Computer Vision and Image Understanding
Vol. 76, No. 3, December, pp. 191–204, 1999
Article ID cviu.1999.0804, available online at http://www.idealibrary.com on IDEAL®

# Three-Class Markovian Segmentation of High-Resolution Sonar Images

M. Mignotte and C. Collet

*Groupe de Traitement du Signal, École Navale, B.P. 600, 29240 Brest-Naval Cedex, France*
E-mail: collet@poseidon.ecole-navale.fr

and

P. Pérez and P. Bouthemy

*IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France*
E-mail: perez@irisa.fr

Received April 15, 1998; accepted September 15, 1999

**This paper presents an original method for analyzing, in an unsupervised way, images supplied by high resolution sonar. We aim at segmenting the sonar image into three kinds of regions: echo areas (due to the reflection of the acoustic wave on the object), shadow areas (corresponding to a lack of acoustic reverberation behind an object lying on the sea-bed), and sea-bottom reverberation areas. This *unsupervised* method estimates the parameters of noise distributions, modeled by a Weibull probability density function (PDF), and the label field parameters, modeled by a Markov random field (MRF). For the estimation step, we adopt a maximum likelihood technique for the noise model parameters and a least-squares method to estimate the MRF prior model. Then, in order to obtain an accurate segmentation map, we have designed a two-step process that finds the shadow and the echo regions separately, using the previously estimated parameters. First, we introduce a scale-causal and spatial model called SCM (scale causal multigrid), based on a multigrid energy minimization strategy, to find the *shadow* class. Second, we propose a MRF monoscale model using *a priori* information (at different level of knowledge) based on physical properties of each region, which allows us to distinguish echo areas from sea-bottom reverberation. This technique has been successfully applied to real sonar images and is compatible with automatic processing of massive amounts of data.** 

*Key Words:* **sonar imagery; unsupervised segmentation; MRF hierarchical model; Weibull law; noise model estimation.**

## 1. INTRODUCTION

In sonar imagery, three kinds of regions have to be identified: echo, shadow, and sea-bottom reverberation areas. The echo information is caused by the reflection of the acoustic wave on the object, while the shadow zone corresponds to a lack of acoustic reverberation behind an object. The remaining information is called the sea-bottom reverberation area. On the pictures supplied by sonar system, the echo features are generally less discriminant than the shadow shapes for the classification of objects lying on the sea bed. For this reason, the detection of each object located on the sea bottom and its classification (as a wreck, a rock, a man-made object, etc.) is generally based on the extraction and the identification of its associated cast shadow [1]. Nevertheless, the echo information may be necessary to detect and then classify objects partially buried in the sea floor. Indeed, in this case, the echo features are the only information that can be used to identify the nature of the detected objects. Let us note that a three-class segmentation can also be interesting in other applications, like medical ultrasound imagery [2], or three dimensional reconstruction of underwater objects [3].

Few studies describe complete approaches allowing such a segmentation of sonar images to be performed automatically, with results that can be efficiently used afterward for object identification [4, 5]. Some of them are based on simple, and often ad-hoc, clustering techniques (such as fuzzy $K$-means) working on luminance mean and variance within small windows [6, 7]. In that case, only a coarse grain classification is obtained, and no statistic modeling of the back-scattered acoustic amplitude within the different types of regions is introduced. This latter aspect induces a lack of robustness for this type of approaches. Nevertheless, some studies include more advanced element of image formation modeling [8–10]. But to our knowledge, no paper in the literature proposes an unsupervised 3-classes segmentation of high resolution sidescan sonar picture entirely based on statistic analysis.

The segmentation task of these images is made difficult by the presence of speckle noise [11]. This peculiar noise makes ineffective any simpler segmentation schemes. In order to extract a reliable and accurate segmentation map, contextual information is important to be taken into account in sonar imagery. This can be done a *posteriori*, using either morphological filters in order to "clean" the classification obtained by a simple clustering technique [6, 7]. This can also be done *a priori*, using MRF models which are appropriate to specify spatial dependencies by *a priori* label field distribution [12]. In the sonar imagery context, a



1077-3142/99 $30.00

MRF-based model is used in [8] but this model is supervised and its monoscale modeling does not ensure a reliable regularization procedure.

To ensure efficient regularization of the set of labels when the sonar image contains strong speckle noise [13], hierarchical MRF models can be used. Among such models, we can cite the sequential maximum a posteriori (SMAP) algorithm introduced by Bouman *et al.*, in which each scale is causally dependent on the preceding coarser one [14], or the hierarchical model introduced by Kato *et al.*, where a pyramidal structure involving a three-dimensional (3D) neighborhood system and a 3D Markovian label field is considered [15]. We can also cite the scale-causal multigrid model that we have introduced in [16] for a two-class segmentation of sonar images (*shadow, reverberation*).

MRF hierarchical models enable the local characteristics of image content to be modeled more accurately thanks to *a priori* specification of spatial and hierarchical dependencies between neighboring sites. Nevertheless, these models usually do not exploit *a priori* knowledge of the spatial relationship between adjacent regions of different nature. In our sonar image context, such information relative to the interactions between different neighboring regions (especially between echo and shadow areas) will be exploited in the segmentation procedure.

We propose herein a hybrid approach defined as a two-step process, associating the efficiency of purely hierarchical modeling (a scale-causal multigrid model) with the spatial prior about relationship between adjacent regions of different nature. For parameter estimation, we adopt an iterative method called iterative conditional estimation (ICE) [17]. It allows simultaneous estimation of the MRF prior model parameters, according to the least squares estimator (LSQR) described by Derin *et al.* [18] (which we have generalized in [19]), and the noise model parameters, according to Maximum Likelihood (ML) estimators. In this paper, we further investigate the issue of noise modeling and we propose to model the distribution of the luminance (acoustic amplitude) within the reverberation and the shadow regions by a Weibull distribution which defines a set of parametrized probability density functions (PDF). Comparisons with a more simplistic noise model previously introduced in [16] is given. For the segmentation step, we use a two-step process which extracts shadow and echo regions separately, exploiting the previously estimated parameters. First, we use a scale-causal and spatial model SCM (scale causal multigrid) [16] along with a multigrid energy minimization strategy to separate the *shadow* class from the others (*echo* and *sea-bottom reverberation*). Second, we propose a MRF monoscale model using *a priori* information about the spatial dependency between each region and allowing us to distinguish between *echo* and *sea-bottom reverberation* labels.

This paper is organized as follows. In Section 2, we detail the parameter estimation step and demonstrate its application on both synthetic pictures and real sonar images. Section 3 describes the segmentation step and the proposed model. Segmentation results on both real and synthetic scenes are presented in Section 4. Section 5 contains concluding remarks.

## 2. ESTIMATION OF PARAMETERS

### *2.1. Problem Statement*

In the unsupervised three-class Markovian segmentation case, we have to estimate the noise model parameters (i.e., the noise distribution parameters associated to each region of the sonar image) and the *a priori* parameters of the Markov model.

In our case, estimation of noise distribution parameters is difficult because of the very small number (or sometimes the absence) of pixels associated to the echo region and by the lack of knowledge about the form of the appropriate conditional noise distribution describing the luminance within the echo region. For these reasons, we cannot estimate efficiently and simultaneously the noise distribution parameters of each region. To circumvent these difficulties and in order to obtain a reliable and proper noise model estimation ensuring an accurate segmentation map, an alternate approach is to decompose the segmentation stage in a two-step process:

- In a first step, the *sea-bottom reverberation* and the *echo* (due to the reverberation of the acoustic wave on the object) classes are not distinguished from a single class; the *reverberation* class, which is considered as capturing both types of regions. An unsupervised hierarchical Markovian segmentation into two classes (*shadow* and *reverberation*) is performed exploiting an ICE [17] estimation procedure which is presented in this section, along with the hierarchical Markovian modeling described in [16].
- In a second step, we perform a supervised Markovian segmentation of the region corresponding to the reverberated signal into two classes. This allows us to distinguish echo areas from sea-bottom reverberation area (see Section 3.2.2). In this step, we search for the echo associated to each shadow region. The block diagram of this unsupervised three-class segmentation scheme is shown in Fig. 1.

In the context of a two-class segmentation, we adopt the ICE procedure to estimate the noise model and the MRF prior model parameters simultaneously. This procedure is described in [17]. In a previous study, this estimation procedure was applied to sonar imagery in the case of a rough noise model [19]; for this noise model, we resorted to a Gaussian law to model the luminance distribution within shadow regions and a Rayleigh law to model the speckle noise phenomenon. Speckle noise modeling based on a Rayleigh law is generally suitable when the length of detectable sea-floor elements (i.e., sea-floor roughness) is much larger than the acoustic wavelength [20] (which is the common case with high-frequency imaging sonar). In the same vein, the Gaussian law seems to be an appropriate model to describe the luminance within the shadow region, if we consider that noise within such a region is essentially of electronic nature. Nevertheless, this noise model can be too simple to cope with

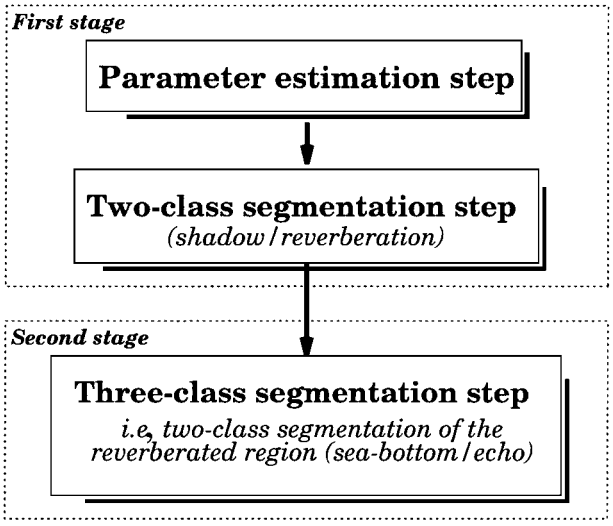



**FIG. 1.** Unsupervised three-class segmentation scheme.

all situations. In order to take into account the minor lobes of the acoustic antenna bringing back-scattered signal within the shadow region and the different sea-floor types likely to occur in sonar images, we investigate in this paper a more general noise model. A more appropriate probability density function (PDF) based on the Weibull distributions mixture is used to describe the luminance distribution within each region of the sonar image.

