Object Detection Using A Background Anomaly Approach For Electro-Optic Identification Sensors (February 2002)

ANDREW NEVIS
JAMES BRYAN
J. S. TAYLOR
BRETT CORDES

Naval Surface Warfare Center Coastal Systems Station, Dahlgren Division
Panama City, FL 32407-7001 USA
(Ph: 850-235-5958, e-mail: nevisaj@ncsc.navy.mil)

I. INTRODUCTION

Electro-optic identification (EOID) sensors[1,2] are transitioning to the fleet and will be used as a short-range identification tool for mine-like contacts from long-range sensors. The present operation of the EOID sensors uses an operator for identification. Whereas the human operator is unparalleled in detecting and recognizing objects of interest, there are still some limitations which may be needed to distinguish between mine types, such as differentiating a 68 inch object from a 72 inch object in a still image or moving waterfall display. To help overcome some of these weaknesses and improve the mine identification process, computer aided identification (CAI) and automatic target recognition (ATR) algorithms are being developed[1]. In addition to building a foundation towards the long-term goal of fully autonomous operation, these algorithms can be used to queue operators of potential mine-like objects within the data as well as to segment and compute vital geometric information on manually flagged objects of interest. The operator can then use this supplementary information for a more accurate identification. The near-term objective is to develop and implement these CAI/ATR algorithms into a real-time console and/or a post mission analysis (PMA) tool that can be used in the FY05 Organic Mine Warfare future naval capability (FNC) demonstration[2].

II. ANOMOLY DETECTION

Due to the highly variable turbid nature of coastal waters, coupled with sometimes heavily cluttered environments and a vast array of sea bottom types (sandy, rocky, muddy, coral reef, etc), underwater electro-optic object detection provides a formidable challenge. To overcome these obstacles, a background anomaly approach has been chosen. This approach is designed to be effective in cluttered and non-uniform background environments, yet makes no assumption on turbidity (sharp edges versus blurry outlines) or on the sea bottom type (other than there is some differentiation between an object and its local background).

[1] Three efforts (CSS, Northrop Grumman Ocean Systems, and Raytheon Electronic Systems) have been funded by the Office of Naval Research, code 322-OP (Dr. Steve Ackleson) under the EOID research program

[2] For more information see http://www.onr.navy.mil/sci_tech/ocean/MCM/

This approach searches for anomalies of a certain specified size and shape that sticks out from the local background. The first step produces an anomaly image using background strips that essentially acts as a filter to remove large objects, although small objects with high SNR may pass through. The next step is to convolve the anomaly image with shape filters. The output image from the convolution is then tested for peak values to obtain candidate detections. After this, the candidate detections are segmented and have basic geometric features computed. To reduce the number of detections passed to a classifier, small and/or irregularly shaped objects (e.g., length to width ratio is extremely skewed) are filtered out to obtain a list of final detections. These key steps are discussed in more detail in the following sections.

A. Anomaly Image

The anomaly image is generated by rotating least squared error (LSE) background strips of specified lengths, excluding the region of interest, computing and testing for the best background fit with minimum error as the criteria. Using the line orientation corresponding to the best background fit, a subsequent LSE fit is next extended across the region of interest, where the background anomaly is defined as the difference between the fitted line and the actual values on a pixel-by-pixel basis throughout the region of interest. These concepts are illustrated in Figure1. This process is applied on a pixel-by-pixel basis throughout the entire image to fill in the anomaly image.

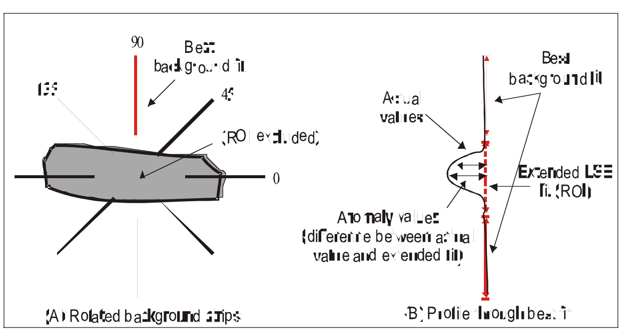

Figure 1. (A) Illustration of the best background fit. The background segments that include the target boundary (0,45,135) will have higher LSE errors, whereas the line segment that includes background only (90) will have minimal error yielding the best background fit. (B) Looking at a profile through the best background line fit, the anomaly values are illustrated as the difference between the extended LSE fit over the region of interest (ROI) and the actual values.

