FUSION OF CONTACTS IN SYNTHETIC APERTURE SONAR IMAGERY USING PERFORMANCE ESTIMATES

Vincent Myers † NATO Undersea Research Centre, La Spezia, Italy.
Øivind Midtgaard Norwegian Defence Research Establishment (FFI), Kjeller, Norway.

1 INTRODUCTION

Automatic Target Recognition (ATR) enables mine countermeasures (MCM) using autonomous underwater vehicles (AUVs). ATR makes non-trivial autonomy feasible by providing information which allows making choices about mission parameters in real time and relying less on human intervention. It also supports networking by on-board decision making to make intelligent use of the limited communications bandwidth underwater.

ATR is typically divided into two stages: detection, which is the process of determining the presence of target signatures in a sonar image, and classification, during which a more refined discrimination is performed[1]. Detection is arguably the most vital step in any automatic target recognition (ATR) system. Without a reliable detection method, the system performance degrades as any further discrimination becomes impossible since only detected objects are afforded the chance to be further processed by any subsequent steps. The application of combining multiple detection algorithms is widespread. For instance Casasent and Ye[2] fused the output of two algorithms to detect objects in infrared images by using a logical "and" method. Fusion of multiple detection algorithms has also been applied to the classification of underwater objects using sonar images; for instance, Dobeck [3] fuses the output of three classifiers by way of a Fisher Discriminant which uses the normalized confidences of the individual detectors as well as the possible combinations of detectors. Aridgides [4] fuses the same three algorithms using a number of logical operators as well as a log likelihood ratio test. Reed et al.[5] applied Dempster-Shafer theory to the fusion of multiple views of the same object in order to increase the overall classification performance; similar methods were also proposed earlier by Stage and Zerr [6]. Fawcett [7] implemented several simple detectors and fused their outputs by using a kernel-based regression to effectively combine them into a single detector.

This study concentrates on the detection stage of the ATR process, with particular emphasis on the combination of multiple detection algorithms. A through-the-sensor method for assessing detector performance is first presented. The evaluation is done by simulating synthetic aperture sonar (SAS) target signatures and embedding them with the real mission data. Sensor performance is assessed using a sonar performance prediction tool and signatures are generated using a ray tracing method. The detector performance is then assessed on the simulated targets and used for fusing four simple detection methods using Dempster-Shafer theory of evidence, which includes a methodology incorporating information provided by missed detections.

The method is tested on a dataset of SAS imagery obtained during sea experiments conducted by the NATO Undersea Research Center, the Norwegian Defence Research Establishment (FFI) and the Royal Norwegian Navy. The sensor used is the Edgetech 4400 SAS mounted on the HUGIN 1000 autonomous underwater vehicle (AUV)[8] which was used in two separate mission areas in the Bay of La Spezia where several realistic dummy mine targets were deployed.

2 DETECTORS

Four detectors are used in this study and are very similar to those described by Fawcett [7]. All of these detectors operate by computing statistics in a window based on a template of a generic target

† The author is currently with Defence Research and Development Canada – Atlantic, Halifax, Canada.

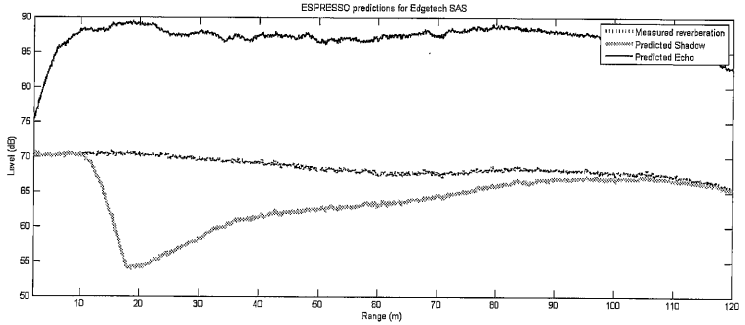

Figure 1. ESPRESSO predictions for highlights, shadows and backgrounds for the Edgetech sensor in approx. 30 m water depth and 14 m altitude. The seabed was composed of silty mud and the wind speed was low.

3 FAST SIMULATION OF SAS TARGETS

The fusion method developed in this paper requires that the performance of the detector be either known or somehow estimated. To arrive at such an estimate, a strategy involving the simulation of target signatures and inserting them into the mission sonar imagery is applied[11,12]. The performance of the detectors can then be evaluated in a controlled manner with the ground truth at hand. The method used to simulate the required target signatures uses ray tracing combined with a sonar performance prediction tool to generate a probabilistic representation of a target, which is then blended into the mission imagery.

