field testing of a proof-of-concept prototype of the multi-camera sensor module. Initial tests were conducted at the Glover's Reef exhibit at the New York Aquarium.

Development of the Prototype

System Hardware: The initial concept for the prototype was an open frame that holds two cameras, various lights, and two 'background' panels that block out the complicated scenes of more distant regions (Fig. 1a). The idea was to study smaller species (less than 20 inches in length) that live close to coral reefs to seek refuge from larger predators. Wide-angle lenses were chosen so that the cameras could be close to the frame for structural rigidity, realizing that there would be a large change in magnification throughout the volume, some portions of the volume not covered, and some minor image distortion. High-resolution GIG-E cameras (supporting up to 100 m cable lengths between cameras and an on-shore PC) were chosen. These cameras (Allied Vision Technologies Inc, formerly Prosilica) support multiple programmable functions, for example, frame rate, integration time, gain, and region of interest, to best interact with a difficult and constantly varying environment. Underwater housings and lights (Ocean Presence Technologies) were matched to the depth requirements at our target reef locations. The completed prototype was tested in a tank at the Rutgers Institute of Marine and Coastal Sciences (Fig. 1b), and installed at the New York Aquarium coral reef exhibit (Fig. 1c), along with its environmentally secured instrumentation (Fig. 1d).

Image Capture: High-resolution (1360 x 1024 pixel) image sequences consisting of 300 precisely synchronized image pairs are captured at 10 frames/sec. and stored at a rate of 5 seq./hour (Fig. 2).

Image Pre-Processing: Background modeling and subtraction is carried out in two steps applied to the entire image: Kernel Density Estimation, and a Graph Cut algorithm (an energy minimization method), followed by a local 'clean-up' re-segmentation using the Graph Cut algorithm applied only to the areas within any bounding boxes extracted in the previous stage (Figs. 3A, B). A particle filter-based tracking algorithm is used to track fish across frames and also to detect overlaps between multiple fish (Fig. 3C). Pose correction and scaling is carried out using information from both cameras (via epipolar analysis), providing depth information, pose correction and disambiguation of overlapping fish.

Feature Extraction and Recognition: We are currently using two sources of features for recognition, shape and color. Fish shape, normalized by translation, rotation, scaling and re-sampling of perimeter points, yields a reduced number of features via principal component analysis (PCA). Color features are extracted from RGB histograms and compared with training histograms using a $\chi 2$ divergence measure. Note that controlled lighting and close-in viewing minimize color variations due to imaging through water. Shape and color distance metrics are fused using weighted linear combinations. Classification is via nearest neighbor detection.

