EVALUATING HUMAN/MACHINE INTERACTION FOR UNDERWATER THREAT CLASSIFICATION: A CASE STUDY

Olga Lopera Tellez[a],

[a] Department of Communication, Information, Systems and Sensors, Royal Military Academy,1000 Brussels, Belgium

olopera@rma.ac.be

Abstract: *In this paper, expert deminers and automated algorithms are charged with the task of analysing sonar images collected during real mine countermeasures exercises in order to classify targets. Images are collected using synthetic aperture sonar (SAS) and side scan sonar (SSS), covering a test area on the Belgian Continental Shelf. A total of 1241 images (with 847 detection opportunities) collected from different sonar systems, each of them covering the entire area, are used. Image resolution is divided in three categories: (1) up to 5cm pixel size, (2) over 5cm until 10cm pixel size, (3) larger than 10cm pixel size. Data are analysed in different ways by the expert operators and the algorithms. Results demonstrate how challenging underwater threat recognition still is, and highlight the utility of considering the human operator as an integral part of the automatic underwater object recognition process, as well as how automated algorithms can extend and complement human performances.*

Keywords: *mine countermeasures; synthetic aperture sonar; side-scan sonar; automatic target classification; human-in-the-loop*

1. INTRODUCTION

The need to maintain maritime freedom of manoeuvre implies a requirement for maritime capability to survey littoral waters, sea line of communications, choke points, structures, ports and harbours. Currently, these maritime capabilities include high-resolution sonar systems such as side-scan sonar (SSS) and synthetic aperture sonar (SAS) for imaging the seabed and the water column in order to detect and identify underwater objects which might threaten safety of navigation. Although automatic object recognition algorithms have been developed for and are applied in different remote-sensing imaging applications, object-recognition from acoustic/optical imagery in this maritime application is still performed almost exclusively by human operators. There is still a long way to see these automatic algorithms being fully trusted to take the place of human analysts [1]. To facilitate this process, [2] suggests the development of methods by which humans and computers can work in concert to achieve improved performance.

In this paper, human experts and algorithms are charged with the task of analysing the same database of acoustic images. The later are collected by unmanned maritime vehicle systems equipped with SSS and SAS, tasked with naval mine hunting and route surveillance operations during real mine countermeasures (MCM) exercises. The group of human experts is composed by 4 deminers with 7 to 25 years of experience in sonar imaging and 2 scientists with 10 years of experience in sonar imaging. Two classification algorithms are used, one based on Markov Chain Monte Carlo [3] and a second one based on Adaptive Boosting Decision Trees [4]. The test area is located on the Belgian Continental Shelf, between the Thorton bank and the Goote Bank. This test area is selected based on the long term stability of its physical characteristics. Different objects (exercise-mines and friendly objects) are deployed at different locations. Data are recorded using different SAS and SSS systems, and data are categorized by their resolution: (1) up to 5cm pixel size, (2) over 5cm until 10 cm pixel size and (3) larger than 10cm pixel size.

This study is a continuation of the analysis presented in [5,6]. In [5], data collected during object-detection trials (mine hunting trials) are analysed in order to create a connection between seafloor characterisation maps and prediction of object-detection performance. Results demonstrate that trends can be established between environmental parameters and the detection performance (in accordance with the analysis presented in [7]). In [6], a preliminary analysis of the target detection performances of human vs machine was carried out with a limited acoustic image database, and the classification performance was not yet evaluated. This paper goes forward with the evaluation of different human/machine interaction strategies to improve classification performance.

This paper is divided as follows. Section 2 describes the test area based on the analysis of ground truthing measurements presented in [8]. The sonar systems and their data are introduced in Section 3. Section 4 introduces the algorithms used to process and analyse the data and a discussion of the results and further work is presented in Section 5.

2. TEST AREA AND ENVIRONMETAL ASSESSMENT

2.1. Test area description

The area is located between the Goote bank and the Thorton bank, a central area on the Belgian Continental shelf, 17 NM off the Belgian coast (Ostend). It is 2NM² and presents 20 to 30 meter of water depth. At the left side of Fig. 1, the location of this area is shown with a dark green rectangle. Its corresponding bathymetry map is shown at the right side of Fig. 1. In the later, sandwaves with height between 2 to 4 meter and wavelength around 150 meters are recognized. They are distributed Northeast and Northwest of the area. A low flat area is present Southworth between those sandwaves and the nearby Goote bank area.

