Robust Image Denoising for Sonar Imagery

Avi Abu and Roee Diamant
Department of Marine Technologies, School of Marine Sciences, University of Haifa, Haifa, Israel.
E-mail: aabu@campus.haifa.ac.il, roeed@univ,haifa.ac.il

***Abstract*—The recent boost in undersea operations has led to the development of high-resolution sonar systems mounted on autonomous vehicles, and aimed to scan the sea floor and detect objects. An important part of sonar detection is the image denoising, where the background is smoothed and noise components are removed while preserving the object's borders. Sonar image denoising is a challenging task, mostly due to the heavy intensity inhomogeneity of the background and the heavy spatial varying background. In this paper, we propose an algorithm for sonar image denoising that is based on the adaptation of the nonlocal means-based filter. The noise in the highlight and background regions is modeled by the exponential distribution, while the noise in the shadow region is modeled by the Gaussian distribution. We estimate the label of each pixel through image segmentation to estimate the parameters of each distribution. Then, the minimum entropy criteria is used to decide which statistics model in the denoising filter gives the best results. Results for synthetic sonar images and over real sonar images demonstrate that the proposed method successfully removes the noise components while preserving the object's edges.**

***Index Terms*—Image enhancement, Speckle filter, NL Means-based denoising, Sonar signal processing.**

I. Introduction

HIGH-resolution imagery of the seabed is mostly provided by sonar systems such as sidescan sonar and synthetic aperture sonar (SAS). The resulting sonar images are fed into the automatic detection and classification (ADAC) process for detecting underwater objects [1]. The process is commonly applied on board an autonomous underwater vehicle for realtime survey of a designated area. ADAC is used for marine applications including seabed archeology [2], pipeline monitoring [3] and offshore oil prospecting [4]. The key for a successful object detection and classification through ADAC is the segmentation process in which the image is clustered into *background*, object's *highlight*, and object's *shadow*. Only then, the object's features are identified and the object is classified. In this work, we focus on the de-noising part of segmentation.

De-noising of a sonar image aims to remove noise components without distorting the object's borders. De-noising is particularly challenging in sonar imaging since the seabed is complex and comprises of many outliers. Wiener filtering and wavelet transform [5] are two typical approaches for image denoising, but mostly apply for mitigating noise transients. The non-local filter [6], performs de-noising by a weighted average of all pixels intensity in the image. The per-pixel weight is determined according to the similarity between the pixel's local neighborhood and the local neighborhood of other pixels. The 3-D filtering [7] performs a similar process but combines the nonlocal filter with a Wiener filter. Wavelet transform can reduce noise with good preservation of edges by thresholding the wavelet coefficients [8]. Alternatively, Coupe *et al.* [9] proposed the nonlocal means-based speckle filtering (NLMSF) method that can preserve the borders of objects in ultrasound images. The filter works well for unique objects, but seems to fail when the background is densely occupied by a large number of reflectors like sand ripples and boulders, which are quite common in near-shore environments [10].

Our approach relies on modeling the intensity of the image's pixels. The model used for the pixel intensity includes an additive noise component that originates from the sonar system [11]. Alternatively, a speckle model was considered in [9], where the pixel's intensity is modeled by an additive noise that is a multiplicative Gaussian noise with the true image. Commonly, the Rayleigh distribution is used to model the statistics of the noise component of the pixel's intensity [12]. Another statistical model is the Weibull distribution [13]. However, while these distributions have been proven to reflect well the statistics of noise for optical images as well as for ultrasound images, they have not been well explored for denoising of sonar images.

The de-noising of sonar images is much different than that for optical or ultrasound images. For one, the inhomogeneity nature of the seabed provides a complex and highly spatial varying image background where both the Rayleigh and the Weibull assumptions may not hold. In fact, the current denoising schemes do not consider well the differences between the highly diverse nature of the sonar background, the relatively fixed pixel's intensity of the object's shadow, and the mildly intensity variation in the highlight region. As a result, the current literature is missing a distribution model that can reflect well the statistics of the pixel's noise in the shadow, background and highlight regions in sonar images, and a denoising procedure that takes these difference into account. Considering these remaining gaps, in this work we propose a novel de-noising algorithm specifically designed for sonar images.

