

# Landmark detection from sidescan sonar images

Mohammed Al-Rawi<sup>1</sup>, Adrian Galdran<sup>2,3</sup>, Fredrik Elmgren<sup>4</sup>, Jonathan Rodriguez<sup>1</sup>, Joaquim Bastos<sup>5</sup>, Marc Pinto<sup>6</sup>

<sup>1</sup> Departamento de Eletrónica, Telecomunicações e Informática (DETI), Universidade de Aveiro, Aveiro 3810-193, Portugal  
al-rawi@ua.pt, jonathan@ua.pt

<sup>2</sup> INESC TEC, R. Dr. Roberto Frias, 4200 Porto, Portugal

<sup>3</sup> Tecnalia Research & Innovation, Derio, Spain  
adrian.galdran@inesctec.pt

<sup>4</sup> DeepVision AB, Linköping, Sweden  
fredrik@deepvision.se

<sup>5</sup> Instituto de Telecomunicações - Pólo de Aveiro, Aveiro 3810-193, Portugal  
jbastos@av.it.pt

<sup>6</sup> ECA Robotics, 262 rue des frères Lumière, 83130 La Garde, France  
mpi@eca.fr

**Abstract**—Sidescan sonars have seen wide deployment in underwater imaging. They can be used to image the seabed to a rather acceptable resolution from a few centimeters to 10 centimeters. Yet, sonar images are still of a substantially lower visual quality as they suffer from quite a few problems, e.g., acoustic shadows that vary according to vehicle heading and sonar grazing angle, speckle noise, geometric deformation due to ping variation and speed of vehicle carrying the sonar, etc. Landmark detection in sidescan sonar images is vital to find objects and locations of interest that are useful in various underwater operations. The objective of this work is proposing novel landmark detection methods for this class of images. Cubic smoothing spline fitted to the across-track signals is proposed as a method to detect the objects and their shadows. To cover a large area, experimental data has been acquired during missions performed in Melenara Bay (Las Palmas/Spain) using autonomous underwater vehicles (AUVs) equipped with Klein 3500 sidescan sonar. The AUVs have been deployed in two missions (one mission performed each day) and a total of 25 large-resolution images have been acquired. The AUV generated 12 parallel path images in the first mission and 13 parallel path images in the second mission with an angle of 70 degrees between the direction of mission #1 and mission #2. This difference in the directions of the two missions was necessary to ensure different acoustic shadows between the two sets of images, each set being generated from a different mission. Results show that the proposed methods are powerful in detecting landmarks from these challenging images.

**Keywords**—sidescan sonar; cubic smoothing spline; autonomous underwater vehicles; underwater landmark detection; AUVs; SLAM

## I. INTRODUCTION

The detection of distinguished landmarks from underwater sonar images is a useful but challenging task. One reason is that landmarks can be used to facilitate the movement of Autonomous Underwater Vehicles (AUVs) that are equipped with imaging sonars, in what is referred to as simultaneous localization and mapping (SLAM) [1]. It is a highly challenging and difficult to achieve task since analyzing sonar images usually implies several difficulties, e.g., the presence of acoustic shadows and their variation with the geometry, the speckle noise

in the images, the echo decay due to sound absorption by water, the artifacts in the image due to the grazing angle, the environmental factors that may modulate the sound parameters, and objects floating close to the seabed. In fact, the acoustic shadows and the speckle noise are severe in any type of coherent imaging, even those based on laser, radar or LiDAR. Nonetheless, these problems are rather generic (true for both sonar and optics) but that the main benefit of sonar is increased range in turbid water at the price of reduced resolution.

One of the main computational tasks associated to sonar-based seabed mapping is the registration of several acquisitions of the same area at different times. Solving this complex problem is a prerequisite to perform robust sonar mosaicking, 3d seafloor modeling, mission control, and other relevant tasks. There exist two main approaches for sonar image alignment: global registration techniques and local matching techniques [1][2]. The goal of global registration strategies is to devise robust similarity metrics capable of identifying a geometrical transformation of an image coordinate system into a second scan. Visual features of both images are extracted at a global level and compared in order to find such transformation. Examples of this approach can be found in [3] where mutual information criterion is applied and in [4] where an adapted particle filtering has been proposed. Unfortunately, the large resolution of modern sidescan sonar scans and their highly textured and noisy content make global registration a considerably complex problem. A reasonable solution to this problem is given by using local techniques that model the scene through the presence of landmarks. Landmarks are defined as small regions that represent meaningful and highly discriminative areas of the seabed. Landmark detection methods aim at extracting these regions in automatic manner and then encoding them in a reference database. When a second scan is analyzed, the same method is applied to extract the landmarks and they are compared with the reference database in order to align the entire image through the matching of corresponding landmarks. To reduce uncertainty, supplementary geographical information [5] as well as information related to the sonar motion are usually used [6][7]. To improve registration, the authors of [8] adopted to augment the sonar scan with

