Autofocus for Circular Synthetic Aperture Imaging

Hayden J Callow, Stig A V Synnes, Torstein Olsmo Sæbø and Roy Edgar Hansen

***Abstract*—Fully coherent circular synthetic aperture (CSA) imaging is challenging. In particular, motion-estimation, sound-speed, and imaging-topography estimates need to be orders-of-magnitude better than for traditional straight-line paths. Such accuracy is required, that data-driven techniques are essential. Current data-driven methods do not work on circular tracks.**

Both topography errors and track errors cause CSA image defects. For autofocus this presents a highly coupled estimation problem—it is difficult to separate topography from motion. To get useful autofocus performance, topography must be accounted for in the autofocus formulation. Traditional autofocus cannot make use of this information.

In order to obtain the benefit of fully coherent CSA, we present here an autofocus algorithm based on maximum a posteriori (MAP) methods. Use of MAP allows regularization of the optimisation with prior information. We chose this framework with the intent that future versions of the algorithm could use echo-strength information to mitigate the effects of multiple reflections.

We present preliminary results from the autofocus algorithm on data collected in the field with a HISAS-1030 SAS. We discuss the difficulties involved in "real-life" autofocus, in particular we discuss some of the limits of CSA imaging. In addition, we discuss the similarities between motion-estimation and interferometry.

***Index Terms*—Synthetic aperture radar, synthetic aperture sonar, interferometry, autofocus, maximum a posteriori estimation, MAP, autofocus, bathymetry, topography, phase gradient, SLAM, simultaneous localization and mapping**

I. Introduction

CIRCULAR synthetic aperture imaging, multi-pass object classification and multi-pass coherent fusion where the target is illuminated from different directions are all variations on the same theme. All methods promise the extraction of additional target information. In particular:

- shadow regions are illuminated
- full angular wavenumber spectrum is collected
- target roughness estimates may be possible

This information is likely to be useful in sonar imaging applications where very high detail levels are required. In particular it is hoped that the additional information will improve target classification [1].

However, navigation constraints, already challenging in synthetic aperture sonar (SAS), become even tighter in circular synthetic aperture imaging (CSA) [2]. In addition, the non-linear collection geometry requires very accurate topographic information for sensible image formation—down to centimeter level accuracies.[1] In this paper, we simplify the problem somewhat by looking into linear subsets of circular aperture data. This partially decouples the imaging defocus problem, but for generating a full image the bathymetric estimation and relative constraints are unaltered.

All authors are with the Norwegian Defence Research Establishment, P. O. Box 25, NO-2027 Kjeller, Norway. E-mail: hayden-john.callow@ffi.no

[1]So accurate, in fact, that unavoidable layover may make traditional imaging infeasible (see section III-B).

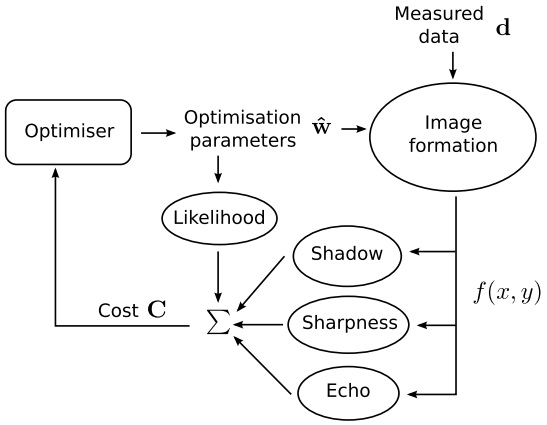

Fig. 1. Overview of the iterative framework. Each time around the iteration, a new image, $f(x, y)$ and set of image likelihoods is generated from the observed SAS data $\mathbf{d}$, and the unknown parameter estimates, $\hat{\mathbf{w}}$ are updated. The optimiser aims to maximise a combined cost function, $\mathbf{C}$.

We address this problem by proposing an autofocus scheme for simultaneously estimating *both* unknown platform motion and scene topography. By allowing the autofocus system to estimate both position correction and topography, the major unknown causes of image defocus are estimated and compensated for.

To implement this autofocus, we use an iterative maximum a posteriori (MAP) estimation. The main benefit of using a MAP structure is that we may explore different information sources to aid the autofocus. In particular, this method allows use of:

- speckle interferometry
- shape from shading
- shadow mapping

In addition, combinations of the above should be possible, further strengthening autofocus performance. The current disadvantage is a computationally heavy iteration strategy.

