_ image correction
- cause for image correction
- segmentation
- feature extraction
- CAD.
- CAC.
- SONAR
- SSS
- shadow & highlight
-

Image Processing Techniques for the Detection and Classification of Man Made Objects in Side-Scan Sonar Images

Esther Dura
Institute of Robotics and Technologies of Information and Communication,
University of de València
Spain

1. Introduction

This is a review chapter that surveys past work in, and the recent status of image processing and other related techniques involved in the detection and classification of man made objects in side scan sonar images. Side scan sonar is a readily, available and cheap device which has found increasing applications, specially for military purposes such as Computer Aided Detection (CAD) and Classification (CAC) of mines. Therefore the main focus of the chapter is on this topic. The list of references is sufficiently complete to include most past and recent publications in the open refereed literature.

Although side scan sonar displays many features similar to an optical sensor from a purely image processing point of view, the basics of the physics and formation of the images are crucial for understanding the difficulties found when detecting and classifying mine like objects (MLO's) in side-scan sonar images. Therefore in the first part of the chapter a brief review of the principles of the side-scan sonar, image formation process and characteristics of the images are explained. Different types of sonar images as well as diagrams showing the the process of generating an image from a single diagram ping will be provided.

The classification and detection of MLO's is traditionally carried out by a skilled human operator. This analysis is difficult due to the large variability in the appearance of the side-scan images as well as the high levels of noise usually present in the images. With the recent advances of Autonomous Underwater Vehicle (AUV) automatic techniques, CAD/CAC of mines, are now required to replace a human operator.

In the literature the computer aided detection/classification (CAD/CAC) problem is not well defined as detection involves an element of classification (mine/not mine), therefore these terms must be defined. For the purpose of this work, we will consider detection as the process of identifying a mine and classification will be a further step where the aim is to determine the shape of such a mine. Therefore the second part of the chapter is divided into two main sections: 1) detection 2) classification of MLO's.

In the last part of this chapter a review will be done on the current state of fusion of multiple algorithms aiming to overcome the limitations and weaknesses of every single CAC/CAD algorithm reviewed in the previous section.

2. Principles of the side-scan sonar: image formation process and characteristics

The section starts by introducing the basic principles of sonar. The following sections present the fundamental side-scan sonar characteristics along with the image construction, characteristics of the images, frequency and resolution of the images.

2.1 Basic principles

Devices which use underwater sound for communication or observation are generally referred to as SONAR systems. This term was coined after the Second World War to provide analogy to the equivalent electromagnetic-echo location system of radar and is an acronym for "SOund Navigation And Ranging".

In general the basic principles of a sonar involves the transmission of a pulse energy into the water medium and the subsequent reception of any returned energy reflected from objects or seabed.

Basically the sonar generates a short electrical pulse, in the form of an acoustic wave centred at particular frequency, length and energy, by the transmitter. This electrical signal is transformed by the transducer, which is normally a piezo-electric ceramic, into mechanical vibration energy. This vibration is transferred into the water as an oscillating pressure, the pulse. The pulse travels trough the water until it is reflected back or scattered by the seafloor or any object. The energy reflected back, which is mechanical energy, is converted by the transducer into electrical energy. This energy is then detected and amplified by the receiver of the sonar.

There is a master unit, with a control function, in charge of synchronizing the operations and control timing for the transmission and reception of the electrical signals. The control unit normally has a unit to display the received data.

It should be noted that what the sonar is measuring is the time that it takes for the transmitter sonar pulse to travel from the transducer to the target and return. It is not measuring the depth or distance.

2.2 Side-scan sonar characteristics

The fundamental purpose of a side-scan survey is to provide images which map a visible representation (intensity of marking) of the strength of acoustic back scattering, from the sea floor onto a two-dimensional image medium, by the process illustrated in 1. These sensors are usually mounted onto a separate body which is towed through the water behind the survey vessel. Alternatively the transducers may be mounted onto Remote Operated Vehicles (ROV) or Autonomous Underwater Vehicles (AUV) allowing more accurate positioning and motion of the vehicle. The characteristic of the side-scan sonar comes from implementing the basic principles mentioned. The main feature of this sensor, as can be seen in figure 2 is that is a side-ways looking device. Each pulse of acoustic energy emitted causing echoes from an area of the sea bottom perpendicular to the direction of travel of the tow fish. The transducers are normally shaped and controlled to produce a beam for each emitted pulse which is narrow in the horizontal direction and wide in the vertical direction as illustrated in figure 2. Due to the narrow horizontal beam, returned energy is received from one strip of the seafloor. The wide vertical across trace beam permits the ensonification of a large area of the seafloor. Another of the characteristics of the side-scan sonar is that sometimes two channels are used to gather information at the same time from the seabed on either side of the tow fish.
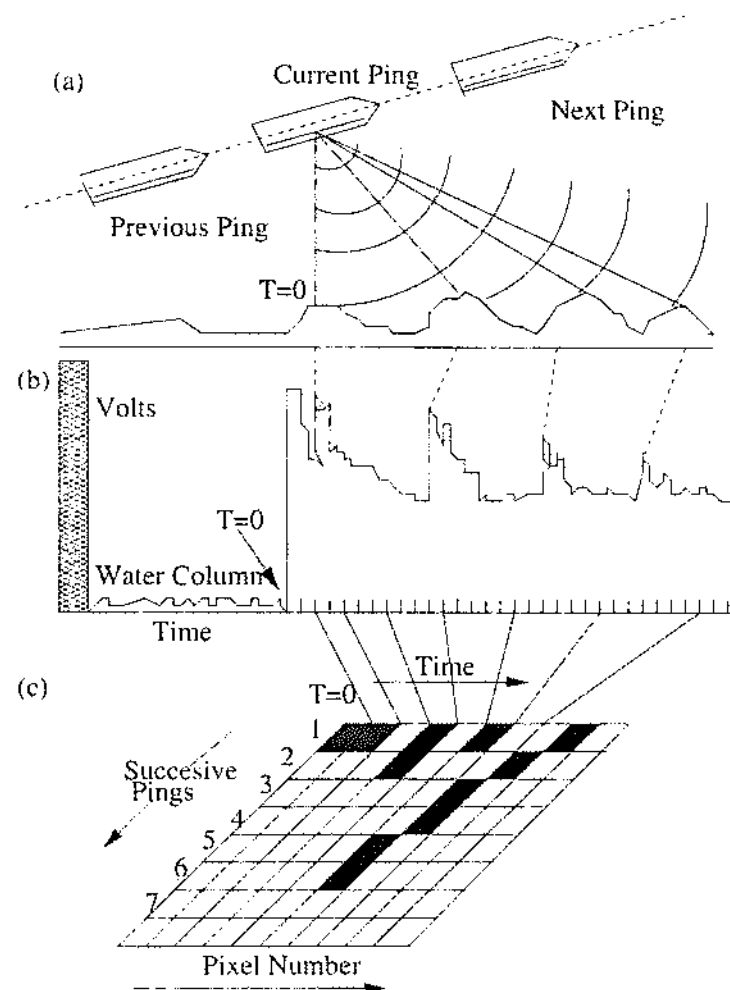
Fig. 1. Three-phase diagram showing the process of generating images from a single side-scan sonar ping. In the top diagram (a) the outgoing pulse from an individual ping is reflected back from the seafloor directly under the fish, and the internal side-scan clock (T=0) is started. (b) The hatched region represents the outgoing pulse, and the low amplitude returns are the time when the pulse is the two-way travel time in the water-column. After the return of the first bottom bounce, subsequent returns appear as peaks and valleys in the transducer voltage. (c) Peaks and valleys are then integrated and translated into pixels values.

