A WAVELET SHRINKAGE APPROACH TO DETECT CANDIDATE POINT SCATTERERS IN SYNTHETIC APERTURE SONAR IMAGES FOR RESOLUTION ESTIMATION

Marc Geilhufe — Norwegian Defence Research Establishment (FFI), Kjeller, Norway
Derek R. Olson — Naval Postgraduate School (NPS), Monterey, California, USA
Roy E. Hansen — Norwegian Defence Research Establishment (FFI), Kjeller, Norway
Stig A. V. Synnes — Norwegian Defence Research Establishment (FFI), Kjeller, Norway

1 INTRODUCTION

Autonomous Underwater Vehicles (AUVs) that are equipped with side-looking synthetic aperture sonar (SAS) have become a popular asset for efficient high resolution imaging and mapping of the seafloor. There are situations where it can be challenging to produce SAS images of good quality. For instance, this depends on the ocean environment, vehicle stability, navigation accuracy or seafloor bathymetry. For running operations more autonomously, an AUV must be able to estimate its sonar performance through-the-sensor such that it can automatically adjust track orientation, processing techniques or other parameters to improve performance. Hence, it is imperative to develop techniques to assess image quality from the imagery itself. Given sufficient signal to noise ratio, we recognize image sharpness as the most important quality metric[1], preferably specified as a number between $[0, 1]$ expressing how close the achieved resolution is to the theoretical resolution of the system.

SAS images are formed by coherently combining echoes gathered over an interval of acoustic frequencies and aspect angles. The standard approach for estimating resolution in an image is to calculate the -3 dB width of candidate point scatterers, but there are various methods on how to find suitable candidate scatterers[2–5]. Point scatterers must be distinguished from larger targets as well as speckle in order to be able to use them to correctly estimate the image resolution. Large targets are undesirable because their width in space is much larger than the system resolution. Speckle is also undesirable because the width of the peaks in speckle are always the same size, independent of the focus quality of the system.

In this paper, for detecting candidate point scatterers in SAS images, we explore the suitability of a wavelet shrinkage method[6] using two non-overlapping looks obtained from a full resolution single look complex (SLC) image. We have applied this method in earlier work[3] and will here focus on a complete description and explanation of how and why it works. The idea is that for point scatterers, the coherence between looks is high, while speckle and noise are uncorrelated. Focusing on the relevant scales and orientations, our approach can enhance point scatterers and de-emphasize texture, thereby facilitating the detection of point scatterers for resolution estimation.

The remainder of this paper is as follows. In Section 2 we give an introduction to the wavelet shrinkage approach. Section 3 interprets the method in light of its application of detection of point scatterers, followed by Section 4 highlighting details on our choices for the implementation. Results for simulated and real SAS data are shown in Section 5. The paper concludes with a summary and outlook over future work in Section 6.

2 WAVELET SHRINKAGE

The wavelet approach used in this study utilizes the maximal overlap discrete wavelet transform (MODWT)[7]. Opposed to the discrete wavelet transform (DWT), the MODWT is valid for all sample

sizes, is shift-invariant and does not involve any subsampling; i.e. for each scale the MODWT coefficients have the same dimension as the input data. We see these properties as favorable for our subsequent analysis since they facilitate interpretation of wavelet coefficients across all orientations and scales for given input data.

Let us denote the one-dimensional wavelet filters as $\{h_{j,l}\}$ and scaling filters as $\{g_{j,l}\}$ of even length L_j with $l = 0, ..., L_j - 1$. Level $j \in \mathbb{N}$ is associated to scale $\tau_j = 2^{j-1}$. We use filters of the Daubechies class[8] of length L, $D(L)$, where the simplest filter $D(2)$ is the Haar wavelet. For 2-D data, the MODWT coefficients are calculated as a tensor product of the 1-D filters with the input data[9]. There are four possible ways to combine the 1-D filters, resulting in four distinct coefficient matrices: wavelet-wavelet W_j, scaling-wavelet U_j, wavelet-scaling V_j, which capture respectively diagonal, vertical and horizontal structure, and scaling-scaling Z_j, which is the j^{th}-level smoothed version of the input data. Well-sampled SAS images can be transformed into the frequency-aspect domain. For the application in this paper, the information in the frequency-aspect domain is divided in two non-overlapping domains and used to form a pair of independent images; i.e. we obtain two looks from the original SLC image. The approach of wavelet shrinkage[10,11] aims at preserving coherent structures, while suppressing incoherent noise. It locally extracts the similarity of the images decomposed into the scales and orientations of the wavelets and suppresses wavelet coefficients with low coherence between looks. The gathered information is used to indicate the location of point scatterers throughout the image. By operating on individual scales, the technique has an advantage over operating directly on the image with the signal being focused on fewer coefficients than speckle or noise. This will be the typical case with focused point scatterers and focused edges.

