

# Evaluation of an Efficient Approach for Target Tracking from Acoustic Imagery for the Perception System of an Autonomous Underwater Vehicle

Regular Paper

Sebastián A. Villar<sup>1</sup>, Gerardo G. Acosta<sup>1,3,\*</sup>, André L. Sousa<sup>2,3</sup> and Alejandro Rozenfeld<sup>1,3</sup>

<sup>1</sup> INTELYMEC-CIFICEN-CONICET Group, Faculty of Engineering, National University of Central Buenos Aires Province, Buenos Aires, Olavarría, Argentina

<sup>2</sup> Federal University of Bahia, Salvador, Brazil

<sup>3</sup> GEE - Department of Physics, University of the Balearic Islands, Spain

\* Corresponding author E-mail: gerardo.acosta@ieee.org

Received 17 Jul 2012; Accepted 30 Aug 2013

DOI: 10.5772/56954

© 2014 The Author(s). Licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract** This article describes the core algorithms of the perception system to be included within an autonomous underwater vehicle (AUV). This perception system is based on the acoustic data acquired from side scan sonar (SSS). These data should be processed in an efficient time, so that the perception system is able to detect and recognize a predefined target. This detection and recognition outcome is therefore an important piece of knowledge for the AUVs dynamic mission planner (DMP). Effectively, the DMP should propose different trajectories, navigation depths and other parameters that will change the robot's behaviour according to the perception system output. Hence, the time in which to make a decision is critical in order to assure safe robot operation and to acquire good quality data; consequently, the efficiency of the on-line image processing from acoustic data is a key issue.

Current techniques for acoustic data processing are time and computationally intensive. Hence, it was decided to

process data coming from a SSS using a technique that is used for radars, due to its efficiency and its amenability to on-line processing. The engineering problem to solve in this case was underwater pipeline tracking for routine inspections in the off-shore industry. Then, an automatic oil pipeline detection system was developed borrowing techniques from the processing of radar measurements. The radar technique is known as Cell Average – Constant False Alarm Rate (CA – CFAR). With a slight variation of the algorithms underlying this radar technique, which consisted of the previous accumulation of partial sums, a great improvement in computing time and effort was achieved. Finally, a comparison with previous approaches over images acquired with a SSS from a vessel in the Salvador de Bahia bay in Brazil showed the feasibility of using this on-board technique for AUV perception.

**Keywords** Autonomous Underwater Vehicles, Side Scan Sonar, Acoustic Image Processing

## 1. Introduction

Perception is one of the key issues in autonomous robotics. It usually involves robot self-perception (position, attitude, remaining energy, faulty situations), as well as perception of the environment (obstacle avoidance, mapping, objects, special waypoints). Hence, the perception system is essential for the robot to succeed in executing any field mission. Particularly in the hostile and unknown underwater world, a high quality perception system is necessary in order to build an AUV robust enough to withstand the main oceanic perturbations. Other important and necessary systems are the dynamic mission planner and the guidance and control systems [1-6].

The use of AUVs has been growing in the last decade, as they are a good tool for the sustainable exploitation of oceanic resources, for example, exploration in the deeper seas. Missions like underwater pipeline inspections and maintenance, prospection studies, mine detection, debris or other object recognition are among the preferred automated tasks to be developed for modern AUVs [7-9]. As seen in the literature, the technology for such task automation has shown a strong improvement in three main areas: 1) AUV technology; 2) perception devices for the underwater world, i.e., SONAR (Sound Navigation And Ranging); 3) novel acoustic image processing techniques. Regarding point (1), AUVs have undergone great improvement regarding constructive aspects and new materials, control algorithms and powerful computation tools [1-2]; [5]; [8-9]. For point (2), many devices like the multi-beam echo-sounder (MBE), side scan sonar (SSS) and synthetic aperture sonar (SAS) appear able to acquire high resolution data [10-11]. SSS is preferred due to its very good quality/cost trade-off. It has been tested in deep water conditions and is one of the most adequate choices for the detection task in underwater environments. The conventional SSS provide lines of acoustic pulses that vary from 200 to 2000 samples. Note also that the bigger quantity of samples implies more computational effort. Finally, with respect to point (3), there is still a great deal of work to be done. In effect, while AUV and sonar technology is mature enough for the aforementioned automated tasks and even though many approaches of acoustic images processing are currently available, they still require a strong on-line computational effort to achieve self and environmental perception. These acoustic image processing approaches can be analysed from different points of view regarding their speed, efficiency, resources needed, precision and robustness [12-24].

Sonar and radar (Radio Detection And Ranging) technologies share similar features in their processing. In addition, radars are used to detect and recognize vehicles with faster dynamics, like airplanes, dealing also with electromagnetic waves that are faster than acoustic ones

[25-26]. The key concept of the present approach is to migrate radar techniques to sonar acoustic data processing. A group of target detection techniques widely used in radar technology is known as CFAR (Constant False Alarm Rate), described in detail in [25]. This group of techniques maintain a constant false alarm rate computed from the last  $n$  samples of the digitalized echoes power, also known as interference power. In this way, an adaptive detection threshold is adjusted to maintain a probability of expected false alarm ( $P_{fa}$ ) by estimating the average of the interference power values of the adjacent  $n$  cells. This approximation is called Cell Averaging-Constant False Alarm Rate, or CA-CFAR for short [27].

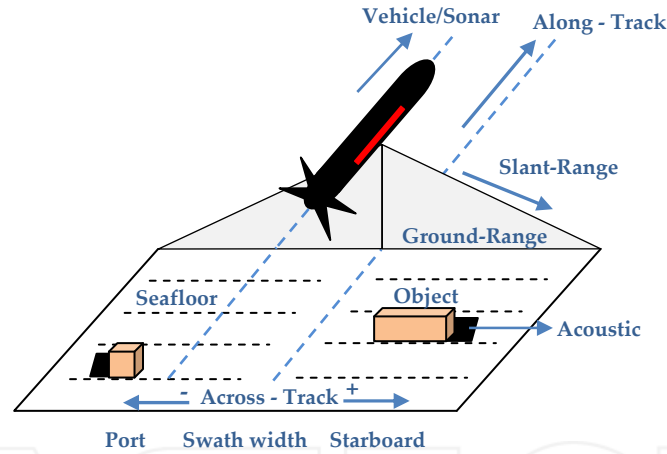
Underwater pipeline and cable tracking is an interesting case study of AUV application with intensive on-board image processing for automatic and autonomous task development. To fulfil this objective, it is necessary to detect the pipeline first, then track it while obtaining other useful information like the pipeline situation (if buried with free-span, with corrosion, with near debris and others).

This article will describe in detail a CA-CFAR based algorithm for the acoustic image processing of SSS data for quantitative analysis of its feasibility to on-line processing. The main objective is to determine if it is efficient enough to be used for on-board and on-line processing in an AUV as an essential input for the AUV's dynamic mission planner. Using a set of data taken from SSS acoustic images of the seafloor of Salvador de Bahia, Brazil, it will be shown that with a refinement in computation, CA-CFAR could make a drastic reduction in time and computational resources.

