SEQUENTIAL FOCUS EVALUATION OF SYNTHETIC APERTURE SONAR IMAGES

Stefan Leier and Abdelhak M. Zoubir

Signal Processing Group
Technische Universität Darmstadt
Merckstr. 25, 64283 Darmstadt, Germany
{leier, zoubir}@spg.tu-darmstadt.de

Johannes Groen

ATLAS ELEKTRONIK GmbH
Sebaldsbruecker Heerstrasse 235,
28309 Bremen, Germany
johannes.groen@atlas-elektronik.com

ABSTRACT

A synthetic aperture sonar (SAS) system borne by an autonomous underwater vehicle is state-of-the-art for high-resolution sea floor mapping. In the application of automatic target recognition, e.g., for naval mine hunting, the high-resolution SAS images serve as input to a detection and classification post-processing stage, which highly relies on excellent image quality. Future autonomous mine hunting systems must include an assessment scheme to ensure sufficient quality for performing target recognition. We propose to assess the focusing capability during the reconstruction of an SAS image to evaluate its quality by probing the instantaneous cross-range resolution of a synthetic sub-aperture and comparing it with its theoretical resolution.

Index Terms— synthetic aperture sonar, image quality metrics, focus assessment, automatic target recognition

1. INTRODUCTION

The objective of an automatic target recognition (ATR) system is to detect candidate objects and classify them according to pre-defined object types. Subsequently, type information and exact object location, e.g., of mines in naval mine countermeasures, are transmitted to an operation base, which decides how to proceed for the disposal. An ATR system in naval mine hunting [1–3] comprises of two stages, namely image reconstruction and automatic detection and classification. In the case of an operator-based system, it is feasible to intervene after the image reconstruction process to assess the image quality and thus, guarantee a reliable input for an automatic detection and classification system. However, in a fully automatic system, the omission of an operator necessitates an assessment to still guarantee that high-quality images are used. Especially, most existing ATR systems assume that the image quality is uniformly excellent for the entire sonar image, which is often not valid in practice [4]. Thus, knowledge about the image quality may help to overcome the degradation in classification rates [5] due to assumption violation.

State-of-the-art ATR systems use a high-resolution synthetic aperture sonar (SAS) borne on an autonomous underwater vehicle (AUV) to produce images of the seabed [1]. Generally, a synthetic aperture system synthesizes a physical aperture by constantly transmitting pulses while moving, ideally, with a constant velocity in cross-range direction. While conventional side-scan sonar images have the drawback of a range-dependent cross-range resolution due to a range proportional beamwidth, this is inherently exploited in SAS imaging by dynamically adjusting the synthetic aperture length to the corresponding range [6]. Thus, neglecting motion and path propagation effects, theoretically, a constant cross-range resolution is maintained for the entire scene of interest, which facilitates the reconstruction of high-resolution images.

A key parameter of any practical mapping system is its area coverage rate. In contrast to the speed of electromagnetic waves in air, the limited speed of sound in water requires an SAS system to consist of an array of hydrophones to achieve reasonable coverage rates [6]. Simultaneously, it provides the capability of constructing low resolution images, which can be exploited to successively evaluate the cross-range focusing of the synthetic aperture and therefore the quality of the resulting images [7]. Besides assessing the reliability of the input for automatic detection and classification, image quality evaluation is also important for adaptive mission planing [8]. Since the quality of SAS images degrades especially at long ranges due to more stringent demands on the motion estimation accuracy or due to multipath effects, information about the image quality has to be considered for *in situ* track planning of the AUV. Thus, guaranteeing sufficiently good image data for the entire mission area.

The contribution of this paper is the development of a technique for evaluating the focusing capabilities of a multiple-receiver SAS system to assess the obtained quality of the resulting image. The applied metric is based on estimating the instantaneous cross-range resolution and relating it to its theoretical resolution. Significant deviations from the theoretical resolution will be considered as an occurrence of defocus and a worsening in image quality. The evaluation over the entire synthetic aperture construction provides a more reliable assessment of image quality than only taking the final SAS image into account as it is often done

for synthetic aperture radar (SAR) images [9]. The proposed approach directly operates in the image domain. This is in contrast to [1,8], who relate image quality to the peak correlation between successive pings and therefore, work in the raw data domain. Moreover, the proposed approach facilitates an evaluation based on the theoretical resolution, while this is not feasible with the energy ratio proposed in [7].

