

Optimal Feature Set for Automatic Detection and Classification of Underwater Objects in SAS Images

Raquel Fandos and Abdelhak M. Zoubir

Abstract—The problem of automatic detection and classification for mine hunting applications is addressed. We propose a set of algorithms which are tested using a large database of real synthetic aperture sonar (SAS) images. The highlights and shadows of the objects in an SAS image are segmented using both a Markovian algorithm and the active contours algorithm. The comparison of both segmentation results is used as a feature for classification. In addition, other features are considered. These include geometrical shape descriptors, not only of the shadow region, but also of the object highlight, which demonstrates a significant improvement of the performance. Furthermore, a novel set of features based on the image statistics is described. Finally, we propose an optimal feature set that leads to the best classification results for the available database.

Index Terms—Active contours (ACs), feature extraction, image segmentation, Markov random fields (MRFs), synthetic aperture sonar (SAS), underwater object classification.

I. INTRODUCTION

BOTH sidescan sonar and synthetic aperture sonar (SAS) technologies provide high-resolution imagery for mine-countermeasure applications. While sidescan sonar has been well developed for several decades, SAS is a fairly new technology [1]. Sidescan sonar systems use antenna arrays to scan the seabed in a narrow beam fashion. The reconstructed images present a resolution that is not only limited by the antenna aperture but that also decreases with range. SAS technology overcomes these limitations: a single antenna system moves along a straight line and pings are sent at different positions. Hence, a synthetic array, theoretically as long as desired, is built up and the resolution increases. Furthermore, it can be demonstrated that the resolution is independent of the range [2].

Traditionally, human operators perform the identification and classification of objects present on a seabed. However, there is a growing demand for computer-aided detection and computer-aided classification (CAD/CAC) operations, which were introduced for sidescan applications in [3] and are also an active field of research in synthetic aperture radar [4]. Due to its strategic relevance for military applications, very few studies

with real sidescan data exist in the literature, and even less with SAS images. This paper presents a complete CAD/CAC system that has been tested with a 10 000 square meter database of real SAS images.

Two main approaches to the problem exist: template fitting and feature description. The former typically applies a set of allowed transformations to a given template, so that it matches the region of interest. The distance between the deformed template and the object is measured in some norm and the object is classified accordingly. Template fitting for sidescan sonar images has been considered in [5]–[8], and in [9]–[11] for SAS applications.

The feature description approach, adopted in this work, requires the segmentation of the regions of interest. That is, the highlight and the shadow of the object under consideration must be extracted. Then, a set of significant features are computed and the object is classified according to the comparison of these features with those of a training set. Sidescan sonar image detection and classification based on feature extraction schemes have been considered in several studies with successful results [12]–[16]. So far, the focus has been on descriptors of the shadow shape [17]–[19]. In [6] and [20], Fourier descriptors are considered and in [18] normalized central moments are used. In this paper, we have tested several kinds of shape descriptors, not only for the shadow but also for the object highlight. While some of them are well-known shape descriptors from the literature, to the best of our knowledge, this is the first time that they are tested with SAS images.

An alternative approach is to focus on the statistical properties of an image. In [21], the mean and variance of different regions are considered. The kurtosis and skewness are used for detection and classification purposes in [22] and [23]. Also the difference of SNR between the different regions has been considered [3]. In this paper, we propose a novel set of statistical features based on a Weibull parametric model of the SAS images.

This paper is organized as follows. In Section II, the segmentation of the regions of interest is considered. Two algorithms, a Markov random field (MRF) model-based algorithm and an active contours (ACs) algorithm are described and results are shown. A statistical classifier is proposed in Section III. Different shape and statistical descriptors for both the shadow and the highlight regions are presented and tested for the available data in Sections IV-A–IV-E. Besides, we propose an original descriptor based on a comparison of the segmentation results provided by both segmentation algorithms. An optimal feature vector is presented, before we provide conclusions and an outlook for some future work. A scheme of the CAD/CAC system for SAS applications that we propose is depicted in Fig. 1.

Manuscript received April 20, 2010; revised August 31, 2010; accepted November 02, 2010. Date of publication November 18, 2010; date of current version May 18, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jocelyn Chanussot.

The authors are with the Signal Processing Group, Institute of Telecommunications, Technische Universität Darmstadt, 64283 Darmstadt, Germany (e-mail: rfandos@spg.tu-darmstadt.de; zoubir@spg.tu-darmstadt.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTSP.2010.2093868

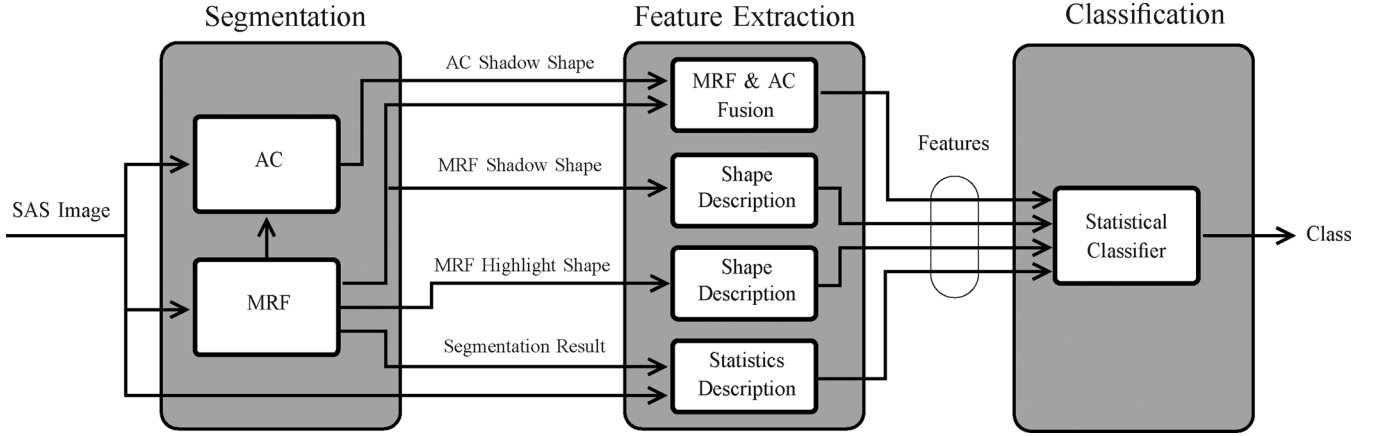


Fig. 1. Scheme of the CAD/CAC system for SAS images proposed in this paper. The sonar image is segmented by two algorithms, an MRF model-based algorithm and the AC algorithm. The latter uses the result of the former for its initialization. Several sets of features are extracted: both statistical and geometrical features for both the shadow and the highlight of the object. Moreover, the comparison of the shadow shape estimated by both segmentation algorithms is exploited. Each set of features is tested independently with an statistical classifier. However, it will be shown that an optimal feature set, consisting of some descriptors from each kind, provides the best classification results.

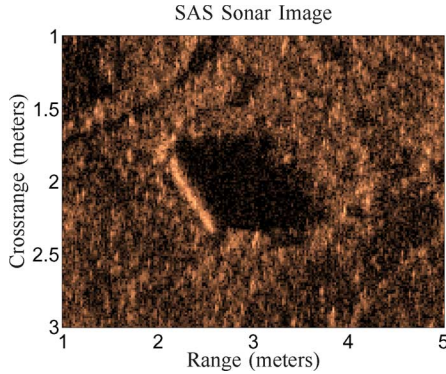


Fig. 2. SAS sonar image of a cylindrical object. The pixels of lower intensity correspond to the shadow of the object, while the object highlight presents an intensity that is higher than the background.

II. SEGMENTATION OF SONAR IMAGES

Segmentation is the process of partitioning an image into several regions [24]. In our application, three different regions occur (see Fig. 2): the highlights H of the objects in the scene, their shadows S and the seabed or background B . The objects cannot only be mines, but also physical features of the terrain such as rocks or sand ripples. In this paper, we consider any element that produces a shadow as being an object.

Several approaches have been carried out to segment sonar images automatically. Some of them are based on simple clustering techniques such as histogram thresholding [18] or fuzzy K -means [25], [26]. These models perform well for flat seabeds with high signal-to-noise ratio (SNR), but fail for more complex environments such as sand ripples. To cope with object detection that is robust to such conditions, we consider two different segmentation approaches. A well-known algorithm based on an MRF model [27] and a second segmentation algorithm, the active contours [28], have been used. Both approaches are modified for the application at hand and tested using a database of real SAS images.

A. Markovian Segmentation Algorithm

Assuming an MRF model for image segmentation is a popular approach for many applications, such as synthetic aperture radar. It has also been used in the context of seabed reconstruction from raw sonar data [29], [30]. MRF models for sidescan sonar image segmentation were introduced in [31] and [32] and later used in [33] and [34], where they were combined with the iterative conditional estimation (ICE) algorithm to achieve an unsupervised implementation. To the best knowledge of the authors, this is the first time that MRF based segmentation is applied to SAS images.

Let us express the sonar image $Y(u, v)$, $u = 1, \dots, U_{\max}$, $v = 1, \dots, V_{\max}$, associated with the lattice \mathcal{L} , as the vector $\mathbf{y} = \{y_i, i \in \mathcal{L}\}$, where y_i denotes the intensity of pixel i . The label field $\mathbf{x} = \{x_i, i \in \mathcal{L}\}$ is the “ground truth” that we want to recover, where each x_i has one of the possible assigned labels $\{S, B, H\}$. The MRF model consists of the two fields (\mathbf{y}, \mathbf{x}) . According to Bayes theorem, the posterior probability of the label field \mathbf{x} given the sonar image \mathbf{y} corresponds to the expression

$$P(\mathbf{x}|\mathbf{y}) = P(\mathbf{x}) \cdot P(\mathbf{y}|\mathbf{x}) \quad (1)$$

where $P(\mathbf{x})$ is the Markovian *a priori* probability and $P(\mathbf{y}|\mathbf{x})$ is the likelihood function of the image. Both are described in the sequel (Sections II-A1 and II-A2). The ICE algorithm, employed to estimate the parameters that define $P(\mathbf{x})$ and $P(\mathbf{y}|\mathbf{x})$, is described in Section II-A4.

The optimal estimate for \mathbf{x} is the one that corresponds to the global maximum of (1). Unfortunately, its calculation is a computationally unaffordable task. In Section II-A5, the Iterated Conditional Modes (ICM) [27], which converges to a local maximum of (1), is proposed as a reasonable and fast approach. For the sonar image \mathbf{y} in Fig. 2, the estimate of \mathbf{x} , $\hat{\mathbf{x}}$, is depicted in Fig. 3.

1) *Markovian Probability $P(\mathbf{x})$* : Given a pixel i , its neighborhood \mathcal{M}_i is formed by a set of pixels such that $i \notin \mathcal{M}_i$, and $\forall i' \in \mathcal{M}_i, i \in \mathcal{M}_{i'}$. Each pair $\{i, i'\} \mid i' \in \mathcal{M}_i$ is

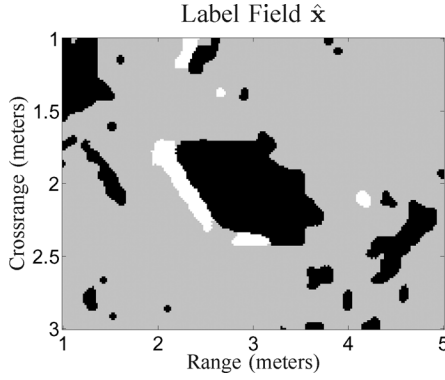


Fig. 3. Estimated label field $\hat{\mathbf{x}}$ associated with the sonar image in Fig. 2. The label for the background pixels B is depicted in gray, the shadow label S in black and the highlight label H in white.

