Detection of Small Targets in Ocean Wave Clutter Using Panchromatic Time Series Imagery

J.P. Dugan, C.C. Piotrowski and D.C. Campion
Areté Associates
1725 Jeff Davis Hwy, Suite 703
Arlington, VA, 22202, USA
dugan@arete-dc.com

***Abstract*- Visual detection of small objects on or near the ocean surface from an aircraft is important for search and rescue, and detection of stealthy boats and other militarily important targets. This is made difficult by a combination of the rapidly moving sensor platform and the typically high level of modulations in scattered light at the surface. To achieve reasonable search rates, the observer must scan through modest grazing angles (i.e., look towards the horizon) where the ocean clutter that limits detection performance is varying light from wave facets and white caps. This paper describes a technique for significantly improving the signal-to-clutter ratio. A time series of images is mapped to a common reference frame at the level of the mean ocean surface, separating time and space variations, and the temporal dwell is utilized. For moored targets, the signal may be simply integrated up in time, while for drifting and powered targets, it must be integrated along the appropriate target velocity vector. Specific examples are provided for navigation buoys, small surface floats and slowly moving boats, most of which are difficult if not impossible to find in any single image frame. Achieved gain using panchromatic images is shown to approach 20 dB in cases where the velocity vector is known or can be calculated from the data.**

I. INTRODUCTION

It is nearly impossible to detect small objects on or near the ocean surface with the unassisted eye from an aircraft at sea because of the high level of clutter from the waves and whitecaps. This problem severely limits the success of many airborne missions such as search and rescue and detection of small buoys, surfaced or near-surface mines, and other navigational hazards. For example, unless accurately cued, a trained, non-fatigued human is incapable of finding these objects by continuous observations from an aircraft window. However, this problem potentially can be reduced considerably because of the simple fact that the target of interest appears at more or less the same location over the observation dwell, while the ambient clutter is a combination of randomly located whitecaps with relatively short lifetimes and wave-related clutter that moves along with the waves. Thus, a relatively straightforward approach to improving the detectability of small targets of interest is to collect a number of images of the target area and co-register them on the surface in order to obtain gain from coherent processing that reduces the ocean clutter relative to the target signature. This approach is developed in this paper, and is shown to provide significant gain in panchromatic images of the ocean collected from a digital camera mounted on a small aircraft.

The data collection system is AROSS, which stands for Airborne Remote Optical Spotlight System, where the word *spotlight* calls attention to its ability to continuously *stare* at the target area independently of the aircraft motions [1]. The camera is continuously trained to the area of the target, and inertial navigation metadata are used to transfer a time series of sequential images to a geodetic coordinate system at the level of the mean ocean surface [2]. The resulting accuracy of this mapping has considerable remaining jitter, which reduces the effectiveness of this approach if too large. This jitter can be reduced for nearshore applications by using stationary control points (tie points) on land to re-calculate the camera orientation for each image and re-registering them at the surface level [3]. These mapped data sets have recently been used to retrieve maps of water depths and surface currents [4, 5] as well as accurately locating military targets of interest [6, 7].

In this detection problem, the object is to find the target in the temporally and spatially varying clutter, and these modulations are so high that one should not be limited by sensor receiver noise. Since one typically must look through some level of marine haze, neither the target nor the clutter has a significant color signature due to absorption and path radiance of the intervening atmosphere. Thus, a panchromatic sensor suffers little disadvantage (and may actually have a modest advantage) over a color-capable sensor, since it collects all the available photons in each pixel and does not split them into the extra dimension of multiple frequency bins, which potentially only increases the receiver noise level. In the following, the panchromatic time series image data from AROSS flights at several experiment sites are used to examine the advantage of temporal gain using multiple looks at areas having subtle targets of interest, much as previously accomplished using a long-range airborne IR imager [8, 9]. Several types of targets are examined, including small surface floats, larger buoys, and a slowly moving jet ski. The gain is evaluated as a function of dwell and resolution on the target.