After the anomaly image has been generated it is convolved with shape filters. These shape filters include rectangular strips (rotated in four orientations) for elongated targets with lengths and widths commensurate with the background strip lengths, and circular regions for near circular targets with diameters corresponding to the background strip diameters. Detection is then achieved by searching for peak values in the output image from the shape-filtered convolution. Figure 2 shows an example of an image with a non-uniform background and its corresponding anomaly image, shape-filtered convolved output image, and peak detection.

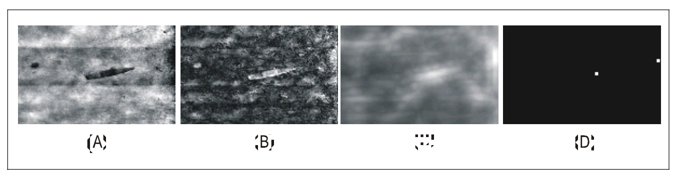

Figure 2. (A) Original image, (B) Corresponding anomaly image, (C) Output image from shape-filter convolution, and (D) Peak values corresponding to detection.

B. Segmentation

After detection, the pixel location of the peak value along with the corresponding orientation is used to generate a background mask, as shown in Figure 3A.

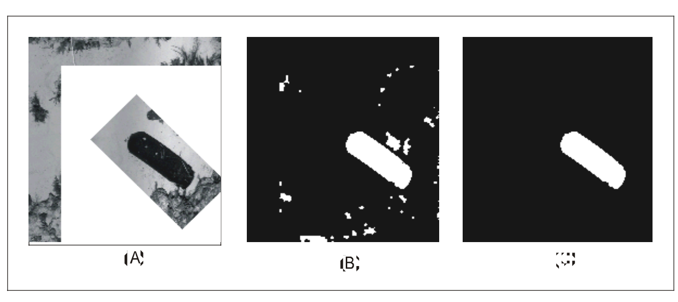

Figure 3. (A) The background mask corresponding

to the centroid and orientation from detection, (B) The region is thresholded into a binary region, and (C) The object is enumerated with extraneous detections removed.

A two-dimensional (2-D) background fit over the background region (excluding the target region) is computed by 1-D LSE line fit, first down the image columns and then across the image rows. This produces a 2-D near planar surface estimate of the local background that extends across the target region. The entire region (background and target regions) is then thresholded into a binary region containing either background pixels (black) or target pixels (white), as shown in Figure 3B. The binary region is computed by declaring actual pixel values that are sufficiently away from the corresponding background fit as target pixels whereas actual pixel values that are sufficiently close are declared background pixels. Customized morphology filters are then implemented across the binary region to remove noise pixels and to fill small interior holes. The detected object connected to the location of the peak value then has its outer boundary pixels enumerated, with all other objects removed as extraneous detections, as shown in Figure 3C.

C. Geometric Features

After segmenting and enumerating the peripheral boundary pixels, second order moments are computed to obtain the detected objects centroid and orientation. Whereas the centroid computation is deemed sufficient, orientation and corresponding extracted length and width from second order moments are not considered accurate enough for the intended purpose of identification. Thus a subsequent procedure was developed that uses the second order orientation as an initial guess and adjusts the orientation based on the best fit of the objects sides at its midsection. Using this adjusted orientation, a more accurate length and width can be computed. The objects length is computed by taking the difference between the endpoints of a line parallel to the adjusted orientation through the centroid, as shown in Figure 4.

