Report no. changed (Mar 2006): SM-294-UU

SACLANTCEN MEMORANDUM
serial no.: SM-294

SACLANT UNDERSEA RESEARCH CENTRE MEMORANDUM

INTERPRETATION AND FUSION OF SONAR IMAGES USING EVIDENTIAL REASONING

B. Stage

December 1995

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.

Report no. changed (Mar 2006): SM-294-UU

SACLANT Undersea Research Centre
Viale San Bartolomeo 400
19138 San Bartolomeo (SP), Italy

tel: +39-187-540.111
fax: +39-187-524.600

e-mail: library@saclantc.nato.int

NORTH ATLANTIC TREATY ORGANIZATION

Report no. changed (Mar 2006): SM-294-UU

Interpretation and fusion of sonar images using evidential reasoning

B. Stage

The content of this document pertains to work performed under Project 25 of the SACLANTCEN Programme of Work. The document has been approved for release by The Director, SACLANTCEN.

Director

Report no. changed (Mar 2006): SM-294-UU

Interpretation and fusion of sonar images using evidential reasoning

B. Stage

Executive Summary: In mine warfare classification of bottom objects is essential. It is also one of the major problems. Especially when there is considerable sea–bottom clutter as in many MCM situations. This report presents a novel approach to automatic interpretation of sonar images of the sea–bottom, which allows combination of information from multiple images of the same site. The method provides increased reliability in the interpretation compared to existing methods, as information from several images is utilized. Additionally, a distinction is made between uncertainty and ignorance. This distinction is relevant as imprecision in an interpretation of a sonar image can be due to both ambiguity and to lack of knowledge. The viability of the technique is demonstrated by combining sonar images acquired with an experimental sonar with two simultaneous views.

Interpretation and fusion of sonar images using evidential reasoning

B. Stage

Abstract: A method and system architecture for automatic interpretation and fusion of multiple sonar images is proposed. The interpretation of a sonar image consists of the creation of a symbolic geographical map indicating the belief in the presence of objects and structures on the sea–bottom. Fusion of information from multiple images of the same site is at the symbolic level. Both interpretation and fusion are based on the Dempster–Shafer theory of evidence. The system can be used for interpretation of images containing man–made as well as natural objects irrespective of model type. The method is applied to fusion of information from images acquired with an experimental sonar system with two simultaneous views. Keywords: Sonar images, interpretation, fusion.

Contents

1 Introduction

Sidescan sonars are used routinely for surveillance and reconnaissance operations over large areas of the sea–bottom with the purpose of locating man–made objects and natural structures of interest. Interpretation of sonar images has traditionally been done by operators, but systems which use automatic processing are evolving. This report describes a method and a system architecture for automatic image interpretation and fusion of information from multiple sonar images of the same site.

A variety of methods has been used for automatic processing of sonar images. Highlights and shadows in sidescan sonar images were used to infer the geometric shape and reflectivity of objects in order to detect and classify bottom mines [1]. A composite method using both shadows and highlights, a histogram based technique and a neural network operating directly on the image data were investigated in connection with an autonomous underwater vehicle for minehunting [2,3]. Waste deposits are another example of man–made objects that have been recognized [4]. This was accomplished without detailed knowledge of the objects' geometry using features derived from the shape of shadows and highlights.

Sea–bottoms can be classified using spatial grey level features of sidescan images [5]. Spectral and cepstral features derived from the signal envelope and moments of the probability distribution of the signal envelope can also be used [6]. Based on these features, classification was accomplished using minimum distance [5] and neural network classifiers [7].

The binary patterns in bitplanes of the images of natural textures can be represented as spatial point processes and used to classify natural textures. A measure of probability can be associated with the hypothesis that a small region in an image belongs to a textured background [8]. In addition to classifying textures the method can be used to detect objects on a textured background. An extension of the method explicitly takes shadow and highlight regions into account to create a map of the probability of the presence of an object [9].

Another probabilistic approach to sonar image interpretation is a rule based system that assigns confidence to the result of the interpretations [10]. The interpretation is based on shape features of regions formed by segmenting the sonar image into background, highlight and shadow regions. In an extension of this work, matching between observation and a stored model is made using coarse characterization of the regions [11].

One system for real–time interpretation is based on extraction of line segments and vertices from a sonar image [12]. Confidence in a match between image features and features derived from an object model is assessed using the Dempster–Shafer theory of evidence.

When multiple sonar images of the same site are available, it is desirable to merge the information in the images. A composite image can be computed as the weighted average of the pixel values in the original sidescan sonar images [13]. The weights are adjusted to emphasize the better quality image. Changes in the visibility of the objects' surfaces and shadow orientation due to different positions of the sonar during acquisition of the original images are disregarded.

The problem considered here is automatic sidescan sonar image interpretation. The applications are creation of sea–bottom data bases during reconnaissance operations in unknown areas and comparison with existing databases during surveillance operations. The capability of current automatic sonar interpretation systems for recognition of man–made objects can be improved in a number of ways. Some of these are increased image quality, the use of better models, the use of multiple images of the same site and the utilization of information on natural objects and structures on the sea–bottom.

The image quality, i.e., the resolution and contrast of an image, is crucial to the success of the image interpretation process whether it is performed automatically or by an operator. High image quality is a prerequisite for the application of improved models for recognition as these models require more detail. Multiple images of the same site taken from different sonar positions increase recognition performance as the chance that an object is recognized in at least one image is increased [14]. By focusing image interpretation only on areas where man–made objects are detected, important information is discarded. If areas containing exclusively natural structures are interpreted and positively recognized as such, the areas where man–made objects could possibly be located are reduced. This will be especially useful in situations where man–made objects are camouflaged either intentionally or by natural processes. An operator examining an image can utilize his knowledge of natural objects to reject an object as being say, a cylindrical man–made object if it has the characteristics of a stone. The performance of an automatic recognition process can therefore approach the performance of an operator only if information on natural objects and structures is used by the system.

Sidescan sonar images are ambiguous two–dimensional representations of a three–dimensional scene. As resolution and contrast are limited, the chance of misinterpretation exists. It is therefore necessary not only to make an interpretation but also to assign some measure of confidence to the interpretation.

This report describes a system concept for automatic interpretation and fusion of multiple sonar images. Interpretation of a sonar image consists of the creation of a symbolic geographical map indicating the belief in the presence of objects and structures on the sea–bottom. Fusion of information from multiple images of the same site is performed at the symbolic level. Both interpretation and fusion are based on the Dempster–Shafer theory of evidence. The system can be used for interpretation of images containing man–made as well as natural objects. No constraints on the types of models that can be used for interpretation or the levels of abstraction of the interpretation are imposed.

Section 2 gives a brief introduction to the Dempster-Shafer theory. In Sect. 3 a framework for interpretation of sonar images is discussed, followed by a description of the system architecture in Sect. 4. Section 5 contains a discussion of the principles for assigning belief to a hypothesis concerning the presence of an object on the sea–bottom based upon evidence provided by a sonar image. The combination of information from multiple images is considered in Sect. 6. The method is demonstrated by combining images from an experimental sonar with two simultaneous views.

2

Dempster–Shafer theory

This section presents a brief overview of elements of the Dempster-Shafer theory of evidence. The presentation and notation is based largely on Shafer [15] and Guan and Bell [16]. The Dempster–Shafer theory of evidence is a mathematical formalism for assigning belief to hypotheses relevant to a problem domain and for combining the belief in a consistent way when the belief is due to evidence from independent sources. The Dempster–Shafer theory distinguishes between uncertainty and ignorance. This distinction is relevant as imprecision in an interpretation of a sonar image can be due to both ambiguity and to lack of knowledge.