The ICE procedure is quite general and can be adapted to different kinds of conditional noise distributions [17]. The only condition is to get an estimator with good asymptotic properties (consistency properties) for completely observed data. As shown in the following, this method can be used efficiently to estimate a mixture of Weibull distributions, i.e., the Weibull distribution parameters associated to each region (shadow and reverberation) of the sonar image.

### 2.2. Iterative Conditional Estimation

Consider a couple of random fields $Z = (X, Y)$, where $Y = \{Y_s, s \in S\}$ represents the field of observations located on a lattice $S$ of $N$ sites $s$, and $X = \{X_s, s \in S\}$ the label field. Each $Y_s$ takes its value in $\Lambda_{obs} = \{0, \ldots, 255\}$, and each $X_s$ in $\{e_0 = \textit{shadow}, e_1 = \textit{reverberation}\}$. The distribution of $(X, Y)$ is defined by the prior distribution $P_X(x)$, assumed to be stationary and Markovian, and by sitewise likelihoods $P_{Y_s/X_s}(y_s/x_s)$, assumed to be Weibull PDFs:

$$P_{X,Y}(x, y) = P_X(x) \underbrace{\prod_s P_{Y_s/X_s}(y_s/x_s)}_{P_{Y/X}(y/x)}. \tag{1}$$

The observable $Y$ is called the *incomplete data*, and $Z$ the *complete data*. Let us note that prior distribution $P_X(x)$ depends on a parameter vector $\Phi_x$, while the conditional likelihood $P_{Y/X}(y/x)$ depends on parameter vector $\Phi_y$. Joint and posterior distributions $P_{X,Y}(x, y)$ and $P_{X/Y}(x/y) \propto P_X(x)P_{Y/X}(y/x)$ thus depends on $\Phi = (\Phi_x, \Phi_y)$.

In the unsupervised Markovian segmentation case, we have to estimate, in a first step (estimation step), parameter vectors $\Phi_x$ and $\Phi_y$. To this end, we resort to the ICE procedure as mentioned above [17]. This method relies on two estimators $\hat{\Phi}_x(X)$ and $\hat{\Phi}_y(X, Y)$ appropriate to completely observed data case. When $X$ is unobservable, this procedure starts from an initial parameter set $\Phi^{[0]}$ (not too far from the optimal parameters) and generates a sequence of parameter vectors $\Phi^{[1]}, \Phi^{[2]}, \ldots, \Phi^{[k]}$ hopefully leading to the optimal parameters ($\lim_{k\to\infty} \Phi^{[k]} = \Phi_{\text{optimal}}$). To this end, $\Phi_x^{[k+1]}$ and $\Phi_y^{[k+1]}$ at step $(k+1)$ are chosen as conditional expectations of $\hat{\Phi}_x$ and $\hat{\Phi}_y$ given $Y = y$, computed according to the current values $\Phi_x^{[k]}$ and $\Phi_y^{[k]}$. These are the best approximations of $\Phi_x$ and $\Phi_y$ in terms of the mean squared error [17]. By letting $E_k$ denote the expectation relative to parameters $\Phi^{[k]} = (\Phi_x^{[k]}, \Phi_y^{[k]})$, this iterative procedure is defined as follows:

- Consider an initial parameter set $\Phi^{[0]} = (\Phi_x^{[0]}, \Phi_y^{[0]})$.
- $\Phi^{[k+1]}$ is computed from $\Phi^{[k]}$ and $Y = y$ by

$$\Phi_x^{[k+1]} = E_k[\hat{\Phi}_x(X) \mid Y = y] \tag{2}$$

$$\Phi_y^{[k+1]} = E_k[\hat{\Phi}_y(X, Y) \mid Y = y]. \tag{3}$$

The computation of these expectations is impossible in practice, but we can approach Eqs. (2) and (3), thanks to the law of large numbers, by

$$\Phi_x^{[k+1]} = \frac{1}{n}\left[\hat{\Phi}_x\left(x_{(1)}\right) + \cdots + \hat{\Phi}_x\left(x_{(n)}\right)\right] \tag{4}$$

$$\Phi_y^{[k+1]} = \frac{1}{n}\left[\hat{\Phi}_y\left(x_{(1)}, y\right) + \cdots + \hat{\Phi}_y\left(x_{(n)}, y\right)\right], \tag{5}$$

where $x_{(i)}$, $i = 1, \ldots, n$ are realizations of $X$ drawn according to the posterior distribution $P_{X/Y,\Phi}(x/y, \Phi^{[k]})$. As explained below, for complete data-based estimator $\hat{\Phi}_y(X, Y)$, we use a maximum likelihood (ML) estimator for the noise model parameter, whereas $\hat{\Phi}_x$ is the least-squares (LSQR) estimator described by Derin *et al.* [18]. Finally, in order to use the ICE procedure, we need:

- An initial value $\Phi^{[0]}$ not *too far* from the optimal parameters.
- A way of simulating realizations of $X$ according to the posterior distribution $P_{X/Y,\Phi}$. To this end, we use the Gibbs sampler algorithm [21].

### 2.3. Estimation of the Noise Model Parameters for Complete Data

The Weibull Probability Density Function is an appropriate distribution to describe the luminance $y$ within the reverberation

and the shadow regions. This PDF offers degrees of freedom which allows a large variety of quite different distributions to be captured. Thus, this PDF is well adapted to sonar images where the speckle distribution is not exactly known and may vary according to experimental conditions. Experiments have demonstrated that this distribution models the speckle noise phenomenon more accurately than a Rayleigh distribution and achieves the best fit to real data (the Weibull PDF is used in other applications such as high resolution radars [22]; experiments in Section 2.4 will show its efficiency).

The Weibull PDF is a two-parameter distribution, of which the Rayleigh and the exponential distributions are special cases. It turns out that, for our application, we have to introduce another parameter to shift this distribution, in order to take into account the different processes forming the final sonar image (automatic control of gain, coding, reduction of the dynamic, offset, etc.) [23]. We propose to consider the expression

$$\mathcal{W}_Y(y; \min, C, \alpha) = \frac{C}{\alpha}\left(\frac{(y-\min)}{\alpha}\right)^{C-1} \exp\left(-\frac{(y-\min)^C}{\alpha^C}\right), \tag{6}$$

with $y > \min$, $\alpha > 0$, and $C > 0$. $\alpha$ and $C$ are the scale and shape parameters, respectively. Figure 2 represents different Weibull laws for several values of the shape parameter $C$. In this example, $\alpha = 1$ and $\min = 0$. For the special cases $C = 1$ and $C = 2$, the PDF corresponds to an exponential and a Rayleigh law respectively. Let $Y = (Y_1, Y_2, \ldots, Y_M)$ be $M$ random variables, independent and identically distributed according a *single* Weibull law $\mathcal{W}_Y(.; \Phi_y)$, and $y = (y_1, y_2, \ldots, y_M)$ a realization of $Y$. The Maximum Likelihood (ML) estimate of $\Phi_y = (\min, C, \alpha)$ consists in finding $\hat{\Phi}_y$, such that

$$\hat{\Phi}_y = \arg\max_{\Phi_y} \ln P_{Y/\Phi_y}(y/\Phi_y), \tag{7}$$

where $\ln P_{Y/\Phi_y}(y/\Phi_y)$ is the log-likelihood function. Assuming independence between each random variable, the log-likelihood function can be written

$$\ln \mathcal{L}(\Phi_y) = \ln P_{Y/\Phi_y}(y/\Phi_y) \tag{8}$$

$$= \ln\left\{\left(\frac{C}{\alpha^C}\right)^M \prod_{i=1}^{M}\left[(y_i - \min)^{c-1} \exp\left(-\frac{(y_i-\min)^C}{\alpha^C}\right)\right]\right\}, \tag{9}$$

with $y_i > \min \forall i$, $\alpha > 0$, and $C > 0$. The maximum value of the log-likelihood function is used to determine a ML estimator of the unknown parameters $\Phi_y = (\min, C, \alpha)$. Setting the partial derivatives of $\ln \mathcal{L}(\Phi_y)$ to zero with respect to each parameter (i.e., solving the system $[\partial \ln \mathcal{L}(\Phi_y)/\partial \Phi_y] = 0$) gives the ML estimators of the *complete data* (all samples are here known to arise from the same Weibull distribution). If $\hat{y}_{\min} = \min_i(y_i)$, is the minimum grey level of the sample $y$, we obtain the following result:

$$\widehat{\min}_{\mathrm{ML}} \approx \hat{y}_{\min} - 1. \tag{10}$$

Denoting now $\tilde{y}_i = (y_i - \widehat{\min}_{\mathrm{ML}})$ and setting to zero the partial derivative $[\partial \ln \mathcal{L}(\Phi_y)/\partial \alpha] = 0$ (which leads to solving the