Wavelets locally extract information on structures of different scales and orientations. While spatially scale-variant texture will be concentrated on a few coefficients, the expected contribution from noise is the same on all coefficients[10,11]:

1. Wavelet coefficients of white noise are (on average) the same for all orientations and scales (though with larger variance at small scales).
2. Wavelet coefficients of point scatterers are independent of orientation, and decay by $1/\tau_j$ for increasing scales τ_j.
3. Wavelet coefficients of edges are strongest for the orientation and scale addressed by the wavelet, and decay with the mismatch ratio, δ, (taken as the largest of edge length and wavelet scale divided by the smallest) as $1/2^{\delta}$.

Defocused scatterers will lead to reduced coefficients at the lowest scales, and increased coefficients at the scales of the blurring.

Our wavelet shrinkage approach for detection of point scatterers is as follows: We perform the MODWT both to a given SLC image and to the corresponding two non-overlapping multi-look images. Next, we calculate the coherence $\rho_{C_j}(x, y)$ between the MODWT coefficients C_j (being the 2-D wavelet coefficient matrices W_j, U_j, V_j or Z_j) from the two looks at position (x, y) and obtain wavelet shrinkage coefficients $\zeta(\rho_{C_j}(x, y))$ by thresholding the coherence of the MODWT coefficients for various scales and orientations. We incorporate the shrinkage coefficients as weights multiplied with the MODWT coefficients of a full resolution SLC image. Then the inverse wavelet transform results in an image with enhanced point scatterers. In order to obtain the location of candidate point scatterers, we check local maxima in the enhanced image following the procedure from previous work[5]. The estimation of along- and across-track resolution for the detected point scatterers will be performed both on the original SAS image and the enhanced image.

3 DETECTION OF POINT SCATTERERS

We adopt the magnitude of the coherence[12] as a metric for estimating linear dependence between looks. The coherence expresses to what degree the contributions (observed in either the spatial or the frequency-aspect domain) combine coherently to form the original image. For the development of our method, we recognize that[13]

1. Focused point scatterers are coherent over all frequencies and angles.
2. Targets are coherent over a span of frequencies and angles.
3. Background speckle is incoherent over all frequencies and angles.
4. Random additive noise is incoherent over all frequencies and angles.

We further assume that defocused point scatterers and targets have reduced coherence, while the frequency and aspect coverage is unchanged. We expect that the -3 dB width of a scatterer will indicate where half the signal is in phase and half the signal is out of phase. Therefore, we expect the coherence to be above 0.5 over the -3 dB width of the target echo response also for defocused point scatterers and targets.

For local estimates of the coherence, the number of samples available will be limited, and the coherence estimates will be impeded by both a significant bias and a large variance[12]. Random peaks in the estimated coherence on noise or speckle could be suppressed by averaging coherence estimates made on different realizations of image pairs[6]. Sparse representations of the data might introduce image artifacts[11], and this can affect the coherence estimates. Moreover, running many realizations of the estimate might be time consuming on large data sets.

Given that a point scatterer or target of interest is significantly stronger than the surrounding background and noise, we expect high coherence between looks for a point scatterer and lower coherence on a defocused scene. For targets, we expect high coherence when both images have their frequency-aspect domain support within the coherent domain, and low otherwise.

By direct application of wavelet shrinkage, we can enhance point scatterers and edges in the image. The technique scales the wavelet coefficients by a weight related to the coherence estimate. In effect, random signals such as noise and speckle are suppressed, while persistent scatterers such as points and edges are preserved. Any high-value coherence estimates indicate the presence of point scatterers or edges in a map. By operating in the wavelet coefficient domain, weaker (or defocused) point scatterers might be detected than when operating in the spatial domain. Furthermore, we might be able to derive the size of the scatterer from the trend of the wavelet coefficients as function of scale.