Consider an observed phenomenon such as a feature in an sonar image for which an explanation is sought. Suppose the cause of the phenomenon is known to be due to one of a finite number of known causes. This knowledge can be used to form a set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ called the frame of discernment. The elements of Θ represent n mutually exclusive and exhaustive hypotheses as to the cause. In contrast to a probability distribution, which assigns belief to the elements of the frame of discernment, the Dempster–Shafer theory assigns belief to all possible subsets of the frame of discernment. The set of all possible subsets of Θ including the empty set $\emptyset$ is the ***power set*** of Θ and is denoted 2^Θ. Given some evidence such as a specific pattern embedded in an image, it is assumed that the total belief due to the evidence can be divided into portions and assigned to various subsets A of Θ. A ***basic probability assignment*** (BPA), or ***mass function***, is a mapping from subsets of the frame of discernment to the unit interval $m : 2^\Theta \to [0, 1]$ satisfying

$$m(\emptyset) = 0 \tag{1}$$

$$\sum_{A \subseteq \Theta} m(A) = 1. \tag{2}$$

The ***basic probability value*** or simply ***mass value***, $m(A)$, represents the portion of belief assigned exactly to the hypothesis represented by the subset A of Θ and not to any subsets B of A. Ignorance is represented by assigning belief to the frame of discernment. Total ignorance is represented by $m(\Theta) = 1$. As belief can be assigned to any subset of Θ including conflicting subsets, uncertainty and ignorance can be distinguished. In addition, information at different levels of abstraction can be represented.

The extent to which the evidence specifically supports belief in a hypothesis represented by the set A is given by the ***belief function*** which is the sum of all basic probability values $m(B)$ in all subsets B of A:

$$Bel(A) = \sum_{B \subseteq A} m(B). \tag{3}$$

The subsets A of Θ characterized by $m(A) > 0$ are the ***focal elements*** of the belief function. Some belief functions have special characteristics. The belief function corresponding to total ignorance is the ***vacuous*** belief function which has Θ as its only focal element. A ***simple support function*** is a belief function which has only a single focal element besides Θ. In a ***consonant support function*** the focal elements A_i, $i = 1 \ldots n$, are nested so the focal elements can be ordered such that each is a proper subset of the following:

$$A_1 \subset A_2 \subset \ldots \subset A_n. \tag{4}$$

This characteristic makes consonant belief functions well suited for representing inferential evidence.

The ***plausibility*** $Pls(A)$ of a hypothesis represented by the subset A is the maximum extent to which the evidence can support the hypothesis.

$$Pls(A) = 1 - Bel(\bar{A}), \tag{5}$$

where $\bar{A}$ is the complement of A. Together $Bel(A)$ and $Pls(A)$ form the ***belief interval*** of A

$$Ivl(A) = [Bel(A), Pls(A)]. \tag{6}$$

The belief interval is a convenient way of representing that which is known or unknown. The following examples illustrate the interpretation of the belief interval [17]:

$$
\begin{array}{lcl}
Ivl(A) = [0.00, 1.00] & \Rightarrow & \text{No knowledge about } A. \\
Ivl(A) = [1.00, 1.00] & \Rightarrow & A \text{ is true.} \\
Ivl(A) = [0.00, 0.00] & \Rightarrow & A \text{ is false.} \\
Ivl(A) = [0.25, 1.00] & \Rightarrow & \text{Partial support is provided for } A. \\
Ivl(A) = [0.00, 0.75] & \Rightarrow & \text{Partial support is provided for } \bar{A}. \\
Ivl(A) = [0.25, 0.75] & \Rightarrow & \text{Support is provided for both } A \text{ and } \bar{A}.
\end{array}
$$

The residual ***ignorance*** is a measure of the lack of knowledge associated with a hypothesis

$$Ign(A) = Pls(A) - Bel(A). \tag{7}$$

In the case of the vacuous belief function, the ignorance $Ign(A)$ is equal to 1 for all A.

Belief functions with the same frame of discernment from independent sources of information can be combined provided they are not totally contradictory. The *orthogonal sum* m of two mass functions m_1 and m_2 is denoted by $m = m_1 \oplus m_2$, and is given by *Dempster's rule*:

$$m(\emptyset) = 0, \tag{8}$$

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)}, \tag{9}$$

provided that

$$\sum_{B \cap C = \emptyset} m_1(B) m_2(C) < 1. \tag{10}$$

The order in which mass functions are combined is irrelevant because the orthogonal sum is commutative, $m_1 \oplus m_2 = m_2 \oplus m_1$, and associative, $m_1 \oplus (m_2 \oplus m_3) = (m_1 \oplus m_2) \oplus m_3$. Given two belief functions, m_1 and m_2, if m_1 is vacuous then $m_1 \oplus m_2 = m_2$.

From this brief overview it can be seen that the Dempster–Shafer theory of evidence is useful for representing and combining incomplete and uncertain information. This makes the theory suitable as a framework for interpretation and fusion of sonar images.

3 The frame of discernment

A principal requirement for the application of the Dempster–Shafer theory is the construction of the frame of discernment. In this section a frame of discernment, useful for the interpretation of sonar images of the sea–bottom, is discussed. The interpretation of a sonar image refers to objects and structures located on the sea–bottom and not to patterns in the image. It is assumed that one and only one type of object or structure is present in a map projection of each sea–bottom location. In this case the elements of the frame of discernment are naturally chosen as the possible objects and structures that might be encountered on the sea–bottom. The frame of discernment must be exhaustive for the problem domain, and its elements must be mutually exclusive. A frame of discernment fulfilling these requirements and emphasizing the distinction between man–made and natural objects is given by

$$\Theta = \{NOS, NAT1, NAT2, \ldots, NATN, MMO1, MMO2, \ldots, MMOM, OU, ART\} \ .$$

The names of elements in the frame of discernment and their meaning are explained below.

NOS : No structure. This characterizes an area of the sea–bottom which does not contain objects or structures at a scale of any interest for the application. Featureless sea–bottoms have traditionally been characterized by their backscattering strength, and for some applications it is desirable to make a further division into sea–bottom types such as mud, sand and gravel. As the interest in the present context is focused on natural and man–made objects or structures, the natural sea–bottom without structure is conveniently treated as a unique element.

$NAT1\,NAT2, \ldots, NATN$: Natural structure. This represents a set of N known natural structures and objects. A distinction is often made between background structures and objects. Background structures are extended, essentially flat sea–bottoms with some form of relief such as sand ripples. Objects are entities protruding from the sea–bottom. In the general case, topographic variation extends over a wide range of scales and the distinction between background and object becomes less useful. *A priori* knowledge of natural structures is inevitably statistical due to natural variation. The local properties that characterize natural structures are shape and backscattering properties. In some cases contextual information describing the spatial ordering of objects and structures is available.

$MMO1, MMO2, \ldots, MMOM$: Man–made objects. This is a set of M known man–made objects. The knowledge can be shape, dimensions, surface properties as well as sharpness of edges and corners.

OU : Object, unknown. Unknown objects and structures must be included in the frame of discernment to make it exhaustive.

ART : Artifact. As stated previously, the elements in the frame of discernment should refer to objects and structures on the sea–bottom. Unfortunately, highlights can be present in sonar images which do not originate from the location of the sea–bottom as it appears in an image. These can be reflections from the sea surface, objects in the water volume or multiple reflections from strongly reflecting objects on the sea–bottom. Since a highlight can have different interpretations, including that of artifact, it is necessary for the sake of exhaustiveness to include artifacts in the frame of discernment.

The frame of discernment described here is a generally applicable basis for interpreting sonar images. To suit a particular purpose the frame can be extended by adding elements corresponding to specific natural or man–made objects.

4

Assigning belief

The purpose of image interpretation in this context is to create a symbolic representation of the sea–bottom in the form of a basic probability assignment map (BPA map). Each point in the BPA map is associated with a mass function representing the knowledge of the location. The mass functions are determined from evidence in the sonar images. The way in which evidence in an image is used to create the mass function determines the architecture of the interpretation system. In this section the traditional method for assigning belief is discussed, and a new method is proposed which leads to a simple modular architecture.

Rule–based expert systems are traditionally used to assign belief in applications of the Dempster–Shafer theory. The first application of the theory was a rule–based system in the field of artificial intelligence [17]. The same approach was taken in an application of the theory to interpretation of sonar images [12]. Formal methods for rule–based assignment of belief have been described [16].

The concept of a rule–based system for assignment of belief is illustrated in Fig. 1. A preselected number of features f_i, $i = 1, \ldots, m$, such as highlights, shadows and edges are extracted from the image by feature extraction modules F_i. Each feature type is examined independently by knowledge sources R_i, and a set of rules is used to assign belief to subsets of Θ conditioned on the evidence. Assuming independence of the features, the composite belief can be determined by combining the mass functions corresponding to each feature by Dempster's rule.