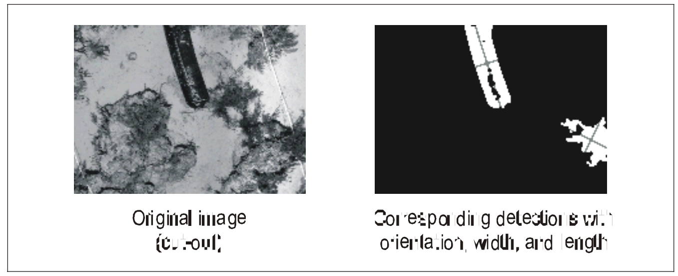

Figure 4. The objects adjusted orientation, along with computed length and width (target and clutter).

The width can be computed in one of two ways. The first width computation (computed width) is the length of the line perpendicular to the adjusted orientation crossing through the centroid, similar to the length computation, which is also seen in Figure 4. This represents the quickest implementation, but may be vulnerable to peripheral artifacts that may happen to be near the objects centroid (such as an indentation), as shown in Figure 5A. The second method (fitted width) is to make a best fit of the objects sides at the midsection (similar to computing the adjusted orientation) and then take the difference between the two sides in a line perpendicular to the adjusted orientation through the centroid, as shown in Figure 5B. However, this fitted width may be vulnerable to circular or elliptical objects that do not have obvious sides that can be easily fitted to, depending on implementation. Regardless of which width computational method is used, the extracted length and width will be used to help determine mine type in the identification process (these computed values will be compared to manually extracted values for accuracy).

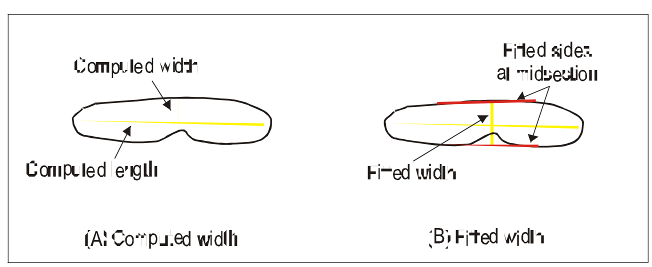

Figure 5. Illustration of (A) The computed width and (B) The fitted width.

D. *Algorithm Implementation*

This section describes the implementation of the background anomaly detection algorithm, using the methods discussed above, as applied to contrast images from EOID sensors (the implementation towards range images will be discussed section III). Figure 6 shows a flow diagram of this implementation.

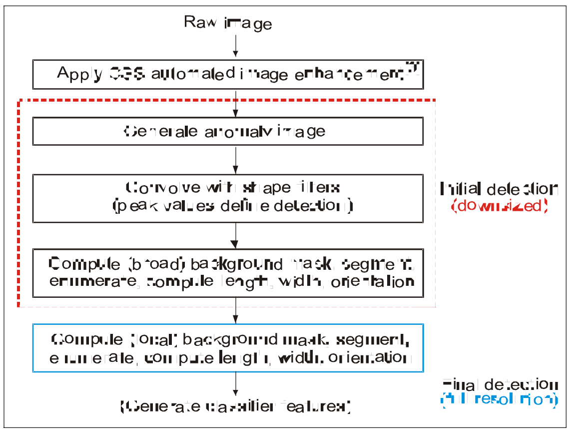

Figure 6. The flow diagram of the background anomaly detection algorithm (implemented on EOID contrast imagery).

Because EOID sensors use such high resolution (needed for identification), the first step in the algorithm is to downsize the image for initial detections. Applying the anomaly detection scheme on downsized imagery allows more flexibility with unknown target shapes and sizes, thus allowing preliminary information (such as length, width, and orientation) to be computed that can in turn be used for a more precise detection and segmentation with the CPU intensive full resolution imagery. This allows more candidate detections to be investigated, with a filter applied to remove small or irregularly shaped detections.

The preliminary information computed during the initial detection on downsized imagery includes object centroid, adjusted orientation, computed length, and computed width. Because of uncertainty of the target characteristics, the background mask used in the initial detection is somewhat broad and away from the target. The process of forming a background mask, thresholding, segmenting and enumerating, and computing target information are all applied in the initial detection. After initial detection, the four computed preliminary parameters (centroid, orientation, length, and width) are used to form a local background mask for final detection and segmentation, where the process of forming a background mask, thresholding, segmenting and enumerating, and computing target parameters is repeated on full resolution imagery. The background mask used in the final detection is more local to the target, and thus more accurate, since target characteristics are now more certain from the preliminary information. Once final detection and segmentation is complete, classifier features can then be computed in preparation for a classifier, as discussed in the next section. Figure 7 shows an example of the initial and final detections.