In addition to filtering the image, high or moderate coherence values can be used separately to map the presence of persistent scatterers. Edges could be separated from points by inspection of the wavelet coefficients. Points could be detected by searching for the expected trend over scale, on each dimension separately. The location of the maximum wavelet coefficient on the scale-axis might also be used to indicate the resolution directly.

4 IMPLEMENTATION CHOICES

4.1 Image pairs

We choose to form an image multilook pair, with half the aspect domain support on each. As result, the image resolution is degraded by a factor of two in the along-track dimension. By avoiding a random gridded division[6], we avoid the risk of introducing image artifacts from holes in the data support. It prevents us from reducing the variance of the coherence estimate through frequency-aspect averaging, but we can still use spatial averaging. Another advantage of the single image pair is that it results in the smallest amount of resolution degradation using multilook.

4.2 Coherence evaluation window

We consider high or medium estimates of coherence (between looks) as indicators of point scatterers or structures (focused or defocused) within the window of evaluation. For estimation of coherence, we would like the window to be as large as possible in order to keep the random high value estimates as rare as possible, and thereby potentially increase the contrast between incoherent background and defocused scatterers. The window must be kept small enough that only one prominent scatterer is included, and also in order to keep the coherence of defocused point scatterers prominent over the impact of the background. This would involve an optimization for

further research. It might also be useful if the coherence metric could indicate the location of candidate point scatterers. We chose a dynamic window size of $9\tau_j + 1_{j>1}$ in each dimension by taking into account the different scales the wavelet coefficients are operating at. The term $1_{j>1} = \begin{cases} 1, & j > 1 \\ 0, & j = 1 \end{cases}$ ensures odd window sizes for all scales. Given a pixel size of 2×2 cm, this corresponds to window sizes of 18×18 cm for $j = 1$, 38×38 cm for $j = 2$ and 74×74 cm for $j = 3$.

4.3 Wavelets

We choose the Daubechies class wavelet with unit-level length 4, D(4), and use $j = 3$ levels for the MODWT. There is a trade-off in length of the wavelet filter and finite sample sizes. The Haar wavelet is shorter and might not capture all details. Longer filters may give better results, but boundary effects are not negligible for the image sizes we are working on. For $j = 3$ the 1-D wavelet filter length L_3 is 8 for the Haar wavelet, 22 for the D(4) wavelet and 50 for the D(8) wavelet. We are using a pyramid algorithm[9] for performing the wavelet transform and its inverse.

Thorough consideration needs to be given in how to choose the thresholds on the coherence between the corresponding wavelet coefficients of the two image looks. We have experimented both with dynamic thresholds based on quantiles, and fixed thresholds, either the same thresholds for all coefficients or specified for a given level and orientation. It is an area of future work to determine the optimal use of coefficients. For simplicity, the results in this paper are calculated with thresholds $\rho_{min} = 0.5$ and $\rho_{max} = 0.8$ following a weight calculating scheme[6]; i.e. the wavelet shrinkage coefficients are

$$\zeta(\rho_{C_j}(x,y)) = \begin{cases} 0, & \rho_{C_j}(x,y) < \rho_{min} \\ 1, & \rho_{C_j}(x,y) > \rho_{max} \\ \frac{\rho_{C_j}(x,y) - \rho_{min}}{\rho_{max} - \rho_{min}}, & \text{otherwise} \end{cases}.$$

5 RESULTS

5.1 Simulated data

In order to test and quantify our wavelet shrinkage method, we have simulated a number of SAS scenes. Our approach is a point-based raw data simulator[14]. The number of points are high per theoretical resolution cell to ensure that correct speckle properties are obtained by the simulation. In addition, we have added texture through random spatial variability following a power law. The essential features such as element directivity, range dependence and frequency dependence are taken into account. The sonar setup is similar to the system used in the real data results in the next section. In addition to textured speckle, we have added a number of point scatterers in a grid in the scene. These serve as coherent point targets for being evaluated by our technique. Finally, we run our simulated raw data through the SAS processor using true navigation and scene geometry. We have made two sets of images per simulation, one where we use the correct sound speed, and one where we deliberately use incorrect sound speed in the SAS imaging such that true defocusing occurs.