II. MAP autofocus

To estimate the scene topography and unknown path-motion in one optimisation we use a MAP estimator in an iterative framework. In general, there is little separating this method from the framework presented in [3, Chapter 5] other than that scene topography is treated as an unknown parameter in this work.

MAP estimators estimate the unknown parameters corresponding to the maximum likelihood image *with prior information*. This prior information regularizes the optimisation problem and gives a set of constraints on the optimisation surface. In addition, one may then use the prior information

as a basis set for the optimisation—greatly accelerating the process.

We define the optimisation problem as in [3, Chapter 5]. We want to estimate the scene topography and motion which correspond best to the measured data. Given the measured data $\mathbf{d}$ we want to estimate the most likely parameters $\hat{\mathbf{w}}$. The maximum likelihood (ML) estimator is then given by

$$\hat{\mathbf{w}}_{\text{ML}} = \max_{\hat{\mathbf{w}}}\{\Pr[\hat{\mathbf{w}}|\mathbf{d}]\}, \tag{1}$$

where $\Pr[\hat{\mathbf{w}}|\mathbf{d}]$ is the probability (likelihood) of the unknown parameters given the data. Using Bayes' theorem

$$\Pr[\hat{\mathbf{w}}|\mathbf{d}] = \frac{\Pr[\mathbf{d}|\hat{\mathbf{w}}]\Pr[\hat{\mathbf{w}}]}{\Pr[\mathbf{d}]}, \tag{2}$$

we obtain the MAP estimator as

$$\hat{\mathbf{w}}_{\text{MAP}} = \max_{\hat{\mathbf{w}}}\{\log\left(\Pr[\mathbf{d}|\hat{\mathbf{w}}]\right) + \log\left(\Pr[\hat{\mathbf{w}}]\right)\}, \tag{3}$$

where the data (and thus its probability) is constant, $\Pr[\mathbf{d}|\hat{\mathbf{w}}]$ is the probability of the data given the parameters and $\Pr[\hat{\mathbf{w}}]$ the probability of the parameters themselves. Essentially then, all the MAP estimator is doing is allowing use of additional information about the likelihood of the unknown parameter set. When no additional prior information about the parameter set is available MAP and ML estimators are the same.

The real challenge in the estimation scheme is in obtaining useful expressions for the likelihood of the data given the unknown imaging parameters, $\Pr[\mathbf{d}|\hat{\mathbf{w}}]$. Typical autofocus schemes use the prior information that the sharpest image is most likely (a contrast optimisation prior).

We choose an autofocus scheme where we include multiple sources of information. The three sources of information used in this work are

Sharpness	the sharpest image is best
Shape from shading	does echo match echo expected from the hypothesis sea-floor slope
Shadow mapping	does the observed shadowing match the hypothesis.

In addition, we show that optimising sharpness in an interferometric system is equivalent to estimation based on minimising interferogram phase. These three methods are described individually in the following subsections.

Our implementation of the iterative autofocus is shown in figure 1. We alter optimisation parameters until the most likely parameters are found and iterate until the chosen cost function, $\mathbf{C}$, is maximised. We discuss our cost function in more detail in section II-D.

Choosing an appropriate set of unknown optimisation parameters has been covered in earlier articles [2]. In this work we optimise using a conjugate gradient solver where $\mathbf{d}$ corresponds to the collected SAS images rendered onto a common Earth-fixed grid and where the optimisation parameters,

$$\hat{\mathbf{w}} = \left[\hat{\mathbf{z}}(x,y), \boldsymbol{\Delta}\hat{\mathbf{X}}_{\mathbf{n}}\right], \tag{4}$$

are a common scene topography, $\hat{\mathbf{z}}(x,y)$, and corrections of navigation positions in 3D space, $\boldsymbol{\Delta}\hat{\mathbf{X}}_{\mathbf{n}}$, for each pass n. Solving the optimisation problem may be seen as Simultaneous Localization and Mapping (SLAM).[2]

A. Image sharpness

Image sharpness is a constraint used often in autofocus. One states simply that the most likely image is the sharpest, or has maximum contrast. These priors are based on the premise that most mistakes in imaging parameters lead to defocus—which blurs the scene and lowers image contrast. As we show in the appendix, the same metric is useful in bathymetric estimation. Thus, the best combined estimate of bathymetry and unknown path parameters is the one generating the sharpest image.

