A MODEL-BASED MULTI-VIEW IMAGE REGISTRATION METHOD FOR SAS IMAGES

Johannes Groen, David Williams, Warren Fox

NATO Undersea Research Centre, Viale San Bartolomeo 400, 19126 La Spezia, Italy.

Contact: J Groen, fax. +39 0187 527 330, groen@nurc.nato.int.

Abstract: *Autonomous underwater vehicles equipped with high-resolution synthetic aperture sonar (SAS), and automatic target recognition (ATR) algorithms show great potential for the task of search, classify and map. The level of detail recorded by the sonar is typically on the order of hundreds of pixels on an underwater object, valuable for improving classification performance. However, a weakness that is still not under control is the object aspect angle, which when unfavourable causes misclassification leading to significantly increased false alarm rates. The natural solution to this problem is to view the object from more than one direction. For this approach, referred to as multi-view ATR, different sonar views are collected, associated, registered and combined. The registration is not trivial, because sonar resolution surpasses object localisation accuracy. The registration approach proposed here is model based. Image templates from a database of objects are matched to the single-view images, which generates both a matching score and optimal registration shift for each object. Model-based registration is shown to be a robust technique, enabling automatic multi-view SAS image fusion.*

Keywords: *Synthetic aperture sonar, detection, classification, mine hunting*

1. INTRODUCTION

Underwater robots will play an increasingly important role in the hunting and neutralization of sea mines. In particular the task of search, classify and map can be conducted by autonomous underwater vehicles (AUVs) equipped with high-resolution sonar rather than personnel on vessels. The automatic classification performance of these systems is currently being improved significantly. Despite this improvement, excessive false alarm rates for complex seabeds are considered the primary remaining issue. A known effective approach to the high false alarm rate problem is to collect and combine

different sonar views of a potential target. In order for this to be successful the different views need to be associated robustly. This stage typically reduces the relative position error of the different views to object size, i.e., from several meters to about a meter. Different approaches have been studied that can obtain the result. For this article the detector described in [1] following by a simple association algorithm based on geo-information of the detections is employed of which an example is shown in Fig. 1. The scheme proved to perform adequately as long as the typical distance between contacts detected is less than the relative navigation error. Techniques such as simultaneous localisation and mapping (SLAM) [2] and association using multiple contacts [3] can be applied to ensure correct contact association even in harsher environments.

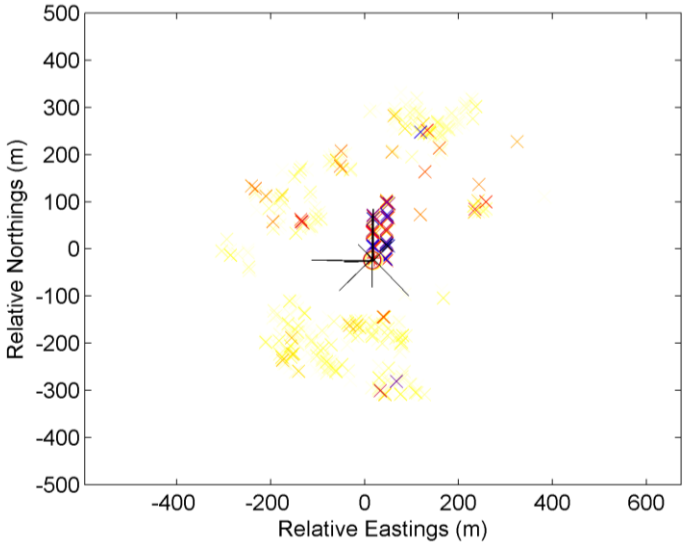

Fig.1: Detection map of a MUSCLE AUV mission, with colours corresponding to detection score. The views (see line between contact and AUV) of one object are collected.

The single-views are preferably fused at image level, providing an unambiguous multi-view image that is valuable in the decision process. A result of a cylinder of about two meters long and half a meter diameter in Fig. 2 shows that the multi-view image is more conclusive and moreover less dependent on aspect angle. Previously fusion of the different views using the single-view classifier output was investigated [4][5]. It is not known yet which approach performs best, or even if there is value in using both approaches.