The remainder of the paper is organized as follows: In Section 2, we introduce the echo signal model for a multiple-receiver SAS system. Section 3 provides a brief description of the image reconstruction process. In Section 4, a phase error model is introduced to deteriorate the focusing. Then, Section 5 describes the proposed focus assessment approach, containing a real data example. Finally, Section 6 provides conclusions.

2. DATA MODEL

Consider a set of D stationary point scatterers located at positions $\mathbf{v}_d = [x_d, y_d, 0]^T$, $d = 1, ..., D$, and reflectivity coefficient σ_d, which is assumed to be independent of frequency and angle of incidence but incorporates spreading losses. Then, the ideal reflectivity function [10] of the scene of interest is

$$f(x,y) = \sum_{d=1}^{D} \sigma_d \delta(x - x_d, y - y_d), \tag{1}$$

where $\delta(x,y)$ is the two-dimensional delta function and x and y denote range and cross-range, respectively. Given a single transmitter and multiple-receiver system consisting of L_u hydrophones, the echo signals $e_p(u,t)$ at ping p can be expressed as

$$e_p(u,t) = \sum_{d=1}^{D} \sigma_d s\left(t - \tau_{p,d}(u)\right), \tag{2}$$

where $u = 1, \ldots, L_u$ is the receiver element index and t denotes fast-time. The echo signals in (2) are a superposition of individually delayed versions of the transmitted pulse $s(t)$ weighted by the respective target reflectivity σ_d. The round-trip delay $\tau_{p,d}(u)$ between the d^{th} scatterer and the imaging platform at ping p is given by

$$\tau_{p,d}(u) = \left(\|\mathbf{v}_d - \mathbf{t}_p\|_2 + \| \mathbf{r}_p(u) - \mathbf{v}_d\|_2 \right)/c, \tag{3}$$

where $\|\cdot\|_2$ denotes the Euclidean distance, and $\mathbf{t}_p$ and $\mathbf{r}_p(u)$ are the transmitter and the u^{th} receiver position, respectively, and c is the speed of sound in water.

3. IMAGE RECONSTRUCTION

This section briefly addresses the image reconstruction stage, which represents the focusing process of the collected raw data of (2) to obtain an intensity representation of the backscattered acoustic energy of the scene of interest. Let $\mathbf{g}_{kl} = [x_k, y_l, 0]^T$ with $k = 1, \ldots, N_x$ and $l = 1, \ldots, N_y$ describe all grid points of the scene of interest. Then the focusing delay $\tau^{\text{foc}}_{p,kl}(u)$ at ping p between the multi-receiver imaging platform and the grid point $\mathbf{g}_{kl}$ is given by

$$\tau^{\text{foc}}_{p,kl}(u) = \left(\|\mathbf{g}_{kl} - \mathbf{t}_p\|_2 + \|\mathbf{r}_p(u) - \mathbf{g}_{kl}\|_2 \right)/c. \tag{4}$$

Then, the corresponding real aperture sonar (RAS) image can be reconstructed by

$$\hat{f}_p(x_k, y_l) = \sum_{u=1}^{L_u} e_p^{\text{MF}}\left(u, \tau^{\text{foc}}_{p,kl}(u)\right) e^{j\omega_c \tau^{\text{foc}}_{p,kl}(u)} \tag{5}$$
$$\text{for } \; k = 1, \ldots, N_x \;\text{ and }\; l = 1, \ldots, N_y,$$

where ω_c is the carrier frequency of the transmitted signal and $e_p^{\text{MF}}(u,t)$ denotes the pulse-compressed echo signals. The synthetic aperture is then reconstructed using the coherent combination of $N_p = P_1 - P_0 + 1$ single ping images where P_0 and P_1 indicate the ping indices for which the grid point $\mathbf{g}_{kl}$ is "seen" for the first and last time by the imaging platform, respectively. Consequently, the image reconstruction to obtain an SAS image is described by

$$\hat{f}(x_k, y_l) = \sum_{p=P_0}^{P_1} \hat{f}_p(x_k, y_l) \cdot B(\mathbf{g}_{kl}, \mathbf{t}_p, \mathbf{r}_p(u)) \tag{6}$$
$$\text{for } \; k = 1, \ldots, N_x \text{ and } l = 1, \ldots, N_y.$$

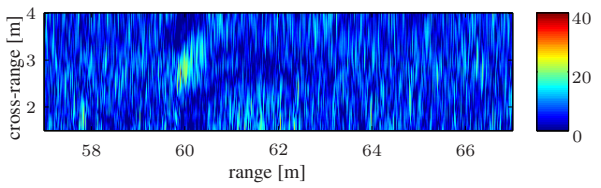
In (6), $B(\cdot)$ is an indicator function, which determines whether a certain grid point $\mathbf{g}_{kl}$ is "seen" by the sonar at the respective ping. It is defined as

$$B(\mathbf{g}_{kl}, \mathbf{t}_p, \mathbf{r}_p(u)) = \begin{cases} 1, & \max\left(\theta^{\text{Rx}}_{p,kl}(u), \theta^{\text{Tx}}_{p,kl}\right) \leq \theta_0 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

