Supervised target detection and classification by training on augmented reality data

E. Coiras, P.-Y. Mignotte, Y. Petillot, J. Bell and K. Lebart

Abstract: A proof of concept for a model-less target detection and classification system for side-scan imagery is presented. The system is based on a supervised approach that uses augmented reality (AR) images for training computer added detection and classification (CAD/CAC) algorithms, which are then deployed on real data. The algorithms are able to generalise and detect real targets when trained on AR ones, with performances comparable with the state-of-the-art in CAD/CAC. To illustrate the approach, the focus is on one specific algorithm, which uses Bayesian decision and the novel, purpose-designed central filter feature extractors. Depending on how the training database is partitioned, the algorithm can be used either for detection or classification. Performance figures for these two modes of operation are presented, both for synthetic and real targets. Typical results show a detection rate of more that 95% and a false alarm rate of less than 5%. The proposed supervised approach can be directly applied to train and evaluate other learning algorithms and data representations. In fact, a most important aspect is that it enables the use of a wealth of legacy pattern recognition algorithms for the sonar CAD/CAC applications of target detection and target classification.

1 Introduction

Fast development of underwater robotics has lead to the possibility of surveying large areas of the seafloor. Remotely operated vehicles and autonomous underwater vehicles (AUVs) are now routinely used for this task, generally equipped with different sets of sensors to image and analyse the seabed. Side-scan sonar is perhaps the most widely used imaging sensor, mainly because of its good compromise between resolution, range and cost.

A common application of side-scan sonar survey missions is target detection and classification for mine countermeasures (MCM). Processing and analysis of the products of these surveys (typically consisting of several hundreds of images per mission) needs to be automated both for practical and operative reasons, especially if used in a real MCM campaign. Automatic processing also opens the possibility for AUVs to adapt the survey plan to the type of objects being observed in real-time.

The task of target detection and classification in side-scan images can be approached in different ways. Model-based methods [1] rely on the development of specific feature extractors, specifically devised to identify characteristic elements of the target geometry or appearance. This *ad hoc* procedure must be performed for every type of target and leads to target-specific detectors, which limits the types of objects that can be detected to those that have been previously modelled. Moreover, the occurrence of new conditions or new types of targets will require re-modelling the detectors or building new ones from scratch.

If a large amount of target examples are available, supervised classification methods [2] are a more practical solution. A classifier can be trained to detect and classify targets using these examples, and new conditions or new types of targets simply require a re-training of the system using new examples. A range of techniques have been considered including nonlinear matched filters [3], thresholding and clustering [4], adaptive filters [5] and template matching [6]. The success of these techniques depend on the similarity of the training data to the test data, and poor results are often observed when the difference between the two is high [4]. It has also been illustrated that their success depends on the data used to train the system [7].

To overcome the limited availability of real target images, synthetic ones could be generated through simulation [8]. Semi-supervised approaches [9] have also been proposed, but these rely upon the availability of an underwater vehicle. The features and characteristics of man-made objects are well known and lead to accurate simulation of synthetic target models. However, general simulation of the seabed is ineffective because modelling the diversity of the natural seafloor is intractable. System noise, environmental inhomogeneities and image artefacts affecting side-scan sonar images in real operational conditions are also difficult to model accurately.

A compromise between using real and synthetic data can be found in augmented reality (AR) simulation [10], where synthetic target models are embedded on a real image of the seafloor. A computer model of the seafloor is constructed from the side-scan image by an inversion process [11], which determines the parameters that characterise the observed scene. Then the computer model for the seafloor and that of the target are combined and rendered to obtain

doi:10.1049/iet-rsn:20060098

Paper first received 13th July and in revised form 28th September 2006

The authors are with the Ocean Systems Laboratory, Joint Research Institute for Signal & Image Processing, School of Engineering and Physical Sciences, Heriot-Watt University, EH14 4AS, Edinburgh, UK

E-mail: coiras@nurc.nato.int

a new AR image that realistically integrates the synthetic target within the observed scene. This way, training images for a future mission can be generated from previous side-scan imagery of appropriate types of seabed and the computer models of the targets to be expected.

The AR technique opens a wealth of opportunities. In this paper, we investigate its applicability to the design of a supervised classification system for automated target recognition (ATR).

2 System description

An overview of the architecture for the proposed supervised classification system based on AR is presented in Fig. 1. A classifier is trained by using a large database of small AR image samples or 'snippets', which are mug shots of a small area of the seabed that may or may not contain a target. Ground-truth is associated to each snippet in the form of a data file that contains the type of snippet (target positive or negative), the type of the local seabed (flat, rippled etc.) and the type of target if present.

Once the classifier is trained, it can be evaluated by classifying a previously unseen test set of AR samples. Comparison of the classification results to the ground-truth gives an estimation of the classifier's performance on unseen data. The trained classifier can also be deployed and used on real data by feeding it with test target samples from a real mission. Note that the classifier evaluation method using an AR database can be used to realistically quantify the performance of any classifier, not only the supervised type presented in this paper, and therefore constitutes a general solution for the evaluation of underwater ATR systems based on side-scan sonar [10].