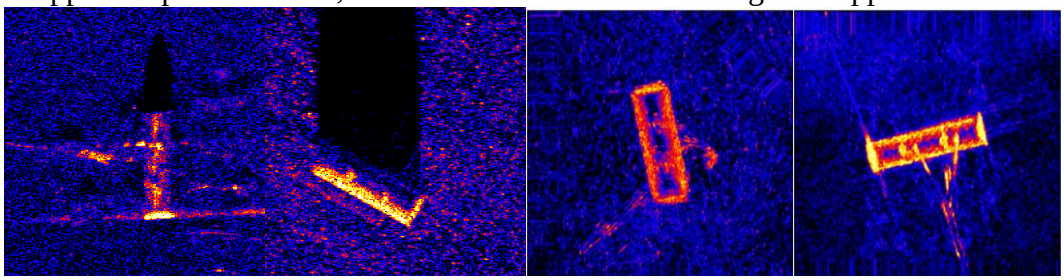

Fig.2: Two single-view images and two multi-view images of a cylindrical object.

The multi-view image can furthermore produce new features that lead to better classification. Co-registration of the single-view images with pixel accuracy is crucial and not trivial, primarily due to inadequate underwater positioning. The registration approach proposed here is model based. Image templates are matched to the single-view images, which generates both a matching score and optimal registration (2D shift) for each object. The template matching approach is well known for mine classification with sonar [6], and has been often employed in single view automatic target recognition (ATR) systems. Consequently one best matching multi-view image is created for each object. In this paper the results of this multi-view registration technique are shown using experimental

synthetic aperture sonar data collected with the MUSCLE AUV. It is also demonstrated that the obtained matching scores are a robust way to identify the correct multi-view image, i.e., the one actually corresponding to the target in the image. Multi-view image classification was previously shown to significantly suppress false alarm rate, but only after robust image registration [7]. It is concluded that model-based registration is a robust technique that enables automatic multi-view SAS image fusion.

The paper is organised as follows. First the template matching applied to the single views is explained in Section 2, followed by the registration that is described and tested on real data in Section 3. Section 4 concludes on the findings in the article.

2. TEMPLATE MATCHING

The case that is studied in this paper is the most south western target that was deployed and surveyed on 29 April 2008 during the Colossus 2 trial conducted by the NATO Undersea Research Centre off the coast of Latvia. The target is a two-meter long cylinder sitting on a flat sandy seabed. Its location relative to other detected objects is shown with the circles in Fig. 2, and the lines indicate where the AUV was. After the simple contact association algorithm we obtain nine views of the object from the complete AUV mission (a lawnmower pattern which was designed to see all objects from different angles). A view is included when it is within 15 m of the centre point of the cloud of detections for this particular object. In the top row of Fig. 3 it can be seen that the single-view images differ a great deal, both in nature and in quality. The second image, for instance, contains some grating lobes and shadow fill-in, and the third and fourth image are rather unclear due to the unfavourable aspect. This unfavourable effect is especially evident in the first image of Fig. 1, which highlights only certain features like a shackle, cable pieces and the end cap of the cylinder, and makes it difficult to classify it as a cylinder.

