

A FRAMEWORK FOR EVALUATING UNDERWATER MINE DETECTION AND CLASSIFICATION ALGORITHMS USING AUGMENTED REALITY.

Yvan Petillot¹, Scott Reed², Enrique Coiras¹.

¹ Heriot-Watt University - School of EPS.

² SeeByte Ltd.

Abstract— This paper presents a novel framework for evaluating Target Detection and Classification algorithms and concepts of operations based on Augmented Reality (AR). Real sonar images and synthetic target models are used to generate a ground-truthed AR theatre of operation. The detection/classification results of the human operator or Automatic Target Recognition (ATR) algorithm to be evaluated are then compared to the mission ground-truth to obtain realistic performance measures. A separate seabed classification module permits to analyze these performance measures in terms of seabed texture complexity and clutter density. This approach permits the evaluation of Detection and Classification algorithms in a large variety of scenarios. The overall system is presented and evaluation on real data demonstrated.

I. INTRODUCTION

A. Background

With the advent of Autonomous Underwater Vehicles (AUV), new mine hunting concepts of operations using very high resolution side scan sonars have appeared [1]. Because these platforms are autonomous, it is critical to develop embedded automatic target recognition (ATR) to enable on-board decision making, collaborative behaviours between heterogeneous platforms and sensors (detection, identification) and ultimately true autonomy with on-board re-planning capabilities. In the context of Mine and Counter Measures (MCM), the ATR performances of such platforms in a largely unknown and unstructured environment is critical. So far, they have been partially validated by various military organizations (NATO Undersea Research Centre, US NAVY, US Marines) and their operational capability in mine hunting demonstrated in a number of real experiments [1]. However, they have not yet been fully characterized and their performances have not been systematically established. This is a critical part of any ATR system evaluation and has proven notoriously difficult in other fields such as Synthetic Aperture Radar where DARPA has funded the MSTAR project specifically to solve this issue [2].

This paper presents a methodology to quantitatively validate Detection and Classification algorithms and concepts of operations for mine hunting in realistic environments. As such our main contribution is one of operational research based on advanced image analysis and pattern recognition techniques. In order to implement the methodology, basic building blocks are required, drawing expertise from image processing and pattern recognition. The techniques presented here implement those building blocks but could be replaced by

other techniques fulfilling the same functionality. However, for clarity, the techniques have been described in some detail.

B. Relevant parameters for Planning & Evaluation

In order to design and evaluate sonar platforms, several system and scene parameters need to be determined. AUVs are currently equipped with high resolution side scan sonar. The resolution of these systems is dictated by the frequency they operate at and the length of their aperture. Low cost systems, prevalent at the moment, operate at frequencies ranging from 600-1800 KHz (good across track resolution-5cm) and a relatively small antenna (25-50cm). They therefore have a comparatively poorer along track resolution (10-20cm). The dynamic range of those system is also limited (6-12 bits) making echo based detection difficult. Synthetic Aperture technology can improve along track resolution without the need for large antennas. Shadows are more defined than echoes and currently used for target classification. The discrimination for such systems is therefore very dependent on the precision of the shadow shape, which is determined by the along and across track resolution and the grazing angle. The shadow is the intersection of the projection of a 3D object with the seafloor surface and therefore also depends on the background on which the object lies. The type of seabed and more critically the 3D variation of the seabed with respect to object height will determine how well the shadows can be defined. Finally, at a given resolution, it might not be possible to distinguish between real targets and mine like objects (MILOCS). The following parameters are therefore of interest when studying evaluating target detection and classification algorithms:

- 1) Sonar along track resolution, determined by the sonar aperture. Synthetic Aperture technology can improve along track resolution without the need for large antennas.
- 2) Sonar across track resolution: this is determined by the operating frequency of the sonar. Higher frequencies will improve resolution at the expense of range. Optimum parameters should be chosen depending on the mission and be part of the planning tool.
- 3) Seabed type: The interaction between the seabed and the targets will depend on the seabed type and bathymetry (marine growth, rocky structures) and the amount of perturbations they will provoke on the target echo and shadow.

C. Methodology

Finding a methodology to consistently assess AUV performances is complex. The performance will depend

on all the parameters described in section I-B. Identifying the influence of each of them is critical to the future development of new systems and should help focusing the resources. Two approaches have been used in the past:

- 1) Experimental evaluation: NURC and the US Navy have performed extensive experiments using real targets and real vehicles [1] (GOATS, BP02/03, various AUV Fests). However, these experiments are expensive to run and only a limited number of targets and situations they can be covered. Furthermore, ground-truthing is difficult, as target positions are not always precisely known. Therefore, extracting meaningful statistics and quantifying as opposed to qualifying the performance of a particular system is difficult. It is also unlikely that performance in one environment will extrapolate easily to another given the strong dependence of the side scan sonar imagery to environmental conditions and sonar systems.
- 2) Evaluation by simulation: Sonar simulators are available to various degrees of realism [3]. Some of them can also simulate the seabed bathymetry. They are an ideal tool to predict the performances of sonar arrays. The environment can be simulated and synthetic mine fields produced. However, they are in general prohibitively computer intensive and sometime lack the realism of real systems as simplifying assumptions are made to render the models tractable [3].

We propose here to use augmented reality to combine the advantages of both approaches. Using real data, we stay close to the real sensor. A new simulation approach to simulate very realistic targets is presented, enabling ground-truth results to evaluate the ATR (computer or human) performances in various scenarios. The proposed solution is best described in Fig 1. The idea behind the proposed framework is to evaluate a *system* rather than its components. Therefore we start from the data produced by a particular system. The sonar data is used by the sonar inversion module to extract an estimate of the bathymetry (Z map) and reflectivity (R map) of the area and the beam pattern (Φ map) of the sonar. The simulation model locally modifies the Z and R map estimates to introduce simulated targets in the real image using the targets CAD model and their associated reflectivity maps. A Lambertian model is assumed for the targets. The output is a ground-truth simulated minefield. Any ATR system can then be used to determine the number of detection and false alarms present in the image and therefore evaluated on realistic data. In parallel, the original sonar images are classified into seabed classes and geo-referenced to produce a class map of the area covered by the sensor. Each detected target from the ATR is correlated with the class map to produce PD/PFA curves for each seabed type.

This framework enables any system to be evaluated using its own data collected on real missions. The main benefits of the approach are:

- 1) Real data is used. This is critical if meaningful evaluation is to be performed.
- 2) Enables the system to evaluate ‘what-ifs’ scenarios:
 - a. The Sonar resolution is changed,
 - b. The reflectivity of targets is different,
 - c. The targets are partially buried.
- 3) Probability of Detection(PD)/Probability of False Alarm (PFA) on existing missions can be calculated
- 4) PD/PFA and percentage clearance on future missions on similar terrain can be predicted.