3 AR training database

The construction of the AR training database starts by selecting appropriate side-scan sonar images and 3D computer models of target geometry. The side-scan images should be representative of the type or types of seabed to be expected in the test data set or final mission, and ideally should have been acquired by the same model of sensor that will acquire the test data. The computer models for the targets should approximate the expected types of targets as accurately as possible. 3D computer models for the selected side-scan images are constructed by inversion of the side-scan images, which obtains a 3D elevation map, a reflectivity map and a beam-pattern map for each of them.

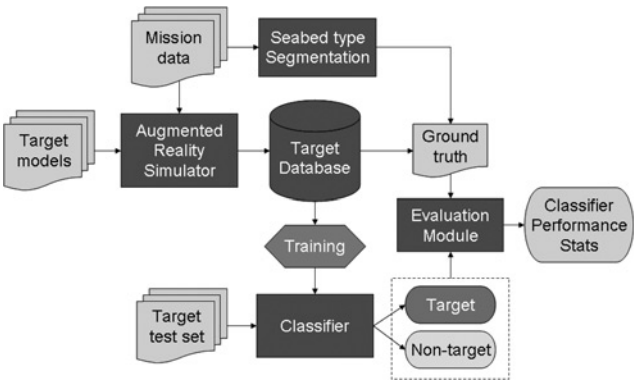

Fig. 1 *Supervised classification system setup*

The system is trained on an AR database, and then used to classify a test set of target samples

With the computer models for the images and the targets, a large number of simulated AR samples can be generated – with their associated ground-truth – and stored in the training database.

3.1 Sonar inversion

In order to generate a 3D computer model of an observed scene, it is necessary to analyse the side-scan image and try to determine what particular characteristics and properties of the seafloor and the observing sensor resulted in the formation of this image. We assume these characteristics to be represented by a set of three parameters per image pixel: seabed altitude Z of the point, reflectivity of the seafloor R at the point and intensity Φ of the illuminating acoustic pulse at the point. The sonar inversion process is the means by which we are able to reconstruct these sets of parameters for a given side-scan image. A more detailed explanation of this procedure can be found in the work of Coiras *et al.* [11].

In order to perform the inversion, a model for the side-scan image formation process is required. We have chosen the classical Lambertian diffuse illumination model [12–14] which assumes that the observed intensity I at a surface point p depends only on the incidence angle θ to the source

$$I(p) = K\Phi(p)R(p)\cos(\theta(p)) \tag{1}$$

where K is a normalisation constant. Assuming a coordinate system centred in the side-scan sensor (Fig. 2), the cosine of the angle of incidence in (1) can be expressed as a function of the surface elevation Z

$$I(x,y) = K\Phi(x,y)R(x,y) \frac{Z(x,y) - x(\partial Z/\partial x)(x,y)}{\sqrt{x^2 + Z^2(x,y)}\sqrt{(\partial Z/\partial x(x,y))^2 + (\partial Z/\partial y(x,y))^2 + 1}} \tag{2}$$

Equation (2) provides a forward model, which permits the rendering of a synthetic side-scan image once the scene parameters R, Z and Φ are known. The inverse problem, that of obtaining the scene parameters from a given real side-scan image, is much more complex and requires the utilisation of statistical optimisation techniques. Using the approach proposed by Coiras *et al.* [11], we can solve the optimisation problem by selecting an adequate initialisation for R, Z

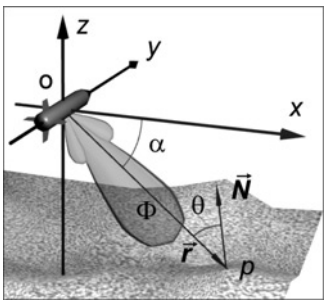

Fig. 2 *Geometry of the side-scan sonar imaging process, showing a coordinate system centred at the sensor*

A seafloor point p is insonified by an anisotropic acoustic pulse of intensity $\Phi(p)$
The intensity scattered back to the sensor will depend on the incidence angle $\theta(p)$ formed by the direction of illumination $\boldsymbol{r}$ and the normal $\boldsymbol{N}$ at point p

and Φ and iterating the following formulas

$$R(x,y) \leftarrow R(x,y) + 2\lambda \frac{\hat{I}(x,y)}{R(x,y)}\Big(I(x,y) - \hat{I}(x,y)\Big)$$

$$\Phi(x,y) \leftarrow \Phi(x,y) + 2\lambda \frac{\hat{I}(x,y)}{\Phi(x,y)}\Big(I(x,y) - \hat{I}(x,y)\Big)$$

$$Z \leftarrow Z - 2\lambda\hat{I}(I-\hat{I})\Bigg(\frac{-(\partial Z/\partial y) - (\partial Z/\partial x)}{1 + (\partial Z/\partial y)^2 + (\partial Z/\partial x)^2} + \frac{1+x}{x(\partial Z/\partial x) - Z} + \frac{Z}{x^2 + Z^2}\Bigg) \quad (3)$$

where λ is a constant that controls the rate of change, and is typically set to a value of 0.25.

Application of this approach using a multi-resolution implementation [11] yields an estimation of the scene parameter maps R, Z and Φ. An example of the results is presented in Fig. 3, which shows an original side-scan image and the three parameter images R, Φ and Z obtained from it.

3.2 Generation of AR images

Once the Z^*, R^* and Φ^* maps from a side-scan image have been estimated, it is possible to realistically introduce simulated mines by locally modifying these maps according to the Z_T and R_T maps that compose the synthetic model of the target (Fig. 4) in order to obtain two new elevation and reflectivity maps Z' and R'.