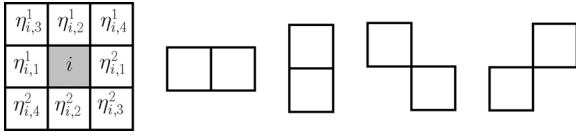


Fig. 4. Second-order neighborhood system and associated clique type 1, 2, 3, and 4.

known as a clique. Several neighborhood systems are commonly used in image modeling [24]. For this application, the second-order neighborhood system is chosen. Fig. 4 shows such a neighborhood configuration and its associated cliques. The first clique type relates the pixel i with the neighbors to its right and left ($\eta_{i,1}^1$ and $\eta_{i,1}^2$), the second one models its dependency with the neighbors above and below ($\eta_{i,2}^1$ and $\eta_{i,2}^2$), the third and fourth one relates it with the pixels in the first ($\eta_{i,3}^1$ and $\eta_{i,3}^2$) and second ($\eta_{i,4}^1$ and $\eta_{i,4}^2$) diagonals, respectively. Hence, the neighborhood of pixel i is defined by $\mathcal{M}_i = \{\eta_{i,1}^1, \eta_{i,1}^2, \eta_{i,2}^1, \eta_{i,2}^2, \eta_{i,3}^1, \eta_{i,3}^2, \eta_{i,4}^1, \eta_{i,4}^2\}$.

According to the Hammersley–Clifford theorem [35], [36], there is a one-to-one equivalence between MRF and the so-called Gibbs Random Fields, which have an associated Gibbs distribution. This kind of representation is very convenient for modeling the *a priori* probability $P(x_i)$, that is, the dependency of a pixel label x_i with the labels of its neighbors \mathcal{M}_i . A random field \mathbf{x} has a Gibbs distribution with respect to a neighborhood system \mathcal{M} if and only if its joint distribution can be expressed as

$$P(x_i) = \frac{1}{Z} e^{-U(x_i)} \quad (2)$$

where Z is a normalizing constant and $U(x_i)$ is an energy function that corresponds to the expression

$$U(x_i) = \Theta_i^T \Omega_{\mathbf{x}} = (\Theta_{i,1}, \Theta_{i,2}, \Theta_{i,3}, \Theta_{i,4}) \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}. \quad (3)$$

Each β_z describes the neighborhood relations of a pixel with its neighbors of clique type z , and

$$\Theta_{i,z} = 2 - \delta[x_i - x(\eta_{i,z}^1)] - \delta[x_i - x(\eta_{i,z}^2)] \quad (4)$$

where $z = 1, \dots, 4$, $\delta(\cdot)$ is Kronecker's delta and $x(\eta_{i,z}^1)$ and $x(\eta_{i,z}^2)$ refer to the labels of the neighbor pixels $\eta_{i,z}^1$ and $\eta_{i,z}^2$,

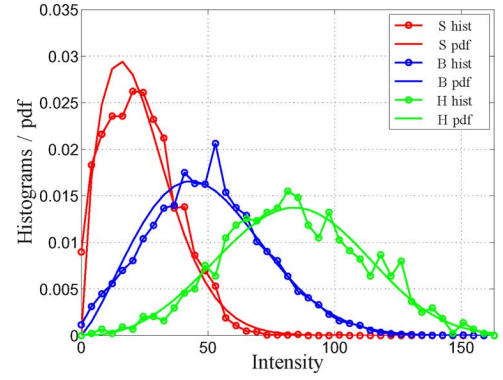


Fig. 5. Histograms and estimated Weibull pdfs for the different regions of the SAS image in Fig. 2 as segmented in Fig. 3. The Weibull parametric model is appropriate for SAS images.

respectively. While Θ_i can be computed directly from $\hat{\mathbf{x}}$, the parameter vector $\Omega_{\mathbf{x}}$ needs to be estimated. Implementation details for the estimation of $\Omega_{\mathbf{x}}$ from $(\mathbf{y}, \hat{\mathbf{x}})$ can be found in [37].

Including *a priori* knowledge about the topology of the sonar images into the algorithm, improves its performance. First of all, highlight regions are always located to the left of shadow regions. Second, the typical size of expected highlight regions is known for a given range. Both effects can be taken into account in the Gibbs energy of (3) [38].

2) *Likelihood Function* $P(\mathbf{y}|\mathbf{x})$: Assuming that y_i are conditionally independent, the likelihood function for the image \mathbf{y} is given by

$$\begin{aligned} P(\mathbf{y}|\mathbf{x}) &= \prod_{i \in \mathcal{L}} P(y_i|x_i) \\ &= \prod_{\{y_i|x_i=B\}} P_B(y_i) \cdot \prod_{\{y_i|x_i=S\}} P_S(y_i) \cdot \prod_{\{y_i|x_i=H\}} P_H(y_i) \end{aligned} \quad (5)$$

where P_B , P_S , and P_H are the probability density functions (pdfs) of the corresponding regions. As proposed in [38], [39], a suitable parametric model for SAS images is the Weibull distribution $\mathcal{W}(\lambda, \xi)$, where λ and ξ are the scale and shape parameters, respectively. The parameters ξ_S and λ_S are estimated maximizing the likelihood function [40]

$$(\xi_S, \lambda_S) = \arg \max_{\xi, \lambda} \prod_{\{y_i|x_i=S\}} \frac{\xi}{\lambda} \left(\frac{y_i}{\lambda} \right)^{\xi-1} \exp \left(- \frac{y_i}{\lambda} \right) \quad (6)$$

and analogously for B and H , which yields the parameter vector $\Omega_{\mathbf{y}} = (\lambda_S, \xi_S, \lambda_B, \xi_B, \lambda_H, \xi_H)$. Fig. 5 shows the estimated pdfs of the SAS image in Fig. 2, as segmented in Fig. 3. For comparison, the histograms are depicted.

3) *Initialization*: The ICE algorithm requires an initialization of the label field $\hat{\mathbf{x}}$ so that initial values for $\Omega_{\mathbf{y}}$ and $\Omega_{\mathbf{x}}$, $\Omega_{\mathbf{y}}^{[0]}$, and $\Omega_{\mathbf{x}}^{[0]}$, can be estimated. The K -means algorithm [41] with $K = 3$ has been tested, leading to poor results due to the high variability of the background region. Instead, we propose the following approach [42], which has shown successful results. A non-overlapping window is shifted along the image \mathbf{y} and the mean m_w of the pixel intensities is calculated for each position, which results in the vector $\mathbf{m} = (m_1, m_2, \dots, m_W)$. The values in \mathbf{m} are sorted in ascending order, and the resulting

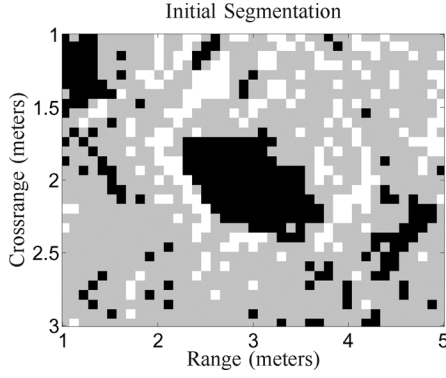


Fig. 6. Initial segmentation of the sonar image in Fig. 2. The label for the background pixels B is depicted in gray, the shadow label S in black and the highlight label H in white.

vector is split into two vectors \mathbf{m}_{low} and \mathbf{m}_{high} of the same length. The K -means algorithm is then applied to divide each vector into $K = 2$ clusters. Hence, \mathbf{m}_{low} splits into $\mathbf{m}_{\text{low}}^1$ and $\mathbf{m}_{\text{low}}^2$, and $\mathbf{m}_{\text{high}}^1$ and $\mathbf{m}_{\text{high}}^2$ stem from \mathbf{m}_{high} . The pixels in $\mathbf{m}_{\text{low}}^1$ are labeled as shadow S . The clusters $\mathbf{m}_{\text{low}}^2$ and $\mathbf{m}_{\text{high}}^1$, that is, those with intermediate intensity, are merged together and assigned to the background B . Finally, the elements that belong to the cluster with the highest intensity, $\mathbf{m}_{\text{high}}^2$, are assigned to the highlight H .

In general, the smaller the window size, the more accurate is the initial segmentation result. However, it also encourages the appearance of small negligible groups of dark pixels being classified as S , or small groups of light pixels being classified as highlight. For the available data, a window of 6×6 pixels has demonstrated to produce good results. The initial segmentation of the sonar image in Fig. 2 is shown in Fig. 6.

4) *Parameter Estimation. The ICE Algorithm:* The parameters Ω_y and Ω_x are estimated iteratively by means of the ICE algorithm [43]. After initializing $\Omega_y^{[0]}$ and $\Omega_x^{[0]}$, the algorithm performs the following steps at each iteration l :

- Use $\Omega_x^{[l]}$ to calculate the Markovian probability $P(x_i)$ for the three labels: $P(x_i = S)$, $P(x_i = B)$ and $P(x_i = H)$.
- Use $\Omega_y^{[l]}$ to calculate the likelihood functions for each pixel: $P(y_i|x_i = S)$, $P(y_i|x_i = B)$ and $P(y_i|x_i = H)$.
- Calculate the *a posteriori* probability for each label:

$$P(x_i|y_i) = P(x_i) \cdot P(y_i|x_i), \quad x_i = S, B, H. \quad (7)$$

- Use the Gibbs sampling algorithm to obtain n samples of the label field, $\hat{\mathbf{x}}_{(1)}, \dots, \hat{\mathbf{x}}_{(n)}$, according to $P(\mathbf{x}|\mathbf{y})$ and using $\Omega_y^{[l]}$ and $\Omega_x^{[l]}$.
- For each sample $\hat{\mathbf{x}}_{(m)}$, estimate the parameter vectors $\hat{\Omega}_x(\mathbf{x}_{(m)})$ and $\hat{\Omega}_y(\mathbf{x}_{(m)})$, $1 \leq m \leq n$ (as described in Sections II-A1 and II-A2, respectively).
- Calculate $\Omega_x^{[l+1]}$ as

$$\Omega_x^{[l+1]} = \frac{1}{n} \sum_{m=1}^n \hat{\Omega}_x(\mathbf{x}_{(m)}) \quad (8)$$

and analogously, calculate $\Omega_y^{[l+1]}$ from $\hat{\Omega}_y(\mathbf{x}_{(m)})$, $1 \leq m \leq n$.

In our implementation, $n = 1$ has been chosen (see [38] for details).

5) *Segmentation. The ICM Algorithm:* In order to estimate $\hat{\mathbf{x}}$, a local maximum of the *a posteriori* probability $P(\mathbf{x}|\mathbf{y})$ is found by the ICM algorithm. First, $P(x_i|y_i)$ is calculated for each pixel and each label using the Ω_y and Ω_x estimates. Then, each pixel is considered in turn and is assigned a label according to the maximum *a posteriori* (MAP) criterion:

$$\hat{x}_i = \arg \max_{x_i} \{P(x_i|y_i)\} = \arg \max_{x_i} \{P(x_i) \cdot P(y_i|x_i)\}. \quad (9)$$

This process is repeated until convergence.

B. Active Contours Algorithm

If the position of the region of interest within the image is known, the AC segmentation algorithm can be applied. Typically used for medical imaging applications [44], the AC algorithm has been successfully applied to sidescan sonar image segmentation [34], [45]. We suggest its application to SAS images in this contribution. Unlike the Markovian segmentation approach, the AC algorithm does not assume any *a priori* probability of the regions, only the intensity values of the pixels influence the segmentation result.

An active contour (or statistical snake) $\mathbf{n} = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_N\}^T$, is a closed curve defined by a list of N nodes that has the ability to evolve in order to match the contour of an object present in an image [28]. Each node, expressed in Cartesian coordinates $\mathbf{n}_j = (u_j, v_j)^T$, corresponds to a pixel i in \mathcal{L} . The image of interest is therefore divided into two regions, the target region \mathcal{T} inside \mathbf{n} and the background region \mathcal{B} outside \mathbf{n} . For segmentation of sonar images, the former corresponds to either the shadow or the highlight of an object.

1) *Cost Function:* The objective of the AC algorithm is to deform \mathbf{n} in such a way that a given cost function $F(\mathbf{n})$ is minimized. Originally, the AC algorithm was based on the gradient so that $F(\mathbf{n})$ is minimum when \mathbf{n} coincides with an edge between two regions. However, this approach performs poorly for noisy images and is very sensitive to initialization. Another approach to the problem consists of using parametric shape templates [46], [47]. Such templates are in general too stiff for an application like SAS, where the target regions present a broad variability.

