Predicting synthetic aperture sonar performance with augmented sonar imagery

John A. Fawcett
DRDC – Atlantic Research Centre
Predicting synthetic aperture sonar performance with augmented sonar imagery

John A. Fawcett
DRDC – Atlantic Research Centre

Defence Research and Development Canada
Scientific Report
DRDC-RDDC-2017-R020
March 2017
Abstract

The prediction of synthetic aperture sonar/sidescan sonar detection performance for a specific sonar survey is a topic of much interest. There are many ways to define performance. In this report, the detection and false alarm rates of specified automatic target recognition (ATR) methods for minelike objects will be the performance statistics of interest. Previous researchers have investigated the expected sonar performance as a function of image and sonar features. The approach taken in this report follows more closely the work of other authors on augmented reality. Synthetic target images are inserted into real sonar images obtained from a specific survey. Thus, the background sonar imagery and the clutter represents a survey exactly. The issue is making the inserted target realistic, in terms of its distortions and the noise characteristics of its shadow and highlight regions. These features need not be exactly correct but they need to be sufficiently correct to yield the correct ATR performance and behaviour. This report will describe the approach taken to insert targets with noise statistics which vary with respect to range and the surrounding image statistics. The approach will be applied to sonar data from a NATO Centre for Maritime Research and Experimentation (CMRE) trial (three different sites) and the predicted ATR performances will be compared to the actual performances. In addition, the performance model can be used to predict performances as a function of other parameters, such as target type, range, etc for which the trial data is limited. Also, the uncertainties in the predicted performance can be estimated through Monte Carlo simulations.

Significance for defence and security

It is shown that the approach of inserting synthetic target images into real sonar images has very good potential for predicting ATR performance for a trial. Given representative sonar images of a region’s seabed, the insertion method can also be used to predict performance for different distributions of target types and target ranges. It can also be used to investigate the statistical variability of the performance measures. There are other possible applications of the synthetic target insertion method including its use in preparing images for training operators.
Résumé

La prévision de la performance de détection d’un sonar à synthèse d’ouverture ou à balayage latéral pour un levé précis suscite beaucoup d’intérêt. On peut définir la performance de plusieurs façons. Dans le présent rapport, les taux de détection et de fausses alarmes des méthodes de reconnaissance automatique de cibles (ATR) prescrites pour des objets de type mine sont les statistiques sur la performance présentant un intérêt. D’autres chercheurs ont examiné la performance attendue d’un sonar en fonction des caractéristiques des images et du sonar. L’approche adoptée dans le présent rapport suit plus étroitement le travail des autres auteurs sur la réalité augmentée. On insère des images de cibles synthétiques dans des images sonar réelles provenant d’un levé précis. Par conséquent, l’imagerie sonar d’arrière-plan et le fouillis représentent exactement un levé. Le résultat donne une cible insérée réaliste sur le plan des distorsions et des caractéristiques de bruit de son ombre et des régions mises en évidence. Ces éléments n’ont pas besoin d’être tout à fait exacts, mais ils doivent l’être suffisamment pour fournir une bonne performance et un bon comportement ATR. Le présent rapport décrit l’approche utilisée pour insérer des cibles dont les statistiques sur le bruit varient en fonction de la distance et des statistiques sur les images de l’environnement. Cette approche est appliquée aux données sonar provenant d’un essai (à trois emplacements différents) du Centre pour la recherche et l’expérimentation maritimes (CMRE) de l’OTAN, et les performances ATR prévues sont comparées aux performances réelles. De plus, on peut utiliser le modèle de performance pour prévoir celles en fonction d’autres paramètres, comme le type de cible, la distance, etc., pour lesquels les données d’essai sont limitées. Par ailleurs, les incertitudes quant à la performance prévue peuvent être estimées au moyen de simulations de Monte Carlo.

Importance pour la défense et la sécurité

Les résultats montrent que la méthode d’insertion d’images de cibles synthétiques dans des images sonar réelles a un très bon potentiel pour prévoir la performance ATR lors d’un essai. Compte tenu des images sonar représentatives du fond marin d’une région, on peut aussi utiliser la méthode d’insertion pour prévoir la performance selon diverses répartitions de types et de distances de cibles. On peut également l’utiliser pour étudier la variabilité statistique des mesures de performance. De plus, la méthode d’insertion de cibles synthétiques peut servir à d’autres applications, dont la préparation d’images pour les opérateurs de formation.
## Table of contents

Abstract .................................................. i
Significance for defence and security ....................... i
Résumé .................................................. ii
Importance pour la défense et la sécurité ..................... ii
Table of contents ........................................ iii
List of figures .......................................... iv
Acknowledgements ....................................... vi
1  Introduction ........................................ 1
2  Insertion of templates into imagery .................... 2
   2.1  Highlight generation ........................... 4
   2.2  Shadow noise generation ...................... 5
3  Detection methods .................................... 8
   3.1  Matched-Filter ............................... 8
   3.2  Haar cascade ................................. 8
4  Synthetically generated images - numerical experiments’ results ... 9
   4.1  Sandy rippled, cluttered site - Site A .......... 9
   4.2  Flat Muddy Site - Site B .................... 11
   4.3  Flat Sandy Site - Site C ..................... 14
5  ROC curve statistical and parameter modelling .......... 19
6  Discussion of Results ................................ 19
References ............................................... 23