Although intuitively appealing, some objections can be raised to this procedure. Since the system operates on a set of preselected features, all methods of recognition are restricted to these features. This constrains the type of models which can be applied. Another problem is that knowledge about a particular object is distributed throughout the system. This makes maintaining and extending the system a difficult task. These problems have led to consideration of an alternative method for assigning belief.

Consider the system architecture shown in Fig. 2. A number of object recognition modules M_i, $i = 1, \ldots, n-1$, where n is the number of elements in the frame of discernment Θ, are used to examine the sonar image. Each module is restricted to knowledge about a single object or structure and can only confirm or disconfirm [18]

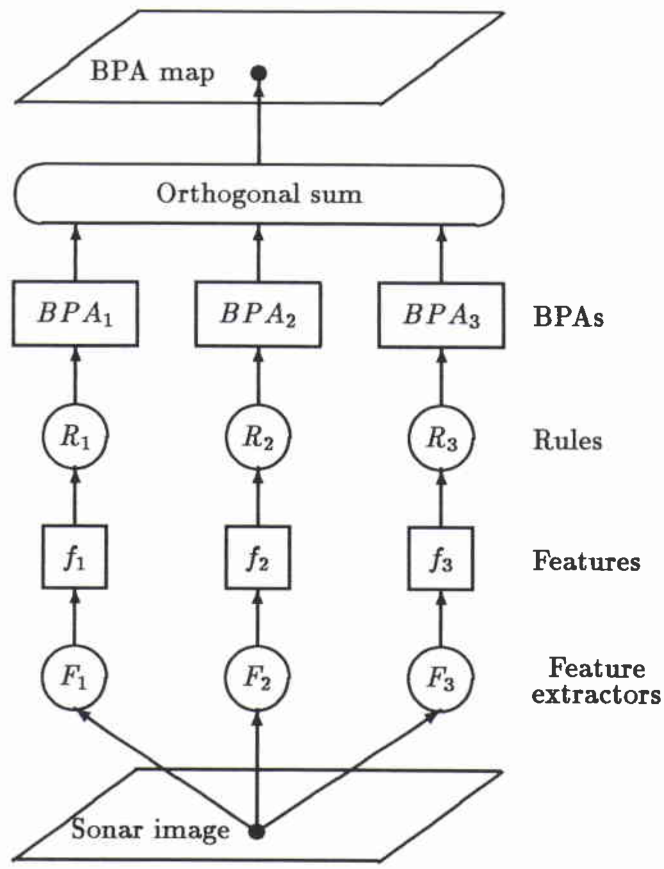

Figure 1: Rule–based image interpretation system architecture for creating a basic probability assignment map (BPA map).

whether the evidence in the image supports a hypothesis concerning the presence of the object. There are only $n-1$ modules because the element $\{OU\}$ corresponding to an unknown object is left unassigned.

The result of the examination of the evidence by a module M_i is a simple support function with degree of support s_i

$$m(A_i) = s_i \tag{11}$$

and

$$m(\Theta) = (1 - s_i), \tag{12}$$

where the subset A_i corresponds the confirmation or disconfirmation of the hypothesis associated with the module M_i.

An ideal system would assign support $s_i = 1$ to one singleton $\{\theta_i\}$ and disconfirm all other hypotheses, however, in practice some confusion will be present, and support

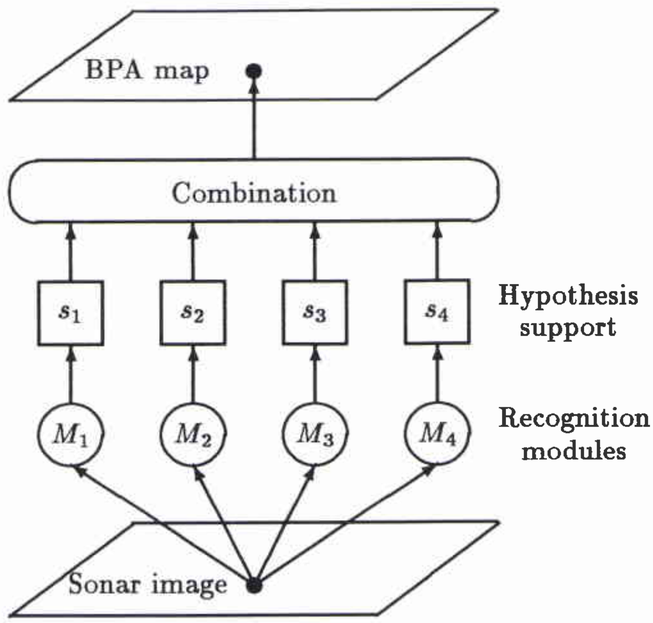

Figure 2: Modular image interpretation system architecture.

given to different sets must be combined. A straightforward way to combine the simple support functions is to assume independence and use Dempster's rule. The independence assumption is justified if the methods used to determine support are based on different principles so that different features in the image are used. In this situation the support for one hypothesis cannot be predicted from the knowledge of support assigned to other hypotheses. But this will not in general be the case. Here, a more cautious approach is taken. The support determined by the recognition modules is considered as inferential evidence, and the simple support functions are combined into a consonant support function.

The features in the neighbourhood of a point in the image are caused by the presence of a particular object on the sea-bottom. If two modules based on these features simultaneously assign support to two different singleton hypotheses, θ_1 and θ_2, it would be natural to conclude that the two objects are to some extent indiscernible on the basis of the available evidence. In the combined mass function some support should therefore be given to the singleton with the highest degree of support, say $\{\theta_2\}$, and some support should be given to the union of the two singleton sets $\{\theta_1, \theta_2\}$.

To determine the consonant support function, the degrees of support in the simple support functions associated with the recognition modules are ordered according to increasing magnitude,

$$s_i \leq s_j, \quad i < j, \quad i = 1, \ldots, n-1. \tag{13}$$

The subsets denoted by $E_1, \ldots, E_{n-1}$ correspond to the ordered degrees of support. For notational convenience $E_0 = E_n = \Theta$ with $s_0 = 0$ and $s_n = 1$ are introduced. The ***differential degree of support*** $\Delta s(F_j)$ for a set F_j is defined as

$$\Delta s(F_j) = s_j - s_{j-1}, \quad j = 1, \ldots, n, \tag{14}$$

and the mass value for the set D in the combined support function is determined by summing all differential degrees of support for F_j

$$m(D) = \sum_{F_j = D} \Delta s(F_j). \tag{15}$$

The set F_j is determined from the set G_j of sets E_i with degrees of support s_i equal to or larger than s_j

$$G_j = \{E_i | E_i \in 2^{\Theta}, s_i \geq s_j\}. \tag{16}$$

If G_j contains any singletons, F_j is determined as the union of these

$$F_j = \cup_{i \in I, I = \{i | E_i \in G_j, E_i \in \Theta\}} E_i. \tag{17}$$

Otherwise F_j is determined as the intersection of the non–singletons

$$F_j = \cap_{i \in I, I = \{i | E_i \in G_j, E_i \notin \Theta\}} E_i. \tag{18}$$

As the E_i are all singletons or complements of singletons, the result of the combination outlined above is a consonant support function.

The creation of a consonant support function is demonstrated by the example illustrated in Fig. 3. Three modules report non–zero degrees of support. The supported subsets are $\{MMO1\}$, $\{MMO2\}$, and $\overline{\{NAT1\}}$ with degrees of support $s_{\{MMO1\}}$, $s_{\{MMO2\}}$ and $s_{\overline{\{NAT1\}}}$.

