Markov random field and fuzzy logic modeling in sonar imagery: application to the classification of underwater floor

M. Mignotte and C. Collet

Groupe de Traitement du Signal, Ecole Navale, Lanvéoc-Poullmic, BP 600, 29240 Brest-Naval, France

and

P. Pérez and P. Bouthemy

IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes cedex, France.

Person to whom correspondence should be sent:

Name: Christophe collet
Address: Groupe de Traitement du Signal, Ecole Navale,
        Lanvéoc-Poullmic, BP 600 - 29240 Brest-Naval, France.
Email: collet@ecole-navale.fr
Phone: (33)(0) 298 234 044
Fax: (33)(0) 298 233 857
This paper proposes an original method for the classification of sea-floors from high resolution sidescan sonar images. We aim at classifying the sonar images into five kinds of regions: sand, pebbles, rocks, ripples and dunes. The proposed method adopts a pattern recognition approach based on the extraction and the analysis of the cast shadows exhibited by each sea-bottom type. This method consists of three stages of processing. Firstly, the original image is segmented into two kinds of regions: shadow (corresponding to a lack of acoustic reverberation behind each “object” lying on the sea-bed) and sea-bottom reverberation. Secondly, based on the extracted shadows, shape parameter vectors are computed on sub-images and classified with a fuzzy classifier. This preliminary classification is finally refined thanks to a Markov random field model which allows to incorporate spatial homogeneity properties one would expect for the final classification map. Experiments on a variety of real high resolution sonar images are reported.

Key Words: high resolution sidescan sonar, sea-bed classification, acoustic shadow, shape analysis, fuzzy classifier, Markov random field.

1. INTRODUCTION

High-resolution sidescan sonar plays an important role in underwater sensing for it provides acoustic “images” of the sea-bed whose quality is much higher than the one of images supplied by optical means [13]. One of the applications of sidescan sonar is the automatic segmentation and classification of the sea-bottom. The segmentation of sea-floor sonar images aims at partitioning the acoustic image into homogeneous regions with respect to certain physical properties or geological characteristics. The goal of the classification task is to assign these different geoacoustic regions to sea-floor types as sand, pebbles, rocks, ripples (or ridges), dunes, etc.

Over the last decades, with the significant advances in mapping techniques and their increasing use, the classification of sea-floor based on sidescan sonar imagery has become an important research topic for marine geophysicists. It plays an important role in understanding the undersea environment, and it is of great interest in a wide range of both military and civilian applications, including geological survey (cartography of sea-floors, geophysical exploration, etc.), ocean engineer-
ing (use of autonomous underwater vehicles, surveillance of pipelines and cables, etc.), military surveillance and simulations, or the detection and classification of manufactured objects lying on sea-floors [16].

A general procedure for sea-floor classification consists of the following steps: (1) data acquisition; (2) possible preprocessing, e.g., geometric correction, reduction of the signal dynamics, contrast correction, noise filtering, etc.; (3) feature extraction over small two-dimensional areas (called sub-images or windows in the following) within the image. This step aims at reducing the information contained in each sub-image to a relevant feature vector; (4) selection of a supervised or unsupervised classification technique; and (5) classification of each sub-image.

For the feature extraction step, a commonly used approach consists in working on the “texture” of sea-floor sonar images. One can use either the raw input image, i.e., the grey-levels themselves [5, 23], or some relevant textural measures to be extracted from the image, such as the grey level co-occurrence matrices [22, 28], the grey level run length difference [21], the autoregressive 1D or 2D parameters [8, 25], spectral estimates [21, 26], fractal measures [6], or some wavelet coefficients [9]. Nevertheless, in many cases, the input sonar image is strongly corrupted by speckle noise [13]. This correlated noise is due to the random interference of the acoustic waves scattered by the micro-structure of the object surfaces within one resolution cell and also to the signal brought by the minor lobes of the acoustic antenna. Depending on the properties of this noise, as well as on the conditions of acquisition (e.g., grazing angle) and the characteristics of the sonar (e.g., sonar gain) [20], images for a same type of sea-floor can exhibit a great deal of variability. As a consequence, the textural cues computed directly on such grey level images are also likely to vary a lot, although a unique type of sea-floor is considered. This lack of consistency between different sonar images for a given type of sea-floor is a critical issue for the classification techniques whose feature extraction module works directly on grey level images.

Instead of using directly the input image, i.e., the grey levels themselves, or some textural features derived from them, we thus propose an alternate approach where acoustic shadows are first extracted (thanks to the technique introduced in [15]). A pattern recognition methodology is then applied to the resulting shadow contours. The underlying rationale is that the morphological elements that compose each type of sea-floor (such as dunes, ripples, pebbles, rocks, etc.) can be identified in a robust and reliable way by simply looking at the shape of associated cast shadows.
Let us recall that in the emission stage, the antenna of the sidescan sonar generates highly directive acoustic waves in the direction orthogonal to the sonar displacement. For each impulse, reverberated signals are collected along with the time they took to get back, in a reception stage. The amplitude of this signal as a function of the time is then processed such as to provide one pixel line of the final sonar image. No acoustic signal is reverberated from behind “objects” lying on the floor, thus resulting in “acoustic shadows” in sonar images (see Fig. 1).

For the classification step, different kinds of methods have been considered in the literature. A first class of approaches resorts to standard statistical techniques such as Maximum Likelihood (ML) classifiers [8], Maximum A Posteriori (MAP) classifiers, etc. An inherent drawback of such statistical approaches is that it is usually assumed that the form of the probability distribution associated to each class is known, and that its parameters can be accurately estimated. This means that the performances of such techniques dependent on how well the selected statistical models are suitable for describing the data and how much data are available for learning step. In addition, estimated models are likely to be sonar-dependent, which we would like to avoid.