bathymetric profiles that have been obtained previously. The authors of [9] fused optical information with sonar data. In [10], landmarks were detected in subsampled sonar scans by a process of image normalization followed by a simple global threshold-based segmentation of highlight, shadow, and background regions. A more advanced segmentation technique, relying on active contours that evolve in the space of Haralick texture features was proposed in [11]. Also in [1], the popular Viola and Jones efficient cascade detection technique has been adapted for landmark detection. Departing from the standard two-dimensional image processing strategy, it has been recently shown in [12] that information drawn only from scanlines, instead of the entire sidescan image, can be enough to detect landmark location, which avoids the need of processing the whole scan at once and can be useful for real-time landmark detection. This work makes use regression methods and relies on 1-dimensional scanlines information to detect the landmarks of seabed images.

## II. METHODS AND DATA

The imaging data used in this work has been collected in the context of the SWARMS project (Smart and Networking Underwater Robots in Cooperation Meshes; [www.swarms.eu](http://www.swarms.eu)). SWARMS' goal is to provide new capabilities for the exploration of the seabed by cooperative teams of underwater robots. These capabilities require the application of several computer vision tasks, e.g. registration, matching, fusion, etc., which can greatly benefit from an advanced data pre-processing techniques. Our tasks in SWARMS have mainly been related to

preparing the seabed maps and to performing various functions on them, e.g., registration, matching, and fusion. For the acquisition of sonar images, AUVs equipped with Klein 3500 sidescan sonar were employed to image the seabed, ECA's A9 AUV is shown in Fig. 1. The missions have been performed in Melenara Bay (Las Palmas/Spain) and have been planned to cover a large area of approximately  $400 \times 250 \text{ m}^2$  with a water depth varying from 10 to 25 meters, which resulted in 25 seabed images. Specifically, two missions have been defined in order to survey the selected area using two different headings, as shown in Fig. 2. According to the sonar range scale, a constant spacing between the rails has been used to allow the full coverage of the nadir gap and the 200% coverage of the area outside the gaps. The missions have been conducted using a vehicle speed of 3 knots and maintaining a 4 meter fixed distance from the seabed (AUV altitude). The sonar has been set with an operative frequency of 900 MHz, a range scale of 40 meters and a pulse length of 1 milliseconds.

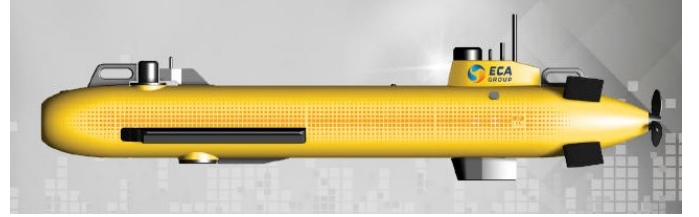
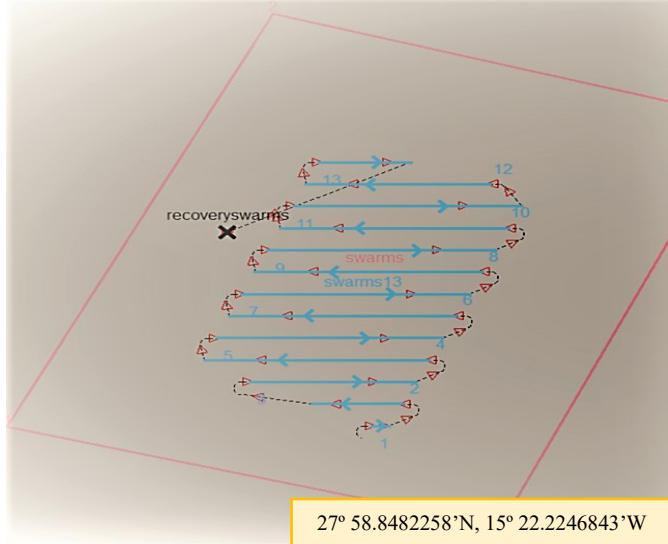
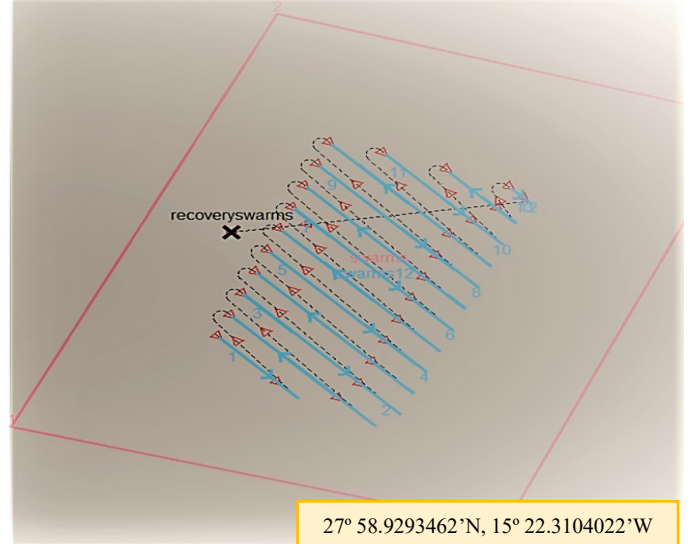