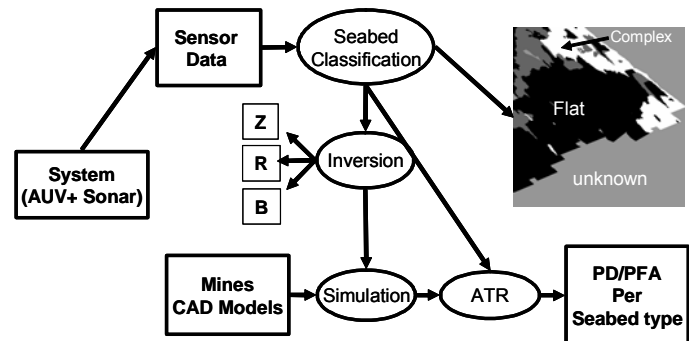


Fig 1: Planning & Evaluation Framework: mixing real data and simulation

D. Datasets

Two large datasets have been used to validate the methodology presented here. Examples of typical images from both data sets are shown in Fig. 2.

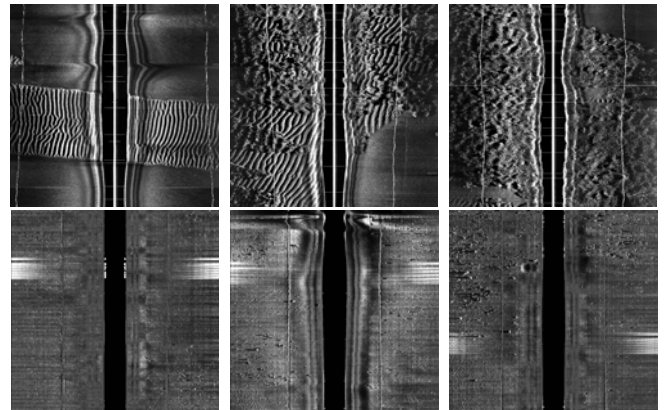


Fig. 2: Example of images from test missions. Top row: BP02 data set examples, bottom row: Baltic data set examples.

The first data set was acquired over the Framura Area (Italy, South of La Spezia) during the BP02 experiment and will be referred as the BP02 dataset. It contains no targets, little clutter (i.e. mine like objects of mine dimensions). The seabed texture is varied: 88% of the mission is flat seabed, 6% is covered by sand ripples and 6% by Complex seafloor (mainly posidonia, complex seafloor is described here as difficult to mine hunt). 235 images covering approximately 1 square nautical mile were used. The second data set was acquired in the Baltic and contains some mine like objects. The seabed is mainly flat but some sections are highly cluttered by

mine like objects. 90 images covering approximately 0.5 square nautical miles were used. These two data sets enable to validate the system as they represent areas of varying mine hunting difficulty and are 'orthogonal' in terms of relevant seabed criteria (complex seabed texture - low clutter versus flat seabed - high clutter).

II. SONAR INVERSION PROCESS

Efforts towards the use of side-scan sonar for the indirect determination of seabed topography have been scarce [4][5][7]. The main reasons are the complexity of the full mathematical projection model and the high number of procedures required for preprocessing the original source data. In most cases where acquisition of seabed topography is important, attention is driven to more straightforward solutions such as multi-beam bathymetric systems. The idea behind the sonar inversion process[7] is to recover the main parameters involved in the sonar image formation. These are bathymetry, reflectivity and sonar parameters. The sonar parameters are here reduced to the beam pattern. The navigation effects, such as pitch and roll are ignored.

A. Image Formation Model

1) Side-scan Sonar

The side-scan image formation process is briefly sketched in Fig 3. The sensor's acoustic source at point o produces an ensonification pulse that illuminates the seafloor. Some of the acoustic energy reaching any seabed point p is scattered back and can be measured by the sensor. The intensity of the corresponding pixel of the side-scan image will depend on the amount of energy scattered back from the surface point. The pulse is not isotropic, but follows a particular beam-profile Φ that depends on the grazing angle α subtended by p . The amount of energy scattered back also depends on the seabed reflectivity $R(p)$ at point p .

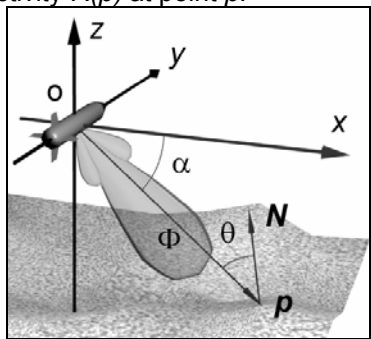


Fig 3: Side-scan image formation

2) Scattering Model

In order to model the scattering process we use the traditional Lambertian model. This allows the returned intensity to be derived from the observed scene parameters. This simple model for diffuse scattering assumes that the returned intensity depends only on the angle of incidence of the illuminating sound pulse, and not on the angle of observation or on the frequency of the pulse. Under these assumptions the intensity I returned

from a seabed point p can be represented by the following expression:

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

where Φ represents the intensity of the illuminating sound wave at point p , R is the reflectivity of the seafloor, θ is the incidence angle of the wave front and K is a normalization constant. In order to simplify the model, all the intensity variations caused by the sensor's beam-profile, the radial intensity decay and the possible Time-Varying Gain (TVG) corrections are supposed to be grouped under the beam-pattern Φ . The dependence on the seafloor's elevation is implicit in the incidence angle $\theta(\bar{p})$, which depends on the grazing angle α from the acoustic source and the orientation of the surface normal $\vec{N}(\bar{p})$. Making this dependence explicit, by expressing the cosine as a dot product of the surface normal and the direction of observation, gives the forward model for the computation of the intensity I at any point \bar{p} , given the model parameters R , Z and Φ :

$$I(x, y) = K \Phi(x, y) R(x, y) \cdot$$

$$\frac{Z(x, y) - x \cdot \frac{\partial Z}{\partial x}(x, y)}{\sqrt{x^2 + Z^2(x, y)} \cdot \sqrt{\left(\frac{\partial Z}{\partial x}(x, y)\right)^2 + \left(\frac{\partial Z}{\partial y}(x, y)\right)^2 + 1}} \quad (2)$$

Where the gradients $\partial Z/\partial x$ and $\partial Z/\partial y$ can be approximated by finite differences, yielding an expression that depends directly on Z .