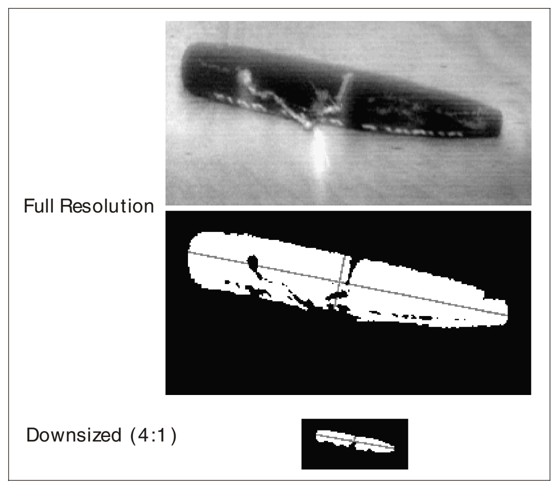

Figure 7. An example of the initial and final detections applied on an image. Note that the final detection appears more accurate than the initial detection since the target characteristics are more certain than in the initial background mask (as well as better resolution).

III. FUTURE EFFORTS/MODIFICATIONS

Section II discussed the current stage of the background anomaly detection algorithm whereas this section describes the (near-term) future efforts. First, the best-fit criteria will be modified allowing the use of shorter background strips. Next, the algorithm will be modified for three-dimensional data. Third, the local background estimate used in conjunction with segmentation will be changed from single line estimates to piecewise overlapping line segments. These three modifications will be discussed in subsections A, B, and C, respectively. Once the detection routine is completed, the next effort will be the development of classifier features for classification, described in subsection D. Subsection E discusses a parallel detection routine that will try to detect the extremely difficult case of bio-fouled targets (near-zero contrast targets), whereas subsection F discusses the development of an identification (classification) scheme using Zernike moments and neural networks.

A. Modified Best Fit Criteria

The key for successful implementation of the background anomaly algorithm is to obtain a best background fit about a targets width orientation, and not along its length orientation (see Figure 1). This can be problematic since man-made objects can often be smooth resulting in a low LSE error for the background strips (where minimum error determines the best background fit). Thus to avoid this condition, the length of the background strips were designed to be longer than the longest expected target (see Figure 1). Not only does this restrict the length of the expected target, but also the longer background strips can extend beyond the local background into nearby clutter, thus reducing the detection performance of this approach. This condition will be circumvented by using a normalizing parameter that indicates the presence of object edges within the region of interest, as shown in Figure 8.

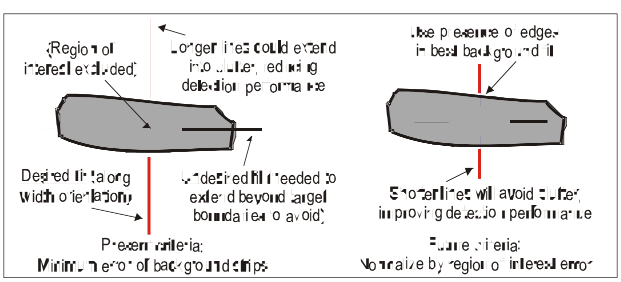

Fig 8. To reduce the length of the background strips (and thus improve detection performance), the criteria for best background fit will be normalized by the LSE error of a Line fit within the ROI.

The best background fit criteria will then be the ratio of LSE errors of the background strip line fit normalized by the LSE error of a line fit within the region of interest. Thus the emphasis of the best background fit will be minimal error on the background strips (indicating the presence of background only in the background strip regions) with maximal error over the region of interest (indicating the presence of object edges within the region of interest). This will allow the use of much smaller background strip lengths closer to the region of interest, removing the effects of nearby clutter and obtaining more accurate local background information, thus yielding improved detection performance.

