Using Object-Oriented Classification of ADS40 Data to Map the Benthic Habitats of the State of Texas

by Kass Green and Chad Lopez

Introduction

The coast of Texas supports a diversity of marine habitats as well as providing an abundance of recreational opportunities and contributing significantly to the Texas economy. However, as populations increase, human impacts to Texas coastal ecosystems are also increasing. Managing and protecting this diverse and sensitive resource requires knowledge of the state's coastal marine habitats and understanding of the causes of change in these habitats over time.

Recognizing this need, the State of Texas recently adopted a Seagrass Monitoring Program, which calls for regional mapping of Submerged Aquatic Vegetation (SAV) for status and trends assessment. To support this program, National Oceanic and Atmospheric Administration's (NOAA) Coastal Services Center is working cooperatively with the Texas Parks and Wildlife Department (TPWD) and the Texas A&M University Center for Coastal Studies to develop benthic (seafloor) habitat data, primarily SAV, for coastal bays and lagoons along the Gulf Coast of Texas. NOAA chose Fugro EarthData, Inc. and a team of professionals from Avineon and The Alta Vista Company to map the bay's benthic habitats based on the Florida *System for Classification of Habitats in Estuarine and Marine Environments* (SCHEME) (Madley *et al.*, 2002). The project consists of two phases. Completed in June of 2007, Phase 1 encompasses 1,400 square miles of the six central Texas Bay systems as illustrated in Figure 1. Phase 2 is comprised of the two large remaining bays and will be completed in 2008.

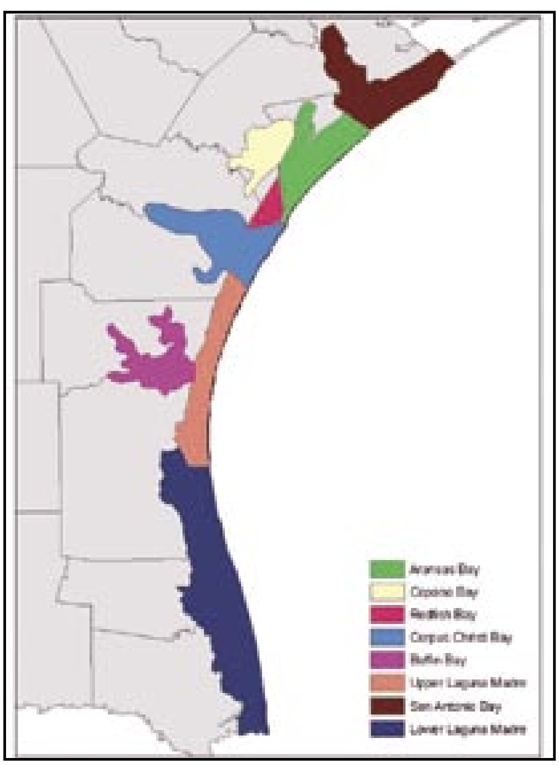

Figure 1. The extent of the entire project area encompassed the bays and lagoons of the Gulf Coast of Texas. Phase 1 included Aransas, Copano, Redfish, Corpus Christi, Baffin, and Upper Laguna Madre Bays. Phase 2 will complete the project by mapping San Antonio and Lower Laguna Madre Bays.

This article reviews the innovative methods developed by the Fugro EarthData team to produce highly detailed maps of benthic habitats. The project is unique in that it successfully utilized multiple new technologies in a production environment including:

- The use of digital airborne imagery rather than film aerial photographs. The project utilized the 2004 National Agriculture Imagery Program (NAIP) imagery collected over Texas for the U.S. Department of Agriculture with the Leica ADS40 airborne digital camera and processed by Fugro EarthData. The 1 meter pixel resolution, 12-bit NAIP imagery was re-sampled to 2 meter, 8-bit data and reprocessed to include the true color and color infrared bands. As the cover of this journal illustrates, although the imagery was collected for agriculture monitoring applications, it was unintentionally collected during a time period with great visibility into the water column and little or no wind or turbidity, which made it ideal for benthic mapping applications.
- Application of automated methods to digital airborne data. While use of automated methods is standard in moderate resolution satellite imagery classification projects, the majority of high resolution imagery thematic mapping projects (whether based on film or digital imagery) are completed using photo interpretation. This project utilized newly developed automated image classification tools that allow for both the management and exploitation of the massive amounts of data available in high resolution imagery.
- Reliance on automated image classification technologies instead of manual photo-interpretation. Traditionally, benthic habitats are created through manual interpretation which is labor intensive and lacks the detail available from automated techniques. However, until now, automated techniques have repeatedly failed when applied to benthic habitats. This project represents one of the first successful applications of automated techniques to benthic habitat mapping, and the first successful application of the methodology in a production environment.