II. DATA, ALGORITHMS, AND RESULTS

Fig. 1 is one of a series of images of the nearshore ocean collected at the U.S. Army Corps of Engineers Field Research Facility (FRF) located on the Outer Banks at

Funding for this work was provided by the Office of Naval Research through the Littoral Remote Sensing Program.

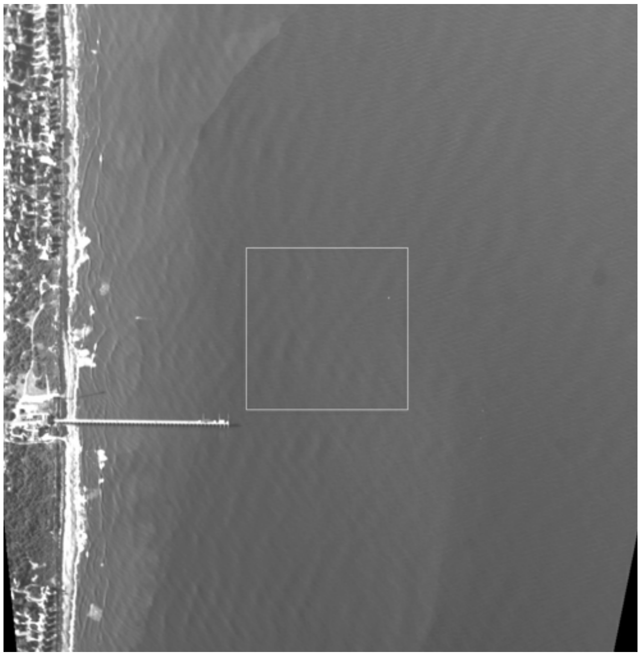

Fig. 1. Mapped image of the nearshore at FRF. Image is about 2 km on a side.

Duck, North Carolina. This location is typical of quasi-linear barrier islands. This data set was collected as part of an experiment called SHOWEX, which stands for SHOaling Wave EXperiment. The aircraft was flying around the target at an altitude of about 10,000 feet and a standoff range of about 6 km, providing a grazing angle at the surface of about 25 degrees. The camera has a one-megapixel array and a 2:1 anamorphic lens, which squares up the mapped image field of view (FOV) and the pixels on the surface. Thus, the mapped image in Fig. 1 is almost square. The effective pixel size on the surface is about 2 m on a side, and the FOV is about 2 km. The data have a dynamic range of about 3500:1, which is almost a full 12 bits, and the water appears very dark in this image because the very bright beach has been retained on the range of the digitizer. There are very few whitecaps in this case. Fig. 2 is the data in the outlined area in Fig. 1, rescaled to better show the waves and more clearly exhibit the targets and clutter outside the surf.

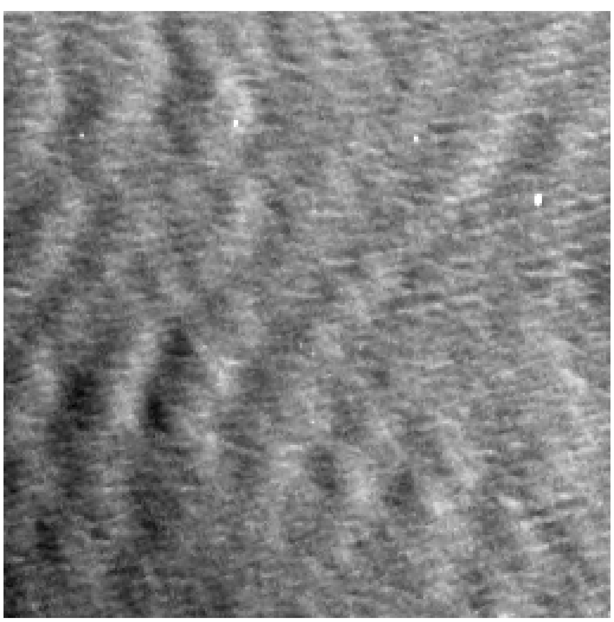

Fig. 2. Expanded scale for a 512 m by 512 m portion of Fig. 1 with multiple targets NE of the seaward end of the pier.