The basic probability values assigned to the focal elements of the consonant support function are determined from Eq. 13–18 to be

$$m(\Theta) = 1 - s_{\overline{\{NAT1\}}} \tag{19}$$

$$m(\overline{\{NAT1\}}) = s_{\overline{\{NAT1\}}} - s_{\{MMO2\}} \tag{20}$$

$$m(\{MMO1, MMO2\}) = s_{\{MMO1\}} - 0 \tag{21}$$

$$m(\{MMO2\}) = s_{\{MMO2\}} - s_{\{MMO1\}}. \tag{22}$$

The system architecture outlined above has several advantages compared to the traditional rule–based approach. All knowledge concerning an object resides within

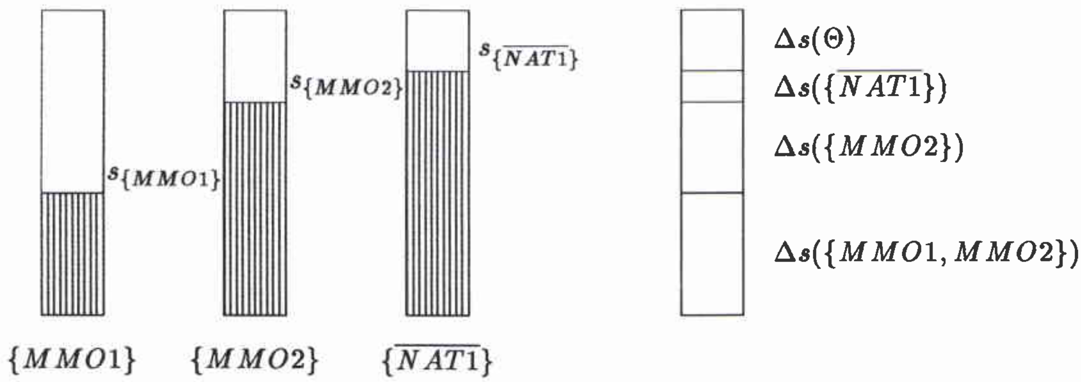

Figure 3: Example of combining simple support functions to a consonant support. Three simple support functions resulting from the recognition modules are illustrated. The singletons $\{MMO1\}$ and $\{MMO2\}$ have degrees of support $s_{\{MMO1\}}$ and $s_{\{MMO2\}}$, respectively. The support to the complementary of a singleton $\{\overline{NAT1}\}$ is $s_{\{\overline{NAT1}\}}$. The degrees of support are illustrated by the cross-hatched areas in the bars which span the interval $[0, 1]$. The rightmost bar illustrates the resulting differential degrees of support.

one module and it can be designed, tested and maintained independently. The performance of the modules is not altered by changing other modules or by changing the frame of discernment. Due to its modularity, the system is suited to distributed processing. The most important advantage is that no constraints are placed on the methods used to recognize particular objects. It is only required that the degree of support for a hypothesis be determined. The principles for determining degrees of support are the subject of the next section.

5

Determining belief

A fundamental problem is to determine the support that a module shall assign to a hypothesis given a particular pattern in a sonar image. The details of determining the support depend on the type of model used for recognition. This section gives a brief overview of the methods available for determining support.

In the simplest case the signal amplitude is less than the noise level of the system. This can be at long range or in shadow regions. In this case no information is available describing the sea–bottom and the result is the vacuous support function. The support assigned should in general increase with increasing quality of the image.

Objects for which detailed shape information is available can be recognized by matching features derived from the image with features synthesized by a model of the object. The similarity between observed and synthesized features can be measured by determining a matching score from a measure of distance between features, or in simple cases by comparing the number of matched features with the number of model features. The support can be determined from the matching score by adjusting the mapping from matching score to support until a desired performance is achieved.

Images of background structures, i.e., extended and essentially flat homogeneous areas such as flat sand and sand ripples, can often be satisfactorily described by a stationary random field model. A large number of techniques are available for determining support in this case. The method based on spatial point processes mentioned previously [8] can be adapted. Textural feature based methods [5,6] can also be used. The support can be determined by obtaining a measure of the distance between an observed feature vector and a template vector. For small distances support is assigned to the confirming hypothesis, and for large distances support is assigned to the disconfirming hypothesis. Linear prediction [19] is another applicable method. The support can be determined from the statistics of the residuals. Other classical methods for modelling texture [20] are potentially useful. For images that can be represented by a stationary random field model, several methods are available for determining the support. Weak nonstationarity can be handled by adaptive methods.

A problem is posed by structures which cannot be represented by a random field model and objects for which no detailed shape information is available. Methods for handling these cases are scarce. One method based on features determined from the shape of shadows and highlights [4] is not sufficiently specific to be useful in a general context. An operator distinguishes the difference between a man–made cylinder and a stone using knowledge about both cylinders and stones. The problem for an automatic system is the representation of knowledge about objects such as stones. In some cases contextual information can be utilized. For instance, the belief that an object is a stone could be conditioned on the number of stones previously identified in the vicinity of the object. Much of the variation in a sonar image can be attributed to variation in topography. Extended non–flat sea–bottoms could be characterized by a combination of large scale topography and local properties.

It is possible with existing methods to determine belief in hypotheses concerning backgrounds of extended homogeneous areas and objects of known shape. Determination of belief concerning objects without detailed knowledge of shape is difficult due to limited image resolution and lack of models for this type of object.

6

Combining belief

A sonar image can be interpreted by means of the method outlined in the previous sections. The result of the interpretation is a basic probability assignment map (BPA map) of known registration with respect to the sea–bottom. When more than one image of a scene is available the corresponding BPA maps can be combined point–by–point using Dempster's rule of combination. In this section the combination of belief functions will be discussed by considering some examples. The intention is to point out the implications for recognition performance when multiple images of the same scene are combined. Finally, the combination of BPA maps will be demonstrated using images from an experimental sonar system with two simultaneous views.

Dempster's rule can be used to combine belief functions based on independent evidence. The independence can be assured only if it is required that different images of an object contain different visible object features. If images are recorded while moving a sonar around an object, it is evident that the number of images containing different visible features depends on the shape of the object. It is thus not possible to establish a theoretical criterion for the number of independent views without considering the details of the object. For many objects of practical interest, the possible number of independent views is small. This limitation should be borne in mind when planning acquisition of the images. If additional images are available that cannot be considered independent, a way to utilize these would be to form the average of the simple support functions from the recognition module before creating the consonant support function. For the sake of simplicity, the following examples consider only hypotheses concerning five types of objects: no structure NOS; natural structure $NAT1$; and man–made objects $MMO1$, $MMO2$ and $MMO3$.

Consider first a situation of concordance. The BPAs from two images both support the hypothesis $MMO1$. The BPAs are:

$$\begin{aligned} m_1(<\{MMO1\},\Theta>) &= <0.70, 0.30> \\ m_2(<\{MMO1\},\Theta>) &= <0.80, 0.20> . \end{aligned} \tag{23}$$

The result of the orthogonal sum $m = m_1 \oplus m_2$ is

$$m(<\{MMO1\},\Theta>) = <0.94, 0.06> . \tag{24}$$

The belief intervals for the singleton hypotheses are

$$Ivl(\{NOS\}) = [0.00, 0.06]$$

$$\begin{aligned} Ivl(\{NAT1\}) &= [0.00, 0.06] \\ Ivl(\{MMO1\}) &= [0.94, 1.00] \\ Ivl(\{MMO2\}) &= [0.00, 0.06] \\ Ivl(\{MMO3\}) &= [0.00, 0.06]. \end{aligned} \tag{25}$$

Combination of concordant mass functions is seen to concentrate the probability mass. Suppose K independent images are interpreted, and that the results are simple support functions with the same focus and with support s_k. The combined support S is then given by

$$S = 1 - \prod_{k=1}^{K} (1 - s_k). \tag{26}$$

If the s_k are considered random variables with mean s, a first–order approximation of the expectation of S is given by

$$E(S) \approx 1 - (1 - s)^K. \tag{27}$$

This expression is an upper bound of the expected gain from using multiple views since not all BPAs can be expected to be consenting.

Consider next a situation of consonance. In one BPA it is not possible to discriminate between $NAT1$ and $MMO1$ so support is given the set $\{NAT1, MMO1\}$. The mass function originating from a second image is focused on $\{MMO1\}$. The mass functions are given by

$$\begin{aligned} m_1(< \{NAT1, MMO1\}, \Theta >) &= < 0.90, 0.10 >, \\ m_2(< \{MMO1\}, \Theta >) &= < 0.60, 0.40 >, \end{aligned} \tag{28}$$

and the composite mass function is determined to be

$$m(< \{MMO1\}, \{NAT1, MMO1\}, \Theta >) = < 0.60, 0.36, 0.04 > . \tag{29}$$

The belief intervals are

$$\begin{aligned} Ivl(\{NOS\}) &= [0.00, 0.04] \\ Ivl(\{NAT1\}) &= [0.00, 0.40] \\ Ivl(\{MMO1\}) &= [0.60, 1.00] \\ Ivl(\{MMO2\}) &= [0.00, 0.04] \\ Ivl(\{MMO3\}) &= [0.00, 0.04]. \end{aligned} \tag{30}$$

Although the probability mass assigned to the frame of discernment is reduced, the mass assigned to the singleton $\{MMO1\}$ is unaltered. The result is a consonant support function.

