TOPICS IN AUTOMATED TARGET RECOGNITION FOR HIGH RESOLUTION SONAR IMAGERY

Marc Pinto, Vincent Myers

NATO Undersea Research Centre,
Viale San Bartolomeo, 400
La Spezia, Italy, 19138
{pinto,myers}@nurc.nato.int

1 INTRODUCTION

Automated target recognition (ATR) in high resolution sonar imagery has been the topic of significant research for more than a decade, with some systems being employed as an integral part of operational minehunting systems. Most systems end up failing by many standard measures of effectiveness, either as operator aids or as fully automated solutions (such as those being considered for AUVs (autonomous underwater vehicles)). The poor ATR performance can be due to poor data quality, due to limitations of the sonar or the environment but sometimes also to inappropriate concepts of use.

After a quick review of basic issues relative to target recognition with high resolution sonar, the issue of performance prediction is addressed. This important issue is often overlooked in both radar and sonar ATR research and, as a result, there is an acute need for the development of a theoretical foundation for performance prediction and evaluation. As an example of this approach, theoretical bounds for the performance of shadow based ATR are recalled. This provides significant insight into the intrinsic performance limitations of existing systems and the operational benefits which can be expected from improved sensors, in particular wideband interferometric synthetic aperture sonar (InSAS). Another promising approach is to work towards increased autonomy, i.e. to exploit platform maneuvers and on-board decision making to collect sufficient information on the target to allow reliable recognition.

2 MINEHUNTING SONAR ATR

Traditional ATRs were designed with the objective of finding signals in noise using a set of techniques known collectively as signal detection theory[1]. When the sensor resolution is low enough so that target can be modeled as a single point scatterer and detection is limited by electrical noise in the system which follows a known statistical distribution (e.g. Gaussian white noise), a hypothesis test can be used to make a decision as to whether or not a target is present with detection performance typically set by the signal to noise ratio and the desired false-alarm rate. Such detectors have been very successful in sonar as well as in radar and are optimal within their assumed theoretical framework.

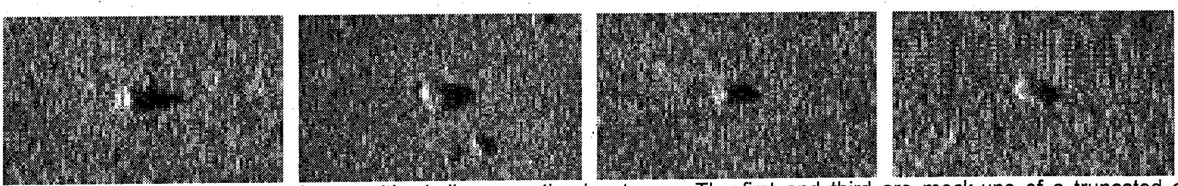

Figure 1 Objects different in nature with similar acoustic signatures. The first and third are mock-ups of a truncated cone, whereas the second and fourth are of naturally occurring rocks. Data gathered using a common off-the-shelf sensor operating at 455 kHz with 20 kHz bandwidth.

However, improvements in sensor technology and signal processing have led to large increases in both resolution and signal to noise ratio with the result that the target is no longer small compared with the sensor resolution cell and that it has now to be detected against a textured and cluttered background (seabed) rather than noise. It should be apparent n the sonar images pictured in Fig.1 that the problem is far more complex than assumed in the simple statistical theory. It is that of finding every object which could potentially be a target (detection) and then determining which of these are of sufficient interest to warrant further attention (classification).

Actually the necessity to go beyond the simple statistical detection theory has been recognized very early on in the field of minehunting sonar, who has defined systems and corresponding procedures to cope with this additional complexity. The traditional minehunting system consists of a maneuverable surface ship with at least two different sonars for detection and classification. Typically a detection sonar is long range forward looking sonar which provides the operator with a fluctuating sequence of echoes for each object under investigation. Detected objects are then classified by another operator. The ship is driven to a specified (much shorter) range to the object, while holding the target on the detection sonar to avoid relocation issues. Then the much higher resolution classification sonar is used to assess additional features of the target (e.g. size, shape as well as its strength). Often multiple aspects on the target are collected to arrive at a better decision.

Errors at each stage in the process, depending on their nature, either increase the residual risk (case of a undetected or misclassified mine) or reduce the efficiency of the operation (case of a non-mine detected and/or classified as a mine). These operator tasks are made increasingly complex by the operational requirement to detect and classify targets which are weaker in strength and smaller in size than before.