DRDC-RDDC-2017-R020
List of figures

Figure 1: Some example windows of background data used to compute statistics (a), (b) and (c) and the corresponding histograms and fit Rayleigh distribution (red curve) for the shadow pixels (d), (e), and (f)................................. 7

Figure 2: Example images, after normalization and conversion to TIFF format, from Site A with real and inserted targets. In the top figure, the synthetic targets are indicated with a $S$. ....................... 12

Figure 3: The ROC curves for a matched filter detector for real and synthetically inserted targets with the sandy, rippled backgrounds (Site A). ................................................................. 13

Figure 4: The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade. ................................................................. 13

Figure 5: The ROC curves for three previously trained Haar cascades on images from the sandy, rippled site. The curves are as in Fig. 4 but the testing set excluded files used in the training of the Colossus trained cascades. ................................................................. 14

Figure 6: Example images, after normalization and conversion to TIFF format, from Site B with real and inserted targets. In the top figure, the synthetic targets are indicated with a $S$. ....................... 15

Figure 7: The ROC curves for a matched-filter for the real and the synthetically-inserted targets for the flat, muddy site (Site B). . . . 16

Figure 8: The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade. ................................................................. 16

Figure 9: Example images, after normalization and conversion to TIFF format, from Site C with real and inserted targets. In the top figure, the synthetic targets are indicated with a ‘S’. ....................... 17
Figure 10: The ROC curves for a matched filter for the real and the synthetically inserted targets for the sandy, benign site.

Figure 11: The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade. Note the very different x-axis scale in this case.

Figure 12: The ROC curves (Site A) for a matched-filter for 21 realizations of the inserted targets — real-target ROC curves in blue and synthetic-target curves in red.

Figure 13: The ROC curves (Site A) for a Haar Cascade (Cascade 1 in Fig. 2) for 21 realizations of the inserted targets — real-target ROC curves in blue and synthetic-target curves in red.
Acknowledgements

The author would like to thank the NATO Centre for Maritime Research and Experimentation (CMRE) for providing the sonar data used in this report.
1 Introduction

In recent years, there has been much interest in the prediction of sidescan and synthetic aperture sonar (SAS) sonar performance for a sonar survey. This could be for an already-completed survey, for a future survey, or for a survey being carried out. In this latter case, it might be of interest for an autonomous underwater vehicle (AUV) to adapt its mission parameters on the basis of the predicted performance [1]. There are various measures of performance. In this report, it will refer to the prediction of Receiver Operating Characteristic (ROC) curves for automatic target recognition (ATR) methods. This requires that the performance in terms of both the false alarm and the target detection statistics is accurately modeled.

Researchers have investigated the sonar performance as a function of seabed/image parameters and synthetic aperture sonar (SAS) processing parameters [2-5]. This performance prediction is often constructed empirically through the use of trial data results from different seabed environments. More qualitatively, a prediction model can be used to indicate to an operator the areas where it is expected that minehunting performance is excellent and regions where the minehunting performance is poor.

In this report, another approach is taken. The actual sonar images from a survey are used as the background images for inserted snippets (called “mugshots”) of target imagery. The target imagery is generated from simple raytracing [6,7] with the object of interest on a flat seabed for a given range, aspect and altitude. In fact, we will use a pre-computed set of target mugshots at a set of discrete ranges and aspects to insert into the images. There has been other work [8–11] which utilizes synthetic data for the training process. In [8-9] an inversion method is used to estimate the elevation and reflectivity of the seabed and the incident intensity of the sonar pulse from a sidescan sonar image. Artificial targets can then be inserted into the reflectivity/elevation data and these maps can be converted back to a sonar image through ray-tracing. In [8] the resulting images were used to train a detector/classifier based upon spatial-filter features. The authors used the trained classifiers and obtained very good results for a test set of real sonar images. In [5], inserted targets were used to investigate the performances of computer detection methods for different seabed types. Ref. 6 considered the insertion of real target snippets into different backgrounds to augment the training set for a Haar cascade detector. The use of inserted synthetic targets is also used in Seabyte’s PATT (planning and evaluation software) [11].

References 8 and 9 report the work most closely related to the work reported here. In this report, instead of using an inversion method to determine seabed elevation and reflectivity, we simply use the existing background imagery and directly insert a target image (template). However, without modifications to the templates, the shadows and highlights are much too perfect to be realistic. There is no noise in the shadow and the target/background boundaries are perfect. Thus, for example, a performance
simulation with this method for ATR methods will likely be overly optimistic as the shadow region is much better defined than in reality. A large portion of this report is concerned with how to add realistic levels of synthetic noise into the shadow regions of the templates and how to randomize the highlight regions.

The accuracy of our performance predictions will be tested by considering imagery with real trial targets and inserting the same target types at approximately the same ranges. Automatic Target Recognition (ATR) methods are then run on this new set of images and the ROC curves for the real targets are compared to those generated with the synthetic targets. This will be done for two types of ATR method, a pre-trained Haar cascade[12-13] and a simple matched-filter [14]. Once one has confidence that the synthetic images are yielding realistic ATR performance statistics, the technique can be used to generate statistics for other scenarios. This report investigates the application of synthetic target insertion to performance prediction. There are also possible applications of the technique to mission planning, operator training, and as reported in [15] ATR training.