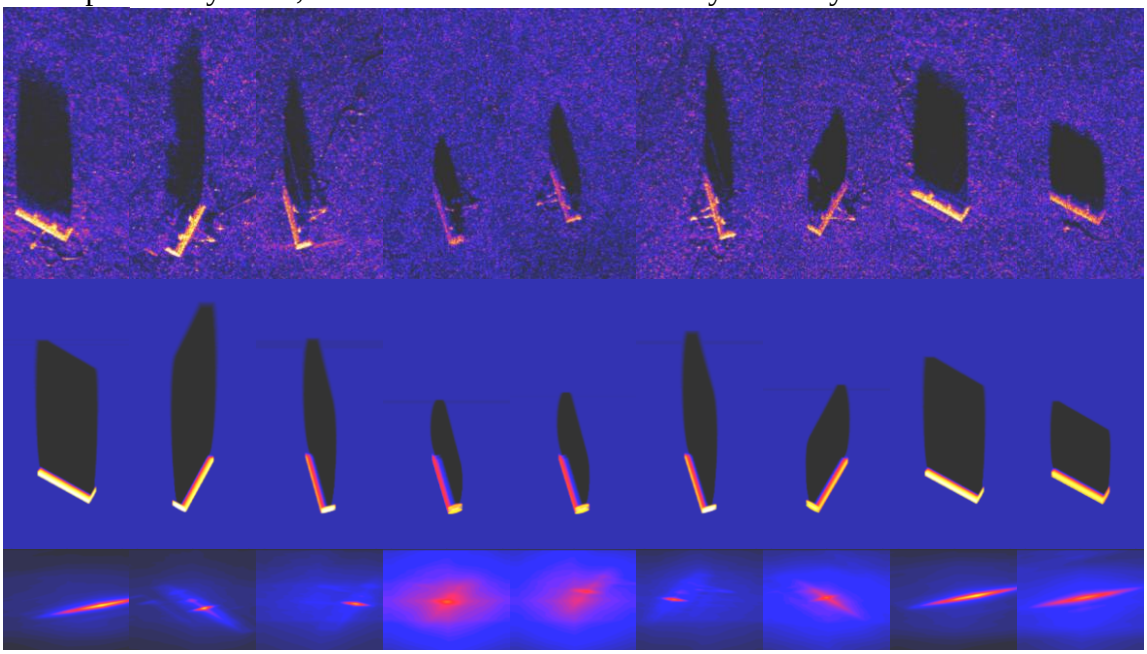

Fig.3: Top row: single views of around 3 by 7 m associated with the object of interest. Middle row: cylinder templates matching best with top row, assuming relative aspects from the AUV compass. Bottom row: corresponding 2D correlation surfaces.

The single-view images are now first matched with templates. In general, it is obviously not known *a priori* that the image contains a cylinder, but the idea here is to repeat this process for a set of targets. In addition it is not known *a priori* what the

orientation of the object is. Therefore, we also have to repeat the process for a set of potential orientation angles. Nevertheless, an important piece of information that is available and valuable is the relative aspect angle. The difference in aspect angle between the different single-views are measured via the heading sensor of the AUV (much less than a degree), which is typically much better than required for the fusion (a few degrees). The templates are generated with the model SIGMAS [8], which generates the signature of an object based on the sonar characteristics and geometry. For the templates in the middle row of Fig. 3 a cylinder was assumed sitting on a sandy seabed, with range taken from the top row SAS images (between 40 and 150 m) and the measured value of the AUV altitude (about 13 m). Each SAS image is subsequently matched with its corresponding template, following a 2D correlation. Due to image pixel statistics and the fact the real SAS images contain ropes and shackles, the resulting correlation coefficients in practice do not supersede 0.8. The bottom row shows the correlation result for the example. The views with a high highlight-to-reverberation give a higher and narrower peak, which indicates a better match. If the correlation peak is exactly in the origin, it means that image and template are aligned. If not the peak position provides the 2D shift between SAS image snippet and template. It is this shift that is the crux of the model-based registration. With these 2D shifts that are corresponding to the template object, each view can be registered to the hypothesised object. Continuing with this assumption the best possible cylinder is created from the views and this process is repeated for the other objects in the database.

As mentioned in the beginning of this section, the correlation processing has to be repeated for several orientation angles. The result for this example is shown in Fig. 4 for orientations between 0 and 180 sampled every degree. The correlation peak for each of the nine views is shown with the dotted lines, which results in curves that have a maximum at approximately the same angle, viz. the correct orientation of the cylindrical object. The width of the peak is around 20 degrees, and is wider for the views that see the cylinder from end-fire. This corresponds well with the findings described in [9]. For symmetric shapes, such as a sphere or a truncated cone, this step in the processing can be disregarded. However, for asymmetric target shapes the process is still straightforward. For each target shape in the database the peak values can be plotted versus orientation angle (dotted lines in Fig. 4), after which the average can be calculated (solid line). The peak of this curve is then taken as the best match, which simply leads to the best matching templates (the middle row of Fig. 3). In this way we already obtain a multi-view score, i.e., the peak of the solid line in Fig. 4, which can be used directly as a classification clue. It tells us how much the best possible cylinder generated from the single views actually looks like a cylinder. However, this is one multi-view score. It does not yet fully exploit the fused multi-view image, as will be explained Section 3.

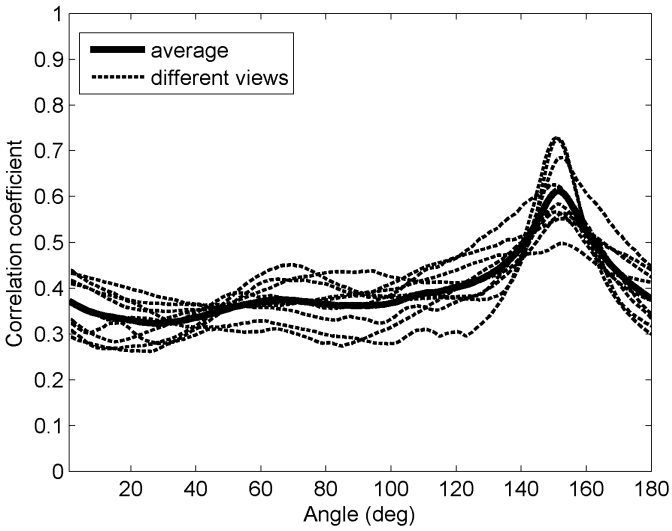

Fig.4: Peak value of 2D cross-correlation of the nine SAS image snippets for several cylinder template orientation angles.