Stating formally, for each imaging pass n. We make a coherent bathymetric image

$$f_n(x,y) = f_{\text{lower},n}(x,y) + f_{\text{upper},n}(x,y) \tag{5}$$

where $f_{\text{lower},n}(x,y)$ and $f_{\text{upper},n}(x,y)$ are the sonar images from the lower and upper banks, and calculate intensity per look, n, at the scene

$$I_n(x,y) = |f_n(x,y)f_n^*(x,y)| \tag{6}$$

We attempt to optimise contrast in the combined intensity image

$$I(x,y) = \sum_n I_n(x,y). \tag{7}$$

And state that the most likely image corresponds to the image with highest contrast

$$\Pr_{\text{sharp}}[\hat{\mathbf{w}}|\mathbf{d}] \propto \frac{\text{std}(I(x,y))}{\text{mean}(I(x,y))}, \tag{8}$$

where the standard deviation and mean operators average over the entire image.

B. Echo strength

Echo strength as a method for estimating local topographical variations is commonly used with structured lighting in normal image processing—so called shape from shading. In sonar it has been used with reasonable success in [4].

To use this source of prior information we state that the most likely height-map is the one where the predicted echo strength matches the observed data:

$$\Pr_{\text{echo}}[\hat{\mathbf{w}}|\mathbf{d}] \propto \frac{1}{I_{\mathbf{d},\hat{\mathbf{w}}}(x,y) - I_{\text{LL},\hat{\mathbf{w}}}(\hat{\mathbf{w}})(\mathbf{x},\mathbf{y})} \tag{9}$$

where $I_{\mathbf{d}}(x,y)$ and $I_{\text{LL},\hat{\mathbf{w}}}(\hat{\mathbf{w}})(\mathbf{x},\mathbf{y})$ are the observed and predicted models respectively.

There are a number of unknown parameters in the prediction model that must be assumed for the method to work. We adopt the model backscattering described in [5] and [6] for a homogeneous sandy sea-floor both for bottom and target.

[2]Typical SLAM implementations for AUVs do not model the actual height measurement process (usually based on interferometric measurements) so cannot provide as accurate results as this higher-order model. SLAM implementations however often model INS drift, something not attempted in this work.

C. Shadow mapping

Raytraced shadows give additional information on topographic maps. In particular this information source corresponds to that used in typical side-scan images to estimate object height.

We state that the most likely scene is that where the shadows match those predicted from a shadow-mapped hypothesis. This may be summarised as

$$\Pr_{\text{shadow}}[\hat{\mathbf{w}}|\mathbf{d}] \propto \frac{1}{S_{\mathbf{d},\hat{\mathbf{w}}}(x,y) - S_{\text{RT},\hat{\mathbf{w}}}(\hat{\mathbf{w}})(\mathbf{x},\mathbf{y})} \tag{10}$$

where $S_{\mathbf{d},\hat{\mathbf{w}}}(x,y)$ corresponds to binary shadow masks created from applying thresholds to images and $S_{\text{RT},\hat{\mathbf{w}}}(\hat{\mathbf{w}})(\mathbf{x},\mathbf{y})$ from predicted line-of-sight for a given parameter set $\hat{\mathbf{w}}$.

Obviously, this technique only gives information where the local-slope is such the self-shadowing of the topography occurs so appropriate weighting needs to be applied.

D. Cost function and optimiser

The cost function we use provides a method for weighting the three information sources discussed earlier.

We use a simple solver where we traverse along the predicted gradient of the cost function, $\mathbf{C}$. We use the derivations from the appendix to obtain an energy-neutral cost function gradient term that corresponds to final image contrast.

Thus we use

$$\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\mathbf{z}(x,y)} \propto \frac{\sum_n w_n(x,y)\angle\big(f_{\text{lower,n}}(x,y)f^*_{\text{upper,n}}(x,y)\big)}{\sum_n w_n(x,y)} \tag{11}$$

where the weights $w_n(x,y)$ are generated by multiplying:

$$w_n(x,y) = \frac{|\gamma_n(x,y)|}{1-|\gamma_n(x,y)|}s_n(x,y) \tag{12}$$

where $\gamma_n(x,y)$ is the interferometric coherence averaged over a 9 by 9 window and

$$s_n(x,y) = \begin{cases} 1, & S_{\mathbf{d},\hat{\mathbf{w}}}(x,y) \neq 1 \\ 0.01, & S_{\mathbf{d},\hat{\mathbf{w}}}(x,y) = 1, \end{cases} \tag{13}$$

is a term to penalize areas predicted by the current parameters to lie in a shadow region. Terms for $\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}\boldsymbol{\Delta}\mathbf{X}_\mathbf{n}}$ are determined by cross-correlating $|f_{\text{upper,n}}(x,y)|^{0.6}$ and for $|f_{\text{upper,n+1}}(x,y)|^{0.6}$ that are 90° from each other.

