

SACLANTCEN MEMORANDUM
serial no.: SM-309

**SACLANT UNDERSEA
RESEARCH CENTRE
MEMORANDUM**



**AUTOMATIC TARGET CLASSIFICATION USING
MULTIPLE SIDESCAN SONAR IMAGES
OF DIFFERENT ORIENTATIONS**

B. Zerr, B. Stage, A Guerrero

September 1997

The SACLANT Undersea Research Centre provides the Supreme Allied Commander Atlantic (SACLANT) with scientific and technical assistance under the terms of its NATO charter, which entered into force on 1 February 1963. Without prejudice to this main task – and under the policy direction of SACLANT – the Centre also renders scientific and technical assistance to the individual NATO nations.

**Automatic Target Classification
Using Multiple Sidescan Sonar
Images of Different Orientations**

B. Zerr, B. Stage and A. Guerrero

The content of this document pertains to work performed under Project 031-3 of the SACLANTCEN Programme of Work. The document has been approved for release by The Director, SACLANTCEN.



Jan L. Spoelstra
Director

NATO UNCLASSIFIED

SACLANTCEN SM-309

intentionally blank page

NATO UNCLASSIFIED

Automatic Target Classification Using Multiple Sidescan Sonar Images of Different Orientations

B. Zerr, B. Stage and A. Guerrero

Executive Summary: Target classification and identification performance is a major factor in the effectiveness of mine counter measure (MCM) operations. In this report, the technique of simultaneously imaging a target from several aspects is investigated. The high frequency (typically > 400 kHz) high resolution (resolution cell typically < 200 cm²) sidescan sonar is the ideal sensor for fast route survey. However, when acoustic response (of the target and of the surrounding seabed) is aspect dependent, using solely the broadside view of the sidescan sonar limits the classification performance.

The study quantifies the target classification improvement for sidescan sonars capable of simultaneously recording a limited number of views, oriented from -45° to $+45^\circ$ to broadside.

The classification performance is evaluated on 10 target shapes, corresponding to proud mines and sinkers of moored mines. Different settings for the number and the direction of the views are compared using shadow based automatic classifiers. The classification results are computed on a larger number of natural and manufactured object images, generated by a dedicated modelling software.

The classification and identification of targets are closely related to the capacity of the feature vectors to discriminate between a given target, viewed with a given orientation, from all the other targets, irrespective of orientation. Using object cross-sections as feature vectors, the capacity to discriminate is established for a single view on target. Different configurations using up to three views with different angular intervals are subsequently compared.

The recommendation for an experimental assessment is a sidescan sonar which records simultaneously three views : the conventional broadside view and two additional views with programmable orientation from -45° to $+45^\circ$ to broadside.

Ana Guerrero was a Summer Research Assistant at SACLANTCEN during the summer of 1996. She is currently at The Faculty of Physics (U.C.M.), Av. Complutense s/n, Madrid 28040, Spain.

NATO UNCLASSIFIED

SACLANTCEN SM-309

intentionally blank page

NATO UNCLASSIFIED

Automatic Target Classification Using Multiple Sidescan Sonar Images of Different Orientations

B. Zerr, B. Stage and A. Guerrero

Abstract: In this report, the target classification performance of a multiple view sidescan sonar is investigated.

The classification statistics are estimated using model based automatic classifiers. The guidelines to the design of efficient classification algorithms are defined. The shadow is retained as the basic information for target classification. The input feature vector of the automatic classifier is the cross-section (or height profile) of the target estimated from its shadow.

The concept of multiple view sidescan sonar is presented and compared to other techniques for recording multiple aspects of a target. Several ways to modify a single view based classifier to process multiple aspects are identified and implemented. The task of the classifier is to recognize 10 target shapes, corresponding to proud mines and sinkers of moored mines. The classification results, expressed by ROC curves and confusion matrixes, are computed on a larger number of natural and manufactured object images, generated by modelling software.

The classification and identification of targets are closely related to the capacity of the feature vectors to discriminate between a given target, viewed with a given orientation, from all the other targets, irrespective of orientation. Using height profiles as feature vectors, the capacity to discriminate is established for a single view on target. Different configurations using up to three views with different angular intervals are subsequently compared.

Keywords: Shadow Based Mine Classification ◦ Automatic Target Classification ◦ Sidescan Sonar ◦ Neural Networks

NATO UNCLASSIFIED

SACLANTCEN SM-309

Contents

1	Introduction	1
2	Automatic Target Classification	2
	2.1 Automatic Target Classification in MCM	2
	2.2 Automatic Classification Techniques	2
	2.3 Preliminary Steps in Automatic Classifier Design	3
	2.4 Automatic Classifier Design : Feature Computation	5
	2.5 Automatic Classifier Design : Classification Algorithm	6
	2.6 Decision Logic for Classification Determination	14
	2.7 Quantification of Classification Performance	15
3	Multiple View Sidescan Sonar	17
	3.1 Using Multiple Aspects for Target Classification	17
	3.2 Acquisition of Multiple Views	18
	3.3 Classification Schemes for Multiple Views	22
4	Classification Performance	26
	4.1 Feature Vector Generation	28
	4.2 Training	29
	4.3 Preliminary Tests on Performance Stability	29
	4.4 Target Classification Performance	30
	4.5 Searching for Ambiguity in the training data	42
	4.6 Synthesis of Classification Performance	47
5	Capability to Discriminate between Target Shapes	49
6	Conclusions	64
7	Acknowledgements	65
	References	66
	Annex A - Simulation	69
	A.1 Overview	69
	A.2 Three-dimensional Objects Model	69
	A.3 Sonar Images Simulation	72

NATO UNCLASSIFIED

1

Introduction

Route survey is the mine counter measure (MCM) operation which takes advantage of the high coverage rate characteristics of sidescan sonar. The survey missions are conducted in operational areas during peace time to collect relevant information. The route survey data are usually acquired by towing a side scan sonar from a MCM vessel. The sidescan sonar can also execute mine hunting missions but the operating mode must be adapted to preserve the safety of MCM crews and vessels. The safety requirement can be complied with by towing the sidescan sonar from a drone [1]. A safer approach consists of mounting the sidescan sonar on an autonomous underwater vehicle (AUV). When the MCM operations are conducted by AUV's, the sidescan sonar images also contain information useful to the navigation of the vehicle. In known areas, the comparison between the sidescan sonar data and the reference data assists the AUV navigator in following an assigned path. In unknown areas, an incremental approach can be used to simultaneously acquire, compile and use the reference data.

The chief objective of route survey is to acquire, compile and maintain a data base containing relevant seabed knowledge information in potential operational areas. This data base is indispensable to the planning of MCM operations and a key component of the the mine warfare data centre (MWDC). The sonar raw data are processed to express the seabed properties at a scale and in a format appropriate to MCM. The efficiency of route survey requires accurate seabed sensing and consistent data management. For a given sonar resolution, recording multiple aspects improves the accuracy of the seabed sensing by taking into account the anisotropic features. Automatic processing of sonar images provides the required consistency for an accurate data base management. The current technology of automatic processing performs worse than skilled human operators. Despite this, automatic processing remains well adapted to the purpose. The lower level of performance of automatic systems is compensated for by the predictability of their outputs.

The objective of this report is to evaluate how multiple views improve the capacity to automatically recognize target shadows. For a given and limited number of views, the report addresses the determination of the angular intervals between them. The results of automatic target classification algorithms on modeled target images define the comparison for different configurations.

2

Automatic Target Classification

2.1 Automatic Target Classification in MCM

Developed in the past two decades by combining pattern recognition and artificial intelligence techniques, Automatic Target Recognition (ATR) has become a key component of the current generation of smart weapons (cruise missile, laser guided bomb, infrared targeting, etc.). Automatic target recognition is the computer processing of data from imaging sensors to identify image locations corresponding to specific targets [2, 3, 4]. The pattern recognition approach to ATR is based on statistical and structural techniques [2]. The basic block diagram of an ATR system is shown in Fig. 1. This ATR system is designed to be autonomous. In the MCM context, it can be applied to mine hunting or route survey missions performed by an AUV. In conventional hunting operations, the classical computer aids for mine detection and classification are obtained by partial use of the ATR system. The outputs of the second and fourth stages (see Fig. 1) correspond to computer aided detection (CAD) and classification (CAC), respectively.

This report focuses on target classification and therefore the assumption is made that the detection of potential mines has already been performed. The basic role of automatic target classification is to assert in if the detected object resembles a mine. Assuming that identification is merely a higher level of classification [5], the automatic classifier may achieve part of the identification task by defining the type of object.

2.2 Automatic Classification Techniques

Recent studies on mine classification exhibit a common approach which consists of applying the classification algorithm on a feature vector rather than directly on the image of the target. Doherty *et al.* [6] present a classifier based on a decision tree. The feature vector is composed of the object-to-shadow average intensity ratio and the estimated diameter of the object. In [7], Schweizer and Petlevich use two neural networks to perform the classification. The first neural network analyses the area surrounding the location to be processed. The result indicates membership of the following categories : background, shadow, highlight, highlight-shadow, texture and

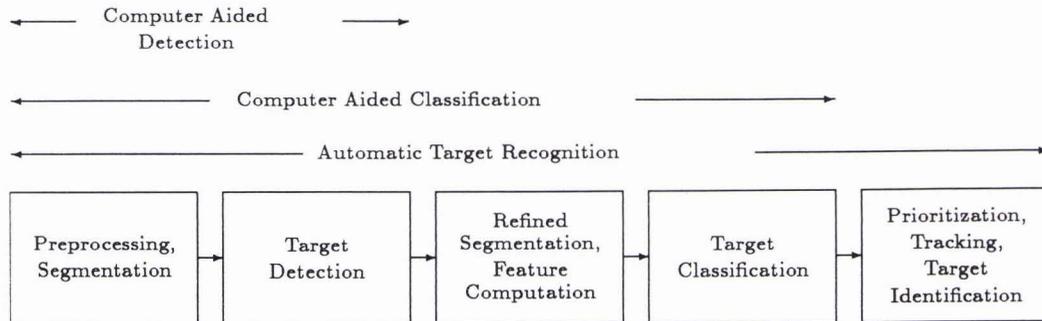


Figure 1: Diagram of an ATR System

anomaly. The input of the second network is the concatenation of three sources of information : (1) the output of the first network, (2) the result of a threshold based shadow-highlight detector and (3) the result of a statistical anomaly detection. Schweizer and Petlevich [8] use the features extracted from the two dimensional Fourier transform of the sonar image. Shazeer and Bello [9, 10] combine grey levels statistics (mean value, standard deviation, skewness and kurtosis coefficients) and object range in their feature vector. The aggregation of dimensional and intensity-based features leads to the comprehensive set of features described by Smedley and Dobeck [11]. Other approaches, based on fractal theory, estimate the fractal dimension of sonar image windows [12, 13].

When the sonar image is accurate, the classification can be based on the object shadow [14, 15, 16, 17]. Preliminary processing extracts the area of interest. Then, feature vectors are extracted from these areas and sent to the classifier. The extraction of feature vectors is based on invariant moments [14, 17], Fourier transform [14, 15], Hough transform [14], and estimation of the object cross-section (height profile) [15, 16].

2.3 Preliminary Steps in Automatic Classifier Design

Automatic mine classifier design requires definition of features extracted from the target response and the type of classification algorithm. To increase the system reliability, the design is enhanced by investigating (1) automatic classifier output, (2) the type of information acquired and (3) the selection of the target acoustic response component most suitable for robust classification.

NATO UNCLASSIFIED

SACLANTCEN SM-309

Automatic classifier output

When the feature vector is from a known target, the output of the classifier is a scalar value indicating the corresponding target class label. In other cases, the classifier output is "object not known". This output can be complemented by the confidence level of the classifier answer. In the first version of the multiple look classifier [18, 19], the output was divided into two subcategories for natural and artificial phenomena. Three classes of man made object (sphere, cylinder and truncated cone) and one class for natural objects (stones) were defined. This approach has now been abandoned, as natural features and mine like objects conform to different description methodologies. Mine like objects can be described in a deterministic way by the geometrical equation - or set of equations - of their shape. Natural objects are better described by stochastic processes. Unless a comprehensive description of the stochastic processes applicable to natural phenomena is available, the output classes correspond only to mine like objects.

Information to be acquired

Information is provided by modelling tools and not by experimental measurements. The modelling approach is particularly well adapted to the generation of target images for all the required points of views. In this study, modelling is restricted to the simulation of object shadows produced by high frequency¹, high resolution² sonars. For these sonars, knowledge of the acoustic target response ensures realistic images. Another advantage of sensor simulation is that parameters such as sonar resolution, target and seabed types and the number and orientation of views may be rapidly changed.

The model based approach has been preferred to more commonly used training methods based on sonar data, collected at sea by MCM or experimental sonars. Usually, experimental data, from a representative set of targets, are divided into two subsets : one for training and one for evaluating the correct classification scores. Because they process both target response and target interaction with the environment, these classifiers locally³ attain high classification performance. This level of performance can be extended to arbitrary locations only if the signatures of all known mines can be recorded for all possible types of environment. Time dependence (e.g. seasonal changes) must also be taken into account.

¹The typical frequency of high resolution imaging sonar for MCM ranges from 400 kHz to 1.3 MHz.

²The typical dimensions of a resolution cell vary from 10 to 40 cm across-range and from 2 to 10 cm along-range.

³the classification results are valid only for a given location (seabed type).

NATO UNCLASSIFIED

Basic Component of the Target Acoustic Response

The processing stages preceding target classification produce small sonar images likely to contain a target. Instead of attributing the same importance to all images pixels, the notion of basic information is introduced to select a pixel subset which carries pertinent information for target classification.

The basic components of the acoustic response are the shadow cast by the target on the seabed and the different physical phenomena contributing to the backscattered signal (specular and diffuse reflections, diffraction, multiple acoustic paths, etc.). The basic components are extracted by segmenting the sonar image into three classes labelled shadow, echo and background. The choice of shadow as basic component results in the study of two properties (1) the degree of interaction with the environment and (2) the stability with respect to changes in the target orientation.

Interaction with the environment. The environment interferes with the echo signal by adding backscattered energy from multiple acoustic paths, volume and side lobes. Waves reflected from the seabed onto the object corrupt the highlight pattern of the object. The environment interacts less with the shadow signal. The definition of the shadow boundaries is altered for objects lying on a seabed which produces shadows according to its geological characteristics (sand ripples, pebbles, rock plates, stone fields, etc.)

Stability with Changes in Target Orientation The basic component must be unaffected to small changes in viewing conditions. When the stability is low, the classifier must process a higher number of configurations to acquire a given target. The highlight information is the sum of energy from specular and diffuse reflection, diffraction and return from multiple path. For narrow beams, the spatial representation of highlights is sensitive to small changes in viewing conditions. The stability requirement can be met by broadening the beam to the detriment of spatial resolution. The shadow shape remains constant for small variations in the viewing conditions.

2.4 Automatic Classifier Design : Feature Computation

The conclusion from the preliminary design assessment is to extract features from the object shadow. The pixels belonging to the shadow are selected by segmenting the sonar image of the object. The height of the object h_i is estimated for the N sonar beams causing the object shadow.

$$h_i = \frac{A_i S_i}{R_i}, i = 1, \dots, N \quad (1)$$

NATO UNCLASSIFIED

SACLANTCEN SM-309

A_i , R_i and S_i are the sonar altitude, the range to the end of the shadow and the shadow length for beam i , respectively (see Fig. 2). The along-track coordinates of the beams⁴ are

$$y_i = (i - 1) \bar{v} P_r, \quad i = 1, \dots, N \quad (2)$$

where P_r is the ping repetition period and \bar{v} the average longitudinal speed of the sonar carrier. The height profile of the object, representing its estimated cross section, is defined by the matrix H

$$H = \begin{bmatrix} y_1 & h_1 \\ \vdots & \vdots \\ y_N & h_N \end{bmatrix} \quad (3)$$

Assuming that the sonar follows a linear motion at nominal speed, the object height profile (OHP) is transformed into a one-dimensional feature vector $P = [h_1, \dots, h_N]^T$. When the speed is different from nominal, interpolation is required to extract the feature vector P from the matrix H . Figure 2 illustrates how the height profile is derived from the sonar image of an object. The segmented sonar image is displayed in Fig. 2(b). Figure 2(a) shows how to estimate the height of the object for each beam and Fig. 2(c) illustrates the OHP. The set of height profiles, recorded when circling around an object, can be processed to reconstruct the three-dimensional shape [20, 21].

2.5 Automatic Classifier Design : Classification Algorithm

The choice of the classification algorithm depends on the way in which the target knowledge will be introduced. The main modes are supervised and unsupervised. Supervised classifiers memorizes a given set of target feature vectors with their corresponding target label. The simplest way to define the target labels is to give a different number to each target. Unsupervised algorithms process only the feature vectors and, consequently, produce their own target labels.

If the mine shapes are known *a priori*, a set of feature vectors labelled with the corresponding target class can be defined. The supervised mode is therefore well adapted to automatic mine classification. Unsupervised classification does however possess interesting properties when used in a supervised classification context namely, (1) the detection of ambiguity in training data (for supervised learning) and

⁴Equation 2 assumes that a single beam is formed for each transmitted ping

NATO UNCLASSIFIED

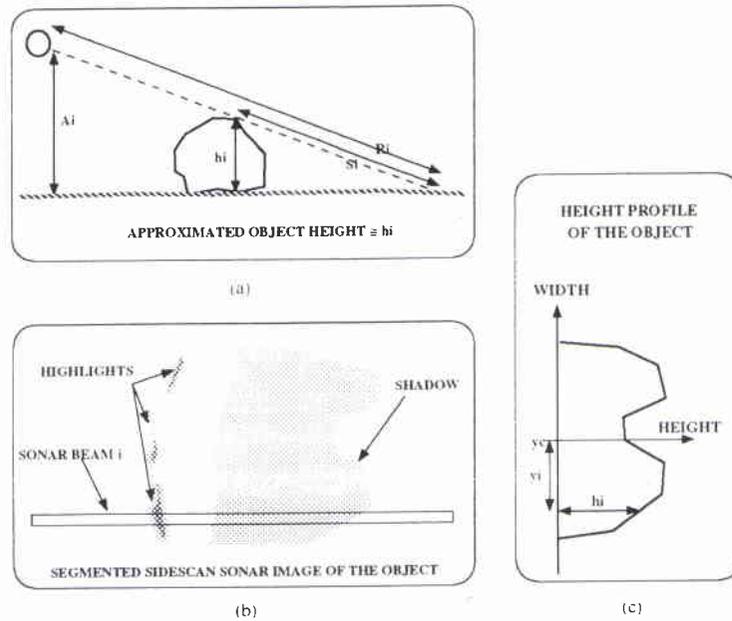


Figure 2: Object Height Profile Estimation

(2) the reduction of the false alarms rate for supervised classifiers based on linear discrimination.

The principle of automatic target classification is derived from Bayes decision theory [22]. The *a posteriori* probability $P(\omega_j | \mathbf{x})$ that the observed feature vector \mathbf{x} is from target ω_j is expressed as

$$P(\omega_j | \mathbf{x}) = \frac{p(\mathbf{x} | \omega_j)P(\omega_j)}{p(\mathbf{x})}$$

where (4)

$$p(\mathbf{x}) = \sum_{j=1}^{N_t} p(\mathbf{x} | \omega_j)P(\omega_j)$$

N_t is the number of targets, $P(\omega_j)$ is the *a priori* probability to encounter target ω_j and $p(\mathbf{x} | \omega_j)$ is called the state-conditional probability density function for \mathbf{x} and corresponds to the probability density function for input feature vector \mathbf{x} given the target is ω_j .

Deciding that observed feature vector \mathbf{x} is from target ω_i if the actual target is ω_j

NATO UNCLASSIFIED

SACLANTCEN SM-309

will incur the loss $\lambda(\omega_i | \omega_j)$. The expected loss, or conditional risk, associated with making the decision that the observed feature vector belongs to class ω_i is

$$R(\omega_i | \mathbf{x}) = \sum_{j=1}^{N_t} \lambda(\omega_i | \omega_j) P(\omega_j | \mathbf{x}) \quad (5)$$

The conditional risk is computed for $i = 1, \dots, N_t$ and the answer of the classifier is the class ω_i for which the conditional risk $R(\omega_i | \mathbf{x})$ is minimum. In the absence of information, a simple way to quantify the loss on a decision is to assign no loss in case of correct decision and a unit loss otherwise

$$\lambda(\omega_i | \omega_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases} \quad (6)$$

The computation of the conditional risk simplifies to

$$R(\omega_i | \mathbf{x}) = 1 - P(\omega_i | \mathbf{x}) \quad (7)$$

Equation (7) shows that minimizing the conditional risk corresponds to maximizing the *a posteriori* probability. If there is no *a priori* information on the target type, all $P(\omega_j)$'s are set to one. Then, maximizing the *a posteriori* probability is equivalent to maximizing the state-conditional probability density function

$$\text{Decide } \omega_i \text{ if } p(\mathbf{x} | \omega_i) > p(\mathbf{x} | \omega_j) \text{ for all } i \neq j. \quad (8)$$

In practice, the classifier has to estimate the state-conditional probability density function for all classes of target. A convenient way of doing this task is to assume that these functions have parametric shapes in the feature space. A typical parametric function is the multivariate normal density. The parameters, mean vector and covariance matrix, are estimated for each target class from reference feature vectors.

The multivariate normal density is unimodal (i.e. it possess a single maximum in the feature space). Therefore, normal densities are well adapted to targets with feature vectors invariant with viewing conditions. But when the feature vectors change according to target orientation, a unique normal density function is unable to correctly represent the state-conditional probability density function of the target class.

NATO UNCLASSIFIED

Targets with feature vectors varying with orientation require more flexible shapes for their state-conditional probability density function. This approach is implemented by using neural network techniques [23]. Three classes are defined in a two-dimensional feature plane. Figure 3 shows the feature vectors defining these three classes. Class #1 (green) is divided into two parts in the feature plane. Class #2 (cyan) and #3 (red) have unimodal shapes.

The most widely employed neural network for supervised classification is the multiple layer perceptron (MLP). Figure 4 shows how the MLP estimates the state-conditional probability density functions for the three class example defined in Fig. 3. The MLP has one hidden layer. The linear discrimination, implemented in the MLP, allows well defined discrimination between classes but the state-conditional probability density functions remain high for regions of the feature space distant from the training data. When the objective of the classifier is to recognise only targets, the MLP produces an unacceptably high false alarms rate.

This problem can be solved by introducing a new class for the rejection of feature vectors which differ too greatly from those defining the target classes⁵. Another way to reduce the false alarm rate is to use a different mechanism for discrimination between classes. Instead of linear discrimination, RBF neural network uses radial basis functions [24, 25] to estimate the state-conditional probability density functions. Each basis function is a normal density function defined, in the feature space, by a mean vector and a scalar standard deviation. The mean vectors are estimated by k-means clustering and the standard deviations are drawn from the Euclidian distance between mean vectors. The state-conditional probability functions obtained with RBF (Fig. 5) allow discrimination between classes and they have a high value only for the regions of the feature space spanned by the classes.

Estimating the state-conditional probability functions can be avoided by using nearest neighbour based classifiers. One famous algorithm implements the *k-nearest-neighbour* (KNN) method [26]. Figure 6 shows the classification results obtained by a five neighbour KNN algorithm. KNN is very sensitive to training data. A single feature vector from the training set can have a significant effect on the results (see the red artifact on the lower right part of Fig. 6).

The high false alarm rate produced by MLP can be reduced by using, in parallel, a Kohonen self organizing feature map [27]. The Kohonen network (KOH) maps the training vectors on a two-dimensional grid of neurons using topological constraints. The mapping is achieved without taking into account the class labels (i.e. unsupervised learning). The combination mechanism, presented by Maillard *et al* [28], has been simplified in the sense that the KOH network does not consider class labels. The fact that a feature vector activates a neuron of the grid, irrespective of location,

⁵The training data set can be extended with random feature vectors, labelled with the rejection class label, and respecting a minimum distance to the nearest feature vector of the target classes

NATO UNCLASSIFIED

SACLANTCEN SM-309

is sufficient to indicate that it belongs to the acquired knowledge. In this working mode, the KOH network estimates the global probability density function of the target set. This probability density function is designated global because it combines in a single function all the state-conditional probability functions. The combination of MLP and KOH, labelled MNK, is achieved by multiplying the MLP output with the global probability density function estimated by KOH. Figure 7 shows the results of MNK combination for the 3 classes example. Compared to MLP, MNK classifier conserves the main advantage (sharp class separation) and attenuates the main shortcoming (state-conditional probability density function decreases for regions distant from the training data).

Using unsupervised techniques with labelled feature vectors, normally suited for supervised classification, allows detection of ambiguities in the training data⁶. The unsupervised network ART2, based on Adaptive Resonance Theory [29], has been chosen for this task. The key characteristic of a ART2 neural network is the incremental creation and modification of prototypes according to a vigilance threshold. At the beginning the network has no information. When the feature vectors sent to ART2 do not match its current knowledge, new prototypes are created. When the feature vectors are close to the ART2 knowledge, the nearest prototypes are updated. Figure 8 shows how ART2 prototypes cover the two-dimensional feature plan for the problem defined in Fig 3. The cross markers indicate the location of the prototypes in the feature plane. The parameter, called "vigilance threshold" in ART2 terminology, determines the size of the region of influence for each prototype. The ambiguities are detected by creating a links table between supervised and unsupervised labels. Unsupervised labels are obtained by giving a different identifier to each ART prototype. If the same unsupervised label corresponds to several supervised labels, it means that different targets share similar feature vectors. Consequently, the corresponding regions of the feature space support ambiguity and no discrimination can be achieved in these regions. No ambiguity has been detected in the results presented in Fig. 8. The prototypes P1,P4,P5 and P7 are linked to class #1, the prototypes P3,P6 and P8 to class #3 and prototype P2 is sufficient to describe class #2.