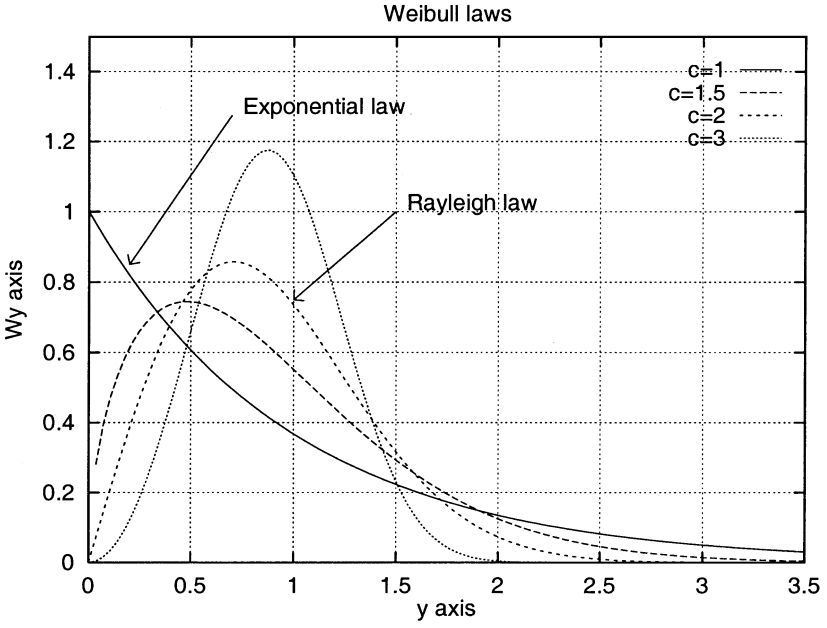



**FIG. 2.** Plot of Weibull laws for several values of $C$.

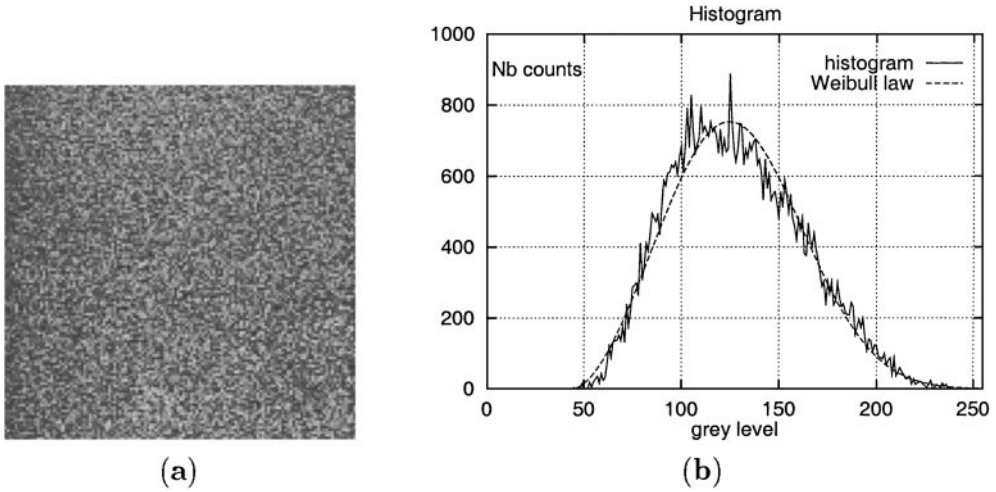



**FIG. 3.** (a) Real sonar image of a sandy sea floor. (b) Image histogram and drawing of the Weibull law with ML estimated parameters (see Table 1 for parameter values).

system $-M + (1/\alpha^C \sum_{i=1}^{M} \tilde{y}^C) = 0$), we then easily obtain $\hat{\alpha}_{\mathrm{ML}}$:

$$\hat{\alpha}_{\mathrm{ML}} = \left( \frac{1}{M} \sum_{i=1}^{M} \tilde{y}_i^{\hat{C}_{\mathrm{ML}}} \right)^{1/\hat{C}_{\mathrm{ML}}}. \tag{11}$$

Setting $\frac{\partial \ln \mathcal{L}(\Phi_y)}{\partial C}$ to zero and using Eq. (11) yields

$$\frac{\sum_{i=1}^{M} \left( \tilde{y}_i^{\hat{C}_{\mathrm{ML}}} \cdot \ln \tilde{y}_i \right)}{\sum_{i=1}^{M} \tilde{y}_i^{\hat{C}_{\mathrm{ML}}}} - \frac{1}{M} \sum_{i=1}^{M} \ln \tilde{y}_i = \frac{1}{\hat{C}_{\mathrm{ML}}}. \tag{12}$$

We have no analytic expression for $\hat{C}_{\mathrm{ML}}$. Nevertheless, Eq. (12) can easily be solved iteratively according to the following scheme. Expression (12) can be written as $F(\hat{C}_{\mathrm{ML}}) = \hat{C}_{\mathrm{ML}}$. Consider now some sequence of the form $U(1), U(2) = F(U(1)), \ldots, U(p) = F(U(p-1))$. If this sequence converges toward a limit $l$, then this limit $l$ is a solution of the equation $l = F(l)$. Inversely, if the equation $F(\hat{C}_{\mathrm{ML}}) = \hat{C}_{\mathrm{ML}}$ has only one solution (in our application, the uniqueness of the ML estimator of $C$ is obviously verified if $F(.)$ is a monotone function), then $\hat{C}_{\mathrm{ML}}$ is also the convergence value of any convergent sequence $U(p)$ defined by the recursion $U(p+1) = F(U(p))$. The convergence of this sequence is then ensured because $F(.)$ is proved to be a monotone function. Consequently, $\hat{C}_{\mathrm{ML}}$ is given by the following relation:

$$\hat{C}_{\mathrm{ML}} = \lim_{p \to \infty} U(p). \tag{13}$$

To speed up the convergence rate of this iterative estimation procedure, we have to take for the first term $U(0)$ of the sequence $U(p)$ a positive value not *too far* from $\hat{C}_{\mathrm{ML}}$. In our application, we take $U(0) = 1$, which corresponds to an exponential distribution (see Fig. 2). Experiments on both real and synthetic sonar images demonstrate that this iterative scheme allows a good approximation of $\hat{C}_{\mathrm{ML}}$ to be obtained quickly, which is then used in Eq. (11) to get $\hat{\alpha}_{\mathrm{ML}}$.

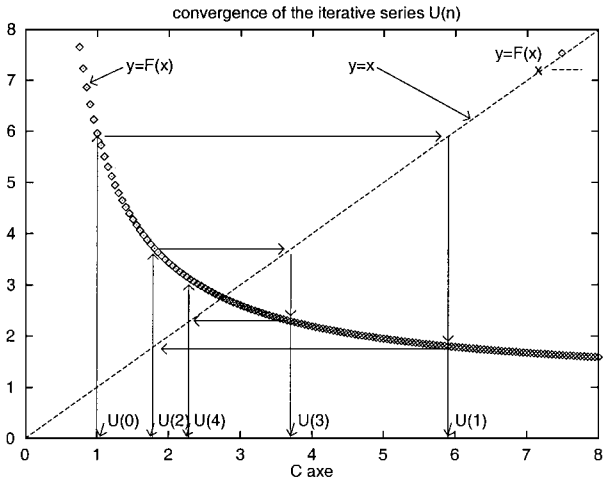
Figure 3.a represents a real sonar image of a sandy sea floor. The ML estimated parameters obtained with this scheme are given in Table 1. Figure 4 shows graphically the convergence of the iterative sequence $U(p)$ for the estimation of the shape parameter $\hat{C}_{\mathrm{ML}}$. The quality of the obtained estimates is difficult to appreciate in the absence of ground-truth values. Nevertheless, one can visually appreciate on Figure 3.b the good match between the image histogram and the PDF corresponding to the estimated parameters. Figure 5 presents a synthetic sonar image of a sandy sea floor in which the speckle noise distribution is a Weibull PDF with specified parameter vector $\Phi_y$ and ML estimates obtained with our scheme. We can notice that estimated parameters are close to the ground-truth parameter.

### *2.4. Parameter Estimation Procedure for the Incomplete Data*

Let us recall that this parameter estimation procedure assumes that the luminance within shadow regions and within reverberation regions follows two different Weibull PDFs. We aim at estimating the parameters of these two PDFs as well as the *a priori* parameters of the Markov model. For the *a priori* model, we adopt a standard anisotropic Potts model with the 8-connexity spatial neighborhood. There are four parameters $\beta_1, \beta_2, \beta_3, \beta_4$ associated to the horizontal, vertical, and right and left diagonal

**TABLE 1**
**ML Estimated Parameters for the Example Reported in Fig. 3**

| | | | |
|---|---|---|---|
| $\hat{\Phi}_y$ | $44_{(\widehat{\min})}$ | $2.75_{(\hat{C}_{\mathrm{ML}})}$ | $95_{(\hat{\alpha}_{\mathrm{ML}})}$ |




| $U(p)$ | $U(0)$ | $U(1)$ | $U(2)$ | $U(3)$ | $U(4)$ | $U(5)$ | $U(6)$ | $U(7)$ | $U(8)$ | $U(9)$ | $U(10)$ | $U(11)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimate | 1.00 | 5.96 | 1.79 | 3.72 | 2.28 | 3.12 | 2.53 | 2.90 | 2.65 | 2.82 | 2.71 | 2.78 |

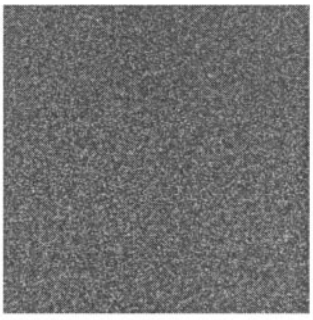
**FIG. 4.** Convergence of the iterative sequence $U(p)(\hat{C}_{\mathrm{ML}} = \lim_{p\to\infty} U(p))$ for the example reported in Fig. 3.

binary cliques, respectively,

$$P_X(x) \propto \exp\left\{-\sum_{\langle s,t\rangle} \beta_{st}(1-\delta(x_s, x_t))\right\} \tag{14}$$

where summation is taken over all pairs of neighboring sites, $\beta_{st} = \beta_1, \beta_2, \beta_3$, or $\beta_4$ only depends on the "orientation" of the clique, and $\delta(.)$ is the Kronecker delta function. The parameter estimation procedure for the incomplete data is outlined below:

• **Parameter initialization**: The initial parameter values have a significant impact on the rapidity of the convergence of the ICE procedure and on the quality of the final estimates. In our application, we use the initialization method described in [19]. This method aims at obtaining a rough two-class segmentation based on features extracted in subwindows and clustered



| Specified parameters ($\Phi_y$) and ML estimates ($\hat{\Phi}_y$) | | | |
|---|---|---|---|
| $\Phi_y$ | $49_{(min)}$ | $2.00_{(C)}$ | $70.7_{(\alpha)}$ |
| $\hat{\Phi}_y$ | $49_{(\widehat{min})}$ | $2.03_{(\hat{C}_{ML})}$ | $71.4_{(\hat{\alpha}_{ML})}$ |

**FIG. 5.** Synthetic sonar image of a sandy sea floor in which speckle noise distribution is a Weibull PDF with specified parameter $\Phi_y$ and ML estimates are obtained with our iterative scheme. In this example, $C = 2$ and the created synthetic noise follows also a Rayleigh law. We can appreciate the quality of the ML estimates; estimated parameters are close to the ground-truth parameters.