High-frequency target modelling has been a topic of research for many years[13,14,15]. For the methodology described here to be useful, the statistical determination of detector performance must be done on-line, which therefore requires that signatures be generated quickly. Less importance is put on modelling seabed reverberation, for instance, since the mission data will be used as background texture. This fusion of simulated targets with mission data has been proposed by Coiras et al.[16,17] who developed a method for adding target signatures to sonar imagery using an inversion technique. Putney et al. [18] employed a physics-based approach to the modelling by using acoustic ray tracing, resulting in greater realism for the target signatures. The targets capable of being modelled are required to have a mathematical description (cylinders, spheres). The method described here uses simple optical ray tracing combined with a sonar performance prediction tool to produce target signatures capable of being used as prediction methods for the detectors.

3.1 Ray tracing

For high-frequencies (≥ 100 kHz), ray tracing is the most widely used method due mainly to the high computational burden of other methods (finite elements, normal modes, etc...). While ray theory permits direct calculation of range and time of returns, it is not able to describe wave effects such as diffraction and usually results in unnaturally crisp shadow zones.

The target is described as a set of interconnected triangular facets designed using a computer program. Rather than compute acoustic scattering from the target this method is concerned with determining mean highlight, background and shadow zone using principles of geometric optics. Once a minimum travel time and corresponding grazing angle for a ray is found, it is simple to verify whether it has intersected the seabed or the target. If it intersects the target then a normalized value of 1 times the sine of the grazing angle scaled by the directivity of the ray is used as a return. After a series of rays has been traced, the times are sorted in ascending order and interpolated at a spacing equivalent to the sampling frequency of the beamformed image with returns occurring at

3.3 SAS issues

The model described above was developed to simulate targets using a planar array. With synthetic aperture sonar, however, several array lengths are combined along the sensor travel path to create an effective array of some length longer than the original. In order for the simulation method to generate realistic SAS target signatures, the imagery must be simulated using the effective synthetic array length. This effective length changes with the range r. The theoretical SAS gain, Q_{SAS} is equal to:

$$Q_{SAS} = \frac{2\lambda r}{L_T L_R}, \tag{4}$$

where λ is the wavelength of transmitted signal and L_T and L_R are the lengths of the transmit and receive arrays. This follows from the ratio of the physical resolution of the array $\lambda r / L_R$ and the theoretical SAS resolution $L_T / 2$ [21]. Therefore, the resulting beam pattern for a receive array of n elements focusing at a range of r is simulated as one having n times Q_{SAS} elements, effectively creating a synthetic aperture at range r.

4 DETECTOR PERFORMANCE

The next step is to determine the detector performance characteristics by using the simulated signatures generated above. Two parameters are required for the fusion algorithm: a measure of detector confidence and the probability of detection versus range.

4.1 Detector Confidence

The confidence of a detection is used to assign values to the mass function when fusing detections with Dempster-Shafer theory, which is described in Section 5. The detection of targets is achieved by thresholding the output of the detection algorithm over the mission sonar imagery. Some filtering is done on the size of the detections to remove contacts which are too small or large. Once found, determination of the likelihood of this detection must be made in order to assign a belief to the hypothesis that the contact is in fact a legitimate one. These likelihood functions are constructed using the simulated target data generated from the signature modelling method described above. A dataset of images is chosen to estimate this performance. Several iterations of simulation followed by the application of the detectors are performed. Let $\gamma_{\bar{T}}$ be the output of the detector over a given SAS image without simulated targets and γ_T be the response on the same image with a number simulated target signatures. Since the images differ only at the pixels where the targets were added, the conditional probability of a detector's response to a target signature can be estimated using:

$$p(\gamma \mid T) = H(\gamma_T), \tag{5}$$

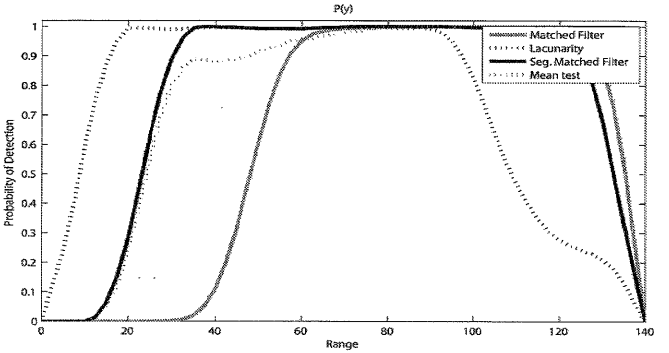

Figure 4. P(y) curves for all detectors for a fixed threshold.