From the above discussion and the results obtained on the three class example, the neural networks retained for the estimation of the classification performance are the RBF and MNK (the combination of MLP and KOH). ART2 will be used to detect ambiguous training data.

⁶It should be noticed that in the presence of ambiguity, the supervised training algorithms will have more difficulties to converge. Despite this, neural networks which have acquired ambiguous knowledge may be able to correctly classify the feature vectors. Minor changes in the feature vectors will however lead to unpredictable behaviour.

NATO UNCLASSIFIED

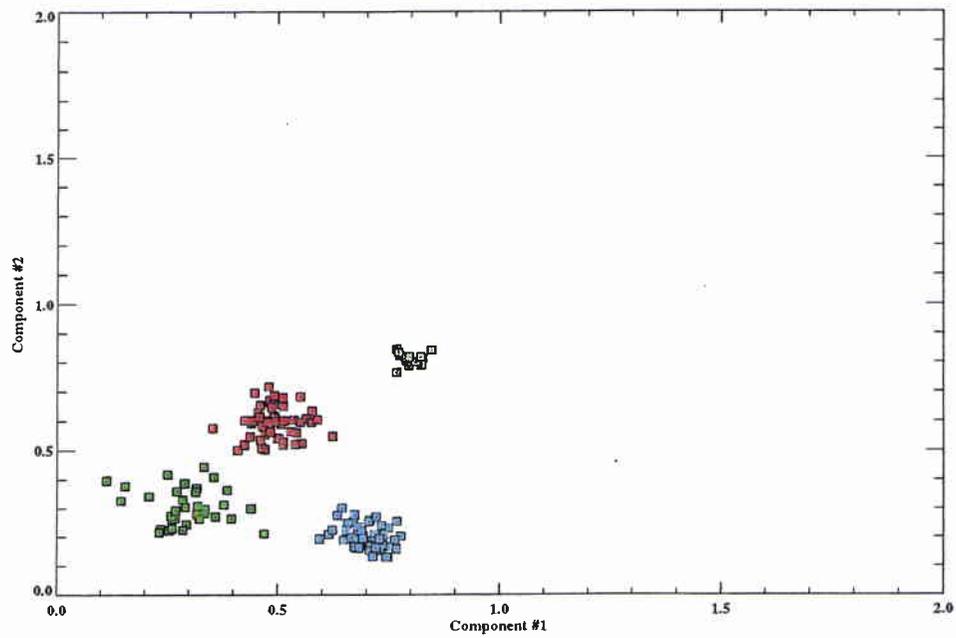


Figure 3: Training Data for Classification Tests : 3 classes of events defined by 2D features vectors

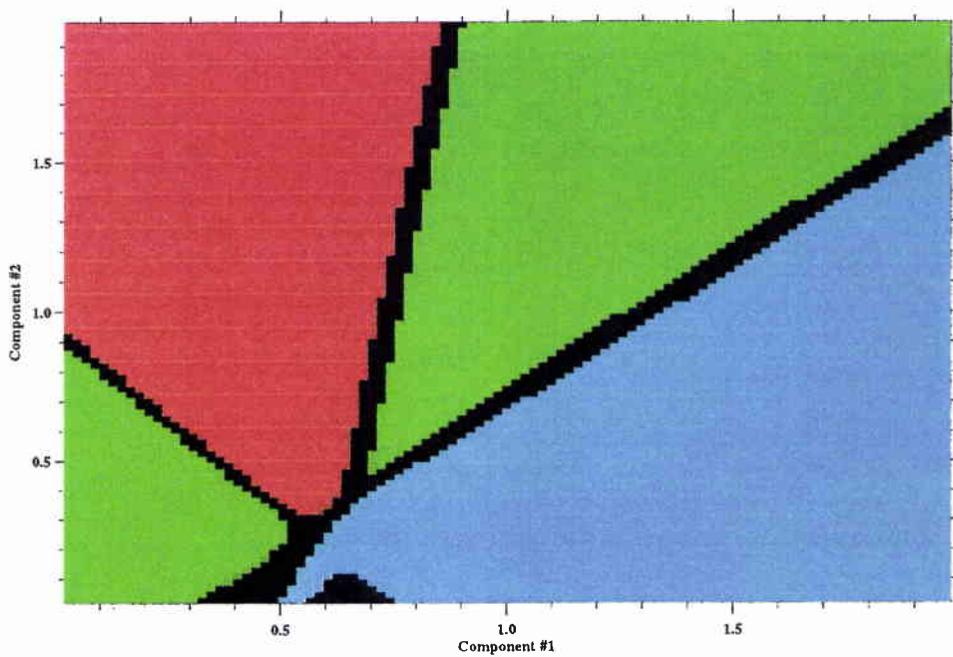


Figure 4: Estimation of the class-conditional probability density function with a multi layer perceptron (MLP)

NATO UNCLASSIFIED

SACLANTCEN SM-309

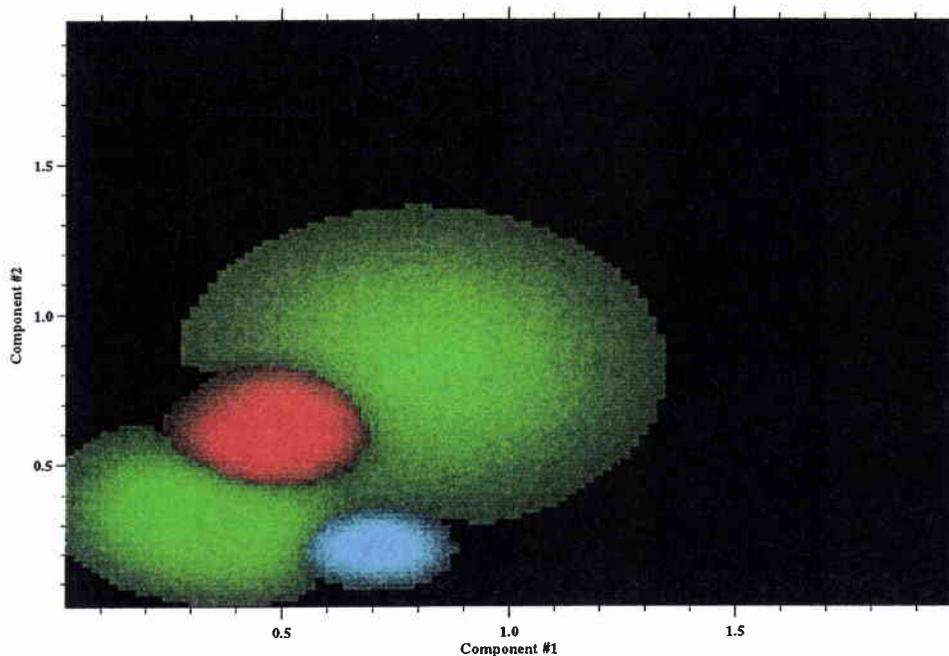


Figure 5: Estimation of the class-conditional probability density function with radial basis functions (RBF)

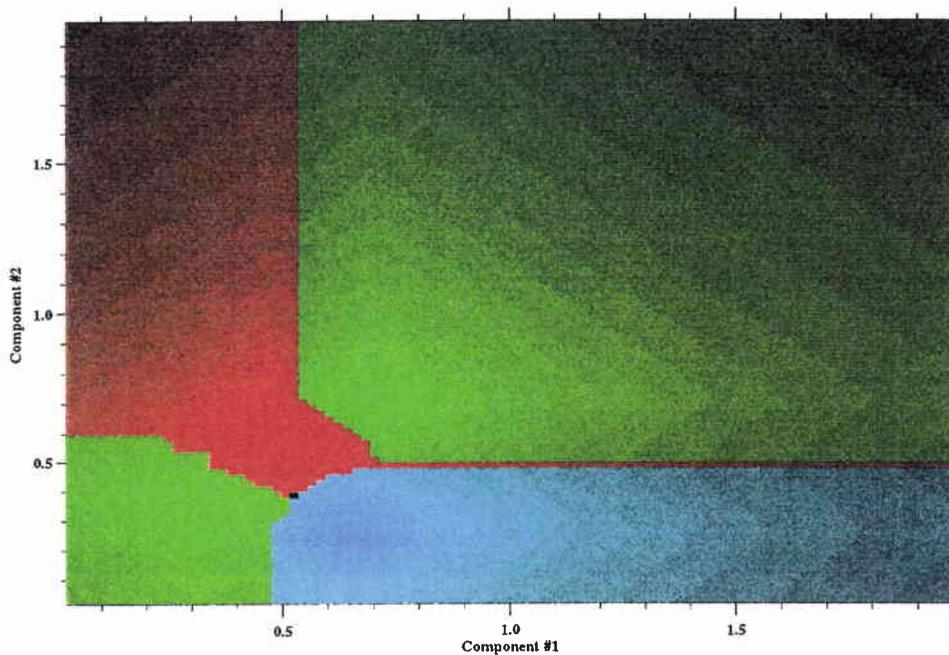


Figure 6: k-nearest-neighbors (KNN) classification results

NATO UNCLASSIFIED

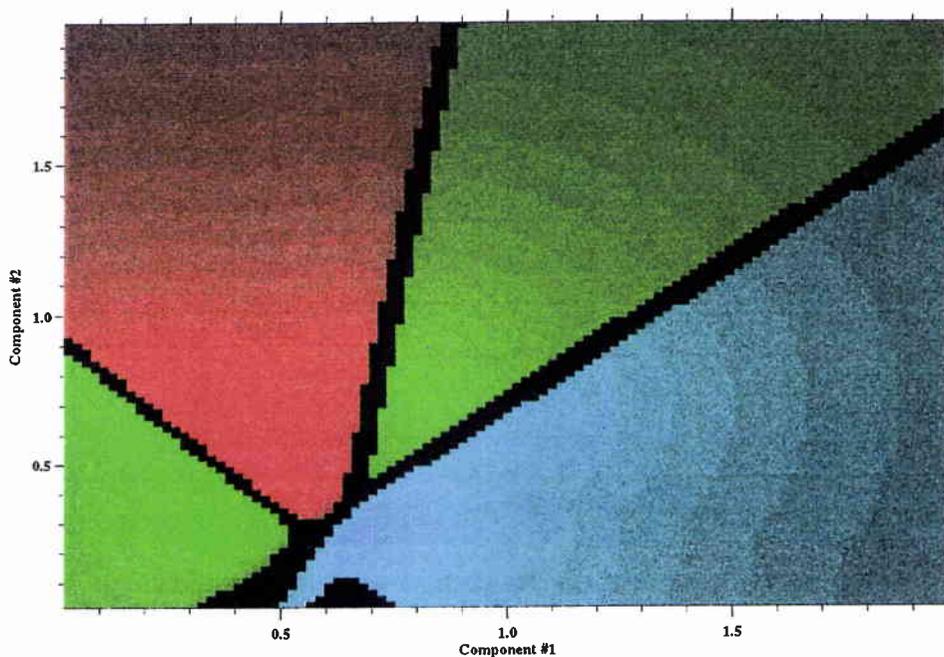


Figure 7: Estimation of the class-conditional probability density function with a combination (MNK) of multi layer perceptron and Kohonen self organizing feature map

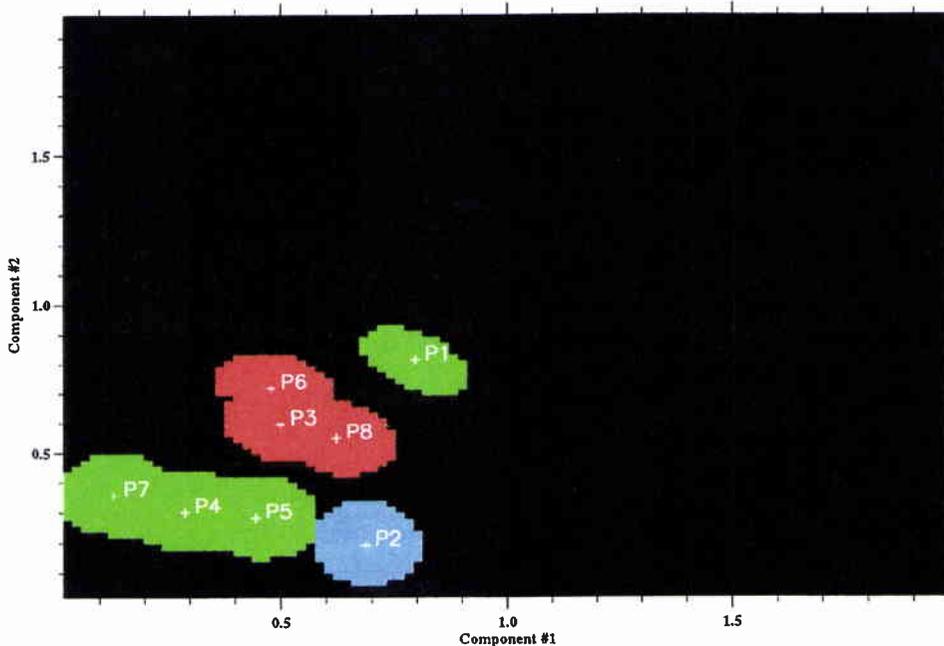


Figure 8: Estimation of the class-conditional probability density function with adaptive resonance theory (ART2)

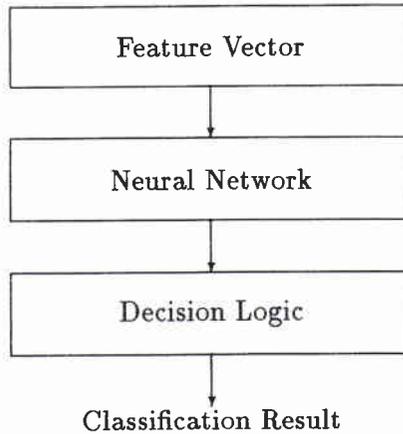


Figure 9: Diagram of the Standard Automatic Classifier

2.6 Decision Logic for Classification Determination

A diagram of the automatic classifier is shown in Fig. 9. The classification of an object results from the analysis of the neural network output which is achieved by the module called Decision Logic. The simplest approach to decision logic is to consider only the output cell i_{cl} of highest activity :

$$O(i_{cl}) \geq O(i) \quad , \quad i = 1, \dots, N_t \quad (9)$$

where $O(i)$ is the activity of output cell number i and N_t the total number of output cells. If each output cell corresponds to a different class of object and assuming that the neural network estimates the class-conditional probability density functions, the decision logic determines that the object belongs to class i_{cl} with an *a posteriori* probability of $O(i_{cl})$. The decision logic can be improved by taking into account the ambiguity of the classification. This is done by introducing the difference of activity ΔO between the two most active cells

$$\Delta O = O(i_{cl}) - O(j_{cl}) \quad (10)$$

The activity of the second most active cell, $O(j_{cl})$, is given by

$$O(j_{cl}) \geq O(i) \quad , \quad i = 1, \dots, N_t \quad , \quad i \neq i_{cl} \quad (11)$$

The rule implemented by the decision logic module is defined by two parameters : A_{MIN} and A_{GAP} . An object is considered as a valid member of class i_{cl} if the neural network output fulfills

$$\begin{cases} O(i_{cl}) \geq A_{MIN} \\ \Delta O \geq A_{GAP} \end{cases} \quad (12)$$

A_{MIN} is the minimum activity required on the most active output cell and A_{GAP} is the minimum activity difference required between the two most active output cells.

2.7 Quantification of Classification Performance

Classification performance is quantified by the percentage of targets correctly classified (or identified) and by the percentage of false alarms. The results are expressed by ROC curves⁷ and confusion matrices. The ROC curves characterize the “mine-like/non-mine-like” classification. The confusion matrices express the performance in target identification.

Establishing ROC Curves

Classification statistics are estimated by processing a large object set, containing M_t target and M_{nt} non target shapes. On the target subset, the percentage of correct classification $ML\%$ is defined as

$$ML\% = 100 \frac{M_{ml}}{M_t} \quad (13)$$

where M_{ml} is the number of targets classified as targets (i.e. as “mine-like”) and M_t is the total number of targets. $ML\%$ expresses the percentage of targets correctly classified, but does not take into account the mismatch between different target types. A second measurement $CL\%$ is introduced for this purpose

$$CL\% = 100 \frac{M_{cl}}{M_t} \quad (14)$$

where M_{cl} is the number of targets classified with the correct target type. In the subset of non target shapes, the percentage of false alarms $FA\%$ describes how non targets are classified as targets

⁷ROC stands for Receiving Operator Characteristics

NATO UNCLASSIFIED

SACLANTCEN SM-309

$$FA\% = 100 \frac{M_{ntcl}}{M_{nt}} \quad (15)$$

where M_{ntcl} is the number of non targets classified as targets. The percentages $CL\%$ and $FA\%$ can be generalized in probabilities P_{cl} and P_{fa} by increasing the number of tests

$$P_{cl} = \lim_{M_t \rightarrow \infty} \frac{CL\%}{100}$$

$$P_{fa} = \lim_{M_{nt} \rightarrow \infty} \frac{FA\%}{100} \quad (16)$$

The statistical measurements (P_{cl}, P_{fa}) define the performance of the classifier. As the decision module is controlled by A_{MIN} and A_{GAP} parameters (Eq. 12), processing the same set of feature vectors with different control parameters will give different (P_{cl}, P_{fa}) scores. These results are merged using ROC curves. The ROC curves display the percentage of correct classifications as a function of the percentage of false alarms.

Confusion Matrices

The confusion matrices synthesize how identification mismatch is spread over the target classes. Each column of the matrix corresponds to a type of target known by the classifier. Each line indicates the classification produced by the algorithm. Element C_{ij} confusion matrix, located in line i and column j , indicates the percentage of object of type j classified as object of type i . An ideal classifier will produce a confusion matrix having non zero values on the main diagonal and zeros everywhere else.

Multiple View Sidescan Sonar

3.1 Using Multiple Aspects for Target Classification

The correct (spatial) classification of an object requires that the visible features of its shape are accurately imaged. This task becomes more difficult in case of asymmetric objects, because their visible features change with viewing conditions. Target classification will give correct results if it exploits the anisotropy of target response and target-seabed interaction.

Assuming that the resolution is higher along range than across range, the worst aspect of an asymmetric object is the aspect at which it intercepts the lowest number of sonar beams. One way to handle the worst aspect of an object is to increase the across range resolution of the sonar so that a higher number of beams hits the object. Alternatively several views of the object are recorded to decrease the probability of viewing the worst aspect an object.

The fact that the target characteristics change with orientation has been investigated for echo-based mine detection [30]. Assuming that, on complex seabeds (with low signal-to-noise ratio), a single track is insufficient to correctly detect mines and considering that the target strength of the mines varies with orientation, the direction of additional tracks is determined in a way which maximizes mine detection probability. Determination is based on the fact that for particular aspects, a higher target strength will increase the signal-to-noise ratio and consequently the detection probability.

The transition from echo-based detection to shadow-based classification can be effected by substituting the scalar target strength by a feature vector, computed on the target image. When using a single view, different objects sharing the same feature vector will confuse the classification algorithm. This is illustrated by using the cross-section of an object as feature vector. The cross-sections are estimated from the target shadow by computing the object height profile. As shown in Fig. 10(a), a rectangular cross-section may come from at least three different shadows (Figs. 10(b), 10(c)) and 10(d)). These shadows are respectively produced by a vertical cylinder Fig. 10(e), a horizontal cylinder Fig. 10(f) and a polyhedron Fig. 10(g). Figures 10(h), 10(i) and 10(j) demonstrate that a second view, orthogonal to the first, allows the objects to be distinguished from each other. This example shows that the worst as-

NATO UNCLASSIFIED

SACLANTCEN SM-309

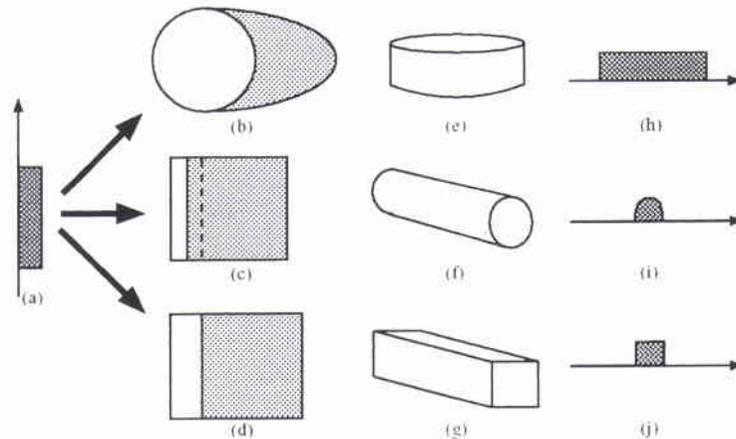


Figure 10: Multiple View Shadow Based Classification. Rectangular cross section (a) estimated from the shadows (b) (c) (d) of three different objects : vertical cylinder (e), horizontal cylinder (f) and polyhedron (g). Cross sections (g) (h) (i) estimated from a second view, orthogonal to the first, allow to discriminated between these three objects.

pect of an object is not only the aspect under which it intercepts the lowest number of beams. It may also be the aspect which introduces confusion in the classification task.

To summarize, the introduction of multiple views in target classification is intended to reduce the probability of processing an object only imaged at its worst aspect.

3.2 Acquisition of Multiple Views

The sidescan sonar (Fig. 11) produces an image the geometry of which is particularly well adapted to rapid acquisition of seabed characteristics. However, a single view of the sea bed affects the reliability of the interpretation. This arises, for example, when the aspect of an investigated location changes with the azimuth or when a large seabed structure hides smaller ones. This means that a single sidescan sonar view of the seabed is insufficient for seabeds with anisotropic characteristics.

The sectorscan sonar, (Fig. 12(a)) which can take images of the same area, with different azimuth and grazing angles, is more suited to acquire anisotropic seabed response.

NATO UNCLASSIFIED

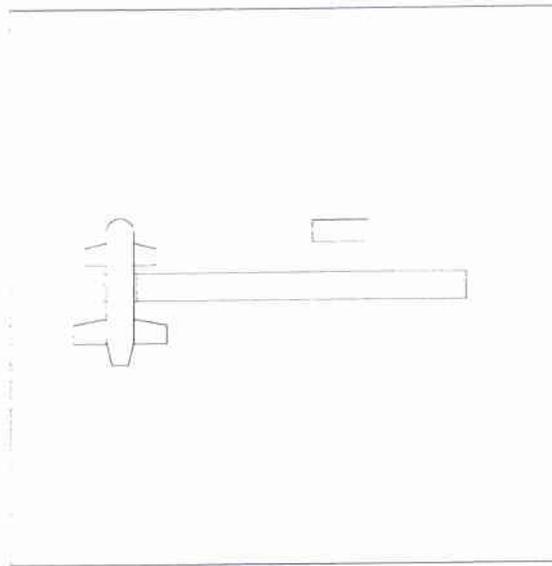


Figure 11: Conventional (Single View) Sidescan Sonar

During the acquisition of the image sequence, precise tracking of the target location is required. When using a hull mounted sonar or a variable depth sonar (VDS), tracking capabilities are limited by operational constraints such as ship maneuverability, positioning accuracy, safety range, obstacle avoidance, etc. Propelled variable depth sonar (PVDS) or AUV mounted sonar allows more flexible tracking. To summarize, the sectorscan sonar is an appropriate device to acquire multiple views of a particular seabed location. However, this system is not suited to recording systematically and precisely, multiple aspects of large areas.

Between sidescan and sectorscan sonars, an intermediate approach consists of using a limited number of views. This approach can be implemented by covering the seabed with several sidescan tracks of different headings (see Fig. 12(b)). However, to be effective, this method requires either an excellent knowledge of sonar navigation or an accurate image registration algorithm. A limited number of views can also be gathered by a multiple view sidescan sonar (Fig. 12(c)). The complexity of image registration is reduced by simultaneously recording several beams pointed in different directions. Though the useful range of the multiple view sidescan sonar is shorter, this system still possesses better coverage rate than multiple tracks by conventional sidescan sonar. The explanation resides in the fact that simultaneous multiple views reduce maneuvering time. Simultaneous coverage of the seabed by a sidescan and a sectorscan sonar is another method to record multiple seabed views (Fig. 12(d)).