This work is organized in the following way: section 2 shows the acoustic input data formation. Then, in section 3, an automatic processing chain is presented for a pipeline detection system, focusing on each of the processes. Section 4 shows the basic concepts of the detection theory with CA-CFAR and accumulated CA-CFAR. In section 5, the experimental result, analyses and comparisons with traditional CA-CFAR [27] and partial sums CA-CFAR [28] are presented. Finally, section 6 discusses the conclusions obtained from the work.

## 2. Acoustic Image Forming from a SSS

The SSS is a very interesting tool for high-resolution mapping of the seabed due to its excellent cost/quality trade-off [10]; [46]. It has been tested in deep water with satisfactory results [29-32]. Though SAS provide higher quality imagery and has been used in numerous works [33-37], it is not yet clear that it is better for automated target detection and recognition purposes. Reports about the use of MBE to explore the sea floor in detail are also given in [29]; [38-40].



**Figure 1.** Idealized operation of a Side Scan Sonar on board of an AUV

The SSS is formed by a group of transducers that are mounted on both sides of the AUV. In each data acquisition cycle, these transducers scan sideways and downward, constituting a plane that advances in the direction in which the vehicle travels, the *along-track*. The direction that is perpendicular to the vehicle's straight movement is called *across-track*. Figure 1 shows an idealized representation of the operation of a SSS mounted on an AUV. The transducers on both sides of the sonar send out oblique acoustic signals in the shape of a fan. These acoustic pulses normally oscillate between 100 and 500 kHz. The port side (left) and starboard (right) sides of the images are scanned separately. The acoustic pulses travel through the water column, hit the seafloor and the echo, also named *backscattering*, is returned to the reception sensor where its amplitude is quantified. This amplitude depends on the angle of incidence and the cover of the seafloor. The echoes coming directly from the seafloor constitute the true returned signal. There are also multiple bounces off the seafloor or the sea surface that constitute reverberation or undesired echoes (multi-path). The regions under (nadir) and above (zenith) the sonar correspond to points of low and high reflection off the surface of the seafloor and the surface of the sea, respectively.

The data acquired are projected on a line traced along the seafloor. This scanning line is known as a *swath*. The acoustic data associated with this exploration line represents an observation of the reflected intensity depending on the range of the SSS and the relative angle between the AUV and the seafloor. If the vehicle is moving in a straight line at a steady speed, the deployment of successive swaths will build an acoustic image of the seafloor [11].

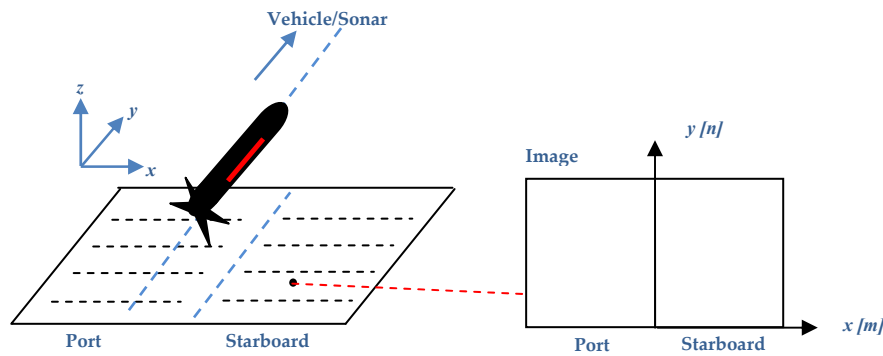
### 3. Underwater pipeline detection system

An *automatic processing chain* is applied to each of the acoustic lines acquired by the SSS. This processing chain consists of a group of serial processes. The input to one

process is the output of the previous process. Figure 3 shows a block diagram of the simple processing chain utilized in the implemented detection system. As shown, the inputs to the whole processing chain are acoustic lines or swaths provided by the SSS and its output is a list of geo referenced (NED) coordinates of the pipeline position.

Before applying the target automatic detection processes, the acoustic data are pre-processed with the objective of improving the input to the detection process.

The first process consists of *geometrical correction* from the distortion caused by the inclination. The SSS acoustic images are prone to numerous unexpected problems, geometrical and natural, which interfere in the detection process [11]; [41-42]. The distortion by inclination corresponds to differences between the relative position of the characteristics of the acoustic image and the actual pipeline location on the sea floor. This distortion is overcome by a process of corrections that have two adjustments, one on the *across-track* direction and another on the *along-track* direction. To carry out this correction, a simple trigonometric relation is applied [10] utilizing navigation data such as altitude, the slant-range and the angle of incidence directly proportional to the *ground-range* (see Figure 1). There are also other sources of distortion that should be considered, such as the AUV's attitude or the water salinity. These factors were not taken into account in this first approach. Hence, the following strong suppositions were assumed for this research [43] for a complete correction: 1) the seafloor is plane and horizontal; 2) the acoustic pulse is propagated through the water at a constant speed; 3) the roll angle of the vehicle is null, because it does not contribute to the geometric distortions; 4) the vehicle is immobile from the moment the acoustic pulse is emitted until the return at maximum range is received. Even though these assumptions are rarely fully satisfied, the correction at this stage yields a much better image for continuing the subsequent processing.



**Figure 2.** Coordinate systems defined on the seabed and on the image. The coordinate system defined on the seabed plane may, to align the  $y$  direction with the trajectory of the vehicle [43].

Thus, an acoustic image is defined as a function of two dimensions of discrete finite values  $s[m, n]$ , where  $[m, n]$  are the coordinates<sup>1</sup> of the image matrix [15]. This intensity, or backscatter force of the sea floor, is defined as  $b(x, y)$ , where  $(x)$  and  $(y)$  are part of a system of rectangular coordinates  $(x, y, z)$  defined of the sea floor, as illustrated in Figure 2 [43].

This coordinates system is defined as follows:  $y$  is aligned with the along-track direction and positive  $x$  represents the starboard across-track direction. Denoting by  $(\bar{s}[l, n])$  the original image and by  $(h[n])$  the height of the water column in pixels at the  $n^{\text{th}}$  line, it can be stated that:

$$s[m, n] = \bar{s}[l, n] \text{ such that } l = \sqrt{h^2[n] + m^2} \quad (1)$$

For  $m = 0, \pm 1, \dots, \pm(N_m - 1)$  and  $n = 0, 1, \dots, N_n - 1$ , where  $(N_n)$  is the maximum number of along-track lines in the slant-range corrected image and  $(N_m)$  is the maximum number of pixels per across-track line. Since the value of  $(l)$  corresponding to  $(m)$  will in general be non-integer; this equation assumes the use of an appropriate technique for interpolating the lines of the original image at non-integer coordinates. Therefore, a linear interpolation was used for this work.

