

# Enhanced Fuzzy-based Local Information Algorithm for Sonar Image Segmentation

Avi Abu, Roe Diamant, *Member, IEEE*,

July 18, 2019

**Abstract**—The recent boost in undersea operations has led to the development of high-resolution sonar systems mounted on autonomous vehicles. These vehicles are used to scan the seafloor in search of different objects such as sunken ships, archaeological sites, and submerged mines. An important part of the detection operation is the segmentation of sonar images, where the object's highlight and shadow are distinguished from the seabed background. In this work, we focus on the automatic segmentation of sonar images. We present our enhanced fuzzy-based with Kernel metric (EnFK) algorithm for the segmentation of sonar images which, in an attempt to improve segmentation accuracy, introduces two new fuzzy terms of local spatial and statistical information. Our algorithm includes a preliminary de-noising algorithm which, together with the original image, feeds into the segmentation procedure to avoid trapping to local minima and to improve convergence. The result is a segmentation procedure that specifically suits the intensity inhomogeneity and the complex seabed texture of sonar images. We tested our approach using simulated images, real sonar images, and sonar images that we created in two different sea experiments, using multibeam sonar and synthetic aperture sonar. The results show accurate segmentation performance that is far beyond the state-of-the-art results.

**Index Terms**—Fuzzy clustering, sonar image segmentation, speckle noise, kernel-induced distance, image de-noising, intensity inhomogeneity.

## I. INTRODUCTION

HIGH-resolution imagery of the seabed is mostly provided by sonar systems for the purpose of object detection. The relatively clean images produced by these technologies have increased the feasibility of automatic detection and classification (ADAC) of underwater objects [1]. ADAC (see Fig. 1) is required for marine applications such as seabed archeology [2], pipeline monitoring [3] and offshore oil prospecting [4]. The process is applied onboard an autonomous underwater vehicle that independently surveys a designated area. In all of these applications, the key to successful object detection and classification is to separate the seabed background from the object's highlight and shadow regions. This separation process is referred to as *image segmentation*. The focus of this work is to develop a segmentation algorithm that combats the main

challenge of intensity inhomogeneity, which poses difficulties in regard to sonar images.

Sonar images consist of a large number of reflectors such as sand and boulders, which may be randomly distributed across the seabed [5]. In relation to object identification, these elements are regarded as background. An example of different seabed backgrounds observed in a sonar image is presented in Fig. 2. Regarding segmentation, the object's highlight is referred to as the direct acoustic reflection from the object; the shadow region is the area behind the highlight (w.r.t to the sonar's location), for which the acoustic signals from the sonar are blocked. The background is the area in the sonar image that includes acoustic backscattering from the seafloor, but not from the object. Once these three regions have been identified, the geometrical properties of the shadow and highlight can be evaluated to carry out object classification [6].

Image segmentation methods include the use of mixture [7], graph cut [8], and active contour [9]. In this work, we consider the application of sonar segmentation using the fuzzy theory. Fuzzy algorithms have been widely applied to the segmentation of optical images, and their simplicity and low complexity offer advantages to sonar segmentation onboard autonomous vehicles, where real-time analysis and low processing needs are of interest. Methods like fuzzy c-means (FCM) [10], fast generalized FCM algorithm (FGFCM) [11] and enhanced FCM (EnFCM) [12] achieve good results on natural images. However, due to the strong intensity inhomogeneity in sonar images, the results of these methods are seriously degraded [13]. Moreover, these methods are very sensitive to initialization inaccuracies and tend towards convergence to local minima. Another challenge of current fuzzy methods is the need to set the fuzzy parameters via a trial-and-error process, e.g., according to the trade-off between the original image and the filtered one, thereby limiting the robustness of the schemes to different sea conditions.

To combat the aforementioned challenges, we propose a new sonar image segmentation algorithm to be used within the ADAC scheme. Considering the need to adapt to complex seabed structures, our solution is based on fuzzy segmentation with a non-Euclidean kernel. Our goal is to achieve stable performance in different environmental conditions and for different objects' shapes. To reduce false segmentation and to improve complexity, our method combines segmentation with a new de-noising algorithm that identifies the highlight,

A. Abu (email: aabu@campus.haifa.ac.il) and R. Diamant (email: roeed@univ.haifa.ac.il) are with the Department of Marine Technology, University of Haifa, 199 Aba Khoushy Ave., Haifa, Israel. Parts of this work have been presented in the Oceans Conference, Kobe, Japan, 2018. This work was sponsored in part by the NATO Science for Peace and Security Program under grant G5293.

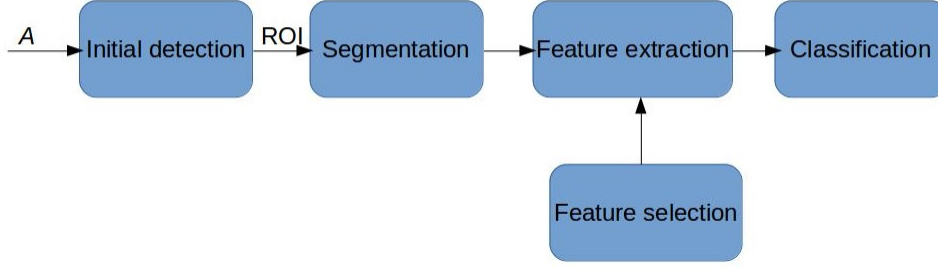


Fig. 1. Automatic detection and classification (ADAC) scheme.

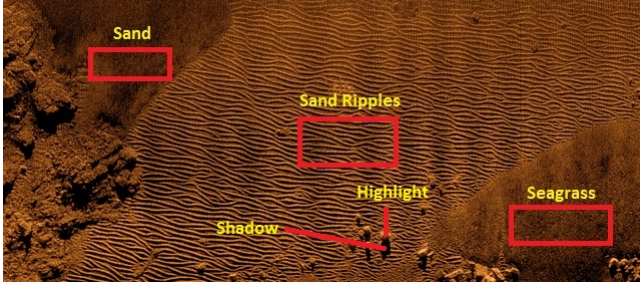


Fig. 2. Complex seabed texture in a side-scan sonar image. An example of shadow and highlight regions are represented by red lines. Examples of background are represented by rectangles.

shadow and background pixels using split-window architecture, and employs an automatic mechanism to evaluate the de-noising performance. To better deal with intensity inhomogeneity in sonar images, our de-noising solution includes a new Bayesian-based filter. To attain accurate sonar image segmentation and fast convergence, as part of the objective function of our fuzzy optimization problem, we add two new fuzzy terms, which we refer to as the *local second moment* and the *between-cluster*. Finally, we note that both our de-noising and segmentation solutions are parameter-free, thereby improving the overall robustness of the segmentation results.

This paper focuses on the segmentation part of the computer-aided detection and classification (ADAC) scheme. Being implemented after the detection algorithm and before feature extraction, (see Fig. 1) segmentation is a key enabling technology towards reducing false positives in the detection scheme while accurately identifying the location of the object within the sonar image. The design of segmentation algorithms has, therefore, drawn attention in the sonar processing community. However, intensity inhomogeneity and complex seabed textures like sand ripples in sonar imagery, cause severe degradation in the performance of existing approaches. This paper is motivated by facing these knowledge gaps to produce a robust and efficient segmentation algorithm.

Two main contributions are identified in this work. The first is image de-noising, as a pre-processing step to image segmentation, and the latter is fuzzy-based image segmentation. We use different statistical models for the shadow and background regions to better smooth the spiky noise in the

image, and improve segmentation accuracy by adding two terms into the objective function that utilize spatial information to reduce the effect of inhomogeneity. These improvements lead to a segmentation scheme that is robust to intensity inhomogeneity and different seabed structure and obtains the best segmentation results with less misclassified regions.

The key idea behind our method, referred to as the *enhanced fuzzy-based with kernel metric* (EnFK) is that, due to the intensity inhomogeneity, the noise in different areas within the sonar image should be treated separately. The flowchart of EnFK is illustrated in Fig. 5. We start with a sonar image de-noising as a pretreatment step. Both the original image and the de-noised one serve as the input to the segmentation procedure. This structure allows for increasing clusters' uniformity with clear boundaries between regions. To increase segmentation accuracy, the proposed de-noising scheme operates in parallel with the segmentation initialization scheme and performs its rough clustering. This operation is controlled through a mechanism to self-evaluate the success of the de-noising scheme. Finally, to reduce the effect of noise and increase the separation of the clustering results, we feedback the segmentation results to refine both the process of image de-noising and the fuzzy formalization. The main benefit of the proposed algorithm is its improved robustness to the non-homogeneous regions in the sonar images, as well as to the different seabed structures. This robustness comes out of the pre-setting of only a few system parameters. While this robustness comes at the cost of a more complex implementation structure, the complexity of EnFK is rather comparable with the benchmarks.

The main contributions of this paper are summarized as follows.

- 1) A new parameter-free fuzzy formalization for image segmentation. The problem is formalized in (14) and its solution is in (24).
- 2) A novel parameter-free de-noising approach that specifically combats intensity inhomogeneity. The algorithm's structure is illustrated in Fig. 5.

Experimental results on synthetic and real sonar images, obtained from different sonar systems such as side-scan, multi-beam and synthetic aperture sonar (SAS), show that the new algorithm is effective and efficient, as well as relatively independent of the sonar system and background type.

This paper is organized as follows. A detailed literature

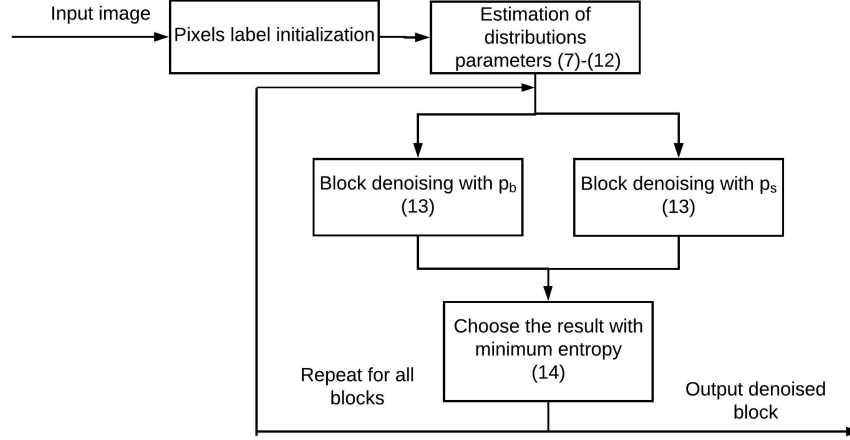


Fig. 3. Flowchart of the proposed de-noising scheme.

survey for fuzzy segmentation and image de-noising is presented in Section II. The system model and assumptions are outlined in Section III. The EnFK method for image de-noising and segmentation are detailed in Section IV, and Section V, respectively. Results and analysis of a database of simulated sonar images, real sonar images, and of sonar images collected during our sea experiments, are discussed in Section VI. Finally, conclusions are drawn in Section VII.

## II. LITERATURE SURVEY

In this section, we survey the state-of-the-art methods for the main components of our contribution, namely, fuzzy-based image segmentation, and image de-noising.

### A. Fuzzy-based Segmentation

Fuzzy algorithms have been widely used for image segmentation. The Fuzzy C-means (FCM) algorithm [10] is a popular method due to its simplicity and fast convergence. Unlike hard-clustering methods, where each pixel is assigned a single label, the FCM allows pixels to belong to multiple labels with different membership degrees, and measures the similarities between each pixel in the image and the center of the clusters. In the geometrically-guided FCM (GG-FCM), proposed by Noordam *et al.* [14], geometrical information is used during the segmentation process. Szilagyi *et al.* [15] proposed the enhanced FCM (EnFCM), which generated a linear weighted image from the original image and the mean filter image. Then, to reduce the computational time, the segmentation is performed on the histogram instead of the pixels. Cai *et al.* [11] proposed the fast generalized FCM (FGFCM). This algorithm generates a nonlinear-weighted image from the original image, local spatial texture, and the gray level neighborhood. To exploit spatial information between the image's pixels, Ahmed *et al.* [16] proposed the FCM<sub>S</sub>, which adds spatial information, at the cost of high complexity, thereby allowing the labeling of a pixel to be influenced by the labels of its surrounding pixels. Instead, Chen *et al.* [17] used mean and median filters in the segmentation process

to incorporate spatial dependency between the pixels. The trade-off between the effectiveness of preserving an object's borders and robustness to noise is controlled by a parameter  $\alpha$ , determined by trial-and-error.

