AUTOMATIC TARGET RECOGNITION FOR AN AUTONOMOUS UNDERWATER VEHICLE (AUV) SIDESCAN SONAR IN COMPLEX ENVIRONMENTS

John Fawcett

DRDC Atlantic, PO Box 1012, Dartmouth, N.S., Canada, B2Y 3Z7

Abstract: In this paper the automated detection and classification of mine-like objects in sidescan (Marine Sonics) data is investigated. In particular, 4 different sites are considered involving sidescan sonar data collected by Australian, New Zealand, and NATO Undersea Research Centre (NURC) REMUS AUVs. Twelve different automated detection methods are proposed and their individual and fused performances are investigated. The different data sets contain different types of seabeds including relatively featureless, very-rippled, pock-marked, and cluttered seabeds. The performances of the detectors are also considered with respect to the different sites. After a detection/fusion process, the extracted small sonar images (mugshots) are further classified using template-matching. Examples of the template-matching are given.

Keywords: sidescan, ATR, autonomous
1. INTRODUCTION

Automatic Target Recognition (ATR) methods are important in the development of autonomous adaptive behaviour that exploits target information. Any Autonomous Underwater Vehicle (AUV) behaviour dependent upon the detection and possible classification of seabed objects will rely heavily upon the ATR performance. Ideally the ATR should work with a high probability of detection yet not produce a prohibitively large number of false alarms. Over the last several years DRDC Atlantic has worked on the development of ATR methods for sidescan sonar data (e.g., [1]-[2]). Most of that work utilized data collected with the DRDC Atlantic Klein 5500 sidescan sonar, often with the DRDC Atlantic semi-submersible vehicle DORADO. Much of this ATR analysis considered trials’ data in regions of relatively featureless seabeds [1] or assumed that the preliminary detection phase has already been carried out (perhaps, manually [2]) and concentrated upon the classification phase.

In this paper we consider data collected with different REMUS AUVs with Marine Sonics (900 kHz) side scan sonars. The data comes from 4 different sites: 2 near La Spezia, Italy [3], one from Australian trials [4], and data collected off Panama City, Florida from a New Zealand REMUS. Different complex seabed sites are considered where the seabed features can cause false alarms and produce highlights and shadows which interfere with those of a mine-like or manmade object. In addition, the sonar itself may have beampattern variations or surface reflections which can cause further complications. The end-to-end problem of preliminary detection(s), fusion, and template classification is considered. In addition, the seabeds considered are quite varied and present challenges to the detection and classification processes.

ATR methods with REMUS/Marine Sonics data have been considered by other authors [3-7]. In [7] trained detection methods are utilized with a data set of synthetic target images inserted into real Remus Marine Sonics images. In the approach of this paper, we consider fairly simple detectors and a classification method (template matching) which requires no training (except perhaps to define a threshold). Our data sets include a number of dummy mine-like or manmade objects: wedge-shapes, truncated-cones, cylinders of different sizes, and other objects. The intent of the detection phase was to detect almost all of these objects while still trying to minimize the number of false alarms.

In this paper, we first describe the data sets and show some example images. The preprocessing of this data (normalization, removal or mitigation of surface reflections, etc) is described. Then the twelve detection methods are very briefly described. Our simple fusion approach is also described. Some example images of the detection/fusion process are given. Finally, the template-matching procedure is described. The performance of the detection process for the data set is analyzed. First, the small images extracted about the detection points were manually inspected by the author and classed as mine-like or clutter. On the basis of this labelling, the performance of the various detectors was investigated by varying the thresholds of the detectors' outputs and Receiver Operating Characteristic (ROC) curves are generated. These curves are shown for different detectors and for the different sites of the data. A subset of the mugshots are then used with the template-classifier. Examples of the resulting template matches are shown.
2. DATA SETS AND PREPROCESSING

A REMUS/Marine Sonic data set was constructed from a composite of 900-kHz Marine Sonic Data taken from the NURC AUVPET2007 trial [3] and also from New Zealand and Australian [4] data. The New Zealand and Australia data consists of sonar files which were selected as their positions were fairly close to nominal dummy target positions or were manually determined to contain dummy targets. In total there are 646 files considered: 459 from the NURC data (this is only a subset of the entire NURC dataset) and 147.5 (this is an equivalent number as we did not analyze the starboard channels from some of the files) from the Australia/New Zealand data. Due to the relative size of the Australia/New Zealand data set, we considered it as one set. The sonar data from the NURC Remus was somewhat different than that of the Australian and New Zealand Remuses as it exhibited stronger nearfield beampattern effects. As well, the reflection from the water/air surface was more pronounced. Of the 459 NURC files, 193 are from Site D and 266 are from Site A. The Site A files contain images with significant sand ripples. Most of the data was obtained with a 30m range setting but in the AUS/NZ data there are number of 20m-range files and one at a 40m setting. Some example images are shown in Fig.1. The beampattern and surface reflection effects can be observed.