A polygonal active contour that minimizes a cost function based on the likelihood function of the image is considered here [48]. Assuming independence among the pixels, the cost function associated with a position of the active contour \mathbf{n} is defined as the negative log-likelihood function of \mathbf{y}

$$F(\mathbf{n}) = -\ln \left(\prod_{y_i|i \in \mathcal{B}} P_{\mathcal{B}}(y_i) \cdot \prod_{y_i|i \in \mathcal{T}} P_{\mathcal{T}}(y_i) \right) \quad (10)$$

where $P_{\mathcal{T}}$ and $P_{\mathcal{B}}$ refer to the pdfs of the target and background regions, respectively, and need to be estimated. An efficient implementation of (10) is possible for certain parametric models of the pdfs [48]. However, the Weibull distribution, which is an appropriate model for SAS images (see Section II-A2), does not lead to this. We have carried out an extensive empirical study to evaluate typical parameter settings for the target and background regions. It could be shown that a Rayleigh distribution

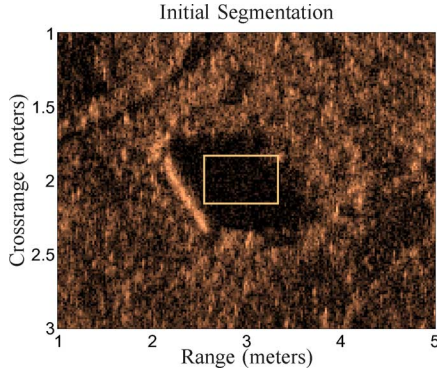


Fig. 7. Initialization of the AC algorithm. The result of the MRF segmentation is used.

is an acceptable approximation of the true pdfs and will be used here for the sake of an efficient AC implementation. Thus, P_T and P_B are modeled by the Rayleigh distributions $\mathcal{R}(\alpha_T)$ and $\mathcal{R}(\alpha_B)$, respectively. The parameter α_T is estimated according to

$$\alpha_T = \arg \max_{\alpha} \prod_{\{y_i | i \in T\}} \frac{y_i}{\alpha^2} \exp\left(-\frac{y_i^2}{2\alpha^2}\right) \quad (11)$$

and analogously for α_B .

2) *Initialization*: The initialization of the AC algorithm is crucial. If \mathbf{n} is initialized without comprising at least a piece of a target, the contour will diverge. Given that the Markovian segmentation result is available, it can be used to initialize the AC. However, if the AC initialization is identical to the Markov based final segmentation, the AC final result in general is not significantly different from it. On the other hand, the purpose of using two segmentation algorithms is their ability to produce different results when the borders between regions are not clear (see Section IV-F). Hence, the initialization of \mathbf{n} is based on the Markovian result but is, at the same time, significantly different: a rectangle centered in the center of mass of the Markovian segmented region has been adopted (see Fig. 7).

3) *The Algorithm*: The number of times that nodes are added to the initial 4-node contour $\mathbf{n}^{[0]}$ is counted by the index l . When the optimal position for $\mathbf{n}^{[l]}$ is achieved, $\mathbf{n}^{[l+1]}$ is initialized: if two consecutive nodes are further apart than a certain distance d_n , N_n nodes are inserted between them. Good results have been found for $d_n = 6$ and $N_n = 1$ or 2.

The index k is set to 0 every time that l increases and counts the number of iterations that are performed for a fixed amount of nodes. At each iteration, a single node \mathbf{n}_j of the current best active contour, $\mathbf{n}^{[l]}$, is shifted by a random distance, resulting in $\hat{\mathbf{n}}^{[l]}$. If $F(\hat{\mathbf{n}}^{[l]}) < F_{\min}$, $\hat{\mathbf{n}}^{[l]}$ is adopted. Furthermore, the so-called crossing test [48] is performed to prevent intersections between the different segments of \mathbf{n} .

A common problem of the AC algorithm is that \mathbf{n} may converge to local minima, leading to a poor segmentation. If a number of iterations N_{\max} elapses after the last time that new nodes were added and no update was produced in $\mathbf{n}^{[l]}$, we might be stuck in a local minimum. Thus, an alternative active contour, $\tilde{\mathbf{n}}^{[l]}$, is calculated by shifting all nodes in $\mathbf{n}^{[l]}$ by Δ pixels. If the final cost function $F(\tilde{\mathbf{n}}^{[l]})$ is smaller than the original cost function $F(\mathbf{n}^{[l]})$, we set $\mathbf{n}^{[l]} := \tilde{\mathbf{n}}^{[l]}$. This modification

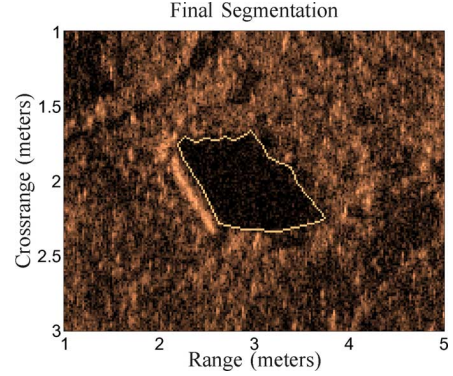


Fig. 8. Final AC segmentation result. A Rayleigh parametric model has been adopted.

of the standard AC algorithm has demonstrated to improve its performance in more than 35% of the cases. Good results have been obtained for $\Delta = 1$ pixel and $N_{\max} = 200$ iterations.

Hence, the implementation of the algorithm reads as follows:

- Initialize $\mathbf{n}^{[0]}$ with 4 nodes, estimate P_B and P_T , and calculate $F_{\min} := F(\mathbf{n}^{[0]})$
- Iterate while nodes in $\mathbf{n}^{[l]}$ are sparse
 - 1) $k := 0$
 - 2) While $k < N_{\max}$
 - $k := k + 1$
 - Choose a node $\mathbf{n}_j \in \mathbf{n}^{[l]}$ (randomly or sequentially) and shift it by a random distance: $\hat{\mathbf{n}}^{[l]}$
 - Estimate P_B and P_T and calculate $F(\hat{\mathbf{n}}^{[l]})$
 - If $F(\hat{\mathbf{n}}^{[l]}) < F_{\min}$, $\mathbf{n}^{[l]} := \hat{\mathbf{n}}^{[l]}$, $F_{\min} := F(\hat{\mathbf{n}}^{[l]})$
 - 3) If $\mathbf{n}^{[l]} = \mathbf{n}^{[l-1]}$ (local minimum?), $\tilde{\mathbf{n}}^{[l]} := \mathbf{n}^{[l]} + \Delta$, else $\mathbf{n}^{[l+1]} := \text{add nodes}(\mathbf{n}^{[l]})$

The AC segmentation result of the example SAS image is shown in Fig. 8.

C. Segmentation Results

Both segmentation algorithms have been applied to a set of real SAS images. We have processed more than 10 000 square meters, from six different images. A broad variability of the seabed is present. The images have a resolution of 2.5 cm in both range and crossrange directions. In this section, exemplary results are shown.

In order to apply the Markovian segmentation algorithm to SAS images of large size, they have been split into subimages of 200×200 pixels. Since the Ω_x and Ω_y parameters vary within big SAS images, this allows a better segmentation result. Hence, the Markovian algorithm is applied to each subimage and the resulting label fields are merged. We discard shadow and highlight pixel clusters that are too small to belong to an object of interest. With the rest, the following database is created. It is known from the literature [18], [34] that an object laying on the seabed produces a shadow but not always a highlight. Hence, an entry is created in the database for each shadow cluster. If there are highlight clusters to the left of a shadow and within a reasonable distance they are associated with that shadow. Highlights that are not associated to any shadow are discarded. Therefore, an object consists of a shadow and, sometimes, one or more highlights. The AC algorithm is applied to each shadow and highlight of the database.

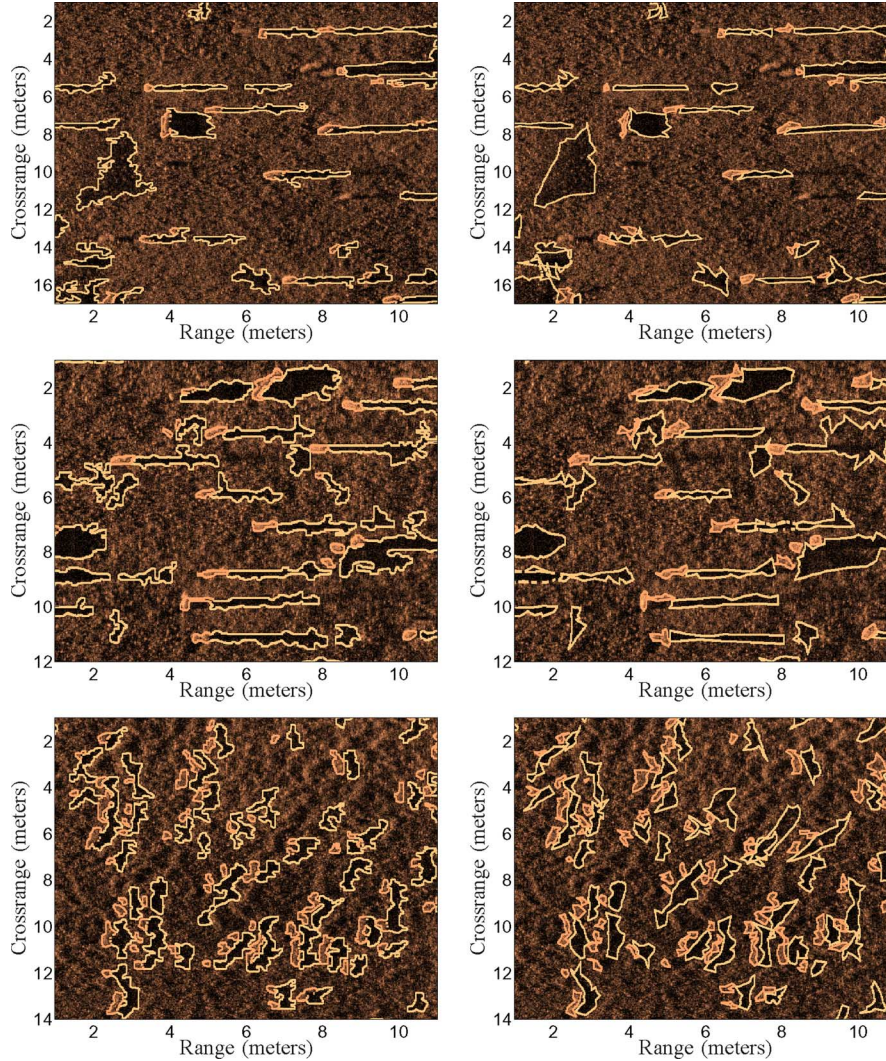


Fig. 9. Markovian (first column) and AC (second column) segmentation result for different SAS images. While the images in the two first images contain several man-made objects, the last image corresponds to a seabed area full of sand ripples.

In Fig. 9, some parts of the segmented images are shown. The top row contains the Markovian segmentation results, while the AC results are shown in the bottom row. The two first images correspond to scenarios where plenty of man made objects are present. They are all successfully segmented, together with several dark areas of the seabed. The last image shows a seabed where deep sand ripples are present. In this case, the Markovian algorithm segments as shadow the dark areas of the sand ripples, and as highlight the lightest parts to the left of the shadow areas.

A total of 1215 objects were segmented. Out of them, 1037 correspond to physical natural features of the terrain, while 178 are man made objects. There are 115 spheres and 63 cylinders. Although the objects in the available SAS database are not mines, their sizes and shapes are comparable, and therefore serve our purpose. Ideal spherical objects consist of a more or less circular highlight and an elongated shadow. Cylindrical objects present an elongated highlight and a rhomboid like shadow. While spheres are insensitive to orientation, the position of the cylinder with respect to the sonar influences the

shadow shape. The natural objects present a great variability. A representative example of each object kind is shown in Fig. 10.

III. STATISTICAL CLASSIFICATION

The object database referred in Section II-C has been used to study the performance of different sets of features to classify underwater objects as man made or natural objects. Visual inspection and classification by several experts was used to generate the ground truth. If the available database is too small, resampling schemes such as the bootstrap may be used [49].