For greater robustness, Stelios *et al.* [18] proposed the fuzzy local information c-means (FLICM) method, which is free of any parameter selection. In this method, a novel fuzzy factor is introduced to replace the above parameter  $\alpha$ . This factor incorporates local spatial and local intensity to improve robustness to noise and outliers. This Euclidean metric may not fit images with high-intensity homogeneity like sonar images. A variant of the FLICM was proposed in [19], which replaced the spatial distance with a local coefficient of variation as a local similarity measure. More recently, Shang *et al.* [13] proposed the clone kernel spatial FCM (CKS-FCM), where the initialization of the cluster centers is set by mimicking the biological process of the acquired immune clone. This method has been found reliable for non-convex optimization and showed robustness to noisy images [20]. However, the complexity of the CKS-FCM may be too high for the online processing of sonar images, since the initial cluster centers using the immune clone process are performed in addition to the fuzzy clustering. In addition, the objective function of [13] also contains free parameters that are fine-tuned in advance, thereby effecting robustness.

### B. Image De-noising

For accurate segmentation, removing noise components without distorting the object's borders is essential. Wiener filtering and wavelet transform [21] are typical approaches for image de-noising. Yet, these approaches are mostly applicable to cases of transient noises. In the *NL-means* filter introduced by Buades *et al.* [22], each pixel value is restored by the weighted average of all pixels' intensity in the image. The weight of a pixel is determined according to the similarity between the pixel's local neighborhood and that of the other images' pixels. A similar approach is performed in 3-D filtering (BM3D) [23], where the non-local filtering is combined

with Wiener filtering. However, our results showed that these filters are appropriate for use in the case of additive white Gaussian noise, which is less compatible with the embedded noise in sonar images. Coupe *et al.* [24] proposed the NLMSF method, which can preserve the objects' borders in ultrasound images.

Our literature survey shows that image segmentation and image de-noising are well-investigated subjects. Yet, for sonar image segmentation, we identify some remaining gaps. These include the sensitivity of the existing segmentation methods to parameter selection, as well as their sensitivity to intensity inhomogeneity and different seabed textures. For image de-noising, we argue that a proper systems model - which can accurately reflect the statistical behavior of the noise in the shadow, background, and highlight regions in sonar images - is still necessary.

### III. SYSTEMS MODEL AND MAIN ASSUMPTIONS

Let  $\mathcal{Y}$  be a two-dimensional sonar image with dataset  $\{y_1, \dots, y_N\} \subseteq \mathcal{Y}$ , where  $y_i$  denotes the intensity of pixel  $i$ . Each pixel  $i$  has one of three possible labels  $l_i \in \{S, H, B\}$ , where  $S$  and  $H$  are the shadow and the highlight of objects found in the image, respectively, and  $B$  is the background. Our aim is to accurately identify the image's shadow and highlight regions. The noise in the sonar image is modeled as an additive component, and the statistical distribution of the noise is region-dependent. We justify this because the noise in the shadow region is the electronic noise from the receiver. Further, as performed in [25], [26], since de-noising is extracted on local blocks, we follow the additive noise model for local statistics. The noisy image is

$$y_i = x_i + \xi_i, \quad (1)$$

where  $\xi_i$  is a conditionally independent additive noise with a probability density function (PDF)  $p_\xi(\xi)$  and  $x_i$  is the true image pixel. In the EnFK method, we model  $p_\xi(\xi)$  according to the pixel's label. Recall that a shadow region is created when the object is blocking the acoustic reverberation. The signal related to the shadow region consists of the electronic noise from the receiver. Thus, in this region, the noise is modeled by a zero-mean Gaussian distribution [26]. Following [27], the noise in the background and the highlight regions is modeled by the exponential distribution. Observing different sonar images, we found that this choice of statistics offers a better distinction between pixels related to the background vs. pixels related to the object, as opposed to, e.g., Gaussian distribution of different parameters per class [28], or the Weibull distribution [6]. Moreover, the results also demonstrate that this distribution model is sufficiently valid to provide accurate de-noising results. We model the PDF of  $\xi_i$  by

$$p_\xi(\xi) = \begin{cases} p_b(\xi) = \lambda \cdot \exp(-\lambda\xi), & \text{for } B \text{ or } H \text{ regions} \\ p_s(\xi) = \frac{1}{\sqrt{2\pi}(\sigma)^2} \cdot \exp\left(-\frac{\xi^2}{(\sigma)^2}\right), & \text{for } S \text{ region} \end{cases}, \quad (2)$$

where  $\lambda$ , is the exponential distribution parameter, and  $\sigma$  is the standard deviation for the shadow regions. Since the noise  $\xi$  is induced by the process of constructing the sonar image, we assume that  $\lambda$  and  $\sigma$  from (2) are constant per image.

### IV. IMAGE DE-NOISING

The image de-noising process is performed to reduce the intensity inhomogeneity of the sonar image. This objective is of importance for the task of object identification, where not only the detection of the object is of concern, but also maintaining its observed shape to ease the classification procedure. The key idea in the proposed de-noising scheme is to use the Bayesian approach in [24] to tie different statistical models to the sonar images' different regions, namely, shadow, highlight, and background. As will be shown in the Results section, this approach leads to improved results in terms of assessment index ( $Q$ ) [24], which reflects the region homogeneity level. Moreover, a novel method to self-evaluate the de-noising accuracy is introduced, thus avoiding the need to initialize the de-noising process.

#### A. The De-noising Filter

Our de-noising is based on the NLMSF in [24]. For the sake of completeness, we briefly describe the main idea of the NLMSF. The NLMSF is a despeckling method that utilizes a dedicated speckle model to handle the spatial speckle patterns in the image. The blockwise Bayesian estimator  $\hat{\mathbf{x}}(B_i)$  is defined as [29]

$$\hat{\mathbf{x}}(B_i) = \frac{\sum_{k=1}^{|\Delta_i|} \mathbf{y}(B_k) p(\mathbf{y}(B_i) | \mathbf{y}(B_k))}{\sum_{k=1}^{|\Delta_i|} p(\mathbf{y}(B_i) | \mathbf{y}(B_k))}. \quad (3)$$

where  $B_i$  is a square block of size  $T$  equals  $(2\alpha + 1)^2$  ( $\alpha \in \mathbb{N}$ ) centered at pixel  $i$ ,  $\Delta_i$  is a square search block centered at pixel  $i$  of size  $|\Delta_i| = (2M + 1)^2$  ( $M \in \mathbb{N}$ ),  $\mathbf{y}(B_i)$  is a  $T \times 1$  vector that contains all observed intensities of the pixels inside block  $B_i$ , and  $\mathbf{x}(B_i)$  is a  $T \times 1$  vector of the unobserved (unknown true image) intensities of the pixels inside block  $B_i$ . By the speckle model in [24], the statistical distribution of  $y_i | x_i$  is

$$p(y_i | x_i) \propto \mathcal{N}(x_i, (x_i)^{2\gamma} \sigma^2). \quad (4)$$

where,  $\gamma$  is the speckle parameter, and  $\sigma$  is the standard deviation of the gaussian noise in the speckle model. Assuming independence among the pixels, the likelihood of  $\mathbf{y}(B_i) | \mathbf{y}(B_k)$  can be factorized as:

$$p(\mathbf{y}(B_i) | \mathbf{y}(B_k)) \propto \exp\left(-\sum_{t=1}^T \frac{(y_i^{(t)} - y_k^{(t)})^2}{2(y_k^{(t)})^{2\gamma} \sigma^2}\right), \quad (5)$$

where  $y_i^{(t)}$  and  $y_k^{(t)}$  are the  $t$ th component in  $\mathbf{y}(B_i)$  and  $\mathbf{y}(B_k)$ , respectively. The restored intensity of pixel  $i$  is given by the mean of all restored values in the blocks  $B_i$  in which  $y_i$  is included. To improve the results and speed up the algorithm, a pixel selection scheme is used [30], which is controlled by the thresholding parameter  $\mu_1$ .

The free parameters  $\gamma, \mu_1$  affect the robustness of the NLMSF to seabed intensity inhomogeneity. This is because of the need to tune these parameters for different background types. Considering this challenge, we add to the NLMSF scheme the capability to include the additional distribution types in (2) and to self-evaluate the distribution parameters. This is described in the following subsection.

### B. Estimation of Distribution Parameters

We make the realistic assumption that the sonar image is intensity inhomogeneous. Under such conditions, we expect  $x_i$  from (1), i.e., the noise-free image components, to be different at various locations of the image. To compensate for the location-dependent  $x_i$ , the image is divided into non-overlapped blocks  $Y_r$  of size  $\kappa_s$  and perform parameter estimation per block. Then, modeling the distribution of the noise components to be the same for the whole image, the estimation from all blocks is fused into a single one. We note the choice of  $\kappa_s$  tradeoffs. Small values of  $\kappa_s$  may enlarge the estimation error of the parameters of the distributions, while large values of  $\kappa_s$  degrade the performance of the despeckling filter in the shadow zone because large blocks contain not only shadow pixels, but also background information. We leave the choice of  $\kappa_s$  to the user based on the size of the object of interest.

As a model (2) reveals, the parameter estimation process must include labeling information. That is, each block must be pre-clustered into one of the possible labels  $\{S, B, H\}$ . Let  $c_r$  be the label of the  $r$ th block. The parameter  $c_r$  is determined based on the majority of the pixels' labels in the  $r$ th block. These pixels' labels are evaluated based on the initialization algorithm from [31]. While the initialization performed well for real sonar images, because it is a model-free algorithm, it may induce some clustering errors. Still, the impact of initialization errors is low because it is used only for estimating  $\lambda$  and  $\sigma$  as an average overall initialized windows. Thus, initialization segmentation errors in some windows would have a small effect.

Once  $c_r$  are determined, we statistically evaluate the parameters in (1). For blocks with  $c_r = \{B, H\}$ , we evaluate parameter  $\lambda$  in the  $r$ th block by

$$\hat{\lambda}_r = \left( \frac{1}{\kappa_s} \sum_{i \in Y_r} (y_i - \bar{y})^2 \right)^{-0.5}, \quad (6)$$

where

$$\bar{y} = \frac{1}{\kappa_s} \sum_{i \in Y_r} y_i. \quad (7)$$

Similarly, for block  $r$  with  $c_r = \{S\}$ , we set

$$\hat{\sigma}_r = \sqrt{\frac{1}{\kappa_s} \sum_{i \in Y_r} y_i^2}. \quad (8)$$

Then, following our assumption that the noise term parameters in (1) are constant throughout the sonar image, we follow the metric in [32], which we find to be the most suitable for dealing with outliers in sonar images, and fuse all per-block estimations as a weighted sum:

$$\hat{\lambda} = \sum_r \varrho_r \hat{\lambda}_r, \quad (9)$$

with

$$\varrho_r = \frac{\exp(-|\hat{\lambda}_r - \bar{\lambda}|)}{\sum_r \exp(-|\hat{\lambda}_r - \bar{\lambda}|)}, \quad (10)$$

and

$$\bar{\lambda} = \frac{1}{\rho_b} \sum_r \hat{\lambda}_r, \quad (11)$$

where  $\rho_b$  is the number of blocks labeled as highlight or background. The fusion of parameter  $\sigma$  is performed in the same fashion.

### C. Setting the De-noising Filter

1) *Formalization*: The flowchart of our de-noising scheme is shown in Fig. 3. The de-noising is carried out separately for each block  $Y_r$ . Since the blocks' NLMSF size,  $T$  and  $|\Delta_i|$  are much smaller than the original image size, for block de-noising, we assume that the labels of the pixels in  $Y_r$  are identical. Based on the model in (1), the enumerator of (3) can be rewritten as

$$p(\mathbf{y}(B_i) | \mathbf{y}(B_j)) = \prod_{t=1}^T p_{\xi}(y_i^{(t)} - y_j^{(t)}). \quad (12)$$

Based on the distribution model in (2)  $p_{\xi}(\xi)$  equals  $p_s$  or  $p_b$  according to the label of the block  $Y_r$ . That is, to calculate (12), we require information about the label of each block. Unlike the parameter estimation process, which tolerates errors in the block's initial clustering, in this case, de-noising is based on the erroneous identification of the block's label; therefore, usage of a wrong distribution model in (2) will likely lead to image distortion. Thus, unlike parameter estimation, regarding block de-noising, we avoid using segmentation initialization. Instead, for each block, we find the distribution that leads to the best de-noising result.