2.3 Beamwidth

It has been pointed out that the side-scan sonar has a beam that is narrow in the horizontal plane and broad in the vertical plane. For a typical system this must be 1 degree horizontal and 40 degree vertical beam.

The beamwitdh can give some idea of the resolution which a sonar will achieve. It is also very important to consider the overall beam pattern of a particular sonar. This will be a truer representation of the the expected behaviour.

The shape of the beam is the result of the transducer design. A side-scan transducer it normally consists of a line array of crystal elements. Each point on the faces of the crystals acts as a sound radiator. It can be though of each infinitesimal point sending out a sound
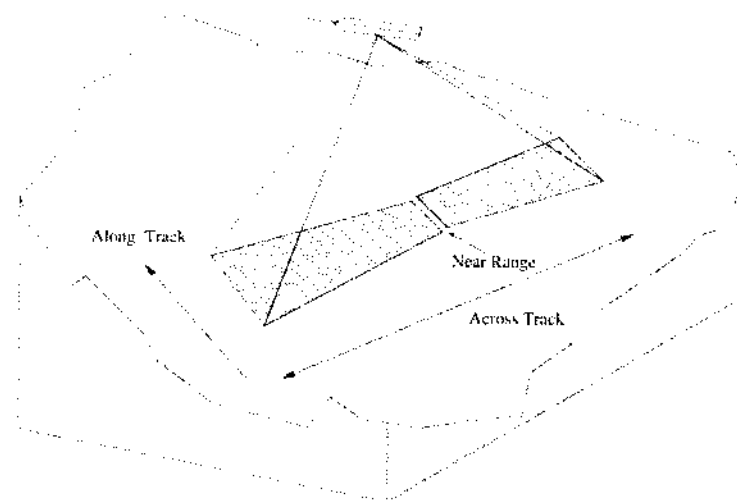
Fig. 2. Schematic diagram of the acoustic 'footprint' of a side-scan sonar system. The gray area represents the swath

pulse that spreads out in all the directions. Looking at just two of these points, as depicted in 3, it can be seen that some distance from the transducer the pressure disturbances from each of the points will meet and will either add or substract from each other depending on the phase. This process is going on for all the points of the transducer face. The net effect of all additions and substractions is to produce the beam pattern. Along the axis of the beam the pressure contributions are reinforced, while on the sides they tend to cancel. For a line of array the beamwidth can be expressed as $50.6\lambda/L$, where λ is the wavelength of the acoustic pulse($\lambda = $ [sound velocity/$frequency$])) and L is the array length.
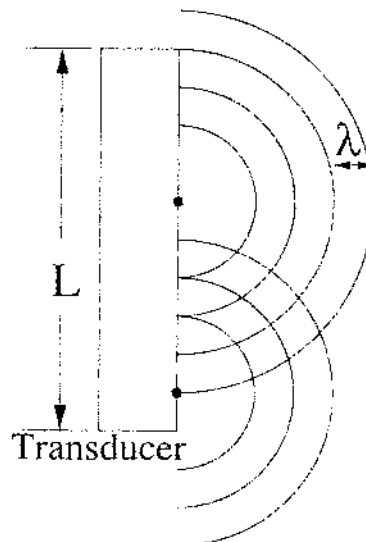

Fig. 3. Beam formation process

2.4 Image construction

The interpretation of the side-scan sonar image, requires an understanding of the image formation process. As the transducers are towed along, they gather sequential lines of data

returned from each pulse, and these lines ('A' scan) are displayed sequentially down a vertical trace to generate an image. This systematic sideways scanning is the basic principle of side-scan sonar. The scanning occurs along track and across track. The data gathered along track is a function of the beam width, the pulse repetition rate and the tow speed. Across track the intensity is received with successively increasing two way travel times or time of flight. The intensity received is dependent on attenuation of sound in water, the direction and angle from which the target was ensonified and the reflectance properties of the seafloor. With rock and gravel acting as stronger reflectors than soft sediments such as mud and sand.The intensity of the return is displayed against the two way travel time, or time of flight.

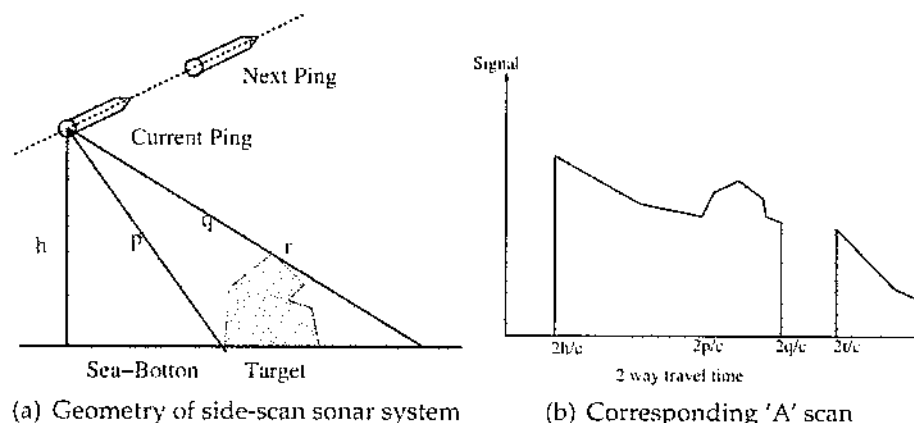(a) Geometry of side-scan sonar system   (b) Corresponding 'A' scan

Fig. 4. Geometry of side-scan sonar system and corresponding 'A' scan

Figure 4(a) illustrates the geometry of the side-scan sonar system and figure 4(b) displays the returned intensity against the two way travel of time in the form of a 'A' scan. An 'A' scan is simply one line of a sonar images corresponding to the returned energy from a narrow strip of the seabed due to the reflections from one emitted pulse.