B. Three Dimensional Detection

Implementation towards EOID STIL[3] 3-D data will be applied in a similar approach to the contrast image (2-D) case. In this case, the 3-D data will be rendered into a 2-D range map and a 2-D contrast image. The same process of detection will then be applied to both the contrast image and the range map (with customization) in parallel, fusing the two separate detection results together as a final step (most likely with an inclusive ORing). Classifier features will then be computed utilizing the range information (3-D features) supplementing the 2-D contrast features, providing a more confident identification.

C. Piecewise background fit

The current background fit in the segmentation process uses 1-D LSE lines, first down the image columns then across the image rows, to produce a 2-D near planar surface. This technique for background estimate may be suitable for initial detection where the background mask is formed somewhat away from the target, but does not accurately represent local nonlinear surfaces near the target. Using overlapping piecewise line segments instead of single line segments will allow nonlinear local variations of the background tangent to the target to be more accurately represented in the

[3] Streak Tube Imaging LIDAR, or STIL, is a 3-D EOID sensor developed by Areté Associates from Tucson, Arizona.

background fit. This in turn will yield a more accurate segmentation of the target from its local background, particularly in the presence of tangent clutter.

D. Classifier Features

The focus after completion of the detection routine will be the development of discriminatory classifier target features. First, features will be developed to discriminate between man-made objects and clutter, such as the measure of convexity shown in Figure 9.

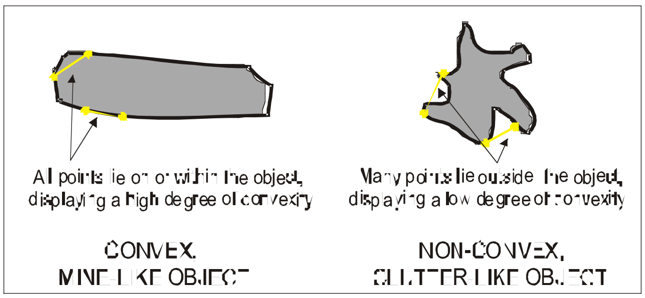

Figure 9. Classifier features will be developed, such as the measure of convexity, to help discriminate between man-made objects and mine-like clutter.

The key to the development of these features is an accurate enumeration (sequence) of the peripheral boundary pixels. These features will be used to reject mine-like clutter (mine-like in size). Next, advanced geometric features (beyond length, width, and area) will be developed that can describe the detected target relative to mine identification. This will include features that describe the targets front-end and back-end (coned shaped versus flat ends), midsection shapes, etc. The key to the development for such features is an accurate orientation that clearly locates the midsections of the front and back ends of the target. The information generated from these geometric features will be compared to a known database of measurements to help determine the identification of a detected mine type. For successful implementation of this approach to identification, it will be necessary to geometrically correct[4] the imagery from rectangular pixels to square pixels before computation of the classifier features.

E. Bio-Fouled Detection

The background anomaly detection scheme searches for objects that can be distinguished from the local background, but will ultimately fail as an objects contrast approaches the contrast of its local background (i.e., the target and background become indistinguishable), a level to be determined by implementation. One case where background anomaly detection will almost certainly fail is for bio-fouled targets. In this case, the target and the local background become covered by some kind of biological or sediment layer causing near zero contrast, making automated detection extremely difficult.

Operators typically detect these targets by faint edges produced from the small difference in altitude between the target and its local background, or by edges produced from the scouring of sediment in the local background at the targets sides. Figure 10 illustrates this in an example of a fresh target versus a bio-fouled target[5]. In an effort to detect bio-fouled targets[6], the background strips and the region of interest will be modified to small widths, searching for the presence of edges instead of the cross-sectional widths of the targets. The output from this will be an edge anomaly image, which will be investigated for the presence of mine-like objects. This approach specifically assumes that the target and its local background have similar contrast, but yield a distinct edge.

[4] Geometric correction here assumes a stable tow-body platform giving negligible roll, pitch, and yaw effects.