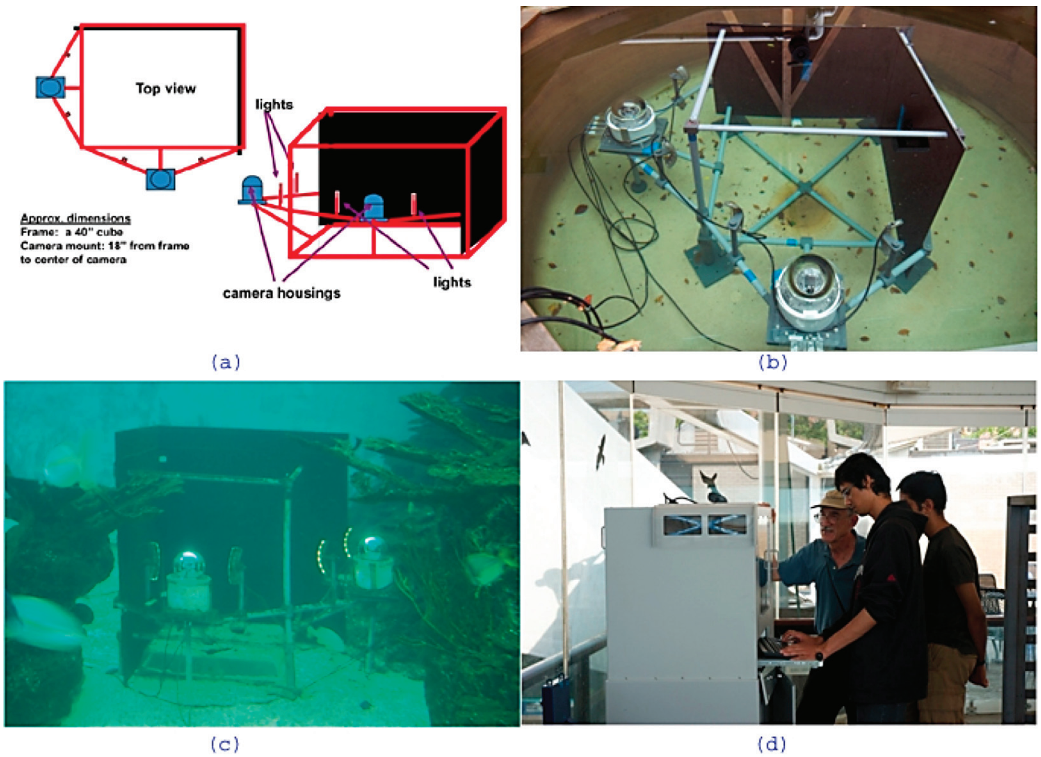


Figure 1. (a) Frame concept design. (b) Initial test of submersible system. (c) System operating at New York Aquarium coral reef exhibit. (d) System instrumentation enclosure.

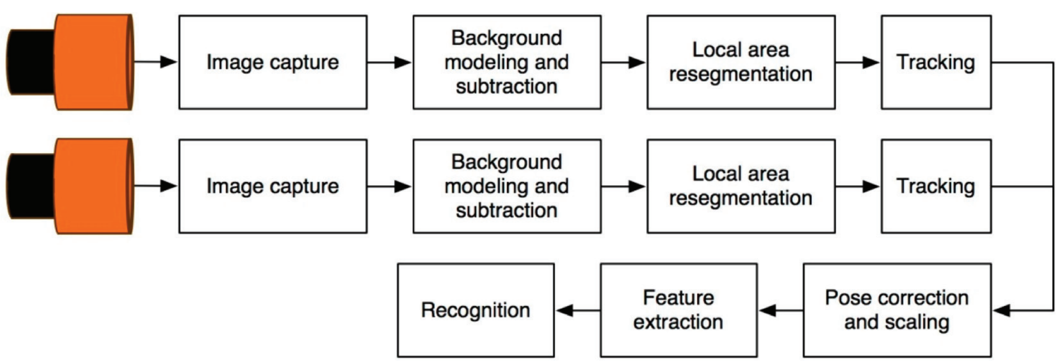


Figure 2. Image processing and pattern recognition flowchart.

A. Image pair captured at New York Aquarium

B. Image pair after segmentation

C. Tracking: red box denotes foreground detection from background modeling, dashed blue-red box denotes the tracker, number indicates the object id

Figure 3. Segmentation of fish from their background (A) and tracking (B).

Preliminary Experimental Results

One hundred seventy-five video sequences containing coral reef fish were recorded at the New York Aquarium, each sequence containing 300 image pairs. In addition to pose and lighting variations, recognition challenges include color variations between members of the same species and small details that distinguish between different species from the same family with similar shapes and colors (Fig. 4). Preliminary recognition results (from 120 fully processed sequences) for eight species are shown in a confusion matrix (Fig. 5). For each species, a test feature vector was selected and compared to a set of vectors obtained by drawing 35 vectors at random from each of the other species. This was repeated 100 times. For each species, this process was repeated over different values of the proportional weighting given to shape and color features to find a species-specific optimal weighting *w*.