B. Parameter Estimation

1) Normalization

The values of the model parameters are limited to particular ranges, which have to be observed during the optimization process. Reflectivity values, by definition, have to lie between 0 and 1, but to avoid the model from collapsing in the darker areas, a lower bound R_{\min} greater than 0 is chosen on initialization (with a typical value of 0.1). Intensity values for the incident sound wave are assumed to lie between 0 and 1, which would normally mean bounding Φ within that range. However, because Φ may include an unknown TVG, we allow it to achieve values greater than 1. In practice this amounts to just a little overshoot for the bigger angles, which naturally correspond to points of the seabed farther away from the sensor and therefore require higher TVG corrections. To ensure that solution parameters are kept within their respective value ranges during optimization, an appropriate normalization scheme is required. This normalization should also take into account that most of the source image intensity values are being affected by some kind of scaling (because of the TVG, the radial decay, etc). In order to do that we use the maximum achievable returns at every surface point to normalize the computed pixel values. Under this assumption we obtain the following normalization constant for the forward

model (1):

$$K(x, y) = \frac{\sqrt{x^2 + Z^2(x, y)} \cdot \sqrt{\left(\frac{\partial Z}{\partial x}(x, y)\right)^2 + \left(\frac{\partial Z}{\partial y}(x, y)\right)^2 + 1}}{-Z(x, y) + \frac{x^2}{-Z(x, y)} \cdot \left(1 + \left(\frac{\partial Z}{\partial y}(x, y)\right)^2\right)} \quad (3)$$

2) Expectation-Maximization

Equation (4) provides a direct formula for estimating the returned intensity given the model parameters. But the inverse problem—obtaining the model parameters from the observed intensities—is clearly under-determined, since we only have one observation (of I) at each point to compute the values of the beam-pattern, height map and reflectivity. In order to solve this problem we use an expectation-maximization [6]. The model will iteratively converge to an optimal set of modeling parameters given a source side-scan image I . The objective is to minimize the absolute difference between the observed intensity I and the one resulting from the application of the model \hat{I} , which we represent by the error quantity E :

$$E = \sum_{x,y} E(x, y) = \sum_{x,y} \left(I(x, y) - \hat{I}(x, y) \right)^2 \quad (4)$$

In the expectation stage, the current estimates for the model parameters are used to compute an estimation of the intensity $\hat{I}(\bar{p})$. This is achieved by substituting the parameters $R(\bar{p})$, $\Phi(\bar{p})$ and $Z(\bar{p})$ from the previous iteration in the forward model presented in expression (4).

In the maximization stage a straightforward gradient descent approach is used to minimize E , by updating the model parameters as follows:

$$\begin{aligned} R(x, y) &\leftarrow R(x, y) + 2\lambda \cdot \frac{\hat{I}(x, y)}{R(x, y)} \left(I(x, y) - \hat{I}(x, y) \right) \\ \Phi(x, y) &\leftarrow \Phi(x, y) + 2\lambda \cdot \frac{\hat{I}(x, y)}{\Phi(x, y)} \left(I(x, y) - \hat{I}(x, y) \right) \\ Z &\leftarrow Z - 2\lambda \hat{I} \left(I - \hat{I} \right) \cdot \left(\frac{-\partial Z/\partial y - \partial Z/\partial x}{1 + (\partial Z/\partial y)^2 + (\partial Z/\partial x)^2} + \frac{1+x}{x(\partial Z/\partial x) - Z} + \frac{Z}{x^2 + Z^2} \right) \end{aligned} \quad (5)$$

where λ is a small constant value used to control the rate of change (typically 0.25). The explicit dependence of the parameters on (x, y) has been removed equation 8 for clarity. The expressions are iterated until the variation in the error E is below a given threshold. Results improve notably when using a multi-resolution version of the same algorithm. This allows the seafloor scene to be recovered in a more progressive as well as reducing the overall error at convergence. Implementation of the multi-resolution version begins with the construction of a multi-resolution pyramid by iterated sub-sampling of the source side-scan image. Processing starts at the smallest level, using the initialization and regularization procedures described in the previous sections. The resulting R , Z

and Φ maps from one level are used as initial maps for the next resolution level. The process finishes when the final stage—corresponding to the full resolution image—is processed.

C. Results

The method has been evaluated on a number of systems and produced very encouraging results on the two datasets described in section I-D. An example of solution obtained using Marine Sonics data is shown in Fig. 4. Full details of this technique can be found in more detailed publications [7].

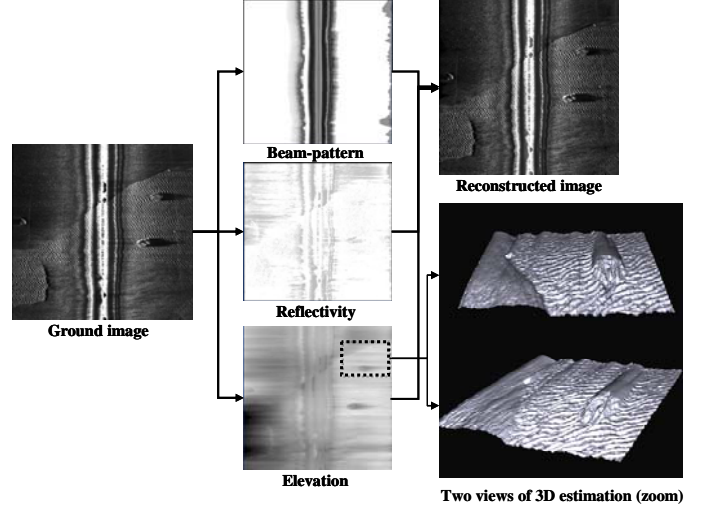


Fig. 4: Example of inversion on REMUS Data (BP02).