**TABLE 2**
**Estimated Parameters for the Picture Reported in Fig. 7.a (Mixture of Weibull Distributions)**

| | | | | |
|---|---|---|---|---|
| $\Phi^{\text{final}}_{y_{\text{shadow}}}$ | $0.06_{(\pi)}$ | $15_{(\min)}$ | $3.29_{(C)}$ | $26.8_{(\alpha)}$ |
| $\Phi^{\text{final}}_{y_{\text{reverberation}}}$ | $0.94_{(\pi)}$ | $42_{(\min)}$ | $1.83_{(C)}$ | $40.6_{(\alpha)}$ |
| $\Phi^{\text{final}}_{x}$ | $1.0_{(\beta_1)}$ | $1.5_{(\beta_2)}$ | $-0.2_{(\beta_3)}$ | $-0.3_{(\beta_4)}$ |

*Note.* $\pi$ stands for the proportion of the two classes within the sonar image. $\alpha$ and $C$ are respectively the scale and the shape parameters of the Weibull law. $\beta_i$ are the *a priori* parameters of the Markov model.

**TABLE 3**
**Kolmogorov Distance and $\chi^2$ Error Made by Approximating the Image Histogram (Reported in Fig. 6) by the Probability Density Mixtures Corresponding to the Estimation Based on Different Noise Models)**

| Noise model | Kolmogorov distance | $\chi^2$ error |
|---|---|---|
| Gauss–Rayleigh | 0.08 | 19,781 |
| Weibull–Weibull | 0.06 | 16,301 |

according to a $K$-means procedure. Once the $K$-means segmentation is obtained, the ML estimator of the *complete data* is used to get $\Phi_y^{[0]}$ (see Section 2.3). The initial parameters of the Gibbs distribution are obtained using the LSQR estimator described in [2] from the ML segmentation based on $\Phi_y^{[0]}$. Let us denote $\Phi^{[0]} = (\Phi_x^{[0]}, \Phi_y^{[0]})$, the obtained result.

- **ICE procedure**: $\Phi^{[k+1]}$ is computed from $\Phi^{[k]}$ in the following way:

  $\triangleright$ Using the Gibbs sampler, $n$ realizations $x_{(1)}, \ldots, x_{(n)}$ are simulated according to the posterior distribution $P_{X/Y,\Phi}(x/y, \Phi^{[k]})$, with the parameter vector $\Phi^{[k]}$.

  $\triangleright$ For each $x_{(i)}, i = 1, \ldots, n$, the parameter vector $\Phi_x$ is estimated by the algorithm proposed by Derin *et al.* [18, 19] and $\Phi_y$ with the ML estimator described in Section 2.3: they are denoted $\hat{\Phi}_x(x_{(i)})$, $\hat{\Phi}_y(x_{(i)}, y)$.

  $\triangleright$ $\Phi^{[k+1]}$ is obtained from $(\hat{\Phi}_x(x_{(i)}), \hat{\Phi}_x(x_{(i)}, y)), 1 \le i \le n$, by averaging these estimates (see Eqs. (4) and (5)).

If the sequence $\Phi^{[k]}$ becomes steady, the ICE procedure is ended and one proceeds to the actual two-class hierarchical segmentation using the estimated parameters [16]. We calibrate the weight of the "stochastic" aspect of the ICE procedure by choosing $n$, the number of realizations of $X$ simulated according to the posterior distribution $P_{X/Y,\Phi}$. When $n$ increases, the "stochastic" aspect of the algorithm decreases. The intentional choice of a small value for $n$ ($n = 1$ in our application) can increase its computation cost efficiency [24].

We can compare the quality and the aptness of this noise model based on Weibull laws over the one previously investigated [19, 23, 25]. In this preceding study, we have considered a rough noise model in which the speckle phenomenon is taken into account by a Rayleigh law and the luminance $y$ within *shadow* regions is described by a Gaussian law. The left part of Fig. 7.a represents a real sonar image of a sandy sea floor with the cast shadow of a man-made object (a cylinder). Figure 6 shows the estimated distribution mixture and the histogram of this sonar image for the different noise models. The quality of the estimations based on the Weibull model over the Gauss–Rayleigh model is difficult to appreciate visually in the absence of ground-truth values. We can roughly perform such an evaluation by comparing the image histogram (dashed curves) with the probability density mixture corresponding to the estimated parameters (dotted curves). We can see that the histogram is closer to the mixture densities based on the Weibull model, especially for the *reverberation* class. Experiments show that the estimate based on this model are also better according to the Kolmogorov distance, or $\chi^2$ criterion [26], than estimates based on the Gauss–Rayleigh models. Table 3 gives the error, in the Kolmogorov distance and $\chi^2$ criterion sense, made by these two noise models in approximating the image histogram. The estimates obtained by the ICE procedure (Weibull model) are given in Table 2.

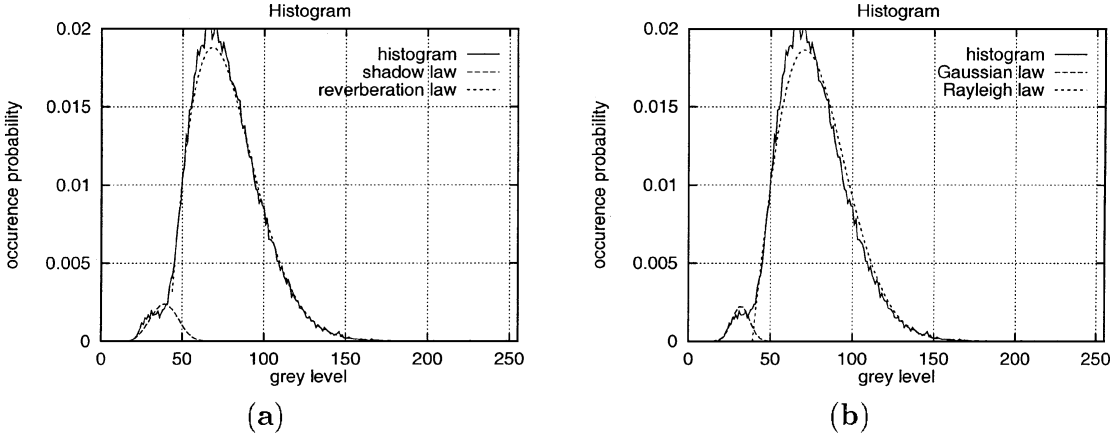



**FIG. 6.** Image histogram of the picture reported in Fig. 7.a and estimated probability density mixture with the ICE procedure. (a) Mixture of two Weibull distributions. (b) Mixture of different distributions (Gaussian law for the *shadow* class and a shifted Rayleigh law for the *reverberation* class) [16].

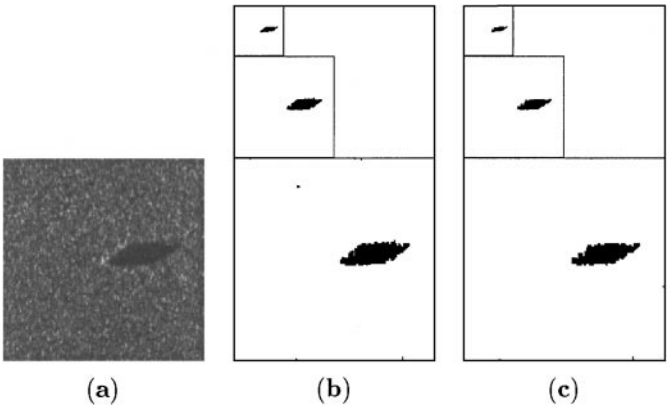



**FIG. 7.** (a) Real sonar image of a sandy sea floor with the shadow of a man-made object. SCM two-class segmentation results obtained with a (b) mixture of Weibull distributions; (c) mixture of different distributions (Gaussian law for the *shadow* class and a shifted Rayleigh law for the *reverberation* class) [16].

## 3. SEGMENTATION

### *3.1. Two-Class Segmentation Step*

In this first segmentation step, we use the hierarchical two-class segmentation (SCM method) introduced in [16]. In order to make this hierarchical Markovian segmentation *unsupervised*, we exploit both the parameters of the mixture-based data model and those of the prior Potts model given by the ICE procedure. This two-class segmentation combines a standard coarse-to-fine multigrid method [27] with a scale-causal model and a multigrid energy minimization strategy. This model allows us to more precisely model the local and global characteristics of image content at different scales. Experiments and comparisons with other related hierarchical approaches (given in [16]) have proved that this scheme is well suited to automatic extraction of shadows from a large variety of sonar images.

Figures 7 and 8 display examples of unsupervised two-class segmentation, exploiting parameters estimated with the ICE procedure, for different noise models. Let us recall that in this segmentation step, the *sea-bottom reverberation* and the *echo* (or the reverberation on the object) classes are merged in a single class: the *reverberation* class. In Fig. 7, segmentation results obtained with two different noise models are comparable. Nevertheless, in Fig. 8, we observe that the Gauss–Rayleigh model does not permit elimination of speckle noise effects inducing reverberation-mislabeled pixels within the cast shadow region of the rock. The proposed noise modeling is appealing since it allows luminance in reverberation areas to be captured with more flexibility: the Weibull model offers a more general framework than the Rayleigh law (which is a particular case of Weibull modeling). Thus it allows a better fit to the data (cf. Table 3). The cost of this augmented model lies in the shape parameter that we now have to estimate. Moreover, the extracted cast shadows of manufactured objects (see Figs. 7, 8, and 9) exhibit, as desired, regular geometric shapes (in contrast to the cast shadows of rocks) without artifacts (i.e., mislabeled pixels within the cast shadow regions, as observed with a more simplistic noise model), which is in excellent agreement with the ground truth provided by an expert. This accuracy in extracting and preserving the borders of the cast shadows is very appealing in the prospect of a further classification step [28]. We therefore do believe that this new noise modeling is worth the pain.