The resolution of ambiguity can be demonstrated by considering a situation with hypotheses concerning three man–made objects $MMO1$, $MMO2$ and $MMO3$. The hypothesis supported based on one image is $MMO1$ or $MMO2$ and the hypothesis supported in another image is $MMO1$ or $MMO3$. The mass functions are

$$\begin{aligned} m_1(<\{MMO1, MMO2\}, \Theta>) &= <0.90, 0.10>, \\ m_2(<\{MMO1, MMO3\}, \Theta>) &= <0.80, 0.20>. \end{aligned} \tag{31}$$

The combined mass function is

$$m(<\{MMO1\}, \{MMO1, MMO2\}, \{MMO1, MMO3\}, \Theta>) = <0.72, 0.18, 0.08, 0.02>, \tag{32}$$

and the belief intervals are

$$\begin{aligned} Ivl(\{NOS\}) &= [0.00, 0.02] \\ Ivl(\{NAT1\}) &= [0.00, 0.02] \\ Ivl(\{MMO1\}) &= [0.72, 1.00] \\ Ivl(\{MMO2\}) &= [0.00, 0.20] \\ Ivl(\{MMO3\}) &= [0.00, 0.10]. \end{aligned} \tag{33}$$

This example shows that although support is not given to a singleton in any of the mass functions the orthogonal sum focuses the probability mass on the singleton $\{MMO1\}$.

Another situation of interest is the situation of conflict. One BPA supports the hypothesis of a natural object $NAT1$ while another BPA supports the hypothesis of a man-made object, $MMO1$.

$$\begin{aligned} m_1(<\{NAT1\}, \Theta>) &= <0.80, 0.20> \\ m_2(<\{MMO1\}, \Theta>) &= <0.60, 0.40>. \end{aligned} \tag{34}$$

The conflict is reflected in the combined mass function by the reduction of the support for both hypotheses.

$$m(<\{MMO1\}, \{NAT1\}, \Theta>) = <0.23, 0.62, 0.15>. \tag{35}$$

The belief intervals show partial support for both hypotheses.

$$\begin{aligned} Ivl(\{NOS\}) &= [0.00, 0.15] \\ Ivl(\{NAT1\}) &= [0.62, 0.77] \\ Ivl(\{MMO1\}) &= [0.23, 0.38] \\ Ivl(\{MMO2\}) &= [0.00, 0.15] \\ Ivl(\{MMO3\}) &= [0.00, 0.15]. \end{aligned} \tag{36}$$

To resolve a conflict additional data must be acquired; if a decision has to be made on the available data, it will be made under conditions of uncertainty.

Finally consider a situation of contradiction. One BPA supports the hypothesis of a man-made object $MMO1$ while another BPA supports the complementary hypothesis, $\overline{MMO1}$.

$$
\begin{aligned}
m_1(<\{MMO1\},\Theta>) &= <0.90, 0.10> \\
m_2(<\{\overline{MMO1}\},\Theta>) &= <0.90, 0.10> .
\end{aligned} \tag{37}
$$

The contradiction is reflected by a reduction of the mass values for the two hypotheses in the combined mass function.

$$
m(<\{MMO1\},\{\overline{MMO1}\},\Theta>) = <0.47, 0.47, 0.05> . \tag{38}
$$

The belief interval for $\{MMO1\}$ is narrow but the plausibility is equal to that of other hypotheses.

$$
\begin{aligned}
Ivl(\{NOS\}) &= [0.00, 0.53] \\
Ivl(\{NAT1\}) &= [0.00, 0.53] \\
Ivl(\{MMO1\}) &= [0.47, 0.53] \\
Ivl(\{MMO2\}) &= [0.00, 0.53] \\
Ivl(\{MMO3\}) &= [0.00, 0.53].
\end{aligned} \tag{39}
$$

This example demonstrates how the effect of a false positive can be removed by using disconfirmation of hypotheses and multiple images.

In order to demonstrate the utility of the method, images acquired with a two–view sidescan sonar system have been combined. The system is based on a commercial sidescan sonar system with the port transducer removed and placed in a wing on the starboard side. In this way two images of the sea–bottom are recorded simultaneously on the starboard side. One image is the normal broadside image and another is a 45 degree forward looking image acquired by the transducer in the wing. The 45 degree separation between views is considered sufficient to justify an assumption of independence between the two views.

Two images of a scene containing two cylindrical objects and a bush of *Posidonia oceanica* on a sand background are shown in Fig. 4. The differences in the images are particularly evident from the orientation of the object shadows. The images have been registered to ground assuming a flat, horizontal sea–bottom. As the sea–bottom is not flat and horizontal, additional registration was obtained by warping one image onto the other before creating the BPA maps.

The BPA maps were created manually by inspection of the sonar images. Color coded representations of the BPA maps are shown in Fig. 5 and the mass functions corresponding to the colors are shown in Table 1. In order to limit the storage required for the BPA maps the resolution is less than the resolution of the sonar

Color	BPA
Blue	$m(<\Theta>) =< 1.00 >$
Red	$m(<\{NOS\}, \Theta>) =< 0.80, 0.20 >$
Grey	$m(<\{NAT1\}, \Theta>) =< 0.80, 0.20 >$
Green	$m(<\{MMO1\}, \Theta>) =< 0.80, 0.20 >$
Cyan	$m(<\{NAT1, MMO1\}, \Theta>) =< 0.90, 0.10 >$

Table 1: Color codes corresponding to BPAs.

images. A further reduction in storage requirements can be obtained by contour encoding.

The belief and plausibility maps of the combined BPAs for the hypotheses $MMO1$, $NAT1$ and NOS are shown in Fig. 6–8. In these images the interval $[0, 1]$ is mapped into the grey scale [Black,White]. The differences between the belief and plausibility maps reflect the lack of information in shadow regions. Other differences are due to conflicts at the border regions between objects and background. The geometry of the man–made and natural objects in the belief maps in Fig. 6 and 7 are seen to reflect the visible geometry of the objects in the sonar images whereas the plausibility maps are more diffuse. The belief maps are useful for making decisions based on what is known, whereas the plausibility maps and the ignorance maps shown in Fig. 9 are useful for planning additional data acquisition.

The results presented here, can be used both for display of information and further processing such as automatic decision making. The method is especially useful in situations where additional data can be acquired during an operation.

The combination of information is not limited to BPA maps originating from sidescan sonar images. Other sensor data such as gap filling sonar images, video images and laser scan images can be combined if an interpretation in terms of a BPA map is performed. This makes the proposed method well suited as a general framework for creating and representing information on the sea–bottom.