III. SIMULATING MINE FIELDS VIA AUGMENTED REALITY

A. Principle

Once the Z, R, Φ maps from a Side Scan Sonar image have been recovered, it is possible to introduce simulated targets by modifying locally the Z and R maps. The Computer Aided Design model of the target and its reflectivity index (if constant) or associated reflectivity maps must be available. Once the Z and R maps have been modified, equation 4 can be used to render the new simulated image. Unlike other simulators which paste a simulated target on top of an existing image, this simulator enables the modeling of the interactions between the topography of the seabed and the target. For instance, if a target 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 target. The principle of the target simulator is summarized in Fig. 5:

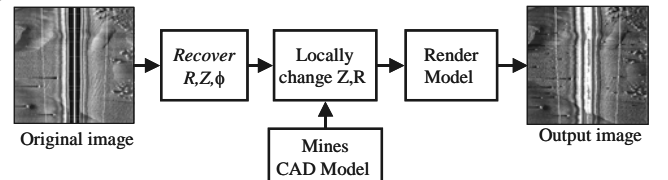


Fig. 5: Augmented Reality Target Simulation Principle

B. Parameters

The parameters of the model are

- 1) Sonar Type: This is driven by the choice of images used in the sonar inversion process and is embedded into the Z,R, Φ maps.
- 2) Target type: at present, two of the most common existing targets can be modeled. They have been named type A and type B for confidentiality reasons. More target types can be readily added if the Computer Aided Design models and reflectivity maps are known.
- 3) Along / across track resolution: can be changed by re-sampling the Z and R maps.
- 4) Burial Depth: This can be changed by modifying the depth at which the target is embedded in the Z map.
- 5) Target Reflectivity: This is simply modified by changing the target's R map. This must be calibrated. At present the reflectivity of the seabed is normalized in the range [0,1] during the sonar inversion process. A calibration process should take place for each sonar to allow the precise setting of targets reflectivity.
- 6) Target aspect: The aspect (angle) of the target with respect to the sensor can be set.
- 7) Echo speckle. Side scan sonar uses coherent image formation. Therefore speckle is present on the target echo. This is modeled as an additive Rayleigh distribution in the reflectivity map. The parameters of the Rayleigh distribution can be adjusted to model realistic levels of speckle for each sonar.

A large range of simulations can be run to explore the impact of those parameters on the mission performances (percentage clearance, number of false alarms, PD, PFA).

C. Augmented Reality Simulation Results

The simulator was tested on Marine Sonics and Klein sonars and the various parameters of the model changed. The Marine Sonics sonar has a 12cm along track and 5.8cm across rack resolution while the Klein has a 10 cm along track and 3 cm across track resolution. The Lambertian model used for the targets is a clear limitation of the actual model as targets (metal) are not Lambertian surfaces. This will be improved in the future to include more complex interaction models.

1) Result for different sonar models

An example of simulation of a two targets for the Klein 5000 and Marine Sonic 900kHz are given in Fig. 6. The difference is resolution between the two sonars is noticeable as well as the difference in clutter level. Both set of images were obtained for a target reflectivity of one.

2) Results for different target reflectivity

The reflectivity of the target can be changed. This is important to model the saturation effect sometimes visible on sonar images due to the reduced dynamic range of the Sonars. FIG. 7 show an example where the resolution of the sonar has been kept to real values (5.8 cm x 12 cm). and the reflectivity for the two targets types is 1 and 3.

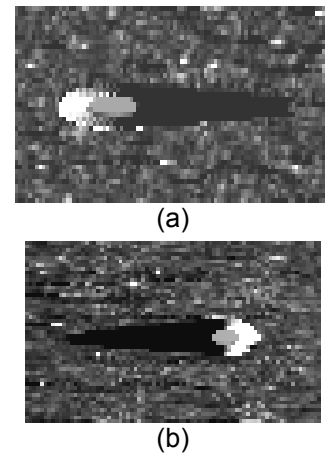


Fig. 6: Example of targets simulation (A) using a Klein 5000 (image a) and a Marine Sonics Sonars (image b).

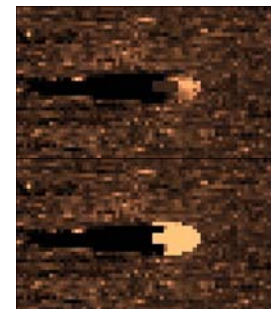


Fig. 7: Target A simulated with reflectivity indexes 1 (top) and 3 (bottom)

3) Target interaction with complex Seabed

The final demonstration shows in Fig 8 a simulated minefield on a typical seabed type containing flat, rippled and complex classes. This demonstrates the interaction with the seabed that this simulation enables through the extraction of the Z map.

IV. CAD-CAC ALGORITHMS

A. Introduction

ATR systems for MCM are generally composed of a CAD (Computer Aided Detection) and a CAC (Computer Aided Classification) component [8] [9][10]. The aim of the CAD module is to detect all possible mine-like objects (MILOC's), often at the expense of detecting many false alarms. The CAC component represents the second level of analysis, and is tasked with providing a measure of how 'mine-like' each of the MILOC's produced by the CAD module are. Based on the information provided by the CAD, the CAC must decide whether the MILOC is a target or a false alarm [9]. The performance of the CAD and CAC modules will be dependent on several parameters. Two important parameters are the sonar type (which will dictate the resolution of the side-scan sonar images) and the seafloor type the AUV is surveying. Assuming the resolution remains sufficient to resolve an object from the background, the CAD module is often able to provide high probability of detection (PD) values regardless of the specific values of these

parameters. However, as the conditions move away from the ideal scenario (high resolution sonar on a flat seafloor), this high PD rating will only be achievable if the probability of false alarm (PFA) is also increased.

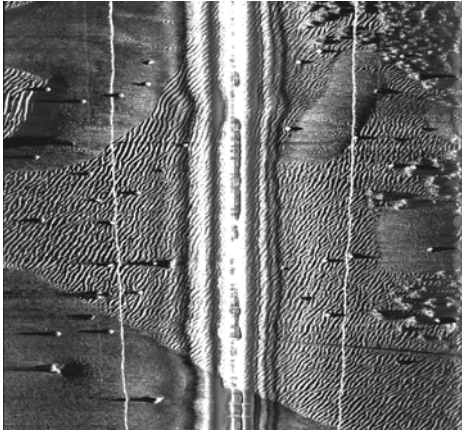


Fig 8: Example of simulated minefield on varied seabed. 10 type A and 10 type B targets have been simulated on each side of an image containing a variety of seabed types (flat, ripples, complex).

Current CAC systems use the shadow of the MILOC to provide a classification decision [17]. The effectiveness of the CAC module will be strongly correlated to both the sonar conditions and the seafloor type. The ability to discriminate between a man-made object and a natural object is ultimately dependant on the number of image pixels present within the MILOC shadow region. The ability to both identify real alarms and remove false alarms will therefore decrease with sonar resolution. Conversely, the ability to classify should increase with across track range from the AUV, as object shadows become larger. Seafloor types containing their own shadow regions (such as rocks and sand ripples) add a further level of complexity. The shadow regions from the background corrupt the shadow regions from the object making harder for the CAC module to determine whether shadow region belongs to the object or to the background. An example of a simulated target on both a flat, 'ideal' seafloor and on a complex seafloor is shown in Fig 9. The object's shadow can be clearly identified on the flat seafloor, but not on the complex one.