[5] This target had been bio-fouled in a controlled underwater environment for two years under the Coastal Benthic Optical Properties (CoBOP) program, sponsored by ONR (Dr. Steve Ackleson, 322-OP). This target has been monitored by the Caribbean Marine Research Center (CMRC) at Lee Stocking Island in the Exuma Cays, Bahamas.

[6] Detection of bio-fouled targets for 3-D STIL data has supplemental range information, which is invariant to bio-fouling effects. Thus the bio-fouled target detection scheme is limited to contrast images only.

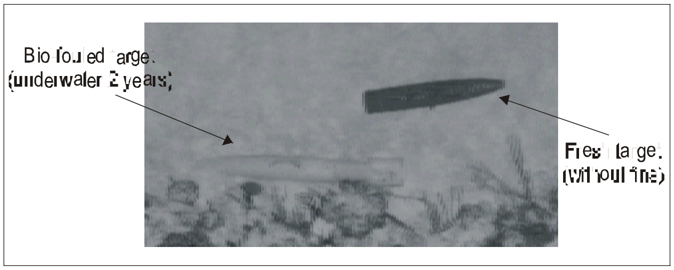

Figure 10. Bio-fouled versus fresh effects on identical targets (including identical paint).

F. Identification

An identification system[4] is being developed based on Zernike Moments and neural networks. This scheme uses Zernike Moments to extract an orthogonal feature set that provides a set of robust features invariant to object rotation, translation and scaling. Zernike moments have been shown to be substantially less sensitive to additive noise than other moments[5]. The target identification system will use shape-dependent extracted features from Zernike moments, along with other target features, in a two-layer back-propagating neural network (BPNN) classification system with 18 inputs, 30 hidden layer neurons and 3 output neurons. Preliminary work using synthetic data has shown promising results.

IV. SUMMARY

A background anomaly approach for object detection has been developed for electro-optic identification (EOID) sensors, which is designed for cluttered and non-uniform backgrounds but makes no assumptions on turbidity conditions or background terrain. This scheme uses best-fit background strips to filter background anomalies that are similar in size and shape to mine-like objects. Once detected, candidate objects are segmented using a two-dimensional (2-D) LSE fit to the local background, followed by geometric feature computation in preparation for classification. The current detection scheme will be modified to include testing for the presence of edges within the region of interest in order to reduce the length of the background strips, thus improving performance.

Future efforts also include the development of advanced geometric features (beyond length, width, and orientation) that can be used as a discriminator in identifying mine types. Other classifier features will also be developed, such as the measure of convexity, which can discriminate between man-made objects and mine-like clutter. A parallel detection effort for bio-fouled targets will modify the background anomaly detection by using shorter background and region of interest strips to search for object edges instead of their cross-sectional widths. The output from the background anomaly detection routine will then be used with a target identification system that has been developed using Zernike moments and a back-propagating neural network classification system.

V. ACKNOWLEGMENTS

This work was funded under the Electro-optic Research Program at the Coastal Systems Station (CSS), a five-year initiative sponsored by the Office of Naval Research (Dr. Steve Ackleson, code 322-OP). The authors would like to thank the CSS program managers, Dr. Sam Taylor and Ms. Mary Hulgan, and others for their steadfast unconditional support.

VI. REFERENCES

[1] Michael P. Strand, “Underwater Electro-optic System for Mine Identification”, *Proceedings for Autonomous Vehicles in Mine Counter Measures Symposium*, Monterey, CA, April, 1995

[2] Anthony Gleckler, Et Al, “Streak Tube Imaging Lidar for Electro-optic Identification”, *Fourth International Symposium on Technology and the Mine Problem*, Monterey, CA, March, 2000

[3] A. J. Nevis, “Image Enhancement for Mine Identification with Laser Line Scan Sensors”, *Third International Symposium on Technology and the Mine Problem*, Monterey, CA, April, 1998

[4] M. R. Azimi, Et Al, “Target Identification using Zernike Moments and Neural

Networks", *SPIE proceedings for Automatic Target Recognition*, Orlando, FL, April, 2001
[5] C. H. Teh and R. T. Chin, "On Image Analysis by the Methods of Moments", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 496-513, July, 1990