The next process consists of the **elimination of irrelevant information**. The acoustic images contain, in the centre, a black track, which is a *blind spot* corresponding to the nadir, which is inherent to this type of sensor. This irrelevant information must be substituted, because it generates a high contrast zone in the image. This will surely generate a false detection in next processing stage. To avoid this, in each acoustic line, the greatest shadow limit is detected on both port and starboard side. Additionally, it is positioned in the centre of the sonar line and is traversed both on the right and the left until the greatest sharp variation of shades is found. Finally, a

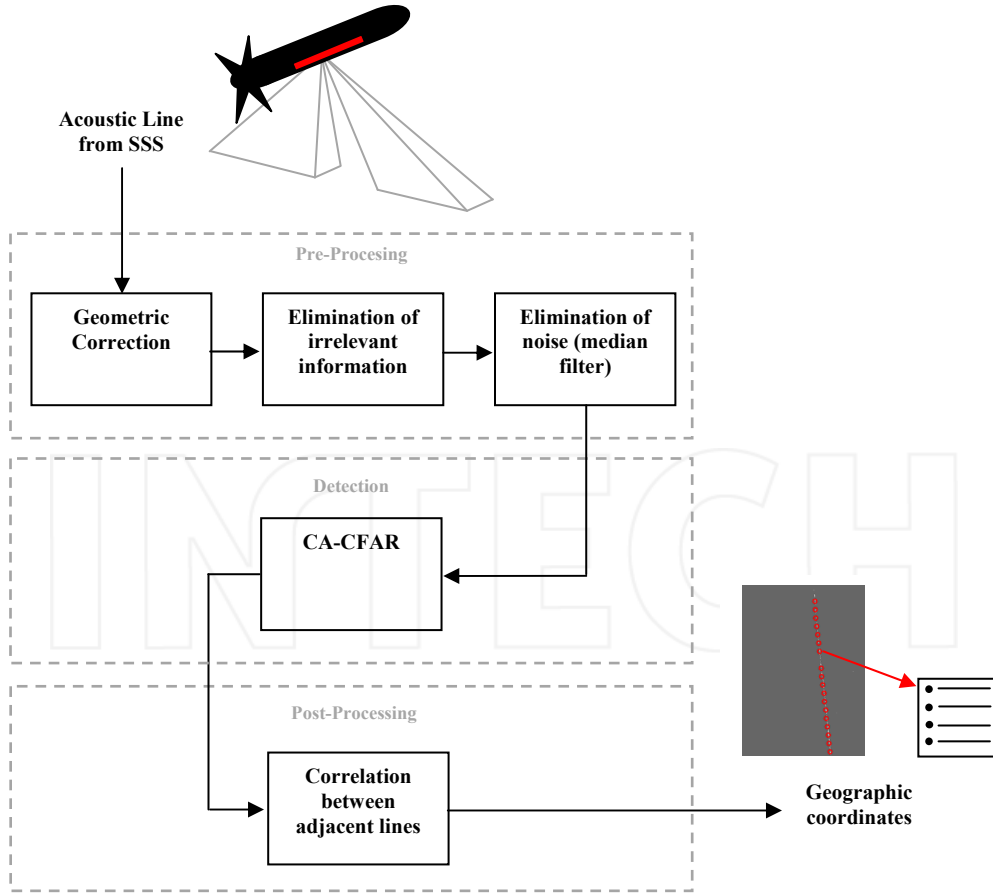
threshold value is calculated, which represents the brightness of the blind spot limit. Thus, the limits for each acoustic line are obtained, a process in which no processing will be carried out.

Another important issue is **image enhancement**. There exists extensive literature about this factor [15-16]; [41]. However, most of the traditional techniques for image enhancement are not adequate enough or cannot be directly migrated to acoustic image processing. Thus, it must be determined for each particular case if the application of this process is particularly useful. For the present approach, very good results were obtained without resorting to this processing stage.

The next step, shown in Figure 3, is **automatic detection**. This consists of labelling the image pixels, classifying the acoustic intensity around a discriminating threshold. In this way, pixels with acoustic intensity above this threshold are labelled with a saturating value (255 in an 8 bits quantization) and minimum value (0 in an 8 bits quantization) if they are below it [46]. As this automatic detection within the processing chain described is a core contribution of this work, it will be explained in further detail in the next section.

The final step in this processing chain is the **correlation between adjacent lines**, which mainly consists of false detections removal, and consequently determining the target's position. This is achieved through a normalized correlation of a set of  $(p)$  preprocessed swaths. The parameter  $(p)$  depended on the technological and physical features of the application. Then, this sub-image  $(pxm)$  was processed as follows. The summation of the  $(p)$  pixels was computed for each column. The maximum value for this summation in the position  $m_j, j \in [0, N_m - 1]$  represented the position of the maximum spatial correlation. The geo-referenced coordinates of this position were the real coordinates for the target, removing the need for considering other spurious pixels in the image analysis.

<sup>1</sup> The square brackets are used to indicate that  $m$  and  $n$  are discrete.



**Figure 3.** Automated processing chain for the detection of the pipeline position coordinates

With two of these geo-referenced points as neighbours, a vector was constructed. This vector pointed from the older geo-referenced detection point to the more recent, successive one. The vector was given as a reference to the guidance system of the AUV.

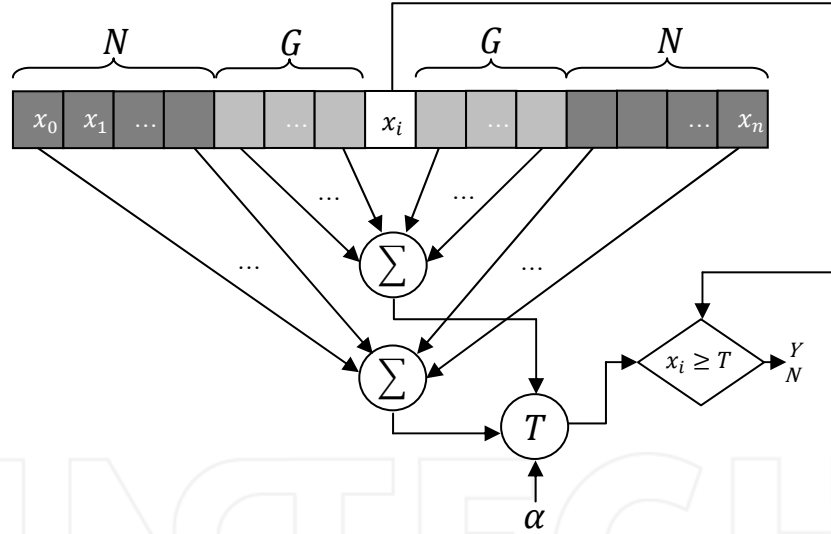
#### 4. Automatic Detection using CA-CFAR

The problem of detection was summed up by analysing each sample with the purpose of detecting the presence or absence of a target. Detection techniques are generally implemented in analysing the information of adjacent samples. In [27], two hypotheses were defined for this analysis: 1) *the sample is the result of interference* ( $H_0$ ) in this case, acoustic reverberation; 2) *the sample was the result of a combination between interference and echoes of a target* ( $H_1$ ), in this case reverberation and backscattering, respectively. Consequently, the detection consisted of examining each sample and selecting one of the above two hypotheses as best fitting. If the hypothesis  $H_0$  was the most appropriate, the detection system declared that the target was not present. On the other hand, if hypothesis  $H_1$  was the most appropriate, the detection system declared that the target was present. Due to the signals being described statistically, the choice between these two hypotheses represents an exercise in statistics decision theory [44].

In the particular case of acoustic images, it was assumed that man-made structures on the sea floor were usually more reflective than the surrounding sediment [46]. For this reason, one of the detection alternatives was centred on finding the backscattering maximum intensities, also called the *acoustic highlight*, which varies considerably according to the relative sonar orientation the target. In fact, it can fall below the detection threshold, causing the target to appear invisible to the sonar.