2 Insertion of templates into imagery

For synthetic target insertion, there are 2 components: the underlying real background sonar image and a smaller rectangular window containing a computer-generated target image or template. For this report, the background images are from the CMRE MUSCLE system used during the COLOSSUS II trial. For the remainder of this report we will simply use the term Colossus. We will only consider the sonar images which contain trial targets. The Colossus trial data is from 2008 and has been extensively used for sonar processing research [4,615,16]. This data was taken from 3 different sites [16], one was sandy with ripples and clutter (we denote as Site A), one was flat and muddy (Site B), and the third was flat and sandy (Site C). Our designation of the sites as A, B, and C are not necessarily the same designation as other publications considering these sites. The MUSCLE data is SAS imagery with an along-track resolution of 2.5 cm and an across-track resolution of 1.5 cm. The images, considered in this report, span the ranges 40 to 150 m.

The use of target templates to train a Haar Cascade detection method was described in [15]. The same basic method is used here, except that more care is taken to modify the basic template to make it appear more realistic. The basic template is generated by ray-tracing [9–10] for a specified target type, range, and aspect. The target types considered are (1) wedge-shaped (2) truncated-cone and (3) cylinder. We use a precomputed library of templates for a set of discrete aspects and ranges. The templates were computed for an altitude of 13 m. If we wish to insert a template in the image at the range \( r \), for an altitude \( a \), we use the precomputed template closest in range.
to \( r_t \) where

\[ r_t = (13/a)r. \]  \hspace{1cm} (1)

This template will have approximately the correct shadow length for the altitude but it is inserted at the range \( r \). The image column index which corresponds to the range \( r \) is used as the starting column for the template insertion. Thus, the range \( r \) refers to the leading edge of the template which also corresponds to the leading edge of the highlight region. The template is a rectangular window consisting of a highlight region (positive values), a background region (zero value), and a shadow region with values of negative one (the acoustic shadow cast by the object on a flat seabed).

It is desired that the inserted template should blend in with the surrounding imagery. This is done in the following manner. There are the original image (background) values \( B(i, j) \) corresponding to the rectangle of the real image where the template is to be inserted. For the template insertion, the template’s shadow and highlight regions are used instead of the corresponding values of \( B(i, j) \) in the synthetic imagery. Where the template is zero (i.e. designated seabed) the original \( B(i, j) \) values are used.

In the model of this report the templates can be resized (using MATLAB \texttt{imresize} with nearest value interpolation) in the along-track and across-track directions by random amounts between 90-110%. This is done randomly 20% of the time. The resulting template may also be flipped vertically with a 50% probability. In addition, the boundary pixels between the shadow and the background region are randomly assigned to be background, thus making the shadow/background boundary imperfect. This is done by convolving a \( 7 \times 7 \) filter of ones with the shadow region. For the resulting convolved image, those shadow pixels with a convolved value of greater than -49 are deemed to be shadow-edge pixels and there is a 50% chance that they are set to zero, indicating a background pixel.

For each real target in a sonar image, a synthetic target is inserted at approximately the same range but at a different random along-track position. A numerical check is performed so that the insertion point has a pixel distance of at least 300 from all other real and synthetic targets’ insertion coordinates to significantly reduce the chance of image overlap between targets. If the insertion coordinates fail the distance check, then another along-track position is randomly chosen. For the density of targets in the sonar images of this report this strategy proved successful.

Once all the targets have been inserted into the sonar data, this data is then processed in a standard manner for subsequent ATR processing. To produce the image files (TIFF format) that we use for the ATR methods, the data is converted into decibel units. This data is then two-dimensionally median filtered (a \( 3 \times 3 \) window) and then subsampled by a factor of 2 in the along-track direction and 5 in the across-track direction. This image is then normalized by the mean column values and then by the
mean of a moving $150 \times 150$ window. This normalized image is then converted into a TIFF format image by mapping the interval between the 0.5 and 99.5 percentiles into $[0,255]$ and the outside values mapped to 0 and 255. In this report, it is these TIFF images which are used by the ATR algorithms. The variations in overall levels are often lost in sonar images because of image normalization. However, we do the target insertion prior to any image normalization and thus, if for example, the normalization increases the overall amplitude of the image pixels, including the noise, this will be properly accounted for in the simulation.

Below, we describe in detail how the original template’s highlight and shadow regions are modified. There were a few basic concepts that were employed in modifying the template before its insertion into a sonar image: (1) randomize somewhat the pixel values in the synthetic highlight region (2) add noise into the shadow regions - the statistics should depend upon the local statistics of the surrounding images and also on the range of the target insertion and (3) make the shadow/background boundary somewhat “ragged”. Many of the values of the modeling parameters were determined by numerical trials and visual inspection of the resulting images. If for a set of parameter values, a number of resulting synthetic images appeared unrealistic, the parameters in the target-insertion code were then adjusted and the process repeated until we were happy with the visual results. It was also determined through numerical experiments that the synthetic images yielded reasonable ATR performances for all the sites. The precise values of these parameters are not unique - similar results would likely be obtained for other sets of similar values. It should be noted that once determined, the parameters of this report were used for all images and for the three different sites. There will be only one parameter $\gamma$, discussed later on, which is site-specific.