Fig. 1. ECA's A9 autonomous underwater vehicle that was used in the missions.



Mission 1



Mission 2

Fig. 2. AUV planned path for acquiring 25 seabed images, generated by ECA's mission planner. The blue numbers denote the image index and each blue arrow denotes the direction of the AUV.

There are various sources of noise and artefacts in sonar images, for example, there is a decay in the received echoes as a function of range due to sound absorption in seawater and the direction of the receiver. In addition, there is also spherical spreading expressed as  $40 \log(r)$  in the same log scale. The seabed backscatter strength varies strongly with grazing angle; hence, range and typical variations are in excess of 30 dB. Hence, the total dynamic range requirements of sonars exceed

100 dB, which is much too large to be pictured/acquired in a single image. Therefore, all sonar manufacturers implement a proprietary dynamic range compression with the objective of reducing the dynamic range to around 40 dB. The goal is not only to acquire images of a better visual quality, but also to enable using computer vision and image analysis techniques on those corrected/normalized images.

### A. Landmark detection

The cubic smoothing spline (which is characterized by a smoothing parameter or weighting parameter  $P$ ) stands out among many others techniques [2]. The cubic smoothing spline combines the ideas of cubic splines and curvature minimization to create an effective data modeling tool for noisy data [2]. Therefore, the cubic smoothing spline technique is adopted in this work and will be applied for each across-track signal, enabling the estimation the across-track signal profile for each image.

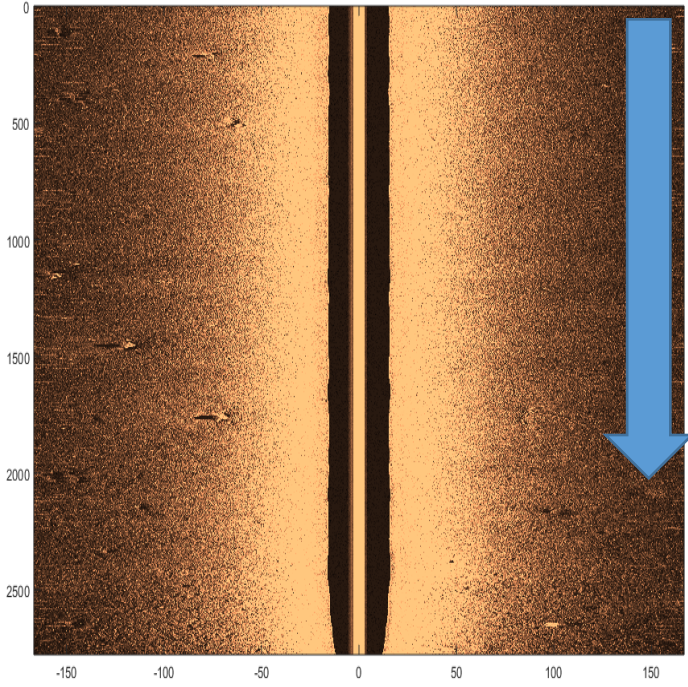


Fig. 3. Direct display of one seabed image. Left and right sidescan sonar images of a seabed location in Las Palmas de Gran Canaria (2777 pings  $\times$  3334 across-track samples for each side). Shown is the ninth image of the first mission. The blue arrow denotes the direction of AUV movement.

In Fig. 4 we show a simple illustration of how the peaks and the acoustic shadows might be formed. A distinguished landmark located around ping 2650 and sample 1741 from the image shown in Fig. 3 is illustrated in Fig. 5. The across-track signals shown in Fig. 5 are considerably noisy and a better representation of these signals is achieved using the cubic smoothing spline fitted on each of these across-track signals, as illustrated in Fig. 6. In what follows, we shall refer to each fitted curve as signal profile. The landmark detection method that is adopted in this work is based on detecting the peaks and valleys in each across-track signal profile. However, not only the detection of the location of peaks and shadows is relevant, but also estimating the width of each of them. The reason is that we can expect the width to be proportional to the object/landmark and/or its shadow size, respectively. Note that the first peak in the across-track signal profile is always a ‘false-peak’ due to the receiver gain, i.e. this peak is not due to some object. Therefore, the first peak should be neglected in all the analysis. The process of finding the peaks in the across-track signal profile is illustrated in Fig. 7. It must be noted that the width of the peak may not well represent the width of the object, thus,

morphological operators can be used to enlarge the object mask that can be extracted from the widths of the adjacent across-track signal profiles that are obtained from the corresponding original across-track signals.