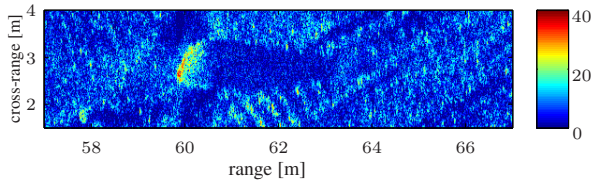
where θ_0 is the minimum physical beamwidth of the transmitter and the receiving elements. Moreover, $\theta^{\text{Tx}}_{p,kl}$ and $\theta^{\text{Rx}}_{p,kl}(u)$ denote the aspect angles between grid points and transmitter as well as receiving elements. An example of an RAS and an SAS image is depicted in Fig.1(a) and Fig.1(b), respectively. Both images are displayed with a dynamic range of 40 dB. The SAS image shows a significant enhancement in the richness of detail due to an enhanced resolution, an improved shadow contour and overall contrast.

4. MOTION ERROR MODEL

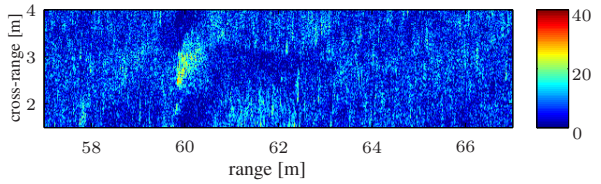
Motion errors of the imaging platform as well as phase errors due to an incorrect knowledge of the speed of sound have a dramatic impact on the SAS image quality [1], since a coherent processing of consecutive pings is no longer feasible. While motion errors result from an inaccurate position tracking by the on-board inertial navigation system (INS), a mismatch in the speed of sound arises due to the fact that water



(a) RAS image



(b) Focused SAS Image



(c) Blurred SAS Image

Fig. 1: Real aperture sonar (a) and synthetic aperture sonar (b) image comparison of a mine-like object. A significant detail gain is apparent for the SAS image. The SAS image in (c) is blurred due to uncompensated motion errors. The images have been constructed using real sonar data provided by ATLAS ELEKTRONIK GmbH.

is an inhomogeneous medium. Both error types lead to a difference between the round-trip delay of the raw data, as given by (3), and the focusing delay of the reconstruction process, as in (4). This yields different phase shifts at the pixels in the complex RAS images. Therefore, the coherent summation of individual RAS images in (6) causes a blurring effect in the overall SAS image, as demonstrated in Fig. 1(c). In the sequel, we only consider errors due to path deviations. In order to model the phase changes between consecutive RAS images, we introduce the phase function

$$\phi_{\text{err}}(p) = \frac{2\pi}{\lambda_c} \cdot \Delta R_{\text{err}}(p) \quad \text{with} \quad p = P_0, \ldots, P_1 \tag{8}$$

where $\Delta R_{\text{err}}(p)$ is a path deviation function in range direction depending on ping p. We model the path deviation as a sinusoidal function due to its frequent occurrence in practice [11] as

$$\Delta R_{\text{err}}(p) = K\lambda_c \sin(2\pi f_p p), \tag{9}$$

where K is a constant representing the error strength in percent of the carrier wavelength λ_c and f_p denotes the cycles per synthetic aperture length frequency. Adapting (6) leads to the motion error affected SAS image as

$$\hat{f}_{\text{err}}(x_k, y_l) = \sum_{p=P_0}^{P_1} \hat{f}_p(x_k, y_l) B(\mathbf{g}_{kl}, \mathbf{t}_p, \mathbf{r}_p(u)) e^{-j\phi_{\text{err}}(p)} \tag{10}$$

The blurred SAS image of Fig. 1(c) has been reconstructed with error parameters $K = 0.5$ and $f_p = 1$. The defocusing causes a loss in highlight as well as background detail, e.g., the small object at range 58 m and the shadow contour of the mine-like object start to vanish.

5. FOCUS ASSESSMENT

Since defocusing occurs successively when coherently combining single RAS images, the proposed approach aims at sequentially assessing the focusing capability by estimating the resolution from the scene of interest for consecutive pings. In the following, we introduce an expression for the instantaneous cross-range resolution.