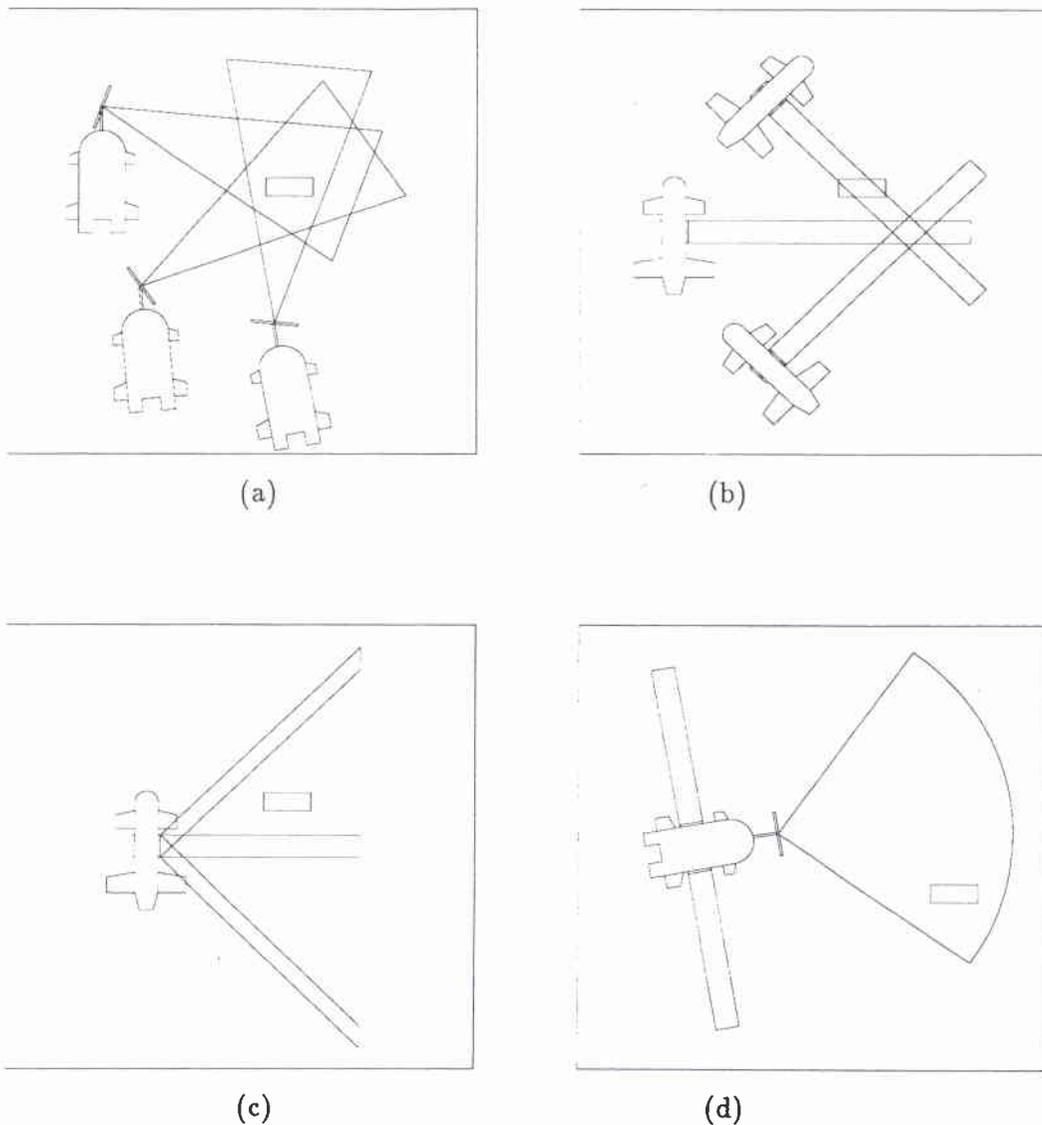


Figure 12: Sonar Systems for the Acquisition of Multiple Aspects

There are two different approaches to deal with multiple views in MCM missions depending on the number of views.

Approach I High Number of Views. The sectorscan sonar, mounted on the MCM vessel or better on a ROV, collects a high number of views. The systematic processing of these views⁸ is intended to improve, in mine hunting missions, the classification and identification of the detected echoes. Within the scope of project 031-3, the systematic processing of these multiple views has been investigated [20, 21]. The unique and short experiment conducted in collaboration with GESMA [31] has given encouraging results. However, the amount of data collected is insufficient to allow firm conclusions to be drawn.

Approach II Limited Number of Views. The multiple view sidescan sonar gathers information for a limited number of directions. For fast route survey missions, this system can be operated like a conventional sidescan sonar. The determination of the viewing directions and the estimation of the classification performance are addressed in this report.

⁸It should be noticed that the multiple view based classification is already (intuitively) performed by the sonar operators.

NATO UNCLASSIFIED

SACLANTCEN SM-309

Individual Classification		Logical Combination	
Classification View #1	Classification View #2	Sum (LO)	Product (LA)
NON-TARGET	NON-TARGET	NON-TARGET	NON-TARGET
TARGET	NON-TARGET	TARGET	NON-TARGET
NON-TARGET	TARGET	TARGET	NON-TARGET
TARGET	TARGET	TARGET	TARGET

Table 1: Definition of the Logical Combination Modes LA and LO

3.3 Classification Schemes for Multiple Views

From One to Two Views

The standard automatic classifier uses a single view of the object. As shown in Fig. 13, the feature vector is sent to the neural network input and the output classification is obtained by applying decision logic to the neural network output. Analysis of this diagram shows three possible modes to introduce a second view in the classifier :

- *Input Concatenation (IC)* : The feature vector extracted from the second view is sent to the neural network input. The size of the input layer is doubled to receive the new feature vector. This vector results in the concatenation of the feature vectors issued by the two views (Fig. 14).
- *Output Averaging (OA)* : The feature vector extracted from the second view is sent to a second neural network, which is a clone of the neural network used for the first view. The output activities of the two networks are averaged. The average output is then processed by decision logic as for the standard single view classifier (Fig. 15).
- *Logical Combination (LA,LO)* : The feature vectors issued by the two views are classified separately by two standard single view based classifiers. The outputs of their decision logic are combined to determine the classification (Fig. 16). Two modes of logical combination have been investigated, the product (LA) and the sum (LO). Table 1 shows the final classification.

For logical combination (LA,LO) and output averaging (OA) modes, the parallel classification of the two views with two neural networks can be replaced by the sequential use of the same neural network.

NATO UNCLASSIFIED

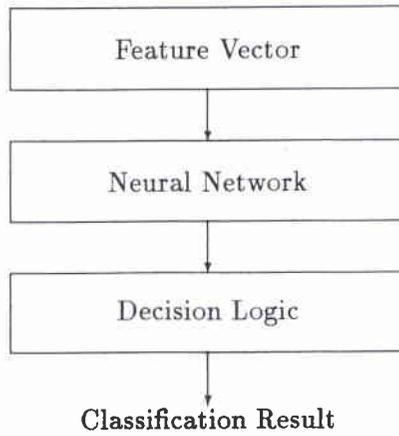


Figure 13: Standard Classifier

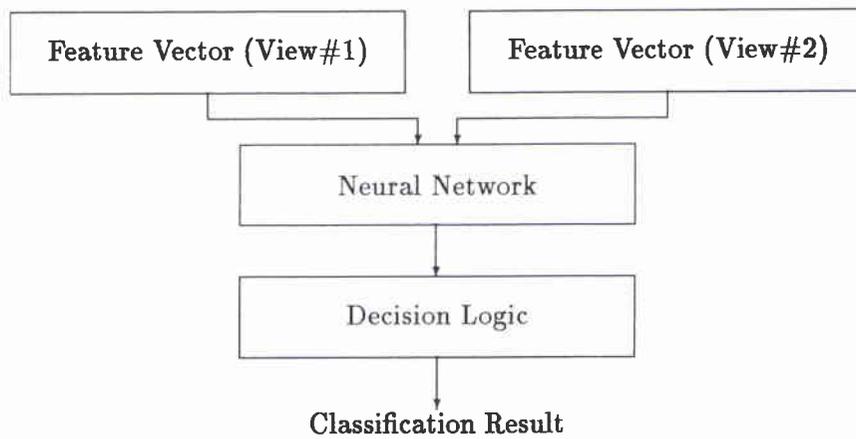


Figure 14: Input Concatenation

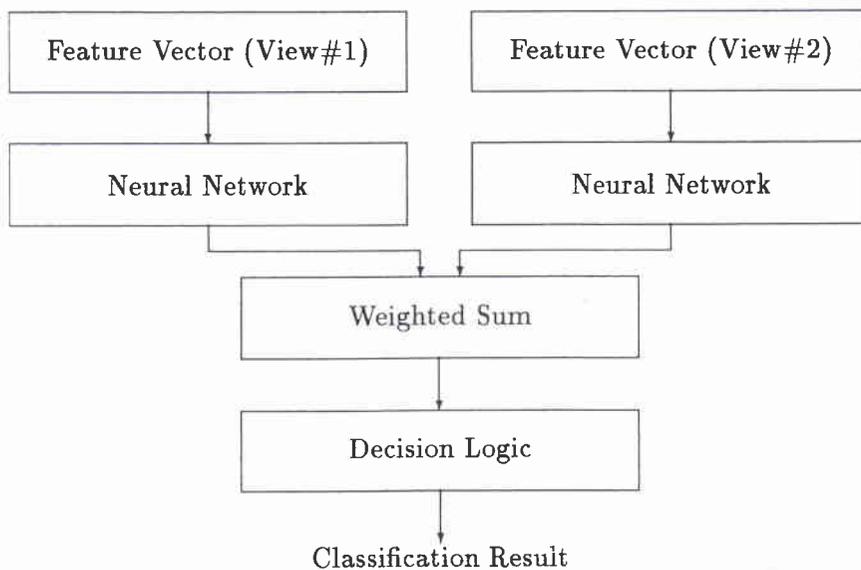


Figure 15: Averaging Neural Network Responses

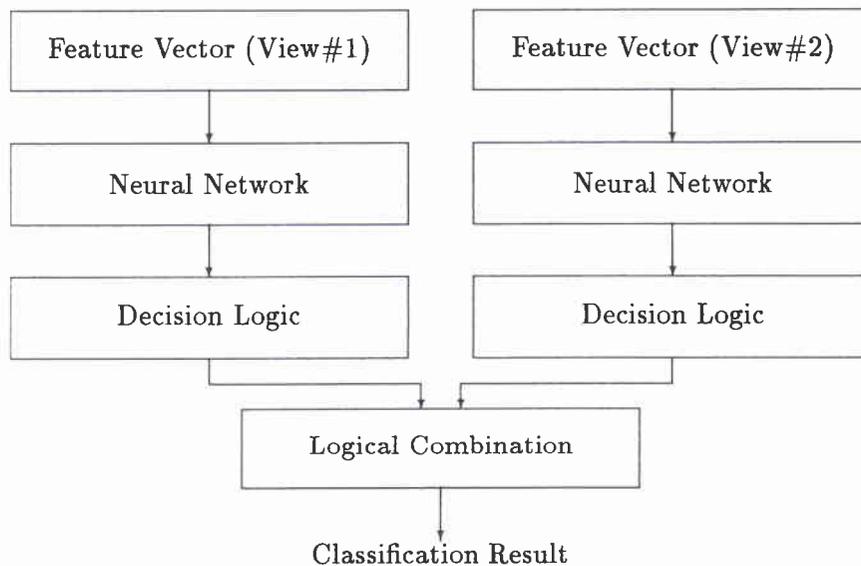


Figure 16: Logical Combination of Individual Classifications

Using More Views

To accept more views, the input concatenation (IC) mode requires an increase in the size of the neural network input. The internal structure of the neural networks is also modified to handle the extended input layer.

The output averaging (OA) mode can process more views without significant architectural modification. The only change is the number of neural network outputs involved in the computation of the average activity (i.e. one output activity per view).

When using more than two views, several ways to logically combine the single classification results may be envisaged. The logical product is extended to three and more views by deciding that an object is a target if all the single classifications state that it is a target. The logical sum is extended by deciding that an object is a target if at least one of the single classifications states that it is a target. An intermediate approach consists in retaining the classification expressed by the majority of the single view classifiers.

4

Classification Performance

The comparison of classification performance is based on a technique widely used in computer science to rank machine performance: the benchmark. Here, the benchmark consists of comparing the scores of automatic classifiers on a predefined set of object for different multiple aspect configurations. The objects have shapes corresponding to proud targets and to sinker of moored targets. The block diagram of the performance estimation software is represented in Fig. 17. The simulation of the feature vectors is controlled by the scenario parameters (object type, across range resolution, along range resolution, distance and relative orientation of the object with sonar track, sonar altitude, etc.). The simulation of the sonar images is briefly described in Annex A.

Two neural network architectures are compared : the radial basis function network (RBF) and the combination of multiple layer perceptron (MLP) and Kohonen self organized feature map (KOH). The combination is labeled MNK. For both MNK and RBF classifiers, several ways to handle multiple views are compared. Table 3 describes these configurations. IC and OA stand for input concatenation and output averaging. LA, LO and LM correspond to logical product, logical sum and majority vote, respectively. In the following discussion, each classifier will be referred to by its mnemonic.

For each type of neural network algorithm (i.e. RBF and MNK) the multiple view

Identifier	Type	Shape	Dimensions (m)
SPHR	Proud	Sphere	$\phi = 1$
VCYL	Sinker	Cylinder	$\phi = 0.5, h = 0.5$
MNLK	Proud	Truncated Cone	$\phi_{\text{Base}} = 1.2, \phi_{\text{Top}} = 0.6, h = 0.5$
H CYB	Proud	Cylinder (Flat Endcaps)	$\phi = 0.5, l = 1.5$
H CYC	Proud	Cylinder (Flat Endcaps)	$\phi = 0.5, l = 2.0$
H CSB	Proud	Cylinder (Spherical Endcaps)	$\phi = 0.5, l = 1.5$
H PPD	Sinker	Polyhedron	$l = 0.8, w = 0.5, h = 0.5$
ROCL	Proud	Polygonal	$l = 1.0, w = 0.8, h = 0.4$
HEMS	Proud	Hemispherical	$\phi = 1.0$
P YRS	Sinker	Polygonal	$l = 1.0, w = 1.0, h = 0.5$

l, w, h and ϕ stand for the object length, width, height and diameter, respectively.

Table 2: Shapes definition for 10 target benchmark

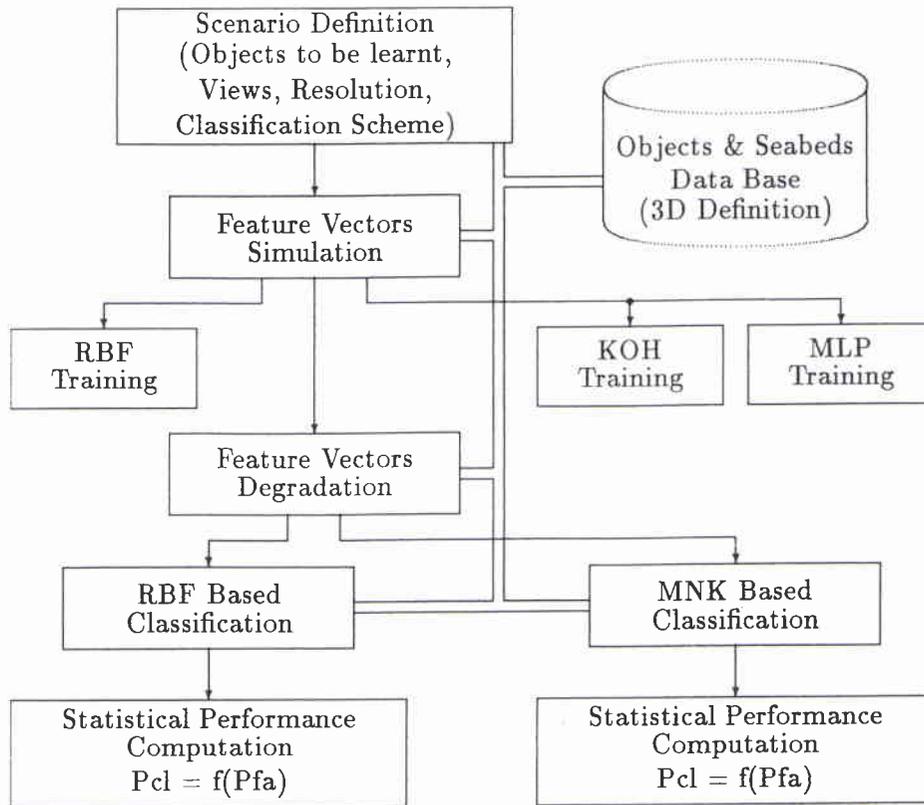


Figure 17: Diagram of the Performance Evaluation Software

Mnemonic	view#1	view#2	view#3	Combin. Mode	Neural network Id.	Net-work
S1	0°	-	-	-	S00	
S2	5°	-	-	-	S05	
S3	45°	-	-	-	S45	
S4	90°	-	-	-	S90	
D12IC	0°	5°	-	IC	D05	
D13IC	0°	45°	-	IC	D45	
D14IC	0°	90°	-	IC	D90	
D12OA	0°	5°	-	OA	S00 & S05	
D13OA	0°	45°	-	OA	S00 & S45	
D14OA	0°	90°	-	OA	S00 & S90	
D12LA	0°	5°	-	LA	S00 & S05	
D13LA	0°	45°	-	LA	S00 & S45	
D14LA	0°	90°	-	LA	S00 & S90	
D12LO	0°	5°	-	LO	S00 & S05	
D13LO	0°	45°	-	LO	S00 & S45	
D14LO	0°	90°	-	LO	S00 & S90	
T134IC	0°	45°	90°	IC	T45	
T134OA	0°	45°	90°	OA	S00, S05 & S90	
T134LM	0°	45°	90°	LM	S00, S05 & S90	

Table 3: Mnemonics for Multiple View Classification Schemes

NATO UNCLASSIFIED

SACLANTCEN SM-309

Neural Network Identifier	Input Feature Vector	Number of Inputs	RBF (Number of Basis Functions)	MNK (Number of MLP Hidden Neurons)	MNK (Size of KOH Neuron Grid)
S00	0° view	21	40	31	11x11
S05	5° view	21	40	31	11x11
S45	45° view	21	40	31	11x11
S90	90° view	21	40	31	11x11
D05	0° and 5° views	42	80	52	17x17
D45	0° and 45° views	42	80	52	17x17
D90	0° and 90° views	42	80	52	17x17
T45	0° , 45° and 90° views	63	120	63	21x21

Table 4: Neural Networks Characteristics

classifiers are implemented according to Table 3 definitions. The identifiers of these classifiers, their input feature vectors and their topology are defined in Table 4. These neural networks differ by the size of their input layer and by the target orientation from which the feature vectors have been extracted. D05, D45 and D90 process feature vectors from two views with an angular interval of 5°, 45° and 90°. T45 processes feature vectors from three views with a 45° angular interval. S00, S05, S45 and S90 process feature vectors extracted from a single view. As the interval between views does not apply for a single view, a unique neural neural network is sufficient to handle any kind of single view based classification. However, these four networks are retained to analyse the accuracy and the stability of the classification statistics. MNK is composed of a single hidden layer MLP and a two dimensional Kohonen grid (KOH). Table 4 gives essential informations on classifiers topology : the number of MLP hidden neurons, the size of Kohonen grid and the number of basis functions. Since change of 20 % on these values have raised to similar results, they must be interpreted as average indicators.

4.1 Feature Vector Generation

For each object of the benchmark, the feature vectors for training the classifiers are simulated for orientations ranging from 0° to 355° with angular sampling of 5°. To establish the classification performance, a large set of objects with random orientation is defined. The parameters controlling the simulation are the noise level and the height resolution. The noise obeys to a zero mean normal density function. The height resolution takes into account the along range resolution and the resolution of the sonar altitude measurement. The objects have shapes corresponding to proud targets and to sinker of moored targets. A set of 10 shapes, defined in Table 2, has been retained. Figures A-2 and A-9 of Annex A show these object shapes rendered by optical ray tracing. The natural objects are stones generated by a

NATO UNCLASSIFIED

fractal technique (annex A). The natural objects are labelled NMML (Non Mine Mine Like). The across range resolution is fixed at 20 cm. The orientation of the first view is randomly defined. The orientations of the 3 subsequent views are at 5° , 45° and 90° intervals from the first one.

4.2 Training

The neural network learns pure feature vectors in the sense that they have neither been degraded by geometrical distortions nor corrupted by noise. Table 5 summarizes learning performance for the classifiers defined in Table 4. On the non degraded feature vectors, MNK and RBF attain a similar level of performance. The results obtained by single view based classifiers (S00, S05, S45 and S90) indicate that RBF has an acceptable stability (the performance deviation is less than 1%). The MNK, with a performance deviation greater than 2%, is less stable. To reach the 2% variation, each MNK network has repeated 10 times the training procedure and the weight configuration giving the best results has been conserved. Without these repeated trainings, the performance deviation is greater than 10%. Irrespective of the type of neural network, increasing the number of views improves performance by reducing data ambiguity. For nominal noise (2.5 cm standard deviation) and height resolution (5 cm), the degradation in performance is more sensitive for single view than for multiple views (see Table 5). In general, MNK is more affected by noise than RBF⁹. A set of 5000 test targets, with random orientation uniformly distributed¹⁰, produces results (columns 6-7 of Table 5) which are very close to the results on the training data with nominal noise (columns 4-5). This demonstrates that, in the shadow based classification context, orientation sampling of 5° is sufficient to correctly acquire the target shapes.

4.3 Preliminary Tests on Performance Stability

Two tests are made before addressing the performance of the multiple view classifiers. These tests are based on the single view classification schemes S1, S2, S3 and S4 using neural networks S00, S05, S45 and S90, respectively.

The first test concerns the stability of the training phases. The noise level and the height resolution are nominal. Each of the four single view classifiers receives

⁹To limit performance decay induced by noise, the solution could be to learn features vectors corrupted by noise. However, this approach gives consistent results only if the noise characteristics are perfectly known. This must be taken into account in practical implementation of automated classification, but here, in the framework of model based study, the degrading effects of noise are only countered by the neural network.

¹⁰the target orientation takes a real random value between 0° and 360°

Neural Network Iden- tifier	Classification Scores on Known Targets					
	Training Data Non Distorted		Training Data Nominal Noise		Testing Data (5000 Targets) Nominal Noise	
	RBF	MNK	RBF	MNK	RBF	MNK
S00	81.9%	80.6%	73.9%	67.5%	74.3%	67.7%
S05	81.9%	81.7%	73.9%	66.9%	73.7%	67.6%
S45	82.5%	79.4%	73.9%	68.1%	73.0%	67.4%
S90	81.9%	80.3%	73.9%	67.5%	72.6%	66.5%
D05	86.1%	78.9%	73.9%	73.7%	74.2%	72.5%
D45	91.4%	87.5%	86.5%	83.7%	86.5%	83.2%
D90	94.4%	92.2%	92.8%	87.9%	91.1%	85.9%
T45	95.8%	95.8%	93.1%	89.7%	92.3%	91.3%

Table 5: Learning Performance with Nominal Height Resolution (5 cm). (*Nominal Noise has a Standard Deviation of 2.5 cm*)

the same feature vectors extracted from the first view of the objects. For RBF based classifiers, the ROC curves in Fig 18 are almost identical, demonstrating that the learning process is stable. For MNK, the deviation seen when classifying the training patterns (Table 5) is also present in the ROC curves (Fig. 19). RBF and MNK classifiers produce different ROC curves (see Figs. 18 and 19). MNK classifies better for very low false alarm rates ($< 10\%$). For high correct classification rate ($> 70\%$), RBF produces less false alarms.

The second test addresses the statistical validity of the classification scores. The results of the four single view classifiers are compared according to their respective feature vectors. If θ is the orientation of the object, the first classifier (S1) receives the feature vector extracted from the view of orientation θ , the second (S2) receives the feature vector issued by the view of orientation $\theta + 5^\circ$ and so on. If the number of objects is high enough, the four classifiers tend to behave similarly and their results can be considered as statistically valid. This is illustrated by the ROC curves in Figs. 20 and 21.

4.4 Target Classification Performance

Multiple view sidescan sonar performance in classification are compared using RBF neural networks on 12000 object images. The performance comparison is achieved on three points : the number of views, the relative orientation of the views and the modes of combining views. The effect of increased noise and decreased resolution is evaluated.

The performance of conventional "mine-like/non-mine-like" classification is not evaluated because the automatic classifier has been designed to discriminate between different object shapes, not between "mine-like" and "non-mine-like" objects. A

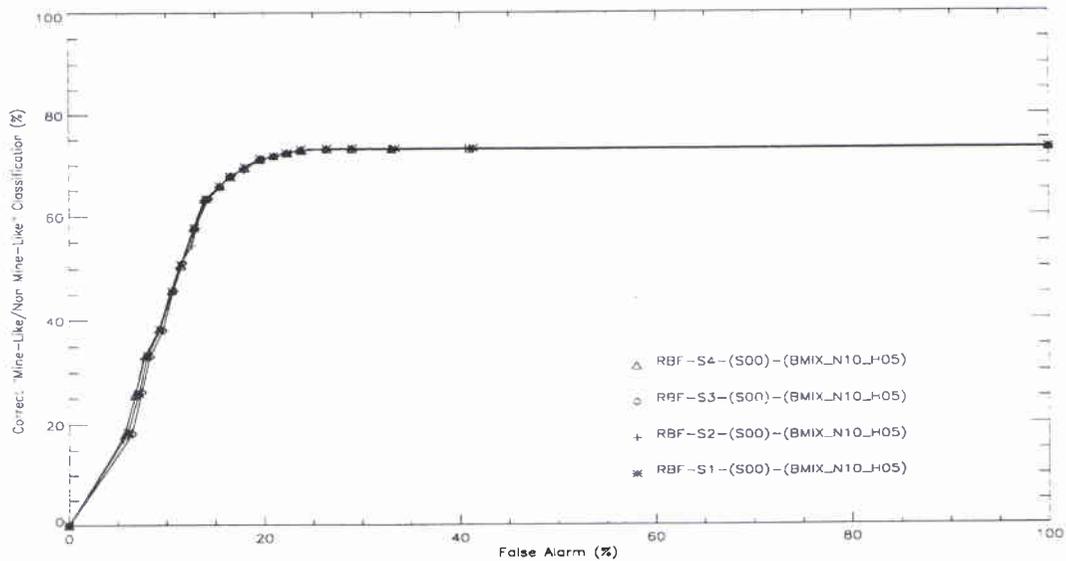


Figure 18: Performance stability for single view based classifiers. 12000 objects are classified using RBF neural networks. (*) marked curve corresponds to S1 classifier, (+) to S2 classifier, (\diamond) to S3 classifier and (Δ) to S4 classifier. Random object orientation is uniformly distributed between 0° and 360° . The object orientation is identical for the four classifiers.