II. Proposed De-noising Solution

Our de-noising approach aims to reduce the intensity inhomogeneity of the sonar image. This objective is of importance for the task of object segmentation, where not only the detection of the object is of concern but also keeping its observed shape is of importance for the classification procedure. The key

A. Abu and R. Diamant are with the Department of Marine Technologies, School of Marine Sciences, University of Haifa, Haifa, Israel, (e-mail: aabu@campus.haifa.ac.il, roeed@univ,haifa.ac.il).

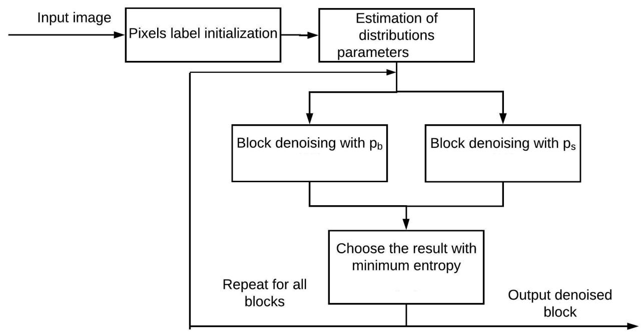

Fig. 1. Flowchart of the proposed de-noising scheme.

idea of our de-noising scheme is to utilize different statistics models for different regions in the sonar images, namely, the shadow, the highlight and the background, and tie it in a single Bayesian analysis. As we will show in the results section, this approach leads to improved results in terms of the assessment index, Q, [9].

The de-noising procedure is performed on non-overlapping blocks of the original sonar image. We model the statistics of the additive noise of each pixel according to the pixel's label. In particular, a shadow region is created due to the lack of acoustic reverberation behind the object. Thus, the noise is mostly electronically driven, and the noise is modeled by a zero mean Gaussian distribution. Differently, the noise in the background and the highlight regions is modeled by the Exponential distribution [14].

The flowchart of our de-noising scheme is shown in Fig. 1. We start by labeling the pixels using our LSM segmentation initialization procedure in [15]. Then, the label of each block is set by the majority labels of pixels in the block, and we use a maximum likelihood estimator to evaluate the distribution parameters within each block. The parameters estimation from each block are fused together to obtain a final distribution estimation per label. As we show in the results section, our modeling of the noise distribution offers a better distinction between pixels related to the background vs. pixels related to the object, as opposed to other distribution models e.g., Gaussian or Weibull.

Our de-noising is based on the Bayesian formulation in [9]. However, different than [9] where the pixel's intensity is assumed effected by a Gaussian distribution, we allow a different distribution type for the noise in each block, which is set based on a decision criteria that maps blocks into shadow or background/highlight region. More specifically, for each block we perform two de-noising attempts: one based on the exponential distribution and one based on the Gaussian distribution. We then select the best result in terms of minimal block entropy. This is because the entropy characterizes the spreading of the restored intensities, and thus very localized regions (as required at the output of the de-noising filter) lead to small entropies, while uniform regions lead to high entropies. In other words, the criteria of minimal entropy leads to a homogeneous intensity after de-noised, thereby significantly relaxing the segmentation effort. To summarize, the process steps are:

1) Estimate the parameters of the two possible distribution functions.
2) De-noise each block assuming Gaussian distribution.
3) De-noise each block assuming exponential distribution.
4) Choose De-noising result that leads to minimum entropy.
5) Repeat (2)-(4) for all blocks in the image.

A. Systems Model and Main Assumptions

Let $\mathcal{Y}$ be a two-dimensional sonar image with dataset $\{y_1, ..., y_N\} \subseteq \mathcal{Y}$, 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.

We model the noisy image as

$$y_i = x_i + \xi_i \,, \tag{1}$$

where ξ_i is a conditionally independent additive noise with a probability density function (PDF) $p_\xi(\xi)$ and x_i is the true image pixel. Note that 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 [11]. Following [14], 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 [16], or the Weibull distribution [13]. Moreover, the results also demonstrate that this distribution model is sufficiently valid to provide accurate de-noising results. We model the PDF of ξ_i by

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

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

B. The De-noising Filter

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 keeping its observed shape to ease the classification procedure. The key idea in our de-noising scheme is to use the Bayesian approach in [9] to tie different statistical models to the sonar images' different regions, namely, shadow, highlight, and background. As we will show in the Results section, this approach leads to improved results in terms of assessment index (Q) [9], which reflects region homogeneity level. Moreover, we introduce a novel method to self-evaluate the de-noising accuracy, thus avoiding the need to initialize the de-noising process.