At the beginning of the trace there is a blank area, as the pulse propagates through the water column without returning any echo. The first bottom return is the first echo to return from the sea bottom closest to the transducer. For a relatively flat seabed the first return is from the seabed directly below the transducer and it occurs at approximately time $2h/c$ seconds where h is the height of the transducer in metres and c is the velocity of sound. The first return is then followed by successive echos at successively increasing slant range, where the slant range is the actual distance from the sonar to the point of the seabed from which the sound was reflected. These points are followed by successive echoes at increasing slant ranges as the sound wave propagates, as illustrated in figure 4 (b).

Once the "spike" of the high-amplitude bottom bounce is received, the side-scan processor begins to divide the transducer voltage time series, which is produced by the subsequent bottom return signals, into equally spaced "time" slices. Because of the geometric effect illustrated in figure 1 and 2, these time slices represent extremely narrow regions of the seabed for the early returns and much wider regions for the later returns. Within each time slice, the varying voltage of the transducer represents the acoustic energy from a fairly large area of the seafloor, and are much larger than that represented by the pixel size of the final image. The Voltage within each individual time slice is averaged (see figure 1 (b)) and then converted to a single digital number that is assigned to a specific pixel location as illustrated in figure 1(c). In practice, the conversion from uncorrected transducer voltage to spatially correct pixel value is more complicated than this description. The signal received from the seabed return

is normally passed through an amplifier with a time varying gain (TVG). This compensates for the effects of absorption of sound by water and the geometric effects of spreading and scattering.
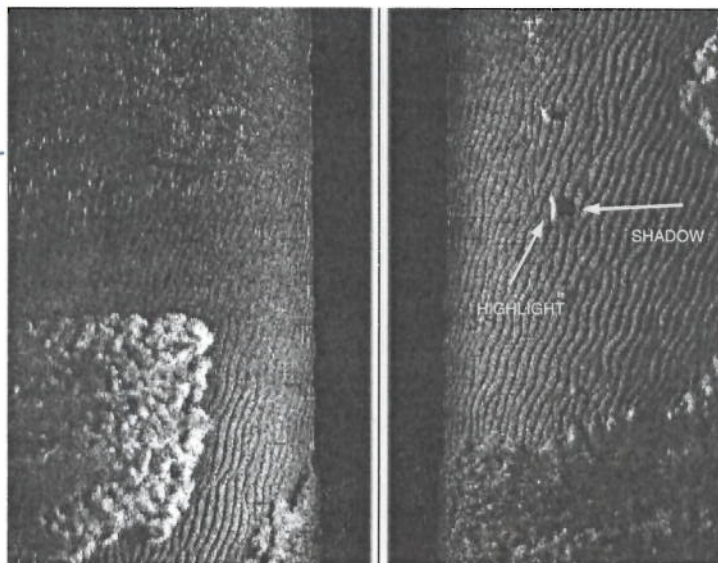
The process described, varies significantly from system to system, and requires a variety of corrections to become an intelligible image. For most side-scan systems, there are approximately 1024 pixels per side, or 2048 total pixels in the full swath. Depending on the system, each pixel value is usually an 8-bit integer (ranging from 0 to 255, or 256 possible shades of gray), which represent the value of the received acoustic echo after detection and the electronic low-pass filtering associated with time slice averaging.

3. Characteristics of side-scan sonar images

As mentioned above the side-scan sonar images are typically displayed as grayscale images, with dark and bright areas representing features of the seabed and water column. The orientation of the target relative to the direction of the incoming pulse will influence the intensity of the reflected signal and consequently the intensity in the image. The closer the inclination of the surface normal of the target to the direction of the incoming pulse, the greater the energy. Objects protruding above the seabed will create high intensity returns, highlight (see figure 5) but will prevent the sound from reaching the seafloor for some distance behind them. This will produce an acoustic shadow in the images (see figure 5) and will appear on the trace as blank area. Shadows can also be generated by depressions on the seafloor or by the self shadowing of the seafloor.

Fig. 5. Example of side-scan sonar images containing man-made objects: a mine like object is identified by a highlight followed by a shadow region

Shadows are one of the primary features which provide three dimensional information and their position and shape contain valuable information for the accurate interpretation of the images.

The side-scan sonar images, as it was mentioned, are essentially a 'picture of the seafloor' but they are usually distorted. In order to become recognizable, image pixels need to be corrected for a variety of effects. These include slant range correction (compensating for the equal time slice interval, which result in unequal distance slice interval), absorption of the sound in sea
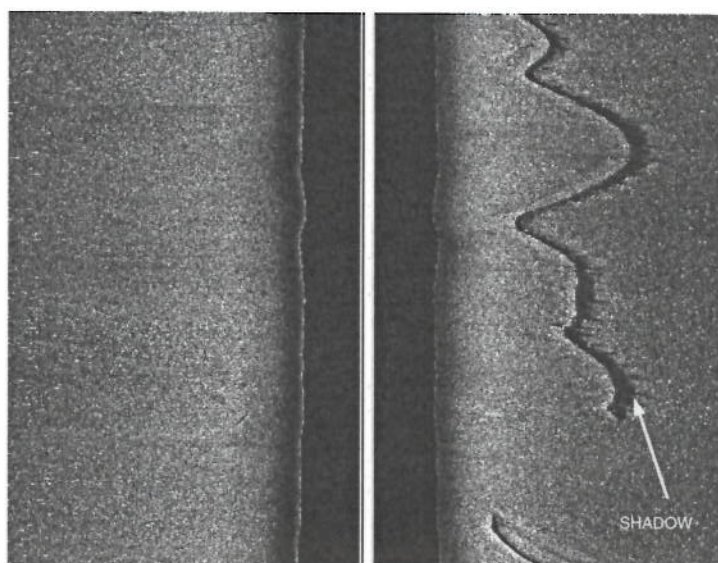
Fig. 6. Example of side-scan sonar images containing a man-made object: a pipe

water, the geometric effect of spreading (compensated by TVG), noise (potentially associated with telemetry errors, multipath effect which occurs when signals arriving at the same time as normal backscatter signals from beyond the target are superimposed resulting in the elimination of the shadow in the record, and side-lobes returns for both horizontal and vertical plane. Horizontal side lobes can direct significant energy in other directions which could then reflect of the seabed/target not perpendicular to the direction of travel and produce reflected energy. On the other hand in the vertical plane we want the main beam to give uniform energy only over the are we are interested in. If we have side lobes they may be for instance be pointing towards seasurface. This can produce backscattered data from sea surface which can interfere with backscattered signal from seabed. Because sea surface is a good reflector this produce noise. This is probably a much larger cause of noise than horizontal effects which may cause more blurring of the image than noise. External interference is caused by other acoustic devices operated at the same time which are added to the process Cervenka & de Moustier (1993), and variable ship speed. The final image after the corrections has a 1:1 aspect ratio and one that has the sonar targets roughly positioned at the same location on the chart as they are in the seafloor. Further details of these corrections can be obtained from Johnson & Helferty (1990)Somers & Stubbs (1984)Mazel (1985)Bell (1995)Cervenka & de Moustier (1993)

4. Detection of mines

The general approach for detecting targets is a two-tier process: 1) Detection of possible MLO's (regions of interest (ROI)). 2) Classification into mine or not-mine like objects with a low detection rate of false alarms. Both stages are crucial in order to get a good detection rate. For the first stage, detection, several approaches based on segmentation techniques and matched filters are reviewed.