After calculating Z', the beam-pattern map Φ has to be recomputed in order to account for the changes in elevation, since $\Phi(x, y)$ is the projection of the sonar beam-profile on the seafloor. This is easily done by obtaining the beam-profile function $\Phi(\alpha)$ which gives the intensity for each insonification angle $\alpha = \tan^{-1}(z/x)$. This beam-profile is then re-projected to obtain the new beam-pattern map Φ'.

Once the modified Z', R' and Φ' maps are obtained, (2) can be used to render the AR image. Unlike other simulators that paste a simulated mine on top of an existing image, this approach enables the modelling of the interactions between the topography of the seabed and the mine. For instance, if a mine is placed behind a 3D structure, it should not be visible. The length of the projected shadow should also depend on the local elevation of the mine.

An example AR image generated using this approach is shown in Fig. 5, where some synthetic targets have been added to a given side-scan image.

3.3 Construction of the training database

Using the AR simulation process described in the previous section, a large number of training samples can easily be generated. Target 3D models are scattered over an inverted side-scan image computer model, and then the final AR image is rendered (Fig. 5). From this AR image, positive

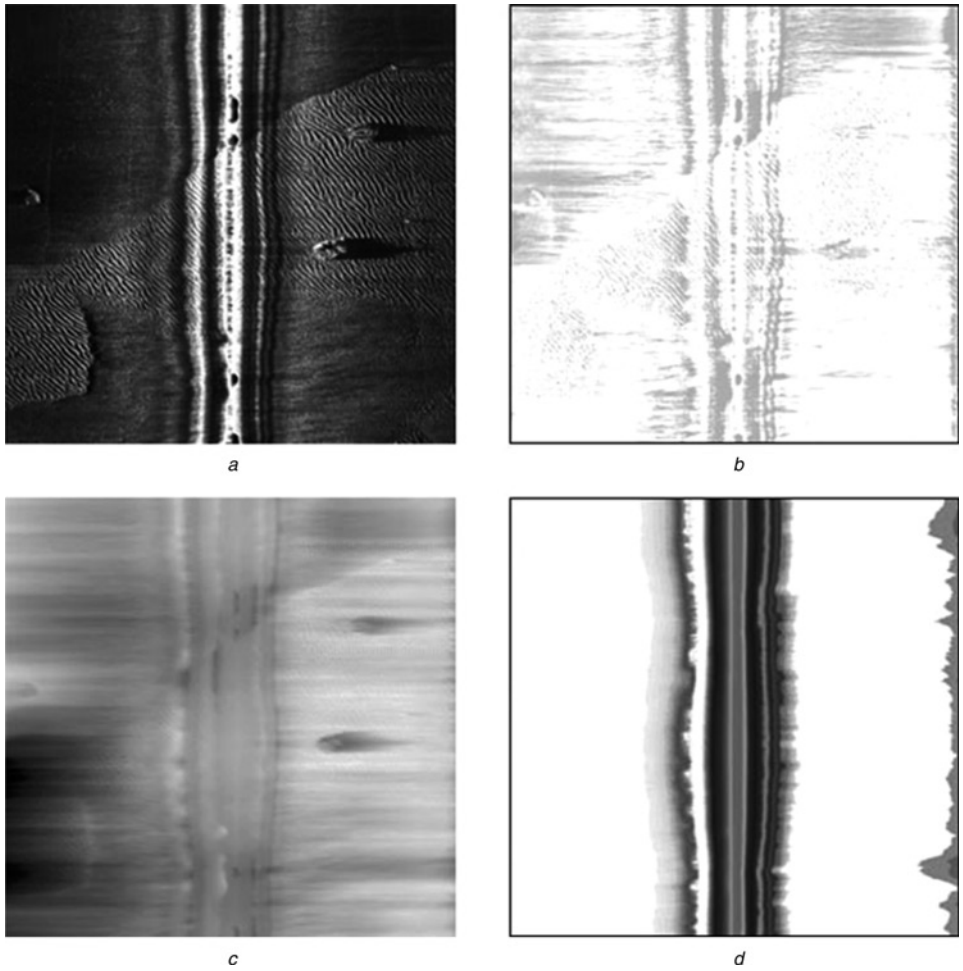

Fig. 3 *Results of the sonar inversion process*

a Original side-scan image
b Reconstructed seabed reflectivity map
c Reconstructed seabed elevation map
d Reconstructed sonar beam-pattern map

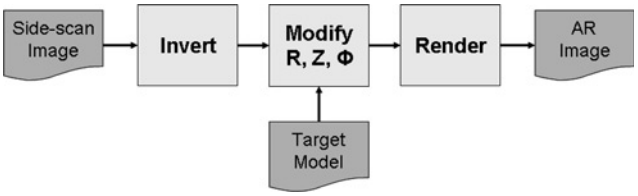

Fig. 4 *AR system*

and negative training image snippets can be extracted, as the positions of the targets are already known. A ground-truth file is then generated for each image snippet specifying the presence or absence of target, its type if present, its geolocation in the source image and so on. If an independent system to classify the seabed type [15] is also available then the type of local seabed can also be indicated in the ground-truth file for the snippet.