continued on page 862

continued from page 861

Methods

Using remotely-sensed data and ancillary information to characterize benthic habitat is effective only because a high correlation exists between variation in the imagery and ancillary data, and variation in benthic habitat as specified by the classification scheme. Hopefully, when the habitat changes, the spectral response of the imagery and the classes of the ancillary data also change. Making a map requires understanding what causes variation in the benthic habitat and understanding how the imagery and ancillary information are responding to those variations. Thus, mapping necessitates

Step 1: developing a rigorous and rule-based classification scheme to specify the type of habitat characteristics to be detected and mapped,

Step 2: understanding how and why the relevant habitat characteristics vary in the field

Step 3: capturing and classifying all the variation in the imagery and ancillary data sets that is related to the variation in the classification scheme.

Following detailed analysis and evaluation of several alternative methodologies in a pilot project, Fugro EarthData chose to pursue an object-oriented image classification approach which relied on a combination of field investigations, image segmentation software, and classification and regression tree (CART) analysis. Object-oriented image classification classifies image objects or segments (i.e. groups of pixels delineated as polygons) instead of individual pixels, allowing for the incorporation of shape and context into the creation of habitat data. While powerful in the classification of moderate resolution data, object-oriented classification is pivotal for automated classification of high resolution satellite or airborne imagery because of the mixture of both shadow and illuminated pixels depicting features in the imagery.

To complete the first step, a dichotomous key based on the Florida SCHEME was developed for the following labels of benthic and coastal habitats.

- Benthic Habitats
  - Continuous SRV (seagrass)
  - Patchy SRV
  - Oysters
  - Unconsolidated Sediments
  - Hardbottom
  - Unknown Habitat
- Land and Land Interface Habitats
  - Emergent Marsh
  - Mangroves
  - Other Land

Possible modifiers to these labels included salt flat, drift algae, shell hash, prop scar, drift wrack, and mat algae.

Next, Fugro EarthData divided the ADS40 imagery into six geographic regions and segmented each region into objects using Definiens Professional software (previously named eCognition Professional). Similar to the manual delineation of polygons in photo interpretation, segmentation algorithms delineate digital imagery into polygons of pixels that have more variation between than within, based on a user defined set of inputs layers. In this project, the four ADS40 raw spectral bands (red, green, blue and near infrared) were used as input layers into the image segmentation algorithm.

To understand how the benthic classes vary (Step 2 above), calibration and validation field trips were conducted. Both trips were supported and attended by NOAA, State of Texas, and Texas A&M personnel, whose contributions of expertise and resources were critical to the success of the project. Involvement of end users in field trips is essential for ensuring that map producers and map users "see" habitats in the same way.