Fig. 1 shows the SAS image of the simulated scene (left) with the corresponding wavelet shrinked image (right). The point scatterers are laid out in a 5×6 grid, at every 10 m along-track and across-track. Our point scatterer detector picks the strongest targets in the shrinked image. The Top 100 candidates are indicated with circles, where red symbolizes those that are ranked as the best candidates followed by yellow and green. The ranking is based on the maximum per 21×21 pixels region of interest in the enhanced image and does not necessarily give the most optimal result. Other possible ranking criteria[5] might be interesting to investigate.

We see that 24 of the 30 point scatterers are detected among the 30 strongest targets (red circles) in the shrinked image in Fig. 1. Only two point scatterers are not among the 100 strongest. In the textured region with large signal intensity (∼10–30 m across-track), there are many extra detections.

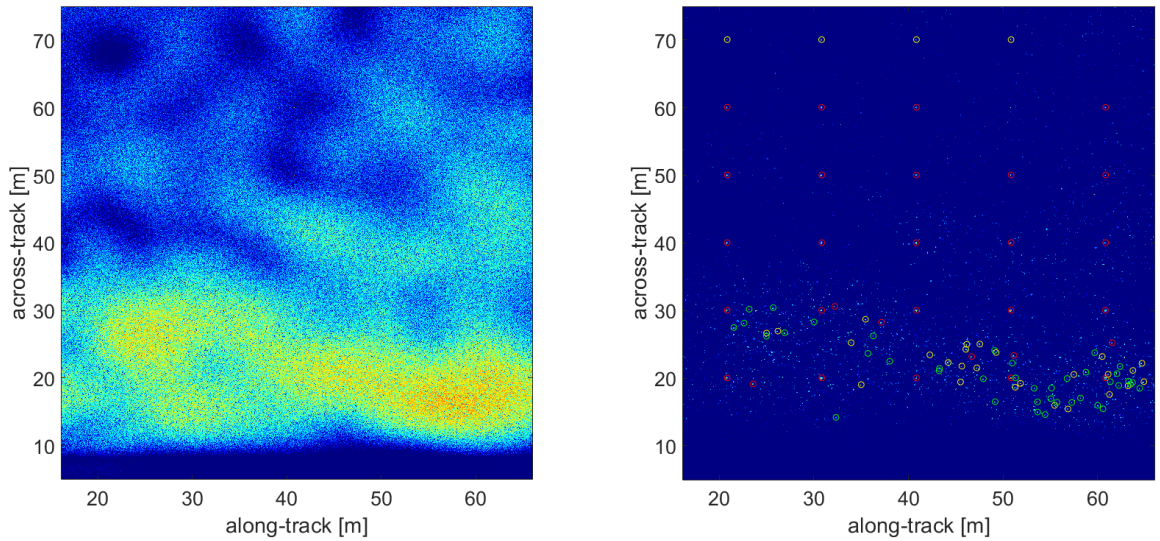

Figure 1: *Left: Simulated image of size 50×70 m with 30 point scatterers, speckle and 3-30 m power law texture. Right: Enhanced image after wavelet shrinkage with Top 100 detections (red, yellow and green circles). The dynamic range in both images is 50 dB.*

Detections in Fig. 1	Enhanced image		Original image	
	Along-track	Across-track	Along-track	Across-track
Red	4.02	4.61	3.08	3.62
Red & yellow	4.12	4.66	3.59	4.21
Red, yellow & green	4.29	4.63	3.63	4.70

Table 1: *Estimated resolution [cm] in enhanced and original image for detected point scatterers from Fig. 1.*

The theoretical resolution for the simulated image from Fig. 1 is around 3.5×3.5 cm. Table 1 contains the estimated resolution for that image using the Top 30 detections (red circles), Top 60 detections marked red and yellow as well as all Top 100 detections. Because 6 of the 30 simulated point scatterers are not among the 30 strongest targets, the estimated resolution is biased. For the Top 30 detections, the along-track resolution is 3.1 cm, while the across-track resolution is 3.6 cm for the case of performing the estimation on the detected locations in the original SLC image. When including more detections, i.e. reducing the ratio of true point scatterers to other scatterers, the estimated resolution deteriorates.