The various techniques used in the literature for the second stage, classification, can be divided into three categories: unsupervised, semi-supervised and supervised algorithms. A comprehensive review of these techniques will be done. These techniques require the extraction of mine features, therefore a review of the most discriminant features used in the literature will also be explained.

4.1 Detection of possible MLO's

As stated before this first step involves identifying regions of interest that may contain a mine. Two main approaches are used in the literature for this purpose: 1) segmentation and 2) matched filtering.

4.1.1 Segmentation

Segmentation is the process of classifying pixels as belonging to a certain class. In side-scan sonar mages the classes of interest normally are: highlight and shadow. Because of the shadow cast by a side-scan sonar appears more consistent than the highlight, some of the most successful algorithms rely only on the shadow information. The main techniques used for segmentation are: 1) thresholding 2) clustering and 3) Markov Random Fields.

Thresholding is the simplest method on image segmentation. During the thresholding process, individual pixels in a side-scan image are marked as shadow pixels if their value is greater than some gray pixel value (assuming shadow to be darker than the background) and as no-shadow pixels otherwise Quidu et al. (2000). Some approaches use two thresholds values to segment images into shadow, highlight and background regions.

Clustering a procedure to determine the intrinsic grouping in a set of unlabeled data, has also been used to segment the images into three categories (shadow, highlight and background). In this technique, a feature vector for each pixel of the image is extracted and then a similarity metric is used to cluster vectors having similar features Guillaudeux et al. (1996).

Unlike previous methods, *Markov Random Fields* provides a reliable framework for incorporating pixel dependencies into the segmentation (i.e a pixel surrounded by a shadow is most likely to belong to shadow itself). This ability to model inter-spacial dependencies between pixels has ensured the use of MRF models for a range of applications. In the context of side-scan sonar images where there is a large variation in the appearance of the images, more complicated models have been used Mignotte & Collet (1999)Reed et al. (2003). These models include a priori knowledge: object highlight generally lies close to shadow regions. One of these studies, Reed et al. (2003), introduced the size and appearance information as a priori information into the model. In this study, after the MRF segmentation, a further post-segmentation step that provided an accurate and robust method for extracting the shadow and highlight was carried out by using a cooperative statistical snake. The model segments the object-highlight and the shadow region by considering the image as being composed of three different statistical regions. The main advantage that this method presented compared to other models, was that using a priori information on the relationship between the object-highlight and shadow, accurate segmentation was achieved on seabed types where other models failed. Details of this implementation can be found in Reed et al. (2003)

An example of MRF segmentation on side-scan images containing mines can be visualized in figure 7(b).

4.1.2 Matched filtering

It is a technique for finding small parts of an image which match a template. This is done by convolving a known template with an image to detect the presence of the template in the image. The identification of mine-size regions in the sonar image has been carried out, as explained inDobeck (1997), by convolving a template that contains four distinct regions:(pre-shadow, highligth, dead zone, shadow and post-target) with the image. After that a threshold is applied to the post-processed images and neighbour pixels over a threshold are grouped together to identify possible MLO's. The threshold varies between the different

(a) Original side-scan image

(b) Segmentation result using the MRF model proposed by Mignotte & Collet (1999)
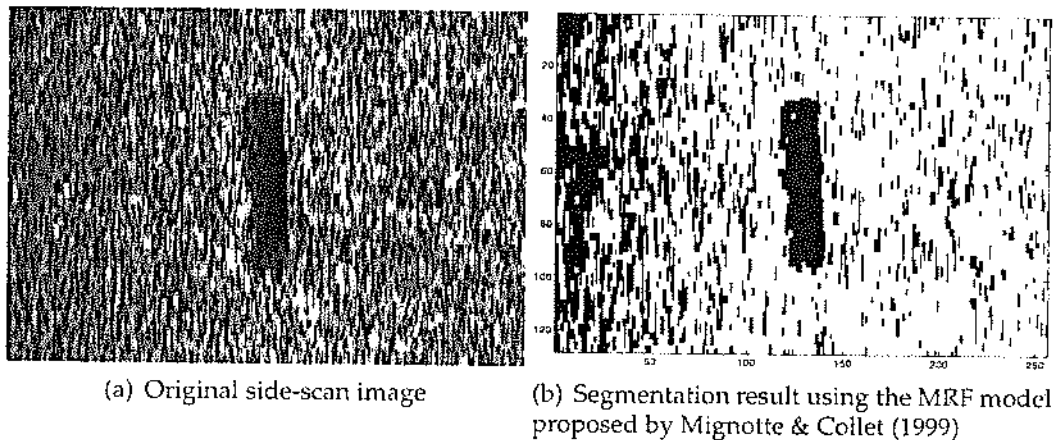
Fig. 7. MRF segmentation

detectors, and is fixed according to the desired sensitivity of each detector. Details of this implementation and similar approaches can be found in several works Dobeck (1997)Dobeck (2000)Hasanbelliu et al. (2009).

4.2 Classification into mine or not-mine

By using any of the techniques explained above the regions of interest of the image, that may contain MLO's, are identified. Afterwards these regions of interest have to be classified into mine or not mines. This classification procedure normally requires the extraction of mine features, therefore a review of the most discriminant features used in the literature will first be explained in the following section.

4.2.1 Feature extraction

In pattern recognition feature extraction is a special form for reducing dimensionality of an image. For side-sonar images the aim of feature extraction is to extract some characteristics that describe a region of interest that may contain a MLO. The main feature used in the literature Aridgides et al. (2001a)Dura et al. (2005)Dobeck (1995)Zerr & Stage (1996)Quidu et al. (2000) for extracting features fall into two categories: 1)shape features 2) gray-level features

Shape features characterize the appearance and geometry of an object. MLO as opposed to non man-made object cast regular shadows and highlight of anticipated dimensions. The following features are mainly used in the literature for extracting some features from the shadow and highlight information:

1) Area: is the surface area of an object , O,(shadow or highlight), defined as:

$$Area = \sum_{i,j} O(i,j)$$

where O(i,j) has a value of one for a pixel in the object and zero if not.