On the other hand, an additional relevant characteristic of SSS images is that the objects that stand out above the seafloor generate shadows; that is to say, areas where the echo intensity is frequently lower than the level coming from the seafloor. Shadow length depends on the vertical height of the object. Thus, there are other detection alternatives that utilize these shadows. Due to the data being acquired from a moving vehicle, the sonar geometry as it concerns the target was variable. In this case, a shadow can be present even when the acoustic highlight is not. Thus, it is desirable to combine both detection approximations, as is proposed in this work.





**Figure 4.** Generic architecture for the CA-CFAR detection process [45]

The dark grey cells in Figure 4 represent the neighbouring data, which will be averaged to estimate the noise parameters. These cells are the *reference cells* ( $N$ ). Note that each cell represents one pixel. Also in Figure 4, a file vector of  $(1 \times n)$  cells is depicted. The length of this file vector depends on the resolution of the SSS. The lighter grey cells, immediately next to the test cell ( $x_i$ ), are called *guard cells* ( $G$ ). These cells are excluded from the average. The reason for this is that if the target is present, then the neighbouring cells will contain similar values. In this case, the acoustic highlight in the cells surrounding  $x_i$  should contain the same values of acoustic intensity and should not be representative only by its own value. The increase in acoustic highlight of the target should tend to increase the estimation of the reverberation parameters.

The total number of reference and guard cells is calculated utilizing the equations (2) and (3), with  $G < N$  (see Figure 4):

$$N_c = 2N + 1 \quad (2)$$

$$G_c = 2G + 1 \quad (3)$$

The procedure for determining the detection threshold ( $T$ ) is described below. Let us consider the case of a Gaussian reverberation with a square law detector. The *probability density function* (pdf) for any cell  $x_i \geq 0$  has only one free parameter, which is the *mean of the reverberation power* ( $\beta^2$ ). Likewise, the process estimates the mean of the reverberation power in the test cell using the adjacent cells' data, using the following expression:

$$f_{\bar{x}_i}(x_i) = \frac{1}{\beta^2} e^{-x_i/\beta^2} \text{ with } x_i \geq 0 \quad (4)$$

It is supposed that the content of ( $N_c$ ) cells, which are neighbouring ones to the cell under test  $x_i$ , will be used to estimate ( $\beta^2$ ). Another supposition is that reverberations are *independent and identically distributed* (i.i.d). Then, the

joint probability density function  $f_{\bar{x}_i}$  for a vector  $\bar{x}_i = (x_1, x_2, \dots, x_{N_c})$  of neighboring cells ( $N_c$ ) is:

$$f_{\bar{x}_i}(x_1, x_2, \dots, x_{N_c}) = \prod_{i=1}^{N_c} f_{x_i}(x_i)$$

Using (4):

$$\begin{aligned} f_{\bar{x}_i}(x_1, x_2, \dots, x_{N_c}) &= \prod_{i=1}^{N_c} \frac{1}{\beta^2} * e^{-x_i/\beta^2} \\ &= \frac{1}{\beta^{2N_c}} * e^{-\frac{1}{\beta^2} * (\sum_{i=1}^{N_c} x_i)} \end{aligned} \quad (5)$$

The equation (5) is the likely function  $\Lambda$  for the vector of observed data  $\bar{x}$ . The maximum estimated likelihood (MEL) of ( $\beta^2$ ) is obtained by maximizing the equation (5) with respect to ( $\beta^2$ ) [44]. Mathematically, it is equivalent to and generally easier to maximize the log-likelihood function thus [27]:

$$\ln \Lambda = -N_c \ln(\beta^2) - \frac{1}{\beta^2} \left( \sum_{i=1}^{N_c} x_i \right) \quad (6)$$

Deriving equation (6) with respect to ( $\beta^2$ ) and equating it to 0 yields:

$$\widehat{\beta^2} = \frac{1}{N_c} \sum_{i=1}^{N_c} x_i \quad (7)$$

The detection threshold ( $\hat{T}$ ) required is estimated as a scalar multiple  $\alpha > 0$  of the reverberation power:

$$\hat{T} = \alpha \widehat{\beta^2} \quad (8)$$

An adaptive threshold can be considered at a constant rate or probability of false alarm; however, the reverberation levels will vary. The threshold ( $\hat{T}$ ) and the

probability of false alarm ( $P_{fa}$ ) are random variables. The CFAR detector is considered if the value of the probability of false alarm does not depend on the current value of ( $\beta^2$ ). Combining equations (7) and (8) yields the expression for the estimated threshold:

$$\hat{T} = \frac{\alpha}{N_c} \sum_{i=1}^{N_c} x_i \quad (9)$$

Defining  $z_i = \left(\frac{\alpha}{N_c}\right) x_i$ ; such that  $\hat{T} = \sum_{i=1}^{N_c} z_i$ , and using the standard result of the probability theory with equation (4) yields the pdf of  $z_i$ :

$$p_{z_i}(z_i) = \frac{N_c}{\alpha\beta^2} e^{-\frac{N_c z_i}{\alpha\beta^2}} \text{ with } z_i \geq 0 \text{ that is exponential } \left(\frac{N_c}{\alpha\beta^2}\right) \quad (10)$$

This pdf of ( $\hat{T}$ ) is known as the *Erlang density* with parameters ( $N_c$ ) and ( $\frac{N_c}{\alpha\beta^2}$ ):

$$p_{\hat{T}}(\hat{T}) = \left(\frac{N_c}{\alpha\beta^2}\right)^{N_c} \frac{\hat{T}^{N_c-1}}{(N_c-1)!} e^{-N_c \hat{T}/\alpha\beta^2} \quad (11)$$

The observed ( $P_{fa}$ ) with the estimated threshold will be  $\exp(-\hat{T}/\beta^2)$ , which is also a random variable. Its expected value was computed as:

$$\bar{P}_{fa} = \int_0^\infty e^{-\frac{\hat{T}}{\beta^2}} p_{\hat{T}}(\hat{T}) d\hat{T}$$

$$\bar{P}_{fa} = \left(\frac{N_c}{\alpha\beta^2}\right)^{N_c} \frac{1}{(N_c-1)!} \int_0^\infty \hat{T}^{N_c-1} e^{-|\left(\frac{N_c}{\alpha}\right)+1|\hat{T}/\beta^2} d\hat{T} \quad (12)$$

Completing the standard integral and carrying out some algebraic manipulation, the final result was obtained:

$$\bar{P}_{fa} = \left(1 + \frac{\alpha}{N_c}\right)^{-N_c} \quad (13)$$

For an expected ( $\bar{P}_{fa}$ ), the required value of the multiplier ( $\alpha$ ) is acquired from solving equation (13):

$$\alpha = N_c \left( \bar{P}_{fa}^{-\frac{1}{N_c}} - 1 \right) \quad (14)$$

Note that ( $\bar{P}_{fa}$ ) does not depend on the reverberation power ( $\beta^2$ ), but on the number ( $N_c$ ) of neighbouring cells and the threshold multiplier ( $\alpha$ ). Thus, the technique of cell average exhibits the CFAR behaviour. This is significant, because a drastic reduction of computation times can be obtained, as will be demonstrated experimentally in the following section.