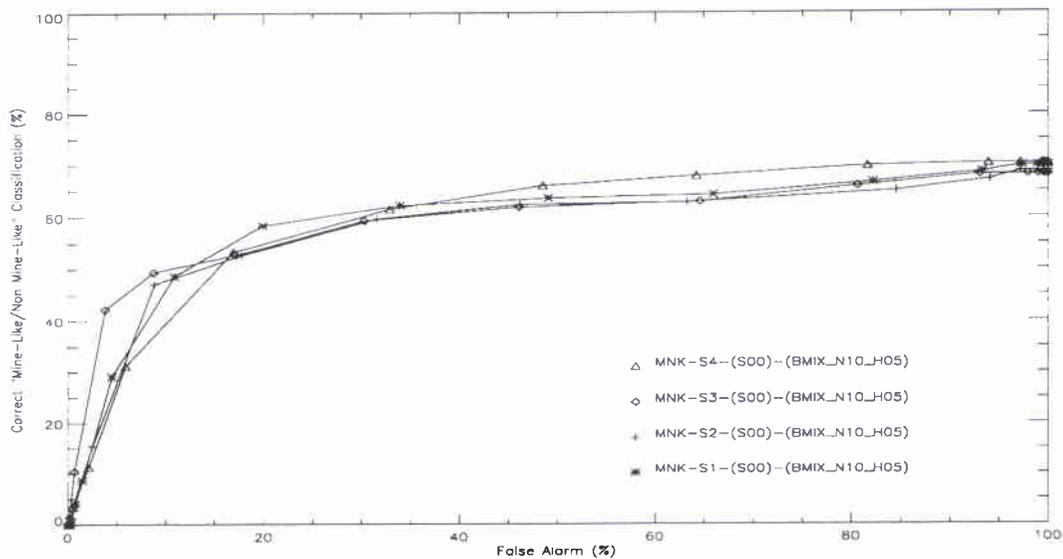


Figure 19: Performance stability for single view based classifiers. 12000 objects are classified using MNK neural networks. (*) marked curve corresponds to S1 classifier, (+) to S2 classifier, (\diamond) to S3 classifier and (Δ) to S4 classifier. Random object orientation is uniformly distributed between 0° and 360° . The object orientation is identical for the four classifiers.

NATO UNCLASSIFIED

SACLANTCEN SM-309

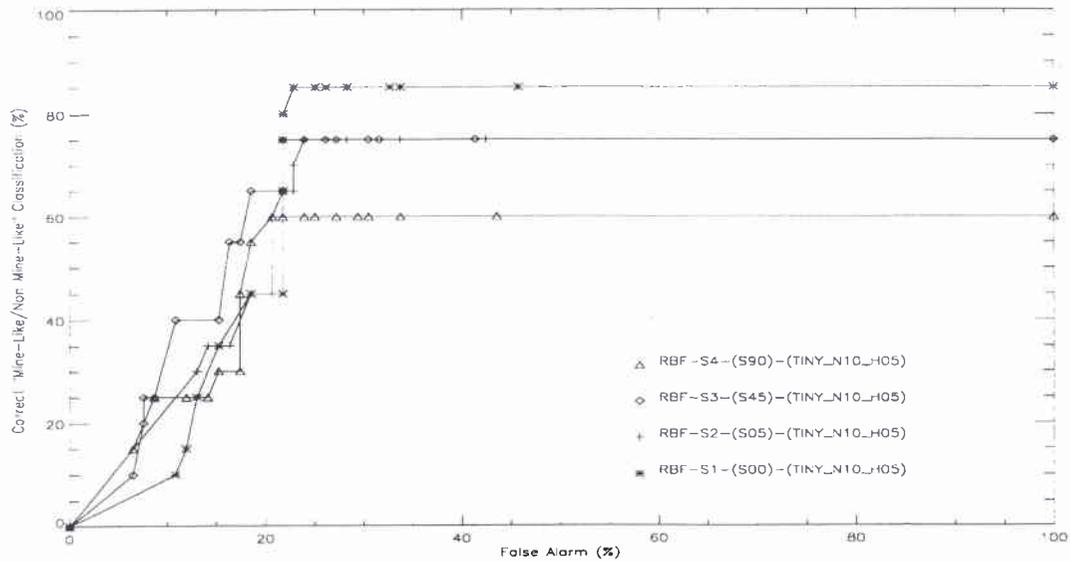


Figure 20: Insufficient number of processed objects (i.e. 112 objects) leads to non valid statistics (RBF neural networks). (*) marked curve corresponds to S1 classifier, (+) to S2 classifier, (◇) to S3 classifier and (△) to S4 classifier. Random object orientation, θ_i , is uniformly distributed between 0° and 360° . The object orientation is θ_i for S1 classifier, $\theta_i + 5^\circ$ for S2, $\theta_i + 45^\circ$ for S3, $\theta_i + 90^\circ$ for S4, ($i = 1, \dots, 112$).

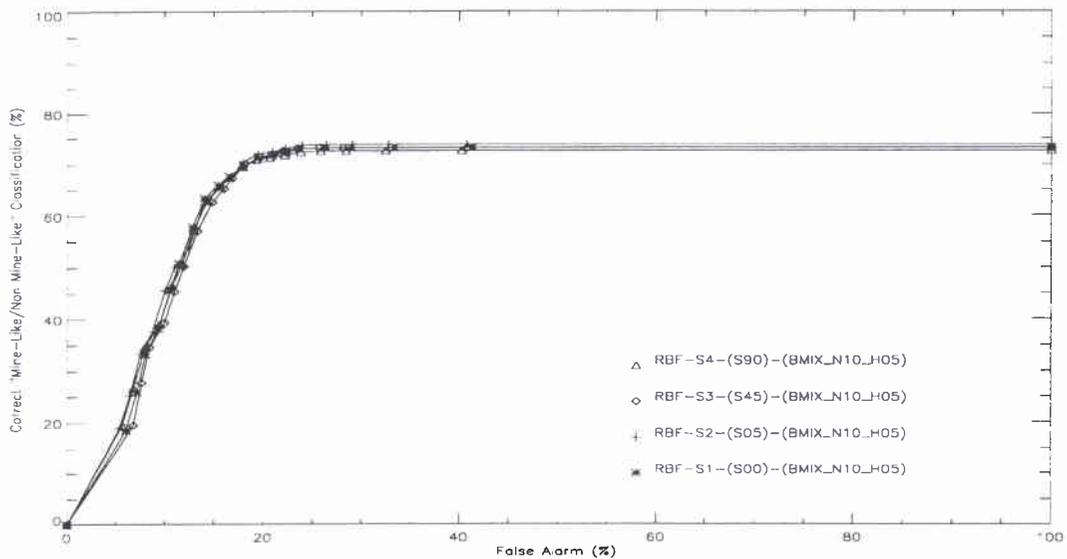


Figure 21: Valid statistics require a high number (i.e. 12000) of processed events (RBF neural networks). (*) marked curve corresponds to S1 classifier, (+) to S2 classifier, (◇) to S3 classifier and (△) to S4 classifier. Random object orientation, θ_i , is uniformly distributed between 0° and 360° . The object orientation is θ_i for S1 classifier, $\theta_i + 5^\circ$ for S2, $\theta_i + 45^\circ$ for S3, $\theta_i + 90^\circ$ for S4, ($i = 1, \dots, 12000$).

NATO UNCLASSIFIED

target is correctly classified if the output of the classification algorithm is the right target type. The percentage of objects correctly classified is defined by Eq. 14 in section 2.7. A natural object which casts a shadow similar to a known target is considered as a false alarm¹¹.

ROC Curves : Two Views

Figure 22 shows the ROC curves produced by input concatenation (IC) mode for two views. Increasing the angular interval between views from 5° to 45° and from 45° to 90° increases the classification performance. For a false alarm rate of 20 %, the classification performance gains are of $\approx 15\%$ and $\approx 20\%$ for 45° and 90° angular intervals, respectively. No gain is observed for 5° interval. For output averaging (OA) and logical product (LA) combination modes, increasing the angle between views lead to higher classification scores, irrespective of the false alarm rate. The corresponding ROC curves are shown in Figs. 23 and 24. In terms of classification performance gain, the combination modes are ranked in the following order : input concatenation, output averaging and logical product.

ROC curves : Higher Noise and Lower Resolution

The influence of noise level and height resolution on classification performance is presented in Fig. 25. The standard deviation of nominal noise is 2.5 cm and the nominal height resolution is 5 cm. When the noise standard deviation is raised to 10 cm or when the height resolution is set to 20 cm , the classification performance decreases considerably. The ROC curve for increased noise has a smooth shape. Conversely, the ROC curve for decreased resolution has a chaotic appearance which implies that single view based classifiers react in a more logical manner to increased noise than to degraded resolution. The classification performance gain for two views (with 90° interval) vanishes when noise increases or resolution decreases (Fig. 26).

ROC Curves : Three Views

Figure 27 shows the classification performance obtained by the RBF neural network when using three views. The angle between the first and the second view is 45° and the angle between the first and third is 90° . The combination modes compared are input concatenation(IC), output averaging (OA) and logical combination (LM). In

¹¹In this case, the feature vector extracted in the object shadow are insufficient for its classification. The correct classification requires the extension of the feature vector to encompass other elements of the object response

NATO UNCLASSIFIED

SACLANTCEN SM-309

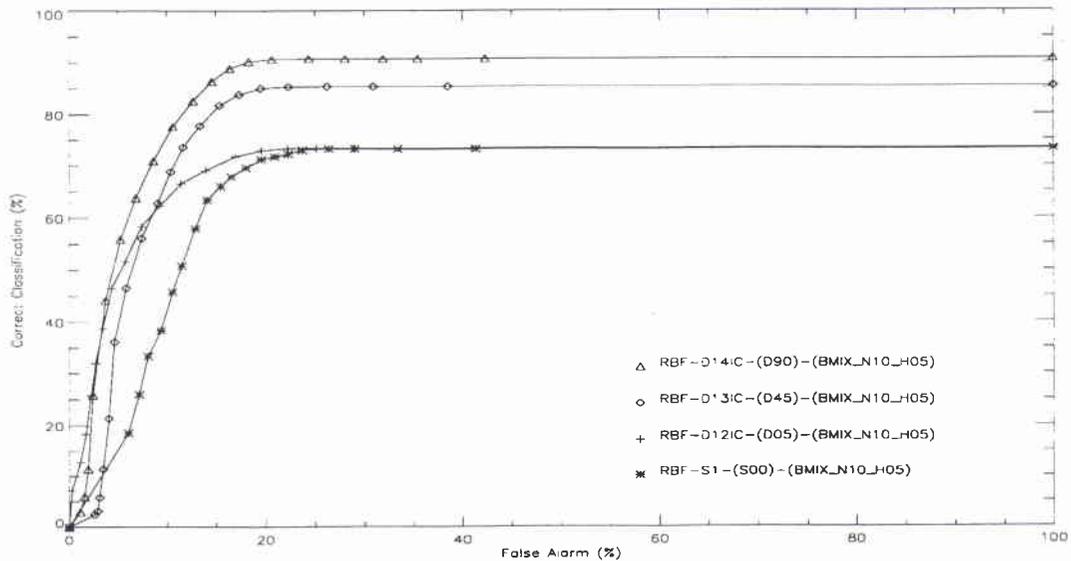


Figure 22: Comparison between single view and two view based classifiers for IC (input concatenation) combination mode (RBF neural networks). (*) S1 classifier using θ_i object orientation. (+) D12IC classifier using $\theta_i + 5^\circ$ object orientation. (\diamond) D13IC classifier using θ_i and $\theta_i + 45^\circ$ object orientations. (Δ) D14IC classifier using θ_i and $\theta_i + 90^\circ$ object orientations. $i = 1, \dots, 12000$.

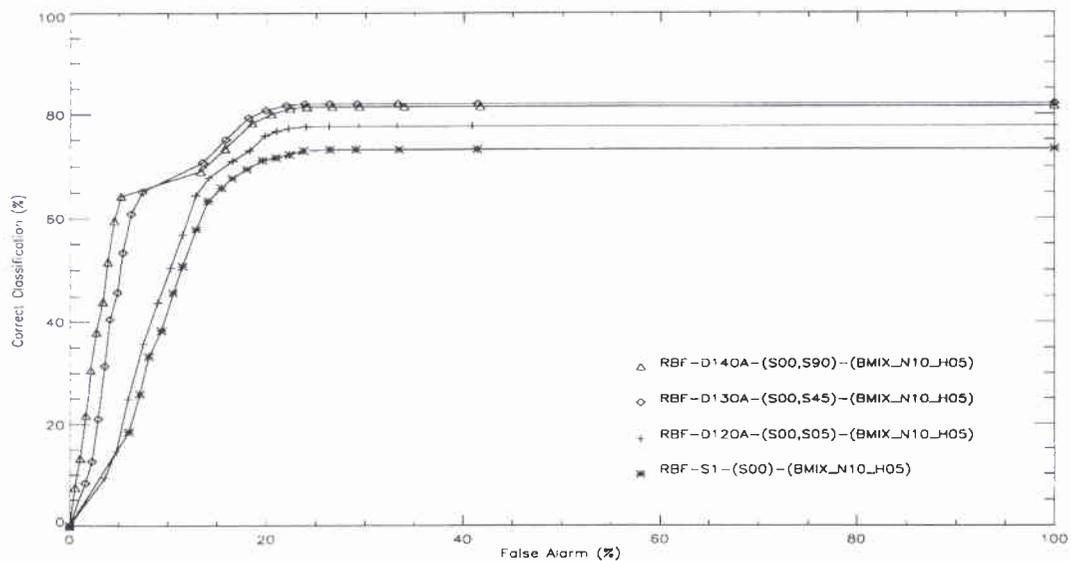


Figure 23: Comparison between single view and two view based classifiers for OA (output average) combination mode (RBF neural networks). (*) S1 classifier using θ_i object orientation. (+) D12OA classifier using $\theta_i + 5^\circ$ object orientation. (\diamond) D13OA classifier using θ_i and $\theta_i + 45^\circ$ object orientations. (Δ) D14OA classifier using θ_i and $\theta_i + 90^\circ$ object orientations. $i = 1, \dots, 12000$.

NATO UNCLASSIFIED

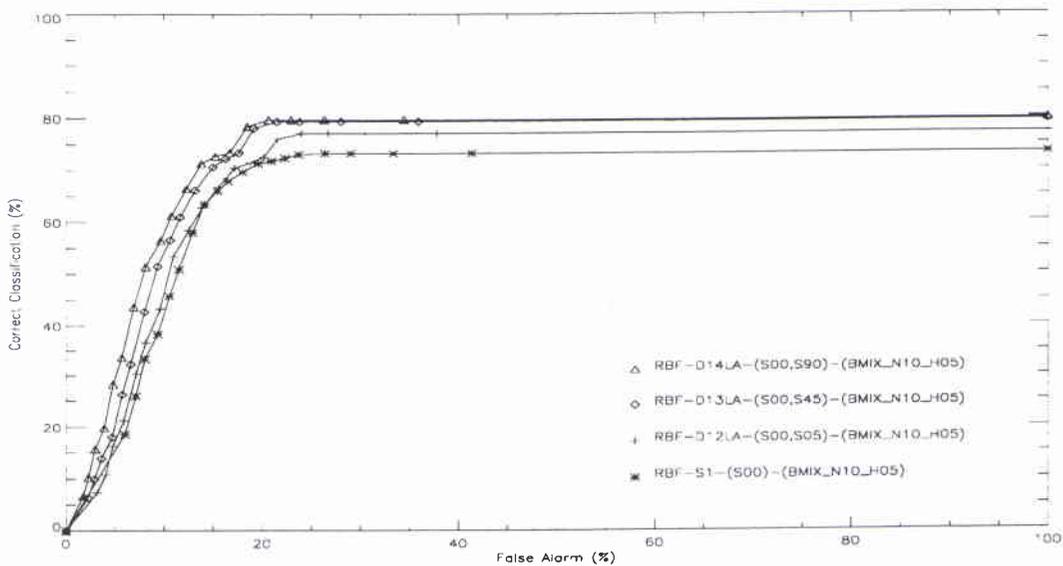


Figure 24: Comparison between single view and two view based classifiers for LA (Logical And) combination mode (RBF neural networks). (*) S1 classifier using θ_i object orientation. (+) D12LA classifier using $\theta_i + 5^\circ$ object orientation. (\diamond) D13LA classifier using θ_i and $\theta_i + 45^\circ$ object orientations. (Δ) D14LA classifier using θ_i and $\theta_i + 90^\circ$ object orientations. $i = 1, \dots, 12000$.

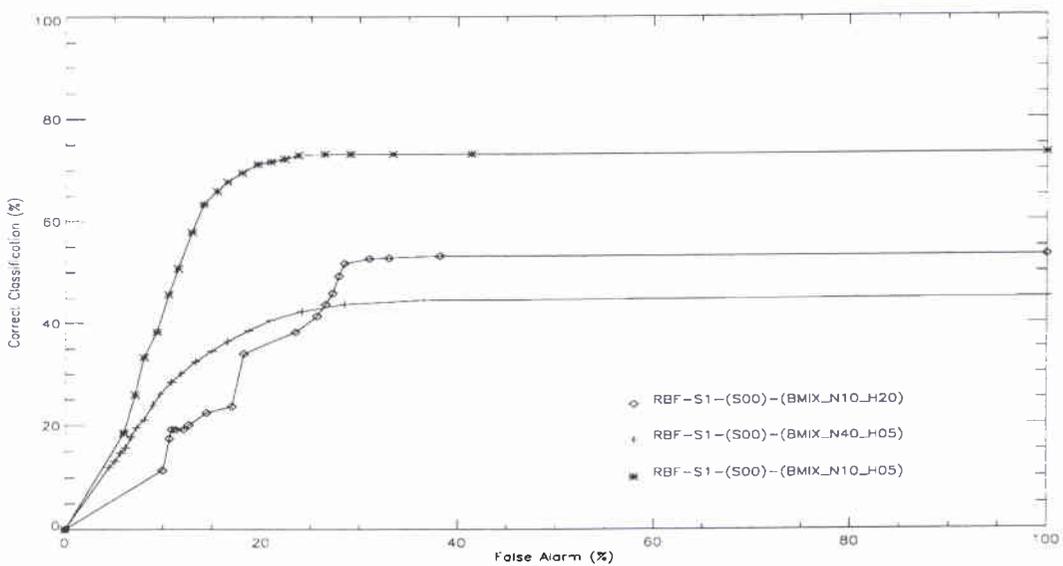


Figure 25: Effect of increased noise and decreased height resolution (H_r) on single view based classification (RBF neural networks). (*) S1 classifier using θ_i object orientation ($\sigma_{\text{noise}} = 2.5\text{cm}, H_r = 5\text{cm}$). (+) S1 classifier using θ_i object orientation ($\sigma_{\text{noise}} = 20\text{cm}, H_r = 5\text{cm}$). (\diamond) S1 classifier using θ_i object orientation ($\sigma_{\text{noise}} = 2.5\text{cm}, H_r = 20\text{cm}$). $i = 1, \dots, 12000$.

NATO UNCLASSIFIED

SACLANTCEN SM-309

terms of performance improvement, the combination modes are ranked in the same order as for two view based classification, namely IC, OA and LM.

For false alarm rates higher than 10 %, two view based IC mode performs better than three view based OA and LM modes. This demonstrates that processing simultaneously the features vectors from multiple views (IC mode) raises to the best classification scores. However, the performance improvement in IC mode results from an increased classifier complexity and a precise respect of angular intervals between views.

NATO UNCLASSIFIED

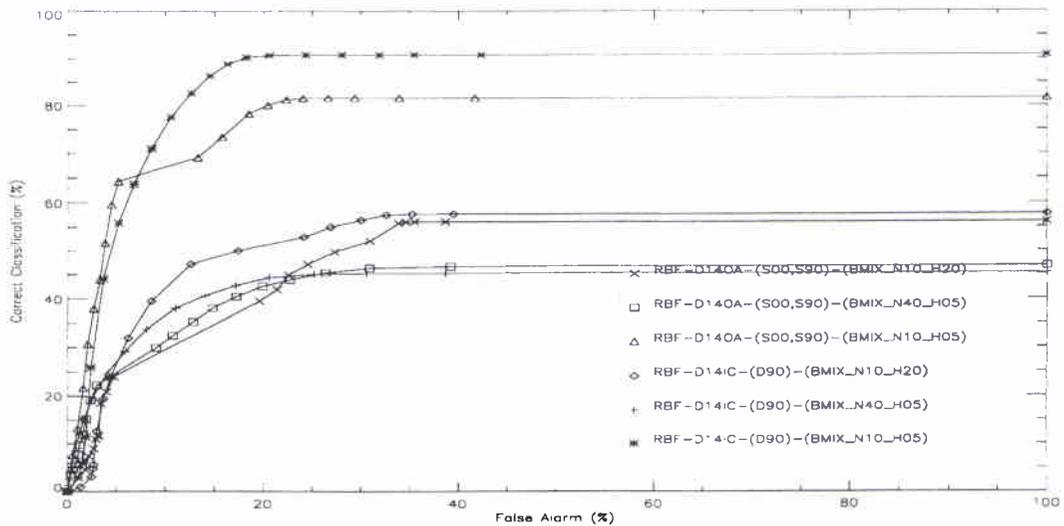


Figure 26: Effect of increased noise and decreased height resolution (H_r) on two view based classification with an angular interval of 90° (RBF neural networks). (*) D14IC classifier ($\sigma_{\text{noise}} = 2.5\text{cm}$, $H_r = 5\text{cm}$). (+) D14IC classifier ($\sigma_{\text{noise}} = 20\text{cm}$, $H_r = 5\text{cm}$). (\diamond) D14IC classifier ($\sigma_{\text{noise}} = 2.5\text{cm}$, $H_r = 20\text{cm}$). (Δ) D14OA classifier ($\sigma_{\text{noise}} = 2.5\text{cm}$, $H_r = 5\text{cm}$). (\square) D14OA classifier ($\sigma_{\text{noise}} = 20\text{cm}$, $H_r = 5\text{cm}$). (\times) D14OA classifier ($\sigma_{\text{noise}} = 2.5\text{cm}$, $H_r = 20\text{cm}$).

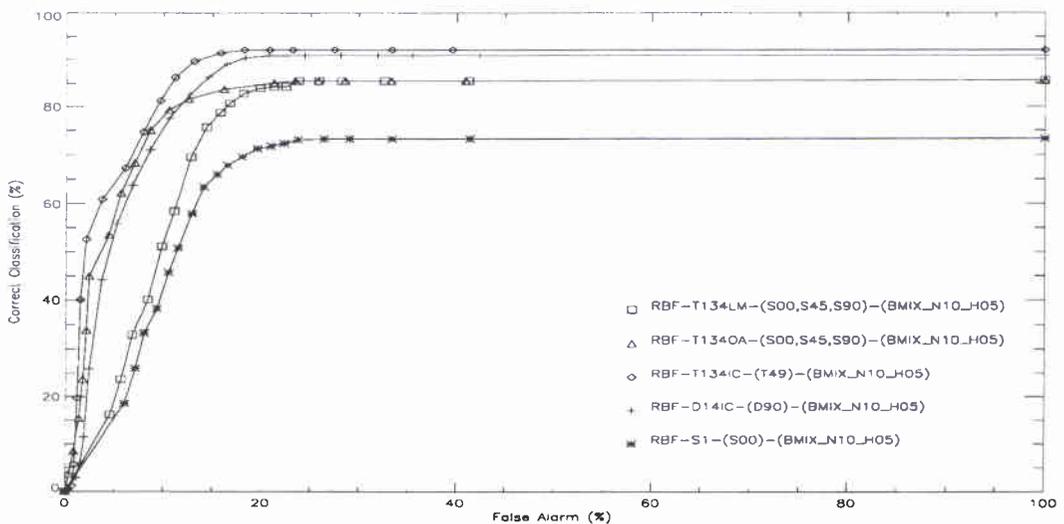


Figure 27: Comparison between single view, two view and three view based classifiers (RBF neural networks). (*) S1 classifier using θ_i object orientation. (+) D14IC classifier using the IC combination of $\theta_i + 90^\circ$ object orientation. (\diamond) T134IC classifier using the IC combination of θ_i , $\theta_i + 45^\circ$ and $\theta_i + 90^\circ$ object orientations. (Δ) T134OA classifier using the OA combination of θ_i , $\theta_i + 45^\circ$ and $\theta_i + 90^\circ$ object orientations. (\square) T134LM classifier using the LM combination of θ_i , $\theta_i + 45^\circ$ and $\theta_i + 90^\circ$ object orientations. $i = 1, \dots, 12000$.

NATO UNCLASSIFIED

SACLANTCEN SM-309

Confusion Matrices

The ROC curves show the classification performance on the 10 target set. The detailed results for each target are given in the confusion matrices. The matrices are computed for a false alarm rate of 20%. The height resolution and the noise level are nominal. The results are presented in Tables 6 to 10.

The automatic classifiers performs when all the figures of the main diagonal tend to 100 % (green color) and all the other figures tend to 0 (black color). Red color corresponds to low values on the main diagonal and high values elsewhere. The blue color corresponds to low, but non zero, values off the main diagonal.