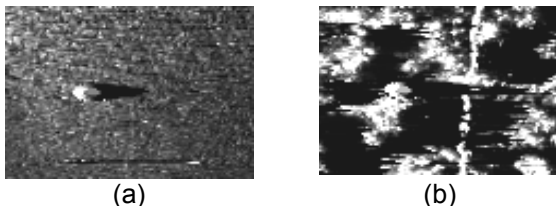


Fig 9: Example of images containing 1 object sitting on a flat (a) and complex (b) seafloor respectively.

The CAD and CAC modules use a model-based approach, where both the sonar process and the physical parameters of the objects being searched for are considered during the analysis process. Therefore as AUV computational capabilities improve, the models can

be continuously adapted to include more realistic information and simulations, resulting in improved performances. Full details on the CAD-CAC models presented here can be found in [13].

B. CAD Model

This CAD module is an adaptation of the model presented in [13]. It is therefore summarized here. The CAD process can be broken down into several key operations. First the image is slant range corrected. Using the AUV navigation information, the across and along track resolutions of the image are determined. The water column region from the centre of the image is removed as it contains no useful information. In a second stage, the surface return is automatically detected and removed and the image is segmented into regions of seabottom-reverberation, object-highlight and shadow regions using a Markov Random Field (MRF) model. Finally, obvious false alarms are removed from the CAD result in a post-processing phase by considering the size and height (determined from the shadow region) of each possible contact.

C. The CAC Model

The CAC module uses a model-based approach and provides a classification decision based on the shadow region of each detected object [9]. The CAC model can be split into key operations :

- 1) The CAD sends the CAC a mugshot of the detected object. This is accompanied with the sonar conditions under which the object was detected (target position, across track distance, AUV altitude, sonar frequency and resolution, estimated object height and width)
- 2) The CAC model generates shadow regions from different possible target objects under the same conditions that the object was detected. The simulation includes the sonar transfer function. Each tested target object is tested through an exhaustive set of possible object parameters such as depth, and orientation angle. The tested objects are cylindrical, spherical, type A and type B targets.
- 3) For each object shadow, the similarity between the real object's shadow and the simulated shadow region is measured.
- 4) The best match from all the tested objects is stored.

The simulated target shadows are produced using a quick line-of-sight sonar simulator and a height map of the object. Simple beam pattern considerations are introduced by convolving the simulated shadow with a filter of variable size dependant on the beam angle and the range to the object. This ensures that the edges of the simulated shadow regions have pixels which have a probability of being shadow OR non-shadow. This information is considered in the classification process.

Fig 10 contains 4 objects from the CAD model. Two of these are real objects (a type A target and a type B target produced using the simulator model described in a previous section) while the last two are real false alarms. The output from the CAC analysis is also shown.

CAD	Actual	Best CAC	CAC
-----	--------	----------	-----









Mugshot	Classification	Model Fit	Likelihood Rating
	Type A		0.94 Type A
	Type B		0.70 Type B
	Clutter		0.48 Type B
	Clutter		0.82 Type B

Fig 10: Examples of the CAC input and output when analyzing the output from the CAD module.

The two real targets both have a high target classification. The correct target type has also been identified correctly. The first clutter objects have low classification values. The final clutter object has a high classification value and under the specific sonar conditions produces a shadow region very similar to that of a Type B Target.

D. Production of Clutter Density Maps

Clutter Density is a term given to the density of objects within a region having the correct sonar response and physical characteristics to be labeled as mine-like by the CAD. Clutter density is a critical parameter to assess the ability of a system to perform target detection and classification. The clutter density depends on two main factors: the specific CAD used and the size and type of targets that the CAD is searching versus the scale of the seafloor

Different CAD models operate using different methodologies to produce their results. Different models will therefore produce different false alarms in different places when analyzing the same data (this and the assumption that the different CAD model methodologies detect most of the real objects are the driving principles behind why multi-CAD fusion is so successful [23]). Different CAD models will locally produce a different measure of the clutter density. However, we argue that the global clutter density order of magnitude will be stable irrespective of the CAD algorithm used and will mostly depend on the seabed characteristics. The size of the objects being looked for by the CAD is another important factor when determining clutter density. This size must be compared to the scale of the seafloor that is being surveyed. Rocky seafloors are generally associated with high false alarm rates and therefore a high clutter density. However, this is only true if the scale of the rocks is comparable to the size of the targets being searched for. This section considers the second requirement – obtaining the clutter density maps. The results show clutter density maps for the CAD model described in this paper searching for objects between 0.25m and 1.5m in height and size. The clutter density is first determined on each individual image. The CAD is run on the individual side-scan sonar image after which

the clutter density is computed per square meters. To produce a large scale mosaic of the clutter density maps, the individual maps are geo-referenced into the frame of the survey region. Many of the individual clutter density images overlap in this frame due to the lawn-mower trajectory traditionally used for MCM surveys. When this occurs, the mosaicing algorithm takes the maximum clutter density value – this ‘pessimistic’ approach ensures that the mosaic retains the worst case scenario and registers the highest possible density of mine-like clutter within each area.

V. SEABED CLASSIFICATION

Seafloor classification information has been identified as being critical for MCM operations. The seafloor type will impact factors such as the overall visibility of any objects present, the number of expected false alarms and the probability of any objects present being buried.

A. What classes for MCM analysis?

Seafloor classification for MCM operations needs to consider both

- 1) The texture of the underlying seafloor
- 2) The clutter density of CAD detections.

Extensive research has been carried out in seabed segmentation and classification using side scan sonar [12][14] but never in the specific context of MCM. Some highly textured area might not present difficulties for a detection algorithm while un-textured area with high level of clutter will. Sonar textures are generally 3D textures and it is the scale of the 3D variations with respect to the objects to be detected that ultimately will dictate the mine hunting difficulty. It is therefore crucial to determine the mapping between 1) and 2) and the MCM performance to generate class-maps which reflect mine hunting difficulty and not just texture (or complexity) or clutter density. For military purposes, seabed complexity has been classified on a scale of A to D while clutter density has been classified on a scale of 1 to 4. For instance a flat, low clutter density seafloor would provide a A0 classification while a high complexity, high clutter density would receive a D4 classification. However, a C3 might be equivalent to an A4 in terms of PD/PFA for a given ATR. The mapping between the somewhat artificial A-D/1-4 classes to ATR performances is our end goal.

This is an initial first step in defining a standard for classification of side scan sonar images where currently only 2 dimensions (texture and clutter density) have been considered. Further dimensions such as sediment type (sand, mud, clay) should be added if this information was available. Whilst sediment information may be retrievable from side scan sonar, it is possible that this would require the fusion of side scan sonar with another sensor.