Contrary to these approaches, the K-means techniques [3] are unsupervised and do not require any parametric modeling of the data. Nevertheless the K-means approach assumes, often wrongly, the presence of spherical clusters of identical volume and low inertia in the feature space.

Neural classifiers have also been considered [5, 28, 21, 25, 26, 1, 2, 14]. In that case, one does not make use of much parametric prior knowledge. This provides a great deal of flexibility, but usually results in the need for a heavy learning process where the learning set must be devised with great care.

A quite flexible framework to combine various degrees of a priori knowledge, while keeping the parameter identification reasonable, is offered by fuzzy classification techniques. In this paper, we introduce such a fuzzy classification technique. We shall see that it allows us to capture in a simple yet efficient way the high level a priori knowledge we have on the shape of acoustic shadows within different types of sea-floors. Another appealing feature of the approach relies in its capability for handling mixtures of classes. This is important when dealing with sub-images which are likely not to exhibit only a unique type of sea-floor.

The main drawback of the various classifications techniques we have just evoked remains the lack of explicit relationship between adjacent regions (or sub-images). In order to obtain a more accurate segmentation map, spatial relationships should
be taken into account. To this end, we use a Markov random field (MRF) model which allows us to specify and handle in a flexible way the spatial dependencies between adjacent sub-images by means of a suitable a priori probability distribution [4].

In this paper we thus address the problem of sea-floor classification in high resolution sidescan sonar imagery, by combining a tailor-made fuzzy classifier working on shadow shapes, an a Markovian segmentation model. The proposed method involves four steps: (1) unsupervised two-class segmentation (shadow and reverberation areas); (2) feature extraction; (3) fuzzy pre-classification; (4) Markovian segmentation. The block diagram of this system is shown in Figure 2. The organization of the paper follows this chain structure: steps 2, 3, and 4 are respectively described in Sections 2, 3, and 4. Experimental results are reported and discussed in Section 5, before we conclude and present further research directions in Section 6.

2. FEATURE EXTRACTION STEP

The feature extraction step we consider does not directly handle the grey level sonar images (or some textural features deduced from local grey level distributions). Instead, it relies on a preliminary two-class segmentation of the sonar images into “acoustic shadows” on the one hand and “reverberation” on the other hand. To this end, we use the hierarchical Markovian segmentation technique that we have introduced in [17, 15].

This Bayesian segmentation method combines a two-fold data model (Gaussian distribution for the luminance in shadow regions, whereas the luminance distribution in reverberation zones is modeled with a Weibull law) with an original hierarchical Markovian prior. This prior, as standard Markovian priors, enables one to introduce spatial coherence in the segmentation process. In our case however, this coherence takes place at various “scales”, which has been experimentally demonstrated as providing improved results at lower cost (as compared, say, to standard spatial Potts model). Also this segmentation scheme has been made totally unsupervised by devising efficient tools for the estimation of the involved parameters. More specifically a so-called iterated conditional estimation (ICE) technique [18] has been devised for this model. It is an iterative method which, at each step, averages parameter estimates computed on segmentation samples drawn from the posterior distribution associated to the previous parameter fit. In our case, the least-squares estimator of Derin and Elliott [10] is used for the estimation of prior
parameters, whereas maximum likelihood estimators are derived for the parameters of the laws involved in the data model. This iterative procedure is initialized thanks to a preliminary K-means clustering of local grey level statistics. The good performances of this unsupervised method for segmenting into two classes high resolution sonar images has been thoroughly assessed on a variety of real images. An example of this hierarchical two-class segmentation is provided in Figure 7a-b. See [17, 15] for a complete account of the method.

The segmentation obtained by this technique is then high-pass filtered and binarized in order to extract the boundary of each detected cast shadow. The resulting edge image is partitioned into small windows (e.g., Figure 7c) from which feature vectors are extracted. The aim of this feature extraction process is to get parsimonious, and hopefully discriminant, information about the acoustic shadows associated to the different sea-floor types. The different cues that make up each of the feature vectors have to be devised carefully, based on one’s “expertise” of the concerned application. They might be of quite different natures (geometrical, spectral, statistical, etc.), but there should be only a limited number of them.

In our application, we have to distinguish between the cast shadows of ripples, those of dunes (which are elongated and roughly parallel), those of pebbles, and those of rocks (which are the most irregular, both in terms of shape and orientation). To this end, we first consider three different parameters which are computed for each individual shadow boundary. They are compactness, elongation, and orientation. Based on them, more global cues will then be defined within each of the sub-images to be classified. Before we come to this issue, we first define each of the three individual parameters for some shadow boundary (i.e., closed curve) $\Gamma$:

- **Compactness.** It is a dimensionless geometrical feature that accounts for the degree of complexity of the delimited region. It is defined as

$$C = \frac{4\pi A_\Gamma}{|\Gamma|^2},$$

where $|\Gamma|$ and $A_\Gamma$ stand respectively for the length of the boundary and for the area of the region delimited by this boundary. This parameter is equal to 1 when the shadow is exactly a circle and it gets close to 0 as the shadow gets more and more complex, or simply more and more elongated.