A single view (Table 6) of ROCL, SPHR and PYRS objects gives acceptable identification scores (99.7%, 97.3% and 86.7%, respectively). The hemispherical target (HEMS) is classified correctly (94.7%) but 31.3 % of truncated cones (MNLK) are classified as HEMS. As expressed by the non mine minelike (NMML) column, the 20% of false alarms are spread over 8 target classes (all target shapes except ROCL and PYRS). This result demonstrates that, in shadow based classification, shapes with specific geometry (e.g. ROCL and PYRS) are less sensible to false alarm than more conventional shapes (e.g. HCYB and HCYC). 23.7% of VCYL are classified as non mine minelike (NMML) objects. The discrimination between HCYB, HCYC, HCSB and HPPD shapes fails.

Tables 7 to 10 show the confusion matrices using IC combination mode. Taking two views with an interval of 5° does not significantly improve the identification scores (Table 7). However, this classifier partially solves the mismatch between MNLK and HEMS classes (MNLK classified as HEMS decrease from 31.3% to 7.3%) and reduces the number of actual target shapes classified as non mines (e.g. from 23.7% to 0.7% for VCYL, from 12.3% to 0.3% for HPPD).

Using two views with an interval of 45° solves most of the confusing situations (Table 8). Only the horizontal cylinders (HCYB, HCYC and HCSB) continue to cause problems. Cylinders with different lengths are better identified (HCYC). But, for a given length, the confusion remains for cylinders with flat endcaps (HCYB) and with hemispherical endcaps (HCSB). The same trend in performance improvement is observed when then angular interval is extended to 90° (Table 9) but with better classification scores.

The IC combination of three views respecting an angular interval of 45° gives optimal identification scores (Table 10) except for the two cylinders of similar dimensions but with different endcaps (HCYB, HCSB).

This demonstrates that, for 20 cm along-track resolution and 5 cm height resolution, classification algorithms based on shadow analysis have severe difficulties to

NATO UNCLASSIFIED

discriminate between HCYB and HCSB cylinders, irrespective of the number and the direction of views.

		Actual Object Shape										
		NMML	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
Classification Results	NMML	30.8	2.3	21.7	2.3	6.7	5.0	13.0	12.3	0.0	0.0	0.3
	SPHR	0.1	97.3	0.0	0.0	4.3	3.3	4.7	0.0	0.0	0.0	0.0
	VCYL	1.9	0.0	75.7	0.0	2.3	3.0	3.7	21.0	0.0	0.0	0.0
	MNLK	0.3	0.0	0.0	62.1	1.0	1.3	3.0	0.0	0.0	3.7	0.0
	HCYB	9.0	0.0	0.0	0.0	62.0	12.7	52.3	0.0	0.0	0.0	0.0
	HCYC	4.5	0.0	0.0	0.0	0.0	63.0	0.0	0.0	0.0	0.0	0.0
	HCSB	0.1	0.3	0.7	0.0	8.7	3.0	8.7	2.0	0.0	0.0	0.0
	HPPD	3.7	0.0	0.0	4.0	11.3	7.0	13.0	64.7	0.0	0.0	0.0
	ROCL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.7	0.0	2.0
	HEMS	0.3	0.0	0.0	31.3	3.7	2.7	1.7	0.0	0.3	94.7	13.0
	PYRS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	96.3

Table 6: Confusion Matrix for S1 Classifier using RBF neural network

NATO UNCLASSIFIED

SACLANTCEN SM-309

		Actual Object Shape										
		NMML	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
Classification Results	NMML	39.3	0.0	0.7	0.3	4.7	1.7	3.0	0.3	0.0	0.3	0.0
	SPHR	0.0	97.7	0.0	0.0	0.7	0.0	1.0	0.0	0.0	0.0	0.0
	VCYL	0.0	0.0	57.7	0.0	0.7	0.0	1.0	12.7	0.0	0.0	0.0
	MNLK	0.5	0.0	0.0	91.5	1.0	0.3	1.0	0.0	0.0	1.3	0.0
	HCYB	0.1	1.0	3.0	0.0	47.0	7.3	48.3	0.7	0.0	0.0	0.0
	HCYC	16.2	1.3	37.7	1.0	33.3	42.0	30.0	10.3	3.0	4.3	14.7
	HCSB	0.1	0.0	1.0	0.0	3.7	2.0	5.0	0.3	0.0	0.0	0.0
	HPPD	2.8	0.0	0.0	0.0	8.3	0.7	10.0	55.3	0.0	0.0	0.0
	ROCL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	97.0	0.0	0.0
	HEMS	0.1	0.0	0.0	7.3	0.7	0.0	0.7	0.0	0.0	92.7	0.3
	PYRS	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	45.0

Table 7: Confusion Matrix for D12IC Classifier using RBF neural network

		Actual Object Shape										
		NMML	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
Classification Results	NMML	39.3	0.0	0.0	0.0	2.0	0.3	4.3	0.9	0.0	0.0	0.0
	SPHR	0.1	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	VCYL	0.7	0.0	100.0	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0
	MNLK	0.1	0.0	0.0	92.0	0.0	0.0	0.7	0.0	0.0	2.0	0.0
	HCYB	0.0	0.0	0.0	0.0	57.3	1.3	39.7	0.7	0.0	0.0	0.0
	HCYC	15.2	0.0	0.0	0.0	24.7	97.7	20.3	0.0	0.0	0.0	8.7
	HCSB	0.3	0.0	0.0	0.0	12.3	0.7	28.7	7.7	0.0	0.0	0.0
	HPPD	3.7	0.0	0.0	1.3	3.7	0.0	5.7	89.3	0.0	0.0	0.0
	ROCL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
	HEMS	0.0	0.0	0.0	6.7	0.0	0.0	0.7	0.0	0.0	98.0	2.7
	PYRS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	53.7

Table 8: Confusion Matrix for D13IC Classifier using RBF neural network

NATO UNCLASSIFIED

		Actual Object Shape										
		NMML	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
Classification Results	NMML	57.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	0.7	0.3
	SPHR	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	VCYL	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	MNLK	0.0	0.0	0.0	95.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
	HCYB	0.0	0.0	0.0	0.0	35.0	0.0	8.3	0.0	0.0	0.0	0.0
	HCYC	10.4	0.0	0.0	0.0	0.7	100.0	0.0	0.0	0.0	0.0	0.0
	HCSB	3.6	0.0	0.0	0.0	60.7	0.0	91.7	0.0	0.0	0.0	0.0
	HPPD	6.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0
	ROCL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
	HEMS	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	97.0	11.3
	PYRS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	88.3

Table 9: Confusion Matrix for D14IC Classifier using RBF neural network

		Actual Object Shape										
		NMML	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
Classification Results	NMML	57.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0
	SPHR	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	VCYL	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	MNLK	0.0	0.0	0.0	99.3	0.0	0.0	0.0	0.0	0.0	1.0	0.0
	HCYB	0.4	0.0	0.0	0.0	47.3	0.0	15.7	0.0	0.0	0.0	0.0
	HCYC	9.3	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
	HCSB	3.1	0.0	0.0	0.0	52.3	0.0	84.3	0.0	0.0	0.0	0.0
	HPPD	7.4	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.3
	ROCL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
	HEMS	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	99.0	9.3
	PYRS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90.3

Table 10: Confusion Matrix for T134IC Classifier using RBF neural network

NATO UNCLASSIFIEDSACLANTCEN SM-309

4.5 Searching for Ambiguity in the training data

Ambiguities are searched for in the training data using unsupervised classification techniques. The ART2 neural network processed the training data from the single view classifier. 28 prototypes have been identified Table 11 indicates the number of the activated prototype for each target and each orientation. If the number is in a box, the prototype is activated by more than one class and the corresponding feature vectors are considered ambiguous. The features are generally ambiguous, except for the ROCL polyhedron, the pyramidal sinker (PYRS) and some views of the cylinders.

The training data for two view based classifier have been processed by ART2 and the results are shown in Table 12. The angle between views is 90° . ART2 identified 66 prototypes. The number of ambiguous feature vector has decreased. The ambiguity has been resolved for SPHR, VCYL, HCYC and HPPD shapes. The feature vectors from cylinders of identical length and diameter, but with different end caps (HCYC, HCSB) remain partially ambiguous. The hemispherical object (HEMS) and the truncated cone (MNLK) can not be distinguished. HEMS and MNLK have a feature vector invariant with the orientation. For these objects, the resolution¹² is not sufficient to produce significant differences in their respective height profiles.

The comparison of RBF (Table 12) and ART2 (Table 9) results shows only minor differences. HEMS and MNLK shapes are declared ambiguous by ART2 but are satisfactorily discriminated by RBF (5% of MNLK are classified as HEMS and 1.3% of HEMS are classified as MNLK). PYRS shape is non ambiguous for ART2 but 11.3% of PYRS shapes are classified as HEMS by RBF.

ART2 based ambiguity search is a promising approach to complement, analyse and interpret the classification performed by supervised neural networks.

¹²The across-range resolution is 20 cm and the height resolution is 5 cm

θ (°)	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
0	0	2	1	0	0	0	2	3	1	4
5	0	2	1	2	2	2	2	5	1	4
10	0	2	1	2	1	2	2	6	1	4
15	0	2	1	7	8	2	2	6	1	9
20	0	2	1	8	8	7	2	10	1	9
25	0	2	1	8	11	8	2	10	1	9
30	0	2	1	8	12	8	8	10	1	9
35	0	2	1	11	12	8	8	13	1	9
40	0	2	1	12	15	8	8	17	1	9
45	0	2	1	12	16	11	8	17	1	9
50	0	2	1	12	16	12	8	17	1	9
55	0	2	1	12	16	12	8	17	1	9
60	0	2	1	12	16	12	8	17	1	9
65	0	2	1	12	18	12	8	17	1	9
70	0	2	1	12	19	12	8	17	1	9
75	0	2	1	12	19	12	8	20	1	9
80	0	2	1	12	19	12	8	20	1	4
85	0	2	1	12	19	12	8	20	1	4
90	0	2	1	12	19	12	21	20	1	4
95	0	2	1	12	19	12	8	20	1	4
100	0	2	1	12	19	12	8	20	1	4
105	0	2	1	12	19	12	8	20	1	9
110	0	2	1	12	19	12	8	17	1	9
115	0	2	1	12	18	12	8	17	1	9
120	0	2	1	12	16	12	8	17	1	9
125	0	2	1	12	16	12	8	17	1	9
130	0	2	1	12	16	12	8	17	1	9
135	0	2	1	12	16	11	8	17	1	9
140	0	2	1	12	15	8	8	17	1	9
145	0	2	1	11	12	8	8	13	1	9
150	0	2	1	8	12	8	8	10	1	9
155	0	2	1	8	11	8	2	10	1	9
160	0	2	1	8	8	7	2	6	1	9
165	0	2	1	7	8	2	2	6	1	9
170	0	2	1	2	1	2	2	6	1	4
175	0	2	1	2	2	2	2	5	1	4

Table 11: ART prototypes on a single view training data. Prototypes marked with a box correspond to ambiguous classification (.../...)

NATO UNCLASSIFIED

SACLANTCEN SM-309

θ (°)	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
180	0	2	1	0	0	0	2	3	1	4
185	0	2	1	2	2	2	2	5	1	4
190	0	2	1	2	1	2	2	22	1	4
195	0	2	1	7	8	2	2	23	1	9
200	0	2	1	8	8	7	2	23	1	9
205	0	2	1	8	11	8	2	24	1	9
210	0	2	1	8	12	8	8	24	1	9
215	0	2	1	11	12	8	8	25	1	9
220	0	2	1	12	15	8	8	26	1	9
225	0	2	1	12	16	11	8	26	1	9
230	0	2	1	12	16	12	8	26	1	9
235	0	2	1	12	16	12	8	26	1	9
240	0	2	1	12	16	12	8	26	1	9
245	0	2	1	12	18	12	8	26	1	9
250	0	2	1	12	19	12	8	26	1	9
255	0	2	1	12	19	12	8	27	1	9
260	0	2	1	12	19	12	8	27	1	4
265	0	2	1	12	19	12	8	27	1	4
270	0	2	1	12	19	12	21	27	1	4
275	0	2	1	12	19	12	8	27	1	4
280	0	2	1	12	19	12	8	27	1	4
285	0	2	1	12	19	12	8	27	1	9
290	0	2	1	12	19	12	8	26	1	9
295	0	2	1	12	18	12	8	26	1	9
300	0	2	1	12	16	12	8	26	1	9
305	0	2	1	12	16	12	8	26	1	9
310	0	2	1	12	16	12	8	26	1	9
315	0	2	1	12	16	11	8	26	1	9
320	0	2	1	12	15	8	8	26	1	9
325	0	2	1	11	12	8	8	25	1	9
330	0	2	1	8	12	8	8	24	1	9
335	0	2	1	8	11	8	2	24	1	9
340	0	2	1	8	8	7	2	23	1	9
345	0	2	1	7	8	2	2	23	1	9
350	0	2	1	2	1	2	2	22	1	4
355	0	2	1	2	2	2	2	5	1	4

Table 11: (continued) ART prototypes on a single view training data. Prototypes marked with a box correspond to ambiguous classification

NATO UNCLASSIFIED

θ (°)	SPHR	VCYL	MNLK	HCYB	HCCY	HCSB	HPPD	ROCL	HEMS	PYRS
0	0	7	1	2	3	2	4	5	1	6
5	0	7	1	2	3	2	8	9	1	6
10	0	7	1	2	10	2	8	12	1	6
15	0	7	1	16	14	2	8	12	1	13
20	0	7	1	16	14	16	8	15	1	13
25	0	7	1	16	17	16	8	15	1	13
30	0	7	1	16	18	16	19	15	1	13
35	0	7	1	20	18	16	19	21	1	13
40	0	7	1	25	23	16	19	28	1	13
45	0	7	1	25	23	24	19	28	1	13
50	0	7	1	25	26	34	19	28	1	13
55	0	7	1	25	29	34	19	28	1	13
60	0	7	1	34	29	34	19	30	1	13
65	0	7	1	34	31	34	32	30	1	13
70	0	7	1	34	33	34	32	30	1	13
75	0	7	1	34	33	40	32	36	1	13
80	0	7	1	40	37	40	32	36	1	6
85	0	7	1	40	38	40	32	39	1	6
90	0	7	1	40	38	40	41	42	1	6
95	0	7	1	40	38	40	32	39	1	6
100	0	7	1	40	37	40	32	44	1	6
105	0	7	1	34	33	40	32	44	1	13
110	0	7	1	34	33	34	32	45	1	13
115	0	7	1	34	31	34	32	45	1	13
120	0	7	1	34	29	34	19	45	1	13
125	0	7	1	25	29	34	19	46	1	13
130	0	7	1	25	26	34	19	47	1	13
135	0	7	1	25	23	24	19	47	1	13
140	0	7	1	25	23	16	19	47	1	13
145	0	7	1	20	18	16	19	47	1	13
150	0	7	1	16	18	16	19	48	1	13
155	0	7	1	16	17	16	8	48	1	13
160	0	7	1	16	14	16	8	48	1	13
165	0	7	1	16	14	2	8	49	1	13
170	0	7	1	2	10	2	8	49	1	6
175	0	7	1	2	3	2	8	50	1	6

Table 12: ART prototypes on a two views training data. Angle between views is 90°. Prototypes marked with a box correspond to ambiguous classification (.../...)

NATO UNCLASSIFIED

SACLANTCEN SM-309

θ (°)	SPHR	VCYL	MNLK	HCYB	HCYC	HCSB	HPPD	ROCL	HEMS	PYRS
180	0	7	1	2	3	2	4	51	1	6
185	0	7	1	2	3	2	8	50	1	6
190	0	7	1	2	10	2	8	52	1	6
195	0	7	1	16	14	2	8	52	1	13
200	0	7	1	16	14	16	8	53	1	13
205	0	7	1	16	17	16	8	53	1	13
210	0	7	1	16	18	16	19	53	1	13
215	0	7	1	20	18	16	19	54	1	13
220	0	7	1	25	23	16	19	55	1	13
225	0	7	1	25	23	24	19	55	1	13
230	0	7	1	25	26	34	19	55	1	13
235	0	7	1	25	29	34	19	55	1	13
240	0	7	1	34	29	34	19	56	1	13
245	0	7	1	34	31	34	32	56	1	13
250	0	7	1	34	33	34	32	56	1	13
255	0	7	1	34	33	40	32	57	1	13
260	0	7	1	40	37	40	32	57	1	6
265	0	7	1	40	38	40	32	58	1	6
270	0	7	1	40	38	40	41	59	1	6
275	0	7	1	40	38	40	32	58	1	6
280	0	7	1	40	37	40	32	60	1	6
285	0	7	1	34	33	40	32	60	1	13
290	0	7	1	34	33	34	32	61	1	13
295	0	7	1	34	31	34	32	61	1	13
300	0	7	1	34	29	34	19	61	1	13
305	0	7	1	25	29	34	19	62	1	13
310	0	7	1	25	26	34	19	63	1	13
315	0	7	1	25	23	24	19	63	1	13
320	0	7	1	25	23	16	19	63	1	13
325	0	7	1	20	18	16	19	63	1	13
330	0	7	1	16	18	16	19	64	1	13
335	0	7	1	16	17	16	8	64	1	13
340	0	7	1	16	14	16	8	64	1	13
345	0	7	1	16	14	2	8	65	1	13
350	0	7	1	2	10	2	8	65	1	6
355	0	7	1	2	3	2	8	9	1	6

Table 12: (Continued) ART prototypes on a two views training data. Angle between views is 90° . Prototypes marked with a box correspond to ambiguous classification

NATO UNCLASSIFIED

4.6 Synthesis of Classification Performance

Different neural networks, different ways to introduce multiple views and different settings for the orientation of the views were tested. Table 13 summarizes the target classification performance for RBF neural networks.

Combination Mode	Correct Classification (%)	
	FA%=10%	FA%=20%
S1	41.6	71.4
S2	45.6	72.2
S3	39.4	71.7
S4	42.1	71.2
D12IC	64.4	73.1
D13IC	66.9	85.2
D14IC	75.7	90.7
D12OA	49.8	76.0
D13OA	69.3	81.1
D14OA	68.4	79.9
D12LA	47.2	72.5
D13LA	55.0	78.4
D14LA	57.7	79.7
D12LO	44.5	75.9
D13LO	29.3	77.3
D14LO	31.8	76.9
T134OA	78.6	84.5
T134LM	52.3	84.1
T134IC	82.4	92.0

Table 13: Synthesis of classification scores for given false alarm rates (RBF neural networks). Columns 2 and 3 indicate the percentage of correct classification for a false alarm rate of 10% and 20%, respectively.

Neural Networks

The combination of MLP and KOH has shown an overall performance inferior to RBF. Tuning a neural network is a difficult task, involving empirical reasoning. This task is more complex for an association of two types of neural networks. The ART2 neural network, used here only to detect ambiguity in the training data, has opened a promising future.

Classification Performance Using Two Views

The ROC curves show better results when the angle between views increases, irrespective of the combination mode¹³. These ROC curves show that target identification benefits from multiple views with large angular intervals between views.

¹³Only the magnitude of the improvement changes with the combination mode.

NATO UNCLASSIFIEDSACLANTCEN SM-309

The IC mode proved to be the most effective way to merge the feature vectors from multiple view in a target classifier. The best classification performance of two view based systems is attained for an interval of 90° . The fact that most of the asymmetrical targets have the highest change in aspect for orthogonal views explains the good performance of a 90° angle between views.

The OA mode performs less well. In this mode, the feature vectors from the two views are processed separately. The angle between views has therefore a lower effect on classification performance. Compared to IC combination mode, OA does not require a precise (*a priori* defined) angular interval between views. The fact that one view can be better than the other is sufficient to improve the target classification. For shadow based classification, a better view can be provided by a different target orientation. If the angular interval varies within range of 45° to 90° , OA combination mode produces similar performance improvement.

The logical combination (LA, LO) modes perform least well.

Multiple View Classifier

Introducing additional views with IC mode increases the complexity of the neural network and requires that the angles between the views precisely respect the values used for classifier initialization. If one view is not correctly oriented, the performance of IC mode decreases. If one view is missing, (e.g. object occluded by another), IC mode cannot achieve the classification. These situations are processed more easily by OA mode. If one view is missing, the output average is computed only on the remaining views.

A reliable implementation of multiple view automatic classifier requires IC and OA combination modes. If the views cannot be acquired with the predefined angular intervals, IC mode is replaced by OA mode. The most suitable configuration for the concurrent implementation of IC and OA modes is based on three views respecting angular intervals of 45° . Practically, the multiple view sidescan sonar is a conventional system with two additional beam orientations at -45° and $+45^\circ$ to broadside.

Feature Vectors

The classification results are strongly linked to the fidelity of the model used for generating natural objects. The performance scores demonstrate that processing multiple views improves the target classification even when using an extremely simple feature vector (object height profile).

NATO UNCLASSIFIED

5

Capability to Discriminate
between Target Shapes

The classification results presented in the former section are strongly associated with the capacity of the model to generate natural objects. As this modelling approach has not been experimentally validated, classification performance is not susceptible to generalization, a limitation which has been surmounted by considering only known targets. In this context, the purpose of the classification algorithm is to discriminate between a given mine and all other known mines, irrespective of orientation. This approach, computationally less demanding than the implementation of automatic classifiers, is designed to rapidly estimate the influence of multiple views and image resolution on classification performance.

Assuming that classification is based on the shadow, the difference between two height profiles is expressed by

$$\Delta s_{ik}(\theta_j, \theta_l) = \sum_{p=1}^{N_p} | P_i(\theta_j, p) - P_k(\theta_l, p) | \quad (17)$$

$$p = 1, \dots, N_p$$

where $P_i(\theta_j, p)$ is the object height profile of target i viewed at azimuth angle θ_j . Similarly, $P_k(\theta_l, p)$ is the height profile of target k viewed at azimuth θ_l . N_p is the feature vector length.

The height profile difference $\Delta T s_i(\theta)$ for which the discrimination of target i at azimuth θ is the least effective is expressed by

$$\Delta T s_i(\theta) = \frac{\min[\Delta s_{ik}(\theta, \theta_l)]}{\Delta h}$$

$$k = 1, \dots, N_t, \quad k \neq i \quad (18)$$

$$\theta_l = l\Delta\theta, \quad l = 0, \dots, N_\theta - 1$$

Δh is the height resolution. The variation of grazing angle is taken into account in the height resolution. The azimuth angle θ is sampled in N_θ steps of $\Delta\theta^\circ$. N_θ is

NATO UNCLASSIFIED

SACLANTCEN SM-309

the number of aspects produced during a circumnavigation of the object. $\Delta T s_i(\theta)$ is the minimum difference, expressed in number of resolution cells, between the height profile of target i seen at azimuth θ and height profiles of the other targets.

A height profile can be discriminated from another if the difference $\Delta T s_i(\theta)$ is greater than a given number of resolution cells (N_{id}). If for each azimuth angles, the difference is greater than N_{id} cells, then the target can always be identified. The capability to discriminate target i seen at azimuth angle θ from the other targets is expressed by

$$T D s_i(\theta) = \begin{cases} 1 & \text{if } \Delta T s_i(\theta) \geq N_{id} \\ \frac{\Delta T s_i(\theta)}{N_{id}} & \text{if } \Delta T s_i(\theta) < N_{id} \end{cases} \quad (19)$$

The function $T D s_i$ is called the discrimination capability of target i .

Multiple aspects of the object are introduced by evaluating the difference between multiple height profiles. For multiple height profiles simultaneously processed, Eq. 17 is replaced by

$$\Delta m_{ik}(\theta_j, \theta_l) = \sum_{v=1}^{N_v} \sum_{p=1}^{N_p} | P_i(\theta_j + \Theta(v), p) - P_k(\theta_l + \Theta(v), p) | \quad (20)$$

$p = 1, \dots, N_p \quad , \quad v = 1, \dots, N_v$

where N_v is the number of views and $\Theta(v)$ defines the angular intervals between the first and the subsequent views. The single view configuration corresponds to $N_v = 1$ and $\Theta = [0]$ The lowest discrimination $\Delta m_{ik}(\theta_j, \theta_l)$ for target i at azimuth θ is now given by

$$\Delta T m_i(\theta_j) = \frac{\min[\Delta m_{ik}(\theta_j, \theta_l)]}{\Delta h} \quad (21)$$

$k = 1, \dots, N_t, \quad k \neq i$

$\theta_l = l\Delta\theta, \quad l = 0, \dots, N_\theta - 1$

As for the single view case (Eq. 19), the discrimination capability $T D m_i$ is determined by

$$T D m_i(\theta) = \begin{cases} 1 & \text{if } \Delta T m_i(\theta) \geq N_{id} \\ \frac{\Delta T m_i(\theta)}{N_{id}} & \text{if } \Delta T m_i(\theta) < N_{id} \end{cases} \quad (22)$$

NATO UNCLASSIFIED

An alternative way to introduce multiple views consists of retaining only the best view. The global sum of the height profiles differences (Eq. 20) is replaced by the maximum difference between individual height profiles

$$\Delta w_{ik}(\theta_j, \theta_l) = \max_{v=1}^{N_v} \sum_{p=1}^{N_p} | P_i(\theta_j + \Theta(v), p) - P_k(\theta_l + \Theta(v), p) | \quad (23)$$

The computation of the discrimination capability is achieved by replacing Δm_{ik} by Δw_{ik} and TDm_i by $TDw_i(\theta)$ in Eqs. 21 and 22, respectively.