2) *Self-evaluation of De-noising Performance*: We measure successful de-noising using the concept of minimum entropy. This is because the block's entropy characterizes the distribution of the restored intensities. In particular, a very localized region will lead to small entropy, while a uniform region will lead to high entropy. Thus, setting the minimum entropy as a quality measure will lead to the choice of the best-localized values, which is the choice with the most homogeneous intensity of the de-noised block. The entropy is calculated by

$$H_{\mathbf{x}} = - \sum_i p(i) \log_2 p(i), \quad (13)$$

where  $p(i)$  is the number of pixels (after normalization) in the  $r$ th de-noised resulting block at the  $i$ th intensity bin.

With the uncertainty of the block label, each block is de-noised using (12) for both  $p_b(\xi)$  and with  $p_s(\xi)$ , to create the de-noised images  $\hat{\mathbf{x}}_b$  and  $\hat{\mathbf{x}}_s$ , respectively. Then, the entropy is calculated for each of the resulting images, and the chosen de-noising result is the one of minimum entropy.

## V. ENHANCED FUZZY-BASED IMAGE SEGMENTATION

The uniqueness of the proposed sonar image segmentation method lies in the introduction of two new terms: the *local second moment* term, and the *between-cluster* term. As shown in the Results section, using these two terms lead to better segmentation accuracy and faster convergence. Inspired by the echo detection process in [26], we use the second moment of the image data in the kernel space to better separate the object's highlight from its background. The *between-cluster* term represents the error between the prototypes and the empirical cluster centers and is incorporated, in the kernel space, into the

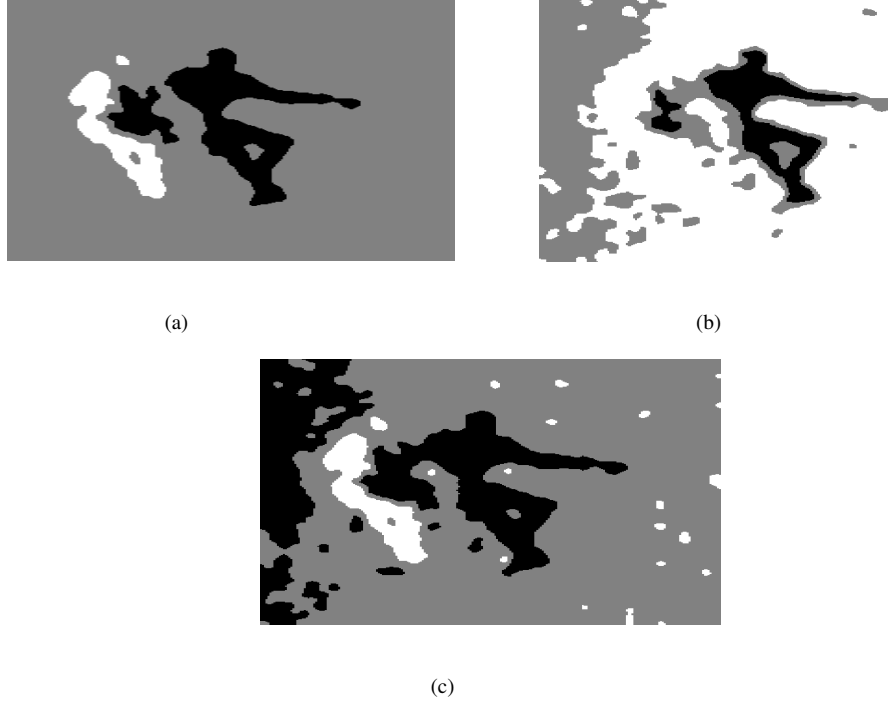


Fig. 4. Final segmentation results of EnFK for the *Drowning Victim* sonar image with different kernel functions. (a) Gaussian radial basis kernel. (b) linear kernel. (c) Polynomial kernel (of order 2).

objective function to reduce false segmentation. Furthermore, as quality metrics, we use the kernel distance [33] rather than the commonly used Euclidean distance between pixels. This is because, in the kernel space, the formed clusters are more spherical and can, therefore, be more easily clustered [34].

The kernel maps a data set  $\mathcal{U}$  into a higher dimensional space  $\mathcal{W}$  (kernel space) via the transform function  $\Phi : \mathcal{U} \rightarrow \mathcal{W}$ . The fuzzy algorithm generates the degrees of membership  $u_{ki}$ , i.e., the probability that the label of pixel  $i$  belongs to the  $k$ -th cluster, by minimizing the objective function

$$(U, V) = \arg \min_{u_{ki}, V_k} J_m(U, V) \quad (14)$$

$$\text{s.t. } \sum_{k=1}^c u_{ki} = 1, \forall i,$$

where,  $c$  is the number of clusters. For the three possible classes  $\{B, H, S\}$ , our objective function is defined as

$$\begin{aligned} J_m(U, V) = & \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m \| \Phi(\hat{x}_i) - \Phi(V_k) \|^2 \\ & + \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m \| \Phi(\bar{x}_i) - \Phi(V_k) \|^2 \\ & + \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m \| \Phi(\tilde{x}_k) - \Phi(V_k) \|^2 + G_{ki} \end{aligned} \quad (15)$$

$G_{ki}$  is referred to as the *fuzzy factor* [18] defined by

$$G_{ki} = \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m \sum_{j \in \mathcal{N}_i, i \neq j} \frac{1}{1 + d_{ij}} (1 - u_{kj})^m (1 - K(y_j, V_k)), \quad (16)$$

where  $d_{ij}$  is the Euclidean distance between pixel  $i$  and pixel  $j$ ,  $m$  is the weighting exponent on each fuzzy membership, which determines the fuzziness of the results,  $\hat{x}_i$  is the denoised image,  $\Phi(\hat{x}_i) - \Phi(V_k)$  is the *local second moment* term in the kernel space and  $\bar{x}_i$  is given by:

$$\bar{x}_i = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} y_j^2, \quad (17)$$

where  $\mathcal{N}_i$  is the local window of size  $(2\beta + 1)^2$  centered at pixel  $i$ ,  $\{V_k\}_{k=1}^c$  are the center of the clusters,  $\Phi(\bar{x}_k) - \Phi(V_k)$  is the *between-cluster* term in the kernel space, and  $\tilde{x}_k$  is the average of all pixels assigned to the  $k$ th cluster, such that

$$\tilde{x}_k = \frac{1}{n_k} \sum_{\{j | l_j = k\}} y_j, \quad (18)$$

where  $n_k$  is the number of pixels with an assigned label equals the  $k$ th cluster. A kernel in the data space can be represented as

$$K(x, y) = \langle \Phi(x), \Phi(y) \rangle, \quad (19)$$

where  $\langle \cdot \rangle$  is the inner product operation. Due to the fact that  $\Phi$  in (15) is unknown, the segmentation problem should be solved using only the kernel function. Since  $k(x, x) = 1$ , the inner product in the kernel space, can be written as

$$\begin{aligned} \| \Phi(x) - \Phi(y) \|^2 = & \langle \Phi(x), \Phi(x) \rangle + \langle \Phi(y), \Phi(y) \rangle - 2 \langle \Phi(x), \Phi(y) \rangle \\ & - K(x, x) + K(y, y) - 2K(x, y) \\ = & 2(1 - K(x, y)). \end{aligned} \quad (20)$$

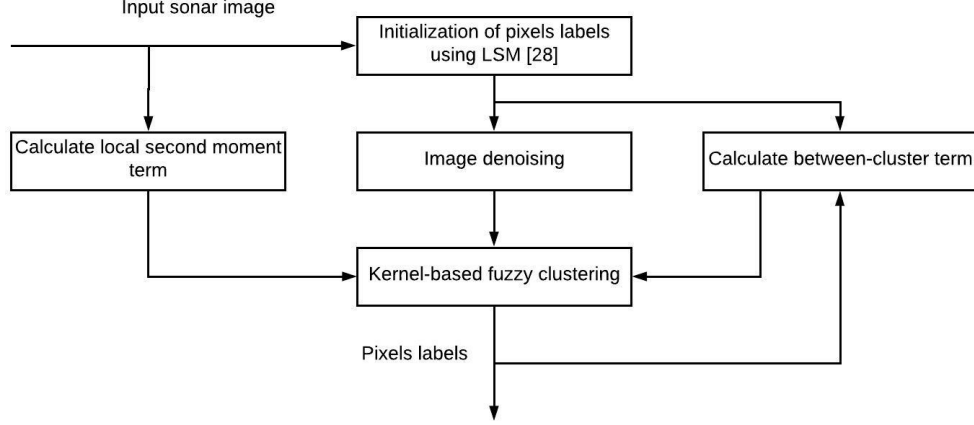


Fig. 5. EnFK segmentation method scheme.

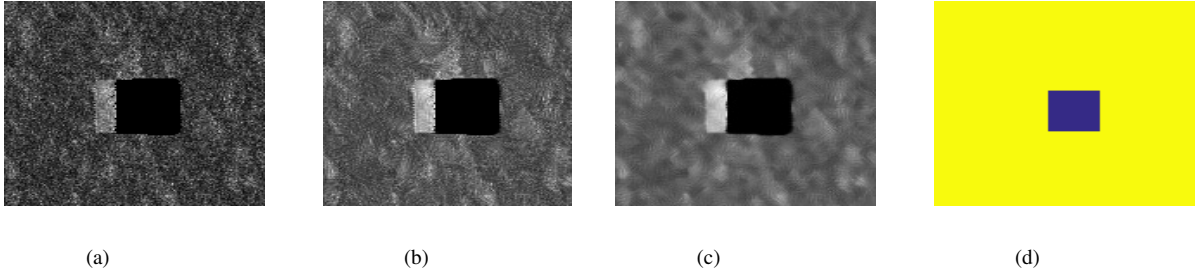


Fig. 6. Comparison of the de-noised images for the synthetic sonar image (sea-grass background) using the proposed filter and the NLMSF filter: (a) Input sonar image. (b) De-noising results of the NLMSF filter. (c) De-noising results of the proposed filter. (d) Decision map of the distributions: blue - Gaussian (for shadow regions), yellow - Exponential (for highlight/background regions). We observe that our approach yields a much better de-noising than the benchmark with an accurate identification of the blocks containing the object's shadow.