Three classes are considered: spheres \mathcal{S} , cylinders \mathcal{C} , and natural objects \mathcal{N} . Some natural objects are much bigger or smaller than typical man made objects and can be directly classified as \mathcal{N} . In our database, 22% of the \mathcal{N} objects fall into this category. For the remaining ones, a statistical three-class classifier based on the Mahalanobis distance has been implemented. If new classes of objects were to be detected, the classifier could be modified accordingly, as long as enough samples are available in order to estimate the features of the new classes. The

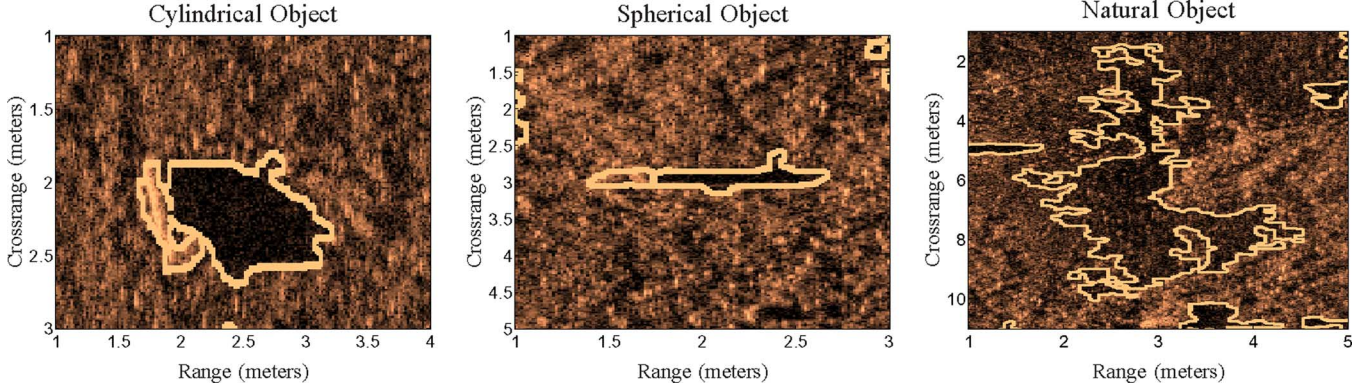


Fig. 10. The first and second images represent the segmentation of the highlight and the shadow of a cylindrical and an spherical man-made object. The last image corresponds to a physical feature of the terrain that has been segmented as shadow.

TABLE I

ESTIMATED MEAN $\hat{\mu}$ AND STANDARD DEVIATION $\hat{\sigma}$ OF THE SHADOW GEOMETRICAL FEATURES (SEE SECTION IV-A), FOR THE \mathcal{S} , \mathcal{C} AND \mathcal{N} CLASSES

Feature	ρ		A		χ		r_S		O		Γ	
Class	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
\mathcal{N}	205	112	1.3e3	1.2e3	35	12	2.1	0.6	8.5	66	0.77	0.08
\mathcal{S}	305	125	1.8e3	0.9e3	52	18	4.2	1.8	1.6	7	0.67	0.11
\mathcal{C}	309	88	3.1e3	1.2e3	31	9	1.9	1	2.3	49	0.82	0.07

TABLE II

NUMBER OF HIGHLIGHTS ASSIGNED TO AN OBJECT AS A FUNCTION OF THE OBJECT CLASS, IN PERCENTAGE

Class	N_H				
	0	1	2	3	> 3
\mathcal{N}	39	48	10	2	1
\mathcal{S}	14	77	8	1	0
\mathcal{C}	1	73	20	3	0

TABLE III

ESTIMATED MEAN $\hat{\mu}$ AND STANDARD DEVIATION $\hat{\sigma}$ OF THE HIGHLIGHT GEOMETRICAL FEATURES (SEE SECTION IV-B), FOR THE \mathcal{S} , \mathcal{C} , AND \mathcal{N} CLASSES

Feature	r_H		W		Δ_O	
Class	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
\mathcal{N}	2.1	0.8	0.45	0.26	8.9	11.6
\mathcal{S}	1.8	0.5	0.48	0.19	10.9	14.2
\mathcal{C}	3.9	1.2	0.75	0.18	3.8	4.9

TABLE IV

ESTIMATED MEAN $\hat{\mu}$ AND STANDARD DEVIATION $\hat{\sigma}$ OF THE SAMPLE NORMALIZED CENTRAL MOMENTS OF ORDER 2 AND 3 (SEE SECTION IV-C). ALL QUANTITIES ARE NORMALIZED BY 10^{-3}

Feature	C_{11}		C_{02}		C_{20}		C_{12}		C_{21}		C_{30}		C_{03}	
Class	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
\mathcal{N}	0.9	58	170	61	90	47	-0.4	12	-0.8	9	-0.1	10	0.4	19
\mathcal{S}	-0.9	48	40	24	450	176	5	12	-1.2	17	41	90	1	11
\mathcal{C}	-1.3	42	90	47	130	102	-0.5	7.6	0.3	7	1	13	-0.2	12

features are considered in Section IV, but we describe the classifier first.

Let us assume a set of M features, $\mathbf{f} = (f_1, f_2, \dots, f_M)$, to describe each object. For simplicity, it is assumed that the pdf of each feature is Gaussian. Then, the pdf of the feature vector is given by the multivariate Gaussian distribution [50]

$$P_t(\mathbf{f}) = \frac{1}{(2\pi)^{M/2}} |\Sigma_t|^{1/2}$$

TABLE V

ESTIMATED MEAN $\hat{\mu}$ AND STANDARD DEVIATION $\hat{\sigma}$ OF THE PRINCIPAL COMPONENTS ζ_b , $b = 1, \dots, 6$ (SEE SECTION IV-D)

Feature	ζ_1		ζ_2		ζ_3		ζ_4		ζ_5		ζ_6	
Class	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
\mathcal{N}	2.8	7.2	-0.5	7	-0.16	8.2	-0.18	4.2	-0.04	4.2	0.2	3.8
\mathcal{S}	-14	8.3	0.6	9	-0.18	2.3	1.2	3.8	0.07	2.2	0.4	2
\mathcal{C}	-3.7	7.6	-2.7	12	-0.28	5.3	0.26	2.9	-0.07	3.5	-0.8	3.8

$$\cdot \exp \left[-\frac{1}{2} (\mathbf{f} - \bar{\mathbf{f}}_t)' \Sigma_t^{-1} (\mathbf{f} - \bar{\mathbf{f}}_t) \right], \quad t = \mathcal{S}, \mathcal{C}, \mathcal{N} \quad (12)$$

where the mean vector $\bar{\mathbf{f}}_t$ and the covariance matrix Σ_t are estimated from the data.

We need to divide the feature space into three exclusive regions, R_S , R_C , and R_N . A given object is classified as \mathcal{S} if its feature vector \mathbf{f} belongs to R_S , and analogously for \mathcal{C} and \mathcal{N} . The probability of classifying a natural object as sphere or cylinder is known as the false alarm rate P_{fa} . The probability of misclassifying \mathcal{S} or \mathcal{C} objects are denoted by $P_m^{\mathcal{S}}$ and $P_m^{\mathcal{C}}$. The misclassification probability P_m is defined as $P_m = \max\{P_m^{\mathcal{S}}, P_m^{\mathcal{C}}\}$. While classifying a sphere as a cylinder or a cylinder as a sphere is a minor problem, classifying them as natural objects is critical. Although this distinction is not considered in the calculation of P_m , it is shown in the confusion matrices (see Tables VII and VIII). The probability of correct classification equals $P_c = 1 - P_m$.

Let us define

$$\Psi_{tr}(\mathbf{f}) = \ln \frac{P_t(\mathbf{f})}{P_r(\mathbf{f})} \quad t, r = \mathcal{S}, \mathcal{C}, \mathcal{N}, \quad r \neq t. \quad (13)$$

The region R_t is then determined by

$$\Psi_{tr}(\mathbf{f}) > Q_t - Q_r, \quad t = \mathcal{S}, \mathcal{C}, \mathcal{N}, \quad r \neq t. \quad (14)$$

The constants Q_t are chosen such that P_c is maximized for a given P_{fa} .

The “leave-one-out” technique has been adopted. For each object, we estimate the parameters $\bar{\mathbf{f}}_t$ and Σ_t using the complete training set but the object under consideration. Then, $\Psi_{tr}(\mathbf{f})$ for $t = \mathcal{S}, \mathcal{C}, \mathcal{N}$, $r \neq t$ are computed. The object is classified as \mathcal{S} if $\Psi_{\mathcal{S}\mathcal{C}}(\mathbf{f}) > Q_{\mathcal{S}} - Q_{\mathcal{C}}$ and $\Psi_{\mathcal{S}\mathcal{N}}(\mathbf{f}) > Q_{\mathcal{S}} - Q_{\mathcal{N}}$, and analogously for \mathcal{C} and \mathcal{N} .

TABLE VI
ESTIMATED MEAN $\hat{\mu}$ AND STANDARD DEVIATION $\hat{\sigma}$ OF THE WEIBULL
PARAMETERS $\xi_S, \lambda_S, \xi_B, \lambda_B, \xi_H$ AND λ_H (SEE SECTION IV-E)

Feature	ξ_S		λ_S		ξ_B		λ_B		ξ_H		λ_H	
Class	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
\mathcal{N}	1.7	0.25	34	4.2	2.19	0.33	53	5.6	3.5	0.63	86	8.8
\mathcal{S}	1.6	0.14	31.6	2.6	2.16	0.14	57.3	4	3.5	0.64	108	11.5
\mathcal{C}	1.5	0.17	27.7	3.2	2.13	0.19	55.8	4.4	3.2	0.44	104	10.2

TABLE VII
CONFUSION MATRIX OF \mathbf{f}^* AT $P_{fa} = 0.07$. THIS TABLE
ILLUSTRATES THE CLASSIFICATION RESULT FOR EACH CLASS WITH
REFERENCE TO THE GROUND TRUTH

Ground Truth	Decision Class		
	\mathcal{N}	\mathcal{S}	\mathcal{C}
\mathcal{N}	711	6	50
\mathcal{S}	2	109	4
\mathcal{C}	1	2	60

TABLE VIII
CONFUSION MATRIX OF $(\mathbf{f}_S, \mathbf{f}_H)$ AT $P_{fa} = 0.07$. WHILE $(\mathbf{f}_S, \mathbf{f}_H)$ REACHES
 $P_c = 0.88, P_c(\mathbf{f}^*) = 0.95$

Ground Truth	Decision Class		
	\mathcal{N}	\mathcal{S}	\mathcal{C}
\mathcal{N}	711	10	46
\mathcal{S}	10	102	3
\mathcal{C}	3	4	56

IV. FEATURE DESCRIPTION OF UNDERWATER OBJECTS

The literature offers a large number of descriptors for shape representation [24], [51]. They can be divided into two main groups: those that model the contour of the region of interest, and those that model the region itself. For SAS images, the latter are more appropriate, since the variability of the contour is not meaningful. Several sets of shape descriptors have been considered. Some simple geometrical descriptors, normalized central moments and principal component analysis haven been applied to the shape of shadow region. We have represented the high-light shape by some geometrical features. Moreover, the statistical properties of the objects have been exploited as well as the comparison of the segmentation results provided by the Markovian and AC algorithms. Finally, an optimal set of features is proposed.

A. Shadow Geometrical Features

Two straightforward features are the area A and perimeter ρ of the shadow. The perimeter tends to reach higher values for the \mathcal{S} and \mathcal{C} classes than for \mathcal{N} (see Table I). The area of \mathcal{C} objects is in average greater than the \mathcal{S} and \mathcal{N} area. Both the perimeter and the area of the \mathcal{N} objects present heavy tailed histograms due to the variability of the natural objects.