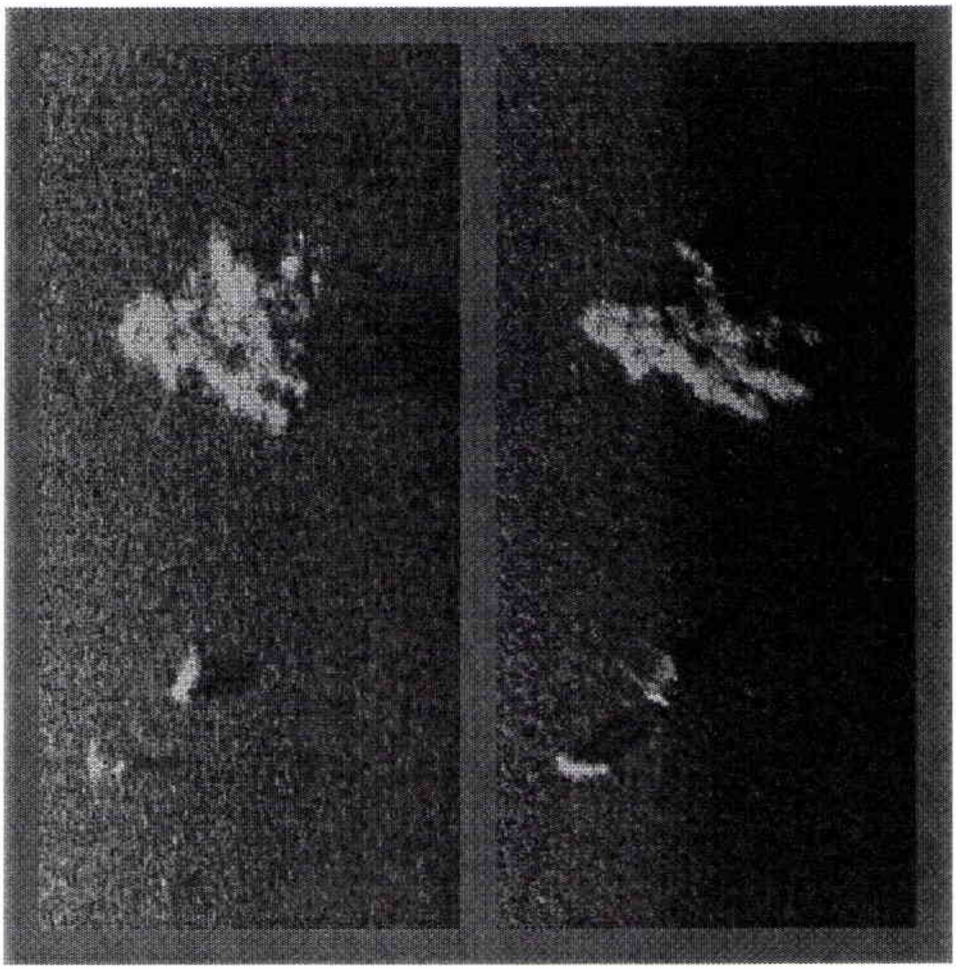

Figure 4: Two view Sonar images.

intentionally blank page

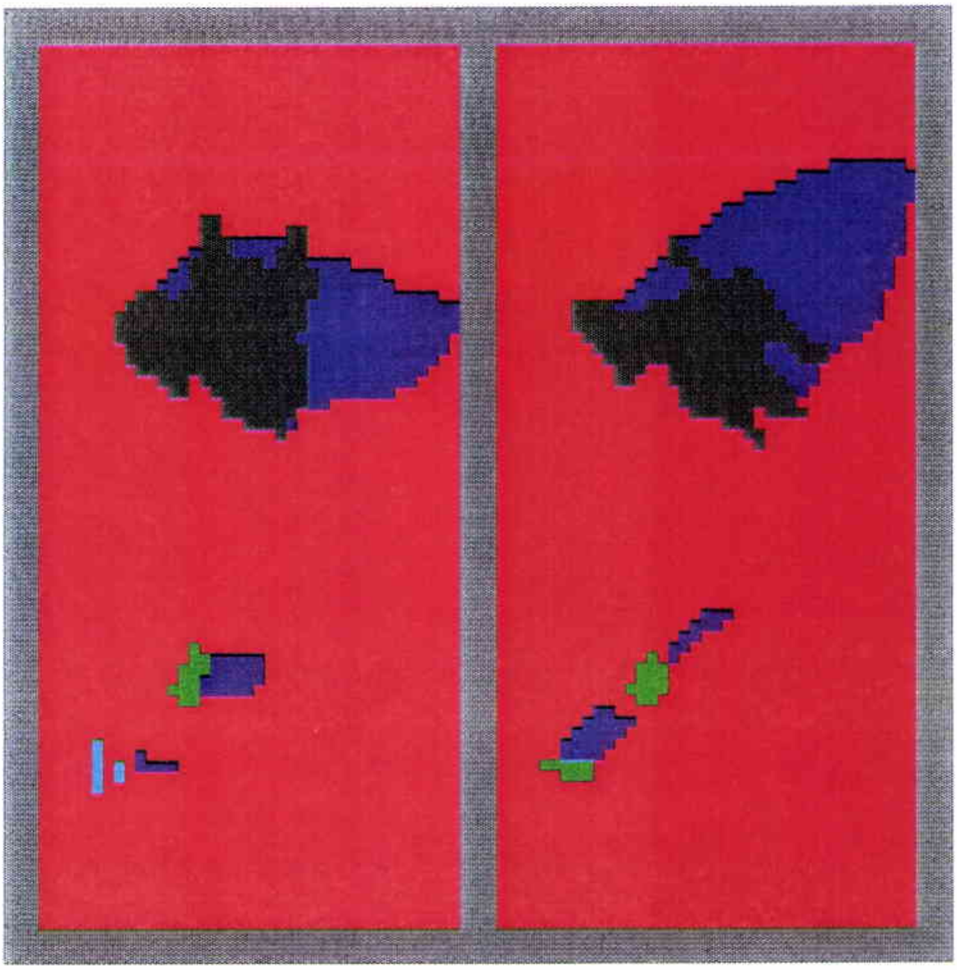

Figure 5: Two view color coded BPA maps.

intentionally blank page

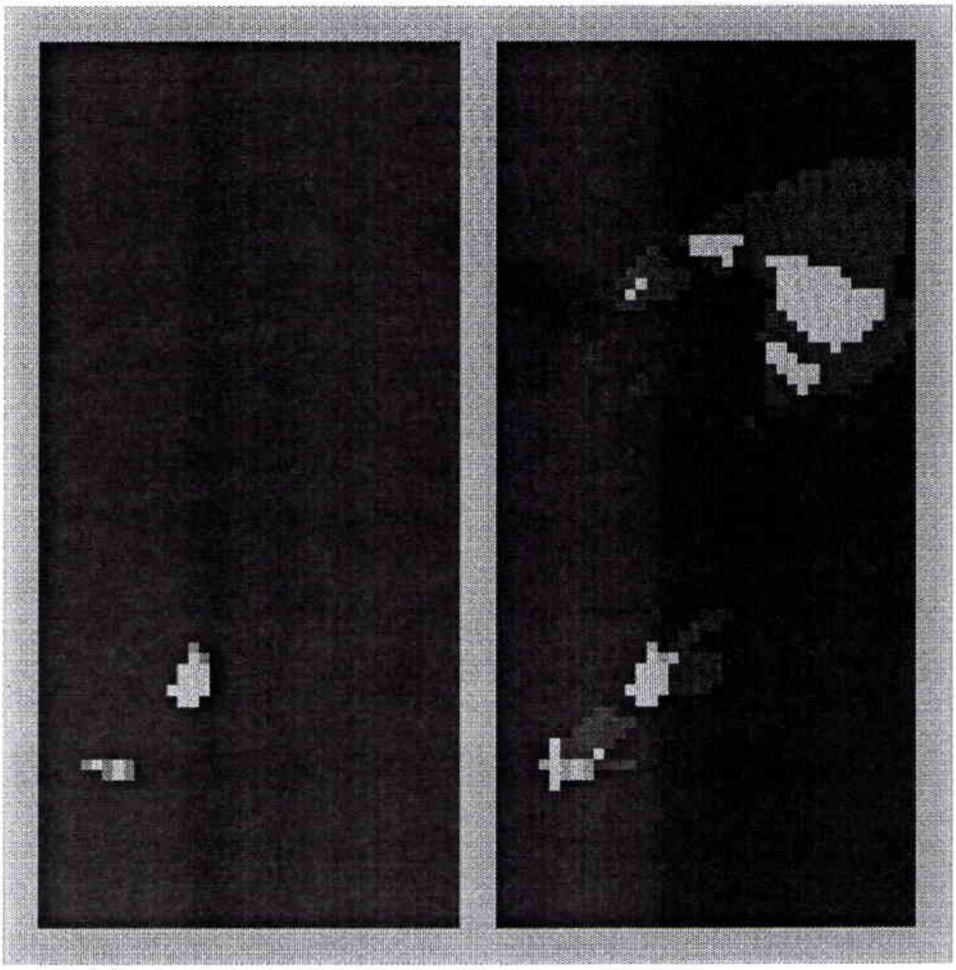

Figure 6: $Bel(\{MMO1\})$ and $Pls(\{MMO1\})$.