The clutter is comprised of wave slopes, which are mostly reflected skylight, and a few whitecaps, which are mostly scattered sunlight. These data were mapped as described in [3], resulting in a temporal time series of the radiance modulations at each 2 m pixel. 100 frames were stacked in time, with the average result shown in Fig. 3a. This exhibits a large number of small targets that were not seen even by careful examination in many of the individual images. Some were occasionally observable by eye in an animation constructed from the mapped frames over this 2-minute dwell. These objects have been identified as a number of small (~20 cm long) surface floats and several larger guard buoys. This figure exhibits considerable low frequency clutter, which may be reduced by applying a matched filter in addition to stacking. A matched filter for unresolved targets was applied to the individual frames, using a 2D Gaussian signal hypothesis with σ = 1.25 m (FWHM ~ 3 m). The frames were normalized by the background mean, convolved with the signal hypothesis, and the result normalized by the standard deviation of the filtered background region. Figure 3b is a 100-frame stack of the resulting matched filter output. The larger noise peaks are the result of whitecaps and some alongshore-oriented streaks.

Fig. 4 is a plot of the effective signal-to-clutter ratio (SCR) as a function of stacking time for several of the surface floats in the 50-second dwell (100 frames at 2 Hz). The average achieved gain is about 18 dB, which is close to the $\sqrt{N}$ gain expected from independent samples, as shown by the dashed line.

The second example is a larger surface buoy, and it was chosen because the data have considerably more and larger whitecaps. The target is a 1 m wave buoy that was further offshore of FRF during SHOWEX. It is near the center of a single frame shown in the top panel in Fig. 5. Since the images were collected much further at sea, and land was not in them, they were mapped from the inertial navigation system data alone, with some minor corrections to reduce jitter by using a frame-to-frame correlator keying on the location of the bright target. The average of 100 frames is shown in the middle panel in Fig. 5, and the stack of 100 frames of matched filter output is displayed in the bottom panel, showing a gain of 14 dB over the single frame SCR. This is significantly less than the previous case, presumably because the noise due to the whitecaps is partially correlated from frame to frame.

A third example is a jet ski that was conducting a bathymetry and current survey to serve as *in situ* truth for hydrographic retrievals using these same data sets [4, 5]. The data were collected over Oregon Inlet, which is the first

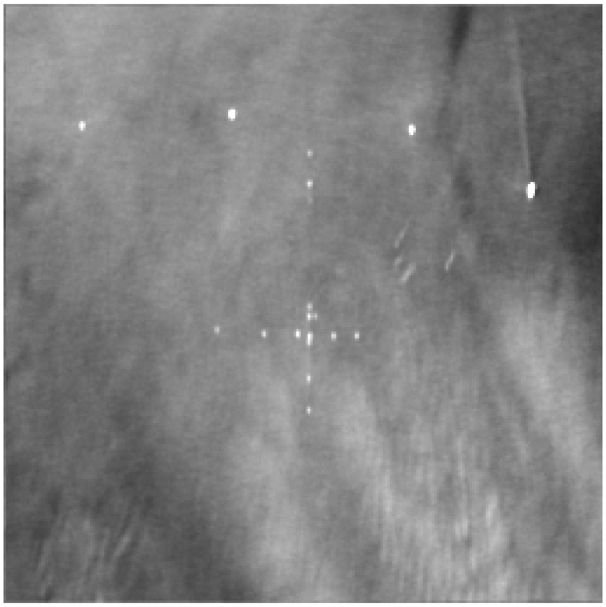

Fig. 3a. 100-frame stack of subpatch in white square in Fig. 1.