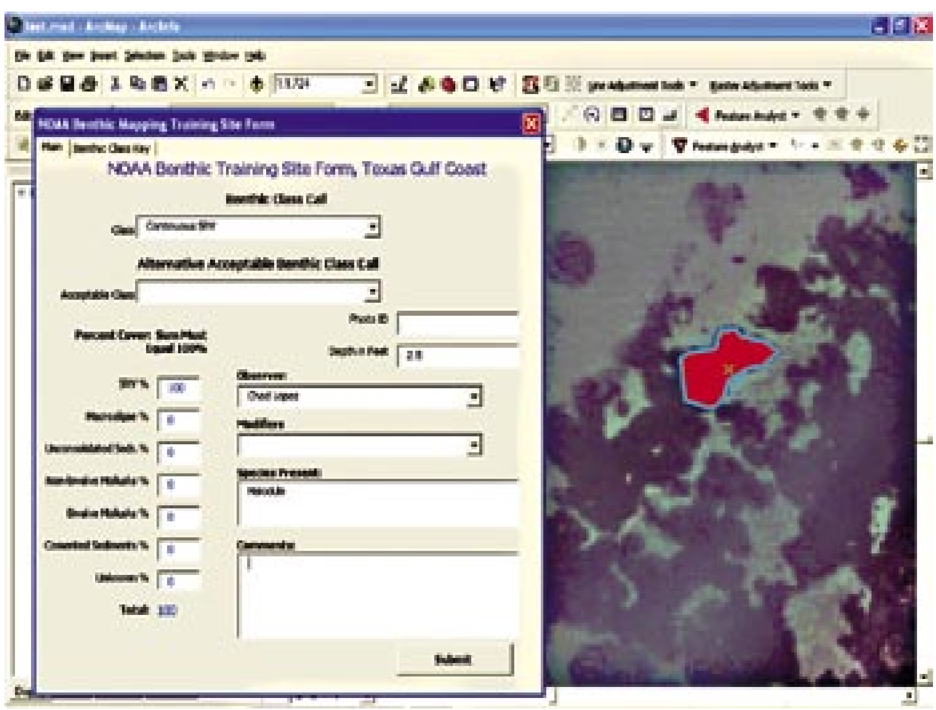

Figure 2. This displays the digital field form developed for the project. Because the classification scheme rules rely on percent cover, each sample site was characterized in terms of percent cover and automatically labeled based on the scheme dichotomous key.

Figure 3. The field data entry set up in one of the Texas Parks and Wildlife boats. Notice the link to the GPS device, which allowed field personnel to use the digital camera imagery to navigate directly to chosen sample segments as displayed on the computer screen.

The calibration trip occurred early in the project and resulted in data collection for 541 sample segments. Field information for each sample segment was entered into an ArcGIS, GPS linked, field form developed specifically for this project by Tukman Geospatial. As displayed in Figures 2 and 3, the form provided access to the digital camera imagery and the segments, and included the dichotomous key rules developed for the habitat classification scheme In addition, extensive field notes were hand written on hardcopy image maps of the project area. Upon return to the office, data on an additional 789 sample segments were collected through manual interpretation of the imagery for a total of 1,330 sample sites. A random number generator was then used to select and set aside 100 sites per benthic class for accuracy assessment.

To understand how the variation in benthic habitat was correlated with variation in the imagery and ancillary data (Step 3, above), a CART analysis was performed on the non-accuracy assessment sample data against the imagery and ancillary data layers using See5 statistical software. CART analysis builds tree diagrams for predicting variables from categorical and continuous data (Breiman *et al.*, 1984). CART analysis "mines" the sample data and builds rules which are if-then statements in hierarchical "trees" that condition the prediction of habitat classes. In the case of this project, the dependent variable was the habitat class (as specified by the classification scheme rules), and the independent variables were the various ADS40 imagery and ancillary data sets (e.g. polygon shape, NWI class, bathymetry, etc.). Using CART to build the classification rules forces consistency and analytical rigor into the segment labeling process.

To assess the efficacy of alternative sets of CART-generated rules, labels of the independent accuracy assessment sites (100 sites for each class) were compared with their CART-generated labels in error matrices (Congalton and Green, 1999) such as the one presented in Table 1. The chosen object-oriented automated techniques obtained overall classification accuracy of 74%, *prior to any editing*. Following the analysis of the error matrices, the Fugro EarthData team applied the input layers, CART rules, and techniques that generated the highest accuracies to the rest of the image segments to generate an initial map of benthic habitat. This map was lightly edited and provided to NOAA and their Texas partners for review.

The initial map was then evaluated in the office and in the field during the second field trip -- the validation trip. Areas to be visited in the field were chosen according to

- class confusion identified in error matrix,
- other confusion identified by the image analysts,
- areas that lacked field data from the previous calibration trip, and
- comments on specific sites from NOAA and Texas personnel.

Using information collected during the field trips and comments from NOAA and Texas end users, the benthic map data was thoroughly manually reviewed and edited. Editing focused on the correction of polygon labels. Only minimal line editing was required. Next, vectors between segments with same labels were dissolved, and vectors were smoothed to reduce the stair step effect caused by the square boundaries of the image pixels. The resulting contractor final map was then provided to NOAA and the Texas partners for review during their independent validation trip.