5 DEMPSTER-SHAFER FUSION

Based on the detector confidence values and P(y) curves, the Dempster-Shafer theory of evidence (DSTE)[22] can be used to combine the detector results. A very brief outline of DSTE is given here.

Suppose an observation can be explained by one of n mutually exclusive and exhaustive hypotheses. This finite set is called the frame of discernment and denoted $\Theta=\{H_1,H_2,..,H_n\}$. Each hypothesis $H_i \subset \Theta$ corresponds to a one-element subset $\{H_i\}$, called a singleton. A composite hypothesis denoting the disjunction $H_i \cup .. \cup H_j$ is represented by the subset $\{H_i,..,H_j\} \subset \Theta$. The collection of all possible subsets of Θ (including the empty set $\emptyset$ and Θ itself) is called the power set of Θ and is denoted 2^{Θ}. Based on evidence such as an observation,belief is distributed within the power set through a basic probability assignment (BPA), also called a mass function. This is a mapping from the power set to the unit interval, m: $2^{\Theta} \rightarrow [0, 1]$, satisfying:

$$m(\emptyset) = 0$$
$$\sum_{A \subseteq \Theta} m(A) = 1, \tag{8}$$

The basic probability value m(A), represents the portion of belief committed *exactly* to subset A and not to any subsets B of A. The extent to which the evidence specifically supports the hypothesis represented by A (i.e. the *total* belief in A), is given by the belief function, Bel(A), which is the sum of all basic probability values m(B) in all subsets of A:

$$Bel(A) = \sum_{B \subseteq A} m(B), \tag{9}$$

Mass functions on the same frame of discernment from independent sources can be combined provided they are not totally contradictory. The orthogonal sum m of two mass functions m_1 and m_2 is denoted by $m = m_1 \oplus m_2$ and is given by Dempster's rule of combination:

$$m(\emptyset) = 0$$
$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)}, \tag{10}$$

provided that:

$$\sum_{B \cap C = \emptyset} m_1(B) m_2(C) < 1, \tag{11}$$

A random SAS image was chosen from the available data set to estimate the detector confidence and P(y) curves. A mix of 40 synthetic targets was generated in the image, repeated over 50 iterations, resulting in a total of 2000 target signatures.

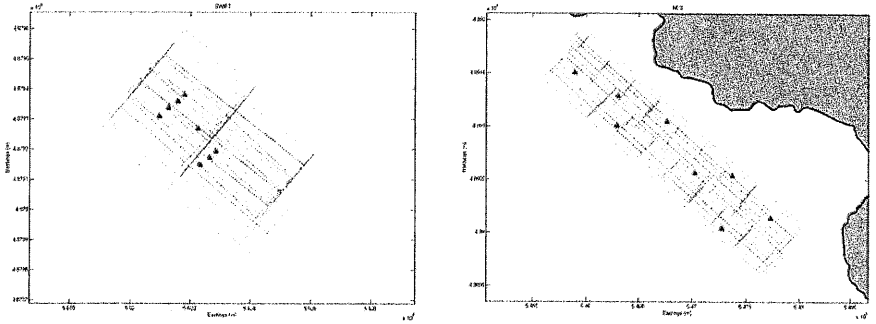

Figure 5 SWIFT and MX3 operational areas. Target fields are shown, as well as HUGIN 1000 vehicle tracks and area coverage.

Figure 6(a) shows the results of processing to obtain ROC curves for each individual detector on the two mission datasets. Dobeck [26] suggested that robust fusion can be performed when adjusting the detector thresholds such that the minimum number of false alarms is obtained for a probability of detection equal to one. Here, the real targets are used to compute these ROC curves in order to obtain the best thresholds for the method. Three of the detectors are able to obtain a 100 % probability of detection and the corresponding operating point used for the fusion is shown. The fourth detector (the matched filter) however, is unable to attain such a high detection probability, with an ROC curve flattening out at Pd = 0.79. Therefore the operating point which gives the least false alarms for this maximum performance is chosen to fuse this detection method.