3. REGISTRATION OF MULTIPLE VIEWS

In the previous section the methodology to obtain the information necessary for registration of the different views was explained and demonstrated on one real data example with a cylinder. The reason why one has to go through a complicated registration was explained in [7]. All the information required is now available: image orientation from the compass in the AUV, and both shift in range and cross-range from the 2D correlation. Shifting and rotating all the views results in true north SAS image snippets with the origin centred on the centre of the object used for the templates. For the example of the nine views on the cylinder this registration result is visualised in Fig. 5.

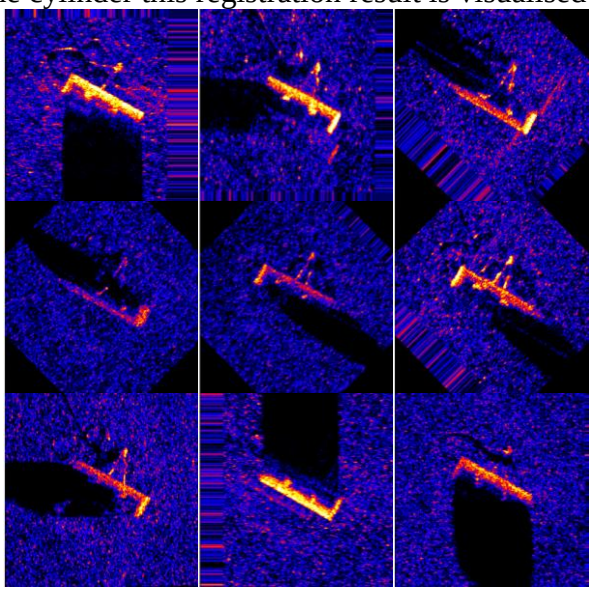

Fig. 5: Nine views of the cylinder rotated and shifted after the model-based registration using a cylinder template at its best matching aspect.

The correct template object, i.e., the one that truly corresponds to the image, is generally not known. However, if the procedure is simply repeated for several objects, the best possible multi-view image for each of those suspected targets will be produced. It is

hoped that the correct object matches best. It should also be noted that even if the object in the image is not in the database, the final result will still be reasonably close. After repeating this calculation for the wedge, the truncated cone, the cylinder and the sphere, the resulting best multi-view reconstructions become the bottom row of Fig. 6 after incoherent image fusion as described in [7].

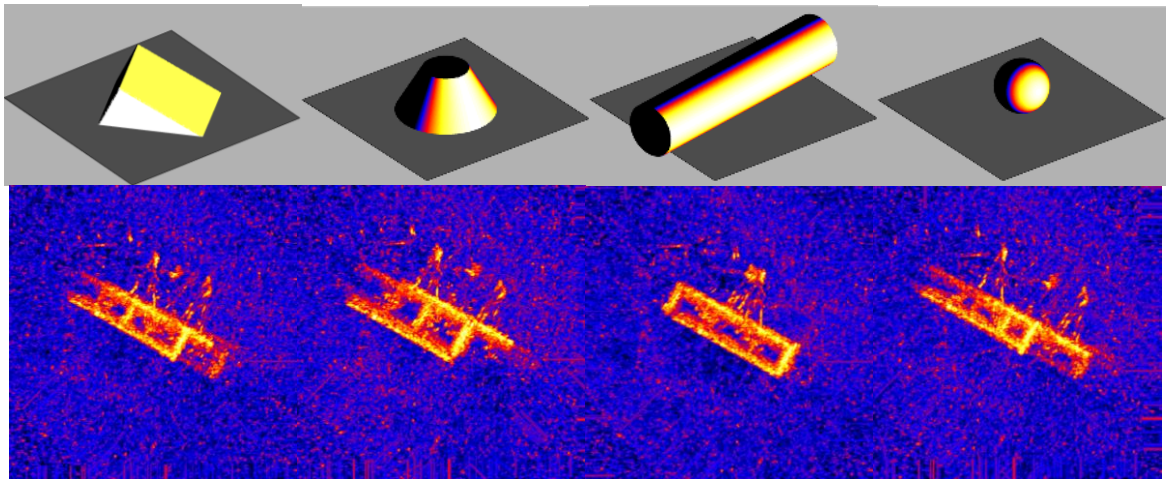

Fig.6: The best possible multi-view reconstruction based on the images collected of the cylinder matched with the four objects of interest.