- **Elongation.** The elongation mentioned in previous item is specifically measured by the ratio of the two inertia moments of the boundary under concern. More
precisely, let us first define the inertia matrix of the boundary $\Gamma$ as

$$
C = \begin{pmatrix}
c_{xx} & c_{xy} \\
c_{xy} & c_{yy}
\end{pmatrix},
$$

with

$$
c_{xx} = \frac{1}{|\Gamma|} \sum_{s \in \Gamma} (x_s - x_G)^2, \\
c_{yy} = \frac{1}{|\Gamma|} \sum_{s \in \Gamma} (y_s - y_G)^2, \\
c_{xy} = \frac{1}{|\Gamma|} \sum_{s \in \Gamma} (x_s - x_G)(y_s - y_G),
$$

where the summations are taken over the $|\Gamma|$ pixels of coordinates $(x_s, y_s)$ that constitute $\Gamma$, and $(x_G = \frac{\sum_{s \in \Gamma} x_s}{|\Gamma|}, y_G = \frac{\sum_{s \in \Gamma} y_s}{|\Gamma|})$ stands for the inertia center. The two eigenvalues of this matrix, $c_{xx} + c_{yy} \pm \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2}$, correspond to the inertia along the maximum inertia axis (principal axis), and along the minimum inertia axis, respectively. The elongation is defined as the square root of the ratio of the largest eigenvalue over the other one:

$$
\xi = \sqrt{\frac{c_{xx} + c_{yy} + \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2}}{c_{xx} + c_{yy} - \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2}}}. 
$$

- Orientation. We measure the overall orientation of the contour as the angle between the direction of its principal axis and the $x$-axis. The slope of the principal axis of inertia being readily obtained as $c_{yy} - c_{xy} + \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2}$, the orientation is defined as

$$
\alpha = \arctan\left(\frac{c_{yy} - c_{xy} + \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2}}{2c_{xy}}\right).
$$

If the window under concern exhibits $M$ different shadows with boundaries $\Gamma_i$, $i = 1 \cdots M$, we thus end up with $M$ parameter triples $(C_i, \xi_i, \alpha_i)_{i=1\cdots M}$. These individual shape parameters have then to be combined into new parameters whose values should allow to infer the type of sea-floor present (or mainly present) in the window. Looking for a compromise between the parsimony of the resulting representation and the knowledge of sonar experts with whom we are working, we came up with four window-wise parameters defined as follows.
1. The *mean compactness*

\[ C_{\text{moy}} = \frac{1}{M} \sum_{i=1}^{M} C_i. \]  \hfill (6)

2. The *directivity*

\[ \sigma^2_{\alpha} = \frac{1}{M} \sum_{i=1}^{M} (\alpha_i - \overline{\alpha})^2, \]  \hfill (7)

with \[ \overline{\alpha} = \frac{1}{M} \sum_{i=1}^{M} \alpha_i, \]  \hfill (8)

is the empirical orientation variance in the window. It will allow us to assess whether the shadows within this window exhibit some sort of privileged direction.

3. The *maximal elongation*

\[ \xi_{\text{max}} = \max_{i \in \{1 \cdots M\}} \xi_i. \]  \hfill (9)

4. The *length of the longest shape boundary* in the window

\[ N_{\text{max}} = \max_{i \in \{1 \cdots M\}} |\Gamma_i|. \]  \hfill (10)

Once these four parameters have been computed for the \( k \)-th sub-image, they are gathered in a feature vector \( \mathbf{x}^k \). The classification process is then performed on the resulting set of such four-dimensional feature vectors.

### 3. FUZZY CLASSIFICATION

We now have to define a classifier on the previously introduced four-dimensional feature space. The characteristics of the classification problem under concern are the following:

- The four features to be used are of different natures;
- Our prior knowledge on how each of these shadow-based features should behave within each of the classes of interest is rather qualitative: no precise parametric prior, whether statistical or not, is available. Instead, as we shall see, the prior knowledge is a collection of qualitative statements of the type “this piece of seafloor is likely to include sand ripples because the detected cast shadows are stretched and exhibit some privileged orientation”;
- The boundaries between the different classes in the feature space cannot be defined in a clearcut way for two reasons: one reason is related to the fact that
the different features are computed over windows which might cover different sea-
floor types at the same time, resulting in a mixture of classes. The second reason
lies within the definition itself of the classification nomenclature. The demarcation
between “pebbles” and “rocks” for instance is something rather imprecise.

In view of these elements, we do think that fuzzy set theory [27] offers the most
appropriate tools for devising a classifier independently of the type of sonar. This
framework actually allows one to easily combine imprecise priors on classifications
with “fuzzy” boundaries within classifiers that are easy to train. For reaching
such goals, this framework seems to us more appropriate than statistical methods
(which require more formal prior, and often lead to difficult parameter estimation
problems), than K-means clustering whose underlying prior is not flexible enough
to fit our problem, and than neural classifiers, which we experimented earlier in the
same context [24] and whose specification and training are difficult for our problem.

We now define the fuzzy classifier we have devised. Recall that the data to be
classified are four-component feature vectors computed on a partition of the image
plane into windows. If the k-th window contains no detected cast shadows, it is
assigned right away to “sand” class. If this window contains at least one detected
acoustic shadow, a feature vector \( x^k = (\sigma^2_k, \xi_{\text{max}}^k, C_{\text{moy}}^k, N_{\text{max}}^k) \) is computed, as
explained in previous section. We then want to assign this vector to one of the
four following classes: ripples (label \( w_1 \)), dunes (label \( w_2 \)), pebbles (label \( w_3 \)), and
rocks (label \( w_4 \)). This assignment is done in a fuzzy way via membership degrees
\( \mu_i(x^k) \in [0, 1], i = 1 \cdots 4 \). The number \( \mu_i(x) \) should capture the strength of our
belief that the sea-floor within a window with feature vector \( x \) is mainly of type \( w_i \).
Extreme value \( \mu_i(x) = 1 \) (resp. \( \mu_i(x) = 0 \)) indicates that one is sure that sea-floor
\( w_i \) is present (resp. not present) in this window.