#### 4.1 Accumulated Cell Average Constant False Alarm Rate ACA-CFAR

As demonstrated in [28], it was possible to achieve pipeline detection from acoustic images of a SSS with the

standard CA-CFAR. In addition, a variation of this approach, Partial Sums CA-CFAR, was introduced and tested experimentally. In this work, a refinement of CA-CFAR was introduced and evaluated with field data. It was named ACA-CFAR for Accumulated Cell Average Constant False Alarm Rate. It consisted of a continuous average of the values of cells with which to calculate the threshold ( $T$ ). Within each step, a reference cells window and a guard cells window were taken and averaged using equation (7). With this value, the threshold was estimated and then the algorithm checks for the presence of a target were conducted. In order to perform this computation, it was necessary to define a window of ( $N_c$ ) reference cells that slid over all samples until the process was complete. Consequently, for each estimation of the adaptive threshold for every sample to be analysed, ( $N_c$ ) access to memory for the calculation of the sum of the reference cells and ( $G_c$ ) new access to memory for the guard cells were required. This calculation required considerable computational resources and time to analyse the entire sample data. For this reason, the proposed improvement focused on the calculation of the sum of reference and guard cells (see Figure 4).

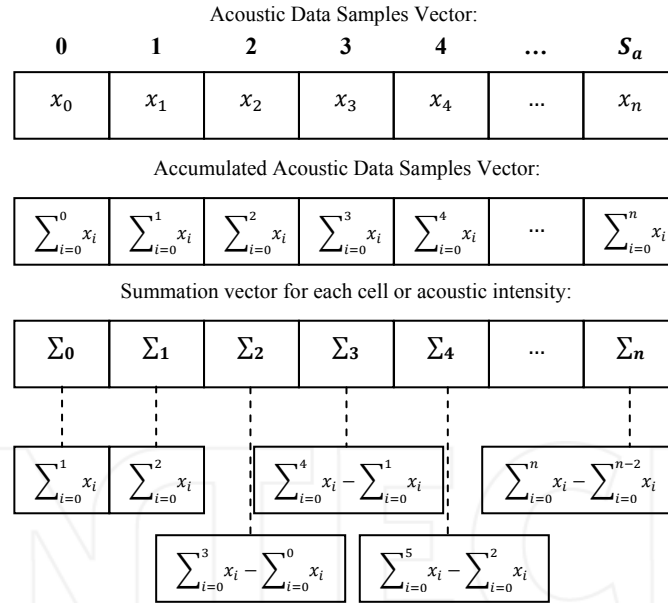
Figure 5 shows an example of a samples distribution ( $x_1, \dots, x_n$ ) and the accumulated summation computation ( $Acc_1, \dots, Acc_n$ ) for the reference cells with  $N = 1$  ( $N_c = 3$ ;  $G_c = 0$ ; please refer to eq. (2) and (3)). The ACA-CFAR technique computed a vector beforehand of absolute accumulated samples ( $\sum_{i=0}^0 x_i, \dots, \sum_{i=0}^n x_i$ ) for each cell by resorting to the following equation (15):

$$Acc_j = \sum_{i=0}^j x_i \quad (15)$$

$$\Sigma_j = \sum_{i=0}^{j+N} x_i - \sum_{i=0}^{j-(N+1)} x_i = Acc_{j+N} - Acc_{j-(N+1)} \quad (16)$$

In this way, the summation for each cell ( $\Sigma_j$ ) or the acoustic intensity value for each cell could be computed using eq. (16). In other words, the acoustic intensity value for each cell was the sum of the accumulated sample values at the present position ( $j$ ) plus the number of reference cells ( $N$ ), varying the sub index  $i$  from  $i=0$  to  $i=j+N$ , ( $\sum_{i=0}^{j+N} x_i$ ), minus the accumulated sample values at the present position, minus the reference cells plus one ( $\sum_{i=0}^{j-(N+1)} x_i$ ).

Notice should be taken of the initialization of the computation over the data samples vector. When the index in the summation ( $\Sigma_j$ ) is smaller than the number of reference cells ( $j < N$ ), the computation of the absolute accumulated sample of the present position plus the number of reference cells ( $\sum_{i=0}^{j+N} x_i$ ) was needed. Then, the second summation of eq. (16) was null.



**Figure 5.** Computation of the sum of the reference cells for the Accumulated CA-CFAR (with  $N = 1 \rightarrow N_c = 3$ )

From analysing the accumulated CA-CFAR, it can be observed that with only two memory accesses at maximum, the value of the summation for any cell could be obtained. This computation was done prior to threshold estimation and detection checking. This method was identically applied to compute the summation of the guard cells.

## 5. Experimental Results

The algorithms were originally developed with MATLAB and were then ported to code written in C++, taking advantage of the data structure within OpenCV. The algorithms were executed on a PC with a CPU 2GHz Intel(R) Core(TM) 2 Duo and 2GB RAM memory, with Linux OS. The SSS was a StarFish 450F, utilizing advanced digital CHIRP acoustic technology. Even when the AUV's on-board CPU facility was a FitPC-2 with different resources, the experiments consisted in the preliminary phases of comparison studies among different detection approaches. It was expected that the best one would be selected to be ported to the run-time environment at the ICTIOBOT AUV prototype [1], travelling at an almost constant speed of 2m/sec.