![Example sonar images from: (a) Site D – NURC data; (b) Australia/New Zealand data, and (c) Site A – NURC data](image)

In Fig.2 we show the results of some of the pre-processing steps. For the NURC data, the location of the surface reflection is predicted from the depth and a window is defined about this location. Within this window, a pixel amplitude is replaced by a ping-dependant local median value if it exceeds a threshold. The normalizing factors are computed along constant grazing angle curves as the altitude of the AUV changes. Two representative curves are shown in Fig.2a. (this is the starboard side of the image from Fig.1 – Site A, with the top/bottom orientation reversed for MATLAB) Also in this figure, the vertical red lines indicate a region of rapid heading change. In these regions (and also in regions about the surface reflection for the NURC data) the outputs from the detectors will be set to
zero. The resulting normalized image is shown in Fig.2b. For some of the detectors, the image is re-binned into 5 values [-1,-0.5,0,.5,1] (deep shadow to high highlight). This re-binned image is shown in Fig.2c. The presence of significant sand ripples can cause false alarms for automated detectors due to their sequence of highlight and shadow regions. In [6] and [8] Fourier-equalization and Dual Tree Complex Wavelet methods are used to reduce the effects of the ripple fields. Our approach is to form segmented images after Fourier filtering [6] and also after a simple version of wavelet filtering, (we use the software from [9] for the wavelets).

![Fig.2](a) the unnormalized data showing 2 representative normalization-integration curves and regions of large course changes (b) normalized image (c) re-binned data (-1,-.5,0,.5,1) and (d) rebinning after Fourier filtering The mine-like object is indicated by the yellow arrows

For example, a new re-binned image is formed from the image of Fig.2c by requiring that for its shadow regions the same pixel values in the Fourier-filtered image also have a relatively small amplitude as well. Thus in Fig.2d it can be seen that many of the shadow regions in Fig.2c have been eliminated; however, the shadow region corresponding to the target is still present.
Twelve detectors are considered in this paper. Some of them are similar to those considered in [1]. They include matched-filters, statistical tests, and highlight and shadow detectors and were typically implemented in terms of the outputs from two-dimensional filters. Most of the methods utilized the sonar images after normalization. In addition, the filter outputs are often normalized by a term involving their average (over positive values) background value. The various detector output images are each thresholded and the resulting regions are region-labelled with the pixel locations of the maximum detector value within the region designating the detection locations. The local density of detections by a detector in a 60 x 60 pixel region is computed and the threshold is increased for values greater than one. Finally, a set of detections, for each detector, is produced for each sonar file (port and starboard considered separately). The position fusion algorithm is simple. The detectors are each considered in turn and if a detection position is sufficiently close in proximity to a previously associated detection D then this detection is considered associated with detection D. If it is not associated with any previous associations, it is considered a new detection. The number of detectors associated with a detection is monitored and for this paper only detections associated with 4 or more detectors are considered. The requirement of 4 detections provided a significant reduction in the number of false alarms over a single detector, with the loss of only a very few mine-like detections. The position of the fused detection is that of the first detector which found it. An example sonar image with the individual detections and the fused detections is shown in Fig.3. One can see that there are many individual detections (we have set the thresholds quite low) and a few fused detections. There tend to be some spurious detections towards the boundaries of the image – for the analysis of this paper, only fused detections with detection coordinates in the range for along-track 32 < j < 970 and across-track km < k < 426 (km is computed in terms of the median altitude of a file) are considered for further analysis. A mugshot 64 x 128 in size is extracted around the detection point. As well, the maximum values of all the detectors in a 9 (along-track) x 59 window centred at the detection point are determined. The overall detection rate of this method was high: there were only a few cases of mine-like objects not detected that we had visually determined. The fusion rule has the nice feature that it lowers the false alarm rates from the individual detectors; on the other hand, there are cases where a particular combination of detectors will detect an object while some detectors will not. The advantages of detector fusion has been previously described by other authors [10].