Since all the parameters involved in the generation of the training samples are controlled, training databases can be tailored to each particular theatre of operation. The statistical composition of the databases can also be controlled, specifying the relative quantities of targets and seafloor types as desired. A typical database will contain several thousand samples.

We have selected a fixed size of 2.5 m by 2 m for our snippet images to accommodate targets of a typical size of 1 m in diameter. Snippet resolution will vary according to the resolution of the side-scan sonar used in the mission. Three kinds of targets are typically available in our databases: cylinders (type A), truncated cones (type B) and trapezoidal (type C). More target models can easily be added if necessary, simply by creating their Z_T and R_T maps.

Some examples of positive and negative training snippets for a Marine Sonics system are presented in Fig. 6. Side-scan image resolution is 0.058 m across-track by 0.12 m along-track, which results in a snippet size of 44 by 17 pixels.

4 Detection and classification

The goal of an ATR system based on side-scan is to analyse a digital side-scan image and determine if a target or targets are present (detection). Some ATR systems might also be able to discriminate the targets by their particular types (classification). A general classifier is able to perform these tasks if provided with enough samples and an adequate feature extractor. If the classifier only discriminates between the target and non-target classes, it will effectively operate as a target detector.

A general data classifier typically works on numerical feature vectors which are computed from the image samples by application of a feature extractor. We have designed a purpose-made set of feature extractors for target classification in side-scan sonar images, which we refer to as the central filters. They are presented in the next subsection of the paper.

In a supervised classification system, a classifier is trained to discriminate between different object classes by presenting it with a vast set of examples. Application of the feature extractor to the training AR database entries will result in a large set of feature vectors and corresponding ground-truthed classes. The goal of the training procedure is to make the classifier estimate the best possible mapping between the feature vectors and their corresponding classes. After training, it is expected that the classifier will be able to generalise, upon presentation of

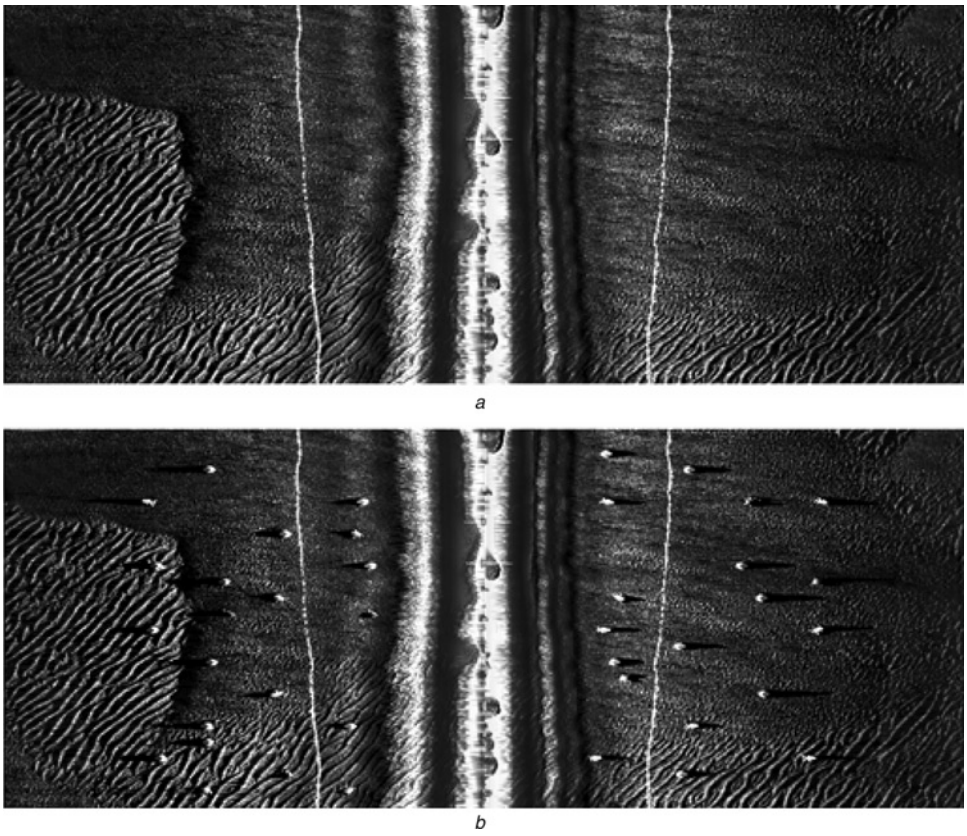

Fig. 5 *AR example*

a Original side-scan image
b Result after addition of synthetic targets to the scene
Notice how the targets cast shadows realistically, depending on their relative positions to the sonar source

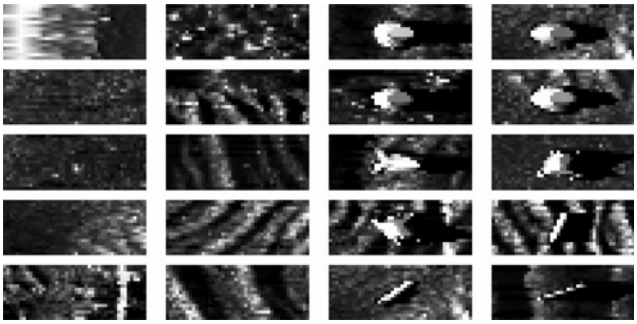

Fig. 6 *AR database snippets for a marine sonics side-scan sonar of 0.058 m by 0.12 m resolution*

