Evolution of a Benthic Imaging System From a Towed Camera to an Automated Habitat Characterization System

[1]Richard Taylor, [4]Norman Vine, [2]Amber York, [2]Steve Lerner, [3]Dvora Hart, [2]Jonathan Howland, [5]Lakshman Prasad, [6]Larry Mayer, and [2]Scott Gallager

[1]www.seascallop.com; [2]Woods Hole Oceanographic Institution; [3]National Marine Fisheries Service; [4]Independent, [5]Los Alamos National Laboratory, [6]Center for Coastal and Ocean Mapping, University of New Hampshire

Abstract

After two generations of development, we have an operational and practical digital imaging system that delivers high resolution overlapping still images to a computer system on the bridge of a commercial scallop fishing vessel for immediate viewing, storage, and onboard image processing. This non-invasive imaging system produces 100 nautical mile long optical transects of benthic taxa, communities, and associated substrate each day, and is intended to provide fisheries managers with accurate scallop population density estimates and habitat characterization within surveyed areas of the continental shelf. We call the instrument HabCam for habitat mapping camera system. Joint ship operations with NOAA vessels conducting annual scallop surveys has allowed for nearly direct comparison between estimates of scallop abundance by survey dredge and the HabCam imaging system. For 47 transects conducted jointly during 2007, dredge efficiency ranged from 10 to 80% with a mean of 40% (SD 23.9%) depending on area, substrate, tow direction relative to current, and mean distance between the dredge tow track and the HabCam imaging track. Integration of synoptically collected acoustical (675 kHz sidescan, 175 kHz synthetic aperture side scan and 300 kHz multibeam) and optical imaging has allowed for direct registration and comparison of sampling modalities, ground truthing of acoustical data, and extrapolation of information gained at small scale (1m) but high spatial resolution (1 mm) with optics to large scale (>200 m) acoustical data sets. What was initially developed as a scallop survey tool has become an instrument system capable of providing information on habitat characterization, estimates of megafauna abundance, biodiversity, and species richness. A project called the Northeast Bentho-pelagic Observatory (NEBO) is using HabCam to evaluate these ecological parameters at sentinel study sites along the northeast continental shelf repeatedly over several years with the intent of documenting mechanistically how and why benthic community composition is changing over time. A key element in the development of HabCam as a tool for habitat characterization is the automated processing of images for color correction, segmentation of foreground targets from sediment and classification of targets to taxonomic category, and in many cases, to species. A test set of images has been developed consisting of about 30,000 images from each of six sites along the northeast continental shelf representing areas differentially impacted by physical, biological and chemical forcing. Each of these 180,000 images has been manually processed for species counts and sizes so as to provide a training set for automated approaches to target classification. All images and data are available on a public website (http://habcam.whoi.edu).

Introduction

The HabCam system was designed and constructed by Woods Hole Oceanographic Institution personnel using off the shelf components including camera, four strobes, CTD, compass, and sonars (Fig. 1), with all components networked subsea, and all data coming up a single optical fiber for topside storage and processing [1] (Fig. 2). The HabCam imaging system is “flown” 2 to 3 meters off bottom while being towed at 4 to 5 knots (~2 m/sec), thus a track approximately 100 nautical miles is imaged each 24 hour day while at sea. Optical imagery is collected at a width of approximately 1 to 1.25 meters (total ~200,000 m^2/24 hr day).

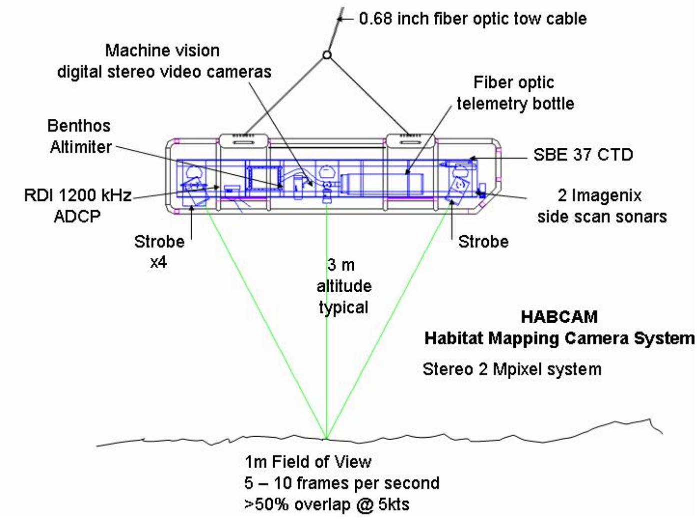

Figure 1. HabCam imaging vehicle with megapixel machine vision camera, four strobes, and a suite of ancillary sensors.