## 2.1 Highlight generation

The highlight of the template has a backscattering value which depends upon the cosine of the angle between the incident ray and the surface normal of the target. The original highlight region of the template is smoothed with a $9 \times 7$ filter of ones. For the sonar considered in this report, a 1 m long object is 40 pixels long in the along-track direction. This smoothing causes the highlight region to extend somewhat beyond its mathematical boundaries. A template highlight value is mapped to a sonar image value in the following way. First, the highlight values are considered relative to the mean background level of the trial site under consideration. This is done so that the highlight echo levels are consistent with sonar return levels for the site. In order to determine the appropriate overall echo level, an amplitude/range curve is computed for each sonar file used for the target insertions. Then for each of the sites, a mean curve is computed from the individual curves. For each of these mean curves, the maximum value $e_{\text{max}}$ (in dB) is determined. A function $E_s(r)$, representing a typical
highlight strength as a function of the target range is then defined as

$$\log_{10} E_s(r) = (e_{max} + 15 - 40 \times 0.4 \log_{10}(r/40))/20$$  \hspace{1cm} (2)$$

This formula is used to model the decay of the target echo with range due to geometrical spreading and attenuation. The factor 40 in the term \(r/40\) in Eq.(2) is because 40 m is the initial range for the sonar images used in this report. Some of the values in Eq.(2) were adjusted (e.g., the factor 0.4) so that the targets’ highlight levels as a function of range agreed empirically with what was observed in the sonar images. The basic template’s highlight values need to be converted into image pixel values which are consistent with the sonar image values. Let us denote the maximum highlight value of the basic template as \(t_m\). Then the template values, \(h_v\) are redefined as new values \(\tilde{h}_v\),

$$\tilde{h}_v = E_s, \quad h_v \geq \alpha_v t_m$$

$$= h_v / (\alpha_v t_m) E_s, \quad h_v < \alpha_v t_m$$  \hspace{1cm} (3)$$

where \(\alpha_v\) is a random number chosen from a uniform interval. Thus the modified highlight has a randomized amplitude. However, the values tend to be higher where the original template value is high. For the ATR methods used in this report, the original sonar images are subsampled by a factor of 2 in the along-track direction and by a factor of 5 in the across-track direction (after median filtering). In order for the randomization approach, just described, to have more of an effect, we apply Eq.(3) to blocks of pixels 3 \times 5 in dimension. Finally, the highlight values are redefined to be the maximum of the highlight value and the underlying background image value. This is done so that at the edges of the highlight region, where the levels maybe very low, the inserted target image does not have a band of values which are less than the surrounding image.

### 2.2 Shadow noise generation

For the highlight region, it was described how the highlight amplitude was randomized and the overall intensity level diminished with range. For the shadow region, Rayleigh-distributed noise will be added. This distribution is defined in terms of a single parameter \(\beta\). For the target insertion, we wish to determine a good value of \(\beta\) to use for our shadow noise generation on the basis of the target insertion range and the surrounding image statistics. The basic concept for determining \(\beta\) is to consider the region of the original sonar data where the target is to be inserted—a window of 123 \times 603 centred at the insertion location. We first convert this imagery into uint8 values with lower and upper limits at the 0.5 and 95\% values for the pixel intensity values within the window. The pixel values of this image are then segmented into 3 regions (2 thresholds) using the MATLAB routine \texttt{multithresh} —the smaller of the two thresholds is denoted as \(\tau_s\) and this is converted into the corresponding threshold.
for the original data. The conversion step into integers is not necessary as the `multithresh` routine can work with floating point numbers, but we found this approach gave slightly better results. It is not clear why. Possibly, the conversion of the values into the range $[0,255]$ using the 0.5% and 95% percentile values of the image pixel values to define the lower and upper bounds mitigated the effects of outlier values. In any event, since this process seemed to yield slightly better results, it was used.

For the shadow pixel values, those pixels with a value less than $\tau_s$, a Rayleigh distribution is fit using a MATLAB routine from the Statistics toolkit. The resulting estimate of $\beta_e$ gives an estimate of the distribution in the local shadow regions. However, it was found from visual inspection of the resulting images, that this value of $\beta_e$ was often too large and the resulting synthetic shadows too noisy. For the portions of the image containing ripples this estimate seemed to yield good results. In Figs.1a-1c, representative sections of imagery are shown and beside them, in Figs.1d–f, the corresponding histograms of the shadow pixel values and the fit Rayleigh distributions. It can be seen that the histogram of Fig.1e is abruptly truncated at the last value and that the probability of the second-to-last bin is more than 1/2 that of the maximum probability. The histograms for the rippled images, Figs.1d and 1f, have a much lower relative level for the higher-value bins. Empirically, we found that for cases such as Fig. 1f where the amplitude of the highest bins is less than or equal to 0.25 of the maximum value (or the bin ratio $b_r \leq 0.25$) then the fit-value of $\beta_e$ was a good value to consider as the noise parameter for the shadow. However, since $b_r$ is not always less than 0.25, a modified strategy is utilized.