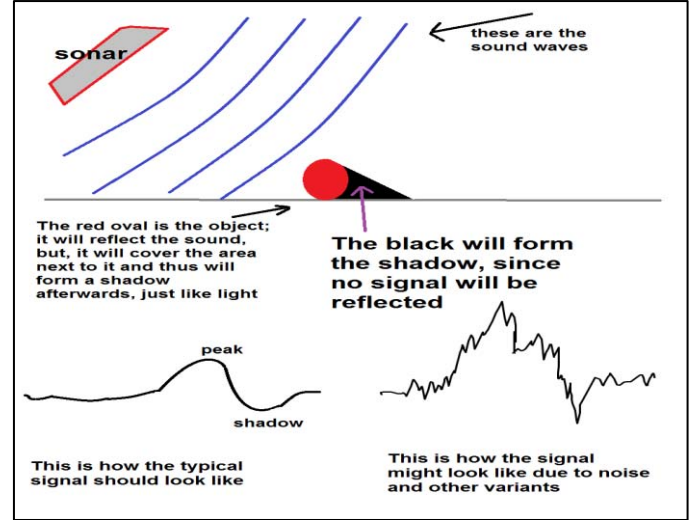


Fig. 4. Formation of peaks and shadows in sonar images.

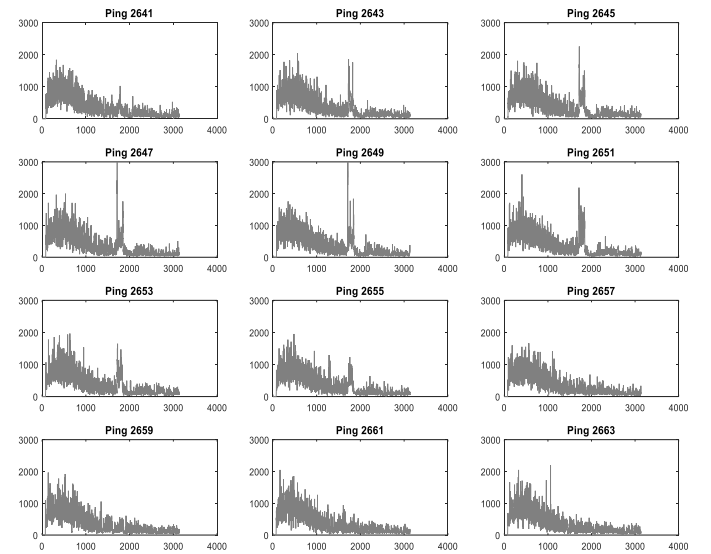


Fig. 5. Across-track signals appearing in the right sonar image between ping 2641 to ping 2663 containing what should be a distinguished landmark or object. The y-axis denotes the signal strength and the x-axis denotes the across-track sample index.

After finding the peaks, the across-track signal profile can be inverted and the method used to find the peaks and their widths can then be used to find the valleys (acoustic shadows) and their widths, as illustrated in Fig. 8. The first valley in the inverted signal, however, still needs to be neglected since it normally corresponds to the false-peak of the original signal. We illustrate finding the valleys in Fig. 8 the method used to detect the acoustic shadows. In all methods, the width is estimated for each peak as the distance between the points where the signal



intercepts a horizontal reference line. The points are found by linear interpolation. The height of the line is selected using the criterion specified by positioning the reference line beneath the peak at a vertical distance equal to half the peak prominence. The prominence of a peak is an estimate that shows how much the peak is significant with respect to its essential height and its location relative to other peaks.

### III. EXPERIMENTAL RESULTS

The proposed methods have been tested on the 25 seabed images. First, the cubic smoothing spline has been applied on each across-track signal and the peaks and shadows have been detected. Firstly, we performed further analysis to exclude peaks and valleys when their prominence value does not pass a certain threshold. Thus, a threshold prominence value will be estimated to enable selecting only significant peaks and valleys.