Classification is generally followed by a third step called identification which is similar in nature but is performed at even shorter range with even higher resolution sensors, usually mounted on a remotely operated vehicle controlled from the ship. Identification can be a very useful stage in the initial phases of operation over unknown areas, since it provides ground truth for both real and false targets, allowing a better understanding a posteriori of their sonar pictures and possible causes for wrong decisions. The lessons learnt are very important to improve future decisions.

Operator aids, termed Computer-Aided Detection and Classification (CAD/CAC) have been developed to assist them in this task. The CAD is basically a multi-ping tracker which is based on statistical detection theory whereas the CAC performs pattern recognition exploiting features in the highlights and/or the shadow when available.

Another class of minehunting systems exploits high resolution sidescan sonars which can either be towed by all types of platforms or even self-propelled. In all current modes of operation the platform performs a pre-determined survey pattern which corresponds to a set of parallel tracks, with a spacing set by the sonar range. The geo-referenced image is formed in this manner is then analysed by an operator either during the mission or post-mission (depending on the bandwidth available on the communication link). This mode of operation was historically used for routine survey operations, to detect changes with respect to prior surveys (e.g. a historical contact data base can be used to detect new contacts). Automated change detection still presents quite a few challenges however, due to navigation errors and distortions in the sonar images induced by unwanted platform motions, which cause ambiguities, and more fundamentally changes in the seabed. .

More recently, with the introduction of autonomous underwater vehicles, there has been a tendency to perform minehunting operations over unknown areas with similar pre-programmed sidescan surveys. Due to the limited bandwidth available for underwater communications, the data is analysed post-mission by an operator whose task is now very challenging. Indeed, unlike the previous ship-based systems, the detection and classification has now to be performed on a single sonar image, with no possibility for the operator to interact with the system to collect additional data on the contact. In addition the identification phase is generally conducted separately, with no feedback into the classification process.

Operator aids, known as post-mission analysis tools, are being actively investigated to assist the operator in this difficult task. They are sometimes also known as CAD/CAC, like the operator aids used for the ship-based forward-looking systems which is rather confusing since they differ significantly in both design and use. The present level of performance achieved by these aids is not considered satisfactory. Meeting the desired goal will require going well beyond improvements in pattern recognition technology alone, to include order of magnitude improvements of the quantity and quality of information provided at the input of the operator aid. In addition to sensor improvement, discussed below, this may include means to restore a much higher level of interaction between the sensing and the decision-making. Due to intrinsic limitations of underwater communication links, it is likely that the largest part of this interaction has to be done on-board the vehicle with little or no human interaction.

The introduction of autonomy is not specific to minehunting systems and similar issues are being addressed for other areas including aerial surveillance by UAVs. However some simplifying features of the minehunting application make it an attractive area to test autonomy concepts. The targets are inherently static which considerably simplifies the information collection process. Indeed by displacing a single AUV at various appropriately chosen spatial locations and memorizing the sensor and navigation data, one forms a "synthetic autonomous sampling network" from which it should be possible to produce a high quality picture of the target. Of course this could also be achieved in principle with a real network of AUVs (for a moving target this would be the only option) but only with a step increase in complexity related to the communications, navigation and water space management overheads.

3 SENSOR LEVEL IMPROVEMENTS

Ultimately, all ATR methods attempt to estimate the joint distribution of the acoustic signatures (or features derived from it) and the target classes of interest. Formally, the marginal distribution $P(X|Y)$ defines the probability that a given d length vector of measurements X will be observed for a target of the class Y. For the sonar problem, the measurements X are obtained from an acoustic image and the class is the target / non-target task or targets subcategorized into finer classes of more specific target types. These class-conditional probabilities $P(X|Y_i)$ are important in determining the possible performance of an ATR system. Any overlap in these distributions will result in unavoidable classification errors. The best achievable probability of error is called the Bayes error rate, defined as [14]:

$$L^* = 1 - \int_{\Re^d} \max_{i=\{1..m\}} P(y_i)p(x \mid y_i)dx$$

where $P(X|Y_i)$ is the marginal distribution of the set of measurements X for class i. This integral is performed over the entire d-dimensional feature space and $L^*= 0$ when all $P(X|Y_j) \geq 0$ for a single class j, with equality for all other classes, i.e. when there is no overlap between the classes in feature space. The Bayes error rate is the theoretical limit for any classification method, automatic or otherwise.