The Phase 1 final map contains over 50,000 polygons depicting the habitats of 1,400 square miles of coastal Texas. Figure 4 illustrates the complexity and detail of a sample 2.3 square mile area of the map. Table 2 displays the contractor final map accuracy using both deterministic and fuzzy accuracy methods (Green and Congalton, 2004). Overall accuracy is quite high at 90% with users and producers accuracies predominantly above 80%.

More importantly, NOAA and their Texas partners are extremely satisfied with the product. As stated by NOAA's Bill Stevenson and Mark Finkbeinder, "The automated methods, supplemented by visual analysis, as demonstrated in this project, have produced outstanding results. The data have exceeded our expectations for classification detail, thematic accuracy, and spatial precision. We

continued on page 864

Table 1. The error matÐ -
able with wÐ
indicates where editing will be best focused (e.g. SRV, unconsolidated sediments).

				Reference Data							
	Algae	SRV	Land	Mangroves	Oysters	Sediments	Unknown	Emergent Marsh	TOTALS	User's Accuracies	
Map Data	Algae	**53**	5	0	0	1	6	0	0	65	0.82
Map Data	SRV	18	**68**	3	0	20	40	11	4	164	0.41
Map Data	Land	1	0	**89**	6	1	3	0	6	106	0.84
Map Data	Mangroves	0	0	0	**88**	0	0	0	0	88	1.00
Map Data	Oysters	2	9	1	0	**80**	11	4	0	107	0.75
Map Data	Sediments	0	9	1	0	1	**27**	2	0	40	0.68
Map Data	Unknown	1	5	0	0	3	7	**83**	1	100	0.83
Map Data	Emergent Marsh	0	4	6	6	4	6	0	**89**	115	0.77
	TOTALS	75	100	100	100	110	100	100	100	785	
	Producer's Accuracies	0.71	0.68	0.89	0.88	0.73	0.27	0.83	0.89		0.74 Overall Accuracy

continued from page 863

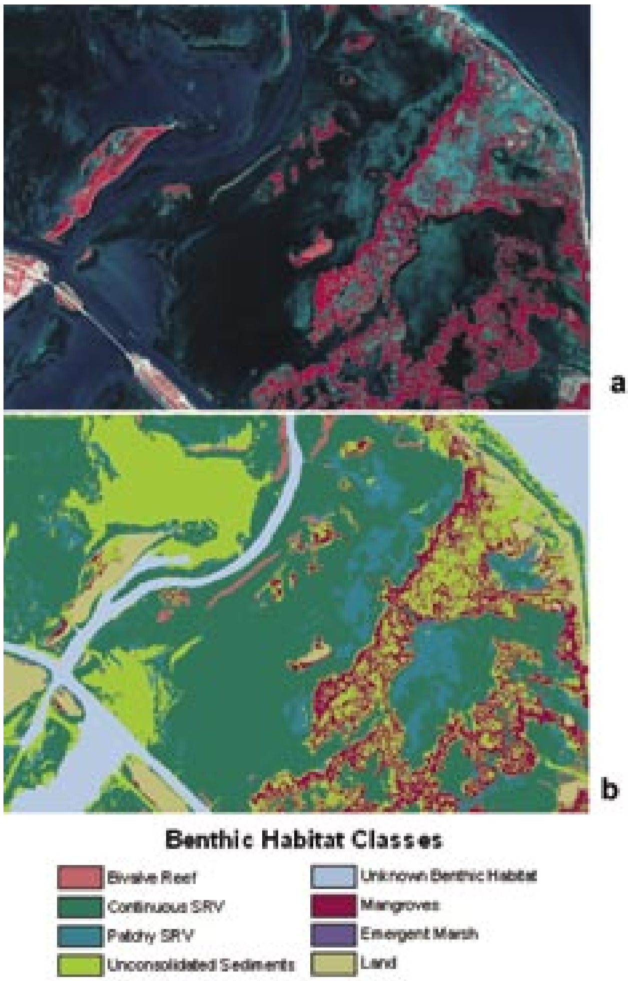

Figure 4a. A color infrared ADS40 image of a 10 square mile portion (less than 1%) of the project area in Redfish Bay.

Figure 4b. An illustration of the detail of the final benthic habitat map for the same area.

see great promise for this method in future operational benthic mapping projects."