B. Single image texture model

Seafloor classification using side scan sonar has been widely studied [12][14]. The model used here is a classical model based on texture features described in more details in [14]. The seafloor classification first segments the image into 3 classes (seabottom reverberation, shadow, highlights). The original side scan

image and the 3-class segmentation are inputs to the seafloor classification module. From these 2 images, 4 features are extracted using sliding windows:

- The Maximum Height of the Seafloor (from the segmented image)
- The Mean Elongation of the shadow regions (from the segmented image)
- Maximum Angular Fourier Energy (from original sidescan image)
- Variance of the Angular Fourier Energy (from the original sidescan image)

The outcome of one classification can be seen in Fig. 11.

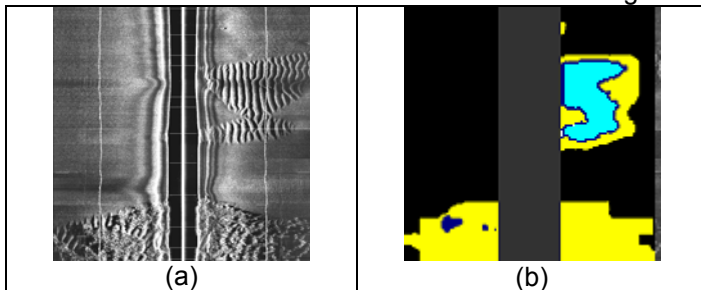


Fig. 11: Example of the seafloor classification (b) on a real side scan sonar image (a). Black represents flat seafloor, yellow (light grey) is the complex class and blue (dark gray) is for sand ripples.

The overall classification accuracy is acceptable. However, the classified image clearly contains miss-classified regions. From testing multiple seafloor classification algorithms on different datasets, the authors believe that an accurate seafloor classification cannot be achieved by considering side scan images in isolation. Even the best classification module will produce isolated incorrect results as the imaging conditions (sonar, seafloor, AUV) are very variable and some areas of the seabed are imaged under poor imaging conditions. To produce an accurate classification result, it is necessary to fuse the individual class maps to produce a class mosaic. This allows correctly classified regions to be verified and miss-classified regions to be identified and re-classified.

C. Production of Class Mosaics

Class mosaics can be produced by geo-referencing each of the individual class maps into the frame of the mosaic. The lawnmower trajectory used in MCM operations ensures that there are multiple classification results for many regions of the seafloor. This different information can be fused together to improve a classification result which is more accurate than any of the images considered in isolation. The mosaic also allows the large scale characteristics of the seafloor to be seen. This is not possible when considering the images in isolation.

Geo-referencing the individual class maps from a MCM survey results in some areas of the seafloor having multiple classification results associated to it. The fusion process uses the multiple classifications and initializes the final fusion map using a voting process. Areas with conflicting classifications are left as unclassified. A

Markov model is then used to smooth the final fusion result and classify the regions initialized as unclassified. Full details of the class map fusion process are available in [14].

For the BP '02 data this produces a final class mosaic shown in Fig. 12. As before yellow (light grey) indicates the complex seafloor, blue (light grey) represents the rippled class and the flat class is transparent (original side-scan mosaic can be seen). There are no unclassified pixels left on the mosaic. The light grey regions seen on the outside of the mosaic are for the unmeasured class – no class data is available for this region.

There are still miss-classified regions, which have been identified as complex. Many of these regions correspond to objects on the seafloor. This problem is addressed in the next section.

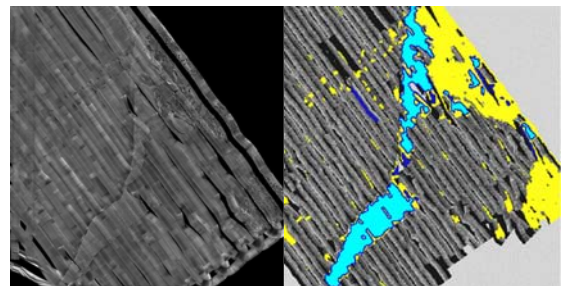


Fig. 12: Contains the fused mosaic obtained from considering over 200 class maps of the region. The mosaic can be seen to contain miss-classified regions of the complex class. The area has been classified as flat where the side scan mosaic can be seen

D. Fusion of Class and Clutter Density Mosaics

Section V-B and V-C described the texture based classification process used to segment the seafloor into regions of flat, rippled and complex. Section IV-D described the process to produce clutter density mosaics. The two models must be fused to produce a classification result, which is meaningful for MCM. As an initial step, the clutter density mosaic is simply used to identify and re-classify the miss-classified complex regions. The technique proposed below is shown to be an effective means of correctly classifying regions where the texture field has been corrupted by small objects lying on the seafloor. The clutter density and class map mosaics are fused at the initialisation phase when producing the class mosaic. After the voting process, the clutter density at each point initialized as complex is considered. Regions with a very low clutter density are re-initialised as unclassified. This is because the clutter density is seen as too low to support the texture analysis decision that the region is complex. The MRF model will again classify the unclassified regions using the surrounding contextual information. If the surrounding area is also complex (but presumably with a higher clutter density), the region can still be re-classified as complex. Otherwise, the area will be classified as flat or rippled, depending on the contextual information.

E. Classification Results

Results of the class mosaics produced from fusing the textural and clutter density information are provided for both the BP '02 data set and the Baltic sea trial data. Fig 13 contains the final classification result produced using the BP '02 data. All miss-classified complex regions have been identified and re-classified. The object field, present within the middle of the mosaic, which previously contained multiple miss-classifications due to the corruption of the texture field, is now correctly identified as being flat seafloor.

In some areas however, the visual evaluation by an operator would classify the seabed as complex while the algorithm classifies it as flat with a clutter density level too low to justify the complex label. This highlights the need for a more formal definition of the class maps in the context of mine hunting and points out the difficulties ahead in finding the appropriate mapping from texture and clutter density to mine huntability. Finally, our notion of clutter density is so far a *posterior* notion as it is directly derived from the ATR system used. It is also dependent on the ATR system whilst clutter should be an *a-priori* notion. We are well aware of this and hope that a more formal definition of clutter can be learnt from the posterior examples and extended to a variety of ATR algorithms.