Simple Example

The computation of the discrimination capability is illustrated in a simple example. Two target shapes are considered : a sphere and a cylinder of identical diameter (50 cm). The cylinder is 1.5 m long. Figure 28 shows the capability to discriminate between the sphere and the cylinder for different sonar resolutions. The horizontal and vertical dimensions of the resolution cell are the across-range and the height resolution, respectively. The number¹⁴ of resolution cells required for target classification is $N_{id} = 30$. The discrimination capability vanishes for the sphere, irrespective of the orientation (Fig. 28(a)). For the cylinder, the discrimination capability varies with the orientation (Fig. 28(c)). When the orientation of the cylinder is 0° or 180° , its height profile is very close to the height profile of the sphere. The capability to discriminate the cylinder is absent at 0° and 180° but discrimination is maximized at 90° and 270° . Increasing the resolution (Figs. 28(b) and 28(d)) leads to better discrimination capabilities but is insufficient to eliminate the confusion for 0° and 180° .

The simultaneous processing of multiple aspects improves the capability to discriminate the sphere from the cylinder (Fig. 29) and *viceversa* (Fig.30). The improvement is minor when, instead of simultaneously processing all views, only the best view is retained (Figs. 31 and 32).

¹⁴In an ideal case, $N_{id} = 1$ should be sufficient to discriminate between two height profiles. Considering that the height profile is extracted from the object shadow, which itself results from the segmentation of the sonar image, several resolution cells will be required to differentiate between two height profiles. Considering that the object intercepts from 5 to 10 beams and that, for each beam, the deviation on the estimated height is 2 to 5 cells, an intermediate value of 30 cells has been chosen.

NATO UNCLASSIFIED

SACLANTCEN SM-309

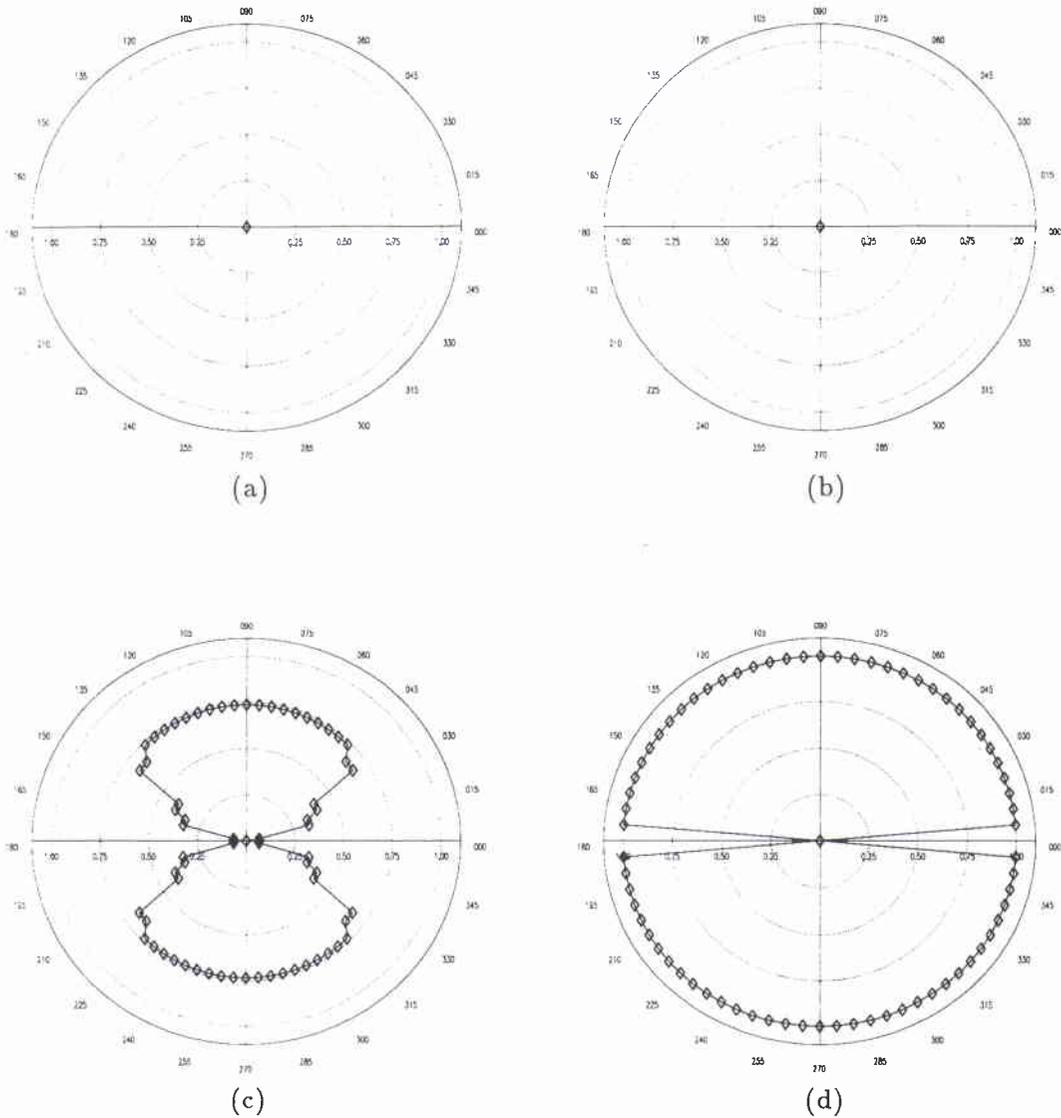


Figure 28: Capability to discriminate between a sphere and a cylinder using a single view. (a) Sphere , Beamwidth = 20 cm, height resolution = 10 cm, (b) Sphere , Beamwidth = 10 cm, height resolution = 2.5 cm, (c) Cylinder , Beamwidth = 20 cm, height resolution = 10 cm, (d) Cylinder , Beamwidth = 10 cm, height resolution = 2.5 cm.

NATO UNCLASSIFIED

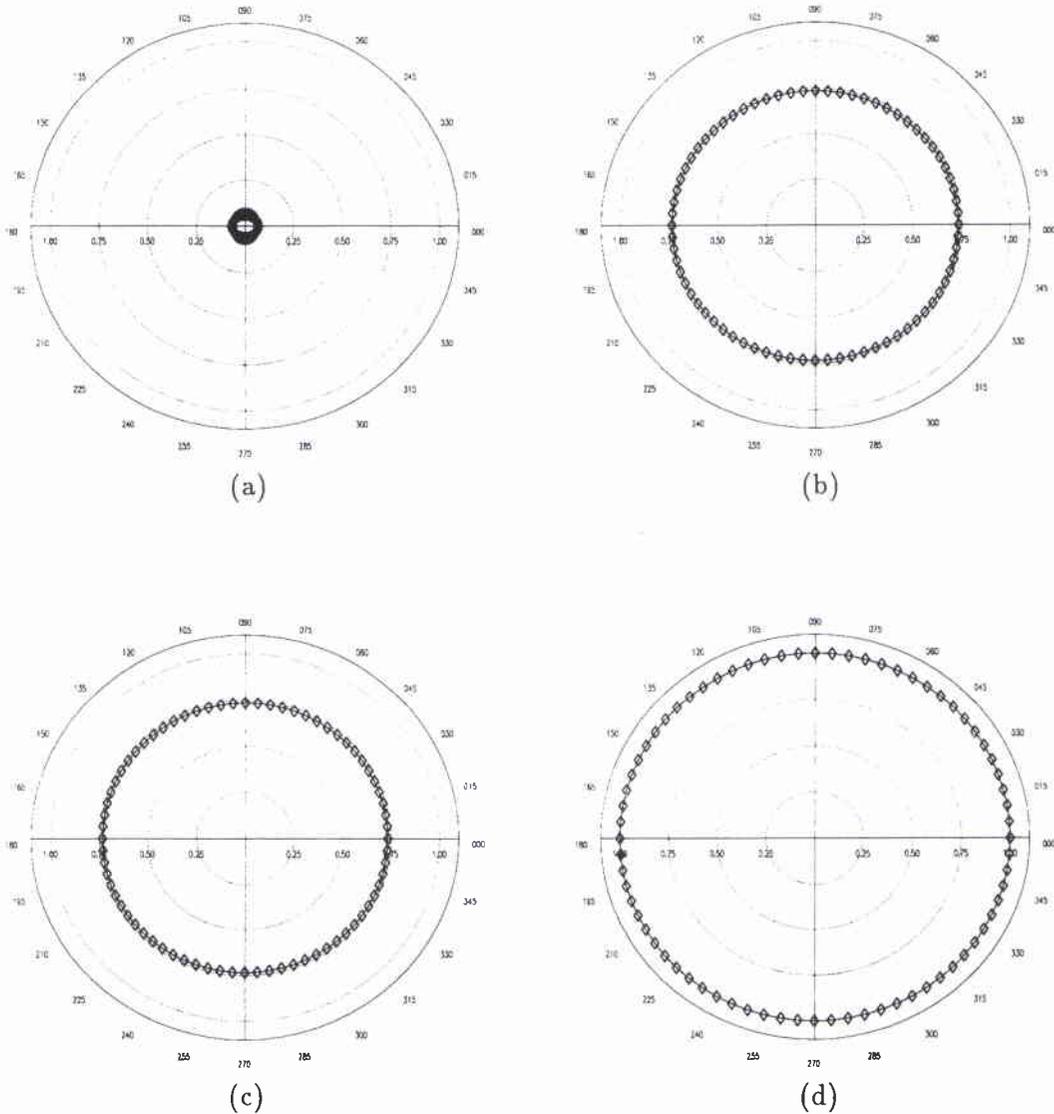


Figure 29: Capability to discriminate between a sphere and a cylinder using multiple views. Using simultaneously all the views. Beamwidth = 20 cm and height resolution = 10 cm. (a) Sphere, two views with an interval of 5° (b) Sphere, two views with an interval of 45° (c) Sphere, two views with an interval of 90° (d) Sphere, three views with intervals of 45°

NATO UNCLASSIFIED

SACLANTCEN SM-309

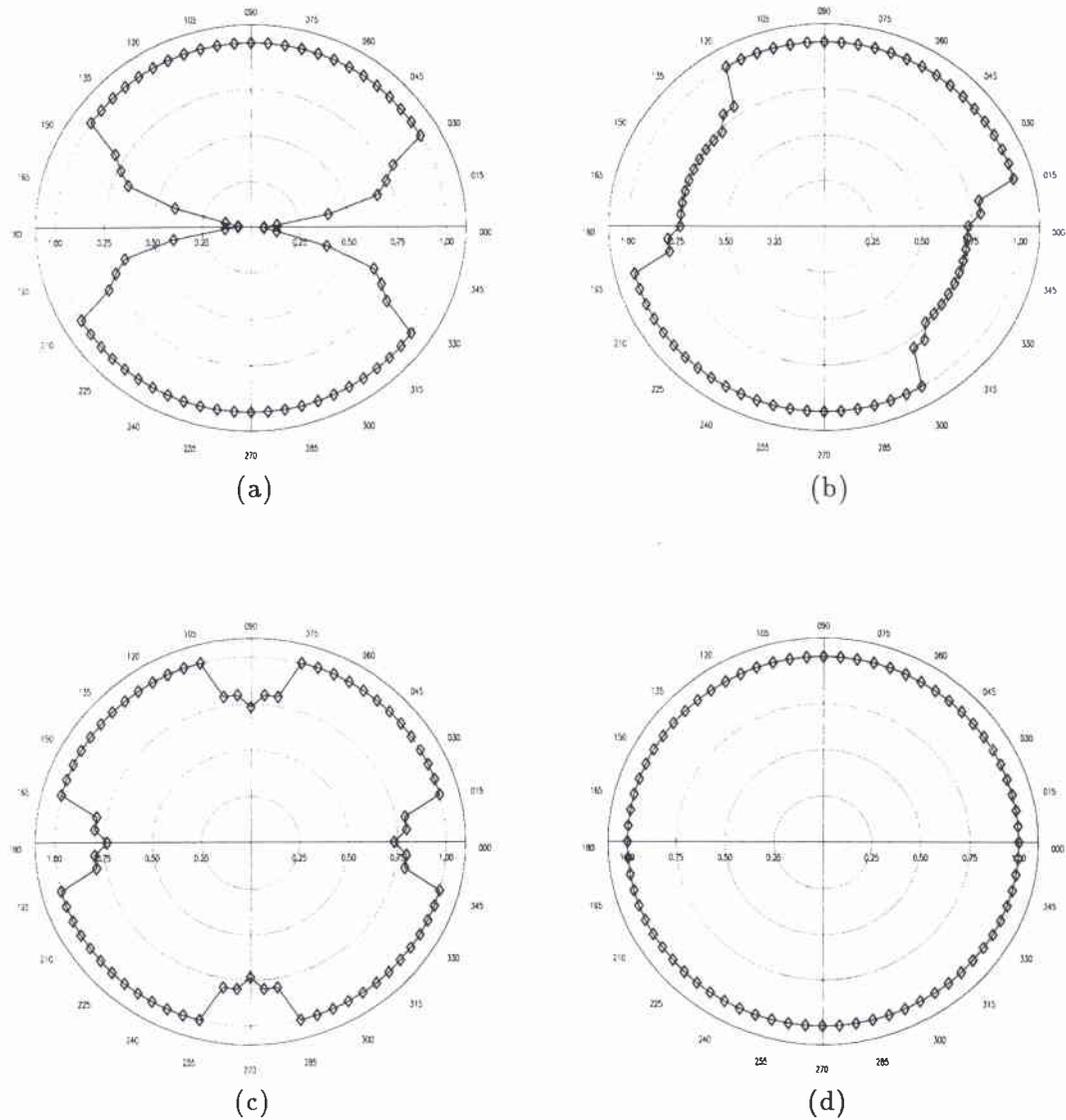


Figure 30: Capability to discriminate between a sphere and a cylinder using multiple views. Using simultaneously all the views. Beamwidth = 20 cm and height resolution = 10 cm. (a) Cylinder, two views with an interval of 5° (b) Cylinder, two views with an interval of 45° (c) Cylinder, two views with an interval of 90° (d) Cylinder, three views with intervals of 45°

NATO UNCLASSIFIED

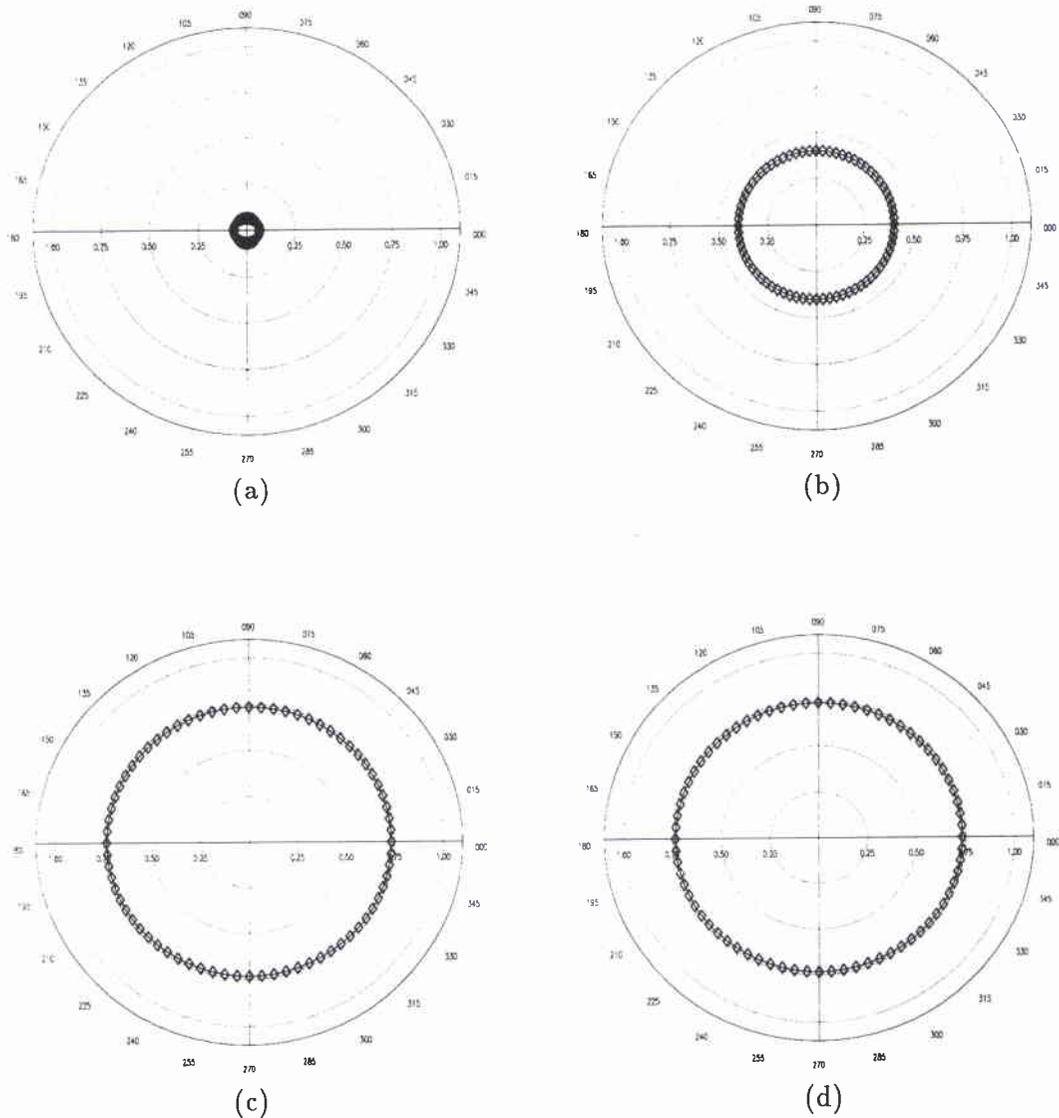


Figure 31: Capability to discriminate between a sphere and a cylinder using multiple views. Using only the best view. Beamwidth = 20 cm and height resolution = 10 cm. (a) Sphere, two views with an interval of 5° (b) Sphere, two views with an interval of 45° (c) Sphere, two views with an interval of 90° (d) Sphere, three views with intervals of 45°

NATO UNCLASSIFIED

SACLANTCEN SM-309

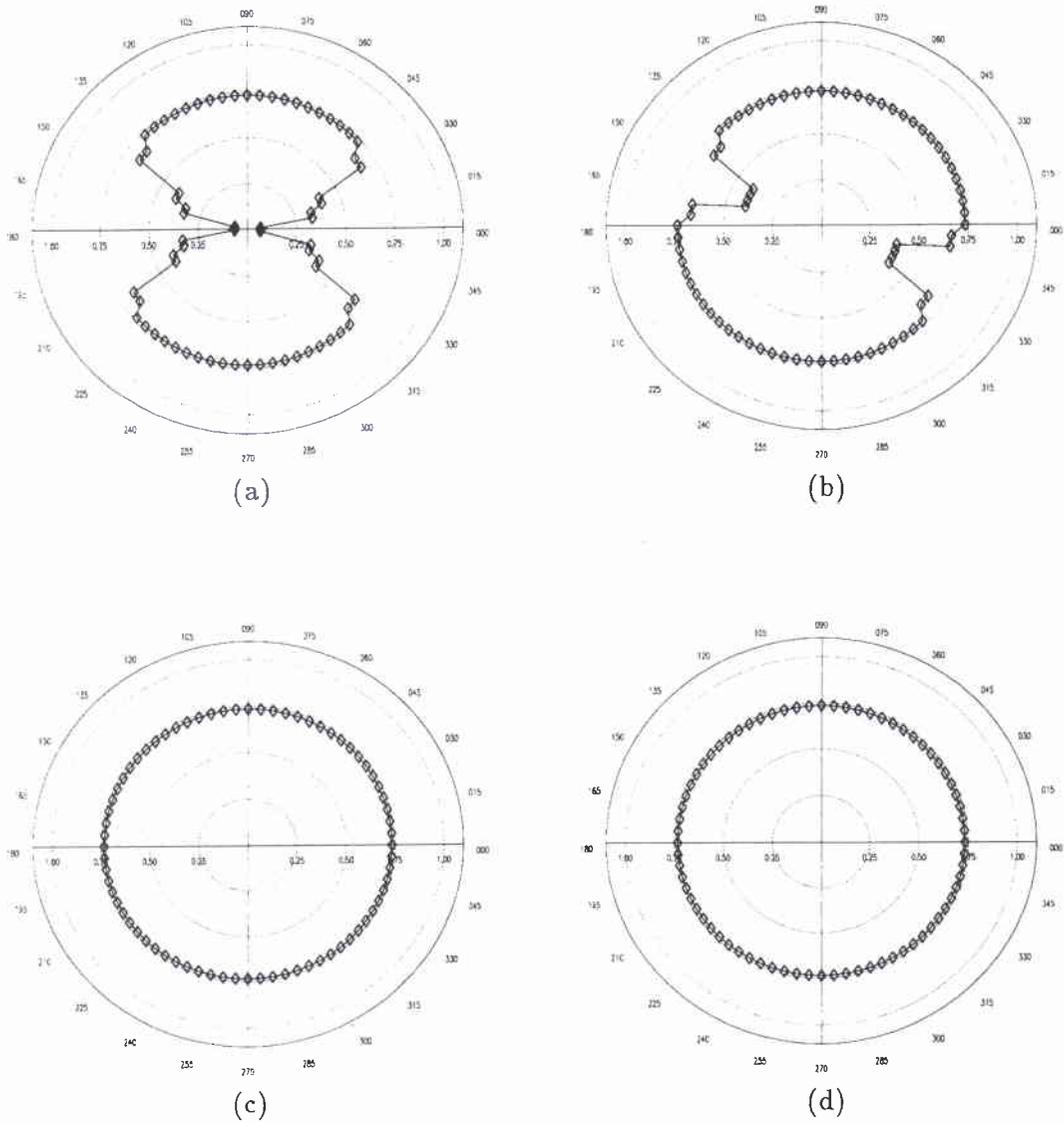


Figure 32: Capability to discriminate between a sphere and a cylinder using multiple views. Using only the best view. Beamwidth = 20 cm and height resolution = 10 cm. (a) Cylinder, two views with an interval of 5° (b) Cylinder, two views with an interval of 45° (c) Cylinder, two views with an interval of 90° (d) Cylinder, three views with intervals of 45°

NATO UNCLASSIFIED

Parametric Study on 10 Target Shapes

A parametric study is conducted in order to balance the effects of across-range resolution (beamwidth), height resolution and multiple aspects on the capability to discriminate between targets. The inputs to this study are ten target shapes (shape and dimensions are defined in Table 2). The across-range resolution varies from 5 cm to 50 cm and the height resolution from 2.5 cm to 10 cm. The aspect configuration assumes the following values : [0°], [0°, 5°], [0°, 45°], [0°, 90°], [0°, 45°, 90°] and [0°, 5°, 10°]. For given across-range and height resolutions, the average discrimination capability TD_m is given by

$$TD_m = \frac{1}{N_t N_\theta} \sum_{i=1}^{N_t} \sum_{l=0}^{N_\theta-1} TD_{m_i}(l\Delta\theta) \quad (24)$$

N_t is the number of targets. N_θ is the number of orientations, $\Delta\theta$ the angular interval between aspects and $TD_{m_i}(\theta)$ is the discrimination capability of target i viewed with orientation θ .

The required number of resolution cells to discriminate between two height profile is $N_{id} = 30$. The height profiles, extracted from the multiple views, are simultaneously processed. Figure 33 shows TD_m for for the 10 targets benchmark. TD_m is plotted versus across-range resolution for several aspect configurations and a height resolution of 5 cm. The aspect configurations are represented with different colors and markers:

Aspect Configuration:	Color:	Marker:
Single View [0°]	Red	★
Double View [0°, 5°]	Green	+
Double View [0°, 45°]	Blue	◇
Double View [0°, 90°]	Cyan	△
Triple View [0°, 45°, 90°]	Magenta	□
Triple View [0°, 5°, 10°]	Orange	×

The best discrimination is obtained for three views with 45° intervals. Then, two views with 90° interval give better results than three views with 5° intervals. Then, two views with 45° interval give better results than two views with 5° interval. Finally, a single view provides the worst discrimination capability.

The main indication from these curves is the preponderant role of across-range resolution. The simultaneous processing of multiple views improves the discrimination

NATO UNCLASSIFIED

SACLANTCEN SM-309

across-range resolution, TD_m is 0.23 using a single view, 0.72 using two views with 90° interval and 0.86 using three views with 45° intervals. For identical aspect configuration, 40 cm across-range resolution leads to TD_m of 0.06, 0.15 and 0.21, respectively. The comparison of these results shows that a single high resolution (10 cm) view of the target and three low resolution (40 cm) views lead to similar discrimination capability.

Figure 34 shows the variation of TD_m with height resolution. The top and bottom subplots of Fig. 34 have been obtained with a height resolution of 10 cm and 2.5, respectively. As for across-range resolution, lack of height resolution attenuates the benefits of multiple view processing.

Figure 35 shows how TD_m varies with object shapes. The top and bottom subplots of Fig. 35 have been obtained with subsets of 4 symmetrical shapes and 4 non symmetrical shapes, respectively. Shapes are symmetrical when their height profile is invariable with observation conditions. For these symmetrical shapes, the discrimination capability increases with the number of views, irrespective of the angular interval between views.

Using the best view only increases discrimination capability (see Fig. 36). However, using simultaneously all the views (see Fig. 33) provides higher improvement. When using the best view only, the discrimination capability is more sensible to the sonar resolution and does not improve for multiple views of symmetrical shapes.