Here we report on use of the HabCam camera system to conduct joint scallop survey work with NOAA vessels, integration of optical and acoustical information for habitat characterization, development of the Northeast Bentho-pelagic Observatory to provide fisheries relevant data in support of Ecosystems Approaches to Management, and initial efforts for automated image processing, target segmentation and classification.

Methods and Results

Joint ship operations

For the past two years, we have been funded by the NOAA sponsored Research Set Aside program to conduct scallop surveys and joint ship operations with NOAA survey vessels in an attempt to “calibrate” survey dredges using the HabCam optical system. Joint ship operations in 2007 with the NOAA vessel R/V Albatross IV and more recently in 2008 with the R/V Hugh Sharp annual sea scallop surveys have allowed one-to-one comparisons between survey dredge catch rates and counts from the HabCam imaging system. During July and August of 2007 ~560 nm of linear imaging transects were conducted in scallop fishing areas, including the Elephant Trunk (ET), Western Great South Channel (WGSC), Nantucket Lightship Closed Area (NLSCA), Closed Area II Habitat Area of Particular Concern (HAPC), and Closed Area I (Fig. 3). Approximately 2,000,000 images were collected and are accessible via the internet at the HabCam website: http://habcam.whoi.edu. Almost one third (170 nm) of these transects were conducted within 1 nm of 105 of the target locations for the R/V Albatross scallop survey stations, with another third imaging the scallop grounds between Albatross stations. While many of the camera transects were made after the Albatross tows, therefore having a precise towpath to replicate, about half were made before the arrival of the Albatross. After close inspection of the data produced from the comparative tows, 47 transects were selected for having track lines closest to the actual Albatross tow, thus being the most

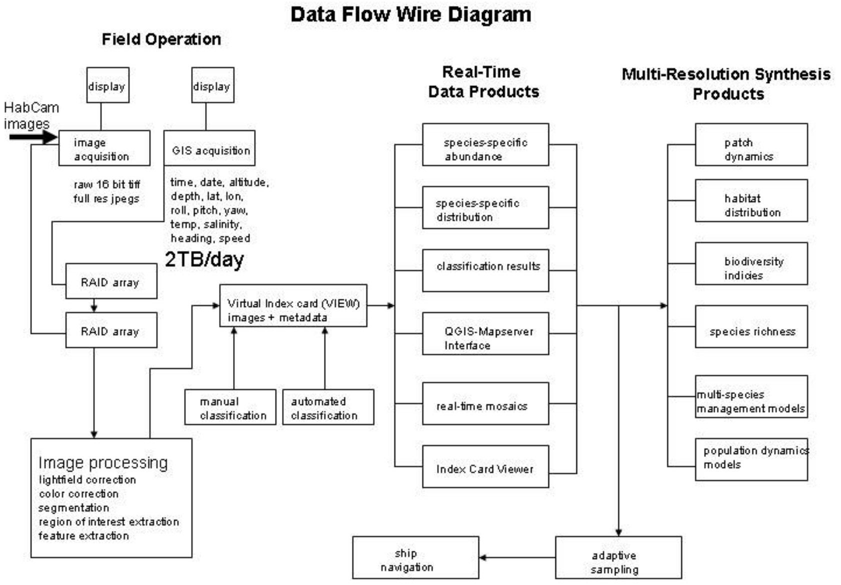

Figure 2. Data flow diagram. Images come up from towed vehicle on top left, processed, viewd in real-time, and stored along with oceanographic data. Manual and automated target classification can proceed in parallel. Real-time data products lead to the development of multi-resolution synthesis products used in population models.