We define a range $r_c$ after which the noise level in the shadows is taken to increase quadratically with range. The formula used for a range-dependent Rayleigh parameter $\beta_r$ is

$$
\beta_r = \begin{cases} 
  a_0 \tau_s + c_r & r \leq r_c \\
  (a_0 + (1 - a_0)(r - r_c)^2/\gamma^2)\tau_s + c_r & r > r_c 
\end{cases} 
$$

Through numerical experimentation, it was found that a value of 0.22 for $a_0$ in Eq(4) yielded good results in the resulting images. The term $c_r$ is zero for $r \leq r_c$ and for $r > r_c$ is defined by the mean amplitude value $m_{100}$ of all the pixels contained within the last 100 columns of the sonar data, $c_r = m_{100}/20$. The parameter $\gamma$ determines how large the Rayleigh parameter is relative to the shadow/background threshold $\tau_s$ at the maximum range of the sonar image. In this report, the maximum range is 150 m and for all 3 sites $r_c = 95$ m (which is the mean range for the sonar images considered). The value $\gamma = 55$ corresponds to $\beta_r = \tau_s$ for the maximum range of 150 m. A larger value of $\gamma$ corresponds to relatively less noise in the shadow at the far ranges. A smaller value, such as $\gamma = 40$ which is used for Site A, means that the Rayleigh parameter $\beta$ will exceed $\tau_s$ for the longer ranges. This results in the shadow levels being close to the surrounding background values. The value of $\gamma$ is
Recall the histogram bin ratio $b_r$. If $b_r \geq 0.25$ then the value of $\beta$ which is used in the Rayleigh noise generation, $\beta_R$, is $\beta_R = \beta_r$ where $\beta_r$ is defined in Eq.(4). For $b_r < 0.25$ then the value used is $\beta_R = \min(\beta_r, \beta_e + c_r)$ where $\beta_e$ is the estimated value from the shadow pixel values. For all shadow noise distributions, if the mean value of shadow noise, $m_s$ is greater than $0.75 \nu$ where $\nu$ is the median value of the background image window, the noise is rescaled by

$$\tilde{n}_i = (0.75 \nu / m_s)n_i$$

where $n_i$ denotes the $i$th shadow noise value. This rescaling is done to ensure that the shadow regions will not appear higher in amplitude (brighter) than the surrounding background in the resulting sonar images. Both the generated Rayleigh noise (scaled by 0.9) and a smoothed version of the local background image (scaled by 0.03) are added to the original shadow region of the template. The addition of a
smoothed version of the background image is to simulate that for SAS images some of the background reverberation can enter into the shadow through the beamforming process.

3 Detection methods

The synthetic imagery will be used to generate ROC curves for specified Automatic Target Recognition (ATR) methods. In particular, we will consider a basic matched-filter detector [14] and a Haar-cascade detector [12-13]. We now briefly describe these methods. More details can be found in Refs.[12-14].

3.1 Matched-Filter

The sonar image is segmented into negative (relatively low pixel values) and positive values (relatively high pixel values) [14]. In this report, the image is rebinned into the discrete values \([-1.0, -0.5, 0.0, 0.5, 1.0]\) based upon 4 specified percentiles of the pixel values. A template of ones (for the highlight) and negative one (for the shadow) is then used to represent a generic target-signature. The length of the sonar shadow region for a scattering object of fixed height increases as a function of the range. This can be emulated in the matched-filtering by allowing the across-track length of the template’s shadow region to increase appropriately as a function of range. The matched-filtering is done using the method of integral images [13]. This method starts by forming a two-dimensional cumulative sum of the image; sum the image cumulatively along the columns and then sum the resulting image cumulatively along the rows. Then it can be easily shown that a sum of image pixels within a rectangular window is equal to a simple combination of the values at the corners of the rectangle in the double cumulative sum. Thus the large number of additions within the original rectangular window which would be required by straightforward filtering is reduced to 3 operations. This approach is even computationally efficient when compared to FFT convolution methods. The integral image method is applicable to our matched filtering where the template is composed of two rectangles with constant coefficients (one and negative one).

3.2 Haar cascade

The matched-filter detector has various parameters which can be varied but it is not trained with sonar data. The Haar cascade is a trained detection method. It is trained with a set of closely cropped images of targets and a set of images (likely full images) with no targets. A Haar feature is the output of cross-correlating various simple filters, consisting of rectangles of ones and negative ones, with an image window.
Using the concept of integral images, discussed in the previous section, this filtering can be done very efficiently. There is a large set of different Haar filters and hence a large number of Haar features for an image window. The training determines a set of stages, each stage with a combination of simple Haar feature outputs and thresholds, which can differentiate image windows with clutter from those containing targets. The cascade concept is that the first stage can eliminate, for example, 50% of the clutter windows while retaining a very high percentage (e.g., 99.8%) of the target windows. The second stage, then considers clutter windows which are not eliminated by the first stage and is trained to reject 50% of these clutter windows, while still retaining a high percentage of target windows. The number of stages continues until one reaches a desired false alarm rate or a specified number of stages. In general, the early stages only require a small number of features to discriminate against clutter and as a detector these stages quickly eliminate most of the possible windows contained within an image. The latter stages use more features. The image windows which pass through all the stages are considered as valid detections. The training algorithm we use is from the OpenCV [17] library as is the cascade detection algorithm. The output from the OpenCV detection algorithm is a set of detection rectangles of varying scales. A local rectangle count is used as the detection threshold.