The, admittedly ad-hoc, optimisation system used here has one important advantage over that discussed in section II—height-estimates from multiple object reflections are discarded. The disadvantage is that the optimiser will not converge to the highest contrast image. It will however, be a better estimator of topography. Future work in defining a better cost function will be needed.

III. Results

At this stage we have implemented a simple descent-based solver which is applied to four 90° passes of HISAS 1030 SAS data collected with a HUGIN1000 AUV [7].

The scene of interest is a rectangular object of approximately 10 m by 3 m. The top-edges of the rectangular section are approximately 1 m over the sea-floor and surface about 0.6 m. HISAS 1030 data has been collected by running in the four cardinal directions at about 70 m range to the object. Both INS and DPCA navigation have been used and over the 15 minute imaging period relative accuracy is estimated to be at 1 m level.[3]

Figure 2 shows the effect of applying the method to obtain a good fused bathymetry and echo strength estimate. We believe that the fused results shown in figure 2 to be more likely to match the actual object in question. This is because:

- Bathymetry edge transition is sharper.
- Bathymetry on the object is flatter.
- Image edges become better co-located. (We believe there is a thin, high-reflectivity, railing around the edge of the object.)
- Effect of likely multiple reflections is reduced.

We admit that there are, as yet, no ground truth measurements so objective comparison is difficult. Section III-A describes the major reason for the improved bathymetric estimates.

The combined echo-strength estimate is not as useful as the individual images. The effects of layover, discussed in the following sub-sections, are a particular challenge.

A. Multi-bounce

For complicated objects, there will often be multiple internal reflections and complicated acoustical interactions. In particular, for man-made objects with straight edges and high reflection coefficients, multiple bounces may cause misleading bathymetric estimates, see figure 3.

Figure 4 shows two height-map estimates of the same object seen from different sides. The side nearest the sonar may suffer from a strong multiple bounce. Interferometry-based height mapping estimates the direction of arrival of the incoming wave. Where the echo goes from object to sea-floor, the bathymetric estimate will be incorrect. These regions are often visible as strong, coherent echoes from an unlikely height.

As may be seen in figure 5, this effect is strongly suppressed by weighting "shadow" regions with a penalty. As the multi-bounce is *below* the sea-floor it looks like shadow and is penalised by the proposed method. The end result in the optimisation is also visible in figure 2 where the surface of the object is more even after five iterations.

B. Layover

Layover, where the acoustic travel time corresponds to multiple ground-plane locations is a particular problem in CSA imaging. We show a simple example of this on the target of interest. The object most likely has vertical walls. Thus the true height, and ground-plane location are entangled and cannot be separated as shown in figure 6.

Circular geometries allow partial estimation of the true height as vertical wavenumber coverage is extended by the new geometrical regime [8]. The sea-floor in front of the object seen from one side is visible from another aspect and thus may

[3]Experimental results shown here appear to indicate 15 cm aided INS drift during the collection period.