### *3.2. Three-Class Segmentation Step*

*3.2.1. Problem statement.* In order to ensure proper detection of the echo information when the picture contains strong speckle noise, a solution consists in taking into account *a priori*

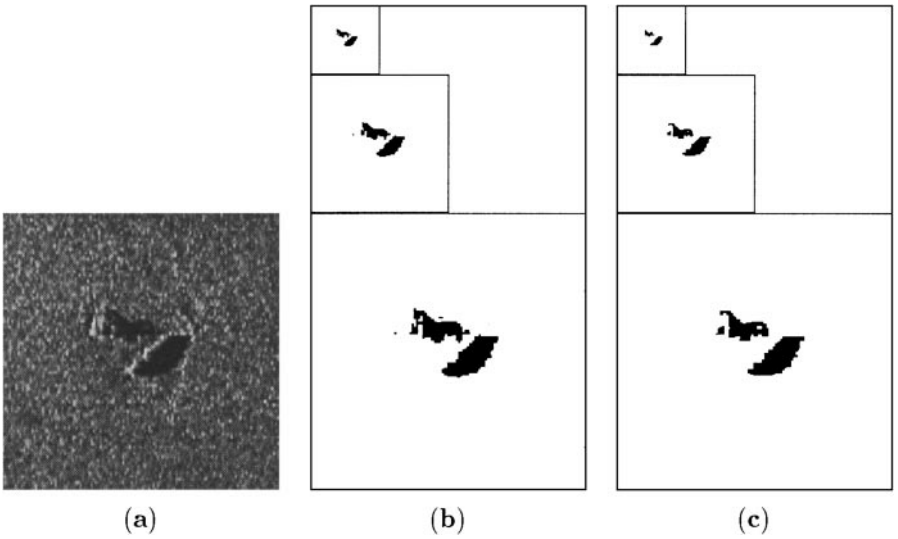



**FIG. 8.** (a) Real sonar image involving object and rock shadows. SCM two-class segmentation results obtained with a (b) mixture of Weibull distributions; (c) mixture of different distributions (Gaussian law for the *shadow* class and a shifted Rayleigh law for the *reverberation* class) [16].

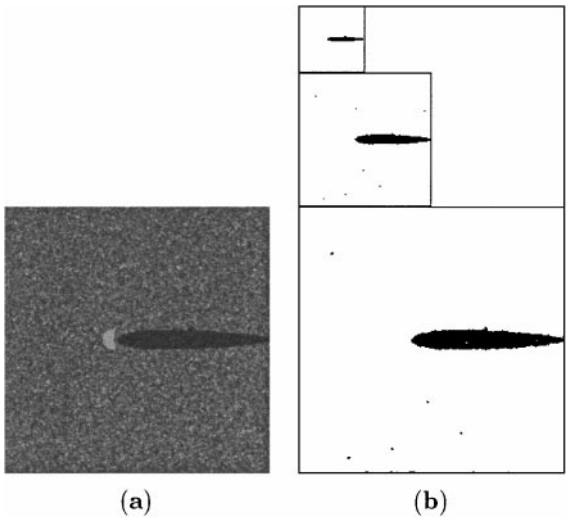



**FIG. 9.** (a) Synthetic sonar image of a sandy sea floor with a sphere lying on the sea-bed. (b) SCM two-class segmentation results.

information about the physical formation of the echo and incorporating it into the MRF model. In sonar imagery, objects lying on the sea floor create cast shadows corresponding to the regions acoustically obscured by the objects and also echoes corresponding to the signals back-scattered by the objects. The spatial dependency between these two areas can be incorporated into the MRF model. In the same vein, similar approaches exploiting *a priori* information relative to the interaction or spatial dependency between neighboring regions have been proposed and applied in image restoration [29] and in the classification problem [30].

In this section, we propose to incorporate *a priori* information about the way echoes physically appear. The adopted model exploits *a priori* information at different levels, pixel and region, to extract echoes from sea-bottom reverberation regions. We will first introduce the notation and the three-class segmentation strategy. Experimental results on real scenes are then presented.

*3.2.2. Three-class segmentation strategy.* Let $\hat{x}^{[1]}$ be the label field obtained after the two-class segmentation stage based on hierarchical Markovian modeling [16]. Label $\hat{x}_s^{[1]}$ belongs to $\{e_0, e_1\}$. Based on $x^{[1]}$, we now consider pixel subsets $S'(S' \subset S)$ such as $S' = \{s \in S : \hat{x}_s^{[1]} = e_1\}$. This set has now to be segmented into two classes, such as to extract reverberated signals. Let $X^{[2]}$ be the corresponding random binary process: $\forall s \in S'$, $X_s^{[2]}$ takes values in $\{e_1 = \textit{sea-bottom reverberation}, e_2 = \textit{echo}\}$. The segmentation will make use of restricted data $y^{[1]} = \{y_s, s \in S'\}$. The distribution of $(X^{[2]}, Y^{[1]})$ is defined, first, by $P_{X^{[2]}/X^{[1]}}(x^{[2]}/\hat{x}^{[1]})$, the distribution of $X^{[2]}$ assumed to be stationary and Markovian, and second, by the sitewise likelihoods $P_{Y_s^{[1]}/X_s^{[2]}}(y_s^{[1]}/x_s^{[2]})$,[1] depending on the class label $x^{[2]}$.

[1] Conditional sitewise data likelihoods of interest should actually read $P_{Y_s^{[1]}/X_s^{[1]},X_s^{[2]}}(y_s^{[1]}/\hat{x}_s^{[1]}, x_s^{[2]})$. By abuse of notation, we omit the conditioning with respect to $\hat{x}_s^{[1]} = e_1$, which is implied by $s \in S'$.

We now search for $x^{[2]}$ such as

$$\hat{x}^{[2]} = \arg\max_{x^{[2]}} P_{X^{[2]}/X^{[1]},Y^{[1]}}\left(x^{[2]}/\hat{x}^{[1]}, y^{[1]}\right). \tag{15}$$

Using Bayes' rule and after elimination of $P_{Y^{[1]}/X^{[1]}}(y^{[1]}/\hat{x}^{[1]})$, which does not depend on $x^{[2]}$, the previous expression can be written as

$$\hat{x}^{[2]} = \arg\max_{x^{[2]}} P_{X^{[2]},Y^{[1]}/X^{[1]}}\left(x^{[2]}, y^{[1]}/\hat{x}^{[1]}\right) \tag{16}$$

$$= \arg\max_{x^{[2]}} P_{Y^{[1]}/X^{[1]},X^{[2]}}\left(y^{[1]}/\hat{x}^{[1]}, x^{[2]}\right) \cdot P_{X^{[2]}/X^{[1]}}\left(x^{[2]}/\hat{x}^{[1]}\right). \tag{17}$$

In accordance with the Hammersley and Clifford theorem, $P_{X^{[2]}/X^{[1]}}(x^{[2]}/x^{[1]})$ is defined as a Gibbs distribution,

$$P_{X^{[2]}/X^{[1]}}\left(x^{[2]}/\hat{x}^{[1]}\right) \stackrel{\Delta}{=} \frac{1}{Z}\exp\left(-U_2\left(x^{[2]}, \hat{x}^{[1]}\right)\right), \tag{18}$$

where $U_2(x^{[2]}, \hat{x}^{[1]})$ stands for the energy function and $Z$ is the normalizing constant. Denoting $P_{Y^{[1]}/X^{[2]},X^{[1]}}(y^{[1]}/x^{[2]}, \hat{x}^{[1]}) = \frac{1}{Z'}\exp(-U_1(y^{[1]}, x^{[2]}, \hat{x}^{[1]}))$, Eq. (16) can be defined in terms of an energy function that has to be minimized,

$$\hat{x}^{[2]} = \arg\min_{x^{[2]}}\left\{U_1\left(y^{[1]}, x^{[2]}, \hat{x}^{[1]}\right) + U_2\left(x^{[2]}, \hat{x}^{[1]}\right)\right\}, \tag{19}$$

where $U_1(y^{[1]}, x^{[2]}, \hat{x}^{[1]})$ expresses the adequacy between observations and labels $\{e_1, e_2\}$ and $U_2$ is the energy term corresponding to the *a priori* model.

- Let us consider the data-driven term $U_1(y^{[1]}, x^{[2]}, \hat{x}^{[1]})$. In order to take into account the speckle noise phenomenon, we have used in Section 2.3 a shifted Weibull law $\mathcal{W}(.; \Phi_y)$ to describe the luminance $Y$ within reverberation regions. Due to the slight number of pixels belonging to the *echo* class, they cannot corrupt this modeling strongly. So we can efficiently approximate the distribution of luminance within the sea-bottom reverberation region by the same distribution. Thus, we define

$$P_{Y_s^{[1]}/X_s^{[2]}}\left(y_s^{[1]}/e_1\right) = \frac{C}{\alpha}\left(\frac{(y_s - \min)}{\alpha}\right)^{C-1}\exp\left(-\frac{(y_s - \min)^C}{\alpha^C}\right) \tag{20}$$

with $y > \min$, $\alpha > 0$, and $C > 0$. The label field $X^{[1]}$ disappears because $Y^{[1]}$ is defined only on $S'$; i.e., the dependence from $X^{[1]}$ in $S'$ brings no more information. $\Phi_y = (\min, C, \alpha)$ is estimated with the ICE procedure described in Section 2.2.