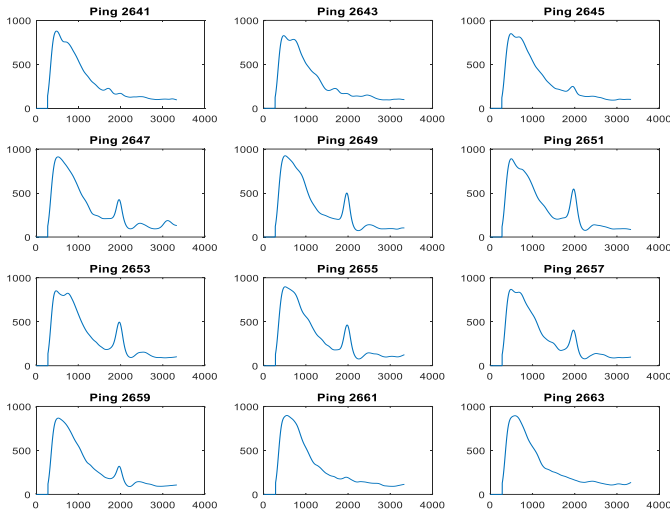


Fig. 6. Landmark appearing in the right sonar image between ping 2641 to ping 2663. These signals have been obtained from the original across-track signals using the cubic smoothing spline. The y-axis denotes the profile of the across-track sample and the x-axis denotes the across-track sample index.

To give an example, the widths of acoustic shadows shown in Fig. 8 are estimated as 482, 90, and 17, and they are located at 2236, 2803, and 3167, respectively. The peaks shown in Fig. 7, on the other hand, have the following widths and locations; {293, 573, 126, 176} and {1971, 2560, 2865, 3265}, respectively. The prominence values for the acoustic-shadows are 87.7661, 0.7430, and 4.28; and for the peaks are 342, 64, 0.74, and 2.9. As we see, the process outlined above corresponds correctly with a visual intuition of what constitutes relevant peaks and valleys. Now, estimating threshold value that can help us decide which peaks/valleys are significant is a critical issue. A simple way to obtain further insights on this problem is by dividing the widths over the corresponding prominence values, where a higher value of this entity may indicate that the segment surrounding the peak is semi flat.

$$E = \frac{2 * width}{prominence} \quad (1)$$

Performing such analysis for the peaks we got  $E = 0.86, 8.96, 170.48$ , and  $59.60$  for the peaks located at {1971, 2560, 2865, 3265}. Hence, the  $E$  threshold should indicate that only peaks at {1971 and 2560} are significant and the other peaks are trivial. Doing a similar analysis for the acoustic shadows we obtained the following values:  $E = 5.50, 122.06$ , and  $39.84002.75$ , for valleys/shadows at {2236, 2803, and 3167} which shows, probably, that only the acoustic shadow at 2236 is significant. In any case, the decision of what to use for the threshold value  $E$  is up to the user by which allowing higher values may produce more false positives.

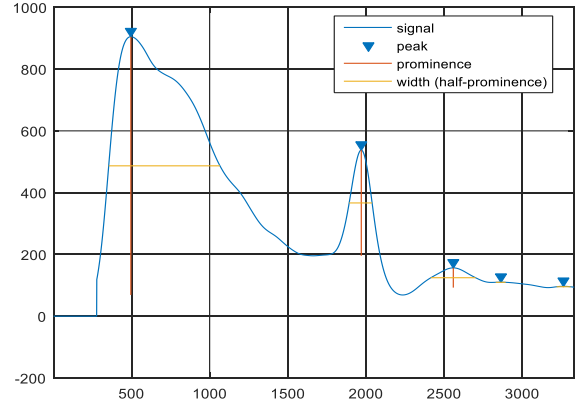


Fig. 7. Detection of peaks in the across-track signal profile. The y-axis denotes the signal strength of the profile of the across-track sample and the x-axis denotes the across-track sample index.

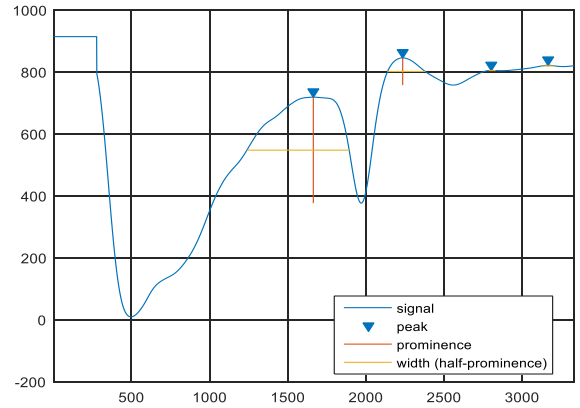


Fig. 8. Detection of valleys/acoustic-shadows for the across-track signal profile shown in Fig. 6, by inverting the across-track signal profile. The peaks and their related widths in this graph are those of the valleys (acoustic-shadows). The y-axis denotes the signal strength of the inverted profile of the across-track sample and the x-axis denotes the across-track sample index.