4 Synthetically generated images - numerical experiments’ results

We now consider the ATR performance for the real targets and for the targets synthetically inserted into the images. The real and predicted ROC curves are presented for each of the three Colossus trial sites. Because of the ripples and the variety of different backgrounds, the sandy rippled site (Site A) was the most difficult case for the ATR methods and will be the first case we consider.

4.1 Sandy rippled, cluttered site - Site A

In Figs.2a and 2b, representative images with real and inserted targets are shown. The concept was to place the inserted targets at the same range as the real targets and also to use the same type of target. However, the aspect angle used for the inserted target was random. Thus, for every real target in a sonar image, there is also a corresponding synthetic target at the same range. For this site, a value of 40 was used for $\gamma$ in Eq.(4) The synthetic inserts are marked in Fig. 2a with a cyan “S”. The insertions, for the most part, appear realistic in terms of the echo strength and the shadow contrast. There are some images and inserted targets for which this is not the case - the shadow noise may appear larger in amplitude than for the real targets in the image or vice versa. It is likely possible to tune various simulation parameters.
to improve these defects. However, on average, the agreement appears visually to be very good.

The ATR detection algorithms are run on a total of 65 sonar images, many with multiple targets. Detections are considered as target detections if they are sufficiently close to the known positions of the targets. The real target positions in the files are known from the ground truth files and the synthetic target positions are saved. The detections are associated with the closest target (real or synthetic) and then divided into those associated with the real targets and those associated with the inserted targets. Other detections are false alarms. One can then consider the real and synthetic target detections separately and generate two ROC curves by varying the detection threshold and determining how many target and false alarm detections exceed the threshold. In the case of the matched-filter, the detection threshold is the matched-filter output and for the Haar detector, it is the local number of rectangles associated with the detection. A perfect ROC curve is one which reaches a probability of detection of one with zero false alarms and then is identically unity as a function of the number of false alarms. In practice, we would like to achieve a probability of detection of greater than 90% for a small number of false alarms. However, this is not always possible. In general, the area under the ROC curve, is used as an indicator of the relative performance of the detectors – larger area being better. For our simulations with real and synthetic targets, the set of false alarms will be the same for the same ATR method, so that the synthetic and corresponding real curves can be compared in terms of which ROC curve has the larger area underneath. In Fig.3 the matched-filter results are shown. The ROC curve for the inserted targets shows poorer performance at higher false alarm rates compared to the detection performance on real targets – the detection rate is offset by approximately 10% for the high detection rates. The agreement between the 2 curves is very good in the region of low false alarm rates (e.g. <100).

In Fig.4 the ROC curves for three Haar cascades are shown. These cascades were previously trained using a portion of Colossus data, or for the curves denoted with green and maroon diamonds, using data from an entirely different trial (MANEX‘13). For the Haar cascade detection methods, the ROC curves for the inserted targets are somewhat optimistic. In the case of the best performing cascade for this site, the agreement between the real-target (blue line) and synthetic curves (red line) is good. The blue and red diamond curves illustrate another Colossus trained cascade for which the real-target performance is poorer but this is not reflected in the predicted performance for the synthetic targets. The final curves (green and maroon diamonds) are for a cascade trained with data from the MANEX’13 trial – in this case the performance on the complex seabed background imagery of this Colossus site is poor, but both the real and synthetic curves show this poor performance. The explanation of the predicted and real performances for the Haar cascades at this site is not completely understood. We conjecture that for a cascade with a good detection rate
for the real data (e.g. not too many stages) then the real and synthetic curves show quite good agreement. In the case, that the cascade has too many stages (likely the case for the blue diamond curve) then the synthetic curve is overly optimistic. It is not obvious what features a Cascade is using for the detection - it may be that the synthetic targets, despite our attempts to degrade them, still possess too many of the features which are consistent with the latter stages of a cascade detector.

For the real/simulated target test set, we use all the files from the site with targets in it. Approximately half of these targets (along with targets from the other sites) were used in the Cascade training process for the Colossus-trained cascades. This is not relevant in the case of the matched filter, since this method is not trained, or the MANEX’13 trained cascade which was trained with an entirely different data set. One might think that the test would be fairer if one used only images not in the training set for the cascade. However, we find that the ROC curves for the entire set of images and those from just the proper testing set are quite similar. The large set of images from the entire set gives a better statistical sampling for the synthetically generated images. In Fig. 5, we have redone the ROC curves using a testing set which excludes the files which were originally used in training the cascades. As can be seen the curves are quite similar in character to those of Fig. 4. For the remainder of this report, we will consider the entire set of images.