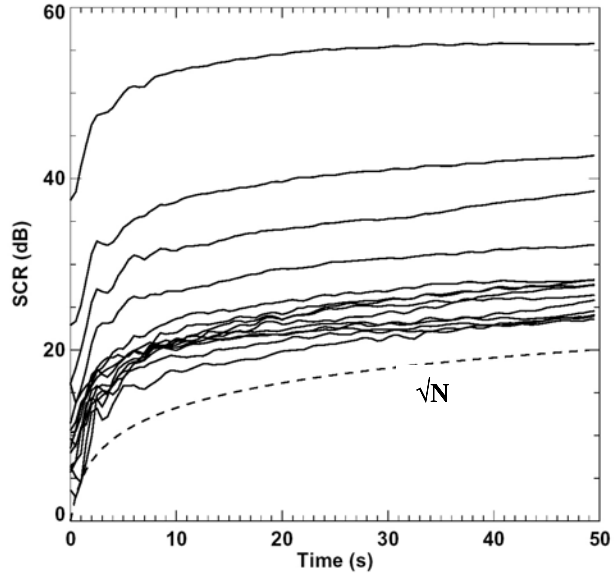

Fig. 4. SCR vs. stacking time (frame rate of 2 Hz) for various targets in Fig. 3. Dashed curve shows expected gain for N independent samples.

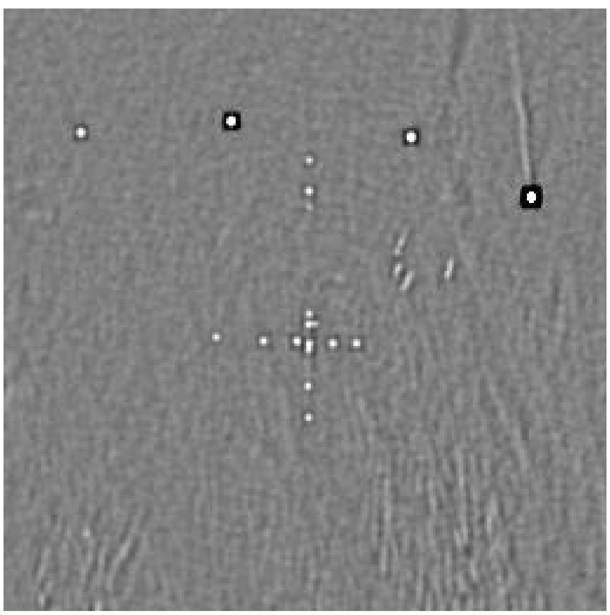

Fig. 3b. Matched filter in addition to stacking.

inlet south of FRF. A single image is shown in the top panel in Fig. 6. Again, the mapped image is about 2 km across and the spatial resolution is about 2 m. The boat may be seen in a single frame, and it is more obvious in a multi-frame animation, though not nearly as obvious as a number of larger, faster fishing boats, due to their bright, white-water wakes. In this case, the velocity vector of the boat was known, so the signal was integrated along this vector. The result for a 10-frame stack of the data in the center of the inlet (the white square) is shown in the lower left panel, and the matched filtered result in the lower right panel. The SCR is improved by about 10 dB in this case, which is the expected $\sqrt{N}$ gain. Only 10 frames were used in this example because the jet ski did not travel in a straight line for very long.

III. DISCUSSION

These examples serve to exhibit significant gain in using a sequence of mapped image frames for detecting small, stealthy targets in the wave clutter of the sea. The specific results should only be considered as representative, as larger or smaller values are achieved under different environments, viewing conditions, and dwell time. For example, the very small buoys are not well resolved in these images, so their signal level is spread out more than is necessary, thereby significantly decreasing the achievable gain from the potential maximum value. For example, the target energy is spread over pixel sizes of 2 m on a side. For targets less than 1 m in size, for example, this reduces the achievable detectability by 12 dB. In fact, we have occasionally collected data having 1 m resolution, mapped the images to 1 m pixels on the water, and have achieved over 30 dB of SCR from stacking. However, in general, there are limits to the resolution that can be obtained on very small targets, since they are not stationary but rather move around as a result of the wave orbitals. Also, our mapping accuracy is not always of the quality to isolate the target locations to within 1 m because of residual jitter and/or inertial navigation drift. Thus, we prefer the more conservative spatial resolution of ~2 m, thereby achieving the gain values quoted. Also, we used the velocity vector of the moving target to integrate up the signal against the