2) Elongation represents the ratio of major axis to that of the minor axis. It is computed from second order central moments as:

$$Elongation = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{\mu_{20} + \mu_{02}}$$

where μ_{pq} stands for the central moment of order $p + q$ which is computed as:

$$\mu_{pq} = \sum_{i=0}^{M} \sum_{j=0}^{N} (i - i_g)^p (j - j_g)^q I(i,j)$$

where (i_g, j_g) is the position of the center of mass of the shadow or highlight and $I(i,j)$ is the digital image. This position is calculated as:

$$i_g = \frac{m_{10}}{m_{00}}$$

and

$$i_g = \frac{m_{01}}{m_{00}}$$

and the two-dimensional moment m_{pq} of order $p + q$ is defined as

$$m_{pq} = \sum_{i=0}^{M} \sum_{j=0}^{N} i^p j^q I(i,j)$$

③ Circularity or shape factor: is a measure of circularity or the compactness of a shape and can be calculated as:

$$Compactness = \frac{4.\pi}{p^2}$$

④ Orientation: the orientation of an object can be defined as:

$$Orientation = \frac{1}{2} tan^{-1} \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}$$

⑤ Eccentricity: Is the ratio of the length of the longest chord of the shape to the longest chord perpendicular to it.

⑥ Rectangularity: This shows how well a region is approximated by a rectangle. The rectangularity measure f_{rect} is the ratio of the area of a region, A, to area of the smallest rectangle, $A_{rectangle}$, that encloses it:

$$f_{rect} = \frac{A}{A_{rectangle}}$$

⑦ Number of zero crossing of the curvature at different scales: for a fixed length, a small number of curvature zero crossing suggests as simple regular contour; a high number, suggests a tortous, irregular, frequently turning contour. Therefore scanning the curvature scale space of a given set of contours from fine to coarse scales, regular shapes can be identified.

The number of zero crossing of an image is obtained by applying the convolution operator $\nabla^2 G$ (which is the Laplacian of a two dimensional Gaussian $G(i,j)$) over the image $I(i,j)$ as:

$$I(i,j)' = \nabla^2 G(i,j) * I(i,j)$$

where

$$\nabla^2 G(i,j) = (\frac{r^2 - \sigma^2}{2\pi\sigma^2}) exp \frac{-r^2}{2\sigma^2}$$

where

$$r = (i^2 + j^2)^{\frac{1}{2}}$$

Some individual features and others that relate shadow and highlight information has also been used such as:

Shadow high profile, ratio of the highlight to shadow area, ratio of highlight to shadow height, minimum distance between highlight and shadow, horizontal aligment between shadow and highlight

When the quality and resolution of the images is low, these are not well characterize by the profile (shape features) of the shadow and highlight. Therefore gray level features extracted from the shadow and highlight are also used for discriminating targets from clutter. Among them the most discriminative features used are:

1) *Average shadow* strength, which is a measure of the object's shadow darkness

2) *Average highlight strength*, which is a measure of the object's highlight brightness

3) The *variance* of the *shadow*

4) The *variance* of the *highlight*

5) *Contrast* between *shadow* and *highlight*, which is the absolute difference of the average shadow strength and average highlight strength

6) *Contrast* between *shadow* and *background*, which is the absolute difference of the average shadow strength and average background strength

7) *Contrast* between *highlight* and *background*, which is the absolute difference of the average highlight strength and average background strength

4.2.2 Classification

At this step classification is the detection of MLO's in side-scan images. There various techniques examined can be broadly divided into three groups: supervised, semi-supervised and unsupervised.

Supervised

With these techniques, one typically requires an a priori-set of training data consisting of a set of features and associated binary labels(mine/clutter) and a testing set to validate the results. A supervised learning algorithm analyzes the training data and produces an inferred function, which is called a classifier. To constitute a training set, known targets (e.g mines) must be emplaced in a given environment and side-scan data collected with all nonemplace scatters are assumed to be clutter. In the context of sonar images, the difficulty of this supervised classification procedure resides 1) in the very large number of mine types, mine deployments and orientations, and mines history; 2)the significance dependence on the nature of the training data, specially the dependence of the imagery on the sea bottom environment. The variability of 1) and 2) makes it impossible to constitute a training set that is robust to all type of mines and environments to be encountered. An algorithm trained for one type of sonar setting may perform poorly when used in another environment. Besides, the point at which a 'training data set' becomes sufficiently large is difficult to define. To overcome this problem some researchers Reed et al. (2004) and Coiras et al. (2007) have generated their data set of synthetic side scan images with inserted random mines at random locations and orientations. The mines inserted had realistic shadows and highlights that took into account the angle of incidence and topography of seabed.

Pioneering research on supervised detection/classification of MLO's was carried out by Dobeck (1997).They used a K-nearest neural network (KNN) and an optimal discriminatory filter classifier (ODFC). The KNN technique involved a two-layer neural network, which

classifies features according to the proximity of the features to a 'feature vector center'. These classifiers were then combined to yield the final result. Results and details can be found in Dobeck (1997)

An adaptive filter has also been used for the detection/classification of mines based on a bayesian classifier (simple probabilistic classifier based on applying Baye's theorem, with strong independence assumptions) known as the log-likehood ratio test (LLRT)Fernandez et al. (1993). A given feature vector was assigned as belonging to either of two classes on the basis of the LLRT. This algorithm addressed the main shortcoming of the bayesian classifiers (the determination of the multidimensional distributions essential for the computation of the LLRT) by mapping the sets of learning vectors to a space of orthogonal features in order to yield histograms. These histograms were then used to get the log-likehood ratio and summed to obtain the final results.

Linear and quadratics classifiers has also been employed for supervised classification by Fawcett (2001).A linear classifier separates objects or events by a linear function whereas the quadratic classifier separate objects or events by a quadratic surface. Unlike previous methods, in the work proposed by Fawcett (2001) instead of extracting features from the regions of interest the whole image was used as a feature vector. Principal components technique was used to identify the most discriminant features of the image. Details of the classifiers, implementation and results can be found in Fawcett (2001)

Recent machine learning techniques based on kernel-based algorithms such as Support Vector Machines (SVM)Vapnik (1995) and Relevance Vector Machines (RVM)Tipping & Smola (2001) have been investigated. These kernel-based learning algorithms are based on mapping data from an original input space to a kernel feature space of higher dimensions to solve a linear problem in that space. The advantages for relevance vector machines over support vector machines is the availability of probabilistic predictions, using arbitrary kernel functions and not requiring to set many parameters. Details of the implementations and results can be found in Dura et al. (2005)Couillard et al. (2008)

It is important to highlight, that is not always necessary to use all the features; sometimes using a smaller is better than using a large set of features which are correlated. Therefore some all the supervised techniques reviewed used some optimisation procedures before the training process to determine the best combination of features.