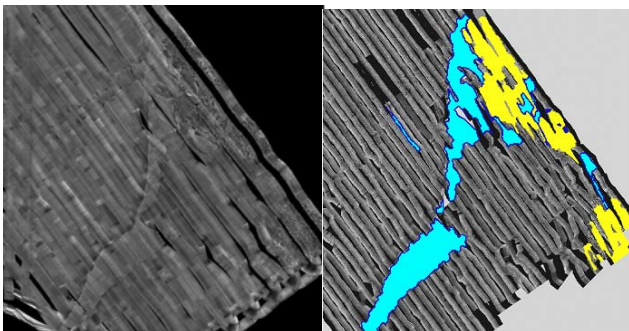


Fig 13: Classification mosaic for the BP '02 data using both textural and clutter density information. There is a large improvement in classification accuracy from the result shown in Fig. 12.

VI. SYSTEM EVALUATION

A. Methodology

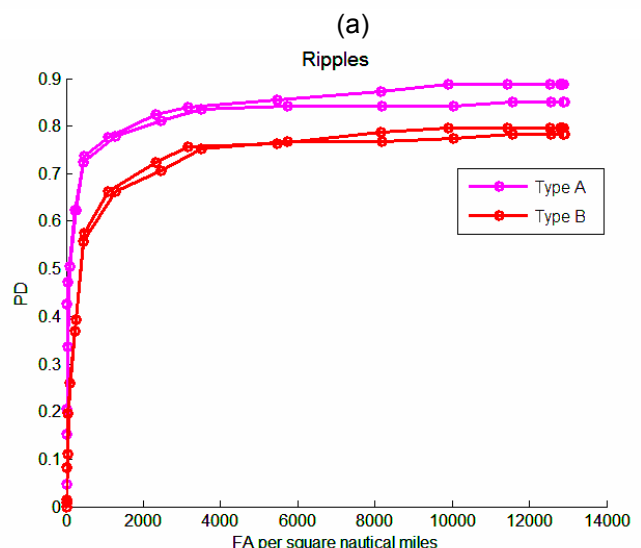
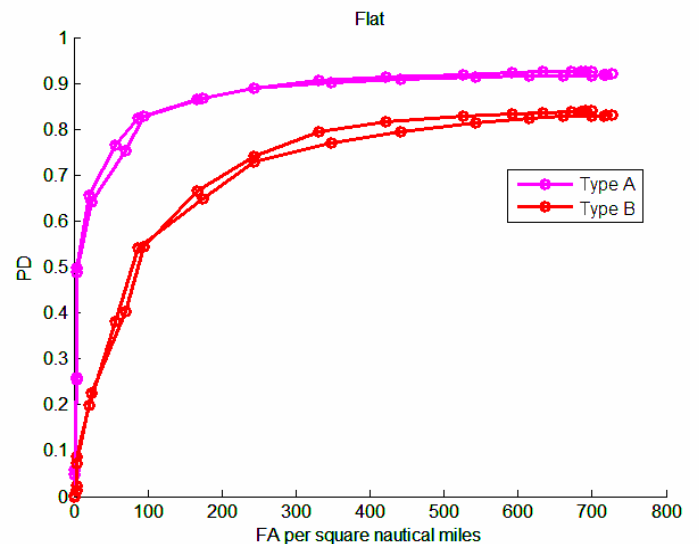
In order to validate the system, the two complete AUV missions described in section I-D were used. The sonar images were inverted to produce the (Z, R, Φ) maps and 40 targets per image (20 type A, 20 type B) were simulated at various angles. The reflectivity was fixed to 1 and there was no burial. The seabed classification modules (classification + mosaicing / fusion) were run on both missions to generate class maps enabling the robust determination of the class of the seafloor under each object. Finally, the ATR system (CAD-CAC) was run and the PD/PFA calculated for each seabed type. Results are presented for the CAD-CAC of section IV but the methodology can be applied to any ATR system. The number of False alarms per square nautical mile was also recorded as we feel it is a figure that can be better

interpreted operationally than the PFA generally reported as the number of false alarm is also an important parameter..

B. CAC Results

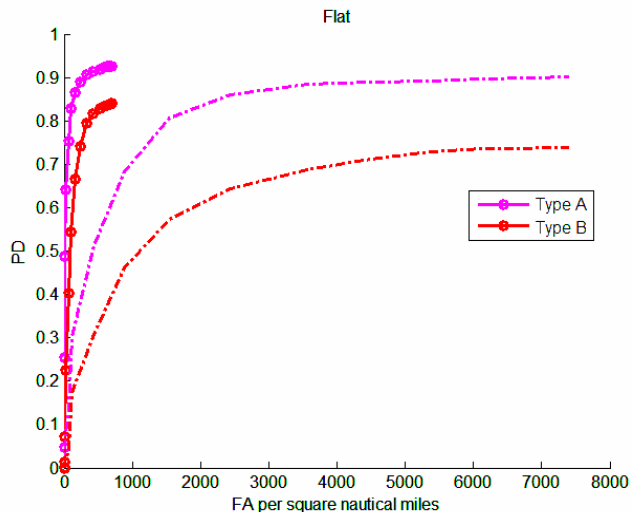
The results for the CAC show the probability of detection versus the number of false alarms per square nautical miles as opposed to the standard PD/PFA graphs. The curves are actually identical and the PD/PFA graph can be obtained by normalizing the FA rate by the maximum FA rate. We however argue that the PD/FA graphs are more informative as the critical information regarding the number of false alarms is not lost. The parameter varied to obtain the ROC curves is the degree of fitness between the model shadows and the segmented shadows. Again, PD/FA curves are obtained for each seabed type.

The first set of results in Fig. 14 shows two runs of the system on the same mission, on flat and rippled seabeds demonstrating the stability of the ROC curves estimation. In Fig. 15, results for two missions on flat and complex seabeds are presented. The Baltic trials are dash lines while the results for the BP02 trials are plain lines.

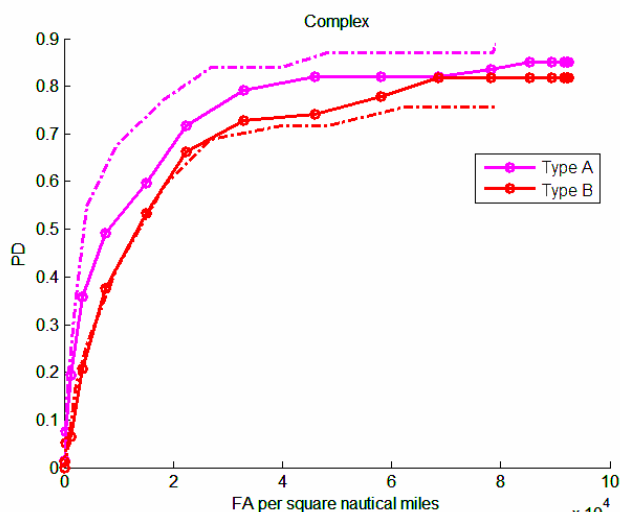


(b)

Fig. 14: CAC Results for two runs of the BP02 data on flat seabed (a) and rippled seabed (b) using the Planning & Evaluation framework proposed.