A convenient way to define these membership functions consists in looking first
how membership degrees can be assigned based on only one of the four feature
parameters. We thus have to define component-wise membership degrees \( \mu_{i,1}(\sigma^2_k) \),
\( \mu_{i,2}(\xi_{\text{max}}) \), \( \mu_{i,3}(C_{\text{moy}}) \), and \( \mu_{i,4}(N_{\text{max}}) \). They are then combined by the minimum
operator [27]:

\[
\forall i \in \{1 \cdots 4\}, \quad \forall x = (\sigma^2_k, \xi_{\text{max}}, C_{\text{moy}}, N_{\text{max}})
\mu_i(x) = \min\{\mu_{i,1}(\sigma^2_k), \mu_{i,2}(\xi_{\text{max}}), \mu_{i,3}(C_{\text{moy}}), \mu_{i,4}(N_{\text{max}})\}. \tag{11}
\]

As usual in fuzzy classification framework, the definition of our individual mem-
bership functions \( \mu_{i,j} \) makes use of shifted exponential functions which are trun-
cated at 1. Let

$$
\phi_{\tau, \nu}(x) = \min\{1, \exp[\tau (x - \nu)]\},
$$
(12)

$$
\psi_{\tau, \nu}(x) = \min\{1, \exp[\tau (\nu - x)]\},
$$
(13)

where \( \tau \) and \( \nu \) are two positive parameters. Note that \( \psi_{\tau, \nu} \) is the symmetric of \( \phi_{\tau, \nu} \) with respect to \( x = \nu \) axis, whereas \( \psi_{\tau, 0} \), which we shall also use, is the symmetric of \( \phi_{\tau, \nu}(x) \) with respect to \( x = \frac{\nu}{2} \) axis (see Fig. 3).

We know review for each feature the pieces of prior knowledge one can simply formulate about each of the four classes:

- **Contribution of \( \sigma^2 \).** In case of ripples (label \( w_1 \)) or dunes of sand (label \( w_2 \)), cast shadows have a privileged orientation, contrary to the ones of pebbles (label \( w_3 \)) or rocks (label \( w_4 \)) whose orientations are equally random. Therefore, we define for parameter \( \sigma^2 \) the following membership functions

$$
\mu_{1,1}(\sigma^2) = \mu_{2,1}(\sigma^2) = \psi_{a,0}(\sigma^2),
$$
(14)

$$
\mu_{3,1}(\sigma^2) = \mu_{4,1}(\sigma^2) = \phi_{a,b}(\sigma^2),
$$

where \( a \) and \( b \) are two positive parameters.

- **Contribution of \( \xi_{max} \).** Cast shadows associated to ripples (label \( w_1 \)) and dunes of sand (label \( w_2 \)) exhibit stretched shapes, by contrast with the shadows cast by pebbles (label \( w_3 \)) or rocks (label \( w_4 \)). Therefore, we define for parameter \( \xi_{max} \)

$$
\mu_{1,2}(\xi_{max}) = \phi_{c,d}(\xi_{max}),
$$

$$
\mu_{2,2}(\xi_{max}) = \phi_{c,e}(\xi_{max}),
$$

$$
\mu_{3,2}(\xi_{max}) = \mu_{4,2}(\xi_{max}) = \psi_{c,d}(\xi_{max}),
$$
(15)

where \( c, d, \) and \( e \) are three positive parameters. The shadows cast by ripples being thinner than those of the dunes, one should set \( d > e \).

- **Contribution of \( C_{moy} \).** Both ripples (label \( w_1 \)) and dunes (label \( w_2 \)) of sand cast thin and complex shadows, whereas those generated by pebbles (label \( w_3 \)) and rocks (label \( w_4 \)) exhibit simple compact circular-like shapes. Therefore, we define for parameter \( C_{moy} \)

$$
\mu_{1,3}(C_{moy}) = \mu_{2,3}(C_{moy}) = \psi_{f,0}(C_{moy}),
$$

$$
\mu_{3,3}(C_{moy}) = \mu_{4,3}(C_{moy}) = \phi_{f,g}(C_{moy}),
$$
(16)

where \( f \) and \( g \) are positive parameters.
• Contribution of $N_{\text{max}}$. The size of the shadows cast by ripples (label $w_1$) and dunes (label $w_2$) may vary dramatically from one window to another. For that reason we set that the membership degrees for these two classes as independent from $N_{\text{max}}$ parameter. For the two other classes, it is a discriminating parameter since rock (label $w_4$) shadows are larger than the ones created by pebbles (label $w_3$). We define for parameter $N_{\text{max}}$

$$
\begin{align*}
\mu_{1,4}(N_{\text{max}}) &= \mu_{2,4}(N_{\text{max}}) = 1 \\
\mu_{3,4}(N_{\text{max}}) &= \psi_{h,i}(N_{\text{max}}) \\
\mu_{4,4}(N_{\text{max}}) &= 1 - \psi_{h,i}(N_{\text{max}})
\end{align*}
$$

(17)

where $h$ and $i$ are two positive parameters.