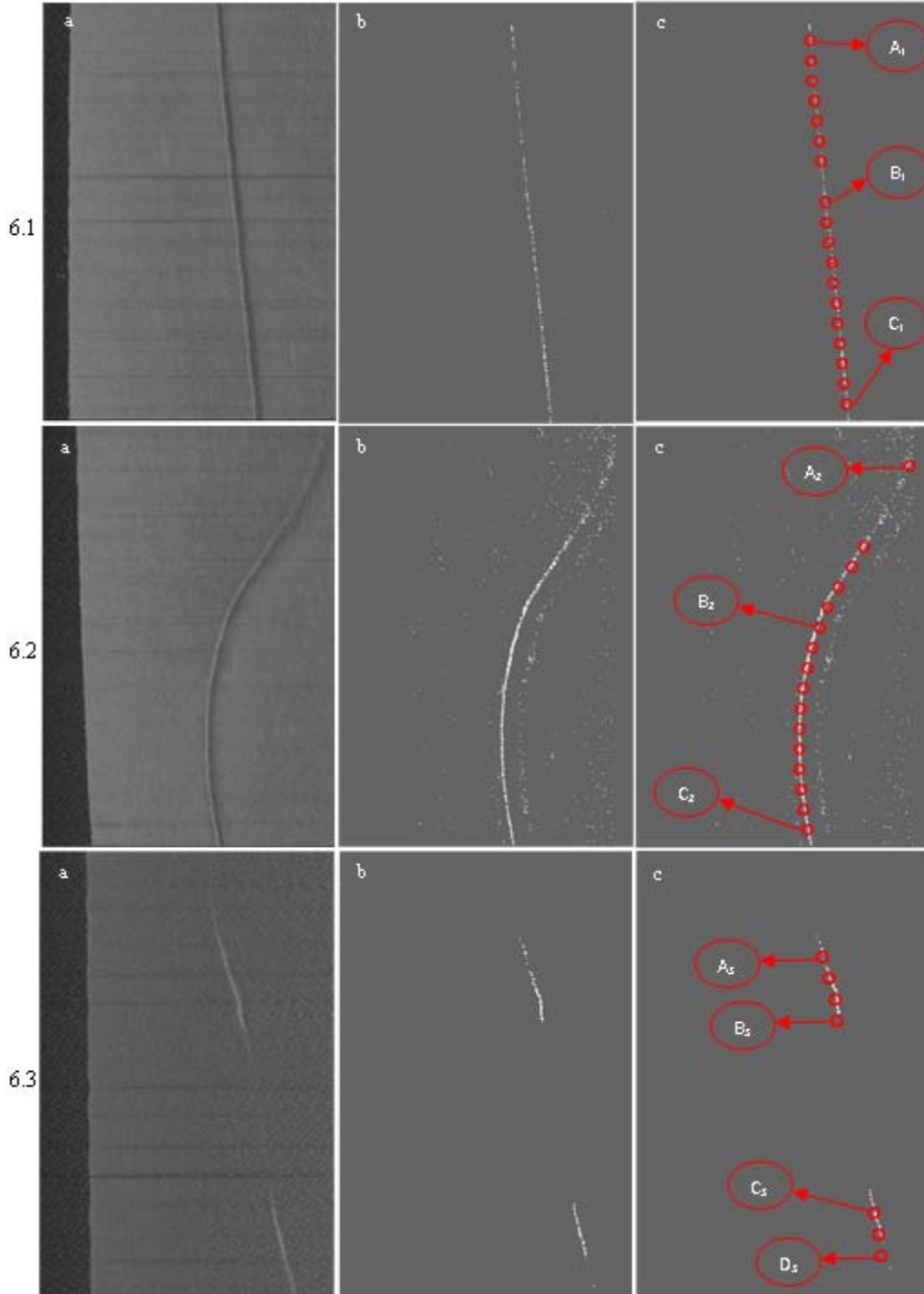
### 5.1 Data

The experimental data employed in this work were acoustic images of a SSS taken from a vessel on the seafloor of Salvador de Bahia, Brazil, where an exposed pipeline has been laid down. For SSS detection, it is necessary that the pipeline be fully or partially exposed. If buried, the perception sensor would have needed a magnetic tracker or a sub-bottom profiler.

The pipeline tracking had two stages: the first was initiated at latitude  $-12^\circ 50' 49,5''$  and longitude  $-38^\circ 31' 23,03''$ , and concluded at latitude  $-12^\circ 51' 33, 28''$  and longitude  $-38^\circ 32' 48,48''$ . 50500 Lines of valid acoustic data were collected, yielding 101 images at 1000x500 pixels for testing the algorithms. The second stage, started at latitude  $-12^\circ 51' 33,04''$  and longitude  $-38^\circ 32' 48,14''$ , and concluded in latitude  $-12^\circ 50' 16,1''$  and longitude  $-38^\circ 30' 37,14''$ , collected 47000 lines of acoustic data totalling 94 images of the same size as the ones obtained for the first test stage.

Figure 6 shows three examples of original SSS images in (a) the output after applying this automatic detection in (b) and the final result after making the correlation of adjacent lines in (c). These images have been cropped for better presentation. In each case, the pipeline can be found on the right side of the SSS. In Figure 6.1, a straight and well-defined pipeline can be observed. In Figure 6.2, the pipeline is slightly curved and a lot of sediment has accumulated on top of the image, which may have produced false detections. Figure 6.3 exhibits an intermittent buried pipeline. In Figure (c), a red circle denotes detection points for tracking, obtained by the algorithm. Details about these detection points are also given in Table 1. As can be seen, the result of this automatic detection consists of spatial coordinates (row and column), as well as the absolute latitude and longitude of the acoustic line, then the pipeline position (point detection for tracking).





**Figure 6.** Experimental results for ACA-CACFAR: (a) the original pre-processed image; (b) the detected pipeline; (c) the detection points for tracking from Table 1

### 5.2 Quantitative comparison of algorithm efficiencies

The quantitative measures selected for comparing the different algorithms are the amount of CPU instructions and the execution time in seconds. They are very descriptive for determining an efficient performance for an on-line automatic target detector and tracker.

Eq. (17) is a performance index representing the number of instructions employed for the standard CA-CFAR algorithm, where ( $S_w$ ) and ( $S_a$ ) represent the amount of acoustic lines or swath and the amount of acquired samples, respectively. Also, ( $N$ ) and ( $G$ ) represent the number of reference cells and the number of guard cells.

$$I_{CA-CFAR} = S_w * (S_a - 2 * (N + G + 1)) * (2 * (N + G) + 1) \quad (17)$$

Point Fig. 6	Image Row	Image Column	Acoustic Line Latitude	Acoustic Line Longitude	Pipeline Latitude	Pipeline Longitude
A1	47	857	-12° 51' 12.14''	-38° 32' 4.52''	-12° 51' 2.21''	-38° 32' 10.67''
B1	239	875	-12° 51' 12.38''	-38° 32' 4.95''	-12° 51' 1.69''	-38° 32' 10.94''
C1	479	900	-12° 51' 12.65''	-38° 32' 5.45''	-12° 51' 1.33''	-38° 32' 12''
A2	47	975	-12° 51' 13.75''	-38° 32' 9.16''	-12° 50' 58.67''	-38° 32' 5.7''
B2	215	878	-12° 51' 13.68''	-38° 32' 9.62''	-12° 51' 1.62''	-38° 32' 7.14''
C2	479	853	-12° 51' 13.75''	-38° 32' 10.22''	-12° 51' 3.71''	-38° 32' 15.91''
A3	119	883	-12° 51' 7.7''	-38° 31' 56.13''	-12° 50' 57.65''	-38° 32' 3.64''
B3	191	889	-12° 51' 7.81''	-38° 31' 56.23''	-12° 50' 59.09''	-38° 32' 6.04''
C3	407	939	-12° 51' 8.17''	-38° 31' 56.52''	-12° 50' 57.72''	-38° 32' 6.45''
D3	455	947	-12° 51' 8.23''	-38° 31' 56.6''	-12° 50' 56.13''	-38° 32' 4.82''

**Table 1.** Computed results after applying the automatic detection with ACA-CFAR. This Table contains the detection points for tracking, which are shown in Figure 6: space coordinates (column 2 and 3), absolute coordinates of the acoustic line (column 4 and 5) and absolute coordinates of the pipeline position (column 6 and 7).

Image	Technique	N	G	$\alpha$	$P_{fa}$	D	I	T (sec.)
1	CA-CFAR	24	6	1.61	.205	428	28609000	1.246
	PSCA-CFAR						15500000	0.617
	ACA-CFAR						500000	0.138
2	CA-CFAR	20	6	1.61	.205	4834	25069000	1.168
	PSCA-CFAR						13500000	0.658
	ACA-CFAR						500000	0.234
3	CA-CFAR	18	6	2.51	.136	220	23275000	0.878
	PSCA-CFAR						12500000	0.527
	ACA-CFAR						500000	0.13

**Table 2.** Data and results of the automatic detection technique, for images of 500x1000 and 500000 samples. Standard CA-CFAR, partial sums PSCA-CFAR and accumulated ACA-CFAR. N: reference cells. G: guard cells. D: number of detections. I: number of CPU instructions (equations 17, 18 and 19).