Our de-noising is based on the NLMSF in [9]. 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 speckle model is given by:

$$y_i = x_i + x_i^{\gamma} \xi_i \,, \tag{3}$$

where $\xi_i \backsim \mathcal{N}(0, \sigma^2)$ is a zero mean Gaussian noise, and γ is the speckle model parameter. The blockwise Bayesian estimator $\widehat{\mathbf{x}}(B_i)$ is defined as [17]

$$\widehat{\mathbf{x}}(B_i) = \frac{\sum_{k=1}^{|\triangle_i|} \mathbf{y}(B_k) p(\mathbf{y}(B_i)|\mathbf{y}(B_k))}{\sum_{k=1}^{|\triangle_i|} p(\mathbf{y}(B_i)|\mathbf{y}(B_k))} \,. \tag{4}$$

where B_i is a square block of size T equals $(2\alpha+1)^2$ ($\alpha \in \mathbb{N}$) centered at pixel i, $\triangle_i$ is a square search block centered at pixel i of size $|\triangle_i| = (2M+1)^2$ ($M \in \mathbb{N}$), $\mathbf{y}(B_i)$ 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 (3), the statistical distribution of $y_i|x_i$ is

$$p(y_i|x_i) \propto \mathcal{N}(x_i, x_i^{2\gamma}\sigma^2) \,. \tag{5}$$

Assuming independence among the pixels, the likelihood of $\mathbf{y}(B_i)|\mathbf{y}(B_k)$ can be factorized as:

$$p(\mathbf{y}(B_i)|\mathbf{y}(B_k)) \propto \exp\left(-\sum_{t=1}^{T} \frac{(y_i^{(t)} - y_k^{(t)})^2}{2(y_k^{(t)})^{2\gamma}\sigma^2}\right), \tag{6}$$

where $y_i^{(t)}$ and $y_k^{(t)}$ are the tth 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 [18], which is controlled by the thresholding parameter μ_1.

The free parameters γ and μ_1 affect the robustness of our de-noising scheme to seabed intensity inhomogeneity. This is because of the need to tune these parameters for different background types. Considering this challenge, we add to our 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.

1) Estimation of Distribution Parameters: We make the practical 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, we divide the image into non-overlapped blocks Y_r of size κ_s, and perform parameter estimation per block. Then, modeling the distribution of the noise components to be the same for the whole image, we fuse the estimation from all blocks into a single one. We note the choice of κ_s tradeoffs. Small values of κ_s may enlarge the estimation error of the distributions parameters, while large values of κ_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 κ_s to the user based on the size of the object of interest.

As 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 rth block. We determine c_r based on the majority of the pixels' labels in the rth block. We evaluate these pixels' labels based on our suggested initialization algorithm from [15]. While our initialization performed well for real sonar images, because it is a model-free algorithm, it may induce some clustering errors. Still, we found that the impact of such errors on the parameter estimation is low. This is because, as we formalize below, both λ and σ from (2) are assumed constant throughout the image's blocks and the final estimation of the distribution parameters is a weighted sum estimation from all relevant blocks.[1]

Once c_r is determined, we statistically evaluate the parameters in (1). For blocks with $c_r = \{B, H\}$, we evaluate parameter λ in the rth block by

$$\widehat{\lambda}_r = \left(\frac{1}{\kappa_s} \sum_{i \in Y_r} (y_i - \overline{y})^2\right)^{-0.5} , \tag{7}$$

where

$$\overline{y} = \frac{1}{\kappa_s} \sum_{i \in Y_r} y_i \,. \tag{8}$$

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

$$\widehat{\sigma}_r = \sqrt{\frac{1}{\kappa_s} \sum_{i \in Y_r} y_i^2} \,. \tag{9}$$

Then, in accordance with our assumption that the noise term parameters in (1) are constant throughout the sonar image, we follow the metric in [19], 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:

$$\widehat{\lambda} = \sum_r \varrho_r \widehat{\lambda}_r \,, \tag{10}$$

with

$$\varrho_r = \frac{\exp(-|\widehat{\lambda}_r - \overline{\lambda}|)}{\sum_r \exp(-|\widehat{\lambda}_r - \overline{\lambda}|)} \,, \tag{11}$$

and

$$\overline{\lambda} = \frac{1}{\rho_b} \sum_r \widehat{\lambda}_r \,, \tag{12}$$

where ρ_b is the number of blocks labeled as highlight or background. The fusion of parameter σ is performed in the same fashion.

[1] Note that the usage of initialization in the process of de-noising is used only for parameter estimation and not for the block-based de-noising.