For each feature parameter, we plot in Fig. 4 the four associated membership functions (with the parameter values used in the experiments. See Section 5). As can be readily seen from these plots, “ripples” and “dunes” classes are rather similar, being only discriminated from each other via the elongation parameter $\xi_{\text{max}}$. Similarly, “rocks” and “pebbles” classes are very much alike, apart from the point of view of $N_{\text{max}}$ size parameter. The combination of all component-wise membership functions via (11) will hopefully allow to discriminate each class from the others. A qualitative description of how this fuzzy discrimination process should work is obtained by establishing the output of the classifier (which assigns to the $k$-th window the class $w_i$ such that $i = \arg\max_{j\in\{1\ldots4\}} \mu_j(x^k)$) in case of gross discretization of the feature space. The range of variation of $\sigma_\alpha^2$, $C_{\text{mgy}}$, and $N_{\text{max}}$ being split into three parts (“low” (L), “medium” (M), and “high” (H)), and the one of $\xi_{\text{max}}$ being split into four parts (“low” (L), “medium-low” (ML), “medium-medium” (MM), “medium-high” (MH), and “high” (H)), the classification result associated to each cell of this partition of the feature space is indicated in Table 1.\footnote{We thank the anonymous referee who established this table.}

The fuzzy classifier we end up with implies nine parameters $a$, $b$, $c$, $d$, $e$, $f$, $g$, $h$, and $i$. It might seem at first sight that the tuning of so many parameters should be an intricate, if possible, task. It turns out they can be easily calibrated as follows:

• In both $\phi_{\tau,\nu}$ and $\psi_{\tau,\nu}$ functions, parameter $\nu$ is of the same nature as the concerned variable $x$. The parameters of this type, which are $b$, $d$, $e$, $g$, and $i$, are some sorts of thresholds which can be heuristically tuned based on our prior knowledge of the classification problem. See Section 5 for the values we selected
for each of these five parameters, using simple considerations on expected features within each class.

- In both $\phi_{r,\nu}$ and $\psi_{r,\nu}$ functions, $\tau$ is a parameter with no physical meaning which tunes the slope of the exponential parts of the functions. When both functions $\phi_{r,\nu}$ and $\psi_{r,0}$ are used in conjunction as membership degrees of two competing fuzzy sets, they intercept each other at $x = \frac{\nu}{2}$, with a common membership degree of $\exp\left\{-\frac{\nu x}{2}\right\}$. To prevent them from overlapping too much, one should keep this value rather small. A good rule of thumb is, if $\nu$ is already selected, to fix $\tau = \nu^{-1}$. This concerns the parameters $a$ and $f$. As for remaining parameters (namely $c$ and $h$) we noticed experimentally that they can be tuned imprecisely without much impact on the performances, provided that they correspond to sufficiently steep slopes.

One has also to find a good compromise for the size of the sub-images involved in the computation of feature vectors $\mathbf{x}^k$. On the one hand, small windows would result into fine resolution classifications. On the other hand, shadow-based features are more reliably computed on larger windows. Besides, too small windows do not allow to account for large cast shadows such as those cast by large dunes of sand (see example in Figure 9). To circumvent this difficulty, we devised a multiscale classification process working on two different sizes of windows. For larger windows, which are first considered, we only look at the regions detected as dunes by the fuzzy classifier. This information is then passed to the finer windows (by duplication), and the fuzzy classifier is only run on remaining windows (cf. Figure 5).

4. SEGMENTATION STEP

In order to obtain a more accurate classification, contextual information (i.e., relationship between features computed on adjacent sub-images) has to be taken into account. To this end, we resort to Markov random field models [4] which allow the specification of such spatial dependencies by means of a proper probability distribution on the segmentation configuration set. More precisely, this Markovian framework allows us to combine a simple spatial statistical prior (about the regularity of the classification map) with the fuzzy classifier previously defined. Note that such a combination of fuzzy classification with MRF formalism has been proposed in a different way (and in the different context of radar image segmentation) by Salzenstein and Pieczynski [19].

The use of Markovian formalism requires to see the unknown class labels as random variables with values in discrete state space $\Lambda = \{w_1, w_2, w_3, w_4\}$. Let
\[ Y = \{Y_s, s \in S\} \] be the resulting label field, where \( Y_s \) is the label random variable associated to window \( s \), and \( S \) stands for the window lattice. A configuration of the label field is denoted as \( y = \{y_s, s \in S\} \), and the set of all possible configurations is \( \Omega = \mathcal{S}^A \).

For each window, a label has already been provided by the fuzzy classifier described in previous section. Let \( y_s^0 \in \Lambda \) be the label thus assigned to window \( s \). Let also \( z_s \in [0, 1] \) be the corresponding membership degree, i.e., if \( s \) is the \( k \)-th window, then \( z_s = \max_{j \in \{1, \ldots, 4\}} \mu_j(x^k) \). We thus have two sets of “observations”, \( y^0 = \{y_s^0, s \in S\} \) and \( z = \{z_s, s \in S\} \).

In this probabilistic framework, we now have to define (and to compute) the “best” classification configuration \( \hat{y} \) given \( y^0 \) and \( z \). There are various ways to define this configuration. A simple and popular way consists in defining it as the most probable configuration knowing the observations. This so-called maximum a posteriori (MAP) inference is thus defined as

\[
\hat{y} = \arg \max_{y \in \Omega} P(Y = y | y^0, z),
\]

where the posterior distribution \( P(Y = y | y^0, z) \) has first to be specified. Under Markovian assumptions, this distribution on a huge number of variates factorizes into “small pieces”, i.e., it amounts to a product of functions that only depend on a few “neighboring” variates at a time. Equivalently, we want to specify a so-called Gibbs distribution \( P(.) \propto \exp\{-U(.)\} \) whose energy function \( U \) splits into a sum of local interaction potentials which depend on a few “neighboring” random variables at a time. Such a distribution will be simple to specify (via the definition the local potentials), and easy to use on a local basis thanks to the conditional independencies that derive from its factorization.