For the environmental assessment of this area, during a scientific survey with the oceanographic Research Vessel Belgica A962, 32 sediment samples were collected. The analysis of these samples is introduced in the next Session.

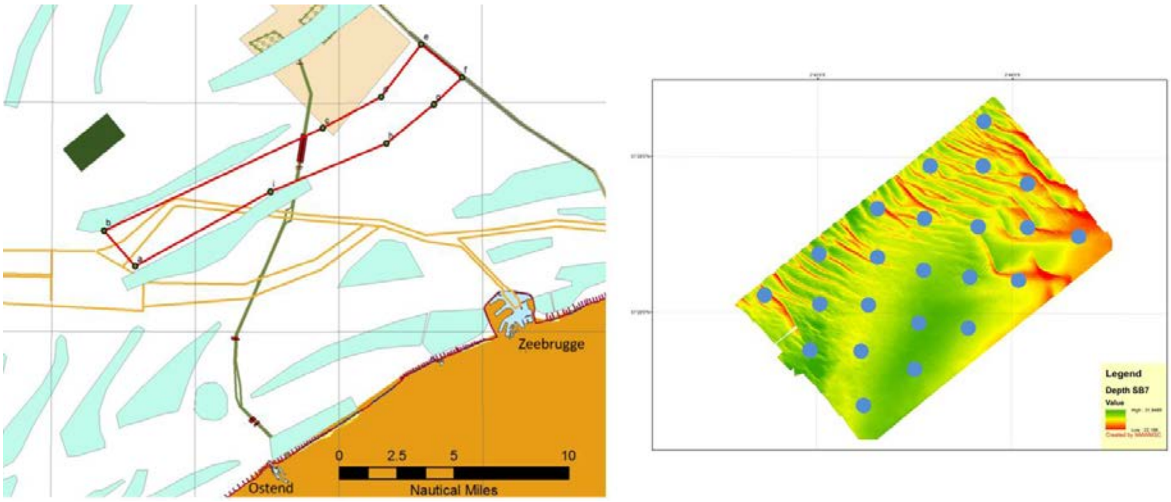

Fig.1: Left: Test area marked by a green box between the Goote bank and Thorton bank. Right: Bathymetric map of the test area, where 32 sampling stations are indicated by blue circles. Distance between the samples varies between 500 and 600m.

2.2. Sampling analysis

In the analysis the collected samples, not only the grains size distribution of the sediments over the area were considered, but also shell distribution, carbonate content distribution, gravel distribution and presence of benthic organisms. Regarding the shells, a distinction was made between entire shells, shell fragments and shell hash. Shell hash is considered being accumulation and abundance of small fragments with dimension less than 4mm.

Considering the shell distribution, abundance of shells is found where bedforms are largely developed, whilst complete absence of shells is noticed at the transition zone between the most western sandwaves and the gully. Shell fragments distribution follows the shells distribution although the fragments are better aligned with the currents directions. Shell hash is mainly founded where abundance of shells is, though a good level of shell hash is found in the gully where nor shells nor shell fragments are present. The hypothesis is that the shell hash in the southern part of the area could be accumulated after being transported by the local currents.

The distribution of the carbonate content is in line with the shell, shell fragments, and shell hash distribution. Regarding tubeworms, they are concentrated where concentration of shell fragments and shell hash is high; this is in response to their nature to construct their main housing with small shell fragments. Echinidae are mainly found where large bedforms are developed and gravels are mainly founded into the gully.

The Grain size analysis of the 32 samples was conducted considering the mixture sand and carbonate into the sediments. Only the organic material was eliminated by chemical treatment. The sediment results had shown a granulometry in a range between 350 and 700 µm corresponding to coarse sand in the Wentworth scale. In the test area the coarseness increases going towards southeast.

The coarseness is accentuated by the presence of the carbonate into the sediments. Proof of that is given by the samples with higher values of granulometry. For those samples, consistent high level of shell hash was found into their mixture.