Left columns: negative samples showing different seafloor types
Right columns: positive samples, showing cylinders, truncated cones and trapezoidal target shapes

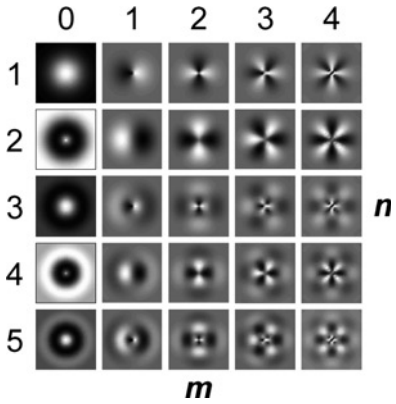

Fig. 7 *Central filters for wave numbers* $n = \{1, \ldots, 5\}$, $m = \{0, \ldots, 4\}$

previously unseen samples. If the interest is in detection, the output classes will simply be target or non-target. In the case of classification, the full set of target classes corresponding to the type of objects in the database will be used.

4.1 Feature extraction using central filters

Supervised classification is used on this database to discriminate between empty snippets and the three classes of targets which were added through the use of the AR simulator. Since each snippet contains several hundred pixels, the dimensionality of the problem is reduced using a feature extractor. The number of features is typically reduced to 16.

Although many filters exist that could be used as feature extractors for target classification (e.g. 2D extension of the wavelet packets used in the work of Sadjadi *et al.* [16]), a special set of filters has been specifically developed for target discrimination in side-scan sonar images. These central filters have been engineered with the idea that the following characteristics need to be accounted for when detecting a target.

- Presence: the filter must differentiate whether an object is present in the image.
- Highlight-shadow dichotomy: objects of interest in side-scan images typically appear as a bright highlight followed by a dark shadow. The filter has to be sensitive to this particular characteristic.
- Details: once an object has been identified, the only way of differentiating a target and a non-target is by examining the particular details that compose its image.

Presence of objects can be determined by using a function with a localised response, such as a Gaussian bell shape. Sinusoidal functions are specially suited for responding to alternating patterns such as the highlight-shadow characteristic. Details can be targeted by modification of orientation and wave numbers of sinusoidal functions.

These desirable properties suggest the use of functions much in the way Gabor filters and Spherical Harmonics work. Inspired in those types of functions, we have designed the following spatial filter function

$$F_{n,m}(x,y) = \mathrm{e}^{-(x^2+y^2)} \cdot \cos\left(n\sqrt{x^2+y^2} + \frac{1+(-1)^n}{2}\frac{\pi}{2}\right) \times \cos(m \arctan(y,x)) \quad (4)$$

where the integer wave numbers $n \geq 1$ and $m \geq 0$ control the spatial frequencies of the radial and angular components, respectively, and where the function arctan is the four-quadrant inverse tangent in the range $\{-\pi, \pi\}$. Fig. 7 shows the form of the central filters for wave numbers $n = \{1, \ldots, 5\}$, $m = \{0, \ldots, 4\}$.

The central filters are used by direct correlation with the target image. Each of the filters will return a number indicating the response of each particular filter shape to the local image region to which the filter is applied. The vector of central features is just the collection of the responses of the filters at a particular location in the image, and is computed as follows

$$V_{N,M}(I, x_0, y_0) = \left\{ \sum_{x,y} \left[I(x - x_0, y - y_0) F_{n,m}(x,y) \right]; \; n = 1, \ldots, N, m = 0, \ldots, M \right\} \quad (5)$$

We commonly use vectors of 16 features, corresponding to wave number ranges of $n = \{1, \ldots, 4\}$ and $m = \{0, \ldots, 3\}$. This number of features represents a good trade-off between the observed performance and the computational cost.

4.2 Supervised classification approach

For each snippet, the 16 central features are extracted. Supervised classification [2] aims at partitioning the feature space into sectors representing the different classes (e.g. target/nontarget). The partition is estimated during the training using ground-truthed samples generated with the AR simulator. This problem can be tackled with a wide range of existing techniques (e.g. support vector machine, neural networks). For this proof of concept, Gaussian classifiers are used. To apply the maximum log-likelihood principle, each class is modelled by a normal distribution. The estimation of the mean and covariance can be pooled or stratified [17], providing, respectively, linear and quadratic classification. These classifiers do not require tuning and give a straightforward evaluation of the selectivity of the features used.

The database of synthetic snippets is split into two sets. The first one is used for the training of the classifier. Its performance is estimated on the second set of unseen data. If the results are suitable (high probability of detection and low probability of false alarm), the trained classifier is then tested with snippets of real targets. This is the key step to validate our approach: the algorithm trained on semi-synthetic data needs to be able to generalise on real

data. Finally, the classifier is deployed on images of a real mission. The following section presents numerical results when this technique is applied to a real data set.

5 Results

5.1 Training performance

Cross-validation can be used to evaluate classification and detection performance. By training on a part of the snippet database and then testing on the remaining elements, the performance of the classifier when presented with unseen data can be estimated. Table 1 shows the average results for 20 runs when training with 66% of a database of 9150 elements and evaluating the performance on the remaining 33%. Results for detection with COTS (Matlab) linear and quadratic classifiers using 16 central features ($N = 4$, $M = 3$) are presented in Table 1.