As in most problems of this type, the $P(X|Y_j)$ are not known and so must be learned using data from past measurements or somehow assumed using a priori knowledge of the problem. A well-known property of learning algorithms is that the number of training samples n must be exponential in d to achieve consistent error rates. Many systems employ feature reduction techniques to reduce d to a manageable number but any such method which does not test all possible subsets of d are suboptimal for some distributions of (X,Y), making for a combinatorial time complexity and so prohibitive for large values of d (see Devroye et. al. pp 562-563). The problem is aggravated by the number of degrees of freedom in the target class. Even in the favourable case when it is known in advance which target one is looking for, the target signature can exhibit a high variability due to changes in the geometry of the scene, or complex interactions between the target and the seabed such as shadows on sand ripple fields.

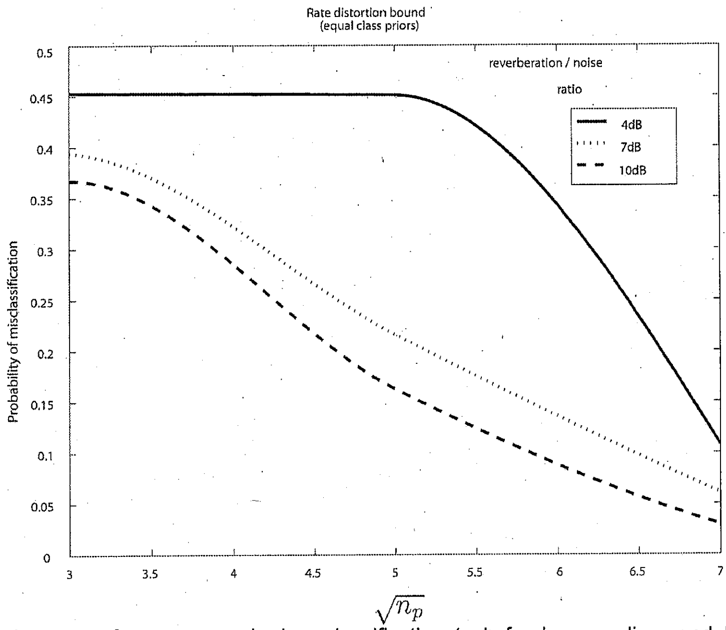

Figure 2 Predicted performance for an easy shadow classification task for low, medium and high contrast ratio with the background. The Bayes probability of error (L*) is shown versus the square root of the number of independent pixels on target (if square pixels are assumed, this is the number of along-track pixels). For a high contrast sensor at least 7 along-track pixels on target are required for reliable performance (95%).

As a result, features of non-isotropic targets become distinct enough to effectively spread the marginal distribution of X, decreasing the predictive ability of any feature.
Results have been reported on object detection and clutter rejection in high resolution sidescan images employing a correlation between a mask of a target-like signature as a detection method, with performance claimed to be similar to that of a human operator for the same task[2]. In areas of high clutter, more discrimination is required, achieved by additional filtering of non-target signatures. Since a parametric model of the clutter is difficult to develop due to the large number of degrees of freedom and consequently number of samples required, non-parametric methods are common: for instance, work has been done with neural networks [3, 4], decision trees [5] and support vector machines [6, 7] differing chiefly in the features being used as inputs. Multiple classifiers are occasionally fused (usually at the decision level) together based on the belief that most classification methods are able to accurately predict the target classes while false positives, appearing more randomly, tend to differ from method to method [8], thus reducing the false alarm rate while maintaining high classification performance. Other learning methods such as active and semi-supervised learning have been employed to increase the convergence of the learning algorithm [9].
Model-based approaches[10, 11, 12]. attempt to remove this dependency, however in addition to the cost of repeatedly of running a forward model over all possible combinations of aspects and burial state in 3D, the variability of the projected 2D model results in being able to match almost any candidate signature.