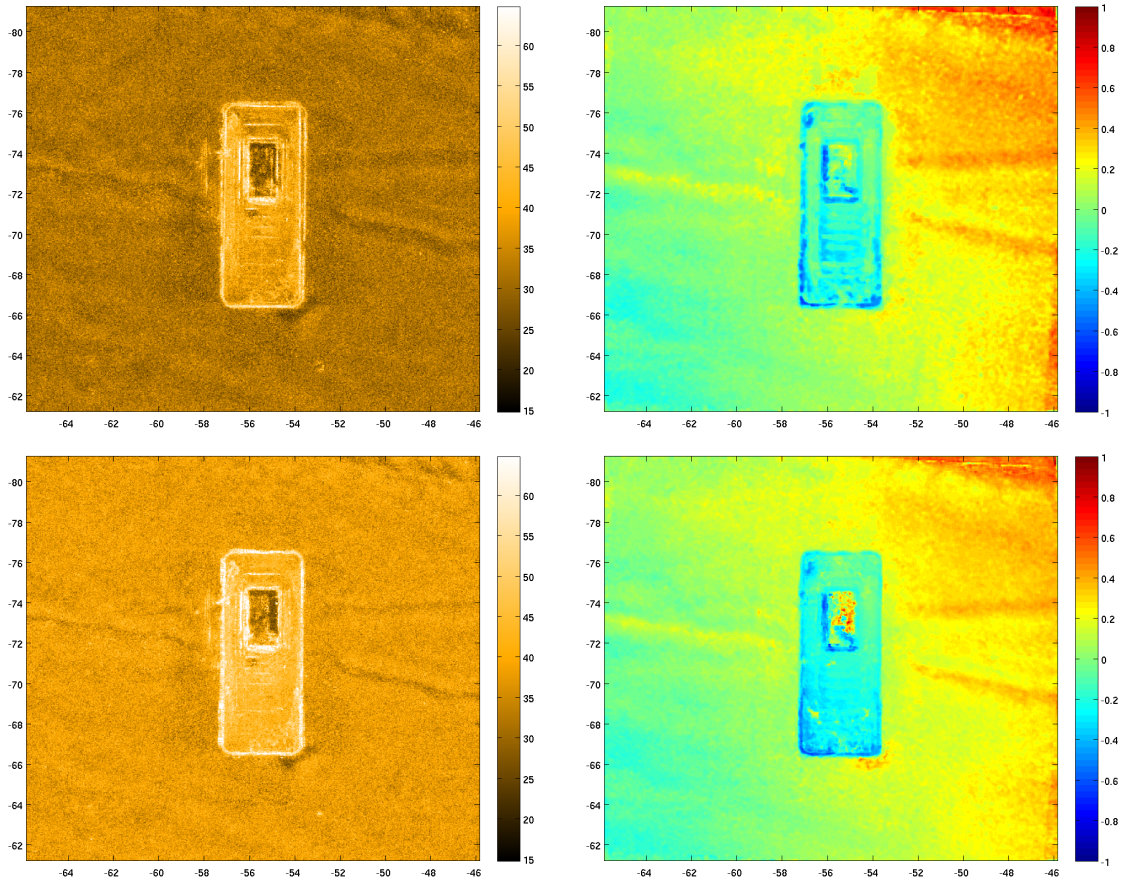

Fig. 2. Comparision of image (in log dB) and height map (relative in m) results at first and fifth iterations, four looks used (eight images). Above, combined image estimate (left) and height-map (right) at end of first iteration. Below, combined image estimate and height-map after 5 iterations. Axes in meters.

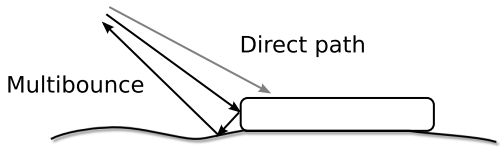

Fig. 3. Multibounce problem. Multiple reflection source has higher signal energy than direct path. Thus direction of arrival for the acoustic wave estimated by interferometry is "incorrect".

be estimated. Rendering in a common height-plane as done in this article is however fraught.

We show an example from the final bathymetric results shown in figure 2 which shows one of the eight combined images. Using a flat topography in imaging results in the first image (first iteration), but using the final estimated topography (fifth iteration) the second image is obtained. A number of image pixels on the right-hand edge are duplicated. This increases image energy, decreases contrast and obscures the sea-floor echo obtained from the other imaging aspects. This is a bad idea.

Due to these problems we suggest for future work using the multiple-aspect imagery as the basis for an inversion instead of attempting to estimate simple echo strength. These regions may then be weighted according to conservation of energy principles.

IV. Discussion

As discussed in sections II-D and III-B, straightforward image contrast is not a very stable measure of scene likelihood. In autofocus, energy is preserved as defocus simply moves image energy around. In imaging onto a predicted height-map this is also true, but instead of imaging many regularly spaced height planes we slice the vertical plane with the predicted height-map. This leads to large energy variations—energy may be "created" or lost by imaging points many times. This is clearly visible in figure 7. Overall this *lowers* contrast.

The iteration scheme derived in the appendix is based on assuming constant energy and therefore attempts to minimise interferogram phase. This strategy still works for producing height maps but will not in the end improve contrast. Better use of prior information is thus required and a new definition of the log-likelihood function used in the iteration framework.