Conclusions

This project has shown that automated image classification techniques can be used to successfully produce detailed coastal habitat maps from a combination of digital airborne imagery and ancillary data layers. Accomplishments of this project include:

- The demonstrated use of digital camera imagery to map benthic habitats illustrates the enormous potential utility of digital imagery. Considering that NAIP digital imagery is available in the public domain, people will rapidly discover more uses for the data.
- Utilization of automated segmentation algorithms to delineate polygons, coupled with CART analysis to develop rules for labeling polygons, brought a level of detail and consistency to the final product that has not been obtainable using manual photo interpretation. While manual interpretation can produce highly accurate benthic maps, inconsistency is inevitable because of the variability of human interpretation: from individual to individual, and from day to day. Use of segmentation algorithms and CART rules imposes consistency on polygon delineation and labeling.
- Comparative assessment of the accuracy of CART inputs and rules allowed us to push the automated techniques to the highest accuracy obtainable, thereby preserving valuable analyst time for focusing on confusion that the automated techniques could not separate. Analyst time is expensive and indispensable to any mapping project. Using automated techniques to map the uncomplicated areas of the project allows us to concentrate analyst time on more difficult areas -- i.e. where the correlation between the habitat class and the imagery and ancillary data is low or unknown.

Table 2. The contractor's final accuracy assessment matrix. Overall accuracy is 90%. The matrix shows both deterministic and fuzzy accuracies. Numbers in the green Ð
be considered acÐ
in the matrix. After production of the initial map, it was determined that all algae in the project area are drift algae. Therefore, all algae sights were changed to unconsolidated sediments with an algae modifier.

				REFERENCE DATA						*User's Accuracy*	
	Combined Matrix	Bivalve Reef	SRV	Land	Mangroves	Emergent Marsh	Unconsolidated Sediments	Unknown Benthic Habitat	Total Map	Deterministic	Fuzzy
	Bivalve Reef	99				1	1		99 2	100%	100%
	SRV	7	144			8	20 33	5	164 53	66%	76%
MAP DATA	Land			106	3	24	1		130 4	79%	97%
	Mangroves				97				97 0	100%	100%
	Emergent Marsh					55			55 0	100%	100%
	Unconsolidated Sediments	3	2 3	1		5	123		125 12	90%	90%
	Unknown Benthic Habitat	7	1	1			6	104	105 14	99%	88%
	Total Reference	116	150	108	100	93	184	109	860		
Producer's	Deterministic	85%	96%	98%	97%	59%	67%	95%		85%	
Accuracy	Fuzzy	85%	97%	99%	97%	85%	78%	95%			90%

References

Breiman, L., J.H. Friedman, R.A. Olshen, C.J. Stone, 1984. *Classification and Regression Trees,* Chapman & Hall, International Thomson Publishing, 358pp.

Congalton, Russell, and Green, Kass, 1999. *Assessing the Accuracy of Remotely Sensed Data, Principles and Practices,* Lewis Publishers, 137 pg.

Green, Kass and Russell Congalton, 2004. An error matrix approach to fuzzy accuracy assessment: the NIMA GeoCover project, A peer-reviewed chapter IN: Lunetta, R.S., and J.G. Lyon (Eds.), *Remote Sensing and GIS Accuracy Assessment,* CRC Press, Boca Raton, FL.

Madley, K.A., B. Sargent, and F.J. Sargent. Development of a System for Classification of Habitats in Estuarine and Marine Environments (SCHEME) for Florida. 2002. Unpublished report to the U.S. Environmental Protection Agency, Gulf of Mexico Program (Grant Assistance Agreement MX-97408100). Florida Marine Research Institute, Florida Fish and Wildlife Conservation Commission, St. Petersburg. 43pp.

Acknowledgements

The authors wish to thank Bill Stevenson and Mark Finkbiener of the NOAA Coastal Services Center, Jim Simons of the Coastal Fisheries Division, Texas Parks and Wildlife, John Wood of the Harte Research Institute at Texas A&M University-Corpus Christi, Harold Remple of Fugro EarthData, Mark Tukman of Tukman Geospatial, and Dan Bubser, Joe Muller, Keith Patterson of Avineon, Inc. for their innumerable and invaluable contributions to the success of this project.

Authors

Kass Green
President
The Alta Vista Company
kassgreen@earthlink.net

Chad Lopez
Digital Imaging Analyst
Fugro EarthData, Inc.
clopez@earthdata.com