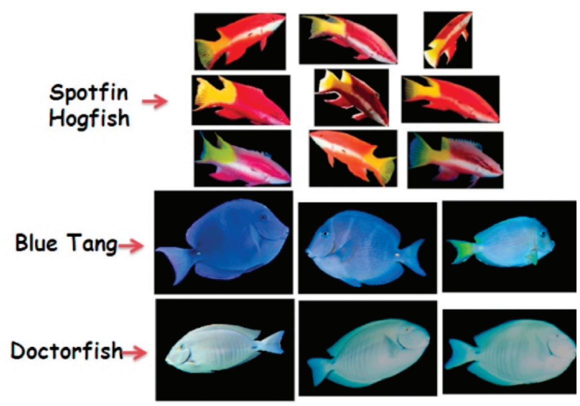


Fig. 4. Recognition challenges: within species (spotfin hogfish) and between species (blue tang and doctorfish).

	1	2	3	4	5	6	7	8	*w*
1	100	0	0	0	0	0	0	0	0.6
2	0	95	0	0	0	0	1	4	0.6
3	0	0	91	0	8	0	0	1	0.5
4	0	0	0	94	0	0	4	2	0.3
5	0	0	0	0	100	0	0	0	0.2
6	0	0	0	0	0	99	0	1	0.1
7	0	0	0	1	8	0	86	5	0.2
8	0	0	0	0	0	0	1	99	0.7

Fig. 5. Diagonal elements represent correct identifications after analysis of eight species. Overall success rate was 95.8%. *w* is the weighting between shape and color features used for each species (higher values indicate stronger weighting towards shape features).

Continuing work includes the analysis of more of our recorded sequences, addition of more species for classification, system integration for fully automatic system operation, data storage and remote access, and design and construction of a more physically rugged system for reliable operation on coral reefs.

Conclusions and Recommendations

Conclusions and Recommendations

This workshop's presentations and subsequent discussions provided many insights into general trends in automated image processing as well as exposing a diversity of problems in marine image-based sampling and approaches to solving them. The following topics illustrate major themes that can guide future research activities and advance the state of automation in marine image-based sampling systems.

1. *Create an international forum or working group for automated analysis of images from marine image-based sampling systems*

   As individual research teams move forward on solving automation challenges, the scientific community within NOAA and beyond working on similar problems could benefit from shared knowledge and experiences, and in a shorter time span than the usual publication cycle. In many cases, operational methods and software developments are not published. With the increasing number of programs seeking new ways of solving common problems, this type of organization could provide a continual and dynamic exchange of knowledge for this fairly specific field.

2. *Inter-disciplinary collaboration*

   Skill sets of biologists and technicians responsible for collecting underwater imagery are generally different from those required for developing automated analysis software. Likewise, computer vision scientists and experts can benefit from involvement in marine projects with clearly defined objectives driven directly by data needs. Rather than supporting graduate students, post-docs, and other research assistantships from marine research fields such as fisheries and oceanography, it may be more beneficial for image-based research programs to reach out to students in computer science and electrical engineering to build interest in applying their knowledge to biological sciences and natural resources, thus fostering future collaborations and partnerships.

3. *Development of a database to facilitate in feature recognition for marine organisms*

   Research on classifying organisms from images generally requires libraries of manually classified training images. Depending on the objectives, a series of individual targets imaged from different aspects and with different backgrounds may be beneficial for creating robust classification algorithms. For this purpose, a shared fish image bank has been proposed, allowing computer vision experts access to validated images otherwise available only within the groups collecting the data. In addition, standardizing image collection and storage would assist in future steps toward automation by providing a common data framework to which algorithms could be applied.

5. *More often than not analysis solutions will have to be customized*

   Image-based systems within NOAA operate in different environments and with different objectives. While there are many common problems that every program faces, often the solutions have to be specific to the situation. For example, some systems may require automated stereo-correspondence analysis coupled with a threshold based segmentation, while others would require pattern-based segmentation of targets followed by target tracking. The best practice may be a modular approach that allows customized analysis structures to built using algorithms, or modules, for sub-processes that may themselves be general. To this end, a medium for exchange of algorithms developed by individual groups but potentially usable by many should be established, providing the building blocks for customized automated analysis systems.