Our goal in devising the energy associated to \( P(Y | y^0, z) \) is twofold. We first want the MAP estimate \( \hat{y} \) to be close to the preliminary fuzzy classification \( y^0 \). At the same time we would like \( \hat{y} \) to exhibit regions that are not too small and with rather smooth boundaries. This goal is hopefully achieved by defining

\[
U(y; y^0, z) = \sum_{s \in S} z_s [1 - \delta(y_s, y_s^0)] + \beta \sum_{<s,t>} [1 - \delta(y_s, y_t)],
\]

where \( \delta \) stands for the Kronecker delta function, \( \beta \) is a positive parameter which tunes the relative importance of each of the two energy terms \( U_1 \) and \( U_2 \), and the second sum is taken over all pairs of neighboring windows for the second-order neighborhood system on grid \( S \) (See Fig. 6).
The first term of energy, $U_1$, favors all the more the identity between any label $y_s$ and the corresponding fuzzy label $y^0_s$ than the confidence within this fuzzy label (measured in terms of membership degree) is high. The second term, $U_2$, corresponds to so-called Potts prior model which is extensively used in MRF-based segmentation techniques. It favors all the more a segmentation than the total length of inter-class boundaries that the segmentation contains is small. As a consequence, it discourages segmentations with isolated labels, and those with complex frontiers between regions.

Setting $P(Y = y | y^0, z) \propto \exp\{-U(y; y^0, z)\}$ with some given parameter $\beta$, the MAP inference then amounts to\(^2\)

$$
\hat{y} = \arg\min_{y \in \Omega} \sum_{s \in S} z_s [1 - \delta(y_s, y^0_s)] + \beta \sum_{s,t} [1 - \delta(y_s, y_t)].
$$

(20)

This global minimization problem is extremely difficult since it is set in a huge discrete set. It could be handled with a stochastic iterative algorithm (simulated annealing) based on the sampling of the distribution proportional to $\exp\{-U(y^0, z)\}$, with $T$ being a decreasing “temperature” parameter [12]. For computation time reasons, we preferred to use a deterministic counterpart known as the iterated conditional mode (ICM) algorithm [4]. This algorithm, which is composed of a succession of component-wise minimizations, converges to a local minima which depends on the initialization. As an initial configuration, we chose the fuzzy classification $y^0$ itself, that is the minimizer of $U_1$.

A last issue concerns the tuning of parameter $\beta$. As demonstrated in [11], the precise value of this parameter does not matter a lot: within certain ranges of variation, different values of $\beta$ yield the same inference results. One thus merely faces a problem of calibration rather than a problem of precise estimation. In this study, we chose to perform this calibration manually. Note however that precise estimation methods can be devised, such as those based on the expectation-maximization (EM) techniques [7] or on the iterated conditional expectation (ICE) techniques [18, 15].

\(^2\)One could legitimately wonder about the usefulness, as well as the statistical relevance, of a distribution which is chosen in a somewhat ad-hoc way and is only used to set the inference problem as a global minimization problem. One can indeed get rid of the statistical aspects, and simply set the inference problem as the global minimization of an ad-hoc objective function which splits conveniently into local terms. We nevertheless decided to stick to the (apparently superfluous) Markovian interpretation for it constitutes in our opinion a rich basis for further statistical treatments, including the estimation of parameters with EM or ICE techniques as evoked at the end of this section.
5. EXPERIMENTAL RESULTS

To validate our method for automatic sea-floor classification, we have carried out experiments with numerous images delivered by different high resolution sidescan sonar systems. Those presented in this Section are only a few examples. Sonar images presented in Figures 7, 8, 9, 10 and 11 are provided by a military sidescan sonar, namely the DUBM41 whose frequency is around 500 KHz. We have no technical precision about the sidescan sonar which has provided the sea-floor images presented in Figures 12 and 13. Note that all these images are quite large, covering from one to several thousands of square meters, and that they correspond to a variety of sea-floors and of acquisition conditions. They thus should allow a fair assessment of the performances of the technique, and of its robustness with respect to the tuning of the parameters.

For all the results reported, we consider windows of $64 \times 64$-pixel for the fine classification. A smaller size of window would lead to finer grain segmentations, but at the risk of loosing robustness at the fuzzy classification level. Besides, for cartography application, the obtained accuracy level is sufficient, as $64 \times 64$-pixel windows amount approximately to $6 \times 6$-meter areas. For the two-level classification strategy described at the end of Section 3, we first used $128 \times 128$ windows to compute the preliminary coarse grain classification.

Based on the discussion on parameter calibration in Section 3, the parameters of the fuzzy classifier were tuned as follows. Expecting that shadows of ripples and dune stick closely to their privileged orientation, we chose $b = 0.1$. We then set $a = b^{-1} = 10$. The typical elongation of ripple and dune shadows have been visually assessed, leading to $d = 7$, and $e = 5$. The typical compactness of rock and pebble shadows has been evaluated yielding $g = 0.2$. We then set $f = g^{-1} = 5$. Finally, the length threshold beyond which compact shadows should be assigned to rocks rather than to pebbles was set to $i = 60$. The two remaining parameters $c$ and $h$ were simply set to 1. As for the unique parameter of the energy-based segmentation model, we chose $\beta = 0.2$. This value has been selected empirically after a set of experiments on our database of real sonar images. It has, in all cases, provided us with a satisfactory regularization of the initial fuzzy classification. Note that with this value the two terms of the energy are of same order.

Figures 7, 8, 9, 10, 11, 12 and 13 represent sea-floor images provided by high resolution sidescan sonars. The classification results obtained with the method we have introduced are superimposed on these images using the following code: an empty window stands for the “sand” class, a window with a small square inside
stands for the “pebbles” class, a bigger square stands for the “rocks” class, a straight line stands for the “ripples” class, and two parallel lines stand for the “dunes” class. Some of these sonar images exhibit only one type of sea-floor (as in Figures 9 and 10), whereas the others combine several types of sea-bed. *Without changing the values of the parameters* we got good results on all these images, as assessed by the sonar experts with whom we are working.