For the total data set there are 5795 valid detections and extracted mugshots. The mugshots were manually examined and flagged as mine-like or clutter. The designation of mine-like was often very clear but there were cases of some objects where the choice was more subjective. The breakdown of the detections is: NURC Site A, 230 mine-like, 2209 clutter, NURC Site D, 335 mine-like, 1694 clutter, AUS/NZ, 150 mine-like, 1177 clutter. For NURC Site A many of the clutter events were due to ripples. The AUS/NZ data contained images with boulder fields and these contributed the majority of the clutter events. In this case, many of the clutter events were quite mine-like in appearance and thus are difficult to discriminate against. From the numbers of clutter events at the sites, it can be seen that the false alarm rate is quite high at this stage. Since we have saved values of the individual detectors, we can investigate if there are further improvements in performance from varying the thresholds of any of the 12 values. In Fig.3 we show the
ROC curves for only 2 of these detectors, a matched filter detector (solid line) and the lacunarity detector (dashed line) for the 3 sites individually and for the total data set. As can be seen, the performance of the lacunarity detector is better than that of the matched filter for all the data sets. In particular it gives very good performance at NURC Site A – the rippled site. We emphasize that these are the performances of these 2 detectors after a first detection phase which required the agreement of at least 4 of the detectors. This means, for example, that some of the false alarms of these detectors have already been eliminated. We have also considered linear combinations of 2 of these features (Fisher discriminant analysis) [10]. In the case of the lacunarity detector there was no advantage in combining it with another feature. Some of the other detectors could be significantly improved by combining them with another. The prefiltered matched filters (Wavelet, Fourier) also performed better than Detector 1 for Site A but not as well as the lacunarity detector.

For the next stage of analysis, we now consider only the original detections with lacunarity \( L > 0.325 \) For the total data set this reduces the number of targets to 651 or 91\% of the original number and the number of clutter objects is significantly reduced to 528 or 10\% of the original number. The associated mugshots are then used by the template-matching algorithm. The method is briefly described as follows. First a library of templates is computed using ray-tracing for detection of ranges \([2,…,30]m\) (increment of 2m) and at different aspect angles (10° increment) for the wedge-shaped target and 4 different-sized cylinders. For the truncated-cone shape, only a single aspect is required. The along-track spacing used was 12cm. Some of the New Zealand data was at a 20 m range setting and the AUV went slower in this case. We used an along-track spacing of 8 cm in this case resulting in larger templates. The mugshot is re-binned into the values using the same basic approach as was described previously for the sonar file. Let us denote this new mugshot as the matrix \( M \). For a given range of detection, the appropriate set of templates are considered – templates from the previous discrete range-index and the subsequent index are also considered. As well, for the truncated-cone and wedge-shapes, we also considered versions which are elongated or shrunk by a factor of 20\%. The templates are such that their shadow values are equal to negative one and the highlight values are positive but usually less than one. In addition, the template is augmented by 7 rows and 10 columns of zeros before and after the extent of the bounding box of the template. Let us denote the augmented template as \( T \) and \( T_Z = 1 \) where \( T=0 \). Then the generalized cross-correlation is defined as: \( C= M * T /(1+5(|M|*T_Z)) \) where \( * \) denotes cross-correlation and \( T \) and \( T_Z \) are normalized to have unit \( L_1 \) norm. In Fig.4 we show the 25 best matches and their corresponding determined templates for the data. As can be seen, the matches are very reasonable and, in general, good template matches were obtained. The aspect of the template is an automatic result of this method. There do remain some seabed features which yield fairly high template match values. Although the template matches may visually appear reasonable, they are not always correct in their classification. For example, the truncated cones are often classed as wedge-shapes, illustrating the fact that without a sufficient pixel resolution, 2 different targets can visually appear very similar, and due to the surrounding seabed features, the variable speed of the AUV, and other reasons, an incorrect template-match may result.

4 SUMMARY

This paper has considered a diverse set of Marine Sonics sonar images. The seabed was often complex and there were sonar-related (beam effects, surface reflection) issues to
contend with. This resulted in a large number of false alarms. We processed the data with 12 detectors. Using a simple detector fusion followed by a lacunarity threshold, we were able to reduce the number of false alarms significantly. The ATR performance did depend upon the seabed type. The best overall performance (in terms of the ROC curves) was for the case of the relatively featureless seabed of NURC Site D and with the appropriate detector (Lacunarity) the rippled seabed of NURC Site A. The poorest overall performance was for the Australia/New Zealand data which had the most mine-like clutter. The mugshots corresponding to the remaining detections were then further processed using template-matching. This procedure gave, in general, visually reasonable matches to the imagery.

Fig. 3 (left) showing individual detections and fused detections (yellow circles) (middle) a zoom of one of the fused detections (right) the ROC curves for the matched filter and lacunarity features after the fused detections for the different sets of data

(a)                                               (b)

Fig. 4 (a) showing the top 25 mugshots and (b) the corresponding determined templates. The vertical axis for each image is along-track and the horizontal axis across-track
ACKNOWLEDGEMENTS

The author would like to thank NURC, DSTO Australia, and DTA, New Zealand for providing the REMUS Marine Sonics data of this paper. Vincent Myers of DRDC Atlantic was very helpful with many template computations and some of the software development used for this study. Dr. Gerald Dobeck of NSWC, Panama City very kindly provided his MATLAB software for large-scale clutter suppression.

REFERENCES