5.1. Theoretical Resolution Aspects

The half-power beamwidth of a uniform linear array of length L_{phy} is $\theta_{3\text{dB}} = 0.89\lambda/L_{\text{phy}}$ [12], which leads to a range dependent cross-range resolution given by

$$\delta_{\text{phy}}(r) = \theta_{3\text{dB}} \cdot r = 0.89\lambda r / L_{\text{phy}}. \tag{11}$$

The advantage of synthetic aperture processing is its aperture length adaption with respect to the focusing range. Given the advance per ping of the physical array Δ_A, the length of the synthetic aperture can be expressed as

$$L_{\text{syn}}(p, r) = \min\left[\Delta_A(p-1) + L_{\text{phy}},\ L_{\text{syn,max}}(r)\right], \tag{12}$$

where $L_{\text{syn,max}}(r) = r\theta_0$ describes the maximum synthetic aperture length as a function of the focus range and θ_0 denotes the minimum beamwidth between transmitter and receiving elements. Replacing the physical array length L_{phy} in (11) by (12) leads to the instantaneous cross-range resolution of the synthetic aperture as

$$\delta_{\text{syn}}(p, r) = \frac{0.89\lambda \cdot r}{\min\left[\Delta_A(p-1) + L_{\text{phy}},\ L_{\text{syn,max}}(r)\right]}. \tag{13}$$

Note that the well known range independent cross-range resolution [10] for synthetic aperture systems of $\delta_{\text{SAS}} = D_{\text{phy}}/2$ is achieved given the maximum synthetic aperture length $L_{\text{syn,max}}(r)$ and a physical aperture extend of the individual receiver of $D_{\text{phy}} = 0.89\lambda/\theta_0$.

5.2. Assessment Scheme

Given the image reconstruction of (6) or (10) in case of phase errors as well as the instantaneous cross-range resolution

as stated in (13), the quality assessment of the SAS image $\hat{f}(x_k, y_l)$ for $k = 1, \ldots, N_x$ and $l = 1, \ldots, N_y$ is achieved by sequentially comparing the estimated and theoretical resolution of a point scatterer. We assume the presence of an isolated point scatterer in the first RAS image and start to estimate its 3 dB width $\hat{\delta}_{\text{syn}}(P_0, r_0)$. Then, we coherently add the consecutive RAS image and repeat this process until the maximum number of pings N_p is attained, for which the extract scatterer is still "seen" by the imaging platform. An example of a changing point spread function (PSF) of an isolated point scatterer in cross-range direction is depicted in Fig. 2 for an increasing synthetic aperture length. The focusing effect is clearly observable by a decreasing mainlobe width. After the

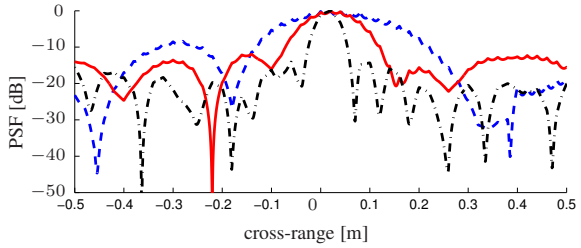


Fig. 2: Changing point spread function (PSF) for an increasing synthetic aperture length of a point scatterer (blue dashed $N_p = 2$, red solid $N_p = 4$, black dash-dotted $N_p = 10$).

estimation of $\hat{\delta}_{\text{syn}}(p, r_0)$, $p = P_0, \ldots, P_1$, we approximate the cross-range resolution of (13) by an exponential model

$$\delta_{\text{mod}}(p) = \alpha_1 \exp\{-\alpha_2 \cdot p\} + \alpha_3, \tag{14}$$

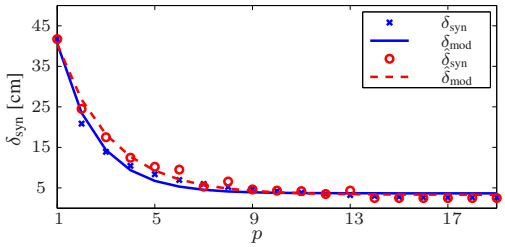
where α_3 describes the convergence parameter approximating δ_{SAS}. In order to estimate the parameters $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]^T$, we minimize the cost function

$$C(p, \boldsymbol{\alpha}) = \frac{1}{2} \left| \hat{\delta}_{\text{syn}}(p, r_0) - \delta_{\text{mod}}(p; \boldsymbol{\alpha}) \right|^2 \tag{15}$$

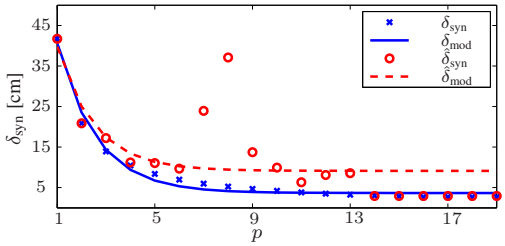
using the Newton-Raphson method.