(a)



(b)

Fig. 15: CAC Results for BP02 and Baltic data. On flat (a) and complex (b) seabed types using the Planning & Evaluation framework proposed.

On complex area, the algorithm performances are comparable while on flat area, the high clutter of the Baltic data set leads to a much higher false alarm rate.

The results clearly demonstrate the need for a more elaborated seabed classification. The Flat Seabed on both missions is of very similar nature but the Baltic mission has a much larger clutter density. This induces a much larger number of false alarms comparable to that obtained for complex seabed. For the Complex seabed class, the algorithm performs in a similar fashion in both missions as the complex seabed is similar to a high clutter density. This suggests that the clutter density – seabed complexity axis are not really orthogonal in terms of ATR performances and that a C1 (high complexity- low clutter) could in fact be equivalent to an A3 (Low complexity-high clutter). We would strongly support a study in this direction in the future to create a more

realistic seabed norm for Side Scan and Synthetic Aperture Sonar.

We can quantify the false alarm reduction rate of the CAC algorithms compared to the CAD for a similar PD. The results obtained are the following for the BP02 data:

- Complex Seabed: 50% reduction. From 100000 to 50000 per square nautical mile
- Flat Seabed: 50% reduction. From 600 to 300 per square nautical mile
- Ripples: 90% reduction. From 10000 to 1000 per square nautical mile.

The higher reduction for ripples can be explained by the fact that the ripples are largely of a different scale than the targets (~15cm for the ripples, 30-45cm for the targets). The false alarms on the complex and flat seabed are of a different nature (sonar artifacts for flat, mine like objects for complex) and more difficult to remove.

C. How can this framework be used operationally?

Within an operational context, the approach presented here will prove very useful. The AUV will first perform a mission, collecting all side scan and navigation information. The Seafloor Classification and Clutter Density will be estimated on each of the side scan sonar images to produce the seafloor classification mosaics fusing the seafloor and clutter density information. Representative Images containing areas of the different seafloor types present within the survey region will be selected to perform the evaluation. On this subset, Z (the Height Map), B (the Beampattern) and R (the Reflectivity) can be extracted. A Monte Carlo Simulation where the simulated targets are placed within the images can then be performed. Various target types, aspect angles and burial depths can be simulated. Finally, the ATR system can be run on the generated images to produce PD/PFA curves for the different seafloor types present in the area.

VII. CONCLUSIONS AND FUTURE WORK

In this paper a framework for evaluation and planning AUV missions for MCM has been presented. It adopted a 'systems' approach of the evaluation. The framework cornerstones are a sonar target simulator based on augmented reality and a seabed classification module enabling the determination of target huntability indexes. The former uses real data through a sonar inversion model to retrieve an estimate of the bathymetry and the reflectivity of the seabed. It enables the simulation of very realistic mine fields. The framework has been assessed on two real AUV missions by evaluating our exemplar CAD-CAC algorithms, but in principle it can be used to evaluate any other ATR system or human operator.

Further research will concentrate on improving the Z, R and B inversions by reducing the parameter space considered and improving the optimization and regularization schemes. The scattering model used for the targets will be enhanced to include specular reflections. The system will also be extended to use bathymetry from other sources such as multibeam echosounders and interferometric side scan. Finally the system could be extended to Synthetic Aperture Sonar.

VIII. ACKNOWLEDGEMENTS

The authors would like to thank The NATO UnderSea Research Centre for providing data and support for this study. Further thanks to the British MoD and SEA for supporting part of this work.

REFERENCES

- [1] Autonomous Underwater Vehicle and Ocean Modeling Networks: GOATS 2000 Conference Proceedings, CP-46, 2000, pp 299-361.
- [2] T.D. Ross, J.C. Mossing, "MSTAR Evaluation methodology", Proceedings of the SPIE, volume 3721, pp705-713, 1999.
- [3] G. Elston and J. Bell., "Pseudospectral time-domain modeling of non-Rayleigh reverberation: Synthesis and statistical analysis of a side scan sonar image of sand ripples. IEEE journal of Oceanic Engineering, vol. 29, issue 2, pp 317-329, 2004
- [4] D. Langer, M. Hebert, "Building Qualitative Elevation Maps From Side Scan Sonar Data For Autonomous Underwater Navigation," in Proceedings of the 1991 IEEE International Conference on Robotics and Automation, vol.3, pp. 2478 -2483.
- [5] B. Zerr, B. Stage, "Three-Dimensional Reconstruction of Underwater Objects from a Sequence of Sonar Images," in Proceedings of the 1996 IEEE International Conference on Image Processing, vol.3, pp. 927-930.
- [6] A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum Likelihood from incomplete data via the EM algorithm", J. Royal Statistical Soc., Ser. B, vol. 39(1), pp. 1-38, 1977.
- [7] E. Coiras, Y. Petillot, D. M. Lane, "Automatic Rectification of Side-scan Sonar Images", UAM'05, 2005.
- [8] G.J.Dobeck, J.C.Hyland, L.Smedley, Automated detection/classification of sea mines in sonar imagery, Proc. SPIE-Int Soc. Optics,2079:90-110, 1997.
- [9] S.Reed, Y.Petillot, J.Bell, A model based approach to the detection and classification of mines in sidescan sonar, Applied Optics, Vol. 43, Issue 2, p237-246, Jan., 2004.
- [10] I.Quidu, Ph.Malkasse, G.Burel, P.Vilbe, Mine classification based on raw sonar data: An approach combining Fourier descriptors, statistical models and genetic algorithms, Proc. OCEANS 2000, Vol. 1, p285-290.
- [11] G.J.Dobeck, Algorithm fusion for automated sea mine detection and classification, Proc. MTS/IEEE Oceans Conf. and Exhibition, 1:130-134,2001
- [12] M.Mignotte, C.Collet, P.Perez, P.Bouthemy, Markov Random Field and fuzzy logic modelling in sonar imagery: application to the classification of underwater floor, Computer Vision and Image Understanding, Vol. 79, p4-24, 2000.
- [13] S.Reed, Automatic detection and classification models for sidescan sonar imagery, PhD Thesis, Heriot-Watt University, Edinburgh, April 2004.
- [14] S.Reed, I.Tena Ruiz, C.Capus, Y.Petillot, The fusion of large scale classified side-scan sonar image mosaics, Accepted for publication in IEEE Trans. Image Processing 2006.