A threshold value  $E = 10$  to select the significant peaks and shadows has been found to be useful. Out of a total of 9563988 possible points in the right scan image shown in Fig. 3, a total of 425 and 137 peak and shadow points have been detected, respectively. Out of a total of 9563988 possible points in the left scan of the same image there were 9128 peaks and 7513 shadows. After thresholding, however, a total of 591 and 228 peak and shadow points have been detected, respectively. The mean of all  $E$  values in the left scan is  $296.4 \pm 2720$ , however, the

thresholded  $E$  values have the mean  $6.67 \pm 2.27$ , Fig. 9 and Fig. 10 illustrate these values visually. To demonstrate the detection and the identification of the exact locations of the peaks and shadows, we show in Fig. 11 and Fig. 12 the detected landmarks for both the left and the right scan. The total number of detected landmarks (shadows and peaks), and the number of the significant landmarks (thresholded with  $E$ ) for all images is illustrated in Fig. 13.

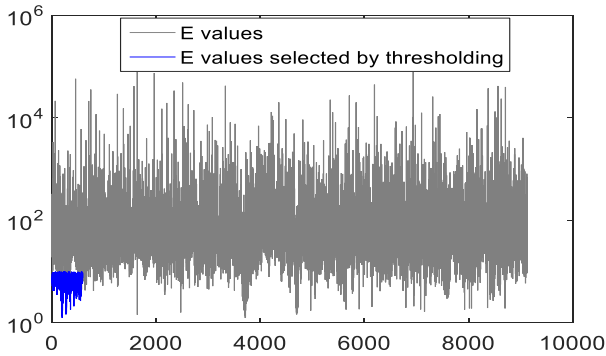


Fig. 9.  $E$  values that measure the significance of the of the left scan of the image shown in Fig. 3. The y-axis denotes the  $E$  values and the x-axis is a trivial value used to indicate the index of the landmark, i.g., first landmark, second landmark, etc.

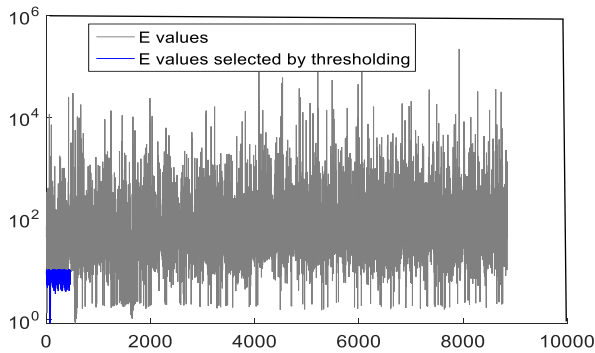


Fig. 10.  $E$  values that measure the significance of the of the right scan of the image shown in Fig. 3. The y-axis denotes the  $E$  values and the x-axis is a trivial value used to indicate the index of the landmark, i.g., first landmark, second landmark, etc.

#### IV. CONCLUSIONS

We note that the proposed method does not detect the landmark from the 2D image, but, the exact location obtained from the peaks/valleys of the cubic smoothing spline fitted on each across-track signal. The use of such method is vital in sidescan sonar images that suffer from signal degradation and various degrees of speckle noise. However, the method still needs further refinements to achieve better results; for example, optimizing the cubic smoothing spline parameter  $P$  and the  $E$  threshold value, as these values control the false positives. The calculated mean of the significant landmarks (via  $E$  threshold) to all detected landmarks ratio, over all images, is  $0.04 \pm 0.02$ . However, the mean of the significant landmarks to the all possible pints ( $2777 \times 34444$ ) ratio, over all images is

$0.00003 \pm 0.00002$ . This clearly indicates that the proposed method reduces the search space needed to detect more vital landmarks compared to the size of the original image. Still, the mean detected landmarks for each image is around 200, which is still high after visual inspection of the images. Hence, a second phase of processing is still needed to refine the number of detected landmarks, e.g., using morphological operators to remove thin landmarks. Further improvements can be achieved by tuning the proposed detection method so that it can be used with the echo-decay normalized signals. Last but not least, deploying all the proposed methods after considering an enhanced speckle filtering of the images may improve the detection of landmarks.

#### ACKNOWLEDGMENT

The research leading to the presented results has been undertaken within the SWARMs European project (Smart and Networking Underwater Robots in Cooperation Meshes; (<http://www.swarms.eu/>), under Grant Agreement n. 662107-SWARMs-ECSEL-2014-1, which is partially supported by the ECSEL JU and the Fundação para a Ciência e a Tecnologia (ECSEL/0002/2014 and ECSEL/0003/2014). The SWARMs project aims at making autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs) further accessible and useful for autonomous maritime operations