In the results of Figure 9, some nice features of the method are highlighted. The combination of versatile fuzzy classifiers, of Markovian regularization, and of two-level hierarchical classification allows the procedure to correctly classify all windows as “dunes” despite the dramatic variability of the shadows (in both shape and size) cast by the dunes present in this image. As concerns more particularly the multi-window aspect, some of the 64 × 64 windows contain either only shadows or only sand. In both cases, if larger windows would not have been used in a first stage, these 64 × 64 windows would have been labeled with “sand” class for they do not exhibit any shadow contours.

A look at Figure 13 further demonstrates the impact of the Markovian *a priori* model described in Section 4. In the classification obtained with the fuzzy classifier alone, a number of windows are obviously mislabeled in the large rock area (Fig. 13a). These spurious classifications are removed as a result of the regularized segmentation, providing a much more correct extraction of the zone of rocks.

Despite the ability of the fuzzy classifier to handle mixed class windows, there still remain errors at the boundaries between zones of different sea-floors (cf. Figure 13b in which some windows containing a mixture of rocks and ridges of sand have been classified either rocks or pebbles). Errors can also be sometimes noticed on the border the input sonar image (cf. Figure 12) due to the lack of contextual information. Nevertheless, experimental results demonstrate the accuracy and efficiency of such a contextual fuzzy segmentation and classification scheme as well as its capability to deal with images from different sonar systems.

The whole classification procedure takes between 10 and 15 seconds on a standard 43P IBM (120 MHz) Unix workstation for a sonar image of size 768 by 512 pixels (e.g., Figures 9, 10, 11). This time does not include the computational time required for the preliminary unsupervised segmentation into two classes (shadows and reverberations areas) whose performances are reported in [15, 17].

6. CONCLUSION
In this paper, we have presented an original approach to sea-floor classification problem. It is based on a window-wise classification of cast shadows, which are extracted beforehand, using a combination of fuzzy logic and Markovian modeling. The fuzzy component of the technique captures in a flexible way simple knowledges on the shapes of the shadows associated to each type of sea-floor. The Markovian part of the technique consists in setting the final classification as the global minimizer of an objective function which combines the fuzzy pre-classification with a standard regularization prior.

The proposed scheme appears as an appealing alternative to classical texture-based neural or statistical classification approaches. It offers the following attractive features:

- This method does not work directly on the input sonar image but on a shadow detection map. This original characteristic provides the method with a first source of robustness with respect to the type of sidescan sonar. The appearance of the shadows cast by each type of sea-floor is indeed quite independent of the precise characteristics of the high resolution sidescan sonar under concern, and of the conditions of acquisition.

- The fuzzy classification allows us to combine various qualitative priors on features of different natures and to deal with classes whose boundaries in the feature space are not clearly defined (due the imprecision in the definition itself of these classes, and due to the fact that we often deal with mixtures of classes).

- The energy-based classification makes all the window-based fuzzy classifiers cooperate with each other in an efficient way via simple local interactions. This allows one to get rid of isolated spurious classifications and to get more correct boundaries between different types of sea-floor (in the limit of the resolution associated to the selected window size).

- Although numerous, the involved parameters can be easily tuned with no need for a real and heavy training. Such an approximate calibration has proved sufficient to cope with various images provided by different high resolution sidescan sonars.

The method has been validated on a number of large images provided by different high resolution sidescan sonars, under various conditions, and over a variety of seabeds. We thus have demonstrated the robustness and the practicability of the method since, with a single set of parameter values, we got good results on all these images, as assessed by the sonar experts with whom we are working. Being
both robust and fast, this technique provides an interesting tool for processing in an automatic way massive amounts of high resolution sonar data.

This study could now be extended to deal with a larger class of sonar images. The proposed technique is indeed specifically devised to classify high resolution sidescan images in terms of a fixed nomenclature of five classes. One can imagine different nomenclatures, more or less detailed than the one we introduced, depending on the aimed application, and on the type of sonar under concern. Remaining in the case of high resolution sidescan sonar images, one could for instance seek different types of sand ripples. More important, in case of other sonar techniques (which we did not consider in this study), such as mono- or multi-beam echosounders, or hull sonars, the resolution and the experimental contexts are significantly different from those of sidescan sonars. The types of sea-floor that can be discriminated from the images obtained with these various techniques are different, and the appearance of a given type of sea-floor may vary drastically from one technique to another. In order to cope with such a variety of situations, new versions of our approach should be devised, in terms of classification nomenclature, shape parameters, fuzzy membership functions, and parameter values. A further step would then to make the resulting general model adapt itself, as automatically as possible, to the type of images under concern.

REFERENCES


TABLE 1
Sketchy output of the fuzzy classifier for a “qualitative” quantization of the range of the four parameters $\sigma^2_a$, $C_{m_{a,y}}$, $N_{m_{a,z}}$, and $\xi_{m_{a,z}}$ (L=“low”, ML=“medium-low”, M=“medium”, MM=“medium-medium”, MH=“medium-high”, and H=“high”).