Each of the current sources of prior information lack full coverage. Some challenges:

- The scene does not look the same from all sides, thus isn't maximally sharp at the correct parameters (e.g. imaging of spheres).
- Multiple target bounces is confusing—particularly for topographic prior information. The ad-hoc system works but better prior information should be used.
- Shadow mapping should not be a binary process.

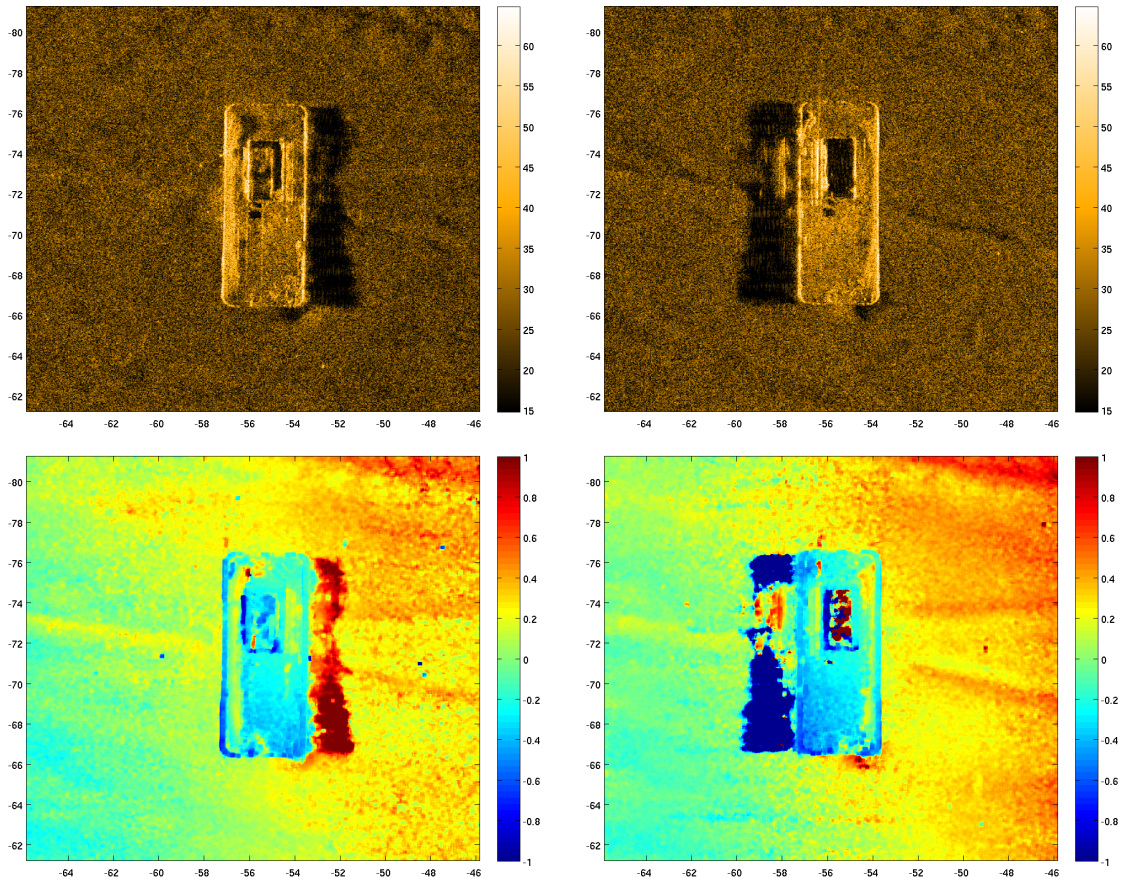

Fig. 4. Comparision of images (log dB) and interferograms (relative depth in meters) obtained from different view directions. View directions correspond to left, and right of the object respectively. Images are above height maps. Note height-map disagreement in region immediately following object leading edge.

- Echo strength predicts slope whereas interferometry estimates height.
- Using a scattering model assuming sandy floor is not sensible on objects.

Future work will concentrate on allowing the image target strength to be an unknown and better integrating the predicted echo strength. We would also like to estimate scattering parameters (surface roughness) as hypothesized in [7], [9].

V. Conclusion

We present a fusion of image based and bathymetry based SLAM in a common framework. This is the first time a combined bathymetry and autofocus framework has been presented.