The ratio between the squared perimeter and the area, known as compactness χ [52], has also been considered. The compactness is minimum for a circle and tends to infinity as the shape approaches a straight line. Hence, it reaches high values for the elongated shadow of the sphere class \mathcal{S} . There are two other features, the ratio of principal axes r_S and the orientation O , which also have distinct values for \mathcal{S} objects. The principal axes of a region are defined as two line segments that cross orthogonally in the center of mass and represent the directions with zero

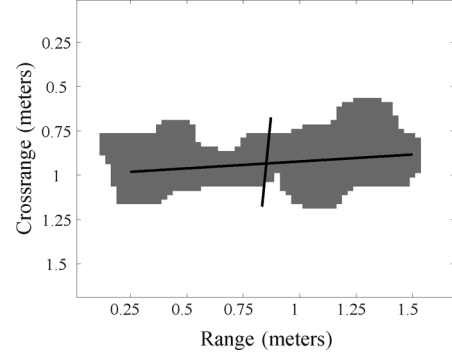


Fig. 11. Principal axes of the shadow of an \mathcal{S} object. The angle between the major axis and the abscissa is the orientation of the region.

cross-correlation (see Fig. 11). For a given region with contour \mathbf{n} and center of mass $\bar{\mathbf{n}}$, the covariance matrix of the contour is defined by

$$\mathbf{\Upsilon} = \sum_{j=1}^N (\mathbf{n}_j - \bar{\mathbf{n}})(\mathbf{n}_j - \bar{\mathbf{n}})^T. \quad (15)$$

The ratio of principal axes r_S equals the coefficient of the eigenvalues of $\mathbf{\Upsilon}$ and is a good measure of the elongation of the region defined by \mathbf{n} . For \mathcal{S} objects, r_S is significantly higher than for the \mathcal{C} and \mathcal{N} classes. The orientation is the angle that the major axis of the shadow forms with the abscissa of the image. Its value is comprised in the interval from -10° to 10° for about 95% of the \mathcal{S} objects, while it is roughly uniformly distributed between -90° and 90° for \mathcal{C} and \mathcal{N} objects.

Finally, the solidity Γ is the coefficient between the area of the region and the minimal convex area that comprises it [52]. Many shadows of \mathcal{N} objects have a very low solidity. Since they have a greater area cylindrical objects are in general more solid than spheres. The solidity of the \mathcal{N} shadow in Fig. 12 equals 0.61. It is 0.85 for the cylinder and 0.72 for the sphere.

The estimated mean and standard deviation of all descriptors, ρ, A, χ, r_S, O , and Γ , are included in Table I. In Fig. 14, P_c is represented as a function of P_{fa} for the feature set $\mathbf{f}_S = (\rho, A, \chi, r_S, O, \Gamma)$. If $P_{fa} = 0.2$, the probability of correct classification is $P_c = 0.92$. The classification of \mathcal{S} objects as \mathcal{C} and vice versa is the reason for $P_c < 1$ when $P_{fa} = 1$.

B. Highlight Geometrical Features

While almost 40% of the \mathcal{N} objects lack a highlight, only 1% of the \mathcal{C} objects and 14% of the \mathcal{S} objects have no highlight. Thus, the presence or absence of a highlight is already a valuable feature of the object. In Table II, the number of highlights associated with each object N_H are specified.

A prominent feature for discerning \mathcal{C} objects is the ratio of principal axes of the highlight r_H . While for \mathcal{S} and \mathcal{N} , its mean value is around 2, it is almost 4 in the case of \mathcal{C} highlights. Another good feature for the \mathcal{C} class is the rate between the highlight and shadow widths along the crossrange direction, W . Again, it reaches higher values for \mathcal{C} than for \mathcal{S} and \mathcal{N} objects. Finally, the difference between the orientation of the right part of the highlight contour and the orientation of the left part of the shadow contour, Δ_O , has been considered (see Fig. 13). The influence of the highlight orientation on the orientation of

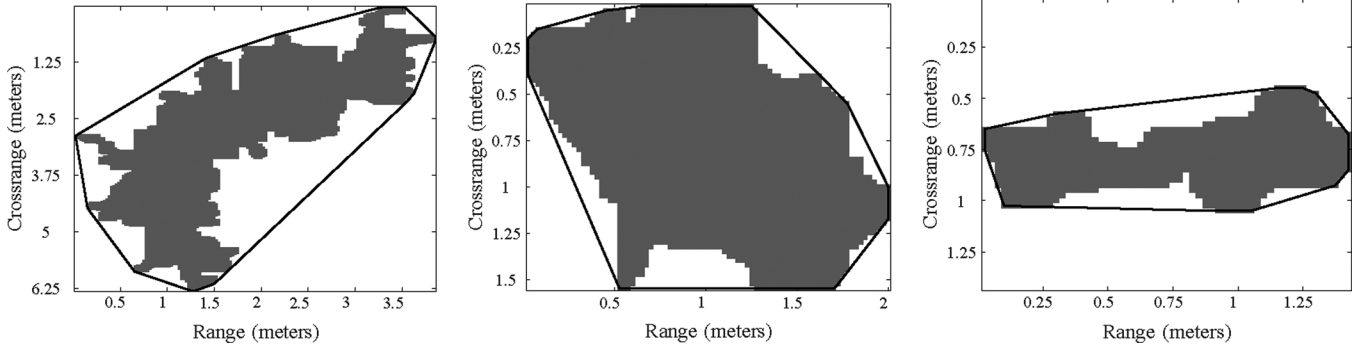


Fig. 12. Solidity of the shadow of a \mathcal{N} , \mathcal{C} , and \mathcal{S} object, respectively. Each shadow region is surrounded by the contour of its minimal convex hull.

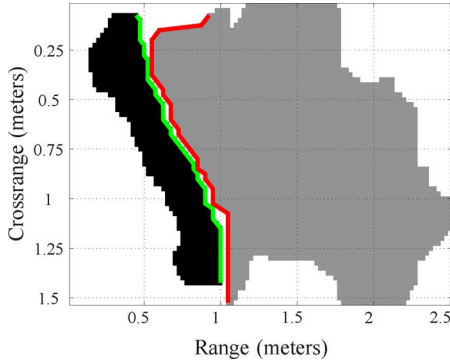


Fig. 13. Difference between the orientation of the right part of the highlight contour (in green) and the orientation of the left part of the shadow contour (in red) of a cylinder. For \mathcal{C} objects, this feature takes smaller values.

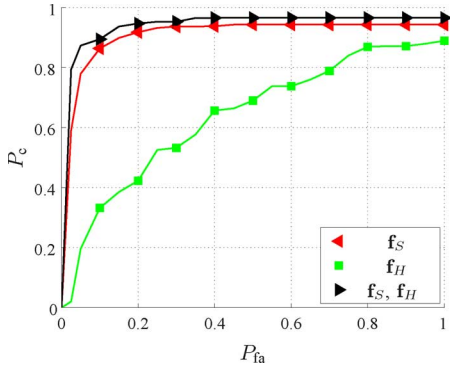


Fig. 14. Classification performance when the shadow geometrical descriptors \mathbf{f}_S , the highlight geometrical descriptors \mathbf{f}_H , and a combination of both are used.

the shadow is stronger for \mathcal{C} objects. The estimated mean and standard deviation of the highlight descriptors r_H , W , and Δ_O are included in Table III. Ideally, the classification should be performed according to $(\mathbf{f}_H, \mathbf{f}_S)$, where $\mathbf{f}_H = (r_H, W, \Delta_O)$. However, \mathbf{f}_H is only defined for those objects with highlight. A possible solution consists of using two different classifiers: one for those objects with highlight and another one for the objects without highlight. This has the disadvantage that not all the elements of the database are used, and the performance consequently decreases. The following classification procedure has been derived to overcome this problem. According to (13), $\Psi_{tr}(\mathbf{f}_S)$ for $t, r = \mathcal{S}, \mathcal{C}, \mathcal{N}$, with $r \neq t$ are computed. In the case where the object under consideration has associated highlights, $\Psi_{tr}(\mathbf{f}_H)$ is also calculated; otherwise, we assume $\Psi_{tr}(\mathbf{f}_H) = 0$. Assuming independence between \mathbf{f}_S and \mathbf{f}_H , the classification can be done according to $\Psi_{tr}(\mathbf{f}_S) + \Psi_{tr}(\mathbf{f}_H)$. In Fig. 14, the per-

formance of \mathbf{f}_H , \mathbf{f}_S , and the combination of both are compared. The highlight feature vector \mathbf{f}_H does not provide a good result on its own, but improves the performance of the shadow feature vector when both are combined. For $P_{fa} = 0.2$, the probabilities of correct classification are $P_c(\mathbf{f}_S) = 0.92$, $P_c(\mathbf{f}_H) = 0.42$, and $P_c(\mathbf{f}_S, \mathbf{f}_H) = 0.94$.

In the case where $N_H > 1$, one of the highlights has to be selected to compute \mathbf{f}_H . The following approach has been adopted. For each highlight h , its area A_h and the rate between its width and the shadow width along the crossrange direction W_h are calculated. We select the highlight that obtains the highest value for the coefficient

$$\theta_h = \tau \frac{A_h}{\max\{A_h\}} + (1 - \tau)W_h, \quad h = 1, \dots, N_H. \quad (16)$$

Good results have been obtained for $\tau = 0.5$.

C. Normalized Central Moments

For an image $Y(u, v)$, $u = 1, \dots, U_{\max}$, $v = 1, \dots, V_{\max}$, the sample central moment of order $(p + q)$ is defined as

$$c_{pq} = \sum_u \sum_v (u - \bar{u})^p (v - \bar{v})^q Y(u, v). \quad (17)$$

The origin of coordinates is placed at the center of mass of the image, (\bar{u}, \bar{v}) . Due to this normalization, c_{10} and c_{01} vanish. For a binary image that takes its values from $\{0, 1\}$, c_{00} equals the area A of the object in Y .

The normalized central moments are obtained normalizing c_{pq} with respect to the area

$$C_{pq} = \frac{c_{pq}}{c_{00}^\kappa}, \quad \text{where } \kappa = 1 + \frac{p+q}{2} \quad (18)$$

with $p = 0, \dots, p_{\max}$, $q = 0, \dots, q_{\max}$.

The moments of order 2 have a simple geometrical interpretation [53]. While C_{11} is related with the covariance of the region, C_{20} and C_{02} correspond, respectively, to the length of the major and minor axes of the ellipse that best fits the object in $Y(u, v)$.

The shadow normalized central moments C_{pq} up to order 4 have been tested for classification, taking c_{00} as an extra feature. In Table IV, the estimated mean and standard deviation of the moments of orders 2 and 3 are listed. While C_{11} is not a good descriptor, C_{02} and C_{20} take distinct values for the different classes, rendering C_{20} specially useful for \mathcal{S} discrimination. Also, C_{30} , C_{40} , and C_{04} are good descriptors. Let us denote the set of normalized central moments of order s by $\mathbf{C}^{[s]}$ so

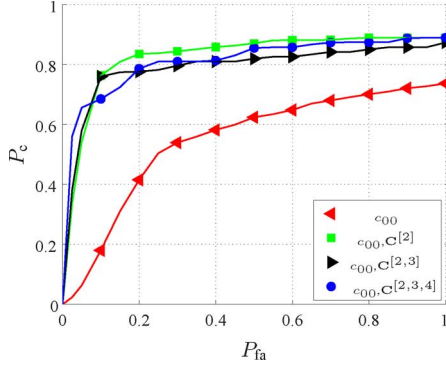


Fig. 15. Classification performance for the normalized central moments. For $P_{fa} > 0.12$, the feature vector that provides the best performance is $\mathbf{f} = (c_{00}, \mathbf{C}^{[2]})$.

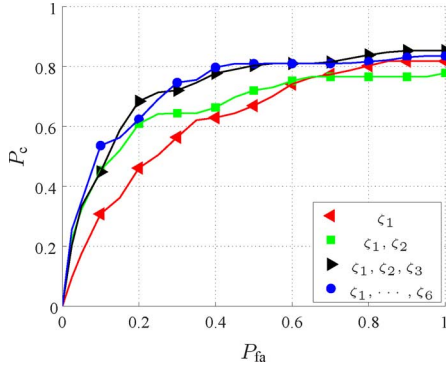


Fig. 16. Classification performance of the principal components. The performance improves when the number of ζ_i increases up to 3. For illustration purposes, the curve that corresponds to the first six principal components is included.