Using (20), (15) is rewritten as

$$\begin{aligned}
 J_m(U, V) = & \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m (1 - K(\hat{x}_i, V_k)) \\
 & + \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m (1 - K(\bar{x}_i, V_k)) \\
 & + \sum_{i=1}^N \sum_{k=1}^c u_{ki}^m (1 - K(\tilde{x}_k, V_k)) + G_{ki}. \quad (21)
 \end{aligned}$$

The solution to (14) is obtained by choosing the values  $u_{ki}$ , which minimize (21). Using a Lagrange multiplier, the solution comes readily as

$$\begin{aligned}
 u_{ki} = & \frac{\left( (1 - K(\hat{x}_i, V_k)) + (1 - K(\bar{x}_i, V_k)) + (1 - K(\tilde{x}_k, V_k)) + G_{ki} \right)^{-\frac{1}{m-1}}}{\sum_{j=1}^c \left( (1 - K(\hat{x}_i, V_j)) + (1 - K(\bar{x}_i, V_j)) + (1 - K(\tilde{x}_j, V_j)) + G_{ji} \right)^{-\frac{1}{m-1}}} \quad (22)
 \end{aligned}$$

$$V_k = \frac{\sum_{i=1}^N u_{ki}^m \left( \hat{x}_i K(\hat{x}_i, V_k) + \bar{x}_i K(\bar{x}_i, V_k) + \tilde{x}_k K(\tilde{x}_k, V_k) \right)}{\sum_{i=1}^N u_{ki}^m \left( K(\hat{x}_i, V_k) + K(\bar{x}_i, V_k) + K(\tilde{x}_k, V_k) \right)}, \quad (23)$$

Finally, pixel  $i$  is assigned to the cluster with the highest membership

$$\hat{l}_i = \arg \max_k \{u_{ki}\}. \quad (24)$$

We adopt the Gaussian radial basis kernel function (GRBF) [35] with the kernel function  $K(x, y) = \exp(-\|x - y\|^2 / r^2)$ , where  $r$  is the kernel's bandwidth, and  $r$  is set on the basis of the distance variance of all pixels [32]. We choose the Gaussian radial basis kernel because of its robust estimation controlled by the single bandwidth parameter  $r$ , and since it is able to catch non-linear connections between the observed data and the classed labels. To demonstrate this, in Fig. 4, we compare the segmentation results of linear, polynomial (of order 2) and Gaussian radial kernels applied on the EnFK with the *Drowning Victim* sonar image. The results show that the Gaussian radial kernel outperforms the linear and polynomial (of order 2) kernels in terms of the number of misclassified

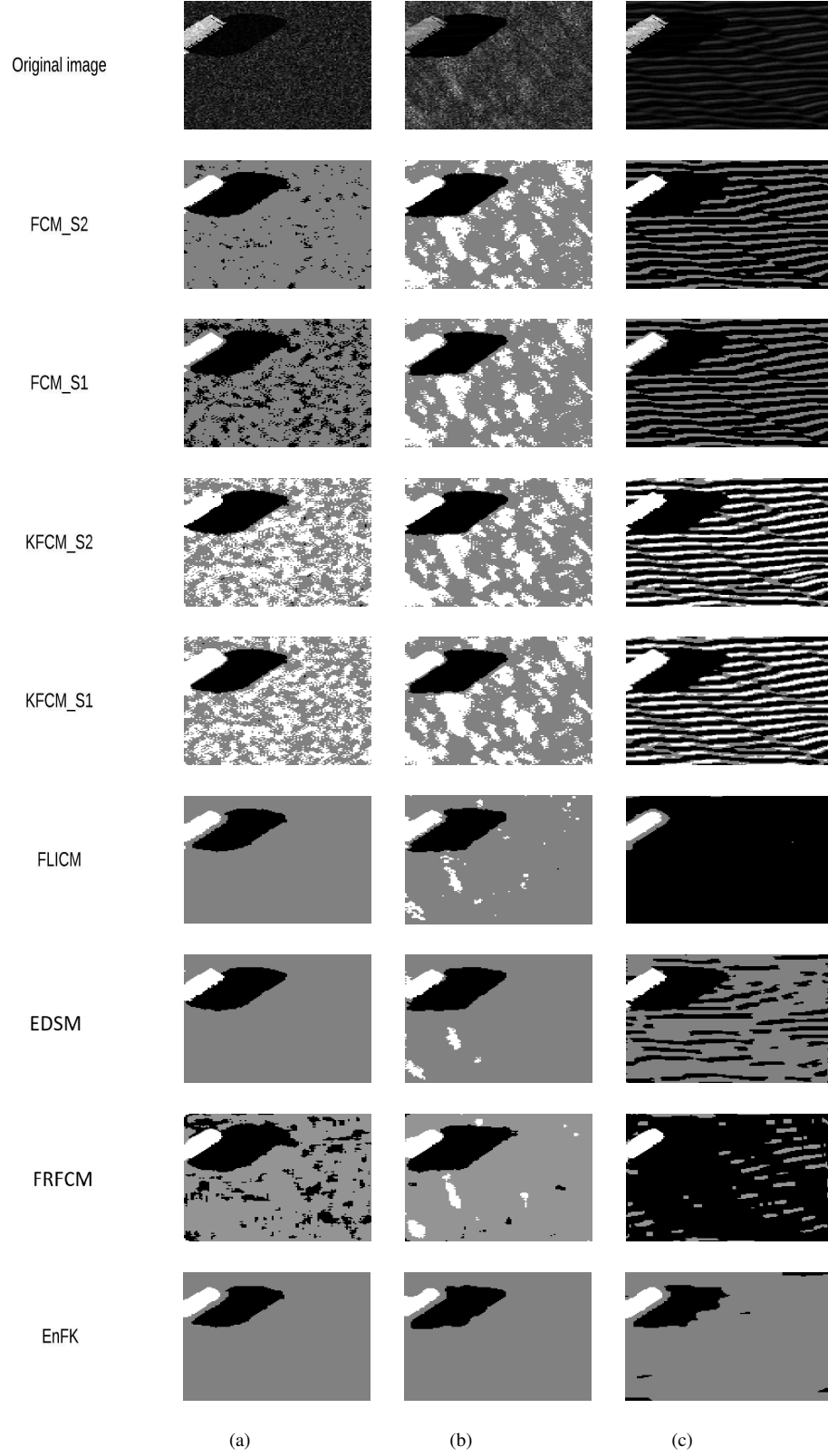


Fig. 7. Segmentation results for synthetic sonar images. The colors in the segmentation maps represent: black for the shadow region, white for the highlight region and gray for the background region. Segmentation results of the benchmarks FCM\_S2, FCM\_S1, KFCM\_S2, KFCM\_S1, FLICM, EDSM, FRFCM and EnFK (from top to bottom, respectively). (a) to (c) for sand, sea-grass and sand-ripples background, respectively. The EnFK obtains the best region uniformity.



---

**Algorithm 1** EnFK Fuzzy Segmentation Algorithm
 

---

**Input:** The original image  $\mathcal{Y}$ ;

**Output:** The pixels labels  $\hat{l}_i$ ;

**Begin**

- 1: Calculate the initial pixels' labels using the LSM algorithm;
- 2: Get the initialized prototypes  $V_i^{(0)}$  using (29);
- 3: Calculate the denoised image  $\{\hat{x}_i\}_{i=1}^N$  using (3), (12) and (13);
- 4: Calculate the local second moment  $\bar{x}_i$  using (17);
- 5: Set the weighting exponent  $m$  and the thresholds  $\epsilon$  and  $N_q$ ;
- 6: Set the loop counter  $q = 1$ ;
- 7: **while**  $q < N_q$  **do**
- 8:   Calculate  $\bar{x}_k$  using (18) with the estimated labels from previous iterations;
- 9:   Update the membership matrix  $u_{ki}$  using (22);
- 10:   Update the prototypes  $V_k$  using (23);
- 11:   Get the pixel labels using (24);
- 12:   **if**  $\max_{k,i} |u_{ki}^{(q)} - u_{ki}^{(q-1)}| < \epsilon$  **then**
- 13:     Stop and output the pixel labels using (24); otherwise, set the loop counter  $q = q + 1$  and go to (8);
- 14:   **end if**
- 15: **end while**
- 16: Get the pixel labels using (24);

**End**


---

regions and better-defined region boundaries. The bandwidth  $r$  is given by

$$r = \left( \frac{1}{N-1} \sum_{i=1}^N (d_i - \bar{d})^2 \right)^{0.5}. \quad (25)$$

The data center  $\bar{u}$  is given by:

$$\bar{u} = \frac{1}{N} \sum_{i=1}^N y_i. \quad (26)$$

The distance from pixel  $i$  to the data center is  $d_i = \|y_i - \bar{u}\|$ . The mean distance  $\bar{d}$  is calculated by

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N d_i. \quad (27)$$

We solve (14) iteratively. In each iteration,  $\bar{x}_k$  is set by (18) according to the segmentation solution from the previous iteration. Then, the prototype  $V_k$  is calculated by (23) and plugged into (22) to yield a new estimation (24). The process stops upon convergence, i.e., when

$$\max_{k,i} |u_{ki}^{(q)} - u_{ki}^{(q-1)}| < \epsilon \quad \text{or} \quad q > N_q, \quad (28)$$

where  $u_{ki}^{(q)}$  is the membership at  $q$ th iteration,  $\epsilon$  is a convergence parameter, and  $N_q$  is the maximum number of allowed iterations. A numerical proof of the convergence of the above procedure appears below. The labels of the pixels in  $\mathcal{Y}$  are initialized using our LSM algorithm [31] and denoted by  $\{l_i^{(0)}\}_{i=1}^N$ . The prototypes  $\{V_k\}_{k=1}^c$  are initialized with

$$V_1^{(0)} = \frac{1}{\rho_s} \sum_{\{j | l_j^{(0)} = S\}} y_j, \quad (29)$$

with  $\rho_s$  equals the number of pixels with initialized label equals  $\{S\}$  and analogously for  $V_2^{(0)}$  and  $V_3^{(0)}$ , which are related to the highlight and the background, respectively.

The proposed segmentation process is illustrated in Fig. 5, and its operation is summarized in Algorithm 1. The EnFK method comprises six main steps: 1) Clusters initialization using the robust LSM initialization method (line 1); 2) Image de-noising using the improved NLMSF-based algorithm (lines 2-3); 3) Calculating the local second moment  $\{\bar{x}_i\}_{i=1}^N$  (line 4); 4) Updating the membership matrix (lines 5-9); 5) Updating the cluster prototypes (line 10); and 6) Stopping criteria (lines 11-16).

## VI. EXPERIMENTAL RESULTS AND COMPARISON

In this section, the performance of the proposed image de-noising algorithm is presented, as well as the overall performance of the segmentation process. For de-noising, we use  $\kappa_s = 255$ , and compare the results of our algorithm with those of the robust NLMSF filter in [24] with  $\gamma = 0.5$ ,  $\alpha = 1$ , and  $M = 3$ , which we found to be the most suitable for sonar image de-noising. In this work, we focus on fuzzy sonar image segmentation. Therefore, the proposed method is compared with the state-of-the-art in fuzzy segmentation. In particular, we choose as benchmarks methods FCM\_S1 [17], FCM\_S2 [17], KFCM\_S1 [36], KFCM\_S2 [36], FLICM [18], and fast and robust fuzzy C-Means (FRFCM) [37]. For completeness, we also add to the benchmark a non-fuzzy segmentation method based on the Dempster-Shafer evidence theory (EDSM) [38]. We choose these methods as benchmarks since they are both a key in image segmentation and are heavily cited, and because of their suitability for sonar image processing. Further, these methods represent different approaches in fuzzy image segmentation. Specifically,

- FCM\_S1: Utilizing the mean filter in the objective function to increase robustness to noise.
- FCM\_S2: Utilizing the median filter in the objective function to compensate the intensity inhomogeneity.
- KFCM\_S1: Extension of FCM\_S1 that maps the original input into the kernel space to increase separability of the data.
- KFCM\_S2: Extension of FCM\_S2 to the kernel space.
- FLICM: Control the influence of the local neighborhood pixels on the labeling without any parameter selection.
- FRFCM: The local spatial information is incorporated into the fuzzy objective function by morphological reconstruction.

For all fuzzy-based segmentation methods, the fuzziness index is set to  $m = 2$ , and  $\beta = 2$ . The maximum number of iterations  $N_q$  is set to 250 and the threshold  $\epsilon = 0.01$ . According to [36], the parameter  $a$  is set to 5 for FCM\_S1, FCM\_S2, KFCM\_S1 and KFCM\_S2. According to [37], the morphological reconstruction parameter is set to 3, and the size of the filtering window is set to 7. For EDSM, we set  $\gamma_1 = 0.1$  and  $\gamma_2 = 1.4$  according to [38].

### A. Sonar Data

Our sonar data comprises two data sets. The first image set consists of three synthetic sonar images, presented in Fig. 7(a)-(c) (top images). The images contain a cylindrical object with different backgrounds: sand, sea-grass, and sand ripples. The

size of the synthetic images is 120×120. Sand and sea-grass seabed textures are generated according to the models in [27], while the sand ripples texture is generated according to [39]. The object's region and background are synthesized separately. The object's intensity level is modeled by a gamma distribution with mean values and standard deviations of 120, 10, for the shadow region, and 1, 0.1 for the highlight region, respectively. Similar to the model in [40], the object and background are superimposed, as follows:

$$A_i = \begin{cases} 0.8\chi_{i,h} + 0.2\chi_{i,b}, & \text{if } i \in \text{highlight region} \\ 0.8\chi_{i,s} + 0.03\chi_{i,b}, & \text{if } i \in \text{shadow region} \\ \chi_{i,b}, & \text{if } i \in \text{background region} \end{cases} \quad (30)$$

where  $\chi_{i,h}$ ,  $\chi_{i,b}$  and  $\chi_{i,s}$  stand for the intensity level of the  $i$ th pixel in the highlight, background and shadow regions, respectively.