An interesting observation is that the resolution estimated directly from the enhanced image is generally worse than using the original image (except for across-track resolution when using all detections). The transformation after wavelet shrinkage can alter the point spread function. We recommend to estimate resolution from the original image with the detections obtained from the enhanced image.

Fig. 2 shows the results where we have degraded the previous image using 40 m/s sound speed error. For this defocused case only 14 of 30 point scatterers are detected among the 30 strongest. None of the reflectors at across-track position 60 m and 70 m are detected, although some of them are clearly visible in the shrinked image. This is due to the defocusing (and corresponding decrease in peak intensity) of the reflectors that increase with range. Note that there is a mispositioning of the reflectors due to the incorrect sound speed[15].

Table 2 shows that estimated resolution is much worse for the defocused case, especially along-track. This is because sound speed errors occur along-track only. The slight deterioration in range resolution might be explained with the fact that we did not manage to correctly detect all the true point scatterers.

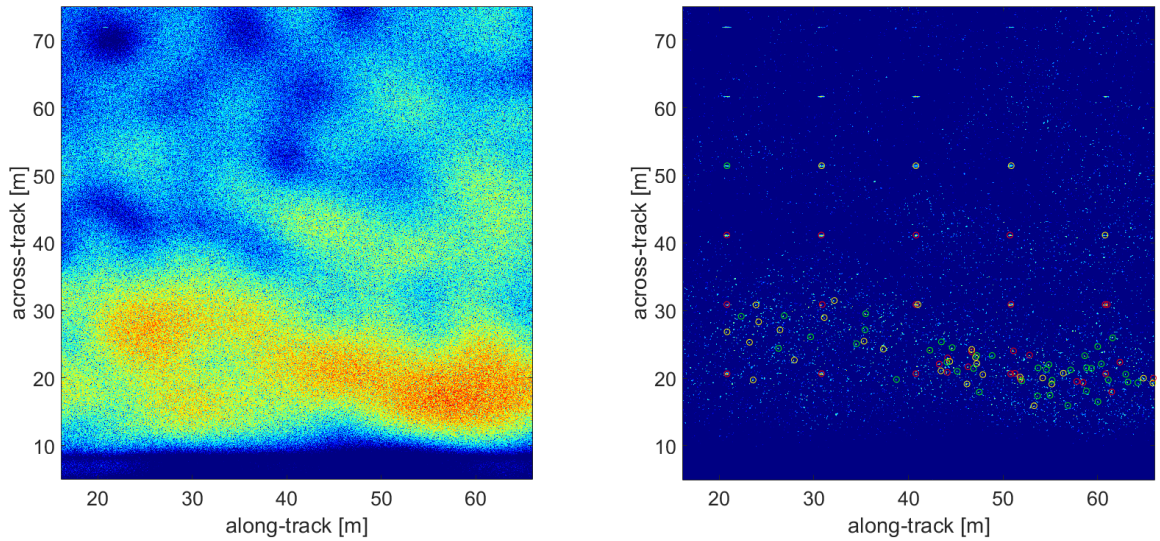

Figure 2: *Left: Simulated image as in Fig. 1, but degraded with 40 m/s sound speed error. Right: Enhanced image after wavelet shrinkage with Top 100 detections (red, yellow and green circles). The dynamic range in both images is 50 dB.*

Detections in Fig. 2	Enhanced image		Original image	
	Along-track	Across-track	Along-track	Across-track
Red	8.84	5.18	10.82	4.97
Red & yellow	7.83	4.83	9.70	5.01
Red, yellow & green	6.81	4.66	7.98	5.05

Table 2: *Estimated resolution [cm] in enhanced and original image for detected point scatterers from Fig. 2.*

5.2 Real data

We have applied the suggested method to real SAS data from FFI's HUGIN-HUS AUV, equipped with the HISAS 1032 sensor with 100 kHz carrier frequency and 30 kHz bandwidth. Fig. 3 (left) shows a SAS image from an area outside Bergen, Norway, collected in 2017. The texture in the image contains larger rock outcrops and small-scale structure from sedimented areas. The theoretical resolution is around 3.5×3.5 cm. In Fig. 3 (right), the corresponding enhanced image after wavelet shrinkage is displayed. It appears that the strongest scatterers in the enhanced image are concentrated around the rocky areas, hence many detections there. The enhanced image shows candidates throughout the whole image, so a different ranking of the detections might give a different result as already mentioned in the simulated data case.