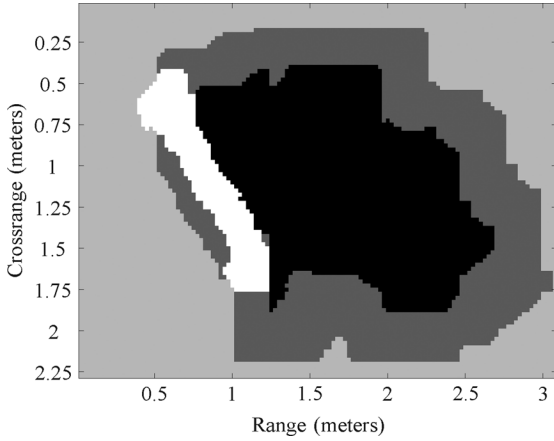


Fig. 17. Estimated label field for the sonar image in Fig. 2, where a stripe around the shadow indicates the pixels used for the background statistics estimation.

that, for instance, $\mathbf{C}^{[2]} = (C_{11}, C_{20}, C_{02})$. In Fig. 15, P_c versus P_{fa} is depicted for different combinations of $\mathbf{C}^{[s]}$. The best performance is achieved by $\mathbf{f} = (c_{00}, \mathbf{C}^{[2]})$, with a probability of correct classification $P_c = 0.82$ at $P_{fa} = 0.2$. For $P_{fa} \sim 0$, the highest P_c is obtained by $\mathbf{f} = (c_{00}, \mathbf{C}^{[2]}, \mathbf{C}^{[3]}, \mathbf{C}^{[4]})$.

D. Principal Components Analysis

Principal components analysis (PCA) [54] is a popular pattern recognition technique. It has been studied for sidescan sonar

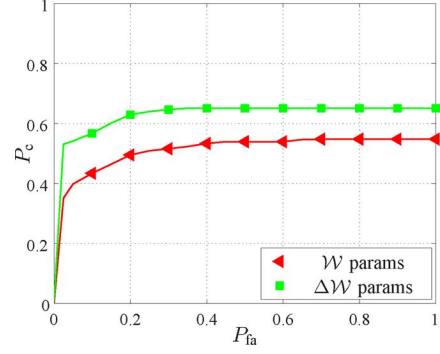


Fig. 18. Classification performance of the statistical features. The curve in red corresponds to the a feature vector formed by the Weibull parameters $\mathbf{f} = (\lambda_S, \lambda_H, \lambda_B, \xi_S, \xi_H, \xi_B)$, and the green line stems from $\mathbf{f} = (\Delta\lambda_{SB}, \Delta\lambda_{HB}, \Delta\lambda_{SH}, \Delta\xi_{SB}, \Delta\xi_{HB}, \Delta\xi_{SH})$.

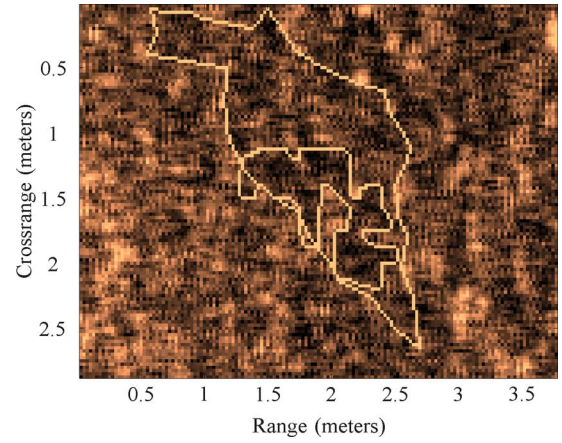


Fig. 19. Segmentation results provided by the Markovian and AC algorithms for an \mathcal{N} object. For shadow regions that are not well differentiated, both algorithms produce significantly different segmentations.

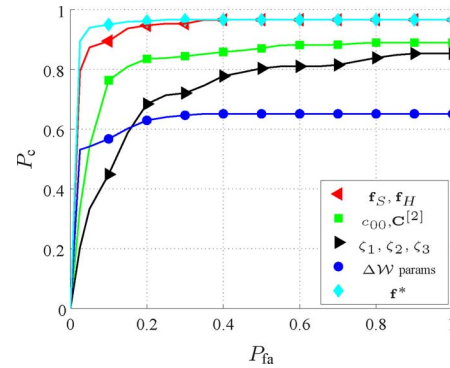


Fig. 20. Comparison of the classification performance for the optimal feature vector \mathbf{f}^* and other feature vectors. At a false alarm rate of $P_{fa} = 0.07$, a probability of correct classification $P_c = 0.95$ is achieved.

detection in [55], where PCA is applied directly to the sonar images. In the current approach, we use it on a binary representation of the shadow region. PCA consists of representing a set of correlated variables by a smaller number of uncorrelated ones, called principal components. In this application, the correlated variables are the pixels of each binary image that contains the shape of a shadow. The first principal component, ζ_1 , represents as much of the data correlation as possible, and each consecutive ζ_b accounts for as much of the remaining correlation as possible. For implementation details see [24] and [56].

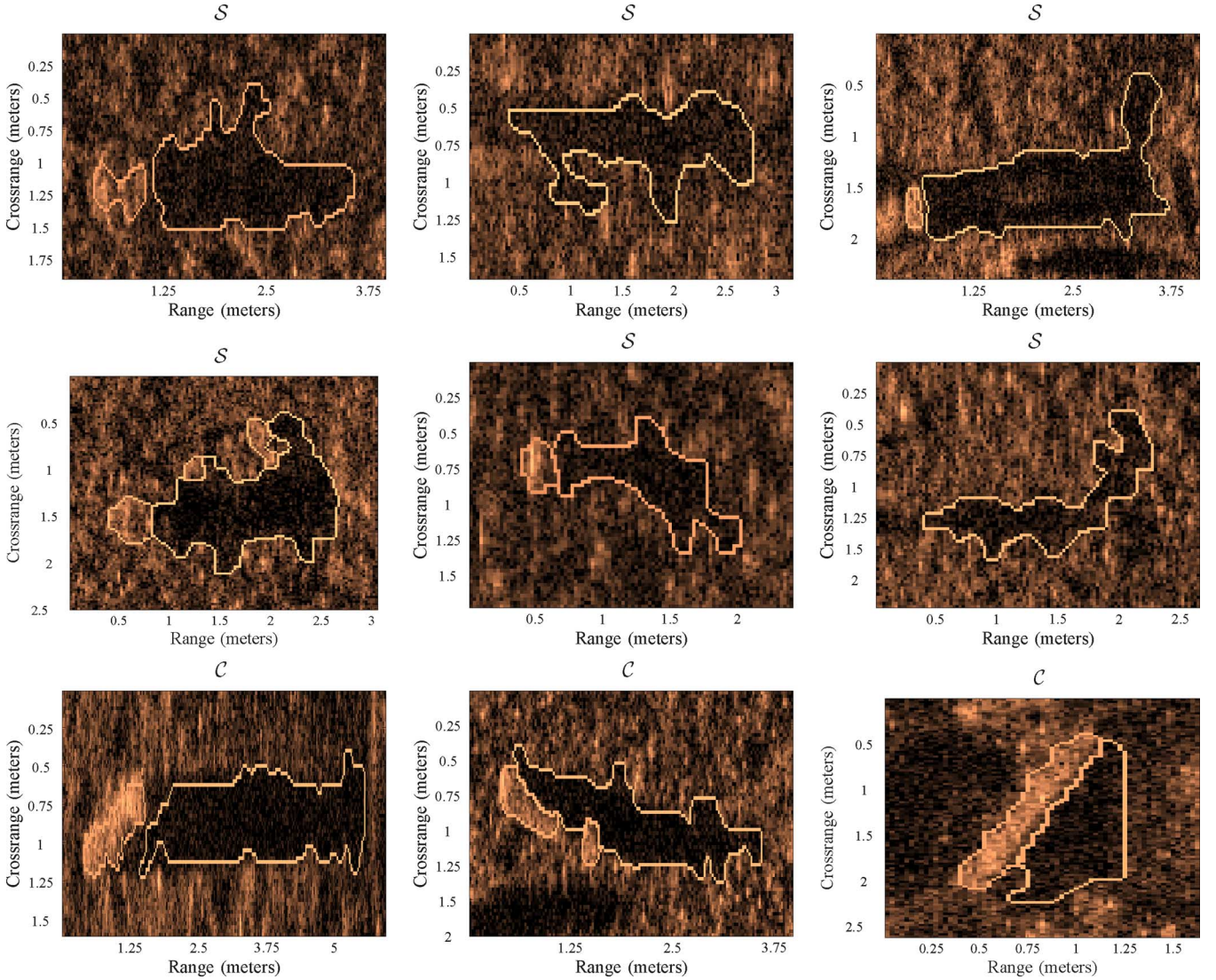


Fig. 21. All objects from the \mathcal{C} and \mathcal{S} classes that are misclassified with the optimal feature set. The four first spheres are classified as cylinders, and the two last ones as natural objects. The first two cylinders are classified as spheres, and the last one as a natural object.

The estimated mean and standard deviation of the first 6 ζ_b are included in Table V. It is ζ_1 the principal component that takes more distinct values from class to class, and also produces better classification results: $P_c = 0.45$ if $P_{fa} = 0.2$, when tested alone. These results are however outperformed when several ζ_b are included in \mathbf{f} (see Fig. 16). If $\mathbf{f} = (\zeta_1, \zeta_2)$, then P_c increases to 0.6, and to 0.65 when also ζ_3 is taken into account. A further increase of the number of principal components has no effect on the performance.

E. Statistical Features

As referred in Section II-A2, the different regions of an SAS image, \mathcal{S} , \mathcal{C} , and \mathcal{H} , are conveniently modeled by the Weibull distributions $\mathcal{W}(\lambda_S, \xi_S)$, $\mathcal{W}(\lambda_C, \xi_C)$, and $\mathcal{W}(\lambda_H, \xi_H)$, respectively. The parameters for the shadow and the highlight are estimated from the segmented regions. For the estimation of the background Weibull parameters, the entire background region

is not used, since the statistical properties might change due to its size. Instead, only a stripe of pixels around the shadow is used. The dilation morphological operation is applied to the segmented shadow in order to build this stripe. See Fig. 17 for an illustration.

The estimated mean and standard deviation for the three classes \mathcal{N} , \mathcal{S} , and \mathcal{C} are detailed in Table VI. While the shape factor ξ does not significantly differ from class to class, the scale factor λ helps to discriminate classes. The value of λ_H for \mathcal{N} objects is significantly lower than for \mathcal{C} or \mathcal{S} objects, while λ_S tends to be higher for \mathcal{N} than for \mathcal{C} or \mathcal{S} objects. This is due to the fact that \mathcal{C} and \mathcal{S} objects tend to have darker shadows and lighter highlights than the natural objects \mathcal{N} .