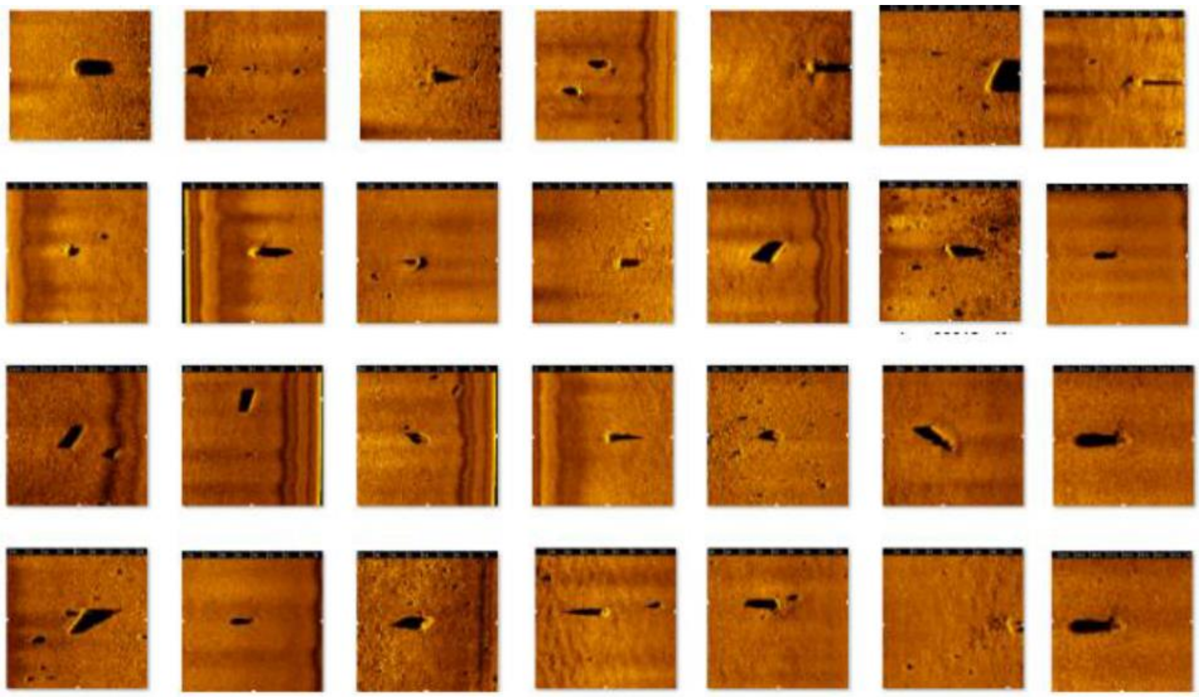

Fig.2: Some thumbnails of SSS data

3. MEASUREMENT SETUP

Data gathered by different SAS and SSS systems are used for this study. The centre frequency of the systems ranges from 100kHz to 300kHz. Some examples of the images provided byone the SSS systems can be seen in Fig. 2.

A total of 31 mine-like objects were deployed (including Manta, Cylindrical, Rockan and Moored exercise mines) on the seabed and in the water column. They were located at different positions distributed all over the test area. In order to give an idea of the dimensions of the objects, their description is presented in Table 1. A few friendly objects were also deployed, including oil barrels and metallic cages. It is worth noting that positions where unknown to the operators and the algorithms during the object recognition process and were released afterwards for the analysis.

Target	Dimensions
Manta mine (MM)	Height: 44 cm Diameter: 98 cm
Cylindrical mine (GM)	Length: 1,60 m Diameter: 50 cm
Rockan mine (RC)	Width : 100 cm Length: 80 cm Height: 40 cm
Moored mine (MM)	Length: 100 cm Diameter: 75 cm

Table 1: Geometrical characteristics of the deployed mine-like objects.

4. METHODOLOGY

4.1. Target detection and classification

In order to detect objects of interest, the algorithm used here focuses mainly on the geometrical configuration of the highlight and shadow cast on the seafloor. The methodology presented in [3] is applied, which segments the shadow and highlight using a combination of fuzzy sets and mathematical morphology. A number of geometrical features are then extracted from these areas, and are then used to classify targets from real data against a previously compiled database of mine-like objects (MLO) and friendly-objects (FO) features using a Monte Carlo Markov Chain approach proposed in [3] and the Adaptive Boosting Decision Trees proposed in [4].