The second set consists of five real sonar images: the first image was imaged with CM2 towfish sonar and is of *Crab Traps* [41]; the second is *Airplane* and was made with the Sea Scan 600 sonar [42]; the third is *Drowning Victim* and was made with the EdgeTech 4125 sonar [43]. The sizes of these three images are 151×301, 182×232, and 182×232, respectively, and are given in Fig. 10(a)-(c) (top images). The last two images consist of two sonar images we sampled during our sea experiment. These images include two steel targets (Fig. 14(b)), and a submerged gas well (Fig. 14(e)), and are of size 200×120, and 120×300 pixels, respectively.

### B. Evaluation Indexes

To calculate the quantitative assessment of the de-noising and segmentation results, ground-truth maps are usually needed. These maps are generated by manual segmentation [4] based on the original sonar images. We use the variation information VI [44], partition coefficient  $v_{pc}$ , partition entropy  $v_{pe}$  [45], and MCR [46] for the segmentation results' evaluation, and despeckling assessment index  $Q$  [24] for the de-noising results' evaluation. The VI measures the dissimilarity between two maps in terms of information entropy. In our case, the first map is the segmentation results map  $S_r$  and the second is the ground-truth map  $G_t$ ,

$$VI(S_r, G_t) = H(S_r) + H(G_t) - 2I(S_r, G_t) \quad (31)$$

where,

$$H(S) = - \sum_{k=1}^3 \frac{N_k}{N} \ln \left( \frac{N_k}{N} \right), \quad (32a)$$

$$I(S, S') = \sum_{k=1}^3 \sum_{k'=1}^3 \frac{N_{k,k'}}{N} \ln \left( \frac{N_{k,k'}}{\frac{N_k}{N} \frac{N_{k'}}{N}} \right), \quad (32b)$$

with  $N_k$  is the number of assigned pixels with  $k$  label and  $N_{k,k'}$  is the number of points in the intersection maps  $S_r$  and  $G_t$  considering the labels  $k$  and  $k'$ , respectively. If both maps are identical, the entropies  $H(S_r)$  and  $H(G_t)$  are equal and  $VI(S_r, G_t) = 0$ .

TABLE I  
COMPARISON OF THE DESPECKLING ASSESSMENT INDEX  $Q$  FOR THE SAND, SAND-RIPPLES AND SEA-GRASS SONAR IMAGES WITH  $\alpha = 1$ ,  $M = 3$ ,  $\gamma = 0.5$  AND  $\kappa_s = 255$ .

| Method | Sand        | Sand ripple | Sea grass   |
|--------|-------------|-------------|-------------|
| EnFK   | 16.1        | <b>16.7</b> | <b>10.1</b> |
| NLMSF  | <b>16.3</b> | 16.3        | 9.51        |

Partition coefficient  $v_{pc}$  and partition entropy  $v_{pe}$  are defined as [45]

$$v_{pc} = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^3 u_{ki}^2, \quad (33a)$$

$$v_{pe} = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^3 u_{ki} \ln(u_{ki}). \quad (33b)$$

Best clustering is achieved when  $v_{pc}$  is close to one and  $v_{pe}$  approaches zero.

Despeckling assessment index  $Q$  is defined as [24]

$$Q = \frac{\sum_{j \neq k} (\mu_j - \mu_k)^2}{\sum_j \sigma_j^2}, \quad (34)$$

where,  $\mu_j$  and  $\sigma_j$  are the mean and the variance of the pixels' intensities with assigned  $j$ th label after de-noising. To calculate the despeckling assessment index, we use the ground-true map for pixels' label information. The higher the value of  $Q$ , the better de-noising results are achieved.

MCR is defined as the number of misclassified pixels normalized by the total number of pixels and is given by [46]

$$\text{MCR} = \frac{\sum_{i=1}^N I(S_r^i \neq G_t^i)}{N}, \quad (35)$$

where  $I$  is an indicator function equals one if its argument is true, and zero, otherwise. Segmentation improves, the lower the MCR is.

### C. Results on Synthetic Images

1) *De-noising Results:* In Fig. 6, the de-noising results obtained for a synthetic image of a cylindrical object in a sea-grass background is presented. It can be seen that the proposed de-noising method achieves higher homogeneous regions compared to the NLMSF. Clear identification of the shadow and highlight/background regions can be seen in Fig. 6(d). Table I shows the values of  $Q$  for the synthetic sonar images. The experimental results reveal that both methods preserve the object's boundaries. A relatively small advantage in terms of the despeckling index  $Q$  is observed for the proposed method. However, as shown later on, for real sonar images, the performance gap significantly increases.

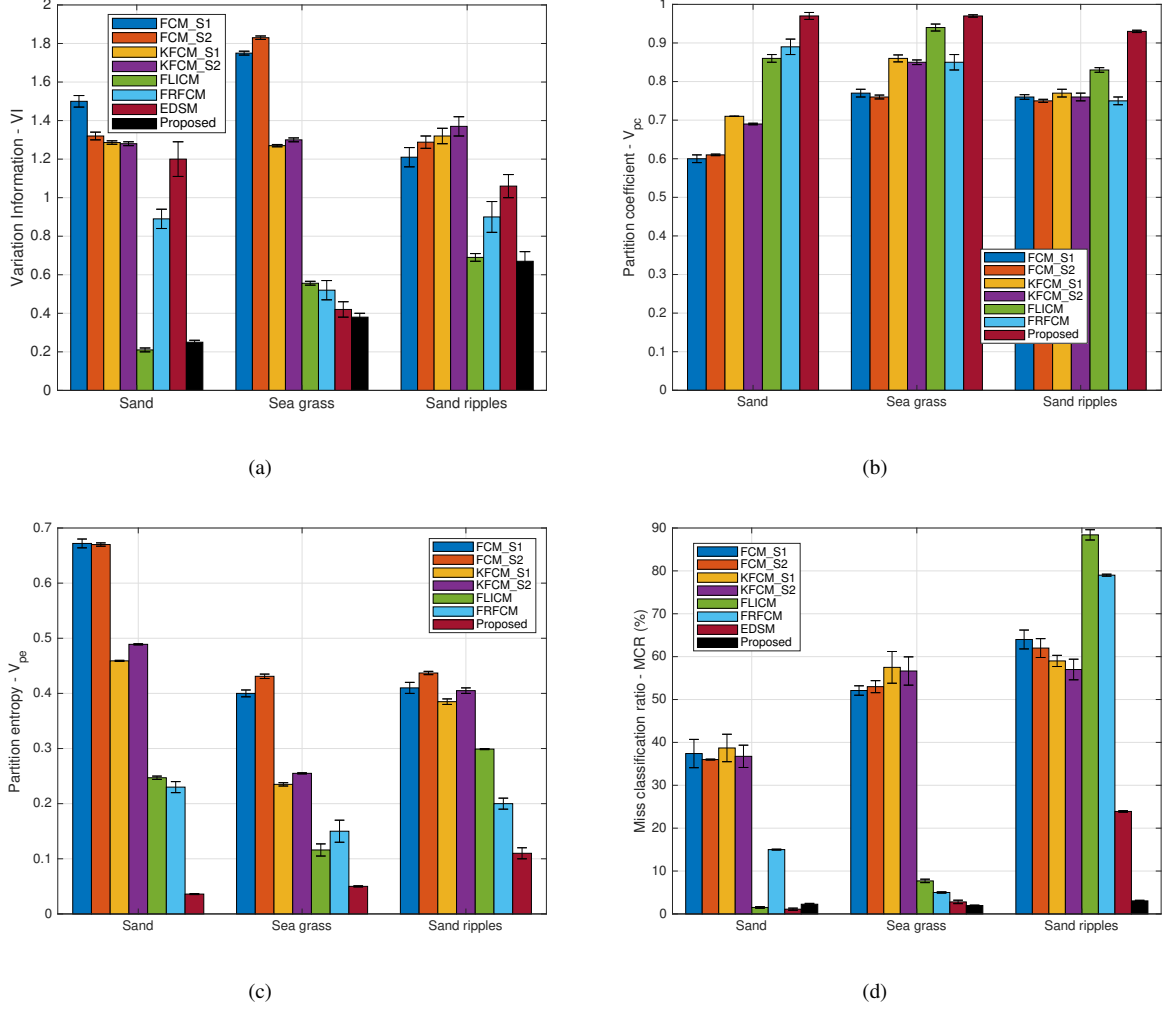


Fig. 8. Variation information (VI) (31), partition coefficient ( $v_{pc}$ ) (33a), partition entropy ( $v_{pe}$ ) (33b), and MCR (35) values for the synthetic sonar images with speckle noise (STD=0.3). The 95% confidence interval is introduced on each bar in the figure. (a) Results for the VI. (b) Results for the  $v_{pc}$ . (c) Results for the  $v_{pe}$ . (d) Results for the MCR (35). The EnFK exceeds the others by 52%, 18%, 64%, and 75% on average over all images for the VI,  $v_{pc}$ ,  $v_{pe}$  and MCR measures, respectively.

TABLE II  
COMPARISON OF THE DESPECKLING ASSESSMENT INDEX  $Q$  FOR THE AIRPLANE, CRAB TRAP AND DROWNING VICTIM SONAR IMAGES WITH  $\alpha = 1$ ,  $M = 3$ ,  $\gamma = 0.5$  AND  $\kappa_s = 255$ .

| Method | Airplane    | Crab Trap    | Drowning Victim |
|--------|-------------|--------------|-----------------|
| EnFK   | <b>9.19</b> | <b>28.77</b> | <b>15.55</b>    |
| NLMSF  | 8.4         | 20.56        | 14.67           |

2) *Segmentation Results*: To statistically test the segmentation performance, we performed 1,000 Monte-Carlo simulation runs. In each run, the image was corrupted by a speckle noise with a variance of 0.09. Fig. 7 shows the segmentation results for the benchmark schemes compared with the EnFK algorithm. The segmentation results are illustrated by three colors: white for highlight, black for shadow and gray for the background. For a sandy background (Fig. 7(a)) FLICM, FCM\_S1, EDSM and the EnFK obtain accurate results. For a sea-grass background (Fig. 7(b)), poor segmentation per-

formances are observed for the benchmark schemes, whereas EDSM and EnFK performance is almost not affected. Still, EnFK has better region uniformity and less misclassified region compared to the results of EDSM. The same result is observed in Fig. 7(c), where we show segmentation results for a sand-ripples background. Here, most benchmarks completely fail to segment the object, whereas the FLICM fails to identify the object's shadow region and wrongly assigns the background region with the shadow label.

The statistical evaluation for the four quality metrics for the Monte-Carlo simulations is presented in Fig. 8. The EnFK achieves the lowest VI for the sand background as it is considered the easiest case for segmentation among the three background types. For all types, EnFK produced the best performance by obtaining the lowest VI. From the results of Fig. 8(b) and Fig. 8(c), we observe that, for the EnFK, the value of  $v_{pe}$  is closer to 1 than that of the benchmarks, and that  $v_{pe}$  is closer to 0, which implies that the EnFK achieves better segmentation accuracy. For a sand background, EDSM

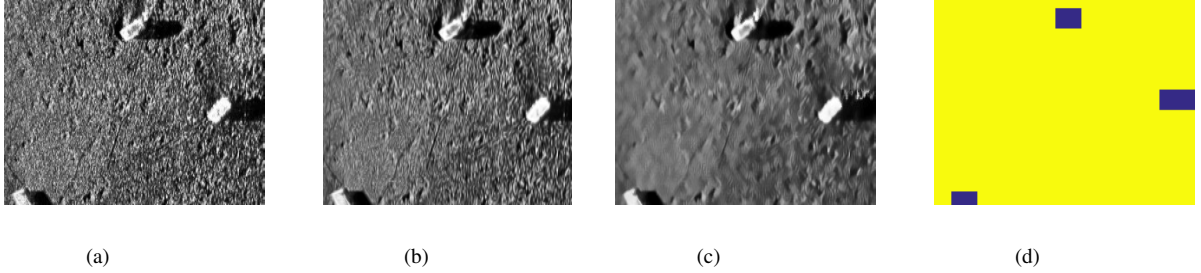


Fig. 9. Comparison of the de-noised images for the real sonar image *Crab Trap* using the proposed filter and the NLMSF filter: (a) Input sonar image. (b) De-noising results of the NLMSF filter. (c) De-noising results of the proposed filter. (d) Decision map of the distributions: blue - Gaussian (for shadow regions), yellow - Exponential (for highlight/background regions). Our proposed de-noising method efficiently reduces the speckle noise, and the background becomes more homogeneous.

achieves the best MCR results. However, for more complex background like the sea-grass and the sand-ripples, EnFK obtains more appropriate and satisfying segmentation results.