An alternative consists of using the distance between the parameters of the different regions as features, that is, $\mathbf{f} = (\Delta\lambda_{SB}, \Delta\lambda_{HB}, \Delta\lambda_{SH}, \Delta\xi_{SB}, \Delta\xi_{HB}, \Delta\xi_{SH})$, where $\Delta\lambda_{SB} = \lambda_S - \lambda_B$ and equivalently for the other elements. The classification performance improves, as can be observed in Fig. 18. For $P_{fa} = 0.2$, $P_c = 0.5$ if the Weibull parameters are used as feature vector and $P_c = 0.62$ when the distance

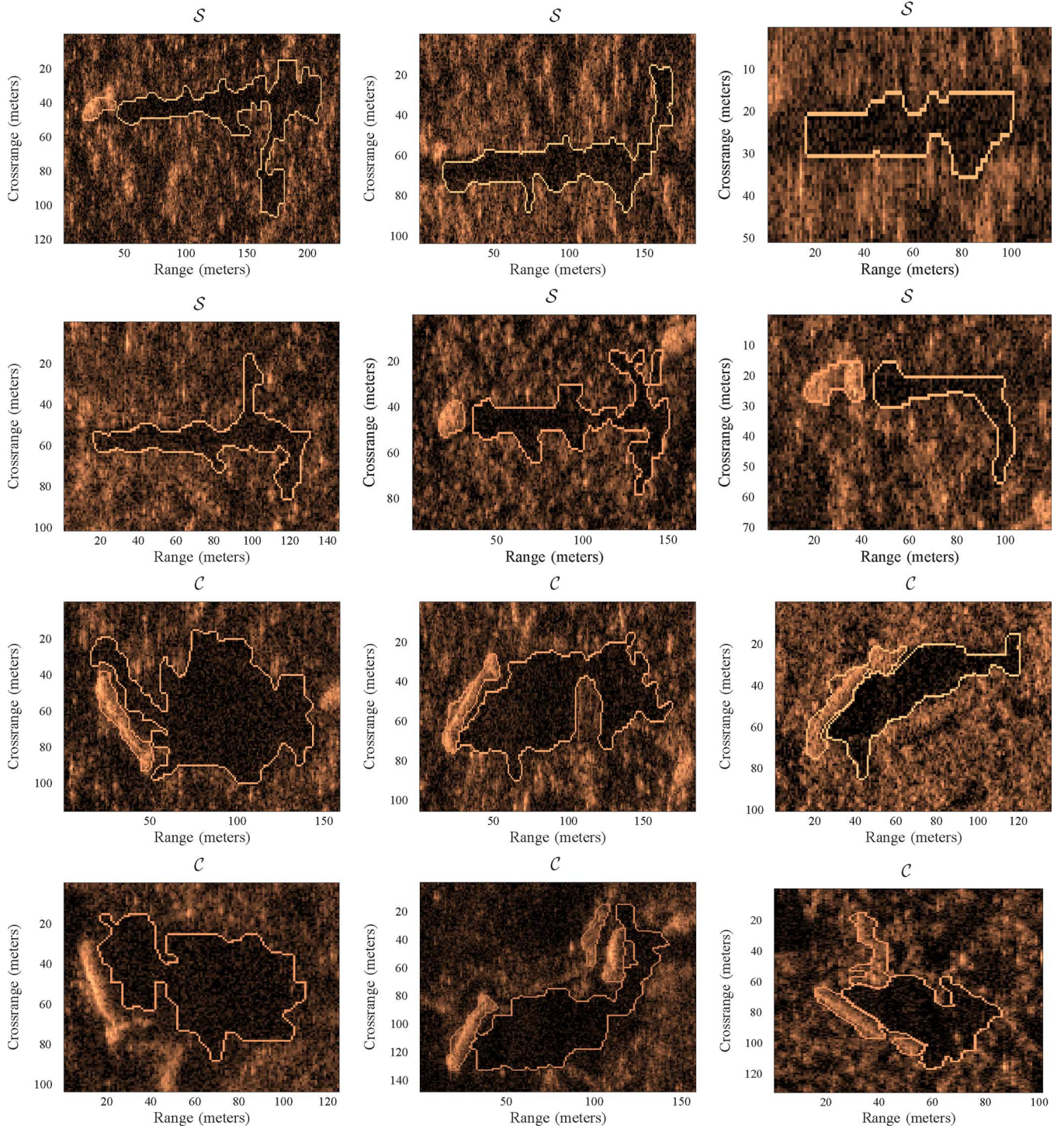


Fig. 22. Some of the most challenging \mathcal{C} and \mathcal{S} objects that are correctly classified with the optimal feature set.

between them is adopted. The performance of the statistical descriptors is quite poor compared with other feature sets but it improves the performance of the classifier when they are all combined (see Section IV-G).

F. MRF and AC Rate

The segmentation results provided by the Markovian and the AC algorithms are in general quite similar (see Fig. 9). However, when no man-made object is present in an image, it can happen that an area that is slightly darker than the background

is segmented as a shadow region. In these cases, the AC algorithm often provides a segmentation result that is significantly different from the Markovian segmentation, as illustrated by the example in Fig. 19. The rate Φ between the area of the region where both segmentation results intersect and the area of the AC segmented region constitutes a good measure of the segmentation reliability.

The mean value of Φ is 0.97 for the \mathcal{C} objects, 0.87 for \mathcal{S} and 0.73 for the \mathcal{N} class. Indeed, while Φ is lower than 0.5 for more than 20% of the \mathcal{N} objects, a negligible amount of \mathcal{S} and

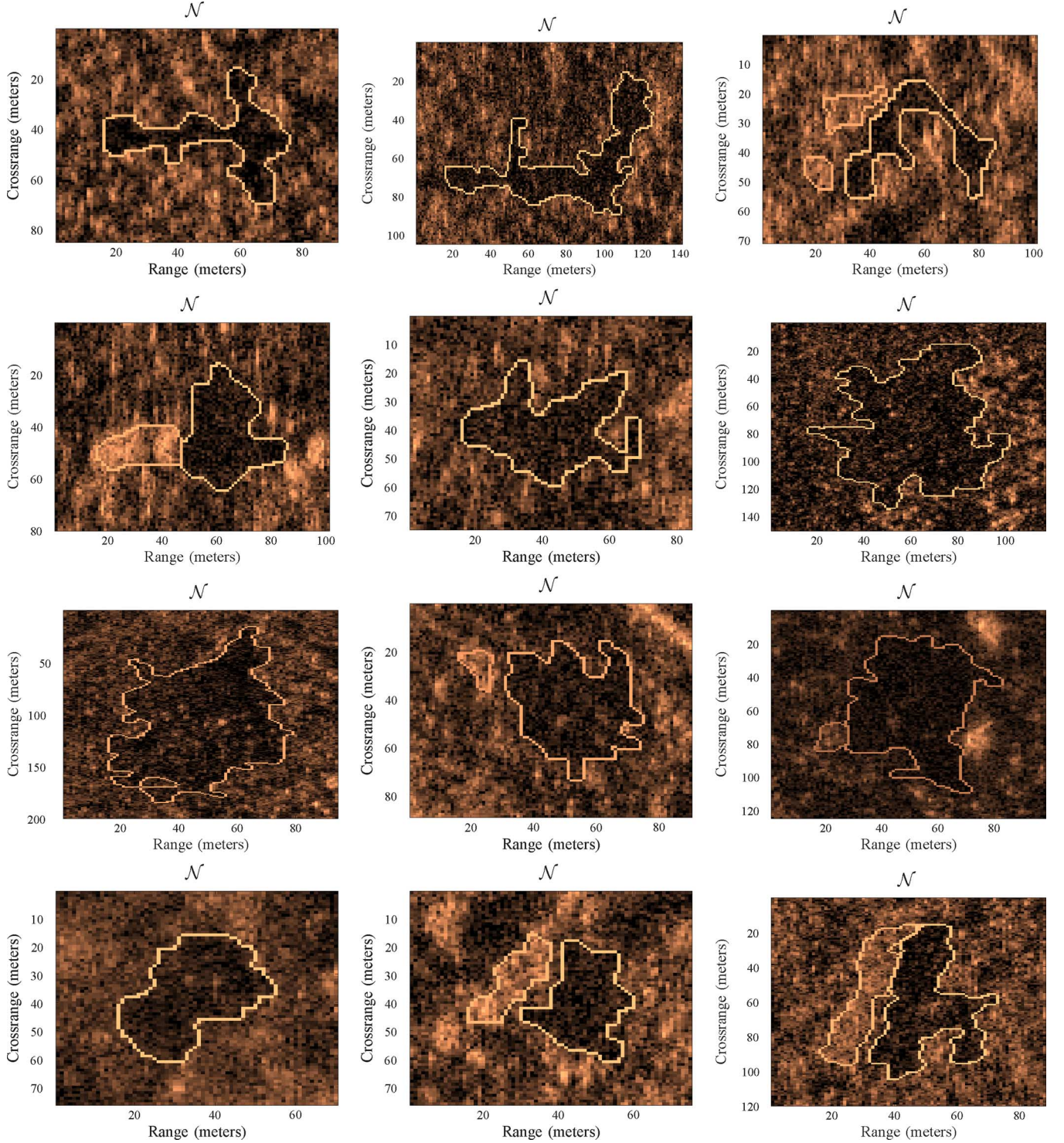


Fig. 23. Some of the \mathcal{N} objects that produce false alarms with the optimal feature set. The objects in the first row are classified as \mathcal{S} and the rest as \mathcal{C} .

\mathcal{C} objects present $\Phi < 0.5$. The estimated standard deviation is smaller than 0.1 for the \mathcal{C} and \mathcal{S} classes and 0.16 for the \mathcal{N} class.

G. Optimal Feature Set

Using all the features described above together does not produce good results due to the well-known curse of dimensionality [57]. Furthermore, some of the descriptors contain redundant information. Therefore, it is necessary to choose an optimal set of features, \mathbf{f}^* , that provides the best performance. For

the available database, the best performance has been observed for a feature vector formed by all shape and highlight geometrical descriptors, the normalized central moments C_{20} and C_{02} , the second principal component ζ_2 , the statistical descriptors $\Delta\lambda_{SB}$, $\Delta\lambda_{HB}$, and $\Delta\lambda_{SH}$, and the parameter Φ , that is,

$$\mathbf{f}^* = (\rho, A, \chi, r_S, O, \Gamma, r_H, W, \Delta_O, C_{20}, \dots, C_{02}, \zeta_2, \Delta\lambda_{SB}, \Delta\lambda_{HB}, \Delta\lambda_{SH}, \Phi). \quad (19)$$

Including other features has either no effect on P_c or degrades it. The comparison of \mathbf{f}^* with some of the feature vectors described above is illustrated in Fig. 20. For $P_{fa} > 0.4$, the performance of \mathbf{f}^* is identical to that of $(\mathbf{f}_S, \mathbf{f}_H)$. However, when a low probability of false alarm is desired, \mathbf{f}^* clearly outperforms $(\mathbf{f}_S, \mathbf{f}_H)$. Hence, for $P_{fa} = 0.07$, the feature vector \mathbf{f}^* reaches a probability of correct classification $P_c = 0.95$ while for $(\mathbf{f}_S, \mathbf{f}_H)$, $P_c = 0.88$. The confusion matrices for both \mathbf{f}^* and $(\mathbf{f}_S, \mathbf{f}_H)$ at $P_{fa} = 0.07$ are included in Tables VII and VIII.

In Fig. 21, all the \mathcal{C} and \mathcal{S} objects that are misclassified by \mathbf{f}^* are shown. Two out of the three misclassified cylinders are classified as \mathcal{S} , and four out of the six misclassified spheres are classified as \mathcal{C} , being still detected as mines. Some of the most challenging \mathcal{C} and \mathcal{S} objects that are correctly classified are shown in Fig. 22. Finally, Fig. 23 shows some of the \mathcal{N} objects that produce false alarms. The first row contains three \mathcal{N} objects classified as \mathcal{S} , and the rest are classified as \mathcal{C} .

V. CONCLUSION

In this paper, two segmentation algorithms have been described. The first one is based on a Markovian model with an optimized initialization scheme. The second is the AC algorithm, whose initialization is based on the Markovian segmentation result. Moreover, we propose a novel modification of the standard AC algorithm that avoids the local minima and significantly improves the performance in more than 35% of the cases. Several kinds of descriptors for underwater mine classification have been investigated and tested with a statistical classifier. An optimal feature vector has been proposed, obtaining a rate of 95% correct classification with a false alarm rate of 7% for a database of real SAS images with more than 1000 objects. This optimal feature vector consists of several geometrical shape descriptors for the shadow and the highlight shapes, two normalized central moments as well as a principal component of the shadow shape, three statistical descriptors based on a Weibull parametric model of the image, and a coefficient that accounts for the similarity between the segmentation results produced by the Markovian and the AC segmentation algorithms. The shadow and highlight shape descriptors are well-known in the literature, but this is the first time that they are tested with SAS images. The statistical descriptors and the fusion of the two segmentation algorithms are novel to this work. Furthermore, objects with and without highlight are compared within the same framework, avoiding the effective reduction of the training set for those objects with highlight.

Some future work on the characterization of the different features is foreseen. The feature robustness in the presence of noise, navigation errors or complicated seabeds such as sand ripples, has to be addressed. A relaxation of the Gaussianity assumption for the feature descriptors might improve the classification performance. Moreover, other approaches to classification such as support vector machines [58], [59] or fuzzy logic [60] will also be considered. Support vector machines do not assume any specific distribution of the features and therefore are more robust to situations where the Gaussianity assumption does not hold. Fuzzy logic techniques can be used to assess the confidence of the classifier, which is extremely useful in real applications.

ACKNOWLEDGMENT

The authors would like to thank Dr.-Ing. R. Engel, Dipl.-Ing. U. Hölscher, Dr. A. Kraft, and Dr. rer. nat K. Siantidis from Atlas Elektronik GmbH for making available the data and for their support and valuable comments.