Table 3 lists the estimated resolution for our real data example, both directly estimated in the enhanced image and also in the original image using the location of the detections from the enhanced image. The estimated along-track resolution is between 6.2–6.3 cm for all cases (also when estimating directly from the enhanced image), which indicates some defocus to the image but not to the same extent than in the case of the degraded simulated image from Fig. 2. The range resolution is between 3.4–3.5 cm when using the location of the detected scatterers on the original image, which indicates close to theoretical resolution. This is expected as defocus generally occurs along-track. Obtaining near theoretical resolution in the one direction that is not impacted by blurring is a good check that the detected scatterers are suitable candidates for resolution estimation. The range resolution estimated directly from the enhanced image is between 4.5–4.6 cm, i.e. more than a cm worse than using the point spread function from the original SLC image. Also for real data, it is more reliable to estimate resolution from the SLC image given the locations of the scatterers detected in the enhanced image after wavelet shrinkage.

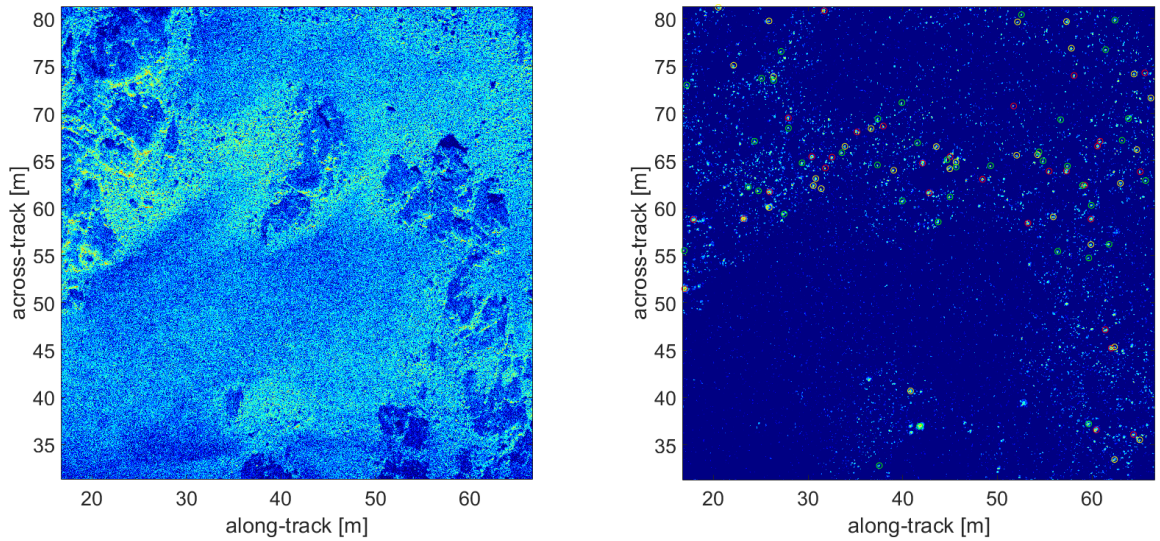

Figure 3: *Left: SAS image of size 50×50 m recorded off the coast of Bergen, Norway, with HISAS 1032 onboard FFI's HUGIN-HUS AUV. Right: Enhanced image after wavelet shrinkage with Top 100 detections (red, yellow and green circles). The dynamic range in both images is 50 dB.*

Detections in Fig. 3	Enhanced image		Original image	
	Along-track	Across-track	Along-track	Across-track
Red	6.23	4.62	6.21	3.37
Red & yellow	6.23	4.53	6.17	3.41
Red, yellow & green	6.34	4.52	6.33	3.47

Table 3: *Estimated resolution [cm] in enhanced and original image for detected point scatterers from Fig. 3.*