5.3. Real data results

A real data example of the evolution of the cross-range resolution is depicted in Fig. 3 for perfect focusing and defocusing in the presence of uncompensated motion errors. A comparison between theoretical (blue cross) and estimated (red circle) cross-range resolution is depicted in Fig. 3(a) together with both estimated model functions. The estimated resolution converges against the theoretical SAS cross-range resolution given by $\delta_{\text{SAS}} = 0.0264$ m. Additionally, the fit of the model function shows a satisfying agreement in a motion error free scenario. The purpose of fitting the model function is to assess the focusing capability and therefore the image quality of the SAS image while synthesizing the aperture. This is done by comparing the estimated, $\hat{\alpha}_3$, to the theoretical, α_3, convergence parameter, which is illustrated by an example in Fig. 3(b). In order to defocus the SAS image, the amplitude of the path deviation function of (9) has been set to $K = 0.3$. The resolution estimate is $\hat{\delta}_{\text{SAS}} \approx \hat{\alpha}_3 = 0.092$ m, which is roughly two and a half times larger than the theoretical resolution given by $\delta_{\text{SAS}} \approx \alpha_3 = 0.037$ m indicating a defocus in the SAS image. Note that the convergence parameter α_3 does not equal the theoretical SAS resolution δ_{SAS}. However, its purpose is only to assess the focusing capability. The importance of fitting a model as well as considering the entire resolution history along the construction of the synthetic aperture becomes apparent for the estimated resolution values starting with ping $p \geq 14$, where the deterioration of the point spread function suggests a perfectly focused SAS image.



(a) Metric evaluation in case of perfect focusing.



(b) Metric evaluation in case of defocus due to motion errors.

Fig. 3: Theoretical and estimated metric evaluation over an increasing synthetic aperture SAS in case of perfect focusing (a) and defocusing (b) due to injected motion errors. Fitted curves are depicted as solid and dashed lines.

6. CONCLUSION

We have introduced a sequential approach to assess synthetic aperture focusing and consequently SAS image quality based on evaluating the instantaneous cross-range resolution. The advantage of a successive evaluation is twofold: First, it allows to exploit the focusing history to judge on the quality rather than assessing only the final SAS image. The latter may suggest a wrong interpretation of the focusing success. Second, the number of pings and therefore a synthetic subaperture length can be determined to construct an image with possibly a lower resolution but without degradations due to motion errors. In the near future, the proposed approach will be embedded in a real-time framework to test it more extensively.

7. REFERENCES

[1] R. E. Hansen, H. J. Callow, T. O. Sabo, and S. A. V. 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.

[2] 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, 2003.

[3] R. Fandos and A. M. Zoubir, "Optimal Feature Set for Automatic Detection and Classification of Underwater Objects in SAS Images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 454–468, 2011.

[4] D. P. Williams and J. Groen, "A fast physics-based, environmentally adaptive underwater object detection algorithm," in *Proc. IEEE Oceans Conf.*, 2011, pp. 1–7.

[5] S. Leier, R. Fandos, and A. M. Zoubir, "Motion Error Influence on Segmentation and Classification Performance in SAS based Automatic Mine Countermeasures," *IEEE Journal of Oceanic Engineering*, 2012, under review.

[6] M. P. Hayes and P. T. Gough, "Synthetic Aperture Sonar: A Review of Current Status," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 207–224, 2009.

[7] S. Leier and A. M. Zoubir, "Quality assessment of synthetic aperture sonar images based on a single ping reference," in *Proc. IEEE Oceans Conf.*, 2011, pp. 1–4.

[8] D. Williams, A. Vermeij, F. Baralli, J. Groen, and F. Warren, "In Situ AUV Survey Adaptation Using Through-the-Sensor Sonar Data," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2012.

[9] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: systems and signal processing*, Wiley & Sons, 1991.

[10] M. Soumekh, *Synthetic Aperture Radar Signal Processing: with MATLAB Algorithms*, Wiley & Sons, 1999.

[11] G. Fornaro, "Trajectory deviations in airborne SAR: analysis and compensation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 997–1009, 1999.

[12] H. L. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, Wiley & Sons, 2002.