Figure 6(b) shows the performance of the DSTE fusion methodology described above for the chosen detector thresholds. By thresholding the belief of the Dempster-Shafer fusion result, it is possible to obtain ROC curves for a given mission. Several curves are shown: First, the dotted line shows results of fusion using only the three best detectors without the matched filter (MF) and without incorporating the information gained by missing detections from multiple passes. Next, by adding information provided by non detections in different legs, the solid line is obtained. Finally, by the adding the MF detector, whose performance is not at the level of the other three, the dash-dot line is obtained. Note that this ROC curve is slightly better than the one obtained without the MF algorithm[§], but what is interesting is that the performance is not degraded. This is due to the fact that the confidence of this detector is lower and therefore the DSTE fusion method is robust to its poor performance. Also, the square shows the performance of "and"-ing the three best detectors, which results, as predicted by Dobeck, in a significant reduction in false alarms and an operating point similar to the ROC curve obtained without incorporating missed detections. The triangle shows the result of "and"-ing all four detectors. Since the MF detector implementation here does not achieve high detection rates, the performance is set to the worst of all the four detectors. Although for the given false alarm rate it appears to be slightly better than the other methods and individual detectors it is key to note that this is the maximum performance for this fusion method. Also shown is the result of "or"-ing the four detectors, resulting in very high numbers of false alarms.

For these results the thresholds were obtained using the real target data in order to fairly compare the proposed fusion method to other methods. In practice, obtaining the thresholds in this way would require a training set, although it should be possible to use the synthetic target signatures

[§] Although perhaps not significantly so

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the work of Karna Bryan, Giampaolo Cimino and Gianfranco Arcieri for their work on the AUV Planning and Evaluation tool, and particularly Gary Davies for help with the ESPRESSO model. Ben Evans was the scientist in charge of the SWIFT trials and Edoardo Bovio was responsible for the organization of the MX3 experiments. Finally, credit is due to John Fawcett, DRDC Atlantic, for his clear and concise work on the detection methods used in this paper.

REFERENCES

1. P.H. Haley and J.E. Gilmour, *Two-stage detection and discrimination system for side scan sonar equipment* (Westinghouse), US Patent 5,493,539, 1996.

2. D. Casasent and A. Ye, "Detection filters and algorithm fusion for ATR," *IEEE Transactions on Image Processing*, vol. 6, no. 1, pp. 114-125, 1997.

3. G. Dobeck, "Algorithm fusion for the detection and classification of sea mines in the very shallow water region using side-scan sonar imagery," in *Proceedings of the SPIE,* vol. 4038, pp. 348-361, 2000.

4. T. Aridgides, M. Fernandez and G. Dobeck, "Side scan sonar imagery fusion for sea mine detection and classification in very shallow water," in *Proceedings of the SPIE*, vol. 4394, pp. 1123-1134, 2001.

5. S. Reed,. Y. Pettilot and J. Bell, "Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information," *IEE Sonar Radar & Navigation*, vol. 151, no. 1, pp. 48-56, 2004.

6. B. Stage and B. Zerr, "Application of the Dempster-Shafer theory to combination of multiple view sonar images," in *UDT '96 Conference Proceedings*, London, UK, 1996.

7. J.A. Fawcett, A. Crawford, D. Hopkin, V. Myers and B. Zerr, "Computed-aided detection of targets from the CITADEL trial Klein data," Defence Research and Development Canada, Tech. Rep., 2006.

8. HUGIN website http://www.ffi.no/hugin

9. G. Dobeck, J. Hyland and L. Smedley, "Automated detection/classification of seamines in sonar imagery," in *Proceedings of the SPIE*, vol. 3079, 1997, pp. 90–110.

10. R. Kessel, "Using sonar speckle to identify regions of interest and for mine detection," *Proc. SPIE Vol. 4742*, 2002, pp. 440-445.

11. Y. Petillot, S. Reed and E. Coiras, "A framework for evaluating underwater mine detection and classification algorithms using augmented reality," in *UDT Europe 2006 Conference Proceedings*, Hamburg, Germany, 2006.

12. V. Myers, D. Davies, Y. Petillot and S. Reed, "Planning and evaluation of AUV missions using a data-driven approach," in *Proceedings of the 7th international conference of technology and the mine problem.* Montery, Ca: MINWARA, May 2006.

13. J. Bell, "A model for the simulation of sidescan sonar," Ph.D. dissertation, Heriot-Watt University, 1995.

14. G. Sammelmann, J. Christoff and J. Lathrop, "Synthetic images of proud targets," in *OCEANS 94*, 1994, pp. 266–271.

15. J.A. Fawcett, "Modeling of high-frequency scattering from objects using a hybrid Kirchoff/diffraction approach," *Journal of the Acoust. Soc. Am.* Vol. 109(4), April 2001.