Note that results in Table 1 include rippled and complex seafloors, which are normally avoided on mine-hunting missions. If we restrict the database to contain only samples located on flat seabed (3392 samples) the classification results are improved (Table 2).

Results for four-class classification using a linear classifier when restricted to flat seabed samples are presented in Table 3. This assumes that three different shapes of object were to be classified and these have been labelled as types A, B and C. Table 4 shows the results for a quadratic classifier.

Table 1: Classifier performance figures for Matlab's linear and quadratic classifiers on the training database, when using ($N = 4$, $M = 3$) central features, estimated by cross-validation (20 runs)

Results for training database	Linear classifier	Quadratic classifier
Probability of false alarm	**0.006 ± 0.002**	0.031 ± 0.003
Probability of detection	0.853 ± 0.016	**0.949 ± 0.008**

Bold indicates the best false alarm rate (lowest value) and best detection rate (highest) that have been found

Table 2: Performance figures as of Table 1 but when restricting the seabed type to flat (3392 samples)

Results for training database (FLAT)	Linear classifier	Quadratic classifier
Probability of false alarm	**0.001 ± 0.001**	0.031 ± 0.006
Probability of detection	0.894 ± 0.022	**0.975 ± 0.009**

Bold indicates the best false alarm rate (lowest value) and best detection rate (highest) that have been found

Table 3: Four-class classification confusion matrix for a linear classifier using ($N = 4$, $M = 3$) central features and restricting the samples to those on flat seabed

Linear	None	Type A	Type B	Type C
None	**0.87**	0.12	0.01	0.00
Type A	0.23	**0.63**	0.13	0.01
Type B	0.05	0.11	**0.71**	0.14
Type C	0.00	0.00	0.09	**0.91**

Total classification error (total number of misclassified samples divided by the total number of samples) is 0.19
Bold indicates the highest probability value for each row

5.2 Evaluation on real data

Performance estimation suggests the system is able to generalise, but a test when training with AR data and testing on real data is nonetheless required to validate the approach. To this end, a database has been constructed from a mission showing real targets. Fifty three different views of several objects of interest are contained in the database (16 type A samples, 20 type B and 17 type C). In addition, 53 empty snippets were created by selecting random pieces of seafloor where no targets were present. Table 5 shows the results for detection when training was undertaken on the same AR database of 9150 elements used above and then testing was performed on the real database. Note that these results are only illustrative, given the small number of targets that have been used to validate the system.

Interestingly, only 0.1% false alarms have been obtained, because of the seafloor being consistently flat for the imagery of the mission containing the real targets. Probability of detection is, however, noticeably smaller than the estimation obtained by cross-validation. This can be explained by the synthetic 3D target models not being very accurate representations of the real objects imaged in this mission. Most of the real targets had cables and frames attached to facilitate their recovery after the

Table 4: Four-class classification confusion matrix for a quadratic classifier using ($N = 4$, $M = 3$) central features and restricting the samples to those on flat seabed

Quadratic	None	Type A	Type B	Type C
None	**0.92**	0.04	0.04	0.00
Type A	0.11	**0.80**	0.09	0.00
Type B	0.01	0.02	**0.97**	0.00
Type C	0.00	0.00	0.06	**0.94**

Total classification error is 0.09
Bold indicates the highest probability value for each row

Table 5: Classifier detection performance figures for Matlab's linear and quadratic classifiers when using ($N = 4$, $M = 3$) central features

Training on ALL seabeds	Linear classifier	Quadratic classifier
Total number of targets in test set	53	53
Total number of non-targets in test set	1452	1452
Number of detected targets	43	36
Number of missed targets	10	17
Number of false alarms	1	8
Total number of matches	1494	1480
Total error on classification	0.007	0.017
Ratio of undetected mines	0.189	0.321
Probability of false alarm	**0.001**	0.006
Probability of detection	**0.811**	0.679

Classifier was trained on the AR database and then tested on the real database
Bold indicates the best false alarm rate (lowest value) and best detection rate (highest) that have been found

Table 6: Classifier detection performance figures for Matlab's linear and quadratic classifiers when using ($N = 4$, $M = 3$) central features

Training on FLAT seabed	Linear classifier	Quadratic classifier
Total number of targets in test set	53	53
Total number of non-targets in test set	1452	1452
Number of detected targets	45	53
Number of missed targets	8	0
Number of false alarms	0	16
Total number of matches	1497	1489
Total error on classification	0.005	0.011
Ratio of undetected mines	0.151	0.000
Probability of false alarm	**0.000**	0.011
Probability of detection	0.849	**1.000**

Classifier was trained on the AR database restricted to samples on flat seabed and then tested on the real database
Bold indicates the best false alarm rate (lowest value) and best detection rate (highest) that have been found

mission, which modified the target appearance in the real snippets.

It must also be taken into account that the training database contained flat, rippled and complex seafloors, whereas the real mission only presented flat seabed. We can modify the training database to reflect this fact by allowing only for the use of samples containing flat seabed. In this case (Table 6), the linear classifier performs better (PD = 84.9%) while causing zero false alarms, and the quadratic classifier actually achieves 100% detection at the expense of a 1.1% false alarm rate.

These results emphasise the importance of taking into account the context where the target is located. Training for the proper context can greatly improve the performance of the classifiers.