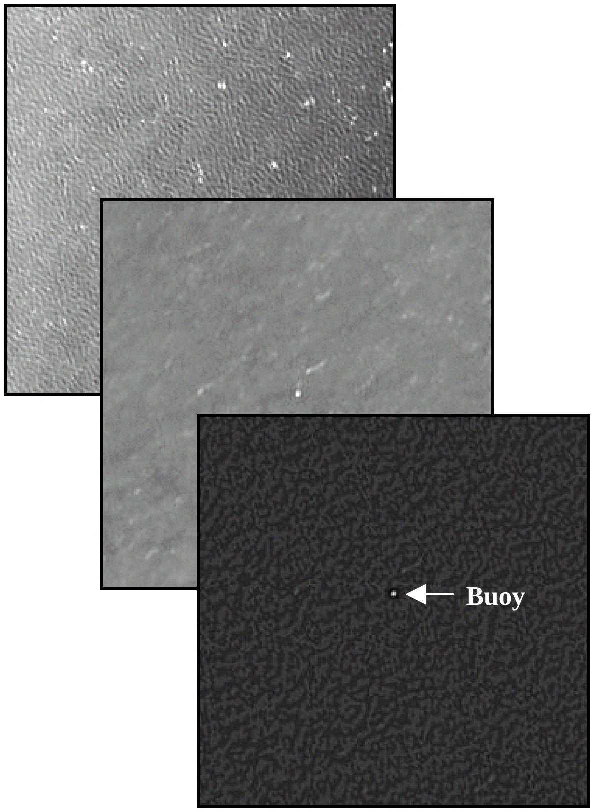

Fig. 5. Wave buoy (near center of frame) in whitecaps. Upper panel is raw image, middle is 100-frame stack, and bottom panel is match-filtered stack.

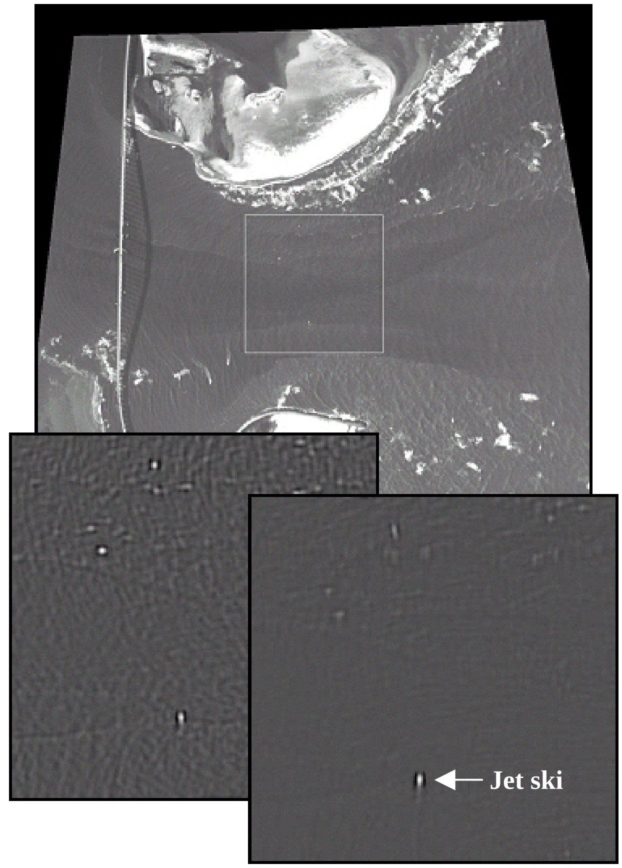

Fig. 6. Jet ski in Oregon Inlet. Mapped image in top panel, 100-frame stack on lower left, match-filtered stack on lower right.

noise. We knew the velocity in this example, so were able to exhibit the potential gain, but in reality, this would not be known for a really stealthy target and the extra information would have to come from the data. This can be done, though it has not been in this study. Obviously, the achievable gain is directly dependent upon the accuracy of the velocity vector.