The sonar imagery itself presents some considerable difficulties in processing which are still somewhat underestimated. For classification, sonar images are segmented into regions of shadow and echo for the computation of features or model-matching [6,10,13]. However even on a flat, featureless seabed, the coherent nature of the acoustic imaging makes segmentation significantly more complex than for incoherent imaging sensors, such as optical or infrared sensors. The pixel amplitudes follow noise-like speckle distributions (e.g. Rice-Rayleigh) which cause estimation noise in the boundary regions and lead to imperfect segmentation, and those errors are propagated into the feature extraction step. The estimation noise can only be combated by increasing the sonar resolution beyond what would be strictly necessary to achieve target recognition, using criteria suited for incoherent imaging systems[16].

Myers and Pinto[15] established bounds on L^* for the sonar ATR based on shadow classification using information theory (see Fig 2). They defined a simple sonar classification task and derived the Bayes error rate for a range of image resolutions and shadow to background contrast ratios. They concluded that for even this simple task with a high contrast approximately 7-8 independent along-track pixels would be required for high probabilities of classification†, with decreasing performance as resolution and contrast worsened. This is much higher resolution than that of most, if not all, sidescan sonars in operation. More complex shape definitions will result in a further decrease in performance. This effect has been confirmed using simulated targets on real data, where automatic detection algorithms performed with varying success and depended greatly on seabed orientation, even for isotropic targets [17]

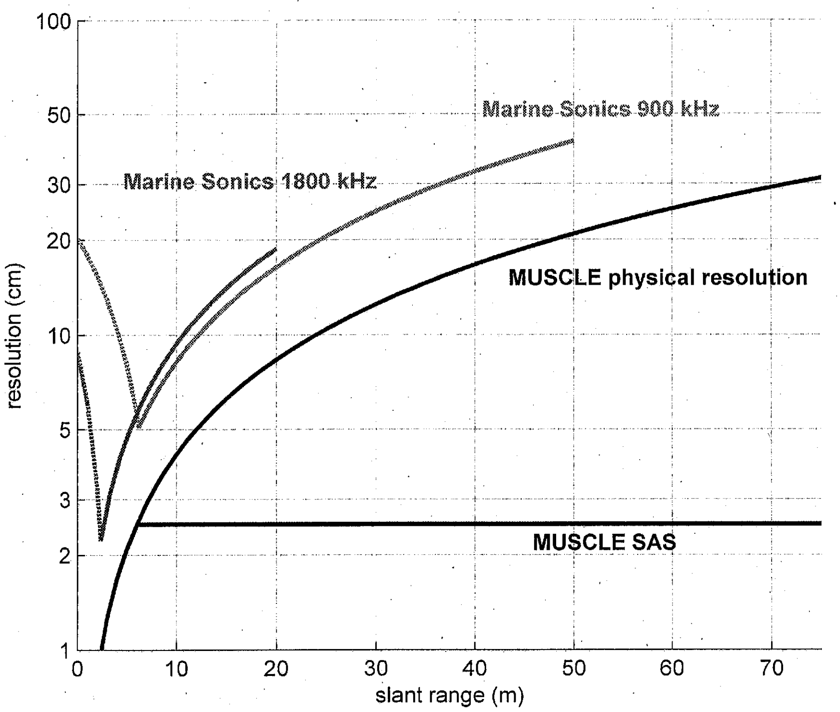

Figure 3 Cross-range resolution as a function of range for a COTS dual frequency sidescan sonar commonly used for AUV-based sidescan surveys. For a target size varying between 0.5m and 1m, meeting the Myers-Pinto criteria requires operation at less than 10m range. In comparison (blue) theoretical resolution of a prototype AUV-based synthetic aperture system under development at NURC called MUSCLE.

4 CONCLUSIONS

† For example Johnson's criteria derived for night-vision systems[16].

Automatic Target Recognition for high resolution sonar imagery presents a challenging task for system developers due partly to some of the issues which have been described above. In addition to being vulnerable to the same difficulties that other automated classification systems face there are additional limitations given by the sensor, the environment and ambiguities in the target classes.
There is evidence that the performance expected from ATR is significantly above what is theoretically achievable with the systems and procedures in place, a fortiori what is achievable in practice. In this case, significant improvement in ATR performance must come from an increase in the quality of information input to the ATR, rather than improvements in the ATR itself. Fortunately recent advances in sonar technology, such as wideband sonar, synthetic aperture sonar, interferometric sonar and combinations thereof, allow a step change in the quality of sonar data which, if appropriately exploited, should enable this ATR improvement. However interpretation of high resolution sonar imagery is not as mature as it may seem and may have to be revisited in connection with sensor developments. Furthermore the excellent navigation performance offered at a reasonable cost by inertial navigation systems is also an important enabler, since this allows precise association of surveys carried out at different headings. It also seems important to restore means to achieve with AUVs a much higher level of interaction between the sensing and the decision-making. This is the challenge facing the next generation of truly autonomous minehunting systems.
Finally, a measure of the performance actually achieved by the system will always be required. There is a tendency to evaluate ATR performance by comparison to that of operators which can then bias the results[18]. The use of objective ground truth would allow a more accurate estimation.