Equations (18) and (19) show, respectively, the computation of the number of algorithm instructions for partial sums CA-CFAR presented in [28] and the ACA-CFAR introduced in this work:

$$I_{PSCA-CFAR} = S_w * S_a * (N + G + 1) \quad (18)$$

$$I_{ACA-CFAR} = S_w * S_a \quad (19)$$

Note that the performance index of equation (19) is constant for the same image, depending only on the amount of samples ( $S_a$ ).

### 5.3 Comparisons

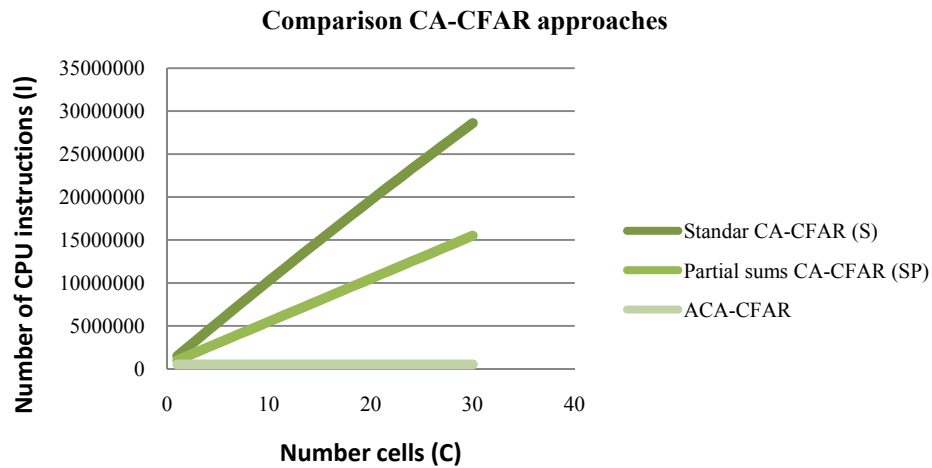
Table 2 shows the settings for the automatic detection process with CA-CFAR, PSCA-CFAR and ACA-CFAR. These all present similar detection results. However, the performance difference regarding the amount of CPU instructions is remarkable.

Analysing Table 2, it can be seen that the improvement of ACA-CFAR introduced in this work is 98,25%  $\approx$  98% for the first image when compared with the standard CA-CFAR technique and of 96,77%  $\approx$  97% when compared with the partial sums CA-CFAR. For the second image,

the reference cells amount decreased and the improvement of ACA-CFAR was 98,01%  $\approx$  98% when compared with the standard CA-CFAR, and about 96,3%  $\approx$  96% when compared with the partial sums CA-CFAR. Finally, for the third image, the improvement of the ACA-CFAR was 97,85%  $\approx$  98% in contrast to the standard technique and 96% better when compared with the partial sum CA-CFAR technique.

A graphical comparison of the algorithms' performance is shown in Figure 7. As can be seen, the ACA-CFAR maintained a constant number of instructions even though the number of reference or guard cells varied. In other words, if the number of neighbouring or contextual cells was increased, this novel technique maintained the same number of CPU instructions, depending only on the sample amount. This is a very significant advantage with regards to previous CFAR techniques, the performance of which does depend on the number of reference or guard cells, which slows down their performance.

The ACA-CFAR algorithmic complexity was  $O(S_w * S_a)$ . Therefore, for a square acoustic image, its complexity was quadratic with value  $O(N^2)$  where the number of samples was  $N^2$ .



**Figure 7.** Comparison of the number of CPU instructions for the three CA-CFAR approaches and for images of 500x1000 and 500000 samples. S: standard CA-CFAR, SP: partial sums CA-CFAR; A: ACA-CFAR. C: number cells. I: number of CPU instructions (equations 17, 18 and 19).

## 6. Conclusions

The main contribution of the work presented here is the proposal of a novel automatic acoustic image processing technique. It was experimentally tested for pipeline detection using acoustic data obtained with a SSS in Salvador da Bahia, Brazil. The image processing technique called cell average constant false alarm rate (CA-CFAR) was borrowed from the radar domain and was strongly improved by changes in the computing algorithm for on-line processing and detection. The accumulated CA-CFAR, or ACA-CFAR for short, gives the same detection results of CA-CFAR, with a significant decrease in the computational effort and time.

This preliminary comparison study was conducted to select the best approach for programming the on-board perception system of the AUV prototype ICTIOBOT. This perception system will be applied to the off-shore industry devoted to pipeline tracking by using images with a higher resolution. These results also showed that it was a good idea to migrate concepts from radar to sonar. The efficient CACFAR image processing technique is a good choice for obtaining on-line and efficient performances also in the acoustic domain.

These features are essential for perception feedback in the dynamic mission planner, the guidance and the control and navigation systems of the aforementioned AUV prototype.

## 7. Acknowledgements

This work was carried out thanks to financing from the following projects: DPI2009-11298, MICINN from Spain, CONICET – PIP 11420090100238 and ANPCyT – PICT-2009-0142 from Argentina. Author Sebastian Villar would

like to acknowledge the IEEE/OES Student Scholarships, which he was awarded for his postgraduate studies.

## 8. References

- [1] Acosta G.G., Curti, H., Calvo Ibáñez, O. and Rossi, S. (2009) *Some Issues on the Design of a Low-Cost Autonomous Underwater Vehicle with an Intelligent Dynamic Mission Planner for Pipeline and Cable Tracking*, Chapter I in *Underwater Vehicles*, pp. 1-18 I-Tech Online Books, Robotics Series, I-Tech Education and Publishing KG, Vienna, Austria, Editor: A. Inzartsev, <Open Access: <http://books.i-tetonline.com/>> (ISBN 978-953-7619-49-7).
- [2] Landay W. E., LeFevre M. A., Spicer R. A., Levitre R. M. and Tomaszewski S. J. (2003). "The USA Navy Unmanned Undersea Vehicle (UUV) master plan". Dept. of the Navy, United States of America.
- [3] Fossen T. (1994). *Guidance and control of Ocean Vehicles*. UK: Wiley Ed.
- [4] Fossen T. (2002). *Marine Control Systems*. Ed. Marine Cybernetics, Norway.
- [5] Acosta, G.G. (2005). "State of the art on trajectory generation for AUV with artificial intelligence techniques". *AUVI Project Internal Report*, 3027-1.
- [6] Acosta G.G., Curti H. and Calvo O. (2005). "Autonomous underwater pipeline inspection in AUTOTRACKER PROJECT: the Navigation Module," *OCEANS 2005 MTS/IEEE*, vol. 1, pp. 389-394.
- [7] Acosta G.G., Curti H., Calvo O. and Rossi S. (2006). "A knowledge-based approach for an AUV path planner development". *WSEAS Transactions on Systems*, vol. 5, no. 6, pp. 1417-1424.
- [8] Balasuriya A. and Ura T. (1999). "Multi-sensor fusion for autonomous underwater cable tracking," *OCEANS 1999 MTS/IEEE*, vol. 1, pp. 209-215.