4. *Optimal allocation of automation in analysis*

   There are “easy” and “hard” problems in automation. It may require many years of development before fish targets can be reliably extracted from moving, heavily patterned backgrounds that often accompany benthic camera work. Partially obscured, occluded, or cryptic objects present difficulties for computer algorithms, while human observers can often easily recognize targets of interest. For some of these situations, a combined approach may be taken that still requires human intervention but could automate certain tasks to make more efficient use of manual processing. As continuing progress is made toward full automation, software for assisting analyses would still find immediate uses in current programs, where image analysis is required in a timely manner for management of living marine resources. In other cases, image collection systems should be designed with automation in mind, providing uniform backgrounds and illumination control to make segmentation more precise and effective.

Participants in the September 4-7, 2010 Automated Image Processing Workshop held in Seattle, Washington.

Name	Institute	Email
Invited speakers		
Amber York	Woods Hole Oceanographic Institution	adyork@whoi.edu
D.J. Lee	Brigham Young University	djlee@byu.edu
Danelle Cline	Monterey Bay Aquarium Research Institute	dcline@mbari.org
Darren White	Scantrol systems	darren@scantrol.no
Euan Harvey	University of Western Australia	euanh@cyllene.uwa.edu.au
Gavriil Tsechpenakis	Indiana University-Purdue University Indianapolis	gavriil@cs.iupui.edu
Jacopo Aguzzi	Marine Science Institute of Barcelona	jaguzzi@cmima.csic.es
Janarbek Matai	University of California San Diego	jmatai@cs.ucsd.edu
Jenq-Neng Hwang	University of Washington	hwang@u.washington.edu
Joe Wilder	Rutgers University	wilder@caip.rutgers.edu
Phil Culverhouse	University of Plymouth	P.Culverhouse@plymouth.ac.uk
Scott Gallager	Woods Hole Oceanographic Institution	sgallager@whoi.edu
Yuri Rzhanov	University of New Hampshire	yuri.rzhanov@unh.edu
NOAA		
Bob McConnaughey	Alaska Fisheries Science Center	Bob.McConnaughey@noaa.gov
Carwyn Hammond	Alaska Fisheries Science Center	Carwyn.Hammond@noaa.gov
Charles Thompson	Southeast Fisheries Science Center	Charles.H.Thompson@noaa.gov
Chris Rooper	Alaska Fisheries Science Center	chris.rooper@noaa.gov
Craig Rose	Alaska Fisheries Science Center	craig.rose@noaa.gov
Deborah Hart	Northeast Fisheries Science Center	dhart@mercury.wh.whoi.edu
Elizabeth Clarke	Northwest Fisheries Science Center	elizabeth.clarke@noaa.gov
Jacob Asher	Pacific Islands Fisheries Science Center	Jacob.Asher@noaa.gov
John Harms	Northwest Fisheries Science Center	john.harms@noaa.gov
Josiah Renfree	Southwest Fisheries Science Center	Josiah.Renfree@noaa.gov
Kevin Stierhoff	Southwest Fisheries Science Center	Kevin.Stierhoff@noaa.gov
Kresimir Williams	Alaska Fisheries Science Center	kresimir.williams@noaa.gov
Matthew Dunlap	Pacific Islands Fisheries Science Center	Matthew.Dunlap@noaa.gov
Michael Parke	Pacific Islands Fisheries Science Center	michael.parke@noaa.gov
Rick Towler	Alaska Fisheries Science Center	rick.towler@noaa.gov
Ruth Hass-Castro	Northeast Fisheries Science Center	rhaas@mercury.wh.whoi.edu
Scott McEntyre	Alaska Fisheries Science Center	Scott.McEntire@noaa.gov
Steve Intelmann	Alaska Fisheries Science Center	Steve.Intelmann@noaa.gov
Victor Simon	Northwest Fisheries Science Center	victor.simon@noaa.gov