intentionally blank page

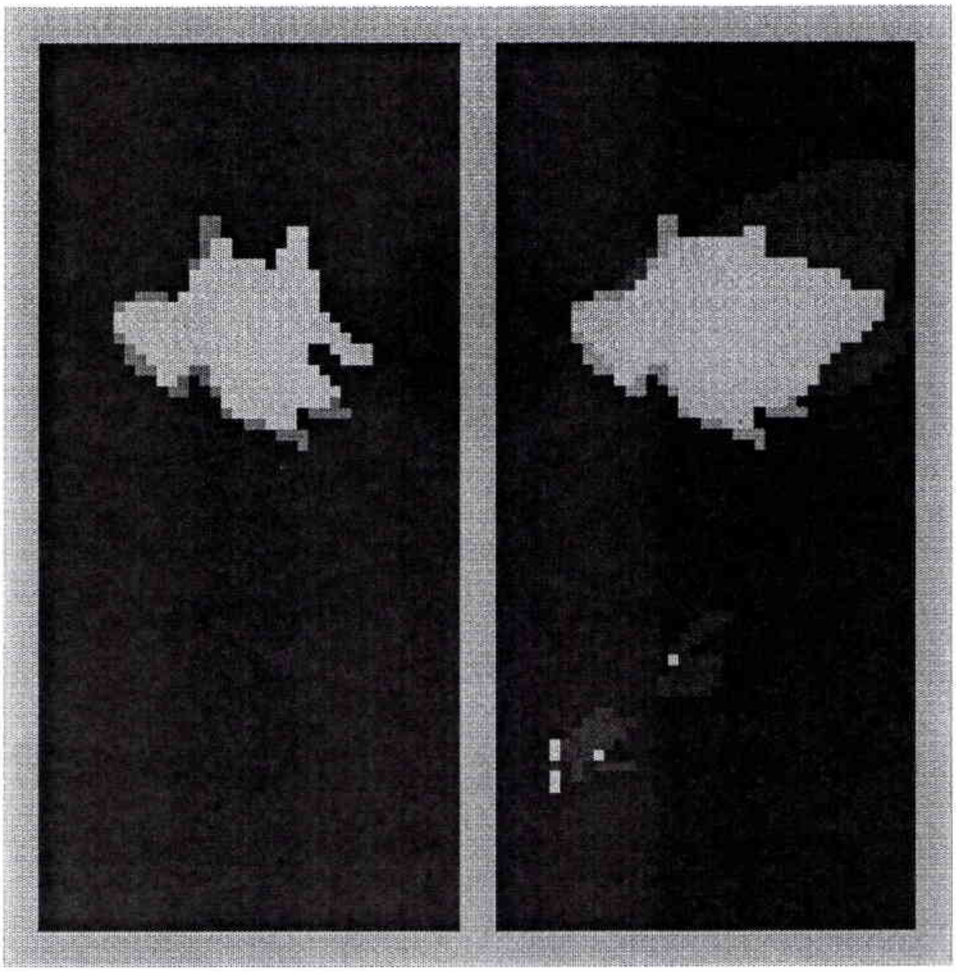

Figure 7: $Bel(\{NAT1\})$ and $Pls(\{NAT1\})$.

intentionally blank page

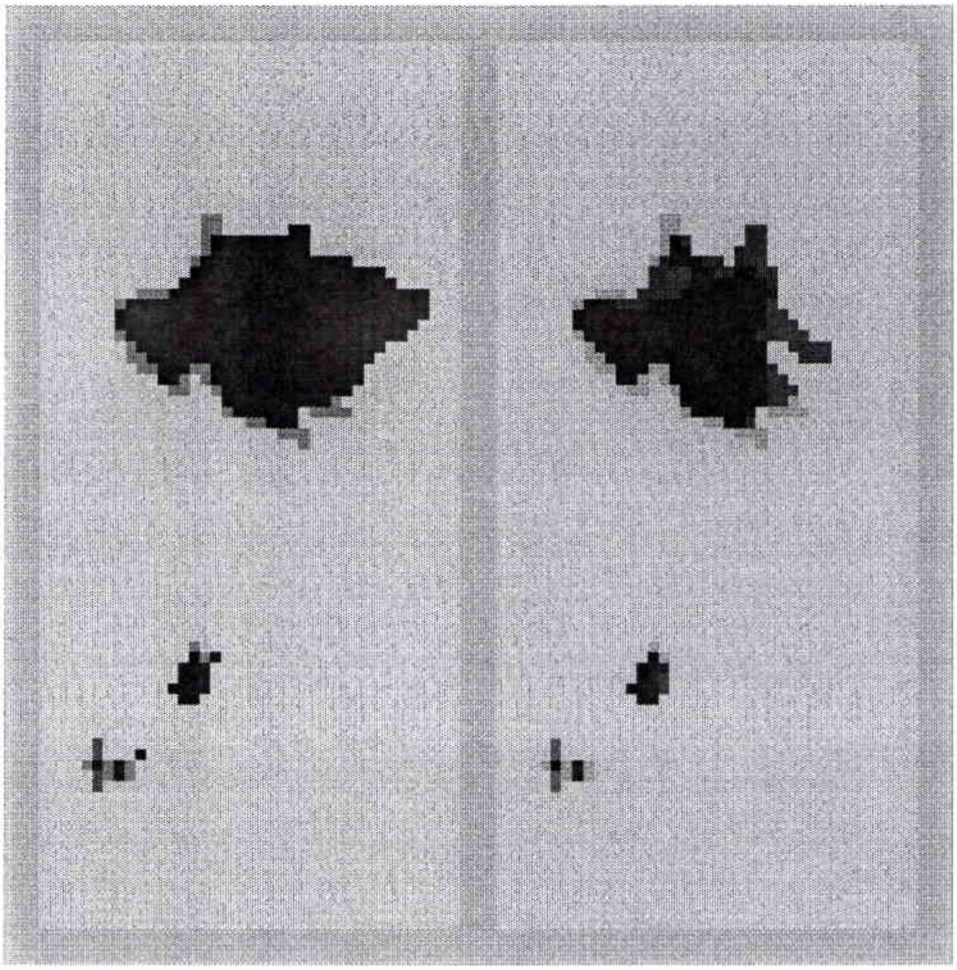

Figure 8: $Bel(\{NOS\})$ and $Pls(\{NOS\})$.

intentionally blank page

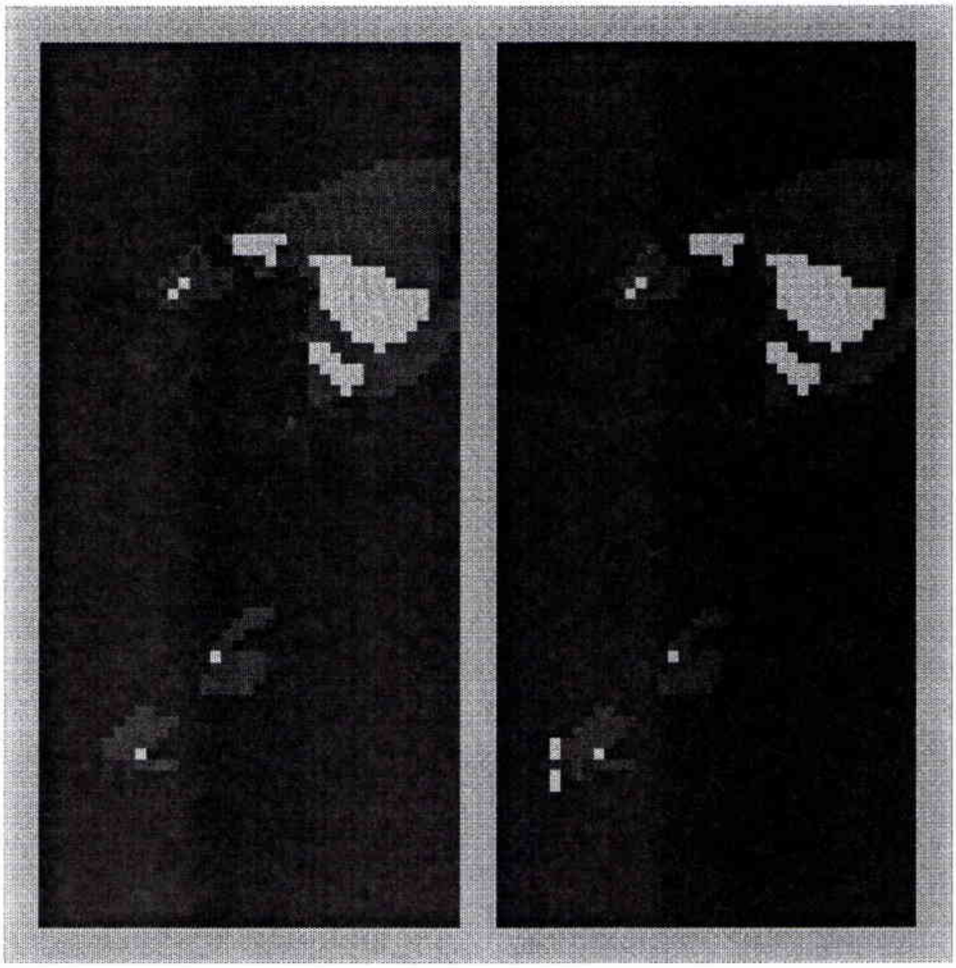

Figure 9: $Ign(\{NAT1\})$ and $Ign(\{MMO1\})$.