#### REFERENCES

- [1] J. Aulinas, A. Fazlollahi, J. Salvi, X. Llado, Y. R. Petillot, J. Sawas and R. Garcia, Robust automatic landmark detection for underwater SLAM using side-scan sonar imaging, Proceedings of the 11th International Conference on Mobile Robots and Competitions, pp. 21-26, 2011.
- [2] P. Craven and G. Wahba, "Smoothing Noisy Data with Spline Functions", Numerische Mathematik, vol. 31, pp. 377-403, 1979.
- [3] P. Y. Mignotte, M. Lianantonakis, and Y. Petillot. Unsupervised registration of textured images: applications to side-scan sonar. In Europe Oceans 2005, vol. 1, pp. 622-627 Vol. 1, June 2005.
- [4] R. Michalec and C. Pradalier. Sidescan sonar aided inertial drift compensation in autonomous underwater vehicles. In 2014 Oceans - St. John's, pp. 1-5, Sept. 2014.
- [5] B. Zerr, G. Mailfert, A. Bertholom, and H. Ayreault. Sidescan sonar image processing for AUV navigation. In Europe Oceans 2005, vol. 1, pp. 124-130 Vol. 1, June 2005.
- [6] J. Guo, W. H. Wang, S. W. Huang, E. Chen, and F. C. Chiu. Map uncertainties for unmanned underwater vehicle navigation using sidescan sonar. In OCEANS 2010 IEEE - Sydney, pp. 1-10, May 2010.
- [7] M. D. Aykin and S. Negahdaripour. On feature extraction and region matching for forward scan sonar imaging. In 2012 Oceans, pp. 1-9, Oct. 2012.
- [8] J. Padial, S. Dektor, and S. M. Rock. Correlation of imaging sonar acoustic shadows and bathymetry for ROV terrain-relative localization. In OCEANS 2014 - TAIPEI, pp. 1-10, Apr. 2014.
- [9] A. Spears, A. M. Howard, M. West, and T. Collins. Acoustic sonar and video sensor fusion for landmark detection in an under-ice environment. In Oceans-St. John's, 2014, pp. 1-8. IEEE, 2014.
- [10] K. Siantidis. Side scan sonar based onboard SLAM system for autonomous underwater vehicles. In 2016 IEEE/OES Autonomous Underwater Vehicles (AUV), pp. 195-200, Nov. 2016.
- [11] S. Stalder, H. Bleuler, and T. Ura. Terrain-based navigation for underwater vehicles using side scan sonar images. In OCEANS 2008, pp. 1-3, Sept. 2008.
- [12] E. Chen and J. Guo. Real time map generation using sidescan sonar scanlines for unmanned underwater vehicles. Ocean Engineering, vol. 91: pp. 252-262, Nov. 2014.

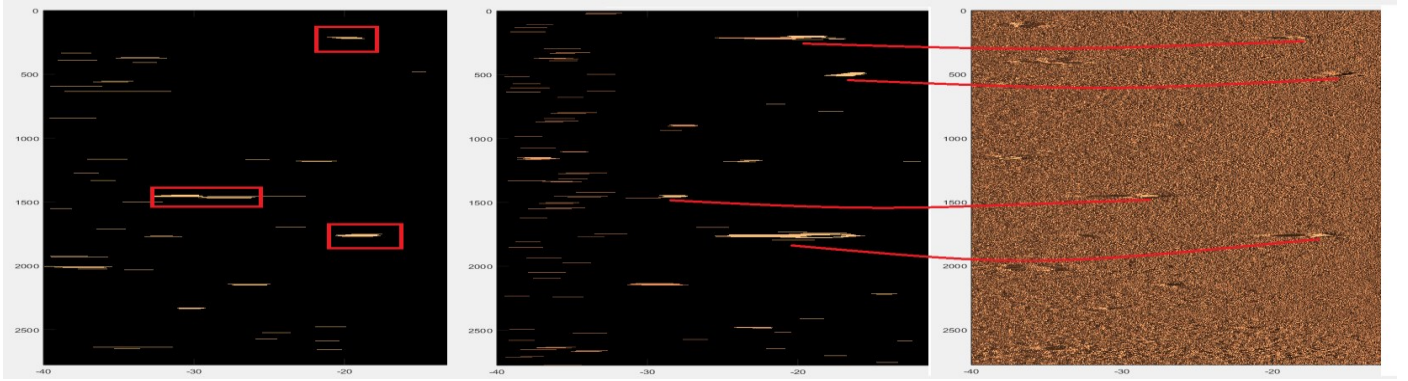


Fig. 11. The peaks and shadows that are detected by the proposed method. The left sonar image, which has been enhanced, is shown to the right, the peaks shown in the middle, and the shadows are shown to the left. The detected landmarks appear brighter than other outlier (false) objects since their values, representing the max peak/shadow value, are higher.