- [9] Ishøy I., Bjerrum A., Calvo O., Acosta G. G., Petillot Y., Evans J., Kyriakopoulos K., Lionis G., Slater T. and Nunn R. (2002). "New challenges for AUTOTRACKER," *Unmanned Underwater Vehicle Showcase*, Southampton, UK, September 2002.
- [10] Lurton X. (2003). *An introduction to Underwater Acoustics*. Ed. Springer Verlag Inc.
- [11] Chapple P. (2008) "Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations". *Technical report*. Maritime Operations Division. DSTO Defence Science and Technology Organisation.
- [12] Reed S., Ruiz I. T., Capus C. and Petillot Y. (2006). "The fusion of large scale classified side-scan sonar image mosaics". *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2049-2060, July 2006.
- [13] Karoui I., Fablet R., Boucher J. M., Augustin J. M. (2009). "Seabed segmentation using optimized statistics of sonar textures," *IEEE Geosciences and Remote Sensing*, vol. 47, pp. 1621-631.
- [14] Williams D. (2009). "Bayesian data fusion of multiview synthetic aperture sonar imagery for seabed classification". *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1239-1254.
- [15] González R. C. and Woods R. E. (2001). *Digital Image Processing*. (2<sup>nd</sup> ed.). New Jersey: Prentice Hall.
- [16] Pratt W. K. (2001). *Digital Image Processing: PIKS Inside*. (3<sup>rd</sup> ed.).
- [17] Mignotte M., Collet C., Perez P. and Bouthemey P. (1999). "Three-class Markovian segmentation of high-resolution sonar images". *Computer Vision and Image Understanding*, vol. 76, no. 3, pp. 191-204.
- [18] Mignotte M., Collet C., Perez P. and Bouthemey P. (2000). "Sonar image segmentation using an unsupervised hierarchical MRF model". *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216-1231.
- [19] Cutter G. R., Rzhonov Y. and Mayer L. A. (2003). "Automated segmentation of seafloor bathymetry from multibeam echosounder data using local Fourier histogram texture features". *Journal of Experimental Marine Biology and Ecology*, vol. 285-286, pp. 355-370.
- [20] Zhou F., Feng J. F. and Shi Q. Y. (2001). "Texture feature based on local Fourier transform". *International Conference on Image Processing*, vol. 2, pp. 610-613.
- [21] Lianantonakis M. and Petillot Y. (2007). "Sidescan sonar segmentation using texture descriptors and active contours". *IEEE Journal Oceanic Engineering*, vol. 32, no. 3, pp. 744-752.
- [22] Chan T., Sanberg B. and Vese L. (2000). "Active contours without edges for vector-valued images". *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 130-41.
- [23] Ye X. F., Zhang Z. H., Liu P. X. and Guan H. L. (2010). "Sonar image segmentation based on GMRF and level-set models". *Ocean Engineering*, vol. 37, no. 10, pp. 891-901.
- [24] Celik T. and Tjahjadi T. (2011). "A Novel Method for Sidescan Sonar Image Segmentation". *IEEE Journal of Oceanic Engineering*, vol. 36, no. 2, pp. 186-194.
- [25] Barton D. K. and Leonov S. A. (1998). *Radar Technology Encyclopedia*. Artech House, London, UK.
- [26] Hansen R. E. (2009). "Introduction to sonar". Course materiel to INF-GEO4310, University of Oslo, October 2009.
- [27] Richards M. A. (2005). *Fundamental of radar signals processing*. Georgia institute of technology: McGraw Hill Electronic Engineering.
- [28] Villar S. A., Sousa Senna A. L., Rozenfeld A. F. and Acosta G. G. (2013). "Pipeline detection system in acoustic images utilizing CA-CFAR". *Proc. of the MTS/IEEE OES OCEANS 2013*, San Diego, CA, USA. ISBN: 978-0-933957-40-4.
- [29] Petillot Y., Reed S. and Bell J. (2002). "Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder". *OCEANS 2002 MTS/IEEE*, vol. 1, pp. 217-222.
- [30] Reed S., Petillot Y. and J. Bell (2003). "An automatic approach to the detection and extraction of mine features in sidescan sonar". *IEEE Journal of Oceanic Engineering*, vol. 28, pp. 90-105.
- [31] Reed S., Petillot Y. and Bell J. (2004). "Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information". *IEEE Radar, Sonar & Navigation*, vol. 151, no. 1, pp. 48-6.
- [32] Perry S. W. and Guan L. (2004). "A Recurrent Neural Network for Detecting Objects in Sequences of Sector-Scan Sonar Images". *IEEE Journal of Oceanic Engineering*, vol. 29, no. 3, pp. 857-871.
- [33] Bell J. M., Petillot Y. R., Lebart K., Mignotte P. Y. and Rohou H. (2006). "Target recognition in synthetic aperture and high resolution sidescan sonar". *Institution of Engineering & Technology Seminar on High Resolution Imaging and Target Classification*, pp. 99-106.
- [34] Hansen R. E., Groen J. and Callow H. J. (2007). "Image enhancement in synthetic aperture sonar". *Detection & Classification of Underwater Targets*, vol. 29, no. 6, pp. 69-6.
- [35] Sæbø T. O., Callen H. J. and Hansen R. E. (2007). "Bathymetric capabilities of the HISAS interferometric synthetic aperture sonar". *OCEANS 2007 MTS/IEEE*, pp. 1-10.
- [36] Sparr T., Hansen R. E., Callow H. J. and Groen J. (2007). "Enhancing target shadows in SAR images". *Electronics Letters*, vol. 43, no. 5, pp. 69-70.
- [37] Hagen P. E. and Hansen R. E. (2008). "Synthetic aperture sonar challenges". *Hydro International*, vol. 12, no. 4, pp. 26-31.
- [38] Pavin A. M. (2006). "The Pipeline Identification Method Basing on AUV's Echo-Sounder Data". *OCEANS 2006*, pp. 1-6.

- [39] Jalving B., Mandt M., Hagen O. and Pöhner F. (2004). "Terrain referenced navigation of AUVs and submarines using multibeam echo sounders". *UDT Europe*, Nice, France.
- [40] Greenaway S. F. and Weber T. C. (2010). "Test methodology for evaluation of linearity of multibeam echosounder backscatter performance". *OCEANS 2010*, pp. 20-23.
- [41] Chavez, Jr. P.S. (1976). "Processing techniques for digital sonar images from GLORIA". *Photogrammetric Engineering, Remote Sensing*, vol. 52, pp. B.
- [42] Stubbs A. R. (1963). "Identification of patterns of asdic records". *International Hydrographic*, vol. 40, pp. 53-68.
- [43] Cobra D.T. (1992). "Geometric distortions in side-scan sonar images: a procedure for their estimation and correction". *IEEE Journal Oceanic Engineering*, vol. 17, no. 3, pp. 252-268.
- [44] Kay S. M. (1993). *Fundamentals of Statistical Signal Processing, Estimation Theory*. New Jersey: Prentice Hall.
- [45] Skolnik M. I. (1990). *Radar Handbook* (2<sup>nd</sup> ed.). McGraw-Hill, New York, USA.
- [46] Blondel P. H. (2007) *The Handbook of Sidescan Sonar*. Springer-Praxis books in geophysical sciences. Springer Praxis Books, Chichester, UK.

ECH

ECH