The framework provides a simple method for including additional prior information sources. These sources can help provide improved bathymetric estimation in difficult cases.

New sources of prior information are desirable and future work will concentrate on finding them. We will also look into obtaining direct recursive solvers to reduce optimisation computational costs.

Acknowledgments

We would like to thank the Hydrography Department of Kongsberg Martime for assisting in collecting the experimental data presented here.

Appendix

Interferometry and Image Sharpness

We note that the combined image is made from two coherently combined interferometric images as

$$f(x,y) = f_{\text{upper}}(x,y) + f_{\text{lower}}(x,y). \tag{14}$$

Making the typical (narrow-band) interferometric assumption that the image pair contains the same complex scene with a superimposed phase shift corresponding to the (as yet unknown) scene topography gives (14) as

$$f(x,y) \approx f_{\text{scene}}(x,y)(1+\exp(-j\phi(x,y))), \tag{15}$$

where $\phi(x,y)$ is the interferogram phase we would normally use to estimate scene topography. The scene intensity is then given as

$$I(x,y) = |f(x,y)f^*(x,y)|, \tag{16}$$

$$= |f_{\text{scene}}(x,y)|^2|2+2\cos(\phi(x,y))|. \tag{17}$$

We make the approximation that only $\phi(x,y)$ varies with the predicted scene topography, and can thus state that maximum likelihood for height-estimation is,

$$\max_{\hat{\mathbf{w}}}\{\Pr[\hat{\mathbf{w}}|\mathbf{d}]\} = \min_{\hat{\mathbf{w}}}\{\phi(x,y)|_{\hat{\mathbf{w}},\mathbf{d}}\}. \tag{18}$$

Thus, the sharpest combined interferometry image corresponds to the image where the interferogram phase is zero.

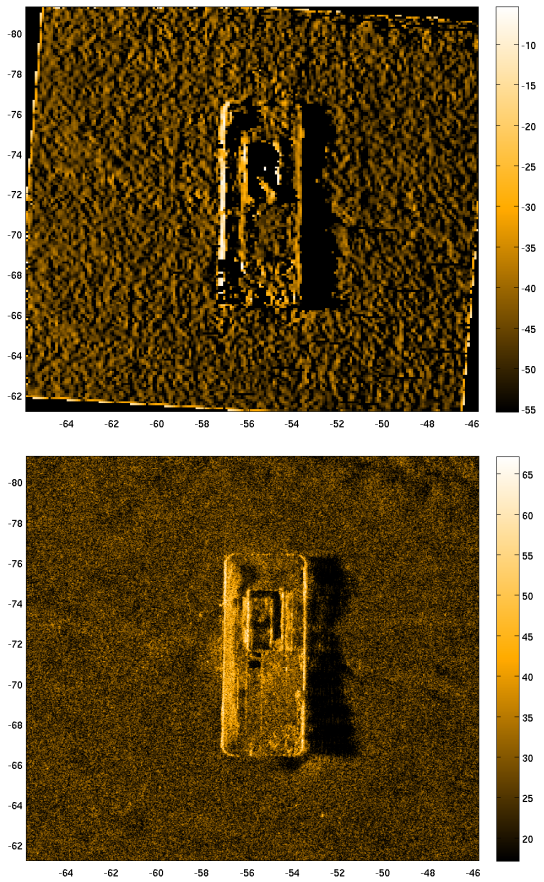

Fig. 5. Comparison of bathymetry predicted and measured echo strength images (in log dB). The sonar is not calibrated so only relative levels are useful for comparison. False estimated texture is due to slope standard deviations in the topography estimate. The bathymetry has been median filtered with a 5x5 window before being used to generate echo strength predictions.

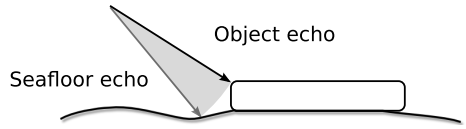

Fig. 6. Illustration of layover geometry. Weak seafloor and strong object echoes are impossible to separate from single look direction. Rendering onto the "correct" ground plane places object energy *both* on the object and on the seafloor.

It is worth noting that this derivation is applicable to multiple-vertically displaced receivers and is just an extension of similar along-track phase gradient constraints to the vertical array. Similar derivations are used in PACE [10] and recursive direct phase estimation techniques based on image contrast [3, Chapter 6].