In MCM, there is an increasing interest in having methodologies for automatic seafloor characterization related to mine-hunting difficulty. Several authors have started developing image processing techniques based on acoustic images [9-11]. Generally speaking, these techniques evaluate the pixel values and classify them following segmentation, feature extraction techniques, statistical models, amongst other approaches. More recently, [7] presented a seafloor characterization approach based on lacunarity with very promising results. The lacunarity of a set of pixels in a grayscale image is the ratio of the variance of the pixel values to the square of the mean of the pixel values. It has several forms of being calculated. In [7], a new, fast technique based on integral- image representations is proposed and validated using thousands of HR sonar images with diverse seafloor conditions. This technique has been applied in [5,8] for the evaluation of the segmentation results compared with ground truthing measurements (sediment analysis, video imaging, sediment profile imagery and diving).

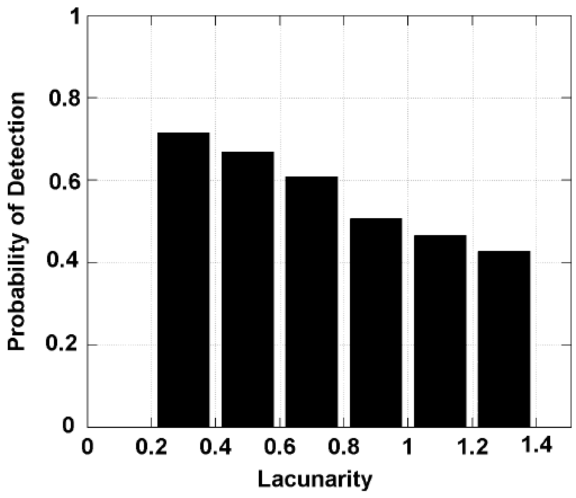

Fig. 3: Probability of detection for the detection program as a function of lacunarity [5].

4.2. Data processing strategies

Three different strategies were tested. In the first one, the complete data set is processed by the detection and classification programs. Time to complete the process is count and the probability of detection and classification are calculated separately for each classification algorithm. For the second approach, human experts are asked to process the entire database. Their timing is also counted and their detection and classification probabilities are also calculated. Finally, the data set is divided into two soil categories: "benign" and "complicated", based on the lacunarity results presented in [5], and the algorithms process one part of the data while human/machine combination processes the other part.

In [7], it was shown that the probability of detection decreases as the lacunarity increases. Results found in [5] are in accordance with the later (see Fig. 3). High values of lacunarity mean that the pixel-intensity variation is high, and after studying probability of detection in such seafloors, they are categorized as "complicated". Seafloors with low values of lacunarity are categorized as "benign".

Sonar data collected in the test area presented a high variation in the lacunarity values [5]. Several zones could be identified and well correlated with the environmental parameters. Where a high concentration of complete shells and benthos were found, high lacunarity values (higher than 1) were calculated. Presence of shell hash also increments the lacunarity values (between 1.1 and 1.4). This could indicate that shell hash increase the sediment (scatterer) size in average and therefore the seafloor becomes more "complicated". The latter is accordance with the hypothesis found in [7] where lacunarity values increased with sediment size.

5. DISCUSSION

The probability of detection is shown in Fig 4 for the operators (Op) and algorithm (ATD). Note that, in order to simplify the graph, the operators are not considered individually but their average performance is used as a single expert. The graph shows results as a function of mine type and pixel size. Not surprisingly, it can be seen that the probability of detection decreases with the resolution. Results drastically decrease for the lowest resolution, with exception of the detection by the operator and the ATD of the cylindrical mine (GM) in such conditions. This can be explained by the geometry and dimensions of the cylindrical mine, which are larger compared to the manta and rockan mines. The detection algorithm performs slightly better than the human expert, except for the manta mine (MA) and the moored mine (MM, not presented in the graph).

The probability of classification is shown in Fig 5 for the operators (Op, averaged as well) and both algorithms (ATC1 and ATC2). Performance of both algorithms is similar and they outperform the operator, except for the manta mine (MA) and the moored mine (MM, not presented in the graph). None of the later mine was correctly classified by the algorithms since it is not included in the training database. Although cylindrical mines (GM) were detected for images with a pixel size larger than 10cm, none of them were correctly classified by the operator nor by the ATC algorithms. For the later, this can be explained by the fact that algorithms were trained using images with pixel size smaller than 5cm. For the cylindrical and rockan mines, it can also be seen that operators performed better for the images with pixel size between 5 and 10cm than for images with better resolution. This can be explained by the fact that the majority of images in the 5-10cm pixel size category comes from the sonar system they are used to work with.