In addition, the MCR results in Fig. 8(d) also confirm that our method is more accurate compared to the benchmark with an MCR level of less than 3% for all images.

#### D. Results on Real Sonar Images

1) *De-noising Results:* The de-noised version of the *Crab Trap* image is shown in Fig. 9. High-intensity inhomogeneity can be clearly seen in the original sonar image in Fig. 9(a). Accurate identification of the shadow and highlight/background regions can be seen in Fig. 9(d). To allow for a quantitative comparison between the de-noising methods, in Table II, we show the despeckling assessment index  $Q$  for Fig. 9. Our method produces the highest values of the assessment index compared to the NLMSF, which reflects the homogeneousness of the image after de-noising.

2) *Segmentation Results:* Fig. 10(a) shows the segmentation results for the *Airplane* image. Obvious false segmentation exists in the results produced with FCM\_S1, FCM\_S2, KFCM\_S1, KFCM\_S2. When using these methods, the homogeneity of the background is corrupted by false segmentation, and the boundaries of the object are not well-defined. The FLICM, EDSM, and FRFCM produce better results, but significant false segmentation still exists. In comparison, the EnFK produces the best results, where the highlight, shadow, and background are well distinguished, and the object structure is maintained. The segmentation results of the *Crab Trap* image are shown in Fig. 10(b). We observe that both the FLICM and EnFK correctly identify the object's shadow as well as its highlight. However, the benchmarks FCM\_S2, FCM\_S1, KFCM\_S2 and KFCM\_S1 contain many cases of false segmentation in the image's background. Moreover, the EnFK has the best segmentation results with better region uniformity and misclassified regions. EDSM and FRFCM fail to identify all objects' highlight. Similar results are obtained for the *Drowning Victim* image in Fig. 10(c). The segmentation produced by the FLICM and FRFCM algorithms includes many cases of false segmentation, and the smoothness of the object's boundaries is corrupted. The boundaries in the segmentation results obtained by EDSM and EnFK are well

defined. But, EnFK captures a few misclassified regions. In comparison, the EnFK produces accurate segmentation with better region uniformity.

A quantitative comparison is presented in Fig. 11. For the variation information VI (Fig. 11(a)), results show that the EnFK has the lowest VI value compared to the benchmark. These results indicate that our method generates more uniform segmented regions and more accurate cluster boundaries than the compared algorithms. Figs. 11(b) and 11(c) show the measures of the partition coefficient  $v_{pc}$  and the partition entropy  $v_{pe}$  from (33). We observe that the partition coefficient value of our method is much closer to one than that of the other algorithms. Similarly, the partition entropy value of our method is significantly closer to zero than that of the benchmark algorithms. Both results indicate that the EnFK generates more separable clusters than the benchmark methods.

3) *Effect of the Between-cluster and Local Second Moment Terms:* These terms are used to improve the convergence rate of our algorithm, as well as the segmentation performances. Fig. 12(a) shows the number of iterations until convergence of two versions of the EnFK: one with the two new terms and another without them. We observe that the use of these terms dramatically reduces the number of iterations for all tested images. Moreover, the variation information VI shown in Fig. 12(b) demonstrates that the VI values obtained when using these terms are significantly better compared to not using these terms.

4) *Effect of De-noising:* The effect of sonar image de-noising on the final segmentation results is analyzed in Table III. We measure the efficiency of the de-noising process in terms of the MCR from (35). We test three versions of our fuzzy segmentation algorithm: In version A, the input for the first term of (15),  $\hat{x}_i$ , is replaced with  $y_i$  (no de-noising); in version B, the de-noised data,  $\hat{x}_i$ , is produced by the NLMSF; version C is provided by the EnFK. The large deviation of MCR among the sonar images is mainly due to the different level of intensity inhomogeneity in the images. While the use of the NLMSF de-noising has a very minimal effect, we observe that the EnFK achieves the lowest values of MCR. This means that false segmentation, caused by the intensity inhomogeneity in the background, is effectively reduced by our method.

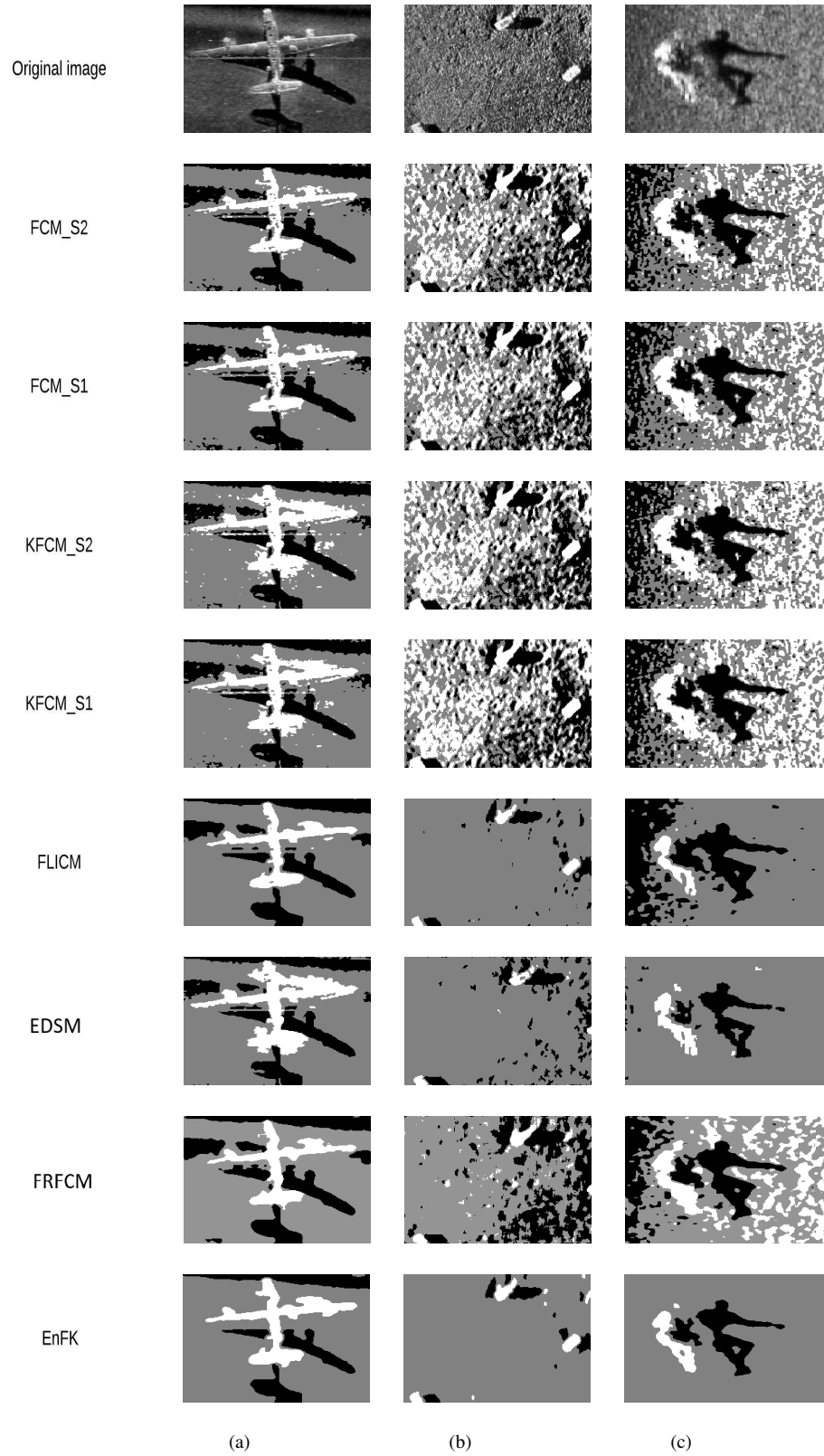


Fig. 10. Final segmentation results for real sonar images (Segmentation results of the second data set). The color in the segmentation maps represent: black for the shadow region, white for the highlight region and gray for the background region. Real sonar images and final segmentation results of the benchmarks FCM\_S2, FCM\_S1, KFCM\_S2, KFCM\_S1, FLICM, EDSM, FRFCM and EnFK. (a) to (c) for *Airplane*, *Crab Trap*, and *Drowning Victim*, respectively. The best segmentation results are obtained by the EnFK.

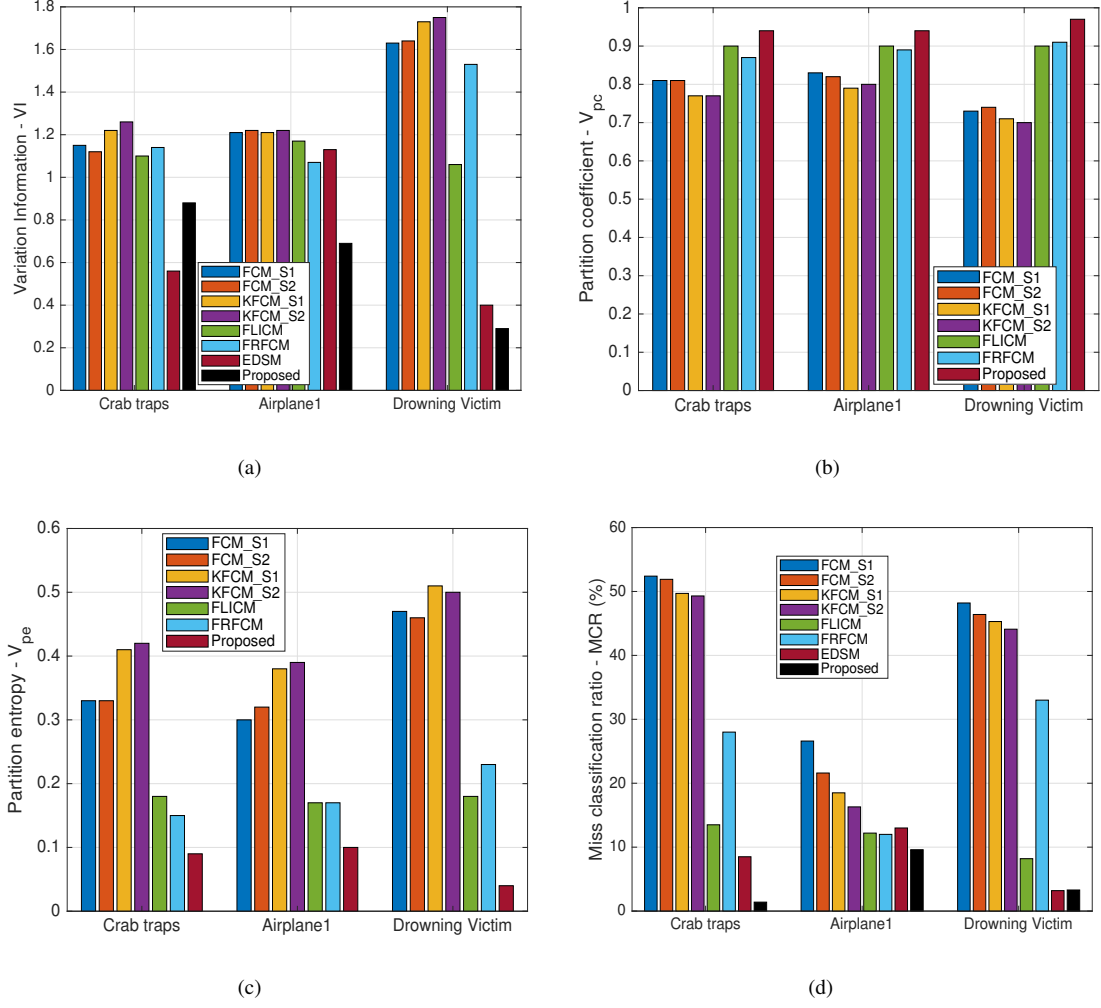


Fig. 11. Variation information (VI) (31), partition coefficient ( $v_{pc}$ ) (33a), partition entropy ( $v_{pe}$ ) (33b), and MCR (35) values for the three cited real sonar images. (a) Results for the VI. (b) Results in terms of  $v_{pc}$ . (c) Results of  $v_{pe}$ . (d) Results in terms of MCR (35). The EnFK exceeds the compared algorithms by 48%, 17%, 76% and 82% on average over all images for the VI,  $v_{pc}$ ,  $v_{pe}$  and MCR measures, respectively.