Results for four-class classification when training on flat seabed using linear and quadratic classifiers are presented in Tables 7 and 8, respectively.

These results show some bias with regard to the training figures presented in Tables 3 and 4. This is most probably caused by the fact that the targets present in the real data were mock-up models which differ in several physical characteristics from the actual targets. In particular, adverse elements are the frames and cables that were attached to the mock-up targets to facilitate their recovery. These resulted in the targets' appearance differing from the one the system was trained to recognise. Note, however, that this has not particularly affected the capability of the system to discriminate these objects as targets, which is particularly reassuring for such a critical task as ATR is.

It is also important to consider the limited number of real target samples that we had access to, which inevitably introduces some uncertainty into the performance values presented here. Extended testing should be performed in order to obtain more accurate performance figures for detection and classification of real targets.

Table 7: Four class classification confusion matrix for a linear classifier using ($N = 4$, $M = 3$) central features when training on flat seabed AR data and testing on the real database

Linear	None	Type A	Type B	Type C
None	**1.00**	0.00	0.00	0.00
Type A	0.25	**0.44**	0.06	0.25
Type B	0.06	0.29	**0.53**	0.12
Type C	0.05	0.20	**0.50**	0.25

Total classification error is 0.07
Bold indicates the highest probability value for each row

Table 8: Four-class classification confusion matrix for a quadratic classifier using ($N = 4$, $M = 3$) central features when training on flat seabed AR data and testing on the real database

Quadratic	None	Type A	Type B	Type C
None	**0.96**	0.04	0.00	0.00
Type A	0.00	0.19	**0.81**	0.00
Type B	0.00	0.12	**0.88**	0.00
Type C	0.00	0.00	0.20	**0.80**

Total classification error is 0.08
Bold indicates the highest probability value for each row

5.3 Results on images

Target detection and classification using the approach presented in the previous sections can be performed directly on side-scan images. A straightforward (but slow) implementation is to use a sliding window, in which a snippet is extracted at each image pixel and then classified as discussed earlier. Our current optimised implementation uses image convolution to accelerate the feature extraction process, and can filter a 1024×1000 side-scan image in 39 s when running on a 3.4 GHz Pentium 4 (16 central features, linear classifier, 3392 training samples).

An example of the type of results obtained with this procedure is shown in Fig. 8, where target detection has been performed on an AR image. Thirteen objects are

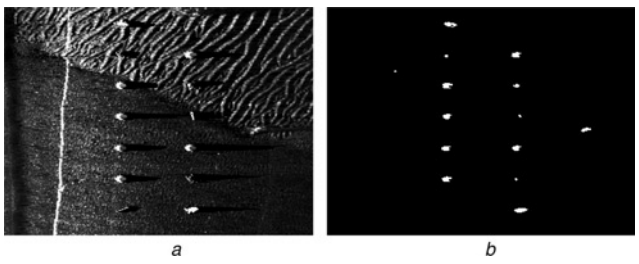

Fig. 8 *Detection on an AR image containing 13 objects, of which 11 are positive targets*
a Input AR image
b Results of detection using the proposed supervised approach, producing 11 detections and three false alarms

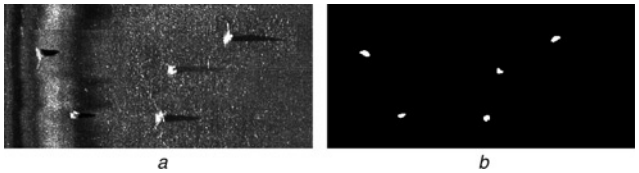

Fig. 9 *Detection results on a collage of real imagery*
a Input image presenting four targets and one non-target object
b Detection results using the proposed supervised approach when training on AR data, producing four detections and one false alarm

present in the AR image (2 type A, 6 type B, 5 type C), and the training database used contained only positive samples of targets of types B and C. The results of the detection show 100% detection of the B and C targets and three false alarms (a type A object, a rock and a false positive on the surface return).

An example of application of the same technique to a collage of real imagery is shown in Fig. 9, where the input side-scan image contains five objects (1 type A, 3 type B and a non-target). Results of detection are 100% on the A and B type targets, and the non-target object produces a false alarm. The system was trained on an AR database containing A, B and C type synthetic targets on flat seabed only.

In order to estimate the performance of this type of approach as an operational ATR system, we have processed a real mission consisting of 51 images and containing 36 target observations. After processing, all 36 targets were detected and 17 false alarms were raised. The 51 images covered an area of 363 477 m^2, which results in a rate of 160 false alarms per square nautical mile for 100% detection rate. This figure compares favourably with other current model-based CAD/CAC systems [10].

6 Conclusions and future work

The viability of the supervised approach for target classification has been demonstrated. A classification system is trained on AR data and then tested against real side-scan images of targets. Unlike model-based CAD/CAC systems, this approach does not require the development of a model for target detection; instead, the system is able to automatically select the feature configurations that define a particular type of target. In addition, it offers a major advantage over traditional supervised approaches to CAD/CAC, which were often constrained by a limited training set in their development.

Different feature extractors can be used as the basis for the classification system. In the particular implementation presented here central features – a feature set specifically designed for targets in side-scan images – have been used. This set of features is well suited for the measurement of the characteristics that define an object of interest for MCM operations (presence, highlight-shadow dichotomy, details).