NATO UNCLASSIFIED

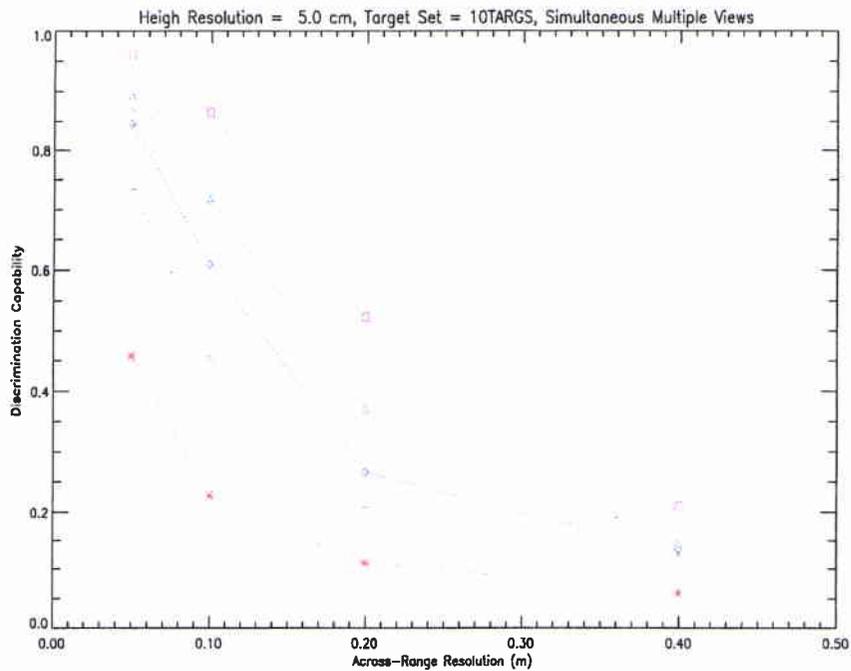


Figure 33: Discrimination Capability versus across-range resolution for a benchmark of 10 target shapes. Multiple views are simultaneously processed. Height resolution is 5 cm. Multiple view configurations are marked as follows: Red *: Single View ([0°]). Green + : Double View ([0° , 5°]). Blue o : Double View ([0° , 45°]). Cyan Δ : Double View ([0° , 90°]). Magenta □ : Triple View ([0° , 45° , 90°]). Orange x: Triple View ([0° , 5° , 10°]).

NATO UNCLASSIFIED

SACLANTCEN SM-309

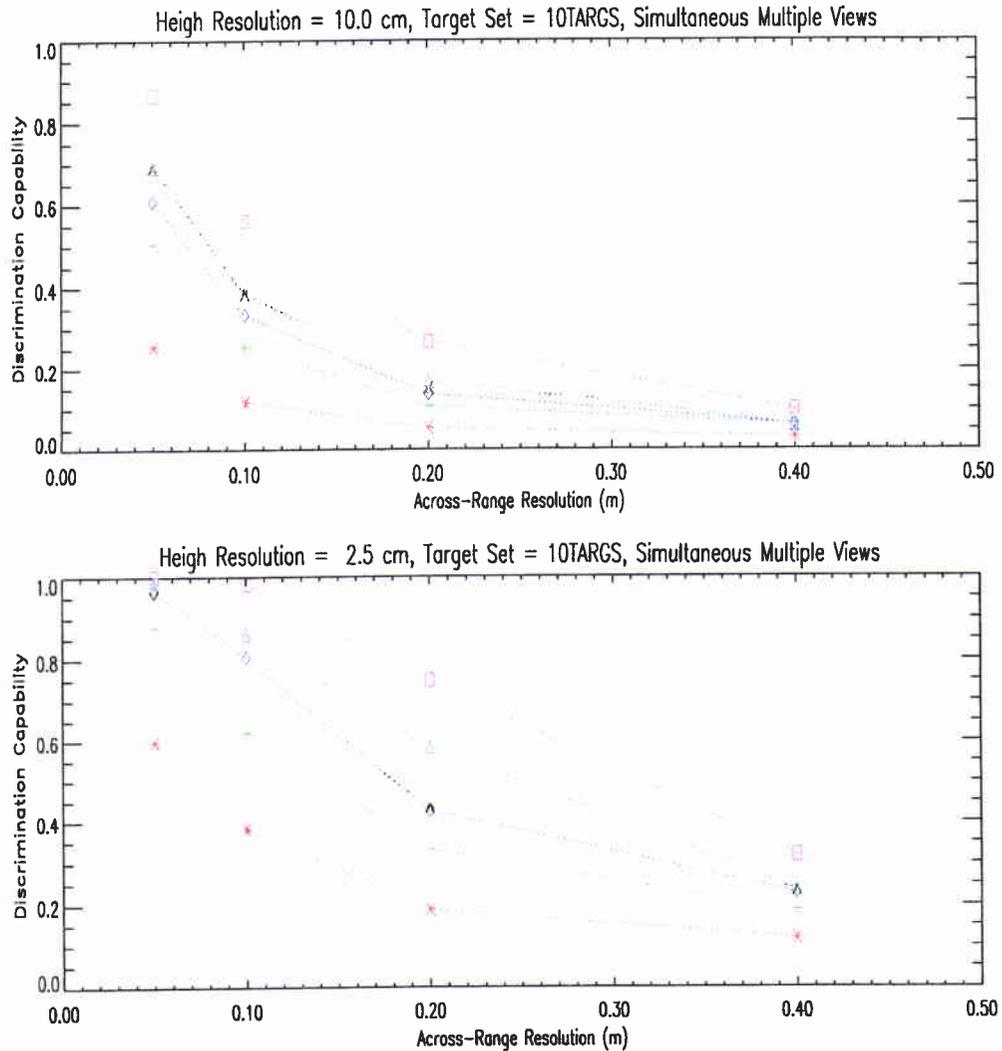


Figure 34: Discrimination Capability versus across-range resolution for a benchmark of 10 target shapes. Multiple views are simultaneously processed. Height resolution is 10 cm and 2.5 cm for top and bottom plots, respectively. Multiple view configurations are marked as follows : Red * : Single View ([0°]). Green + : Double View ([0° , 5°]). Blue ◇ : Double View ([0° , 45°]). Cyan △ : Double View ([0° , 90°]). Magenta □ : Triple View ([0° , 45° , 90°]). Orange × : Triple View ([0° , 5° , 10°]).

NATO UNCLASSIFIED

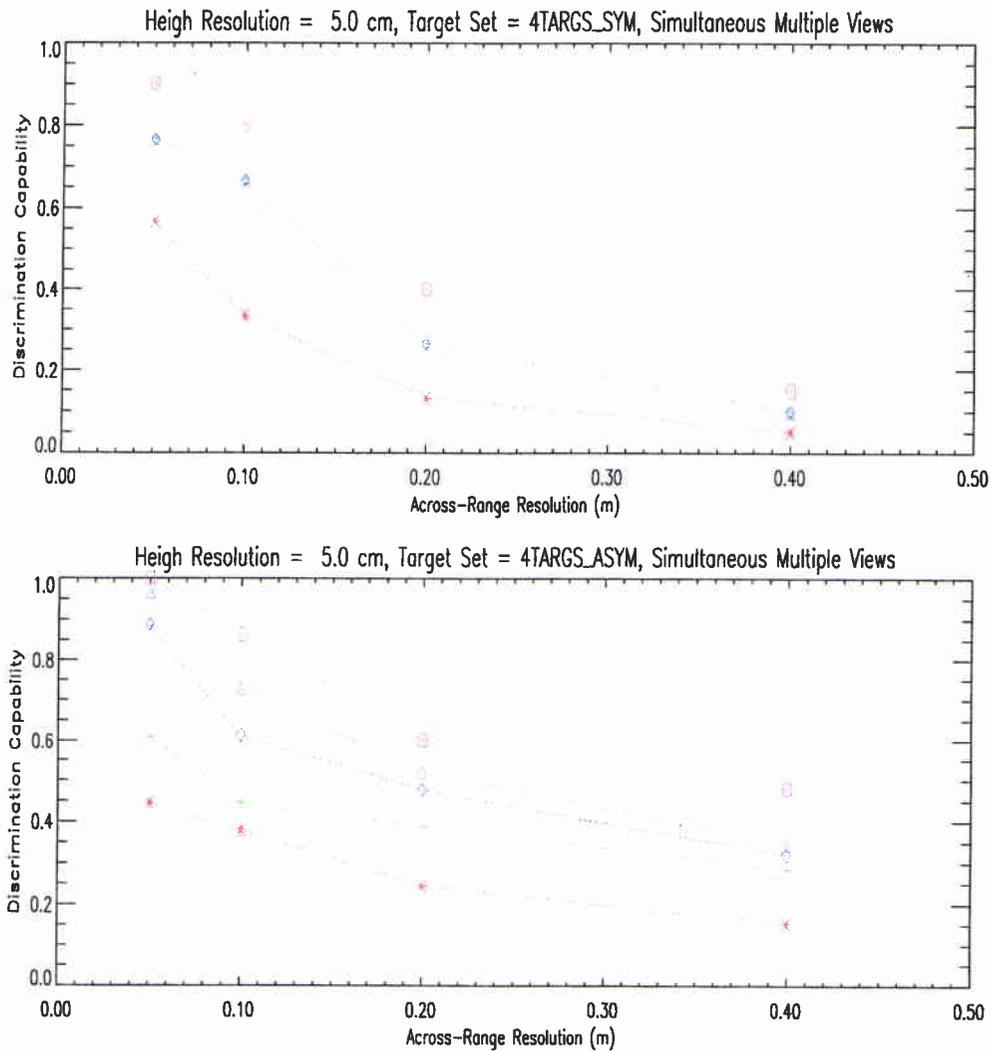


Figure 35: Discrimination Capability versus across-range resolution with simultaneous processing of the multiple views. Height resolution is 5 cm. Top and bottom plots show the results for a subset of 4 symmetrical shapes and a subset of 4 asymmetrical shapes, respectively. Multiple view configurations are marked as follows : Red \star : Single View ($[0^\circ]$). Green $+$: Double View ($[0^\circ, 5^\circ]$). Blue \diamond : Double View ($[0^\circ, 45^\circ]$). Cyan \triangle : Double View ($[0^\circ, 90^\circ]$). Magenta \square : Triple View ($[0^\circ, 45^\circ, 90^\circ]$). Orange \times : Triple View ($[0^\circ, 5^\circ, 10^\circ]$).

NATO UNCLASSIFIED

SACLANTCEN SM-309

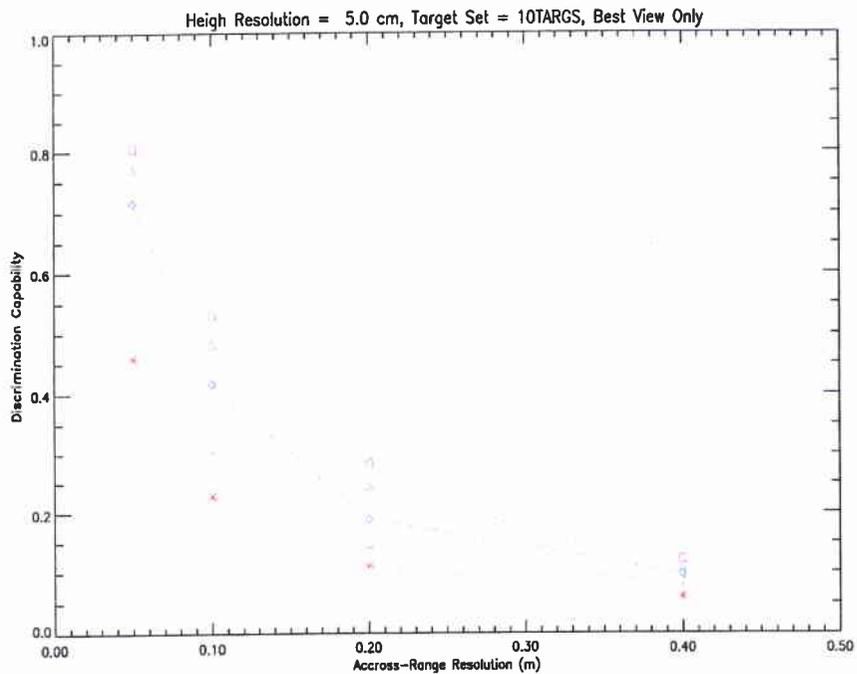


Figure 36: Discrimination Capability versus across-range resolution for the 10 target shapes. Height resolution is 5 cm and only the best view to the best view is used. Multiple view configurations are marked as follows : Red \star : Single View ($[0^\circ]$). Green $+$: Double View ($[0^\circ, 5^\circ]$). Blue \diamond : Double View ($[0^\circ, 45^\circ]$). Cyan \triangle : Double View ($[0^\circ, 90^\circ]$). Magenta \square : Triple View ($[0^\circ, 45^\circ, 90^\circ]$). Orange \times : Triple View ($[0^\circ, 5^\circ, 10^\circ]$).

NATO UNCLASSIFIED

These curves demonstrate that image resolution, irrespective of the mine shapes, plays a key role in classification performance. Multiple views significantly improve the capability to discriminate provided that the resolution is sufficient and the angular intervals between views exhibit the asymmetrical features of the targets. Another important conclusion is that the simultaneous use of the information gathered by the multiple views performs better than using only the best view.

As these curves depend on the definition of the targets composing the benchmark, no definitive conclusions can be drawn from these results. The fact that using three views with angular intervals of 45° always gives the best results does however permit the conclusion that this configuration is the most promising for an experimental system.

NATO UNCLASSIFIED

SACLANTCEN SM-309

6

Conclusions

This study, based on modelled images, demonstrates that the multiple view sidescan sonar significantly improves target classification performance, for a given rate of false alarms.

The performance is better when the information from multiple views is combined in the early stages of the classification process. Building a common vector with the feature vectors extracted from the multiple views produces the best results. Averaging the neural network outputs performs less well but can be used when one view is missing or when the target images cannot be recorded for *a priori* defined angular intervals between views. Combining the multiple views by applying a logical function to the classifier outputs has not given satisfying results. The angular intervals of 90° for two views and 45° for three views produce the best classification performance.

Similar conclusions are drawn from the parametric study on discrimination capability. The simultaneous processing of multiple views (equivalent to building a common feature vector) gives the best results. The parametric study demonstrates that the improvement is significant only when the sonar possesses sufficient resolution. In other words, using additional views does not compensate for poor spatial resolution.

Automatic target classification has been performed by neural networks. The radial basis functions (RBF) have performed better than the combination (MNK) of multiple layer perceptron and Kohonen self organizing feature map. ART2 neural network has proved to be an useful method to search for ambiguities in the supervised training data. Stable classification statistics have been attained by processing 12000 objects images. The model based approach has been extremely useful for the generation of these 12000 object images.

The recommendation for a future experimental assessment is a sidescan sonar which simultaneously records three views : the conventional broadside view and two additional views with programmable orientation from -45° to $+45^\circ$ to broadside. The across-range and along-range resolution must be better than 20 cm and 5 cm, respectively.

NATO UNCLASSIFIED

7

Acknowledgements

Many thanks go to René Laterveer for the careful and precise technical review of this document.

NATO UNCLASSIFIED

SACLANTCEN SM-309

References

- [1] Carson, L., Weddle, M.J.H., Horwitz, A., and McIntyre, L. Data processing for remote mine hunting. *Proceedings of the Undersea Defence Technology Conference*, 1994: pp. 446-449.
- [2] Bhanu, B. Automatic target recognition: State of the art survey. *IEEE Transactions on Aerospace and Electronic Systems*, **22**, 1986: pp. 364-379.
- [3] Brown, W.M. and Swonger, C.W. A prospectus for automatic target recognition. *IEEE Transactions on Aerospace and Electronic Systems*, **25**, 1989: pp. 401-410.
- [4] Augustyn, K. A new approach to automatic target recognition. *IEEE Transactions on Aerospace and Electronic Systems*, **28**, 1992: pp. 105-114.
- [5] McKinney, C.M. Acoustic sensor for minehunting. *Proceedings of the Autonomous Vehicles in Mine Counter Measure Symposium*, 1995: pp. 6.30-6.47.
- [6] Doherty, M.F., Landowski, J.G., Maynard, P.F., Uber, G.T., Fries, D.W., and Maltz, F.H. Sidescan sonar object classification algorithms. *Proceedings of the 6th International Symposium on Unmanned Unthethered Submersible Technology*, 1989: pp 417-424.
- [7] Schweizer, P.F. and Petlevich, W.J. Automatic target detection and cuing system dor an autonomous underwater vehicle (AUV). *Proceedings of the 6th International Symposium on Unmanned Unthethered Submersible Technology*, 1989: pp. 359-371.
- [8] Schweizer, P.F. and Petlevich, W.J. Procede pour detecter et classifier des particularites dans des images sonar. *US Patent 628116*, 1991.
- [9] Shazeer, D.J. and Bello, M.J. Minehunting with multi-layer perceptrons. *IEEE Conference on Neural Networks for Ocean Engineering*, 1991: pp. 57-68.
- [10] Shazeer, D.J. and Bello, M.J. Detection and classification of undersea objects using multilayer perceptrons. *SPIE Applications of Artificial Neural Networks II*, **1469**, 1991: pp. 622-636.
- [11] Smedley, L.D. and Dobeck, G. Improved automated detection and classification of hpss sonar targets with neural network supplement. *NSWC Report CSS/TR-94/34*, 1994.

NATO UNCLASSIFIED

- [12] Tuovila, S.M. and Nelson, S.R. Target classification using fractal based analysis of sonar images. *NSWC Report CSS/TR-94/37*, 1995.
- [13] Linnet, L.M., Clarke, S.J., Graham, C., and Langhorne, D.N. Remote sensing of the seabed using fractal techniques. *Electronics & Communication Engineering Journal*, October 1991: pp. 195-203.
- [14] Zerr, B. Automatic classification of objects in sonar images. *Proceedings of the Undersea Defence Technology Conference*, 1990: pp 137-142.
- [15] Delvigne, J.C. Shadow classification using neural networks. *Proceedings of the Undersea Defence Technology Conference*, 1991: pp. 214-220.
- [16] Zerr, B. Automatic classification en imagerie sonar - approche neuromimetique. *GESMA Report 2022*, 1992.
- [17] Johnson, S.G. and Deaett, M.A. The application of automated recognition techniques to side-scan sonar imagery. *IEEE Journal of Oceanic Engineering*, **19**, Jan 1994: pp. 138-144.
- [18] Stage, B. and Zerr, B. Multiple view sidescan sonar. *Proceedings of the Undersea Defence Technology Conference*, 1995, pp. 235-238.
- [19] Stage, B. and Zerr, B. Multiple view sidescan sonar. *SACLANTCEN SM-291*, 1995.
- [20] Zerr, B. and Stage, B. Three-dimensional reconstruction of underwater objects from a sequence of sonar images. *IEEE International Conference on Image Processing*, Lausanne 16-19 September, 1996.
- [21] Zerr, B. and Stage, B. A method for the three-dimensional reconstruction of underwater objects. *SACLANTCEN SM-310*, 1996.
- [22] Duda, R.O. and Hart, P.E. Pattern classification and scene analysis. *Wiley*, 1973: pp 10-39.
- [23] Lippman, R.P. An introduction to computing with neural nets. *IEEE Acoustic Speech and Signal Processing Magazine*, April 1987: pp 5-22.
- [24] Ng, K. and Lippman, R.P. A comparative study of the practical characteristics of neural network and conventional pattern classifiers. *MIT technical report TR-894*, March 1991.
- [25] Reynals, S. and Rohwer, R. Phoneme classification experiments using radial basis functions. *1989 IEEE Int. Joint Conference of Neural Networks*, 1989: pp I-461:I-467.
- [26] Cover, T.M. and Hart, P.E. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, **13**, January 1967: pp. 21-27.

NATO UNCLASSIFIED

SACLANTCEN SM-309

- [27] Kohonen, T. Self-organization and associative memory. *Springer Verlag, Berlin*, 1984.
- [28] Maillard, E., Zerr, B., and Merckle, J. Classification of texture by an association between a perceptron and a self organized feature map. *EUSIPCO, Signal Processing IV*, 1992: pp 1173-1176.
- [29] Carpenter G., A. and Grossberg, S. Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, **26**, December 1987: pp. 4919-4930.
- [30] Best, J.P. Exploitation of the aspect dependence of mine target strength in mine hunting search tactics. *Maritime Operations Division (Melbourne) Report DSTO-RR-0051*.
- [31] Zerr, B. Topographic sidescan sonar 1 TSS1. *SACLANTCEN Cruise Report*, September 1996.

NATO UNCLASSIFIED

Annex A

Simulation

A.1 Overview

The main task of the simulator is to accurately model the shadow cast by minelike objects on the seabed. The shadow information in a sonar image depends on the sonar characteristics, the geometry of the objects, the seabed topography and the viewing conditions. After defining the characteristics of the sonar, the simulator proceeds in three phases. The first phase consists of a three-dimensional definition of the objects shapes and the seabed relief. The result is a data base containing geometrically defined objects, natural objects and seabed surface elements. In the second phase these items are selectively extracted from the data base and combined to build an underwater scene¹⁵. The selection of the items is based on the coordinates of the track followed by the sonar. Finally, the third phase is the simulation of the sonar image of the scene.

For the sake of homogeneity, the simulator uses an identical three-dimensional description for objects and seabeds. The acoustic ray tracing technique has been chosen for its capability to accurately produce the boundaries of acoustic shadows. Since the main focus is on shadow generation, the backscattering model has been simplified. This section describes the key features of the simulator.

A.2 Three-dimensional Objects Model

The three-dimensional description model divides the object surface into polygonal facets. Each object is described by two tables : the first dealing with the vertices coordinates and the second with the polygons defined by these vertices. Figure A-1 shows how such a model represents the elements of a three-dimensional surface. First, the vertices table defines, for each vertex, the x,y and z coordinates. Then, each polygon in the polygon table is defined by a number of vertices and a list of identifiers pointing to the vertices definition table. For a triangle based description, the number of vertices is of little interest, but it is kept both to conform with the

¹⁵For the current report, the underwater scene, described in a specific language, can only be processed by the high frequency sonar simulator. However, a translation to VRML (Virtual Reality Modeling Language) standard has been initiated to "open" these virtual landscapes to other applications (AUV simulator, optical underwater imaging, ...)

NATO UNCLASSIFIED

SACLANTCEN SM-309

IDL three-dimensional description scheme and also for future extensions. For each triangle, the normal vector is computed by taking the vectorial product between units vectors of two edges. The direction of the normal is set to look outside the object according to an inside indicator. On simple convex shapes, the center of gravity is used as the inside indicator.

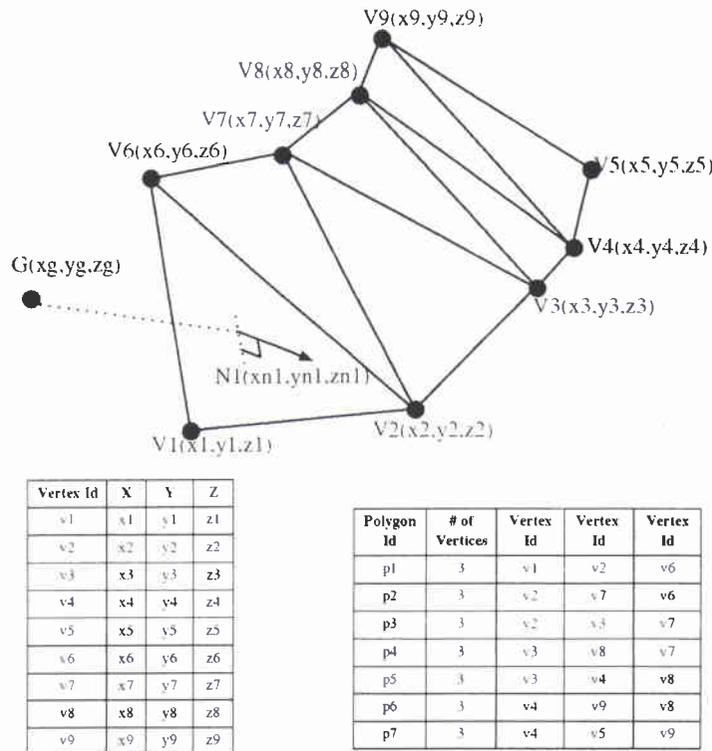


Figure A-1: Generic Three Dimensional Object Description

More complex shapes result from the combination of simple ones by applying the CSG (Constructive Solid Geometry) union operation¹⁶. In this case, each simple shape used in the definition of the complex object has its own inside indicator. For height fields, the inside indicators are a set of points, spread under the field surface. Some additional characteristics, such as the acoustical properties of the surface, may be added to the polygon table to refine the description of the three-dimensional object. All the generic parameters and algorithms which define objects are stored in a database. For a polygonal object (Fig. A-2), the vertices and polygons descriptions are shown in Fig. A-3. The simulator can also import shapes defined externally, by a computer aided design tool. Natural objects are derived from simple shapes using fractal based process. An example of how to create a stone from a polyhedron is illustrated in Fig. A-4.

The seabeds are defined using small square elements, or patches, from 5 m² to

¹⁶Glassner A.S., An Introduction to ray Tracing, Academic Press, 1989

NATO UNCLASSIFIED

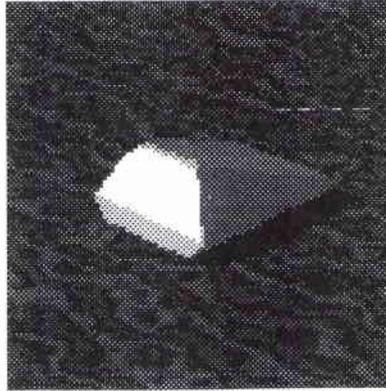


Figure A-2: Polygonal Object (ROCL)

200 m², depending on their nature. For each of these elements, the height at the boundaries is constrained to obey a wrap around criteria. Hence, defining a large area of the same nature by the juxtaposition of small elements will not produce height jumps at the boundaries. Various kind of synthesis may be used to define patches with boundaries satisfying the wrap around criteria. A method which implicitly verifies this constraint is the Fourier synthesis. Such a method consists of “editing” the Fourier plane by defining the amplitude and the phase for each location.