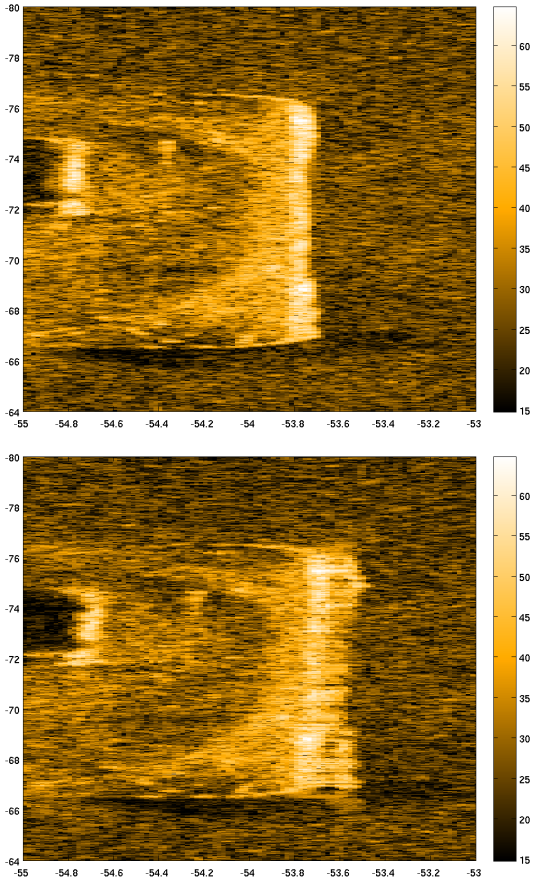

Fig. 7. Illustration of the layover issue. Top image has been rendered onto a flat topography. Bottom image has been rendered onto the height-map at the fifth iteration. Note in particular the "extension" of the right edge of the object response. The response is drawn out by approximately 10 cm. This results in an increase in image energy but a decrease in image contrast.

References

[1] D. P. Williams and J. Groen, "Multi-view target classification in synthetic aperture sonar imagery," in *Proceedings of the 3rd International Conference & Exhibition on Underwater Acoustic Measurements (UAM)*, (Nafplion, Greece), pp. 699–704, June 2009.

[2] H. J. Callow, R. E. Hansen, S. Synnes, and T. O. Sæbø, "Circular synthetic aperture sonar without a beacon," in *Proceedings of the 3rd International Conference & Exhibition on Underwater Acoustic Measurements (UAM)*, (Nafplion, Greece), June 2009.

[3] S. A. Fortune, *Phase Error Estimation for Synthetic Aperture Imagery*. PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, June 2005.

[4] E. Coiras and J. Groen, "3D target shape from SAS images based on a deformable mesh," in *Proceedings of the 3rd International Conference & Exhibition on Underwater Acoustic Measurements (UAM),*, (Nafplion, Greece), June 2009.

[5] S. A. Synnes, R. E. Hansen, and T. O. Sæbø, "Assessment of shallow water performance using interferometric sonar coherence," in *Proceedings of the 3rd International Conference & Exhibition on Underwater Acoustic Measurements (UAM)*, (Nafplion, Greece), June 2009.

[6] A. P. L. (APL), "High-frequency ocean environmental acoustic models handbook," tech. rep., APL, University of Washington, Seattle, Washington, 1994. TR 9407.

[7] R. E. Hansen, H. J. Callow, T. O. Sæbø, and P. E. Hagen, "Interferometric synthetic aperture sonar for pipeline inspection," in *Oceans pacific 2010*, (Sydney, Australia), 2010.

[8] M. Pinheiro, P. Prats, R. Scheiber, M. Nannini, and A. Reigber, "Tomographic 3D reconstruction from airborne circular SAR," in *Geoscience and Remote Sensing Symposium*, vol. 3, pp. 21–24, IEEE International, IGARSS, 2009.

[9] T. O. Sæbø, H. J. Callow, and P. E. Hagen, "Pipeline inspection with synthetic aperture sonar," in *Proceedings of the 33th Scandinavian Symposium On Physical Acoustics (SSPA)*, 2010.

[10] H. J. Callow, T. O. Sæbø, and R. E. Hansen, "Autofocus of synthetic aperture sonar data using the PACE algorithm," in *Proceedings of the ninth European Conference on Underwater Acoustics (ECUA) 2008*, (Paris, France), June-July 2008.