Semi-supervised

Semi-supervised is a class of machine learning technique that make use of both labeled and unlabeled data for training. The amount of labeled data required for training is tipically very small compare to the amount of unlabeled data. The cost of acquiring data with the associate label data is expensive and may make a set of data infeasible, whereas the adquisition of adquiring unlabeled data is inexpensive. Therefore semi-supervised techniques introduce and important advantage: the cost for mine hunting operations is reduced. In the context of side-scan sonar this is very important as labeling the data is very expensive, a diver or unmanned underwater vehicle with a camera has to label it.

An active-learning algorithm based on semi-supervised techniques was first proposed by Dura et al. (2005). The algorithm, kernel-based, was developed with the goal of enhancing mine detection/classification of mines without requiring a priori data set. It was assumed that divers or unmanned underwater vehicles with a camera were used to determine the binary labels of a small set for a given side-scan collection. This set of data and associated label were used to train the algorithm. Information-theoric concepts were used to adaptively construct the kernel classifier and guide which data and associate label were most informative

in the context of of algorithm training (this information content is computed without a priori knowledge of the labels itself). In this work authors demonstrated that the number of data for which the associate label was required was very small relative to the number of potential targets in a given image.

Unsupervised

Most of the current automated systems, as stated before, require training data and thus produce poor results when the training data differ from the test set. The success of this systems depend on the similarity of the training and testing set of data. This has led research into unsupervised techniques that requires no training data. The main advantage of these systems is that they are able to cope with the large variability in conditions and seabeds seen in side-scan sonar images. Also, very important, the cost of mine-hunting operations is reduced. One of the most complete unsupervised and sucesfull systems implemented so far is the one implemented by Reed et al. (2003)Reed et al. (2004). The system was composed of two consecutive and complementary phases: 1) a MRF algorithm was employed to segment the raw side-scan image into regions of object highlight, shadow and background. A post-processing procedure was then applied to remove false alarms. Objects that were too large or small were removed. The height (calculated by using the shadow length and navigational data) was also taken into account to remove false alarms. 2)In a second phase a cooperating statistical snake model was use to consider each of the detected MLO's. The model assumed the highlight and shadow regions to be statistically separated, therefore it was enforced a dependency between the two snakes. Also their movement was constrained. If snakes expanded beyond MLO dimensions the MLO was identified as false alarm and removed. Good detection rates was obtained with this two-step unsupervised algorithm.
Another approach with tackled this problem was the one proposed by Mignotte et al. (2000). In this work a set of deformable template model which allow linear transformation were used to separate natural objects from man made objects in an image. The detection was based on a objective function measuring how well a given instance of a template fits the contents of the segmented image (previously the image was segmented using a MRF). If the result of the objective function was less than a certain threshold then the desired object was assumed to be present and the final configuration revealed the shape of the object.

5. Classification of MLO's

Once the mine has been the detected the following step is the classification. As stated before, classification is the process of recognizing the shape of a mine (type of mine). In sonar imagery, MLO's produce a shadow which represents a regular geometry shape. In particular the shadow cast by spherical mine almost always is an ellipse with different vertical and horizontal axis lengths. For cylindrical mines the associate shadow may be a rhomboid, rectangle or ellipse.
The classification problem has not been widely addressed in the literature. The few approaches that deal with this problem fall into two different groups: mono-view and multi-view classification depending on whether they make use of a single view or several views for determining the shape.

5.1 Mono-view classification
In general classical models Castellano & Gray (1990)Quidu et al. (2000)Delvigne (1992) consisting of feature extraction and classification have widely been used for mono-view classification purposes. First using a presegmented shadow a mine a set of features are

extracted from the shadow. Afterwards the set of features are normally presented as input to the classifier. Althougth feature based are appealing the performance of the classifiers depend to a significant extent upon the feature extraction.

A totally different approach based on available properties of the shape (as a prior model) and an observation model (likehood model) was proposed by Mignotte et al. (2000). In such terms they proposed two prototype templates, square and ellipse, along with a set of transformations, to take into account the shape variability of different for every type of mines. The classification of an object was based on an objective function measuring how well an instance template fitted the content of the segmented image.

Along similar lines, Balasubramania & Stevenson (2001) did some interesting work on model fitting. In this work it was assumed that the shadows from targets such as cones, cylinders and rocks were close to an ellipse. Hence the shadow shapes were modelled as ellipses. To this end the edges of the shadow regions were extracted and the elliptical parameter fitting was performed using Karhonen-Loeve method. Then the parameters were used as features to describe the ellipse. Although this approach is relevant for spherical mine-like shapes, it is not the best to provide good separation class.

In the work proposed by Dura et al. (2008)they also advocated for a model-fitting approach by modelling the mine-like shadow with a superllipse. Superellipse provide a compact and interesting approach for representing a variety of shapes. By simply varying the squarness of the function shapes such as ellipses, rectangles, parallelograms, ovals and pinched diamonds can be easily generated. Thus, based on these observations, a classification procedure was proposed based on the squareness parameter. The procedure extracted the contour of the shadow given by an Unsupervised Markovian segmentation algorithm. Afterwards a superellipse was automatically fitted by minimising an appropiate metric with the Nelder Mead Simplex optimization technique. Some results can visualized in figure 8

Another approach has recently been investigated by Reed et al. (2004) and Coiras et al. (2007). A synthetic database of side-scan sonar images was generated with a sonar simulation under different conditions: seabed types, mine orientations and sizes. Then a classifier Coiras et al. (2007) was trained on the features extracted from the synthetic images generated. Afterwards real side-scan sonar images were classified. In Reed et al. (2004) instead of using a supervised classifier, the Hausdorff technique was implemented to measure the resemblance between the features of the synthetics images generated and the real images.