Thus, the detection performance decreases with uncertainty in the target velocity vector, decreased spatial resolution on very small targets, or the accuracy of the mapping. For freely drifting targets, the velocity vector can be estimated directly from the data by fitting the theoretical dispersion relation for gravity waves to 3-D frequency-wavenumber spectra of 3-D data cubes [5,10]. Further analyses of performance dependence upon these various parameters is of interest since freely drifting naval mines have been a problem for U.S. forces in the past, and this problem would be an important application of this technology.

The examples all used data that were collected in a staring mode, which is fine for a small search area. However, the approach can be modified to be just as effective for a significantly larger search area, as the camera FOV can be dragged along at the speed of the aircraft. In this case, the search area is a direct function of the size of the target of interest and the number of pixels in whatever camera is used. These, along with the aircraft speed and the required dwell on any one location on the surface, determine the effective search rate. For ~1 m size targets, aircraft speed of 100 m/s, and a commercially available 16 megapixel camera, the search rate could be as large as 5000 km^2/hr, while still enabling more than one minute of dwell on each potential target. This is a respectable search rate.

The potential utility of spectral data remains to be examined. Multi-spectral cameras are available and, if the received photons are not split into too many color bins, as per our previous comments, there is expected to be some additional detection capability. On the other hand, a much larger number of color bins such as in a hyperspectral camera are not expected to be beneficial, even if it were implemented in a focal plane array imager to achieve the temporal gain, since the photon noise will be a limiting factor. Thus, this is not a very promising direction of future enquiry.

Also, the stacking did not perform as well for the second example having the large number of whitecaps as it did for the other two cases. An alternative to performing a simple stack of the matched filtered, registered images is to form a likelihood ratio and combine the images in the likelihood sense. A key advantage to this approach is the ability to better model the background noise statistics. In the cases where there are a significant number of whitecaps, the background noise statistics can no longer be accurately modeled by a simple Gaussian noise distribution, which is inherent in the simple stacking approach. However, our colleagues [11] have shown that the likelihood ratio can be formed with the background noise statistics modeled using a two-Gaussian noise model. The two-Gaussian noise model presumes that a Gaussian distribution models most of the data probability distribution. However, a second Gaussian distribution that models large, but rare, outliers is added to the primary distribution. The simple addition of this second distribution for the outliers prevents a single large value in the data from "capturing" the likelihood function. This approach allows small but persistent signals to more effectively build likelihood through the image stacking process while the large but transient signals from whitecaps only temporarily build likelihood, but are quickly suppressed as each white cap dissipates.

This approach is being used in our current line of investigation. Initial results are very encouraging, and these will be documented in a future paper. In addition, we now are flying a multi-color version of AROSS, and it will be collecting ocean data this next month during the Nearshore Canyon Experiment (NCEX) where there will be many targets of opportunity, just as we found at SHOWEX.

IV. CONCLUSIONS

We have shown that the temporal domain provides significant gain for the detection of stealthy targets embedded in imagery of the ocean surface, as expected. As much as 20 dB gain has been achieved for the small surface floats, and the performance for the other target types was respectable as well. This improvement is roughly proportional to the square root of the number of looks at the target for simple integration. The accuracy of the mapping, the resolution on the water relative to target size, and the accuracy of the target motion vector all affect the performance, although we have not exhaustively evaluated parametric results to date. Because of the rather inexpensive investment and small required real estate on the aircraft involved in turreted, stare-capable EO sensor technology, the above-mentioned missions all can be conducted effectively, even if not at night nor through cloud cover. The cloud cover is not a fundamental issue for some missions, as small aircraft such as naval organic UAVs that are envisioned for these missions (e.g., [12, 13]) can fly under them and only suffer a concomitant decrease in effective search rate. Utilization of an IR sensor could be effective for 24-hour operations, but this has not been pursued beyond some early results [8, 9].