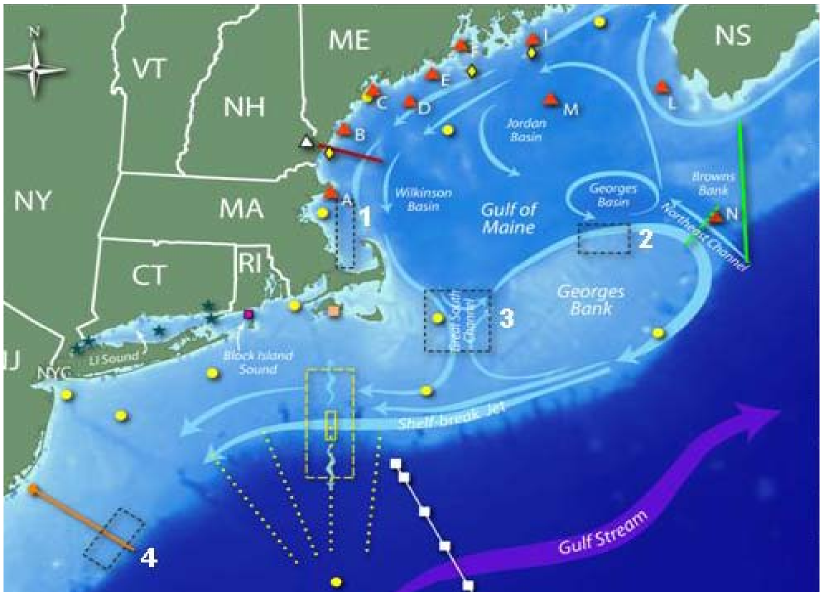

Figure 3. Study sites. 1. Stellwagen Bank including Jefferies Ledge. 2. Closed Area II, HAPC. 3. Closed Area I. 4. Hudson Canyon and Elephant Trunk. 5. Nantucket Lightship Closed Area.

representative for direct comparison. In addition approximately 60 nm miles of 300 kHz multibeam (Kongsberg 3002) data were collected in the NLSCA, and ~125 nm of high resolution 175 kHz sidescan were collected in the HAPC, Closed Area I, and WGSC. Two hundred nm of combined sidescan and multibeam data were collected in the area of outer Cape Cod, Stellwagen Bank National Marine Sanctuary SBNMS and Jeffreys Ledge.

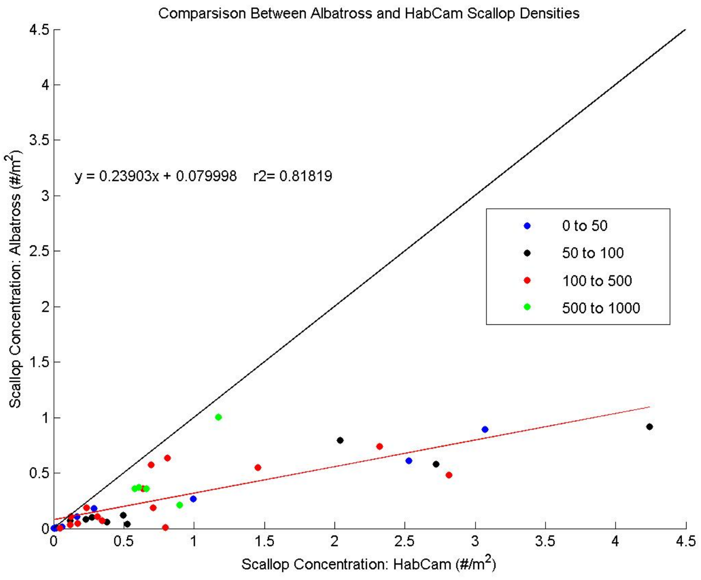

Figure 4. Direct comparison between scallop densities obtained by paired dredge tows and transects made by HabCam. See text for explanation.