7

Conclusion

A method and system architecture for automatic interpretation and fusion of multiple sonar images of the sea–bottom based on the Dempster–Shafer theory of evidence is proposed. The main advantages of the Dempster–Shafer theory are that it allows distinction between uncertainty and ignorance and provides a consistent way of combining information. The system can be used for interpretation of images containing man–made as well as natural objects and does not impose constraints on the type of models that can be used for the interpretation. The proposed system is modular, extendible and well suited to distributed processing.

The utility of the concept of using multiple images of the same site has been demonstrated by means of examples, and the viability of the technique has been demonstrated by combining sonar images acquired with an experimental sonar with two simultaneous views.

8 Acknowledgements

The following persons deserve thanks for their contribution to this work. The experiment for acquiring the two view sonar image was coordinated by P. Guerrini. Experimental equipment was designed by L. Gualdesi and V. Manzotti. The experiment was performed on the "R/V Manning" with the assistance of A. Spairani, G. Bertoli, G. Ciuffardi, F. Prezioso, M. Vitozzi , D. Galletti, D. McGowan and M. Cassola. D. Sheldon and B. Zerr provided helpful suggestions to improve the quality of the paper.

References

[1] Doherty M.F., Landowski, J.G, Maynard P.T., Uber G.T., Fries D.W. and Maltz F.H. Side Scan Sonar Object Classification Algorithms. Proceedings of the 6th International Symposium on Unmanned Untethered Submersible Technology, Elliot City, MD, 1989.

[2] Schweizer,P.F. and Petlewitch, W.J. Automatic Target Detection and cueing system for an autonomous underwater vehicle. Proceedings of the 6th International Symposium on Unmanned Untethered Submersible Technology, 1989: 359–371.

[3] Schweizer, P.F., W.J. Petlewitch *et al.* Image Processing Architecture for autonomous underwater vehicles in mine detection and classification operations. Proceedings of the 7th International Symposium on Unmanned Untethered Submersible Technology, 1991: 280–293.

[4] Johnson, S.G. and Deaett, M.A. The Application of Automated Recognition Techniques to Side–Scan Sonar Imagery. IEEE Journal of Oceanic Enginering, Vol. 19, no. 1, 1994: 138–144.

[5] Pace, N.G. and Dyer, C.M. Machine Classification of Sedimentary Sea Bottoms. IEEE Transactions on Geoscience and Remote Sensing, Vol. 17, No. 3, 1979: 52-56.

[6] Pace, N.G. and Gao, H. Swathe Seabed Classification. IEEE Journal of Oceanic Enginering, Vol. 13, no. 2, 1988: 83-90.

[7] Stewart W.K., Min Jiang, and Marra, M. A Neural Network Approach to Classification of Sidescan Sonar Imagery from a Midocean Ridge Area. IEEE Journal of Oceanic Enginering, Vol. 19, no. 2, 1994: 215-225.

[8] Carmichael, D.R., Clarke, S.J. and Linnett, L.M. Spatial point process models for texture analysis and object detection. Undersea Defense Technology Conference Proceedings, 1994: 200–203.

[9] Carmichael, D.R., Linnett, L.M., Clarke, S.J. and Calder, B.R. Texture measures for seabed classification and object detection. Undersea Defense Technology Conference Proceedings, 1995: 200–203.

[10] Russel, G.T. and Lane, D.M. A knowledge based system framework for environmental perception in a subsea robotic framework. IEEE Journal of Oceanic Enginering, Vol. 11, no. 3, 1986: 401–412.

[11] Lane, D.M. and Stoner, J.P. Automatic Interpretation of Sonar Imagery Using Qualitative Feature Matching. IEEE Journal of Oceanic Enginering, Vol. 19, no. 3, 1994: 391–405.

[12] Trimble, G.M. Vilaro, J., Okamura, D., Lum, R., and Dutta, K. Underwater Object Recognition and Automatic Positioning to Support Dynamic Classification. Proceedings of the 7th International Symposium on Unmanned Untethered Submersible Technology, 1991: 273–279.

[13] Kuwahara, R.H. and Poeckert, R.H. Accurate Seafloor Mapping Using Precise Navigation and Side–scan Sonar. Oceans 89 Conference Proceedings, 1989: 1148-1152.

[14] Stage B., and Zerr, B. Multiple View Sidescan Sonar. Undersea Defence Technology Conference Proceedings, 1995: 235–238.

[15] Shafer, G. A mathematical theory of evidence. Princeton University Press, 1976.

[16] Guan, J.W. and Bell, D.A. Evidence Theory and its Applications, vol. 1, Elsevier Science Publishers, 1991.

[17] Garvey, T. Lowrance, J. and Fischler, M. An inference technique for integrating knowledge for disparate sources. Proceedings 7th International Joint Conference on Artificial Intelligence, Univ. British Columbia, Vancouver, BC, Canada, 1981.

[18] Wilkins D.C. and Yong Ma The refinement of probabilistic rule sets: sociopathic interactions. Artificial Intelligence, 70 , 1994: 1–32.

[19] Jain, A.K. Advances in mathematical Models for Image Processing. Proceedings of the IEEE, Vol. 69, no. 5, 1981: 502–528.

[20] Haralick, R.M. Statistical and structural approaches to texture. Proceedings of the IEEE, Vol. 67, no. 5, 1979: 786–804.

Report no. changed (Mar 2006): SM-294-UU

Document Data Sheet

NATO UNCLASSIFIED

Security Classification		Project No.
NATO UNCLASSIFIED		25
Document Serial No.	**Date of Issue**	**Total Pages**
SM-294	December 1995	41 pp.

Author(s)

B. Stage

Title

Interpretation and fusion of sonar images using evidential reasoning.

Abstract

A method and system architecture for automatic interpretation and fusion of multiple sonar images is proposed. The interpretation of a sonar image consists of the creation of a symbolic geographical map indicating the belief in the presence of objects and structures on the sea-bottom. Fusion of information from multiple images of the same site is at the symbolic level. Both interpretation and fusion are based on the Dempster-Shafer theory of evidence. The system can be used for interpretation of images containing man-made as well as natural objects irrespective of model type. The method is applied to fusion of information from images acquired with an experimental sonar system with two simultaneous views.

Keywords

Fusion – interpretation – sonar images

Issuing Organization

North Atlantic Treaty Organization
SACLANT Undersea Research Centre
Viale San Bartolomeo 400, 19138 La Spezia, Italy

[From N. America: SACLANTCEN CMR-426 (New York) APO AE 09613]

Tel: +39 (0)187 540 111
Fax:+39 (0)187 524 600

E-mail: library@saclantc.nato.int

NATO UNCLASSIFIED

Report no. changed (Mar 2006): SM-294-UU

Initial Distribution for SM-294

SCNR for SACLANTCEN		*National Liaison Officers*	
SCNR Belgium	1	NLO Canada	1
SCNR Canada	1	NLO Denmark	1
SCNR Denmark	1	NLO Germany	1
SCNR Germany	1	NLO Italy	2
SCNR Greece	2	NLO Netherlands	1
SCNR Italy	1	NLO UK	3
SCNR Netherlands	1	NLO US	4
SCNR Norway	1		
SCNR Portugal	1		
SCNR Spain	1		
SCNR Turkey	1		
SCNR UK	1		
SCNR US	2		
French Delegate	1		
SECGEN Rep. SCNR	1	Total external distribution	34
NAMILCOM Rep. SCNR	1	SACLANTCEN Library	26
SACLANT	3	Total number of copies	60