We know that the reflection of the acoustic wave on the object returns a high-amplitude signal and we have no *a priori* knowledge about the distribution of the gray levels within the echo region. The high amplitude signal at concerned pixel locations often induces a saturation of the reception captor of the sonar and consequently a maximal gray level for most of the pixels within these echo regions. Sites with gray level $y_{\max}$ ($y_{\max}$ designating the maximal gray level on the sonar image) are therefore

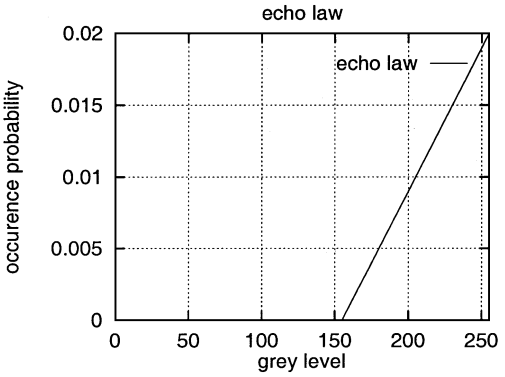



**FIG. 10.** Probability density function associated with the echo law.

likely to belong to the *echo* class. Thus we empirically model the conditional density function of the *echo* class by the simple law

$$P_{Y_s^{[1]}/X_s^{[2]}}\left(y_s^{[1]}/e_2\right) = \frac{2}{\gamma}\Lambda(y_s - y_{\max})\mathcal{U}(y_{\max} - y_s) \qquad (21)$$

where $\mathcal{U}(.)$ is the Heaviside function, $\Lambda(.)$ stands for the triangular function, $y_{\max}$ corresponds to the maximal grey level on the sonar image due to signal quantification, and $2/\gamma$ is a normalizing constant to ensure that the above function integrates to 1. This modeling can also be justified by the fact that the echo signal more often induces, a saturation of the reception captor of the sonar and consequently a maximal gray level for most of them. Figure 10 shows the plot of this law for $y_{\max} = 255$. Given these sitewise data likelihoods, the data energy term is expressed as

$$U_1\left(x^{[2]}, y^{[1]}, \hat{x}^{[1]}\right) = -\sum_{s \in S'} \ln P_{Y_s^{[1]}/X_s^{[2]}}\left(y_s^{[1]}/x_s^{[2]}\right). \qquad (22)$$

• Let us now consider $U_2(x^{[2]}, \hat{x}^{[1]})$, the energy term corresponding to the *a priori* model. We adopt an 8-connexity spatial neighborhood in which $\beta_1, \beta_2, \beta_3, \beta_4$ represent the *a priori* parameters associated to the horizontal, vertical, and right and left diagonal binary cliques respectively and $\beta_5$ stands for the unary clique parameter (cf. Fig. 11). This energy allows us to express constraints on the desired solution. In our application, we want to favor homogeneous regions. To do that, we use an anisotropic Potts model that associates to the binary clique $\langle s, t\rangle$, the

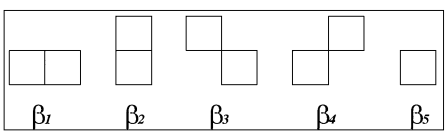



**FIG. 11.** One-site and two-site cliques for the 2nd order neighborhood, and associated parameters.

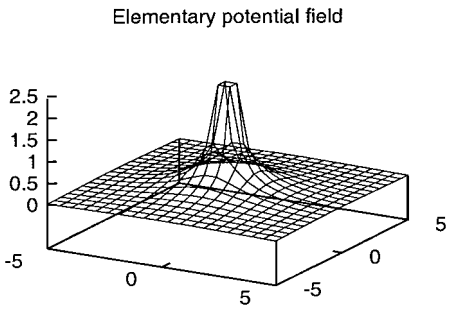



**FIG. 12.** Elementary potential field $\psi_{s_0}(r)$ (with $\sigma = 1$) created by a site $s_0$ labeled *shadow* in the two-class segmentation $\hat{x}^{[1]}$.

potential

$$\beta_{st}\left(1 - \delta\left(x_s^{[2]}, x_t^{[2]}\right)\right), \qquad (23)$$

where $\beta_{st} = \beta_1 = \beta_2 = \beta_3 = \beta_4$, depending on the orientations of the clique. Potential of the singleton clique with parameter $\beta_5$ is defined in order to disadvantage the choice of the *echo* label for a site that is too far away from a shadow region. To define this potential, we introduce a potential field $\Psi_{\hat{x}^{[1]}}(t), t \in S$, as follows. Each site $s_0$ labeled *shadow* in $\hat{x}^{[1]}$ creates an elementary potential field $\psi_{s_0}(r)$ such as

$$\psi_{s_0}(r) = \frac{1}{r}\exp\left(-\frac{r}{\sigma}\right), \qquad (24)$$

where $r (r \neq 0)$ is the distance to the pixel $s_0$ and $\sigma$ is a standard deviation parameter controlling the interaction distance between echo and shadow regions. Figure 12 shows an example of this elementary potential field. The set of pixels labeled *shadow* in $\hat{x}^{[1]}$ create a global potential field $\Psi_{\hat{x}^{[1]}}(t)$, combining the different elementary potential fields $\psi_{s_0}(r)$ as follows,

$$\Psi_{\hat{x}^{[1]}}(t) = \inf\left\{\sum_{s \in S : \hat{x}_s^{[1]} = e_0} \psi_s(d(s,t)), 1\right\}, \qquad (25)$$

where $d(s, t)$ is the distance between pixels $s$ and $t$. The unary potential with parameter $\beta_5$ is ass to disadvantage the choice of the *echo* label $e_2$ for a site too far away from a shadow region. Its form is given by

$$-\beta_5 \ln \Psi_{\hat{x}^{[1]}}(s) \cdot \delta\left(x_s^{[2]}, e_2\right). \qquad (26)$$

The corresponding energy term induces an area within which the *echo* label will not be discouraged. Figure 13 displays an example of the potential field, computed for different values of the parameter $\sigma$ from the segmented image presented in Fig. 13.b. Finally, the global energy function to be minimized is defined

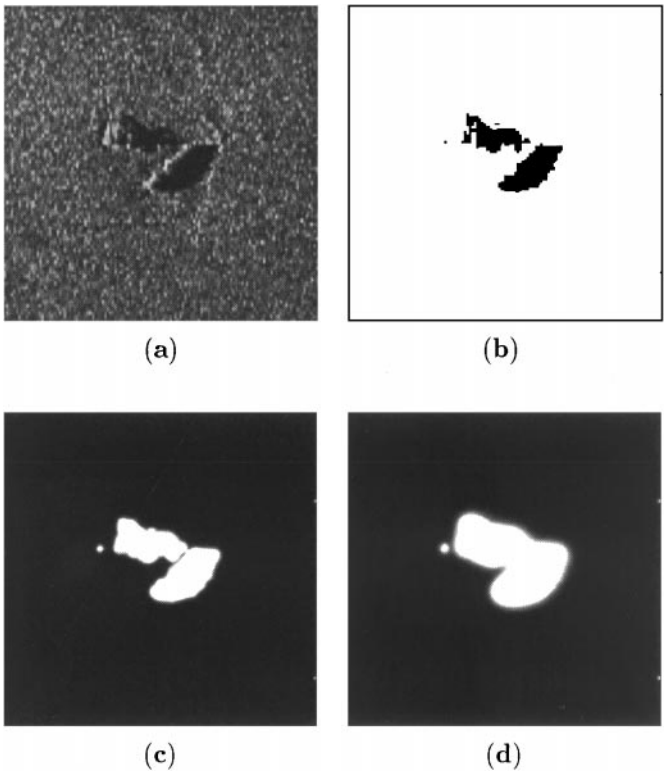



**FIG. 13.** (a) Real sonar image involving object and rock shadows. (b) Two-class segmentation results ($\hat{x}^{[1]}$). (c) Magnitude of the potential field $\Psi_{\hat{x}^{[1]}}(t)$ combining the different elementary potential fields $\psi_{s_0}(r)$ created by each site $s_0$ labeled *shadow* in the two-class segmentation $\hat{x}^{[1]}$ (with $\sigma = 2$). (d) Magnitude of the potential field with $\sigma = 4$.

as follows:

$$U\left(x^{[2]}, \hat{x}^{[1]}, y^{[1]}\right) = \underbrace{-\sum_{s\in S'} \ln P_{Y_s^{[1]}/X_s^{[2]}}\left(y^{[1]}/x_s^{[2]}\right)}_{U_1\left(y^{[1]}, x^{[2]}, \hat{x}^{[1]}\right)} + \underbrace{\sum_{\langle s,t\rangle \subset S'} \beta_{st}\left(1-\delta\left(x_s^{[2]}, x_t^{[2]}\right)\right)}_{U_{21}\left(x^{[2]}, \hat{x}^{[1]}\right)} + \underbrace{\sum_{s\in S'} -\beta_5 \ln \Psi_{\hat{x}^{[1]}}(s)\cdot\delta\left(x_s^{[2]}, e_2\right)}_{U_{22}\left(x^{[1]}, x^{[2]}, \hat{x}^{[1]}\right)}.$$

We use the deterministic relaxation algorithm ICM [12] to minimize this global energy function. For the initialization of this algorithm, we exploit the segmentation map obtained by a ML segmentation.

## 4. EXPERIMENTAL RESULTS

For the three-class segmentation step, we use the following parameters: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 1$ for the *a priori* parameter associated to the binary clique (the Potts model is here isotropic) and the unary clique respectively, and $\sigma = 2$ for the standard deviation parameter controlling the interaction distance between the echo and shadow regions (see Eq. (24)).

Sonar images presented herein are provided by a side-scan sonar (namely, the DUBM41, whose frequency is around 500 KHz). The size of these pictures is 256 by 256 pixels corresponding to a sea floor surface of 25 by 25 m.

We observed that the final segmentation results are not very sensitive to the value of the $\sigma$ parameter (within the range $[1, \ldots, 4]$). If this parameter is low (below 1, for example), the elementary potential field created by each site labeled *shadow* in $\hat{x}^{[1]}$ is too sharp. As a result, the area within which the *echo* label will not be discouraged is too small and, consequently, the *echo* region can be not entirely determined. Conversely, a large value for $\sigma$ could let appear in the segmentation results false alarms, i.e., small *echo* areas due to the speckle noise. On one hand, this parameter could be estimated optimally and automatically if the depth of the sea floor and the mean size of the objects to be visualized were known; but on the other hand observed robustness for $\sigma \in [1, 4]$ did not drive us to consider such an estimate.