A fractal algorithm is used to define seabed elements of controlled roughness. Peitgen and Saupe¹⁷ state that a surface of fractal dimension, D_f , is obtained if, in the Fourier plane, the phase is uniformly distributed and the amplitude $S(u, v)$ follows

$$S(u, v) = \alpha * \frac{1}{(u^2 + v^2)^{H+1}} \quad (1)$$

where u and v define the Fourier plane, α is a random coefficient with Gaussian distribution and H is related to the fractal dimension of the surface by

$$D_f = 3 - H \quad (2)$$

Figure A-5 illustrates seabeds obtained by varying the fractal dimension (D_f) from 2.2 to 2.8. Combining two or more height fields allows the construction of composite seabeds. For example, a rock and sand seabed can be defined as the maximum of two height fields : a fractal one and a flat one.

¹⁷Peitgen H.O. and Saupe D., The science of fractal images. Springer Verlag, New-York, 1998

NATO UNCLASSIFIED

SACLANTCEN SM-309

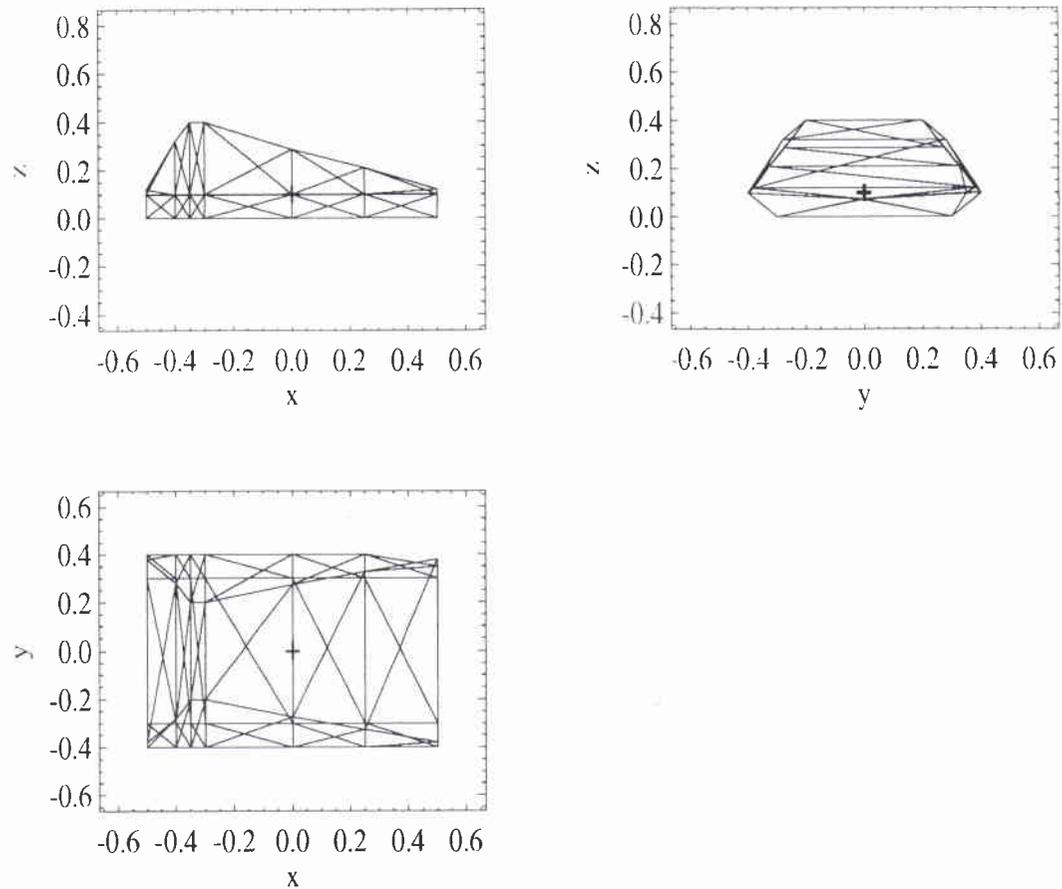


Figure A-3: Vertices and Polygon Description

Another important property of the synthesis in the Fourier plane is the inherent capability to address periodic phenomena. Perfect sinusoidal ripples are simulated by placing a Dirac pulse in the Fourier plane. Simply altering the Fourier plane, in the vicinity of the perfect ripples peak, leads to surprisingly realistic ripples fields (see for example the simulated sonar image in Fig. A-6).

A.3 Sonar Images Simulation

A simple and generic model for the three-dimensional definition of objects reduces the complexity of ray tracing algorithms. With the current three-dimensional object definition, only one intersection algorithm between lines and triangles is required. The realism of sonar simulation depends on the accuracy of the sonar model. For the current study, the main requirement is to produce precisely the theoretical shad-

NATO UNCLASSIFIED

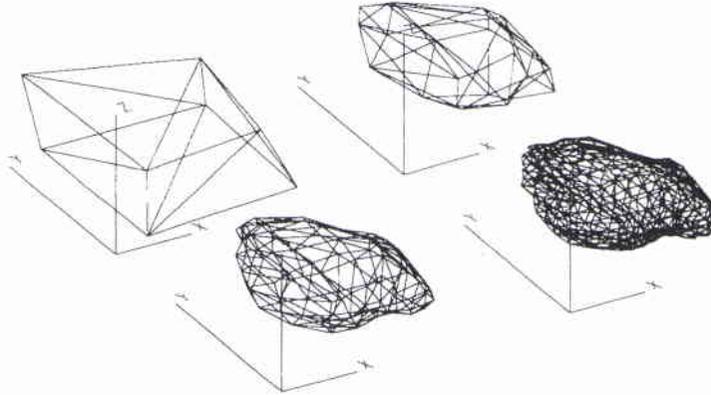


Figure A-4: Iterative Construction of a Fractal Stone

ows cast by the objects under different viewing conditions. This is carried out by simulating precisely the seabed and the objects lying on it and by defining a sonar model dedicated to shadow production.

The sonar simulation technique implemented here is close to optical ray tracing. The intersection between acoustic rays sent by the sonar and three-dimensional objects is tested. Secondary rays may be then created from the hit point using classical laws for reflection and refraction. Assuming a Lambertian reflectance function, the intensity I_r , scattered back by a surface element of unit area, is approximated by

$$I_r = \mu I_i \cos \Psi$$

where:

$$\Psi = \vec{n} \cdot \vec{i} \quad (3)$$

I_i is the intensity of the incoming ray and the constant μ , varying with the type of seabed, controls the backscattering strength. \vec{n} is the unit vector normal to the surface element and \vec{i} the unit vector in the direction of the acoustic incident ray. The simulator computes high resolution beams with an ideal beam pattern (i.e. without introducing the side lobes effect). The resolution is degraded, if required, by summing high resolution beam.

The capabilities of the simulation tool are shown in the following examples. Figure A-6 shows the sonar images of an underwater landscape containing an aircraft wreck, a truncated cone, a wedding cake and some stones. The seabed type is

NATO UNCLASSIFIED

SACLANTCEN SM-309

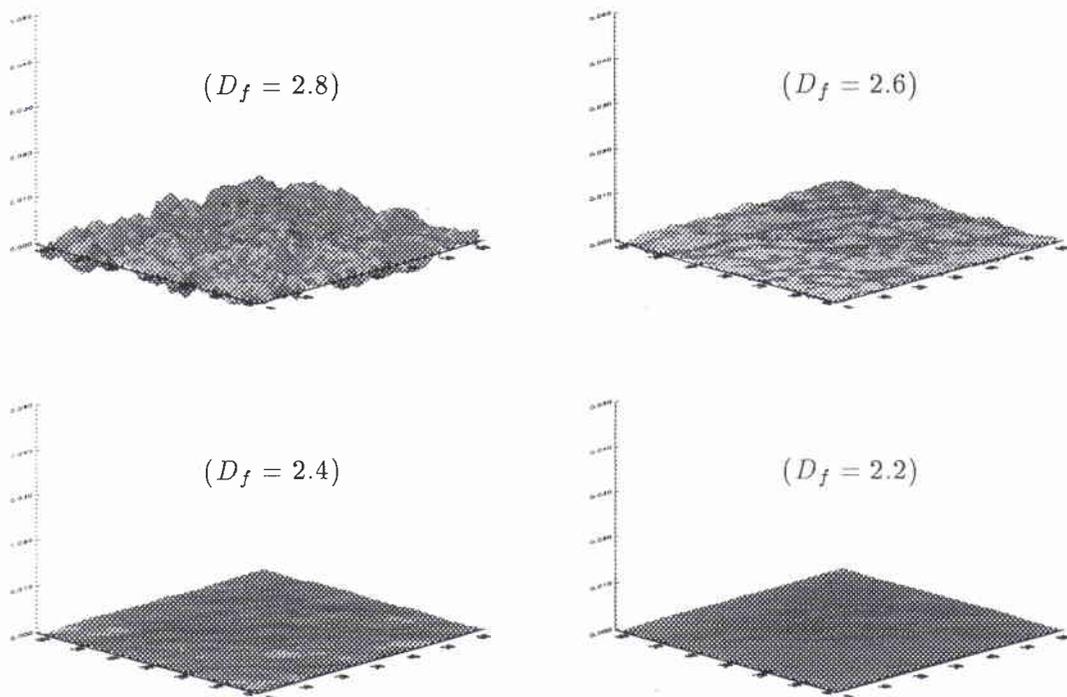


Figure A-5: Fractal Seabeds of Varying Roughness. D_f (Fractal Dimension) = 2.8, 2.6, 2.4 and 2.2.

sand ripples. Figure A-6(a) represents the conventional sidescan image of the scene, while Fig. A-6(b) highlights the interest of an additional view when investigating the wedding-cake shadow. The images are projected from slant range to a square grid on the seabed.

The same region of the seabed, imaged by a sector scanning forward looking sonar, is displayed in Fig. A-7. The six images of this figure correspond to six different sonar bearings : 0° , 30° , 60° , 90° , 120° and 150° in a left to right and top to bottom order. In Fig. A-8, the viewing conditions are identical to those of Fig. A-7, but the resolution of the sector scanning sonar is lower.

In addition to the polygonal shape of Fig. A-2, nine shapes, corresponding to proud targets and to sinker of moored targets, are used in the classification benchmark. Figure A-9 shows these object shapes rendered by optical ray tracing. The dimensions of these objects are defined in Table 2 (Section 4).

NATO UNCLASSIFIED

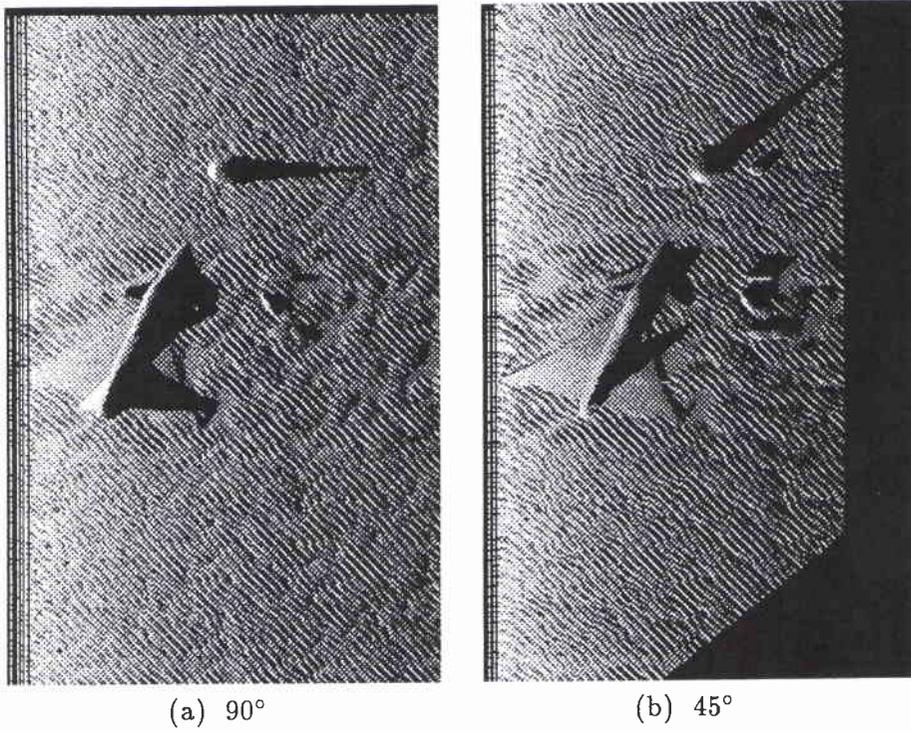


Figure A-6: Two Look Sidescan Sonar Simulation

NATO UNCLASSIFIED

SACLANTCEN SM-309

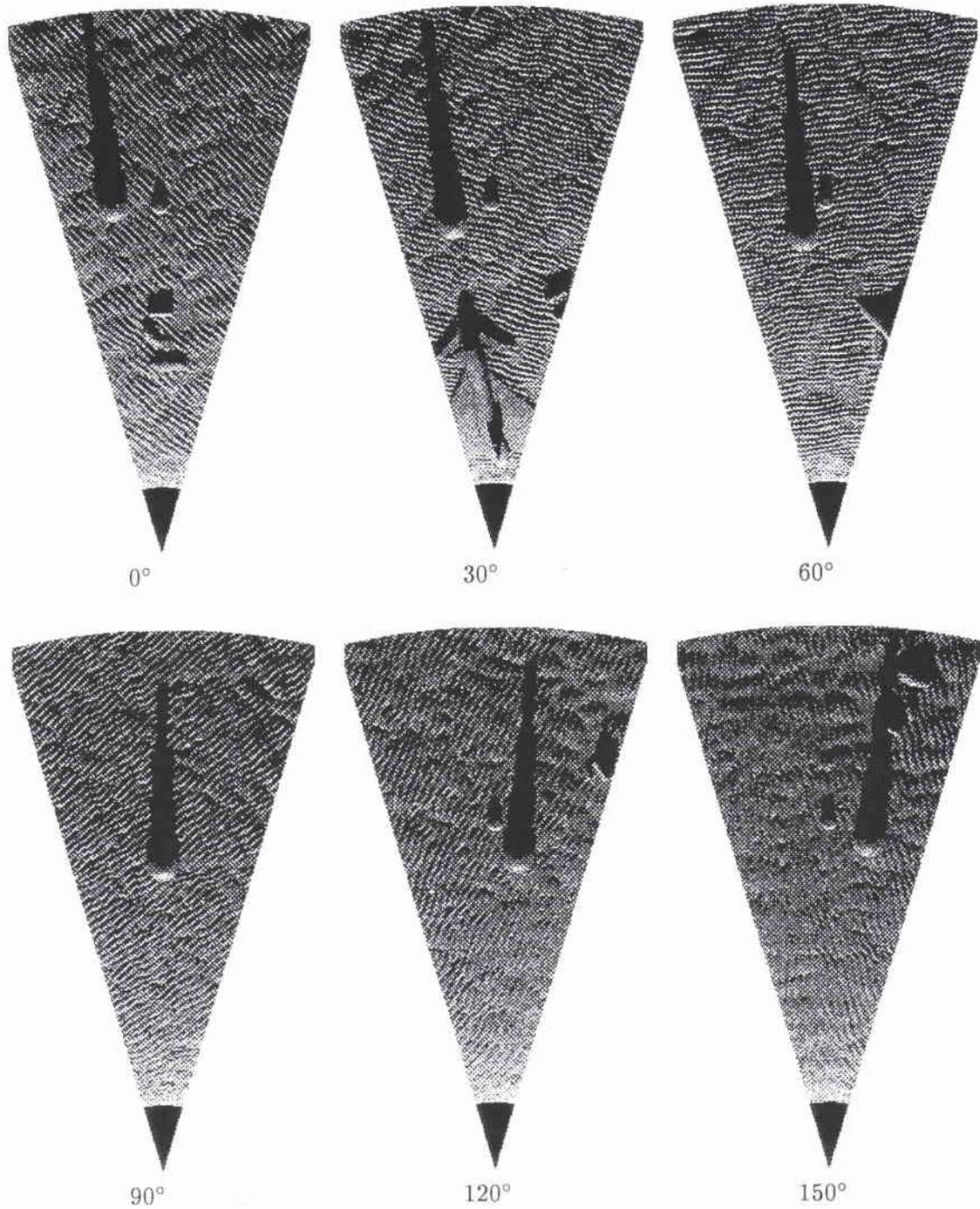


Figure A-7: Simulation of High Resolution Sector-scan Sonar Images

NATO UNCLASSIFIED

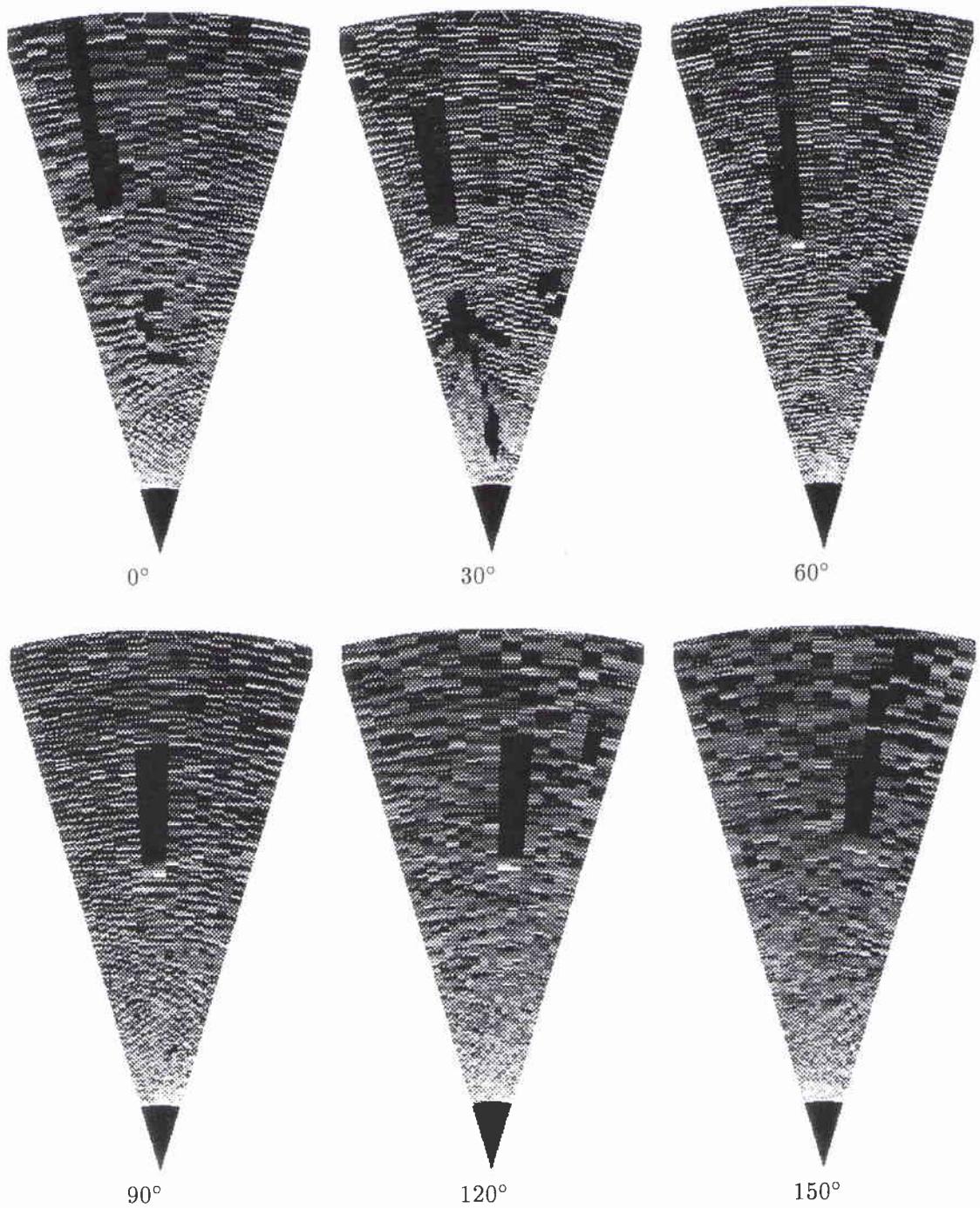
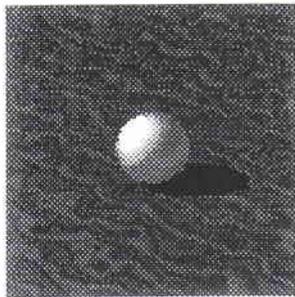


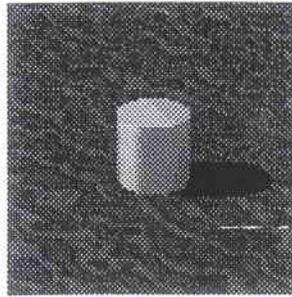
Figure A-8: Simulation of Lower Resolution Sectorscan Sonar Images

NATO UNCLASSIFIED

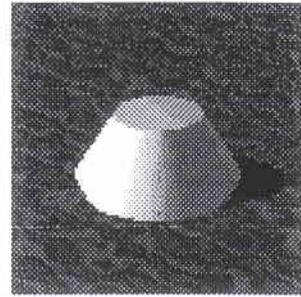
SACLANTCEN SM-309



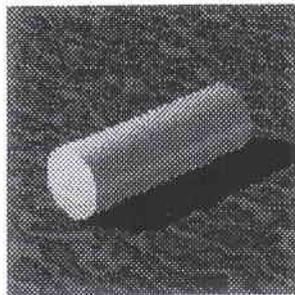
SPHR



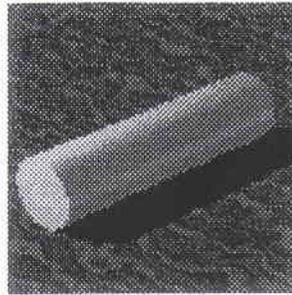
VCYL



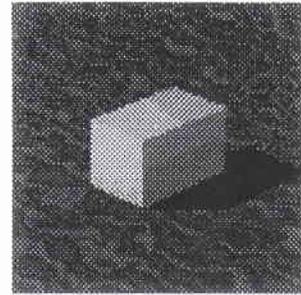
MNLK



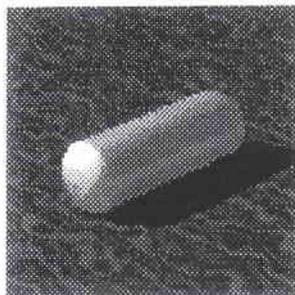
HCYB



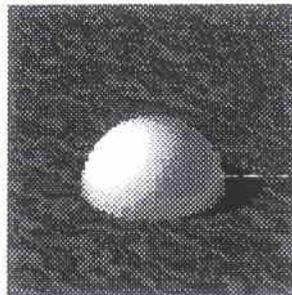
HCYC



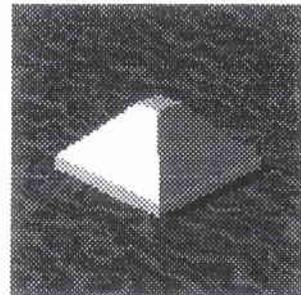
HPPD



HCSB



HEMS



PYRS

Figure A-9: Target Shapes for Classification Benchmark

NATO UNCLASSIFIED

*Document Data Sheet***NATO UNCLASSIFIED**

<i>Security Classification</i>		<i>Project No.</i>
NATO UNCLASSIFIED		031-3
<i>Document Serial No.</i>	<i>Date of Issue</i>	<i>Total Pages</i>
SM-309	September 1997	84 pp.
<i>Author(s)</i>		
B. Zerr, B. Stage, A. Guerrero		
<i>Title</i>		
Automatic target classification using multiple sidescan sonar images of different orientations.		
<i>Abstract</i>		
<p>In this report, the target classification performance of a multiple view sidescan sonar is investigated.</p> <p>The classification statistics are estimated using model based automatic classifiers. The guidelines to the design of efficient classification algorithms are defined. The shadow is retained as the basic information for target classification. The input feature vector of the automatic classifier is the cross-section (or height profile) of the target estimated from its shadow.</p> <p>The concept of multiple view sidescan sonar is presented and compared to other techniques for recording multiple aspects of a target. Several ways to modify a single view based classifier to process multiple aspects are identified and implemented. The task of the classifier is to recognize 10 target shapes, corresponding to proud mines and sinkers of moored mines. The classification results, expressed by ROC curves and confusion matrixes, are computed on a larger number of natural and manufactured object images, generated by modelling software.</p> <p>The classification and identification of targets are closely related to the capacity of the feature vectors to discriminate between a given target, viewed with a given orientation, from all the other targets, irrespective of orientation. Using height profiles as feature vectors, the capacity to discriminate is established for a single view on target. Different configurations using up to three views with different angular intervals are subsequently compared.</p>		
<i>Keywords</i>		
Shadow based mine classification – automatic target classification – sidescan sonar – neural networks		
<i>Issuing Organization</i>		
North Atlantic Treaty Organization SACLANT Undersea Research Centre Viale San Bartolomeo 400, 19138 La Spezia, Italy		Tel: +39 (0)187 540 111 Fax: +39 (0)187 524 600 E-mail: library@saclantc.nato.int
<i>{From N. America: SACLANTCEN (New York) APO AE 09613}</i>		

NATO UNCLASSIFIED