2) Setting the De-noising Filter: The process of de-noising is carried out separately for each block Y_r. Since the blocks' size of the sonar image, T and $|\triangle_i|$ are much smaller than the original image size, for the purpose of 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 (4) can be rewritten as

$$p(\mathbf{y}(B_i)|\mathbf{y}(B_j)) = \prod_{t=1}^{T} p_{\xi}\left(y_i^{(t)} - y_j^{(t)}\right). \tag{13}$$

Based on the distribution model in (2), to calculate (13), we require information about the label of each block. The process of de-noising does not tolerate an 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, in regard to block de-noising, we avoid using segmentation initialization. Instead, for each block, we find the distribution that leads to the best de-noising result.

3) 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

$$H_{\mathbf{x}} = -\sum_i p(i)\log_2 p(i), \tag{14}$$

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

With the uncertainty of the block's label, each block is de-noised using (13) for both $p_b(\xi)$ and $p_s(\xi)$ to create the de-noised images $\widehat{\mathbf{x}}_b$ and $\widehat{\mathbf{x}}_s$, respectively. Then, the entropy is calculated for each of the resulting images, and the chosen de-noising result is the one of minimum entropy.

III. Performance Results

We compare the results of our de-noising algorithm with those of NLMSF in [9] and Co-occurrence filter (CoF) in [20]. The NLMSF is a de-noising algorithm designed for sonar images, while CoF is a common de-noising technique in optical images. We consider two sonar images. The first image is a 201×201 synthetic sonar image composed of a *cylindrical* object with sea-grass as background (Fig. 2). The second is a 151×301 image taken from a CM2 tow-fish sonar and includes *crabs traps* (Fig. 3). For both figures, image (a) shows the original image, image (b) shows our de-noised solution, image (c) shows the benchmark de-noising using NLMSF, image (d) shows the benchmark de-noising using CoF, and image (e) is the decision map that is based on the minimal entropy criteria. We observe that our proposed de-noising performs better than the NLMSF filter, i.e., the background is more homogeneous. Moreover, from the decision map, we observe a clear identification of the shadow and highlight/background regions.

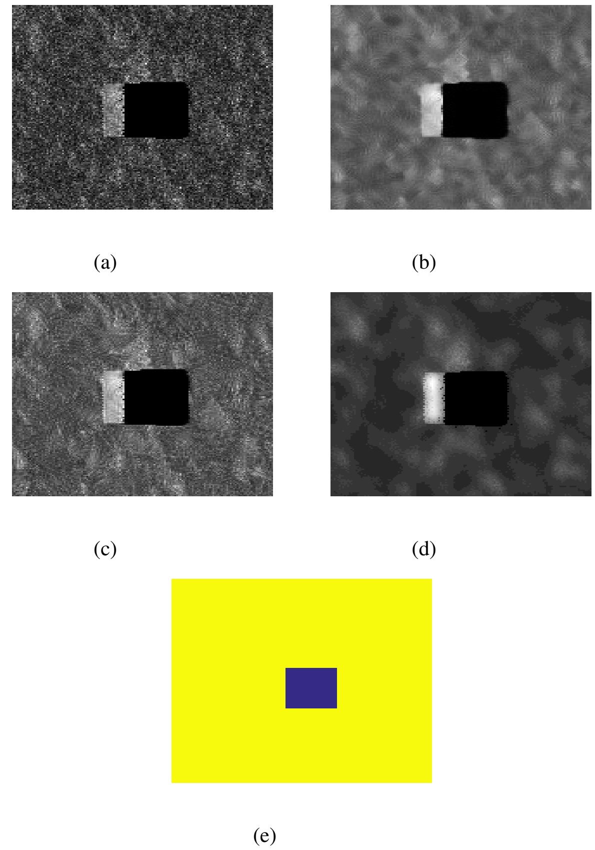

Fig. 2. De-noising of a cylindrical object. (a) Input sonar image; (b) Proposed approach; (c) NLMSF, (d) CoF, (e) Decision map: blue – shadow, yellow – highlight or background.

To quantitatively evaluate the performances of our proposed denoising method, we use the despeckling assessment index Q. Despeckling assessment index Q is defined as [9]

$$Q = \frac{\sum_{j \neq k} (\mu_j - \mu_k)^2}{\sum_j \sigma_j^2}, \tag{15}$$

where μ_j and σ_j are the mean and the variance of the pixels' intensities with assigned jth label after de-noising. To calculate the despeckling assessment index, we use the ground truth map for pixels' label information. The higher the value of Q, the better de-noising results are achieved.