5 ACKNOWLEDGEMENTS

The authors are grateful to A. Bellettini for many useful discussions and in particular for producing Figure 3.

6 REFERENCES

1. H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I, Wiley-Interscience, (2001).
2. T. Aridgides, D. Antoni, M.F. Fernandez, G.J. Dobeck, "Adaptive filter for mine detection and classification in side-scan sonar imagery," Proc. SPIE Vol 2496, pp 457-486, (1995).
3. R. Balasubramanian and M. Stevenson, "Pattern recognition for underwater mine detection," in CAD/CAC 2001 proceedings, November (2001).
4. W. Weiming and W.G. Szymczak, "Multiresolution neural networks for mine detection in side scan sonar images", Proc. SPIE, Vol. 3392, pp 297-305, (1998).
5. V.L. Myers, "Decision trees for computer-aided detection and classification (CAD/CAC) of mines in sidescan sonar imagery," DRDC-Atlantic TM 2002-144, Halifax, Canada, (2002).
6. J.A. Fawcett and V.L. Myers, "Computed-aided classification for a database of images of minelike objects," DRDC-Atlantic, TM 2004-272, Halifax, Canada, (2004).
7. G.J. Dobeck and J. Cobb, "Fusion of multiple quadratic penalty support vector machines (QPSVM) for automatic sea mine detection and classification," in Proceedings of the SPIE, vol. 4742, (2002).
8. G.J. Dobeck, J.C. Hyland and L. Smedley, "Automated detection / classification of seamines in sonar imagery," Proceedings of the SPIE, vol. 3079, pp. 90-110, (1997).
9. E. Dura, Yan Zhang; Xuejun Liao; G.J. Dobeck, and L. Carin, "Active learning for detection of mine-like objects in side-scan sonar imagery", IEEE Journal of Oceanic Engineering, Volume 30, Issue 2, pp 360 – 371, (2005).
10. S. Reed, Y. Pettilot, and J. Bell, "An automatic approach to the detection and extraction of mine features in side scan sonar imagery," IEEE Journal of Oceanic Engineering, vol. 28, pp. 90-105, January, (2003).

11. S. Reed, Y. Petillot and J. Bell, "Model-based approach to the detection and classification of mines in sidescan sonar," Applied Optics, Vol. 43, Issue 2, January, pp 237-246 (2004).
12. Quidu, J. Ph. Malkasse, G. Burel and P. Vilbé, "Mine classification based on raw sonar data: An approach combining Fourier descriptors, statistical models and genetic algorithms," OCEANS MTS/IEEE conference proceedings, Vol. 1, pp 285-290, (2000).
13. M. Mignotte, C. Collet, P. Perez and P. Bouthemy, Sonar image segmentation using an unsupervised hierarchical MRF model, IEEE Transactions on Image Processing, Vol. 9, Issue 7, pp 1216 – 1231, (2000).
14. L. Devroye, L. Györfi, and G. Lugosi, A Probabilistic Theory of Pattern Recognition, New York, NY: Springer, (1996).
15. V.L. Myers and M.A. Pinto, "Information theoretic bounds on the performance of ATR algorithm performance for sidescan sonar target classification," in Proceedings of the SPIE, vol. 5807, (2002).
16. J. Johnson, "Analysis of image forming systems," in Image Intensifier Symposium, U.S. Army Research and Development Laboratories, pp. 244–273, (1958).
17. V.L. Myers, G. Davies, Y. Pettilot and S. Reed, "Planning and Evaluation of AUV missions using data-driven approaches," in 7th International Symposium on Technology and the Mine Problem, May, (2006).
18. R.T. Kessel, "Apparent Reliability: Conditions for Reliance on Supervised Automation," DRDC-Atlantic, TM-2005-155, Halifax, Canada, (2005).