We compare the result obtained for a real sonar image with (1) a classical ML segmentation, (2) our segmentation model without *a priori* information on the spatial dependency between echo and shadow regions (by setting $\beta_5 = 0$), and finally, (3) our scheme. Figure 14.a shows a real sonar image involving object

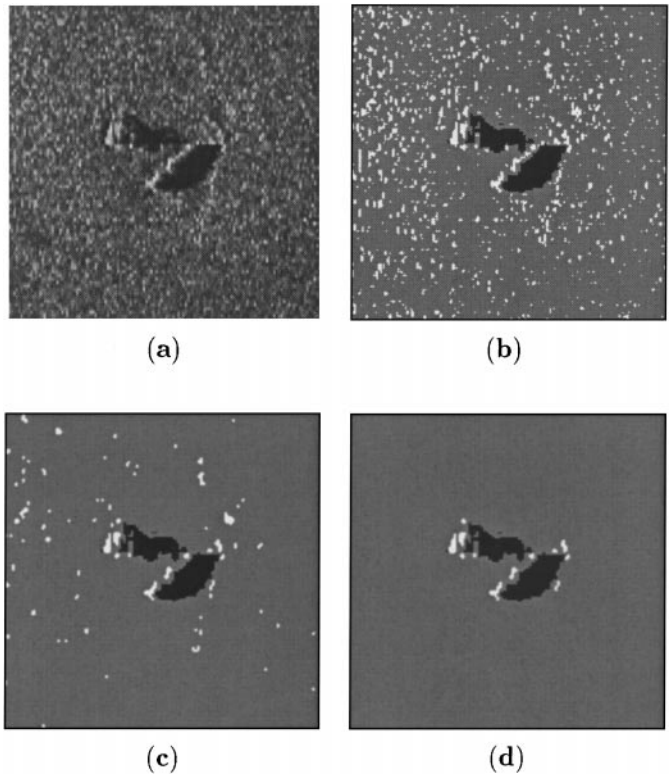



**FIG. 14.** (a) Real sonar image involving object and rock shadows. (b) ML segmentation. (c) proposed MRF model without *a priori* information on the spatial dependency between echo and shadow regions ($\beta_5 = 0$). (d) Proposed MRF model with *a priori* information on the spatial dependency between echo and shadow regions.

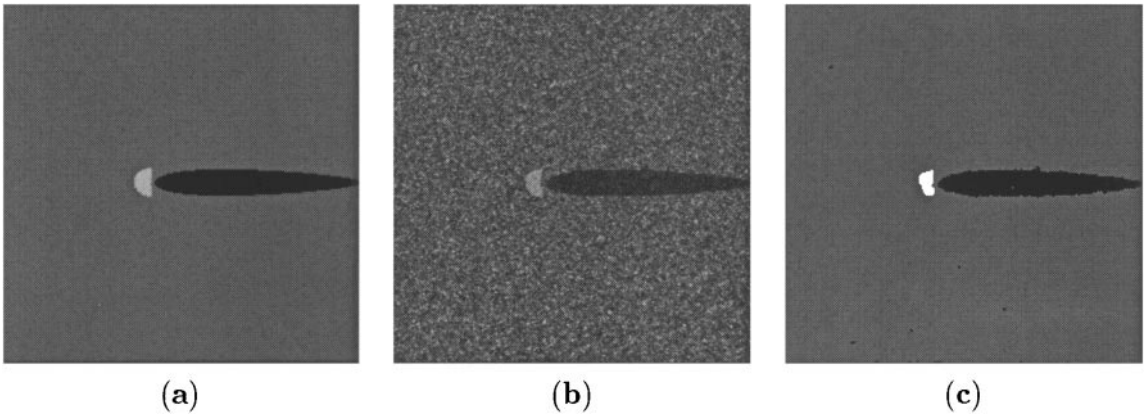



**FIG. 15.** (a) Synthetic shadow shape of a sphere lying on a sea-bed (obtained with a ray tracing procedure). (b) Synthetic sonar image (with synthetic speckle noise) of a sphere lying on a sandy sea-bed. (c) Three-class segmentation results obtained with the proposed scheme.

and rock shadows and Figs. 14.b, 14.c, and 14.d present the segmentation results obtained with these three approaches. Compared to a ML segmentation, a segmentation model integrating the *a priori* energy term allows the spatial coherence of the obtained partitions to be ensured, i.e., it favors homogeneous regions (see Figs. 14.c and 14.d). Nevertheless, without the *a priori* term expressing the spatial dependency between echo and shadow areas, the segmentation scheme does not permit totall eliminatly the speckle noise effect, inducing false small echo areas (see Fig. 14.c). We can notice that our approach leads to better result and can remove efficiently undesired echos induced by speckle noise effects (see Fig. 14.d).

Due to the "stochastic" aspect of the ICE iterative estimation procedure, the resulting estimates are less sensitive to the initial parameter values than those obtained by other estimation algorithms of the distribution mixture parameters (like the well-known Expectation Maximization (EM) algorithm, for example). There is no theoritical proof of convergence for ICE algorithm; nevertheless good behavior is generally noticed. A bad initialization can lead the ICE procedure to discard wrongly one class, the shadow class in our application (by estimating as zero the proportion of this class and by considering wrongly that all the pixels of the image belong to a single class, which is only true for the sonar image of a sandy sea-floor). But the initialization method we used [18] seems to be good enough to avoid this problem. It allows to get a good convergence of the ICE procedure in all tested cases (300 images in the data base).

Experiments have been carried out on both synthetic and real sonar images. Figures 15.b and 16.b present synthetic sonar images (with synthetic speckle noise) of a sphere and a metallic core (or a metallic tire) lying on a sandy sea floor. In these two examples, the echo shape and the acoustic cast shadow of these objects are obtained by a ray tracing procedure. Figures 15.c and 16.c present the three-class segmentation results obtained with our approaches. These results can be compared to the ground-truth segmentations given in Figs. 15.a and 16.a. These experiments demonstrate that the recovered segmentations are close to the ideal one.

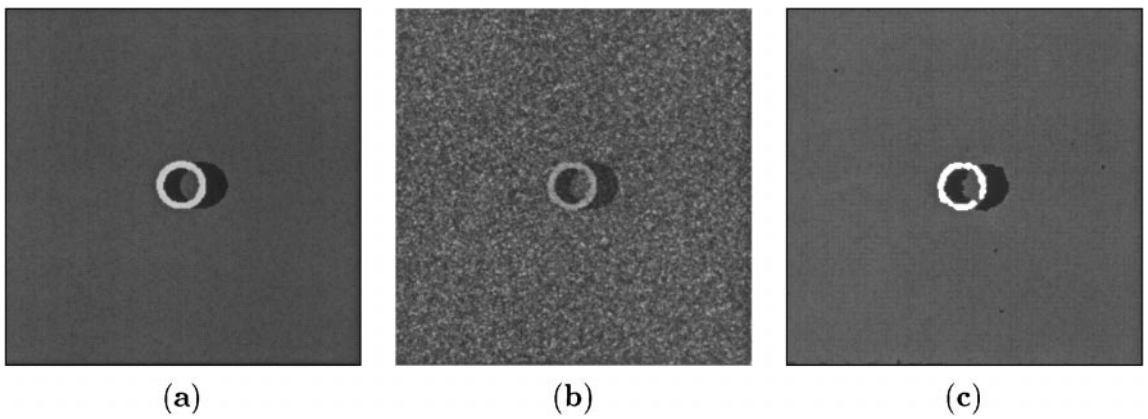



**FIG. 16.** (a) Synthetic shadow shape of a metallic core lying on a sea-bed (obtained with a ray tracing procedure). (b) Synthetic sonar image (with synthetic speckle noise) of a metallic core lying on a sandy sea-bed. (c) Three-class segmentation results obtained with the proposed scheme.

We present now three-class segmentation results obtained on real sonar images. Figures 17.a, 17.c, and 17.e display three metallic manmade objects (two cylinders and a trolley) lying on the sea bed. In these two examples, the echo features are discriminant and easily identifiable. Figures 18.a and 18.c display sonar images showing respectively a tire and a cylindrical manufactured object lying on a sandy sea floor. In these examples, due to the nature of the object, echo regions are not very large. Experiments indicate that the obtained segmentation maps are close to the expected results. The echo and shadow regions are well segmented and the proposed algorithm exhibits good robustness against speckle noise. Boundaries of each object have been well preserved and numerous false alarms corresponding to spurious small shadow and echo areas due to speckle noise have been correctly eliminated.

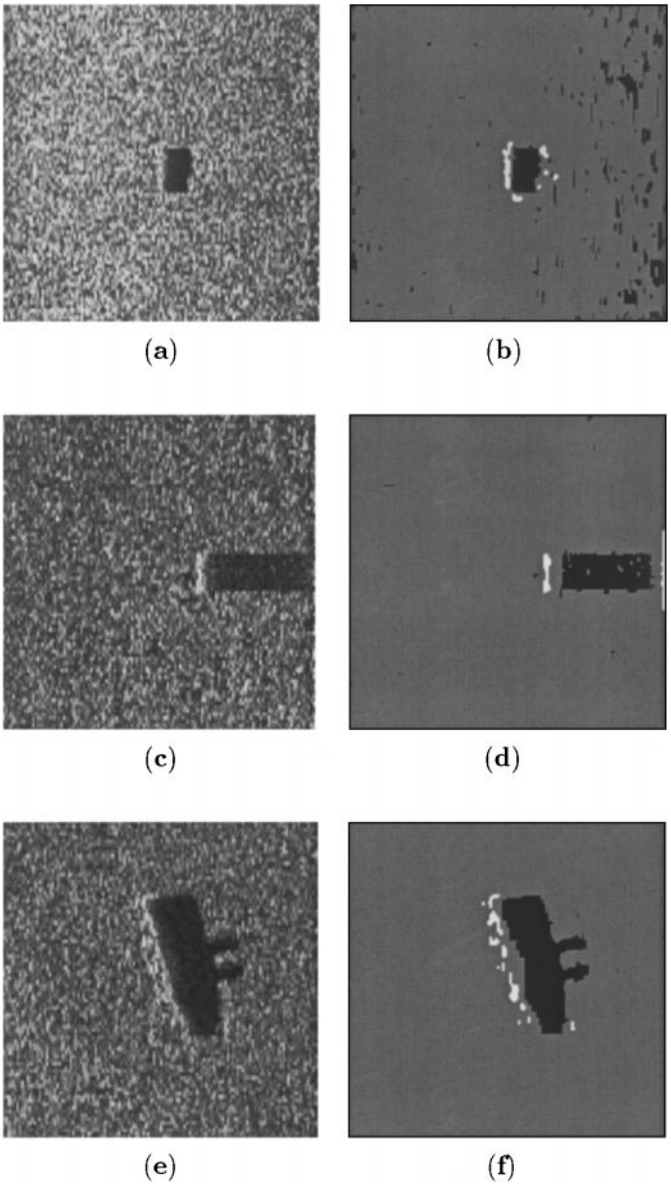



**FIG. 17.** (a, c, e) Real sonar images involving a sandy sea floor and a manmade object (two cylinders and a trolley). (b, d, f) Three-class segmentation results obtained with the proposed scheme.