5.2 Multi-view classification

Sometimes is possible to obtain an accurate classification relying on a single view of an object. However some uncertainty of the object true class remains. In particular for sonar images if more than view of an object is provided at different angles this uncertantity can be reduced. The fusion of multiples images for classifying an object has been investigated by various authors. One of the most extensive work on multi-view classification was the one undertaken by Fawcett et al. (2010). In this work they extracted two features sets corresponding to two different view side-scan images and they investigated two approaches for fussing this information: 1)fuse-feature and 2)fuse- classification. In the first approach the two feature sets were combined to form a large feature vector (CF). Then a kernel based classifier was trained a tested with the resulting extended feature vector. The second approach consisted of fussing the two individual-aspect classification of the two feature vectors using Dempster-Shafer (DS) Theory. DS, frequently used as alternative to Bayesian theory and fuzzy logic for data fusion, allows to combine evidence from different sources and arrive to a degree of belief (represented

(a) Rectangular shape   (b) Romboid shape
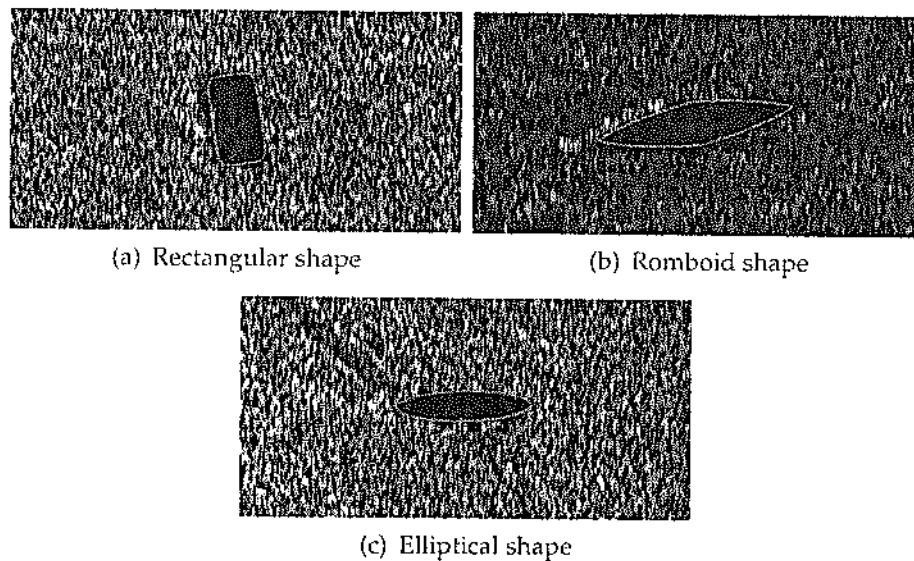

(c) Elliptical shape

Fig. 8. Results of the Superellipse fitted algorithm proposed by Dura for different shadows shapes

by a belief function) that takes into account all the available evidence. The belief function is derived from a mass function, which is analogous to well known probability density function. In this work, three different approaches were examined to calculated DS masses for each of the two looks of the classifier. In the first approach the histogram of the output labels for a single-aspect kernel regression classifier were used to empirically determine a simple analytical function which converted the values of the multiclas outputs into a set of masses. In the second approach they used a confusion matrix obtained from the single aspect classifier to specify the DS masses for various objects, for each of the looks, given the single aspect classifications. For the last approach, a nonempirical mass assigment based upon the relative values of the classifier outputs were considered.

Along similar lines Zerr et al. (2001) Stage & Zerr (1996)andReed et al. (2004) have also investigated the classification of a target by fusing several views using DS theory. However in the work investigated by Reed et al. (2004) the mass functions were generated from a fuzzy functions membership algorithm based on fuzzy logic.

6. Fusion of detection algorithms

The detection algorithms described in section 4.2.2 have their own weakness and strengths. This is due to the fact that each algorithm is based on different statistical properties and therefore emphasizes different characteristics of the data. Thus a combination of them may increase the probability of detection of MLO's and consequently reduce the number of false alarms.

Various methods of fusion of algorithms have been studied. In the work presented by Aridgides et al. (2001b), the results of different detection algorithms developed by three research teams (Naval Surface Warafe (NSWC), Coastal Systems Station (CSS), Raytheon and Lockeed Martin) were fused. Three different strategies were examined: 1)Logic-based fusion , 2)m-out-of-n fusion. 3) Log-Likehood Radio Test (LLRT)-based fusion algorithm. The logic-base fusion strategy was based on a set variety of rules including Boolean operators, AND, OR and their combinations. The m-out-of-n fusion was based of a particular instance

of this, '2-out-of-3'. This means that a target was included in the overall result if at least 2 out of the 3 algorithms tested also detected it. The LLRT-based strategy, originally developed by Fernandez et al. (1993)to perform perfusion, utilized the three confidence output vectors to form a three dimensional vector which were then processed through an orthogonalization and matrix extraction procedure to yield histograms for each orthogonal feature. These histograms were then used to obtain the log-likehood ratio and sum to obtain the overall result (detection). Best perfomance was obtained utilizing LLRT-based fusion which resulted in a 3:1 false alarm reduction improvement over the '2-out-of-3' strategy and 4:1 improvement over logic-based function.

Another different approach which also combined the output of the three detection algorithms previously mentioned, was the one investigated by Ciany & Huang (2000). The fusion algorithm received the two dimensional coordinates and confidence value for the detection mines and a geometrical clustering algorithm was applied. The resulting clusters were then processed via cluster confidence processing to produce the final fused results, which were the position of the mines. This procedure was applied as it was assumed that valid mines would be close by whereas false alarms would appear in random position of the image

Fusion detection algorithms based on score results of each individual algorithm have also been proposed. This can be performed by a number of ways such as was suggested by Dobeck (2005) 1) majority voting where the detections can be conditioned on thresholds applied to scores, 2) computing the sum of the algorithms scores and comparing the sum to a threshold, 3) computing a linear combination of the scores and comparing the weighted sum to a threshold. In this work was demonstrated that one can afford to run individual algorithms with higher probability of detection and higher probability of false alarm that would normally tolerate, in the knowledge that the fusion process will bring the false alarm rate down.

7. Conclusions

In this chapter the techniques involved in the detection and classification of MLO's on side-scan sonar images have been reviewed. The main components of CAD/CAC systems have been examined. These components are: 1)Image formation and characteristics of the image 2)Detection 3) Classification and 4) Fusion of different algorithms for detection of mines. For each component successful image processing techniques as well as related areas were examined.

However some questions remain: are the current automatic systems reliable enough to detect and classify mines without the assistance of a human operator?Do they perform well under different environment conditions?

8. Acknowledgments

9. References

Aridgides, T., Fernandez, M. & Dobeck, G. (2001a). Fusion of adaptive algorithms for the classification of sea mines using high resolution side scan sonar in very shallow water, OCEANS, 2001. MTS/IEEE Conference and Exhibition, Vol. 1, pp. 135 –142 vol.1.
Aridgides, T., Fernandez, M. & Dobeck, G. (2001b). Fusion of adaptive algorithms for the classification of sea mines using high resolution side scan sonar in very shallow water, OCEANS, 2001. MTS/IEEE Conference and Exhibition, Vol. 1, pp. 135 –142 vol.1.