ACKNOWLEDGMENTS

We acknowledge the diligent work of many other Areté Associates scientists that assisted in the collection and analysis of these data. We also greatly appreciate the fine flight support provided by the NPS CIRPAS and the valuable ground support and ocean *truth* data provided by the Corps of Engineers Field Research Facility personnel.

REFERENCES

[1] J.P. Dugan, G.J. Fetzer, et al. "Airborne optical spotlight system for remote sensing of ocean waves," *J. Atmos. and Oceanic Technol*, vol. 18, no. 7, pp. 1267-1276, 2001.

[2] J.P. Dugan, A.G. Evans, C.C. Piotrowski, and U.V. Bhapkar, "Accurate navigation of airborne image sequences for rapid surveys of water depths and currents," *MTS/IEEE OCEANS '01 Conf.*, Honolulu, Hawaii, pp. 559-564, 5-8 Nov 2001.

[3] D.C. Campion, J.P. Dugan, C.C. Piotrowski, and A.G. Evans, "Direct geo-referencing techniques for rapid positioning of targets and environmental products using tactical grade airborne imaging data," *MTS/IEEE OCEANS '02 Conf.*, Biloxi, MS, pp. 1603-1608, 29-31 Oct 2002.

[4] J.P. Dugan, C.C. Piotrowski, and J.Z. Williams, "Rapid environmental assessment of nearshore METOC fields using motion imaging techniques applied to surrogate UAV data," *MTS/IEEE OCEANS '02 Conf*, Biloxi, MS, pp. 1956-1961, 29-31 Oct 2002.

[5] J.P. Dugan and C.C. Piotrowski, "Surface current measurements using visible image time series," *Remote Sensing of Environment*, vol. 84, pp. 309-319, 2003.

[6] J.P. Dugan and S.C. Lubard, "Airborne determination of the undersea environment in the littoral zone with a spotlight optical camera system," *Joint Undersea Warfare Technology Conf.*, Monterey, CA, 18-20 Mar 2003.

[7] J.P. Dugan and D.J. Wurzel, "Exploiting UAV imagery: Advanced techniques promise significant benefits for littoral warfare," *ISR Journal*, vol. 1, no. 4, pp. 36-39, 2002.

[8] H.H. Suzukawa, Jr. and M.S. Farber, "Long range airborne detection of small floating objects." *SPIE Conf. on Detection Technologies for Mines and Minelike Targets, AeroSense'95*, vol. 2496, Orlando, FL, pp. 193-204, 17-19 Apr 1995.

[9] P.J. Davis, E.P. Krumrey, and R. Myers, "Simulation and data analysis of infrared ocean clutter at grazing angles," *SPIE Conf. on Characterization, Propagation, and Simulation of Sources and Backgrounds III*, Orlando, FL, pp. 355-367, Apr 1993.

[10] C.C. Piotrowski and J.P. Dugan, "Accuracy of bathymetry and current retrievals from airborne optical time series imaging of shoaling waves," *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 12, pp. 2606-2618, 2002.

[11] R. Lindgren and L. Taylor, "Bayesian field tracking", *SPIE Conf. on Signal and Data Processing of Small Targets*, Orlando, FL, vol. 1954, pp. 292-303, 1993.

[12] J.P. Dugan and C.C. Piotrowski, "Developmental system for maritime rapid environmental assessment using UAVs," *Oceanology International 2000, Session on Rapid Environmental Assessment*, Brighton, UK, pp. 201-211, 7-10 Mar 2000.

[13] J.P. Dugan, "UAV-based real time data for coastal ocean forecasting systems," *Nearshore Prediction*

Workshop, SACLantCentre, La Spezia, Italy, 15-17 Feb 2000.