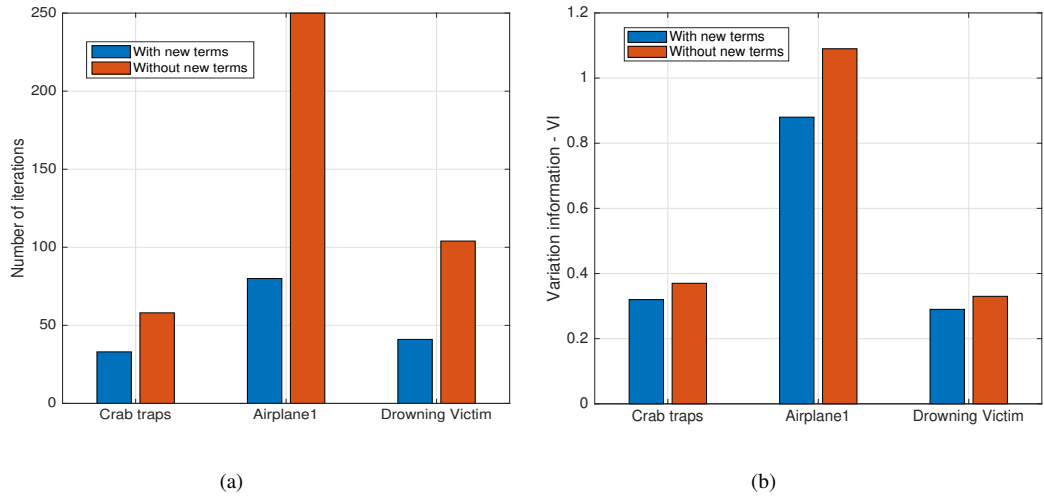


Fig. 12. (a) Convergence rates with the two new terms and without them. Maximum number of iterations  $N_q$  is set to 250. For all images, the novel *between-cluster* and *local second moment* terms dramatically reduce the number of iterations. The large deviation in the results among the sonar images are mainly due to the different sizes of the objects. (b) Variation information VI of the EnFK for the second data set, with the two new terms and without them. Clearly, VI is closer to zero with the new terms.

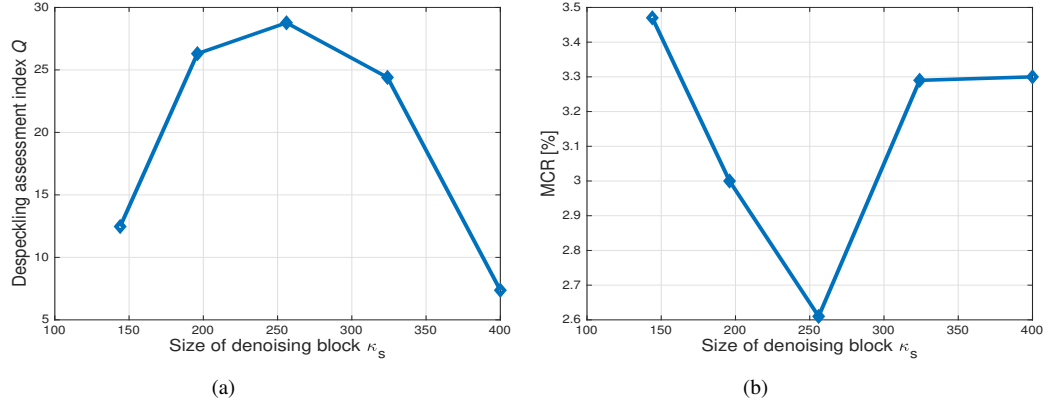


Fig. 13. (a) Despeckling assessment index  $Q$  results in different denoising block size. (b) MCR results.

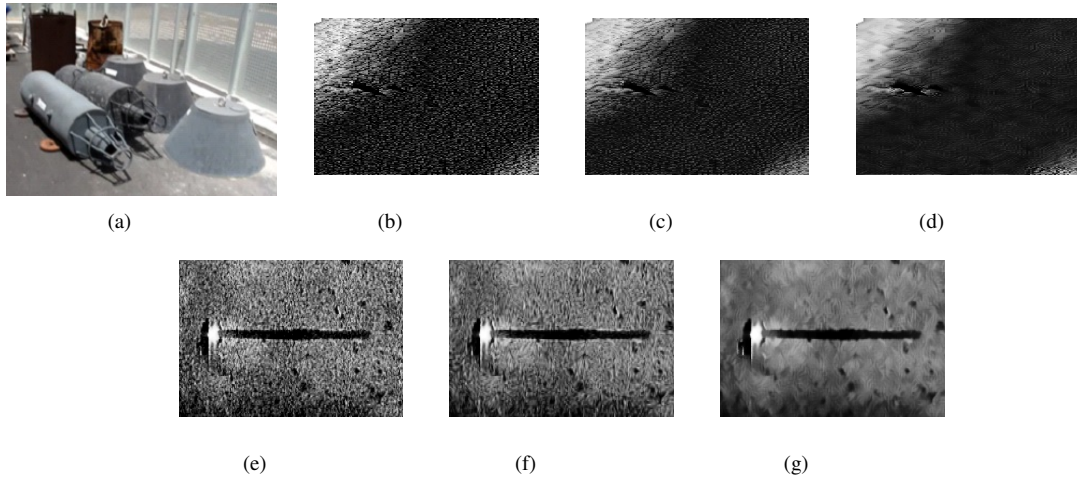


Fig. 14. Comparison of the de-noising results on the two sonar images: (a) Photographs of the targets considered in the sea trial study. Two targets were used: one was cone-shaped (right side of the figure), and one was tube-shaped (left side of the figure). (b) First original sonar image ( $Q=4.3$ ), (c) De-noising results of the NLMSF ( $Q=7.11$ ), (d) De-noising results of our proposed method ( $Q=11.1$ ). (e) Second original SAS image ( $Q=5.8$ ), (f) De-noising results of the NLMSF ( $Q=7.3$ ), (g) De-noising of our proposed method ( $Q=10.2$ ). The EnFK efficiently removes the intensity inhomogeneity in the three sonar images, while preserving the targets' structures.

5) *Effect of De-noising Block Size*: To compensate for the location-dependent in the sonar image, we divide the image into non-overlapped blocks of size  $\kappa_s$  in the de-noising process. The effect of  $\kappa_s$  on the final segmentation results and the de-noising performance in terms of MCR and  $Q$  respectively is analyzed in Fig. 13. To that end, we use the *Crab Traps* from Fig. 9(a). The tradeoff in the choice of  $\kappa_s$  is observed in Fig. 13(a). While small values of  $\kappa_s$  may enlarge the estimation error of the parameters of the distribution, large values of  $\kappa_s$  degrade the performance of the despeckling filter because large blocks contain pixels from different regions. For de-noising, we observe a deviation of 30% in the despeckling assessment index  $Q$ . For the overall segmentation results, we notice a deviation in MCR results of about 1% for different values of  $\kappa_s$ .

#### E. Sea Trial Results

To further validate the performance of the EnFK, we now introduce results from two sea experiments that we performed with our own sonar system. In contrast to the previously

TABLE III  
EFFECT OF DE-NOISING ON SEGMENTATION ACCURACY. COMPARISON OF THE MCR (%) FROM (35) FOR THE AIRPLANE, CRAB TRAP AND DROWNING VICTIM SONAR IMAGES.

| Version           | Airplane   | Crab traps | Drowning Victim |
|-------------------|------------|------------|-----------------|
| A (No de-noising) | 10.1       | 1.6        | 4.3             |
| B (NLMSF)         | 9.7        | 1.4        | 4.1             |
| C (EnFK)          | <b>9.4</b> | <b>1.3</b> | <b>3.8</b>      |

explored sonar images, this data allows us to compare performance with real sea conditions.

1) *Experiment Description*: In the first experiment, a 400 kHz multi-beam sonar, EM 2040, mounted on the RV Bat-Galim, was used to scan two targets of truncated cones, shown in Fig. 14(a). This experiment was performed roughly 2 miles west of Haifa Port in Israel. The targets were placed on a rocky seafloor at a depth of 25 m. The second experiment was conducted about 10 Km west of northern Israel at a water



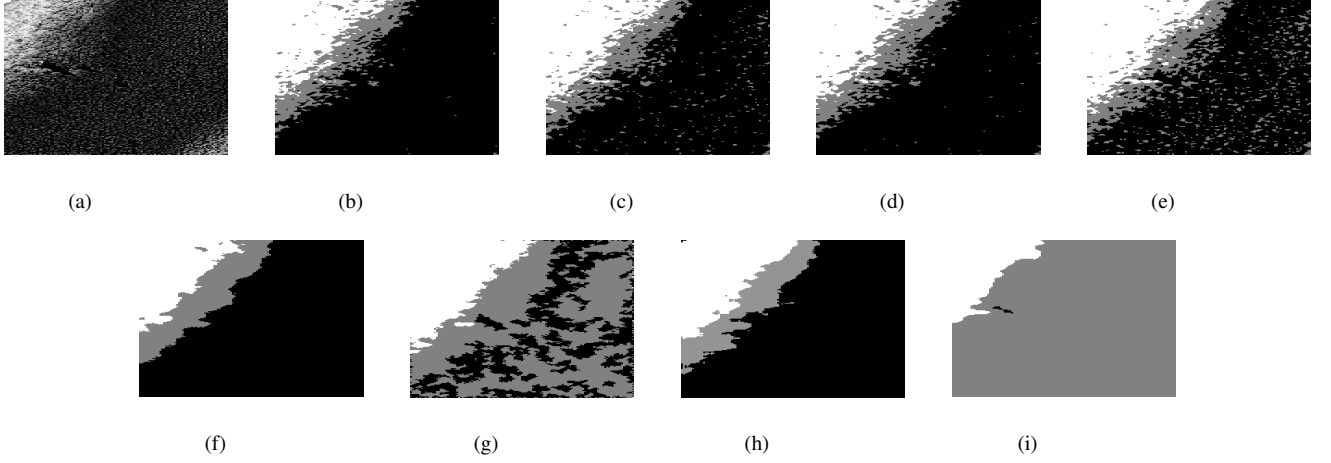


Fig. 15. Comparison of the segmentation results on a multi-beam sonar image with a single target. (a) Original sonar image (b) FCM\_S1 result. (c) FCM\_S2 result. (d) KFCM\_S1 result. (e) KFCM\_S2 result. (f) FLICM result. (g) EDSM result. (h) FRFCM result. (i) EnFK result. EnFK successfully identifies the target's shadow region.