4.2 Flat Muddy Site - Site B

In order to insert targets into the background for this site, the value of $\gamma$ is set to 70, larger than the value used for the previous case of Site A leading to less relative shadow noise. Two example images are shown in Fig. 6. It can be seen that the synthetic targets appear to have the correct noise characteristics. The seabed is uncluttered; however, there are linear features which are prominent. The deterioration of the shadow/background contrast at the longer ranges is evident and is accurately simulated by the inserted targets. The ROC curves, corresponding to the matched-filter output are shown in Fig. 7 and the curves for the Haar Cascades (the same cascades as for Site A) are shown in Fig. 8. There were a total of 96 files used in the analysis. In all cases, the ROC curves for the real and synthetic targets are similar. For the Haar cascades, all the cascades perform well, even the one trained with MANEX’13 data – although, it is a little difficult to see from Fig. 8, the largest difference between the real and synthetic curves occurs for this case where the synthetic prediction is somewhat optimistic. For all methods, the improved detection and relatively low false alarm rates for this site compared to the complex rippled site are accurately predicted.
Figure 2: Example images, after normalization and conversion to TIFF format, from Site A with real and inserted targets. In the top figure, the synthetic targets are indicated with a $S$. 
Figure 3: The ROC curves for a matched filter detector for real and synthetically inserted targets with the sandy, rippled backgrounds (Site A).

Figure 4: The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade.
4.3 Flat Sandy Site - Site C

This site was likely the easiest site from a minehunting point of view. In order to insert targets into the background for this site, the value of $\gamma$ is set to 110, which is significantly larger than the values used for the previous 2 sites (40 for Site A and 70 for Site B). Two example images are shown in Fig. 9 and it can be seen that the synthetic targets appear to have the correct noise characteristics. The background seabed is generally uncluttered. The ROC curves corresponding to the matched-filter output are shown in Fig. 10 and the curves for the Haar Cascades (the same ones as for Sites A and B) are shown in Fig. 11. There were a total of 86 images considered. In all cases, the ROC curves for the real and synthetic targets are similar. For the Haar cascades, all the cascades perform very well, even the one trained with MANEX’13 data. For all methods, the high detection and low false alarm rates for this site are accurately predicted.
Figure 6: Example images, after normalization and conversion to TIFF format, from Site B with real and inserted targets. In the top figure, the synthetic targets are indicated with a S.
Figure 7: The ROC curves for a matched-filter for the real and the synthetically-inserted targets for the flat, muddy site (Site B).

Figure 8: The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade.
Figure 9: Example images, after normalization and conversion to TIFF format, from Site C with real and inserted targets. In the top figure, the synthetic targets are indicated with a ‘S’.
**Figure 10:** The ROC curves for a matched filter for the real and the synthetically inserted targets for the sandy, benign site.

**Figure 11:** The ROC curves for various previously trained Haar cascades on images from sandy, rippled site. The blue and red lines and diamonds are the performance of cascades trained using Colossus data. The green and maroon diamonds are for the MANEX’13 trained cascade. Note the very different x-axis scale in this case.
5 ROC curve statistical and parameter modelling

For the three Colossus sites, good qualitative and quantitative agreement was found between the synthetic target ROC curves and the real target ROC curves. One can now use the synthetic target insertion to predict results for new scenarios or to predict the statistics of the performance predictions. For example, we create 21 sets of inserted targets for Site A. Each set follows the insertion rules described in Section 2 but because the insertion locations with respect to row, the target aspect, and the randomization used in template values varies with the simulation realization, the resulting ROC curves are somewhat different for each realization. The curves for the real targets also vary somewhat – this is because the inserted targets will effect the normalization of the imagery somewhat (recall that the synthetic targets are inserted into the sonar data before the subsequent normalization takes place). Also, an inserted target’s pixel values will replace the original sonar data in the window corresponding to the inserted target. This window in the original data may have contained or contributed to a false alarm detection(s). Thus replacing it with the inserted target window can result in a small variation in the false alarm statistics. In Figs. 12 and 13, the resulting curves for the real targets (blue curves) and the synthetic targets (red curves) are shown for the matched filter detector and for the first Haar cascade. There is a noticeable variation in the predicted curves. Interestingly, the variation is greater for the matched-filter ROC curves. The synthetic matched-filter curve of Fig. 3 corresponds to the upper portion of the multiple curves of Fig. 12.

6 Discussion of Results

A method for inserting artificial target images into real synthetic aperture sonar background images was presented. Care was taken to have the highlight return levels and the shadow noise levels appear realistic. In particular, the shadow noise levels were based upon the range of the target and the local pixel statistics. A window of the original image data about the insertion point was used to determine the local shadow threshold using an image segmentation algorithm and a range-dependant multiple of this threshold was then used to generate Rayleigh-distributed shadow noise. The range dependency was characterized by a specified range $r_c$ which defines where the shadow noise starts to grow, a quadratic scalar $\gamma$ which defines how the noise grows in level for ranges greater then $r_c$ and a constant term which is based upon a specified fraction of the mean pixel level for the maximum ranges. In this report, only $\gamma$ was varied between the 3 sites. Given a knowledge of the seabed type and the water depth, it should be possible to use a sonar/propagation model to predict good values for these parameters. In this report, we simply chose values of $\gamma$ which led to
Figure 12: The ROC curves (Site A) for a matched-filter for 21 realizations of the inserted targets —real-target ROC curves in blue and synthetic-target curves in red.

Figure 13: The ROC curves (Site A) for a Haar Cascade (Cascade 1 in Fig. 2) for 21 realizations of the inserted targets —real-target ROC curves in blue and synthetic-target curves in red.
simulated targets with realistic levels of shadow noise for the particular site.