6 CONCLUSION

The proposed wavelet shrinkage procedure leads to an enhancement of point scatterers by reducing incoherent signal in SAS images. This transformation facilitates detection of candidate point scatterers. We recommend to first use the enhanced image for obtaining suitable candidate point scatterers and then estimate the resolution at the detected locations in the original SAS image. This paper does not fully explore the possibilities of the presented approach as we only have demonstrated wavelet shrinking with fixed coherence thresholds across all scales and orientations. The optimal use of the different parameters is a topic for further research. Defocus generally occurs along-track, hence the shrinkage weights should be tuned more towards the wavelet coefficients that capture this orientation. The optimal ranking method in the detection step is also subject for future work. Finally, it is of interest to compare different methods for detecting point scatterers with a suitable metric, especially those that are also based on multilook images[5].

7 ACKNOWLEDGMENTS

The authors would like to thank Torstein O. Sæbø and Ole J. Lorentzen at FFI for fruitful discussions and inputs during the preparation of this paper. Further, the authors thank FFI's HUGIN-HUS operator group for collecting the SAS data used in this study.

8 REFERENCES

1. Ø. Midtgaard, I. Alm, T. O. Sæbø, M. Geilhufe, and R. E. Hansen, "Performance Assessment Tool for AUV based Mine Hunting," in *SAS & SAR conference 2014, Proc. IOA*, vol. 36 Pt. 1, Lerici, Italy, September 2014, pp. 191–200.

2. J. L. Prater, J. L. King, and D. C. Brown, "Determination of image resolution from SAS image statistics," in *Proceedings of Oceans 2015 MTS/IEEE*, Washington DC, USA, Oct. 2015.

3. M. Geilhufe, R. E. Hansen, Ø. Midtgaard, and S. A. V. Synnes, "Through-the-sensor sharpness estimation for synthetic aperture sonar images," in *Proceedings of Oceans 2019 MTS/IEEE*, Seattle, WA, USA, October 2019.

4. D. J. Pate, D. A. Cook, and B. N. O'Donnell, "Estimation of synthetic aperture resolution by measuring point scatterer responses," *IEEE Journal of Oceanic Engineering*, vol. 47, no. 2, pp. 457–471, 2021.

5. R. E. Hansen, M. Geilhufe, S. A. V. Synnes, T. O. Saebo, and S. H. Thon, "Detection of coherent scatterers in synthetic aperture sonar using multilook coherence," in *EUSAR 2022; 14th European Conference on Synthetic Aperture Radar*. VDE, 2022, pp. 1–6.

6. A. J. Hunter and R. van Vossen, "Sonar target enhancement by shrinkage of incoherent wavelet coefficients," *The Journal of the Acoustical Society of America*, vol. 135, no. 1, pp. 262–268, 2014.

7. D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*. New York, USA: Cambridge University Press, 2000.

8. I. Daubechies, *Ten Lectures on Wavelets*, ser. Cbms-Nsf Regional Conference Series in Applied Mathematics Series. Society for Industrial and Applied Mathematics (SIAM), 1992.

9. M. Geilhufe, D. B. Percival, and H. L. Stern, "Two-dimensional wavelet variance estimation with application to sea ice SAR images," *Computers & Geosciences*, vol. 54, pp. 351–360, 2013.

10. D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

11. A. Borsdorf, R. Raupach, T. Flohr, and J. Hornegger, "Wavelet based noise reduction in CT-images using correlation analysis," *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1685–1703, 2008.

12. G. Carter, C. Knapp, and A. Nuttall, "Statistics of the estimate of the magnitude-coherence function," *IEEE Transactions on Audio and Electroacoustics*, vol. 21, no. 4, pp. 388–389, 1973.

13. J. C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

14. D. C. Brown, S. F. Johnson, and D. R. Olson, "A point-based scattering model for the incoherent component of the scattered field," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. EL210–EL215, 03 2017. [Online]. Available: https://doi.org/10.1121/1.4976584

15. R. E. Hansen, H. J. Callow, T. O. Sæbø, and S. A. Synnes, "Challenges in seafloor imaging and mapping with synthetic aperture sonar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3677–3687, 2011.