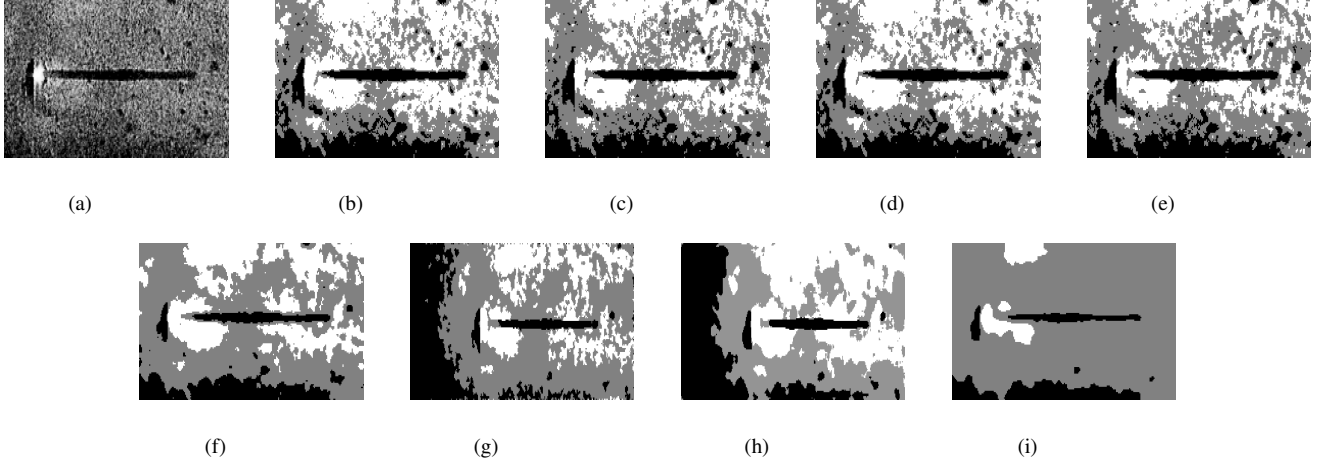


Fig. 16. Comparison of the segmentation results on a SAS sonar image. (a) Original sonar image. (b) FCM\_S1 result. (c) FCM\_S2 result. (d) KFCM\_S1 result. (e) KFCM\_S2 result. (f) FLICM result. (g) EDSM result. (h) FRFCM result. (i) EnFK result. Results show that the EnFK yields the cleanest segmented shadow region.

depth of 1,000 m. In this experiment, we deployed our A18 5.5m Eca Robotics Inc. AUV [47], and scanned the ground using the vehicle's Kraken-made two-sided synthetic aperture sonar (SAS) to scan the seabed for opportunities objects. A SAS image of a found gas well is presented in Fig. 14(e).

2) *Experiment Results:* Fig. 14 shows the de-noising results for the NLMSF and EnFK for the two sonar images. Both methods preserve the edges of the targets. However, the EnFK better smooths the background and, as observed from the resulting despeckling index  $Q$ , the final de-noised image is more homogeneous.

In Fig. 15, the segmentation results are introduced for the multi-beam image. This image contains one cone-shaped target. We observe that the segmentation results obtained by FCM\_S1 (Fig. 15(b)), FCM\_S2 (Fig. 15(c)), KFCM\_S1 (Fig. 15(d)) and KFCM\_S2 (Fig. 15(e)) yield poor performance in terms of region uniformity, and that the FLICM ([Fig. 15(f)]) fails to identify the target's shadow regions.

In FRFCM and EnFK, the region uniformity is good, and the boundaries between regions are clear. But, the EnFK (Fig. 15(g)) manages to separate the target's shadow from its background entirely. For the SAS image, Fig. 16 shows that the shadow region is successfully segmented by all methods. However, the considerable intensity inhomogeneity in this image leads to many false segmentation regions using all methods. Still, the EnFK produces, by far, the lowest false segmentation rate. The simulation and experimental results reveal that the EnFK dramatically reduces the misclassified regions in sonar images, and can generate clean and accurate clusters of shadow and highlight. This result has a tremendous effect on object classification performance, which is the next step in the ADAC detection chain.

## VII. CONCLUSIONS

In this paper, we proposed a fuzzy-based image segmentation method for sonar imagery, referred to as EnFK. To



improve background homogeneity, the EnFK includes a novel image de-noising step which, together with the original image, feeds into the segmentation process. To reduce false segmentation, as part of the fuzzy objective function, we introduced two novel terms. Simulation results show that the use of these two terms dramatically reduces the number of iterations until convergences and that the false segmentation rate decreases. Experiments over a broad set of Monte Carlo simulated sonar images, three cited real sonar images, and two self-measured sonar images from a multi-beam and a SAS show that EnFK produces high segmentation accuracy. The experimental results show that the highlight region in the segmented image of objects in multi-beam sonar images is barely detected. Future work will deal with this problem to enable more accurate segmentation.

## REFERENCES

- [1] D. P. Williams, "Fast target detection in synthetic aperture sonar imagery: A new algorithm and large-scale performance analysis," *IEEE J. Ocean. Eng.*, vol. 40, no. 1, pp. 71-92, Jan. 2015.
- [2] H. Singh, J. Adams, D. Mindell and B. Foley, "Imaging underwater for archeology," *J. Field Archeol.*, vol. 27, no.3, pp. 319-328, 2000.
- [3] Y. Petillot, S. Reed and J. Bell, "Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder," *Proc. IEEE OCEANS conf.*, pp. 217-222, 2002.
- [4] T. Celik and T. Tjahjadi, "A novel method for sidescan sonar image segmentation," *IEEE J. Ocean. Eng.*, vol. 36, no. 2, pp. 186-194, Apr. 2011.
- [5] F. Schmitt, M. Mignotte, C. Collet, and P. Thourel, "Estimation of noise parameters on sonar images," in *Proc. SPIE Conf. Signal Image Process.*, Denver, CO, vol. 2823, pp. 1-12, Aug. 1996.
- [6] R. Fandos and A. M. Zoubir, "Optimal feature set for automatic detection and classification of underwater objects in SAS images," *IEEE J. Sel. Topics Sig. Proc.*, vol. 5, pp. 454-468, Jun. 2011.
- [7] T. M. Nguyen and Q. M. J. Wu, "Fast and robust spatially constrained gaussian mixture model for image segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 621-635, Apr. 2013.
- [8] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124-1137, 2004.
- [9] U. Javed, M. M. Riaz, A. Ghafoor and T. A. Cheema, "SAR image segmentation based on active contours with fuzzy logic," *IEEE Trans. Aero. and Elec. Sys.*, vol. 52, no. 1, pp. 181-188, Feb. 2016.
- [10] J. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32-57, 1974.
- [11] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825-838, March 2007.
- [12] L. Szilagyi, Z. Benyo, S. Szilagyi, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proceedings of 25th Annual International Conference of IEEE EMBS*, 2003, pp. 17- 21.
- [13] R. Shang *et al.*, "A spatial fuzzy clustering algorithm with kernel metric based on immune clone for SAR image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 4, pp. 1640-1652, Apr. 2016.
- [14] J. Noordam, W. van den Broek, and L. Buydens, "Geometrically guided fuzzy C-means clustering for multivariate image segmentation," in *Proceedings of International Conference on Pattern Recognition (ICPR00)*, vol. 1, pp. 462-465, 2000.
- [15] L. Szilagyi, Z. Benyo, S. Szilagyi, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proceedings of 25th Annual International Conference of IEEE EMBS*, pp. 17- 21, 2003.
- [16] M. Ahmed, S. Yamany, N. Mohamed, A. A. Farag and T. Moriarty, "A modified fuzzy C-means algorithms for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 21, no. 3, pp. 193-199, Mar. 2002.
- [17] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 34, no. 4, pp. 1907-1916, 2004.
- [18] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328-1337, May 2010.
- [19] M. Gong, Z. Zhou and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141-2151, Apr. 2012.
- [20] S. Xie, F. Khomh, Y. Zou and I. Keivanloo, "An empirical study on the fault-proneness of clone migration in clone genealogies," in *Proc. IEEE Conf. Software Maint. Reeng. Reverse Eng.*, pp. 94-103, Feb. 2014.
- [21] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613-627, May 1995.
- [22] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simulat.*, vol. 4, no. 2, pp. 490-530, Jul. 2005.
- [23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007.
- [24] P. Coupe, P. Hellier, C. Kervrann, and C. Barillot, "Nonlocal means based speckle filtering for ultrasound images," *IEEE Trans. Image Process.*, vol. 18, no. 10, pp. 2221-2229, Oct. 2009.
- [25] R. James and M. H. Supriya, "Despeckling of sonar images based on a naive homogeneity index," *OCEANS 2016 MTS/IEEE Monterey*, pp. 1-6, Sept. 2016.
- [26] 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, no. 4, pp. 956-970, Oct. 2007.
- [27] J. T. Cobb, "Sonar image modeling for texture discrimination and classification," Ph.D. dissertation, Department of Electrical Engineering, University of Florida, Florida, 2011.
- [28] S. Reed, Y. Petillot and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 1, pp. 90-105, Jan. 2003.
- [29] C. Kervrann, J. Boulanger, and P. Coupe, "Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal," in *Proc. Conf. Scale-Space and Variational Methods*, Ischia, Italy, pp. 520-532, Jun. 2007.
- [30] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise non local means denoising filter for 3D magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 27, pp. 425-441, 2008.
- [31] A. Abu and R. Diamant, "Unsupervised segmentation of underwater objects in sonar images," *Proc. IEEE OCEANS conf.*, June, 2017.
- [32] M. G. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma, "Fuzzy C-means clustering with local information and kernel metric for image segmentation," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 573-584, Feb. 2013.
- [33] D. Zhang and S. Chen, "Kernel-based fuzzy clustering incorporating spatial constraints for image segmentation," in *Proc. IEEE 2nd Annu. Int. Conf. Mach. Learn. Cybern.*, pp. 2189-2192, Oct. 2003.
- [34] L. Chen, C. L. P. Chen, and M. Lu, "A multiple-kernel fuzzy C-means algorithm for image segmentation," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 41, no. 5, pp. 1263-1274, Apr. 2011.
- [35] K. R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181-202, Mar. 2001.
- [36] R. J. Hathaway, J. C. Bezdek, and Y. Hu, "Generalized fuzzy cmeans clustering strategies using norm distances," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 576-582, Oct. 2000.
- [37] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng and A. K. Nandi, "Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Trans. on Fuzzy Systems*, vol. 26, no. 5, pp. 3027-3041, Oct. 2018.
- [38] T. Fei and D. Kraus, "Dempster-shafer evidence theory supported EM approach for sonar image segmentation," *Advances in Systems, Signals and Devices*, vol. 4, pp. 115-129, Dec. 2017.
- [39] Miao, Mu, and Wu, "Computer Simulation of Aeolian Sand Ripples and Dunes," *Physics Letters A, Elsevier Science B.V.*, 2001.
- [40] T. Loupas, W. McDicken, and P. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasound image," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 129-135, 1989.

- [41] [Online - last visited on 06-07-2017]. Available: <http://www.cmaxsonar.com>
- [42] [Online - last visited on 08-09-2017]. Available: <http://www.marinesonic.com>
- [43] [Online - last visited on 08-09-2017]. Available: <http://www.edgetech.com>
- [44] M. Meila, "Comparing clustering by the variation of information," Proc. 6th Ann. Conf. Comput. Learning Theory, pp. 173-187, 2003.
- [45] K. Chuang and H. Tzeng, "Fuzzy c-means clustering with spatial information for image segmentation," Comput. Med. Imag. Graph., vol. 30, pp. 9-15, 2006.
- [46] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm," IEEE Trans. Med. Imaging, vol. 20, no. 1, pp. 45-57, Jan. 2001.
- [47] [Online - last visited on 15-10-2017]. Available: <http://www.ecagroup.com>



**Avi Abu** received his B.Sc and M.Sc degrees from Tel Aviv University, Tel Aviv, Israel, in 1998 and 2003, respectively. He is currently pursuing his Ph.D. degree at Haifa University's Underwater Acoustic and Navigation Laboratory (ANL), Haifa, Israel. His research interests include sonar image segmentation, underwater object detection and sonar image de-noising.



**Roe Diamant** received his PhD from the Department of Electrical and Computer Engineering, University of British Columbia, in 2013, and his B.Sc. and M.Sc. degrees from the Technion, Israel Institute of Technology, in 2002 and 2007, respectively. From 2001 to 2009, he was employed in Rafael Advanced Defense Systems, Israel, as a project manager and systems engineer, where he developed a commercial underwater modem with network capabilities. In 2015 and 2016, he was a visiting Prof. at the University of Padova, Italy. In

2009, he received the Israel Excellent Worker First Place Award from the Israeli Presidential Institute. In 2010 he received the NSERC Vanier Canada Graduate Scholarship. Dr. Diamant received two best paper awards, and serves as an associate editor for the IEEE Ocean Engineering. He currently leads the Underwater Acoustic and Navigation Laboratory (ANL) as a senior lecturer (equivalent to assist. Prof.) at the Dept. of Marine Technology, University of Haifa. His research interests include underwater acoustic communication, underwater navigation, object identification, and classification.