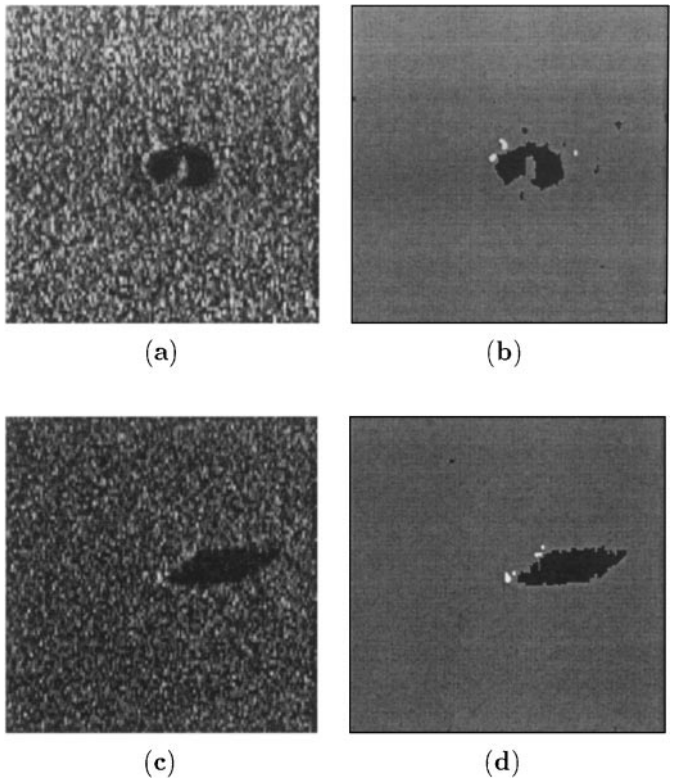



**FIG. 18.** (a, c) Real sonar images of a sandy sea floor with the cast shadow of a tire in (a) and the cast shadow of a cylindrical manufactured object in (c). (b, d) Three-class segmentation results obtained with the proposed algorithm.

Nevertheless, we can notice on some sonar image segmentations the presence of some echo points "behind" the cast shadow shapes of some objects lying on the sea floor. These artifacts could be easily discarded by ordering the spatial relationship between echoes and shadows, i.e., by exploiting the *a priori* (approximate) information about the sonar position, which is not available on our data base. Thus, in our model this interesting *a priori* information is not taken into account because it is not available.

## 5. CONCLUSION

We have described an unsupervised three-class segmentation method, based on an estimation step and a segmentation step, which seems well adapted and efficient for sonar image segmentation issues. The estimation step offers an appropriate estimation of the model parameters, and takes into account the diversity of the laws in the distribution mixture of sonar images by modeling each noise distribution with a Weibull PDF. In order to obtain an accurate segmentation map in spite of the presence of speckle noise, the proposed segmentation is articulated in two stages and

exploits the previously estimated parameters. In the first, a scale-causal and spatial model is used to separate the *shadow* class from the others (*echo* and *sea-bottom reverberation*). Then, in the second stage, a MRF-based monoscale scheme integrating a priori information at different levels of representation (pixel and region) allows the echo regions to be distinguished from the sea-bottom reverberation ones. This scheme is computationally simple and well suited to automatic three-class segmentation on a large variety of sonar images. This method has been validated on a number of real sonar images. Obtained results demonstrate the efficiency and the robustness of this scheme.

## ACKNOWLEDGMENTS

The authors thank GESMA (Groupe d'Études Sous-Marines de l'Atlantique, Brest, France) for providing numerous real sonar pictures and DGA (Direction Générale pour l'Armement, French Ministry of Defense) for partial financial support of this work (student grant).

## REFERENCES

1. C. Collet, P. Thourel, P. Pérez, and P. Bouthemy, Hierarchical MRF modeling for sonar picture segmentation, in *Proc. 3rd International Conference on Image Processing, Lausanne, Sept. 1996*, Vol. 3, pp. 979–982.
2. I. L. Herlin, C. Nguyen, and C. Graffigne, Stochastic segmentation of ultrasound images, in *Proc. 11th IAPR International Conference on Pattern Recognition, The Hague. Sept. 1992*, pp. 289–292.
3. B. Zerr and B. Stage, Three-dimensional reconstruction of underwater objects from a sequence of sonar images, in *Proc. 3rd International Conference on Image Processing, Lausanne, Sept. 1996*, pp. 927–930.
4. P. Cervenka and C. de Moustier, Sidescan sonar image processing techniques, *IEEE J. Oceanic Eng.* **18**(2), 1993, 108–1229.
5. L. M. Linnett, S. J. Clarke, and D. R. Carmichael, The analysis of sidescan sonar images for seabed types and objects, in *Proc. of Second European Conf. on Underwater Acoustics, 1994*, pp. 733–738.
6. S. Guillaudeux and S. Daniel, Optimization of a sonar image processing chain: A fuzzy rules based expert system approach, in *Proc. OCEAN, Fort Lauderdale, 1996*, pp. 1319–1323.
7. S. Daniel and S. Guillaudeux, Adaptation of a partial shape recognition approach, in *Proc. IEEE Conf. On Systems, Man, and Cybernetics, Oct. 1997*, pp. 2157–2162.
8. S. Dugelay, C. Graffigne, and J. M. Augustin, Deep seafloor characterization with multibeam echosonders by image segmentation using angular acoustic variations, in *Proc. SPIE*, Vol. 2847, 1996.
9. M. Gensane, A statistical study of acoustic signal backscattered from the sea-bottom, *IEEE J. Oceanic Eng.* **14**, 1989, 84–93.
10. M. Jiang, W. K. Stewart, and M. Marra, Segmentation of seafloor sidescan imagery using Markov random field and neural networks, in *Proc. of OCEANS'93*, Vol. 3, No.3, 1993, pp. 456–461.
11. J. W. Goodman, Some fundamental properties of speckle, *J. Opt. Soc. of Am.* **66**(11), 1976, 1145–1150.
12. J. Besag, On the statistical analysis of dirty pictures, *J. R. Statist. Soc.* **B-48**, 1986, 259–302.
13. F. Schmitt, L. Bonnaud, and C. Collet, Contrast control for sonar pictures, in *Signal and Image Processing, SPIE'96*, Technical Conference on Application of Digital Image Processing XIX, Vol. 2847, pp. 70–82, Aug. 1996.
14. C. A. Bouman and M. Shapiro, A multiscale random field model for Bayesian image segmentation, *IEEE Trans. Image Process.* **3**(2), 1994, 162–177.
15. Z. Kato, M. Berthod, and J. Zerubia, A hierarchical Markov random field model and multitemperature annealing for parallel image classification, *Graph. Models Image Process.* **58**, 1996, 18–37.
16. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, Sonar image segmentation using an unsupervised hierarchical MRF model, *IEEE Trans. Image Process.* to appear.
17. F. Salzenstein and W. Pieczinsky, Unsupervised Bayesian segmentation using hidden Markovian fields, in *Proc. International Conference on Acoustics, Speech and Signal Processing, Detroit, May 1995*, pp. 2411–2414.
18. H. Derin and H. Elliot, Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Trans. Pattern Analy. Mach. Intell.* **9**(1) 1987, 39–55.
19. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, Unsupervised segmentation applied on sonar images, in *Proc. International Workshop EMM-CVPR'97: Energy Minimisation Methods in Computer Vision and Pattern Recognition, Venice, Italy, May 1997*, Lecture Notes in Computer Science, Vol. 1223, pp. 491–506. Springer-Verlag, Berlin/New York, 1997.
20. R. C. Gonzalez, *Digital Image Processing*, 2nd ed., Addison–Wesley, Reading, MA, 1987.
21. S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 1984, 721–741.
22. E. Conte, M. Lops, and G. Ricci, Optimised radar detection in correlated Weibull clutter, in *14th Colloque GRETSI, Juan-les-Pins, France, Sept. 1993*, pp. 161–164.
23. F. Schmitt, M. Mignotte, C. Collet, and P. Thourel, Estimation of noise parameters on sonar images, in *Signal and Image Processing*, Proc. SPIE Vol. 2823, 1996 pp. 1–12.
24. B. Braathen, P. Masson, and W. Pieczynski, Global and local methods of unsupervised Bayesian segmentation of images, *Graphics Vision* **2**(1), 1993, 39–52.
25. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, Unsupervised Markovian segmentation of sonar images, in *Proc. International Conference on Acoustics, Speech and Signal Processing, Munich, May 1997*, Vol. 4, pp. 2781–2785.
26. J. A Cornell, *Experiments with Mixtures: Designs Models and the Analysis of Mixture Data*, Wiley–Interscience, New York, 1990.
27. F. Heitz, P. Pérez, and P. Bouthemy, Multiscale minimisation of global energy functions in some visual recovery problems, *CVGIP: Image Understand.* **59**, 1994, 125–134.
28. M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, Statistical model and genetic optimization: application to pattern detection in sonar images, in *Proc. International Conference on Acoustics, Speech, and Signal Processing, Seattle, May 1998*, Vol. 5, pp. 2471–2475.
29. C. Regazzoni, F. Arduini, and G. Vernazza, A multilevel GMRF-based approach to image segmentation and restoration, *Signal Process.* **34**, 1993, 43–67.
30. Y. Kim and H. S. Yang, Efficient image labelling based on Markov random field and error backpropagation network, *Pattern Recognition*, **26**(11), 1993, 1695–1707.