Table I shows the values of Q for the *cylindrical* and the *Crab traps* sonar images. The experimental results reveal that both methods preserve the object's boundaries. However, our method is more efficient at smoothing the noise in the images, and more suitable for sonar imaging.

IV. Summary and Conclusions

In this paper, we focused on the de-noising process of sonar images. This process is required in applications requiring detection and/or classification of objects, and its aim is to smooth the image's background without distorting the object. We offered a new de-noising scheme, aimed particularly for sonar images. The scheme uses different statistics for the

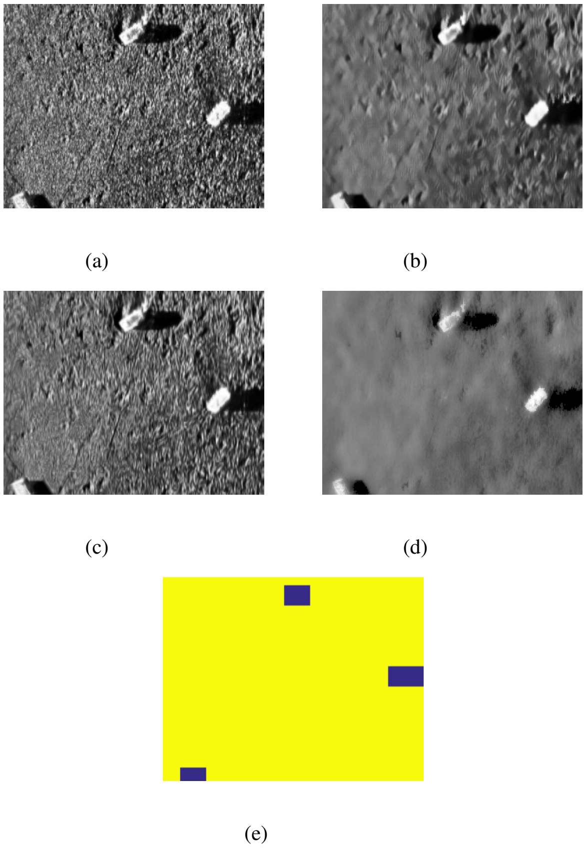

Fig. 3. De-noising of a crab trap image. (a) Input sonar image; (b) Proposed approach; (c) NLMSF, (d) CoF, (e) Decision map: blue – shadow, yellow – highlight or background.

TABLE I
COMPARISON OF THE DESPECKLING ASSESSMENT INDEX Q FOR THE CYLINDRICAL AND CRAB TRAPS SONAR IMAGES WITH $\alpha = 1$, $M = 3$, $\gamma = 0.5$ AND $\kappa_s = 255$.

Method	*cylindrical*	*Crab traps*
Proposed method	**10.1**	**28.77**
NLMSF	9.51	20.56
CoF	7.37	8.87

shadow and highlight/background regions that exist in sonar imagery. Based on the concept of minimum entropy, we also offered a quality measure to self-evaluate the de-noising result. Synthetic and experiments results show that our de-noising method successfully removes the noise components while preserving the object's edge. Future work will incorporate the de-noising scheme within the process of image 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 sonat image segmentation," IEEE J. Ocean. Eng., vol. 36, no. 2, pp. 186-194, Apr. 2011.
[5] D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inf. Theory, vol. 41, no. 3, pp. 613-627, May 1995.
[6] 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.
[7] 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.
[8] D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inf. Theory, vol. 41, no. 3, pp.613-627, May, 1995.
[9] 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.
[10] 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.
[11] 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.
[12] W. I. Stewart, D. Chu, S. Malik, S. Lerner, and H. Singh, "Quantitative seafloor characterization using a bathymetric sidescan sonar," IEEE Journal of Oceanic Engineering, vol. 19, no. 4, pp. 599-610, 1994.
[13] 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.
[14] J. T. Cobb, "Sonar image modeling for texture discrimination and classification," Ph.D. dissertation, Department of Electrical Engineering, University of Florida, Florida, 2011.
[15] A. Abu and R. Diamant, "Unsupervised segmentation of underwater objects in sonar images," Proc. IEEE OCEANS conf., June, 2017.
[16] 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.
[17] 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.
[18] 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.
[19] 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.
[20] R. J. Jevnisek, and S. Avidan, "Co-occurrence Filter," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3184-3192