A standard scallop survey dredge is towed for 15 min along a 1 nm track line at a ship speed of 3.8 kts, thereby sweeping 4,500 m^2 per tow. The HabCam vehicle was towed as close as possible to the dredge track line at a ship speed of about 5 kts imaging at 5 images per second providing about 50% overlap for mosaicing purposes. Total survey area for the HabCam is approximately 1,852 m^2 per tow. For this study, all scallops were counted by hand from each image track line and abundance calculated by dividing number of scallops by the survey area in m^2 . The correspondence between the concentration of scallops as estimated by dredge tows made by the Albatross and compared with direct observation from the HabCam images is shown in Fig. 4. Each data point represents a single paired 1 nm tow. Since the two tows were never exactly superimposed on each other, the points are color coded by the average distance between the two tows (0 to 50m, 50 to 100m, 100 to 500m, 500 to 1000m). Data are from the Elephant Trunk, Nantucket Lightship, Western Great South Channel and Closed Area II HAPC. The black line indicates a 1:1 correspondence between sampling modalities. Note that about half of the tows were in areas where the scallop concentration was 0.5 m^2 or less. While the overall r^2 value was 0.818, the lowest correlation was for comparisons where the distance between tows was greater than 500m indicating that the scallop populations were patchy. There was one paired tow in WGSC where no scallops were detected in either the dredge or by HabCam. The overall range of concentrations based on scallop counts and area swept for HabCam and survey dredge was from 0 to 4.4 and 0 to 0.9 scallops/m^2, respectively. In all areas and conditions, HabCam always produced higher scallop densities than captured by the dredge. Given the direct comparison between the dredge and HabCam we can calculate a dredge efficiency. Fig. 5 shows that the relative efficiency ranged between 10 and 80% with a mean of 40.1 (+/- 23.9, n= 46). Dredge efficiency will vary as a function of tow angle, ship speed, current direction and speed and bathymetry. The intent was to sample in a variety of areas under a number of conditions to estimate overall dredge efficiency as correlated with these other factors. Data from a single paired tow is given in Fig. 6 as an example of correspondence between the two track lines, comparison between size distributions, and dredge efficiency. Note that the size distributions are quite similar between the dredge and HabCam measurements with mean shell heights within 5 mm.

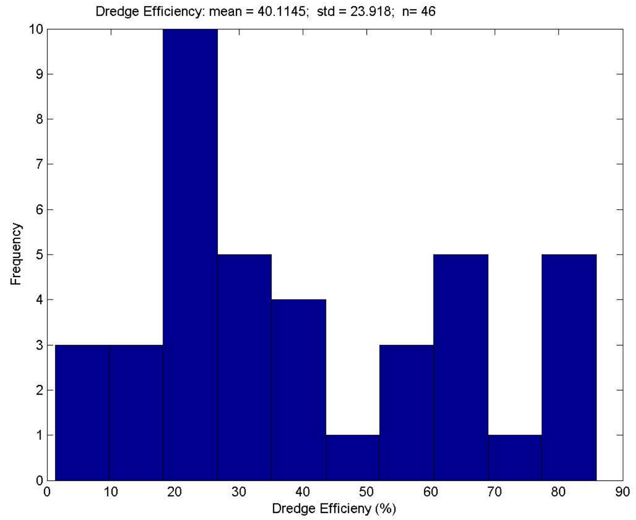

Figure 5. Frequency distribution of dredge efficiencies from all areas sampled. Mean 40.1 +/- 23.9, n= 46.

Fusion of optics with acoustics

Because of the physical limitations of light penetration in water, optical imagery provides high spatial resolution (1 mm), but only over relative small areas (1 m^2) when flying at an altitude of 2 to 3 m. At higher altitudes image quality suffers because of light scatter from particulates. Mulitbeam and sidescan acoustic techniques offer the opportunity to rapidly survey much larger areas of the seafloor albeit at reduced resolution relative to optics. One of our objectives has been to develop approaches for using quantitative acoustic measurements to extrapolate, with confidence, habitat-relevant parameters between sparse lines of high resolution optical imagery. Optical mosaics are registered with sidescan and multibeam bathymetry to provide spatial resolution of millimeters to kilometers and used to ground truth and verify the acoustic data. Our approach involves the collection of high-resolution (300 kHz) bathymetery and backscatter from a multibeam sonar, sidescan (175 and 200 kHz), and forward or side looking (675- 1000 kHz) data. At these frequencies, multibeam sonars can produce bathymetry with resolution on the order of 0.02 - .05% of the water depth (or height above the bottom in the case of a system on a towed or autonomous vehicle) and lateral resolution on the order 1 – 2 % of the water depth (or height above the bottom). While attaining these levels of resolution, these sonars also typically cover an area of between three and five times the water depth (depth below transducer, whether towed or hull mounted). For example, in 70m of water a swath of 350m wide is attained using a Kongsberg EM 3002. When mounted on our optical imaging vehicle, a Teledyne Benthos C3D sidescan/bathymetry sonar covers 15m to either side providing cm resolution at an altitude of 3m. Embedding and registering data from the two acoustic systems provides a powerful, multi-scale backdrop into which optical imagery may be registered.