Simple classifiers were used as a proof of concept but more advanced ones could be employed instead. Our future and current work tackles the fusion of different weak classifiers [18, 19]. This has the desirable advantage that none of the classifiers needs to excel at the classification task, as the combination of several weak classifiers will generally yield better results if the classifiers are not perfect. This permits the use of very simple and varied approaches, rather than fine tuning a complex classification scheme. Using our AR-based approach, we are able to use, at a very low development cost, a number of legacy algorithms from the supervised pattern recognition literature. By choosing the algorithms to be fused, and the rule to fuse their outputs, we will also have some control on the trade-off between the final probability of detection and false alarm rate.

Besides the development of a model-less CAD/CAC system, the good matching between synthetic and real targets shows the proposed approach can be used as a versatile CAD/CAC system evaluation tool. This way a fair comparison of several algorithms can be performed on ground-truthed and calibrated target databases.

7 Acknowledgments

This work was funded by DSTL grant RT/COM/4/028. Some of the data used in this paper was gathered during the CITADEL 05 experiment, a collaboration between the NATO Undersea Research Centre (NURC), GESMA and Defence R&D Canada, Atlantic. The authors would like to thank the Scottish Funding Council for their support to the Joint Research Institute in Signal and Image Processing, which is a part of the Edinburgh Research Partnership. The authors wish to thank the editor and reviewers for their useful comments and suggestions on the early versions of this paper.

8 References

1 Reed, S., Petillot, Y.R., and Bell, J.: 'An automated approach to the detection and extraction of mine features in sidescan sonar', *IEEE J. Ocean. Eng.*, 2003, **28**, (1), pp. 90–105
2 Duda, R.O., Hart, P.E., and Stork, D.G.: 'Pattern classification' (Wiley Interscience, 2001, 2nd edn.)
3 Dobeck, G.J., Hyland, J.C., and Smedley, L.: 'Automated detection/classification of sea mines in sonar imagery', *Proc. SPIE – Int. Soc. Opt*, 1997, **2079**, pp. 90–110
4 Ciany, C.M., and Zurawski, W.: 'Performance of computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines'. CAD/CAC Conf., Halifax, Canada, November 2001
5 Aridgides, T., Fernandez, M., and Dobeck, G.: 'Adaptive three-dimensional range-crossrange-frequency filter processing string for sea mine classification in side-scan sonar imagery'. Proc. SPIE, 1997, **3079**, pp. 111–122
6 Doherty, M.F., Landowski, J.G., Maynard, P.F., Uber, G.T., Fries, D.W., and Maltz, F.H.: 'Side scan sonar object classification algorithms'. Proc. 6th Int. Symp. on Unmanned Untethered Submersible Technology, 1989, pp. 417–424
7 Balasubramanian, R., and Stevenson, M.: 'Pattern recognition for underwater mine detection'. Proc. CAD/CAC Conf., Halifax, NS, Canada, November 2001
8 Bell, J., and Lianantonakis, M.: 'Analysis of target scattering using a finite difference time domain model'. Proc. Institute of Acoustics Conf. on Sonar Transducers and Numerical Modelling, March 2005, vol. 27, Pt. 1
9 Dura, E., Yan, Z., Xueiun, L., Dobeck, G.J., and Carin, L.: 'Active learning for detection of mine-like objects in side-scan sonar imagery', *IEEE J Ocean. Eng.*, 2005, **30**, (2), pp. 360–371
10 Petillot, Y., Reed, S., and Coiras, E.: 'An augmented reality solution for evaluating underwater sonar MCM systems'. Proc. of the 7th Int. Symp. on Technology and the Mine Problem, 2006
11 Coiras, E., Petillot, Y.R., and Lane, D.M.: 'An expectation-maximization framework for the estimation of bathymetry from side-scan sonar images'. Proc. Oceans 2005 Europe, June 2005, vol. 1, pp. 261–264
12 Langer, D., and Hebert, M.: 'Building qualitative elevation maps from side scan sonar data for autonomous underwater navigation'. Proc. 1991 IEEE Int. Conf. on Robotics and Automation, vol. 3, pp. 2478–2483
13 Zhang, R., Tsai, P., Cryer, J., and Shah, M.: 'Shape from shading: a survey', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1999, **21**, (8), pp. 690–706
14 Ragheb, H., and Hancock, E.R.: 'Surface radiance correction for shape from shading', *Pattern Recognit.*, 2005, **38**, pp. 1574–1595
15 Reed, S.: 'Automatic detection and classification models for sidescan sonar imagery'. PhD Thesis, Heriot-Watt University, Edinburgh, UK, April 2004
16 Sadjadi, M.R.A., Yao, D., Huang, Q., and Dobeck, G.J.: 'Underwater target classification using wavelet packets and neural networks', *IEEE Trans. Neural Netw.*, 2000, **11**, (3), pp. 784–794
17 Krzanowski, W.J.: 'Principles of multivariate analysis: a user's perspective' (Oxford Science Publications, 1988)
18 Dobeck, G.: 'Algorithm fusion for the detection and classification of sea mines in the very shallow water region using side-scan sonar imagery', *Proc. SPIE*, 2000, **4038**, pp. 348–361
19 Freund, Y., and Schapire, R.E.: 'A decision-theoretic generalization of on-line learning and an application to boosting'. Proc. European Conf. on Computational Learning Theory, 1995, pp. 23–37