REFERENCES

- [1] M. P. Hayes, "Synthetic aperture sonar: A review of current status," *IEEE J. Ocean. Eng.*, vol. 34, no. 3, pp. 207–223, Jul. 2009.
- [2] D. Massonnet and J. C. Souyris, *Imaging with Synthetic Aperture Radar*. Boca Raton, FL: CRC, 2008.
- [3] S. G. Johnson and A. Deaett, "The application of automated recognition techniques to side-scan sonar imagery," *J. Ocean. Eng.*, vol. 19, no. 1, pp. 138–144, Jan. 1994.
- [4] C. Debes, M. Amin, and A. Zoubir, "Target detection in single- and multiple-view through-the-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1349–1361, May 2009.
- [5] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Hybrid genetic optimization and statistical model based approach for the classification of shadow shapes in sonar imagery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 2, pp. 129–141, Feb. 2000.
- [6] R. Balasubramanian and M. Stevenson, "Pattern recognition for underwater mine detection," in *Proc. CAC/CAD Conf.*, 2001.
- [7] E. Dura, J. Bell, and D. Lane, "Superellipse fitting for the classification of mine-like shapes in side-scan sonar images," in *Proc. OCEANS Conf.*, 2002, vol. 1, pp. 23–28.
- [8] S. Reed, Y. Petillot, and J. Bell, "Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information," *IEE Proc. – Radar, Sonar, Nav.*, vol. 151, no. 1, pp. 48–56, 2004.
- [9] J. Groen, E. Coiras, and D. Williams, "Detection rate statistics in Synthetic Aperture Sonar images," in *Proc. Underwater Acoust. Meas. Conf.*, 2009.
- [10] E. Coiras and J. Groen, "3D target shape from SAS images based on a deformable mesh," in *Proc. Underwater Acoust. Meas. Conf.*, 2009.
- [11] H. Midelfart, J. Groen, and O. Midtgaard, "Template matching methods for object classification in synthetic aperture sonar images," in *Proc. Underwater Acoust. Meas. Conf.*, 2009.
- [12] M. Doherty, J. Landowski, P. Maynard, G. Uber, D. Fries, and F. Maltz, "Side scan sonar object classification algorithms," in *Proc. 6th Int. Symp. Unmanned Untethered Submersible Technol.*, 1989, pp. 417–424.
- [13] G. J. Dobeck, J. C. Hyland, and L. Smedley, "Automated detection and classification of sea mines in sonar imagery," in *Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) Conf. Series*, 1997, vol. 3079, pp. 90–110, 1.
- [14] S. Perry and L. Guan, "Pulse-length-tolerant features and detectors for sector-scan sonar imagery," *IEEE J. Ocean. Eng.*, vol. 29, pp. 138–156, 2004.
- [15] T. Aridgides, M. F. Fernandez, and G. J. Dobeck, *Side-Scan Sonar Imagery Fusion for Sea Mine Detection and Classification in Very Shallow Water*, A. C. Dubey, J. F. Harvey, J. T. Broach, and V. George, Eds. Bellingham, WA: SPIE, 2001, vol. 4394, pp. 1123–1134.
- [16] C. Ciany and W. Zurawski, "Performance of fusion algorithms for computer aided detection and classification of bottom mines in the shallow water environment," in *Proc. OCEANS Conf.*, 2002, vol. 4, pp. 2164–2167.
- [17] A. Castellano and B. Gray, "Autonomous interpretation of side scan sonar returns," in *Proc. Symp. Auton. Underwater Veh. Technol.*, 1990, pp. 248–253.
- [18] I. Quidu, J. Malkasse, G. Burel, and P. Vilbe, "Mine classification using a hybrid set of descriptors," in *Proc. OCEANS Conf.*, 2000, vol. 1, pp. 291–297.
- [19] J. C. Delvigne, "Shadow classification using neural networks," in *Proc. 4th Undersea Defence Conf.*, 1992, pp. 214–221.
- [20] D. Boulenguez and A. Quinquis, "Classification of underwater objects using Fourier descriptors," in *Proc. Int. Conf. Image Process. and its Applicat.*, 1999, pp. 240–244.
- [21] I. Tena Ruiz, D. Lane, and M. Chantler, "A comparison of inter-frame feature measures for robust object classification in sector scan sonar image sequences," *IEEE J. Ocean. Eng.*, vol. 24, no. 4, pp. 458–469, Oct. 1999.
- [22] F. Maussang, J. Chanussot, A. Hétet, and M. Amate, "Higher-order statistics for the detection of small objects in a noisy background application on sonar imaging," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, pp. 25–25, 2007.

- [23] F. Maussang, M. Rombaut, J. Chanussot, and M. Amate, "Fusion of local statistical parameters for buried underwater mine detection in sonar imaging," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–19, 2008.
- [24] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Boston, MA: Addison-Wesley Longman, 2001.
- [25] S. Guillaudoux, S. Daniel, and E. Maillard, "Optimization of a sonar image processing chain: A fuzzy rules based expert system approach," in *Proc. OCEANS Conf.*, 1996, vol. 3, pp. 1319–1323.
- [26] S. Daniel, S. Guillaudoux, and E. Maillard, "Adaptation of a partial shape recognition approach," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.: 'Comput. Cybern. Simulat.*, 1997, vol. 3, pp. 2157–2162.
- [27] J. Besag, "On the statistical analysis of dirty images," *J. R. Statist. Soc.*, vol. B-48, pp. 259–302, 1986.
- [28] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.
- [29] V. Murino, A. Trucco, and C. Regazzoni, "A probabilistic approach to the coupled reconstruction and restoration of underwater acoustic images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 1, pp. 9–22, Jan. 1998.
- [30] V. Murino and A. Trucco, "Edge/region-based segmentation and reconstruction of underwater acoustic images by Markov random fields," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, 1998, p. 408.
- [31] C. Collet, P. Thourel, P. Pérez, and P. Bouthemy, "Hierarchical MRF modeling for sonar picture segmentation," in *Proc. Int. Conf. Image Process.*, 1996, vol. 3, pp. 979–982.
- [32] P. Thourel, C. Collet, P. Bouthemy, and P. Pérez, "Multiresolution analysis and MRF modeling applied to the segmentation of shadows in sonar pictures," in *Proc. 2nd Asian Conf. Comput. Vis.*, 1996, vol. 2, pp. 81–85.
- [33] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical MRF model," *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1216–1231, Jul. 2000.
- [34] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 90–105, Jan. 2003.
- [35] J. Besag, "Spatial interaction and statistical analysis of Lattice systems," *J. R. Statist. Soc.*, vol. 36, pp. 192–236, 1974.
- [36] F. Spitzer, "Markov random fields and Gibbs ensembles," *Amer. Math. Monthly*, vol. 78, no. 2, pp. 142–154, 1971.
- [37] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 1, pp. 39–55, Jan. 1987.
- [38] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Three-class Markovian segmentation of high resolution sonar images," *Comput. Vis. Image Understanding*, vol. 76, no. 3, pp. 191–204, 1999.
- [39] F. Maussang, J. Chanussot, A. Hetet, and M. Amate, "Mean/standard deviation representation of sonar images for echo detection: Application to SAS images," *IEEE J. Ocean. Eng.*, vol. 32, pp. 956–970, 2007.
- [40] A. C. Cohen, "Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples," *Technometrics*, vol. 7, no. 4, pp. 579–588, 1965.
- [41] S. Banks, *Signal Processing, Image Processing and Pattern Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
- [42] R. Fandos and A. M. Zoubir, "Enhanced initialization scheme for a three-region markovian segmentation algorithm and its application to SAS images," in *Proc. Eur. Conf. Underwater Acoust.*, 2010.
- [43] F. Salzenstein and W. Pieczynski, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation," *Graphical Models Image Process.*, vol. 59, no. 4, pp. 205–220, 1997.
- [44] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "Use of active shape models for locating structures in medical images," *Image Vis. Comput.*, vol. 12, no. 6, pp. 355–365, 1994.
- [45] F. Langner, C. Knauer, W. Jans, and A. Ebert, "Side scan sonar image resolution and automatic object detection, classification and identification," in *Proc. OCEANS Conf.*, 2009, pp. 1–8.
- [46] Y. Amit, U. Grenander, and M. Piccioni, "Structural image restoration through deformable templates," *J. Amer. Statist. Assoc.*, vol. 86, no. 414, pp. 376–387, 1991.
- [47] C. Kervrann and F. Heitz, "A hierarchical statistical framework for the segmentation of deformable objects in image sequences," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, 1994, pp. 724–728.
- [48] C. Chesnaud, P. Refregier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1145–1157, Nov. 1999.
- [49] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [50] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ: Wiley, 2003.
- [51] Luciano and R. M. Cesar, *Shape Analysis and Classification: Theory and Practice (Image Processing Series)*. Boca Raton, FL: CRC Press, Taylor & Francis, 2000.
- [52] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recogn.*, vol. 37, no. 1, pp. 1–19, 2004.
- [53] M. Timothy, *Signal and Image Processing with Neural Networks*. New York: Wiley, 1994.
- [54] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philos. Mag.*, vol. 2, no. 6, pp. 559–572, 1901.
- [55] J. A. Fawcett, "Image-based classification of sidescan sonar detections," in *Proc. CAC/CAD Conf.*, 2001.
- [56] P. M. Mather and T. Brandt, *Classification Methods for Remotely Sensed Data*, Second ed. Boca Raton, FL: CRC Press, Taylor & Francis, 2009.
- [57] R. E. Bellman, *Adaptive Control Processes*. Princeton, NJ: Princeton Univ. Press, 1961.
- [58] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [59] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ: Springer-Verlag, 2006.
- [60] L. A. Zadeh, "Fuzzy sets," *Information Control*, vol. 8, pp. 338–353, 1965.



Raquel Fandos received the M.Eng. degree from the University of Zaragoza, Zaragoza, Spain, in 2004. She is currently working toward the Ph.D. degree at the Technische Universität of Darmstadt, Darmstadt, Germany.

While a student at Zaragoza, she visited KTH University, Stockholm, Sweden, for one year. From 2005 to 2008, she was with CERN, Geneva, Switzerland, as an RF Engineer for the CLIC project. Her main interests are detection and classification of underwater objects for sonar applications.



Abdelhak M. Zoubir received the Dr.-Ing. degree from Ruhr-Universität Bochum, Bochum, Germany.

He was with Queensland University of Technology, Brisbane, Australia, from 1992 to 1998, where he was an Associate Professor. In 1999, he joined Curtin University of Technology, Bentley, Australia, as a Professor of Telecommunications and was Interim Head of the School of Electrical and Computer Engineering from 2001 to 2003. In 2003, he moved to Technische Universität Darmstadt, Darmstadt, Germany, as Professor and Head of the

Signal Processing Group. His research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation, and sensor array processing applied to digital communication, sonar, radar, car engine monitoring, and biomedicine. He published over 300 papers in these areas.

Dr Zoubir is currently an IEEE Distinguished Lecturer (2010–2011). He was Technical Chair of the 11th IEEE Workshop on Statistical Signal Processing held in Singapore in 2001, General Co-Chair of the 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003) held in Darmstadt, Germany, and General Co-Chair of the IEEE Sensor Array and Multi-Channel Signal Processing Workshop (SAM 2008), held in Darmstadt in July 2008. He is the technical Co-Chair of ICASSP'14 to be held in Florence, Italy. He also served on Organizing and Technical Committees for numerous IEEE workshops and conferences. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2005 and has been a Member of the Senior Editorial Board of the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING since 2009. He currently serves as an Associate Editor for the European Association for Signal Processing (EURASIP) journals *Signal Processing* and the *Journal on Advances in Signal Processing*. He is the Chair of the IEEE SPS Technical Committee Signal Processing Theory and Methods (SPTM), (he was its Vice-Chair from 2008 to 2009 and a member from 2002 to 2007). He has been a Member of the IEEE SPS Technical Committee Sensor Array and Multi-Channel Signal Processing (SAM) since 2007. He was a Member of the IEEE SPS Technical Committee on Signal Processing Education (SpEd) from 2006 to 2008. He is an elected member of the EURASIP Board of Directors.