| $\sigma^2_a$ | $\xi_{m_{a,z}}$ | $C_{m_{a,y}}$ | $N_{m_{a,z}}$ | Class | $\sigma^2_a$ | $\xi_{m_{a,z}}$ | $C_{m_{a,y}}$ | $N_{m_{a,z}}$ | Class | $\sigma^2_a$ | $\xi_{m_{a,z}}$ | $C_{m_{a,y}}$ | $N_{m_{a,z}}$ | Class |
|-------------|----------------|--------------|--------------|-------|-------------|----------------|--------------|--------------|-------|-------------|----------------|--------------|--------------|-------|-------|
| L           | L              | L            | L            | Pe    | M           | L              | L            | L            | Pe    | H           | L              | L            | L            | L     | Pe    |
| L           | L              | L            | H            | Ro    | M           | L              | L            | H            | Ro    | H           | L              | L            | H            | L     | Ro    |
| L           | M              | L            | Pe           | M     | M           | L              | M            | L            | Pe    | H           | L              | M            | L            | Pe    |       |
| L           | M              | M            | H            | Ro    | M           | L              | M            | H            | Ro    | H           | L              | M            | H            | Ro    |       |
| L           | H              | L            | Pe           | M     | M           | L              | L            | Pe           | H     | ML          | L              | Pe           |       |       |       |
| L           | H              | L            | H            | Ro    | M           | ML             | L            | H            | Ro    | H           | ML             | L            | H            | Ro    |       |
| L           | M              | M            | L            | Pe    | M           | ML             | M            | L            | Pe    | H           | ML             | M            | L            | Ro    |       |
| L           | M              | H            | Ro           | M     | ML          | L              | H            | Ro           | H     | ML          | L              | H            | Ro           |       |       |
| L           | M              | H            | Du           | M     | MM          | L              | L            | Du           | H     | MM          | L              | Pe           |       |       |       |
| L           | H              | H            | Du           | M     | MM          | L              | H            | Du           | H     | MM          | L              | H            | Ro           |       |       |
| L           | M              | M            | Du           | M     | MM          | M              | L            | Du           | H     | MM          | M              | L            | Pe           |       |       |
| L           | M              | H            | Du           | M     | MM          | M              | H            | Du           | H     | MM          | M              | H            | Ro           |       |       |
| L           | M              | H            | Ri           | M     | MH          | L              | L            | Ri           | H     | MH          | L              | L            | Pe           |       |       |
| L           | M              | H            | Ri           | M     | MH          | L              | H            | Ri           | H     | MH          | L              | H            | Ro           |       |       |
| L           | M              | M            | Ri           | M     | MH          | M              | L            | Ri           | H     | MH          | M              | L            | Pe           |       |       |
| L           | M              | H            | Ri           | M     | MH          | M              | H            | Ri           | H     | MH          | M              | H            | Ro           |       |       |
| L           | M              | H            | L            | Pe    | M           | MH             | L            | L            | Pe    | H           | MH             | L            | Pe           |       |       |
| L           | H              | H            | Ro           | M     | MH          | H              | H            | Ro           | H     | MH          | H              | H            | Ro           |       |       |
| L           | H              | L            | Ri           | M     | MH          | L              | L            | Ri           | H     | H           | L              | L            | Ri           |       |       |
| L           | M              | L            | Ri           | M     | MH          | M              | L            | Ri           | H     | M           | M              | L            | Ri           |       |       |
| L           | H              | M            | H            | Ri    | H           | M              | H            | Ri           | H     | M           | M              | Ri           |       |       |       |
| L           | H              | H            | L            | Pe    | M           | H              | H            | L            | Pe    | H           | H              | L            | Pe           |       |       |
| L           | H              | H            | H            | Ro    | M           | H              | H            | H            | Ro    | H           | H              | H            | Ro           |       |       |
FIGURES

FIG. 1. Formation of acoustic shadows in sidescan sonar imagery.

FIG. 2. Overview of the sea-bed classification scheme.
FIG. 3. Plot of the truncated exponential functions $\phi_{\tau, \nu}$, $\psi_{\tau, \nu}$, and $\psi_{\tau, 0}$ ($\tau = 1$, $\nu = 1$).
FIG. 4. Plot of the parameter-wise membership functions $\mu_{j,1}(\sigma^2_{\text{m}})$, $\mu_{j,2}(\xi_{\text{m}})$, $\mu_{j,3}(C_{\text{m,xy}})$, and $\mu_{j,4}(N_{\text{m,ee}})$, $j = 1 \cdots 4$, for parameter values $a = 10$, $b = 0.1$, $c = 1$, $d = 7$, $e = 5$, $f = 5$, $g = 0.2$, $h = 1$, and $i = 60$.

FIG. 5. The multiscale classification strategy.
FIG. 6. Different types of pairs of neighboring blocks for the second-order neighborhood system on window grid $S$.

(a) real sidescan image of a sea-bed with sand and dunes (b) hierarchical two-class segmentation (shadow vs. reverberation) obtained by the method introduced in [15,17]; (c) contours of the cast shadows extracted at the finest level; (d) window-wise classification obtained with the complete method described in the paper (empty windows stand for “sand”, and windows marked with parallel lines stand for “dunes”).
FIG. 8. Classification of a sidescan sonar image (24 × 24-meter sea-floor area) including sand (empty windows), pebbles (windows with small squares inside), and rocks (windows with bigger squares inside).

FIG. 9. Classification of a sidescan sonar image (72 × 48-meter sea-floor area) including only dunes of sand (windows marked with two parallel segments).
FIG. 10. Classification of a sidescan sonar image (72 x 48-meter sea-floor area) including only ripples of sand (windows marked with one segment of line).

FIG. 11. Classification of a sidescan sonar image (72 x 48-meter sea-floor area) including sand (empty windows) and pebbles (windows with small squares inside).
FIG. 12. Classification of a sidescan sonar image (42 × 54-meter sea-floor area) including sand (empty windows), rocks (windows with squares inside), and dunes (windows marked with parallel line segments).
FIG. 13. Classification of a sidescan sonar image (42 x 78-meter sea-floor area) including sand, ripples, rocks, and pebbles: (a) result obtained by the fuzzy classifier alone; (b) final result obtained by adding the MRF-based regularization.