It is clear from Fig. 6 that even after optimal shifting and matching, the best possible image for the wedge, truncated cone and sphere do not resemble the image that one would expect from such a target. The cylinder, although not perfect, looks very much like an image one would expect from a cylinder. There are tiny artefacts visible for the rope and the shackle, but those are not really supposed to be solved, because the template was not prepared with those.

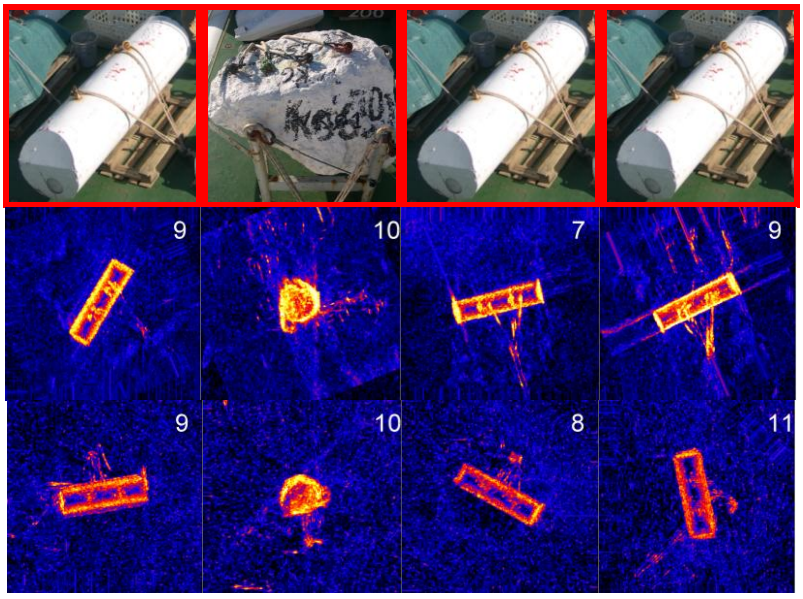

Fig.7: Target shapes (top row) and corresponding best reconstructed multi-view SAS images, deployed on a muddy seabed (middle row) and sandy seabed (bottom row).

After this test case the procedure was automated and run for trial data that was obtained during the Colossus 2 trial conducted by NURC off the coast of Latvia in 2009. The two experiments that were analysed here were on the same target set, which were deployed on a muddy seabed in the first experiment, and on a sandy seabed in the second experiment,

visible in the bottom row. The target-highlight-to-reverberation ratio was noticeably higher in the case of mud. The computation to generate the multiple views was fully automated, which is a requirement due to its role in the autonomous MCM system. No tuning parameters exist in the computation, and views were automatically selected from an automated detector [1], and corresponding geo-referenced information. The number in the multi-view images in Fig. 7 corresponds to the number of views that were fused. The overall result is considered promising for a number of reasons. The computation is already very robust and generates reliable results. When observing these results it can be concluded that the multi-view images indeed look like the objects that were imaged. Even if the rock was not in the database, this 'difficult' target still resembles the rock. The disadvantageous aspect and range dependence (present for single-view images) is removed, and images of all the cylinders are indeed very similar and ideal for classification.

The images also directly provide strong classification features, viz. the individual matching scores from the single-view images. The normalised average of these values was used to determine which object is the correct one. Fig. 8 shows all the average matching scores. For example, the first image of the middle row in Fig. 7 corresponds to the square in Fig. 8, which is the highest match, 1.5 standard deviations higher than the average other matches. So it is rather safe to say that middle row, image number 1, is a cylinder.