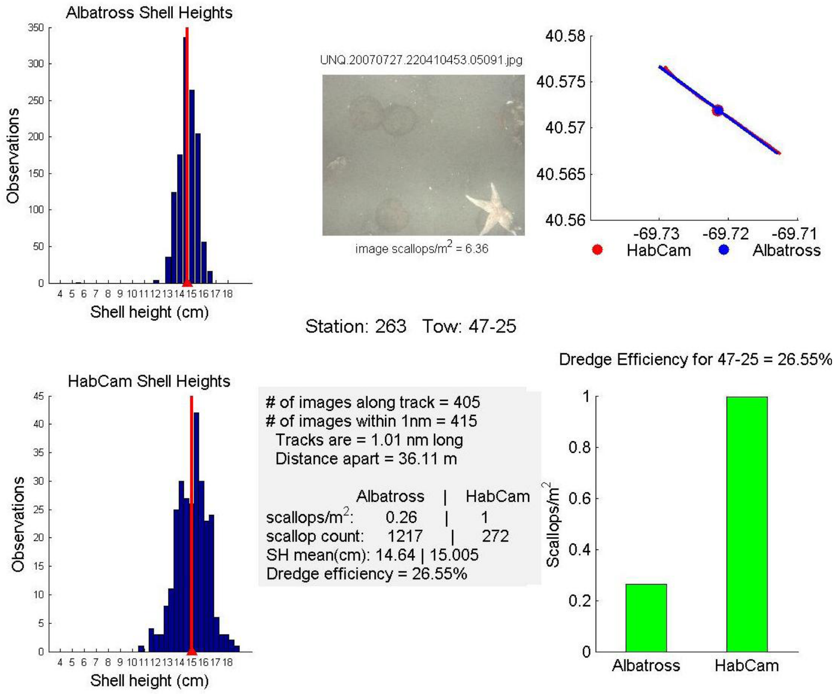

Figure 6. Data from an example paired tow between the Albatross dredge and the HabCam imaging system from Nantucket Lightship Closed Area. Both transects are 1 nm (top right). Areas swept are given in Table 1. Scallop concentrations quantified by the dredge are 26.5% of that observed by HabCam (Bottom right). Shell heights are within 5mm for both systems (Top and bottom left). Image in center is representative of the area.

The starting point in the process of using backscatter for quantitative seafloor characterization is the correction of the backscatter for radiometric and geometric factors [2]. These corrections involve the removal of the effect of changes in gain, power level and pulse width, and the effect of any residual beam pattern. If the detailed bathymetry is known (as is the case for multibeam sonars, which collect both bathymetry and backscatter), then the effective incident angle and true position on the seafloor can be calculated. If such corrections are applied properly, then the observed differences in acoustic backscatter are related only to differences in seafloor properties and the backscatter values will represent the actual backscatter cross-section returning from the seafloor. The acoustic backscatter values from different acquisition lines can be reduced to a near-calibrated scale of scattering strength, and can be compared directly to ground truth (HabCam images) or to a mathematical model [3].