Overall the simulations were promising in terms of both qualitative and quantitative predictions. The fidelity of the synthetic ROC curves with respect to the ROC curves corresponding to real targets depended upon the site. For Site C, the targets were in a region of benign seabed with little clutter - all detection methods worked very well in this environment and the ROC curves were very well predicted by the target insertion method. Site B had a slightly more complex seabed but once again all detection methods worked well and the real ROC curves were well predicted by target insertion – the predicted matched filter performance was a little overly optimistic. Not surprisingly, the most difficult case to simulate was the most complex environment of Site A. At the other sites, the trial targets were, with a few exceptions, not difficult to detect so the character of the ROC curves is dominated by the number of false alarms as the detection threshold is varied. For Site A, there were real targets which were difficult or very difficult to detect. We managed to simulate the increased detection difficulty due to range and environment but it is difficult to simulate those occurrences where real targets are nearly impossible to detect because of the surrounding environment or beamforming issues. Also, the approach of this report was to place the synthetic targets at the appropriate ranges but in random positions along-track. On the other hand, the trial targets are at fixed absolute positions on the seabed for a given site. Particularly for a cluttered complex seabed, this could result in some statistical differences between the simulated and the real results.

It is interesting to note that for Site A, the predicted performance of a Haar cascade trained with data from another site was quite accurate. Also, the predicted performance of the matched filter was accurate for a small to moderate number of false alarms. The worst prediction was for the case of a cascade trained with Colossus data which did not perform well at Site A (Haar 2). In this case, the synthetic ROC curve was significantly overly optimistic. We surmise that this cascade was perhaps over-trained (too many stages) for this data and that the synthetic data possessed too many of the features, relative to the real imagery, which were consistent with the latter stages of the cascade.

An issue for all sites is that the real target images sometimes suffer from imperfect SAS beamforming – they may appear unfocussed or have sidelobe energy. This is also difficult to capture in a simulation without going to a more complicated motion/beamforming simulation.

We will continue to investigate this type of performance modelling. The simulations, in terms of noise and degrading the templates, can likely be improved. For synthetic aperture sonars, a measure of the correlation between the time series from successive pings of the sonar is a good predictor of sonar performance [5]. This could perhaps be used in the synthetic modelling. For systems which also yield bathymetry, the
bathymetry could be used to improve the fidelity of the inserted target. For example, the shortening (lengthening) of the shadow region for an upslope (downslope) region could be accounted for in the template insertion. We would also like to use the same methodology for other sonar systems including conventional sidescan sonars such as the Marine Sonic, Klein, and Edgetech sonars.
References


| 1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor’s report, or tasking agency, are entered in section 8.) | 2a. SECURITY MARKING (Overall security marking of the document, including supplemental markings if applicable.) |
| DRDC – Atlantic Research Centre PO Box 1012, Dartmouth NS B2Y 3Z7, Canada | UNCLASSIFIED |
| 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC DECEMBER 2013 |
| 3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) | Predicting synthetic aperture sonar performance with augmented sonar imagery |
| 4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.) | Fawcett, J. A. |
| 5. DATE OF PUBLICATION (Month and year of publication of document.) | March 2017 |
| 6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.) | 34 |
| 6b. NO. OF REFS (Total cited in document.) | 17 |
| 7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) | Scientific Report |
| 8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) | DRDC – Atlantic Research Centre PO Box 1012, Dartmouth NS B2Y 3Z7, Canada |
| 9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) | 01cb |
| 9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.) |
| 10a. ORIGINATOR’S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) | DRDC-RDDC-2017-R020 |
| 10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.) |
| 11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) (X) Unlimited distribution |
| ( ) Defence departments and defence contractors; further distribution only as approved |
| ( ) Defence departments and Canadian defence contractors; further distribution only as approved |
| ( ) Government departments and agencies; further distribution only as approved |
| ( ) Defence departments; further distribution only as approved |
| ( ) Other (please specify): |
| 12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.) | unlimited |
The prediction of synthetic aperture sonar/sidescan sonar detection performance for a specific sonar survey is a topic of much interest. There are many ways to define performance. In this report, the detection and false alarm rates of specified automatic target recognition (ATR) methods for minelike objects will be the performance statistics of interest. Previous researchers have investigated the expected sonar performance as a function of image and sonar features. The approach taken in this report follows more closely the work of other authors on augmented reality. Synthetic target images are inserted into real sonar images obtained from a specific survey. Thus, the background sonar imagery and the clutter represents a survey exactly. The issue is making the inserted target realistic, in terms of its distortions and the noise characteristics of its shadow and highlight regions. These features need not be exactly correct but they need to be sufficiently correct to yield the correct ATR performance and behaviour. This report will describe the approach taken to insert targets with noise statistics which vary with respect to range and the surrounding image statistics. The approach will be applied to sonar data from a NATO Centre for Maritime Research and Experimentation (CMRE) trial (three different sites) and the predicted ATR performances will be compared to the actual performances. In addition, the performance model can be used to predict performances as a function of other parameters, such as target type, range, etc for which the trial data is limited. Also, the uncertainties in the predicted performance can be estimated through Monte Carlo simulations.

automatic target recognition; sonar; sidescan