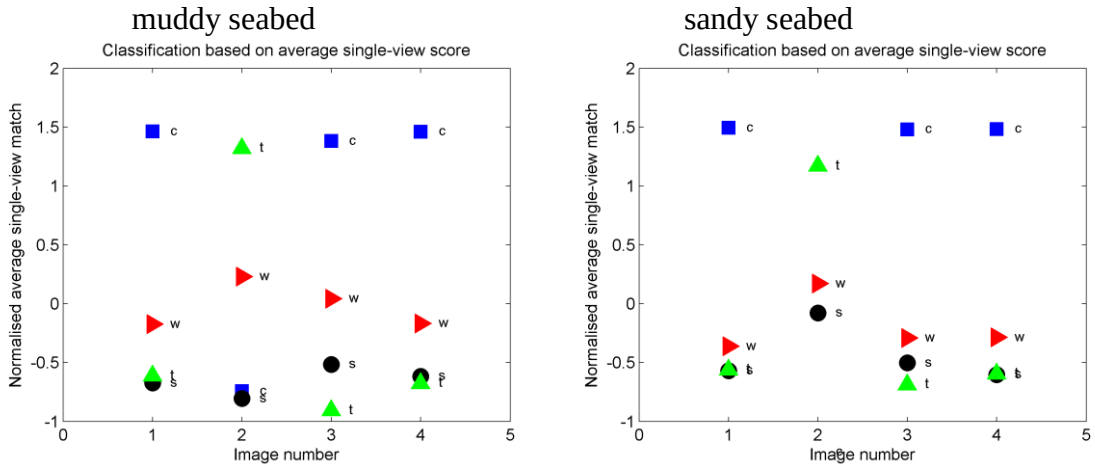

Fig.8: Classification scores for four template shapes, cylinder (square), wedge (triangle right), truncated cone (triangle up) and sphere (circle).

The results plotted in Fig. 8 are encouraging. The cylinder is always correctly selected, and the rock was classified as a truncated cone, because there was no rock in the database. It should be noted that the classification feature discussed here, the average of the single-view matching scores, does not fully use the potential of multi-view analysis yet. The actual multi-view image fits everything together into one image, which should in theory look very much like the object it corresponds to. Model-based matching based on the actual multi-view images and multi-view templates is currently being investigated and shows an even higher performance gain. The other important issue that deserves attention is value of multiple views in situations that are far more complex than the ones in this paper, because this is where multi-view performance gain is needed the most. It was found that underwater positioning is an important variable here, due to the fact that which view belongs to which object becomes more and more ambiguous when detection rate rises. Better navigation, mosaicking techniques and SLAM are expected to help here.

4. CONCLUSION

In this paper a new method for multi-view SAS image registration was described and demonstrated on data collected at sea with an AUV. The method proved successful on real data, is automated and is promising for multi-view classification methods. The next step in the multi-view classification is to use this multi-view image and extract classification features from it directly, which is expected to be much more powerful than fusion of single-view features as used so far.

REFERENCES

[1] **J. Groen, E. Coiras, and D. Williams**, Detection rate statistics in synthetic aperture sonar images, *Proc. of the UAM*, pp. 367–374, 2009.

[2] **M. W. M. G. Dissanayake, H. Durrant-Whyte, S. Clark, and M. Csorba**, A solution to the simultaneous localization and map building (SLAM) problem, *Univ. Sydney, Sydney, Australia, Tech. Rep. ACFR-TR-01-99*, Jan. 1999.

[3] **E. Coiras, F. Baralli, B. Evans**, Rigid Data Association for Shallow Water Surveys, *IET Radar, Sonar and Navigation*, Vol.1, No.5, pp. 354-361, 2007.

[4] **J. Fawcett, V. Myers, D. Hopkin, A. Crawford, M. Couillard, B. Zerr**, Multiaspect Classification of Sidescan Sonar Images: Four Different Approaches to Fusing Single-Aspect Information, IEEE JOE, Vol.35, No.4, pp.863-876, Oct. 2010.

[5] **G. Dobeck, J. Hyland and L. Smedley**, Automated detection/classification of seamines in sonar imagery, Proc. SPIE, vol. 3079, p.90, 1997.

[6] **D. Williams**, Bayesian Data Fusion of Multiview Synthetic Aperture Sonar Imagery for Seabed Classification, *IEEE Transactions on Image Processing*, Vol. 18, No. 6, pp. 1239-1254, June 2009.

[7] **J. Groen, E. Coiras, and D. Williams,** False-Alarm Reduction in Mine Classification using Multiple Looks from a Synthetic Aperture Sonar, *IEEE OCEANS 2010*, Sydney, Australia, May 2010.

[8] **E. Coiras, J. Groen**, Simulation and 3D Reconstruction of Side-looking Sonar Images, *book chapter in "Advances in Sonar Technology", In-Tech Books, Vienna, Austria*, pp.1-14, 2009. http://intechweb.org/book.php?id=142.

[9] **J. Groen, E. Coiras, J. Del Rio Vera, B. Evans,** Model-Based Sea Mine Classification with Synthetic Aperture Sonar. *IET Radar Sonar and Navigation*, 4(1), 62, 2010.