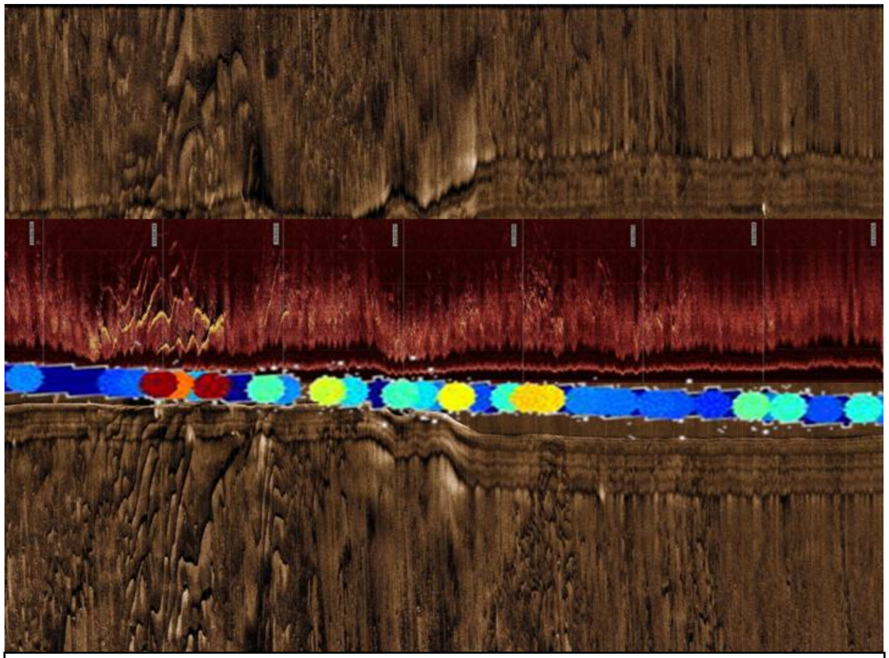

Figure 7. Example of 125 kHz side scan 200m wide by 400 m long with superimposed 600 kHz side scan 50 m wide and data from HabCam on ground fish abundance. HabCam data are binned at intervals of 50 m. Color represents abundance and ranges from 0.03 (dark blue) to 0.3 (red) fish/m^2. See text for further explanation.

Our approach is to collect multibeam, sidecsan and down-looking sonar data that is co-located with HabCam image data. We then run Angular Range Analysis (ARA) analyses on the sonar data and compare these results with the range of products derived from the analysis of the HabCam images. We use visualization and data fusion techniques to explore the relationships between the acoustically derived parameters and those derived from the HabCam. As an example of image optical and acoustic data fusion, Fig. 7 shows a acoustic track taken in Great South Channel east of Cape Cod using a 175 kHz synthetic aperture side scan (SASS) sonar. The track is 200 m wide and 400m long and shows large bathymetric features such a sand waves. Superimposed on the SASS image are data from a 675 kHz Imagenix side scan revealing fine details of the sand waves. Superimposed on top of that are data extracted from the HabCam optical imagery as the abundance of ground fish such as cod, haddock, plaice and other flat fish aggregated into 50m long bins. Note that ground fish abundance is relatively low where the seafloor is flat but increases by a factor of 10 where the sand waves increase in size. Sand waves and other bottom structures are known to provide habitat for fish.

The Northeast Bentho-pelagic Observatory (NEBO)

What was originally envisioned as a tool to survey sea scallops and other ground fish has evolved into a system for characterizing benthic community structure, sediment characteristics, and water column properties. The NOAA funded NEBO project is designed to produce unique data products for fisheries and marine protected area managers and to foster development of ecosystem approaches to management (EAM)(project website, http://nebo.whoi.edu). We are observing and quantifying key taxa, benthic community structure, species diversity, seafloor habitat characteristics, and coincident water column properties with repeated measurements in multiple, sentinel sites on time scales of weeks to years. At sentinel sites along the U.S. Northeast coast that have both high fisheries and conservation value (Fig. 3), we are quantifying how communities respond to system change (climate events, fishing activity, position of oceanographic features [fronts], etc). This approach requires fusion of disparate, synoptically acquired data sets, including high-resolution acoustic bathymetry and backscatter (on scales of meters to kilometers), stereo optical imagery (on scales of millimeters to meters), water column plankton distributions (microns to millimeters), and the development of image bioinformatic tools for classifying targets and substrates. Integrated data products are being developed using advanced visualization tools so key fishery target species and non-target community responses to regulatory practices can be observed and quantified at multiple, relevant space and time scales, in relation to variations in seafloor habitat and boundary layer conditions.

While the HabCam system has achieved rapid progress, significant challenges remain particularly in the realm of fully automating the segmentation and identification of targets and substrate, automating the mosaicing of contiguous imagery for determining inter-animal distances, and the fusion of all data into a near real time display where the high level goal is to be able to navigate the vessel using the incoming data.

Image test sets

A key element in the development of HabCam as a tool for habitat characterization is the automated processing of images for color correction, segmentation of foreground targets from sediment and classification of targets to taxonomic category, and in many cases, to species. A test set of images has been developed consisting of about 30,000 images from each of six sites in the northeast representing areas differentially impacted by physical, biological and chemical forcing (fig. 3). Each of these 180,000 images has been manually processed for species counts and sizes so as to develop a training set for automated approaches to target classification. Data are available from the HabCam website.

Automated segmentation and classification

As described in Gallager et al. [4] and Howland et al [1], images are acquired at 5Hz and processed in real-time on the ship. Raw 16 bit Tiff images are light field corrected, color corrected, and converted to 24 bit jpegs. The jpegs are then segmented into regions of interest (ROI) which removes foreground targets from the background sediment. The ROIs are then subjected to feature extraction and classification into known categories such as flatfish, sponge, scallop, starfish, etc. Our most challenging problem in image processing by far is segmentation of foreground targets from background sediment.