Balasubramania, R. & Stevenson, M. (2001). Pattern Recognition for Underwater Mine Detection, *CAC/CAD 2001 Conference, Halifax, Canada*.

Bell, J. M. (1995). *A model for the simulation of sidescan sonar*, PhD thesis, Department Computing and Electrical Engineering, Heriot Watt University, Scotland, UK.

Castellano, A. R. & Gray, B. C. (1990). Autonomous Interpretation of side-scan sonar returns, *Proceedings of the IEEE Symposium on Autonomous Underwater Vehicle Technology Conference*, pp. 248–253.

Cervenka, P. & de Moustier, C. (1993). Sidescan sonar image processing techniques, *Oceanic Engineering, IEEE Journal of* 18(2): 108 –122.

Ciany, C. & Huang, J. (2000). Computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines, *OCEANS 2000 MTS/IEEE Conference and Exhibition*, Vol. 1, pp. 277 –284 vol.1.

Coiras, E., Mignotte, P.-Y., Petillot, Y., Bell, J. & Lebart, K. (2007). Supervised target detection and classification by training on augmented reality data, *Radar, Sonar Navigation, IET* 1(1): 83 –90.

Couillard, M., J.A.Fawcett, Myers, V. & Davidson, M. (2008). Support vector machines for classification of underwater targets in sidescan sonar imagery, technical memorandum drdc atlantic tm 2008-190, *Technical report*, Defence Reserach and Development Canada.

Delvigne, J. C. (1992). Shadow classification using neural networks, *Proceedings of 4th Undersea Defence Conference*, pp. 214–221.

Dobeck, G. (1995). Sea mine detection and classification using side-looking sonar, *Proc. SPIE*, Vol. 2496, pp. 442 – 453.

Dobeck, G. (1997). Automatic detection/classification of sea bed mines in sonar imagery, *Proceedings of SPIE. In detection and Remediation Technologies for Mines and Minelike Targets V* 3079: 90–110.

Dobeck, G. (2000). Algorithm fusion for the detection and classification of sea mines in the very shallow water region using side scan sonar, *Proceedings of SPIE. In detection and Remediation Technologies for Mines and Minelike Targets V* 4038: 348–354.

Dobeck, G. (2005). A probabilistic model for score-based algorithm fusion, *OCEANS, 2005. Proceedings of MTS/IEEE*, pp. 2429 – 2434 Vol. 3.

Dobeck, G. (2006). The k-nearest neighbor attractor-based neural network and the optimal linear discriminatory filter classifier, *OCEANS 2006*, pp. 1 –6.

Dura, E., Bell, J. & Lane, D. (2008). Superellipse fitting for the recovery and classification of mine-like shapes in sidescan sonar images, *Oceanic Engineering, IEEE Journal of* 33(4): 434 –444.

Dura, E., Zhang, Y., Liao, X., Dobeck, G. & Carin, L. (2005). Active learning for detection of mine-like objects in side-scan sonar imagery, *Oceanic Engineering, IEEE Journal of* 30(2): 360 – 371.

Fawcett, J. A. (2001). Image-based classification of sidescan sonar detections, *CAD/CAC 2001 Conference, Halifax, Canada*.

Fawcett, J., Myers, V., Hopkin, D., Crawford, A., Couillard, M. & Zerr, B. (2010). Multiaspect classification of sidescan sonar images: Four different approaches to fusing single-aspect information, *Oceanic Engineering, IEEE Journal of* 35(4): 863 –876.

Fernandez, M., Aridgides, A. & Bourdelais, J. (1993). Algorithm for sonar-based signal identification, *OCEANS '93. Engineering in Harmony with Ocean. Proceedings*, Vol. 3, pp. 438–443.

Guillaudeux, S., Daniel, S. & Maillard, E. (1996). Optimization of a sonar image processing chain: a fuzzy rules based expert system approach, *OCEANS '96. MTS/IEEE. 'Prospects for the 21st Century'. Conference Proceedings*, Vol. 3, pp. 1319 –1323 vol.3.

Hasanbelliu, E., Principe, J. & Slatton, C. (2009). Correntropy based matched filtering for classification in sidescan sonar imagery, *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, pp. 2757 –2762.

Johnson, H. P. & Helferty, M. (1990). The Geological interpretation of Side-Scan Sonar, *Review of Geophysics* 28(4): 357–380.

Mazel, C. (1985). *Sidescan Sonar Training Manual*, Klein Associates, Salem.

Mignotte, M. & Collet, C. (1999). Three-Class Markovian Segmentation of High-Resolution Sonar Images, *Computer Vision and Image Understanding* 76(3): 191–204.

Mignotte, M., Collet, C., Perez, P. & Bouthemy, P. (2000). Hybrid Genetic Optimization and Statistical Model-Based Approach for the Classification of Shadows Shapes in Sonar Imagery, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(2): 689–700.

Quidu, I., Malkasse, J., Burel, G. & Vilbe, P. (2000). Mine Classification using a Hybrid Set of Descriptors, *Proceedings of OCEANS 2000 MTS/IEEE Conference and Exhibition*, pp. 291–297.

Reed, S., Petillot, Y. & Bell, J. (2003). An automatic approach to the detection and extraction of mine features in sidescan sonar, *Oceanic Engineering, IEEE Journal of* 28(1): 90 – 105.

Reed, S., Petillot, Y. & Bell, J. (2004). Model-based approach to the detection and classification of mines in sidescan sonar, *Applied Optics* 43(2): 237– 246.

Somers, M. L. & Stubbs, A. (1984). Sidescan sonar, *IEE Proceedings* 131(3): 243–257.

Stage, B. & Zerr, B. (1996). Application of the dempster-shafer theory to combination of multiple view sonar images, *Proceedings. Undersea Defence Technology*, pp. 148–152.

Tipping, M. & Smola, A. (2001). Sparse bayesian learning and the relevance vector machine, *Journal of Machine Learning Research* 1(1): 211–24.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory.*, Springer-Verlag.

Zerr, B., Bovio, E. & Stage, B. (2001). Automatic mine classification approach based on AUV manoeuverability and cots side scan sonar, *Proceedings of Goats 2001 Conference, La Spezia, Italy.*

Zerr, B. & Stage, B. (1996). Three-dimensional reconstruction of underwater objects from a sequence of sonar images, *Proceedings of the IEEE International Conference on Image Processing*, pp. 927–930.