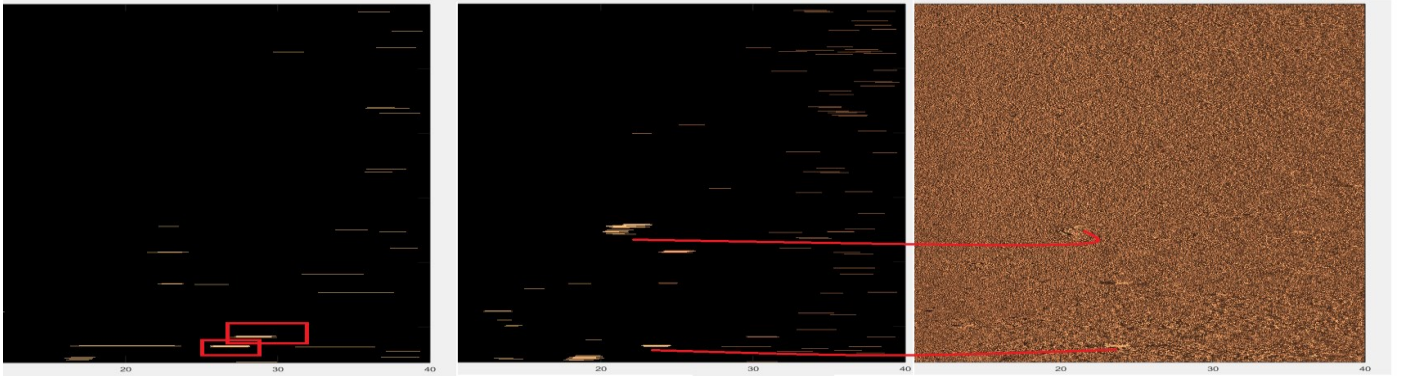


Fig. 12. The peaks and shadows that are detected by the proposed method. The right sonar image, which has been enhanced, is shown to the right, the peaks shown in the middle, and the shadows are shown to the left. The detected landmarks appear brighter than other outlier (false) objects since their values, representing the max peak/shadow value, are higher. The landmark width seems correctly approximated by the width of the peak in the center and left images.

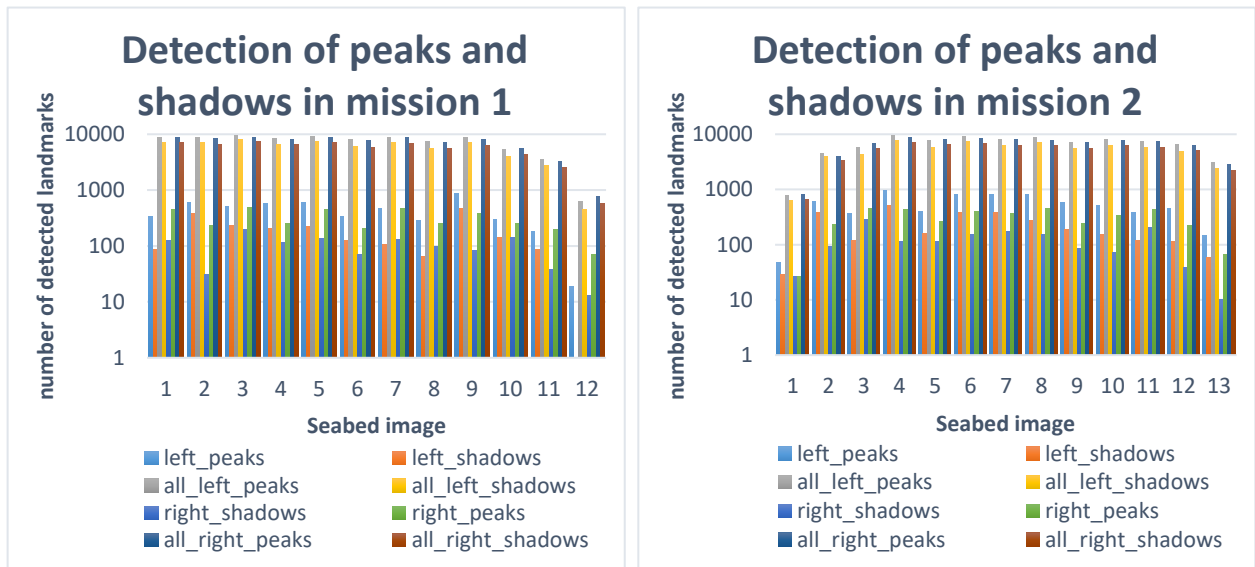


Fig. 13. Peaks and shadows detected for the 25 seabed images acquired during the two missions performed in the demonstration trial of the SWARMs project in Las Palmas de Gran Canaria. All\_left\_peaks are all the detected peaks, at the left scan, before thresholding with the  $E$  value; and left\_peaks are those obtained after thresholding with  $E$ . The same argument applies to shadows and the right scan.