We are developing segmentation algorithms based on image texture, color and edge features that will operate efficiently in real-time following transmission of the images onto the deck of the ship using the approach being developed at the Los Alamos National laboratory [5;6]. Briefly, the image is first filtered through a Canny edge detection algorithm to extract edge pixel chains followed by application of a constrained Delaunay triangulation of the edge contour set, which yields triangles that tile the image without crossing edge contours. Each triangle is given a color based on pixel content with in it. Triangles are then merged into polygons based on a set of rules such as proximity, continuity, color, etc. The resulting polygons are represented in the segmented image by a variety of targets (Fig. 8). The process essentially converts a raster image into a vector image with polygons in the vector image representing targets of interest. A variety of grouping filters or rules may be implemented during the triangle aggregation stage to achieve desired results. Characterization of sediment (sand, gravel, cobble, boulder) actually falls out of the segmentation process since the process of triangulation and development of polygons encompassing regions of similar texture and color provide a measure of surface area and grain size. Grain size is then mapped back onto a simplified substrate characterization scheme as that provided by Valentine et al. [7].

Figure 8. Segmentation of image taken in Closed Area II, HAPC of nibe scallops on a sandy bottom. Raw raster image (left), fully segmented vector image (center), vector image where polygons have been aggregated based on similar color and texture features. Note that the scallops have been successfully segmented.

Classification of the segmented polygons begins with feature extraction followed by both unsupervised and supervised classification. Features currently being extracted include surface texture (entropy, energy, correlation, and homogeneity) and color (ratio of red, blue, green image planes). Morphological descriptors such as size, excentricity, elipticity, Fourier descriptors, etc. are also used when the target of interest has a defined shape, but these are not useful features for low growing colonial forms such as tunicates, sponges and bryozoans. The features are then subjected to a principle components analysis to reduce dimensionality of the data set. The first three principle components are used as features in a Support Vector Machine classifier, which has been trained with manually classified image data from the image test sets. Table 1 shows the results of a classification run using five target categories. Percentage accuracy varied from 77% for seastars to 87% for razor clams. The high accuracy for razor clams is attributable to their very distinct elongated shape compared with the other categories. Detailed results for classification of a variety of taxonomic groups will be reported in future publications.

Table 1. Results of classification of 5 taxonomic groups using a Support Vector Machine.

				Manual			
							% false +
		scallop	seastar	sediment	Razor clam	other	
	Scallop	5984	8	240	43	193	8
	Seastar	397	842	111	39	38	
Automated	Sediment	197	49	7839	31	2837	
	Razor clam	54	129	62	1825	927	
	Other	467	58	1132	138	15736	
% false -		15					
total manual		7009	1086	9384	2076	19731	
% accuracy		84	77	83	87	79	

References

[1] Howland, J. S Gallager, H Singh, A Girard, L Abrams, C Griner. 2006. Development of a Towed, Ocean Bottom Survey Camera System for Deployment by the Fishing Industry. Oceans06, Boston

[2] Fonseca, L. and Calder B. (2005), "Geocoder: and Efficient Backscatter Map Constructor
Proceedings of the U.S. Hydrographic 2005, San Diego, CA.

[3] Fonseca, L. and Mayer, L. A., in press, Remote estimation of seafloor properties through the application of Angular Response Analysis to multibeam sonar data, Marine Geophysical Researches

[4] Gallager, SM, H Singh, S Tiwari, J Howland, P Rago, W Overholtz, R Taylor and N Vine. 2005. High resolution underwater imaging and image processing for identifying essential fish habitat. National Marine Fisheries Service Workshop on Underwater Video analysis. DA Somerton and CT Glendill (eds) NOAA Technical Memorandum NMFS-F/SPO-68. pp. 44-54.

[5] Prasad, L. A. N. Skourikhine, *Vectorized Image Segmentation via Trixel Agglomeration*, Pattern Recognition vol. 39 pp 501-514, 2006

[6] Prasad, L. A. Skourikhine, “Raster to Vector Conversion of Images for Efficient SVG Representation”, *Proceedings of SVG Open 2005*, Enschede, The Netherlands, August 2005.

[7] Valentine, OC, BJ Todd, and VE Kostylev. 2005. Classification of marine sublittoral habitat with application to the northeastern north America region. American Fisheries Society Symposium 41:183-200