For the last part of the analysis and following the lacunarity, data are divided between ‘benign’ and ‘complicated’. For the test area considered here, the area with the lowest lacunarity values (between 0.2 and 0.4), i.e., scoring the highest probability of detection (0.7) is categorized as “benign”. This corresponds to almost 30% of the dataset. The rest of the area is considered as “complicated” (probability of detection lower than 0.7). Data included inside the “benign” soils are treated by the detection and classification algorithms directly, while detection in the “complicated” soils is first performed by the human operators and their results used as input for the classification algorithm. Following this strategy, the number of false alarms is drastically reduced by 73% and the number of misdetection is reduced by 48%, compared with the first strategy. It is worth noting that for this strategy, time of processing is doubled during the analysis of the human operator in combination with the algorithm, compared with processing using only the algorithms. The extra time added when passing from one strategy to another also depends on the experience of the operators with the type of images (working with previously unseen images, i.e., not previously used for training operators, which is the case in this study) as well as the size of the data to be processed by the operator only (corresponding to the “complicated” soils). The selection of the threshold between ‘benign’ and ‘complicated’ soils could therefore play an important role in the timing. These are crucial factors that would have to be considered for the planning and evaluation of a mission.

Further work needs to be done to improve the detection and classification algorithms, and to consider a different approach such as deep learning. Future MCM capabilities will integrate different systems for autonomous detection and classification of targets. This study opens a discussion for the importance of human-in-the-loop, in conditions such as presence of targets unknown for the ATC algorithm or difficult environmental conditions.

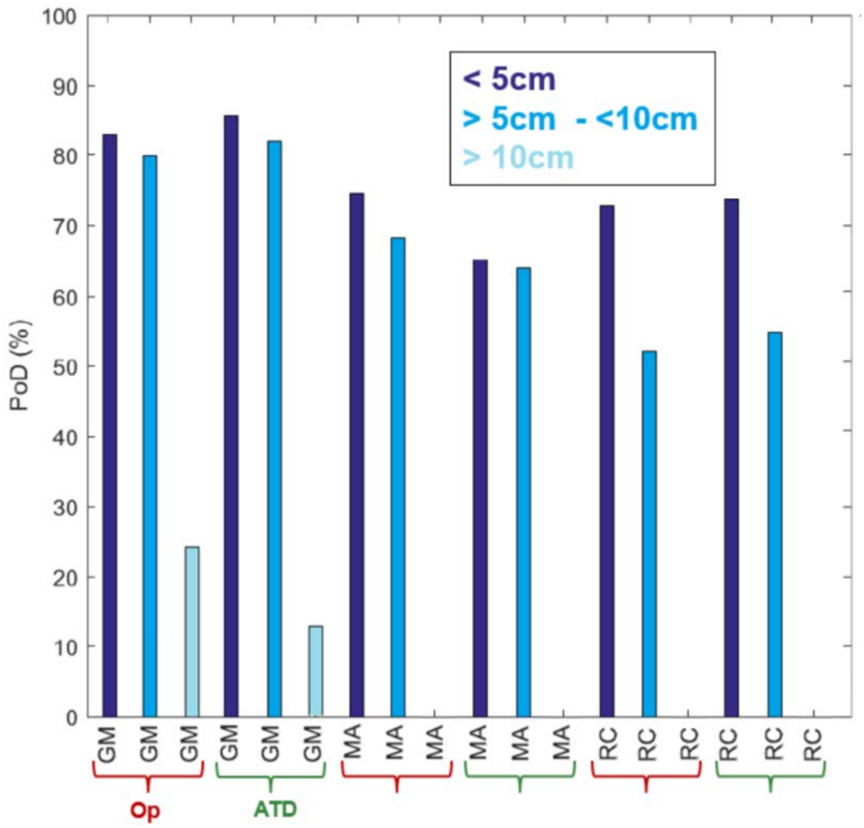

Fig. 4: Probability of detection for operators (BE Op) and algorithm (ATD) as a function of mine type and image resolution.

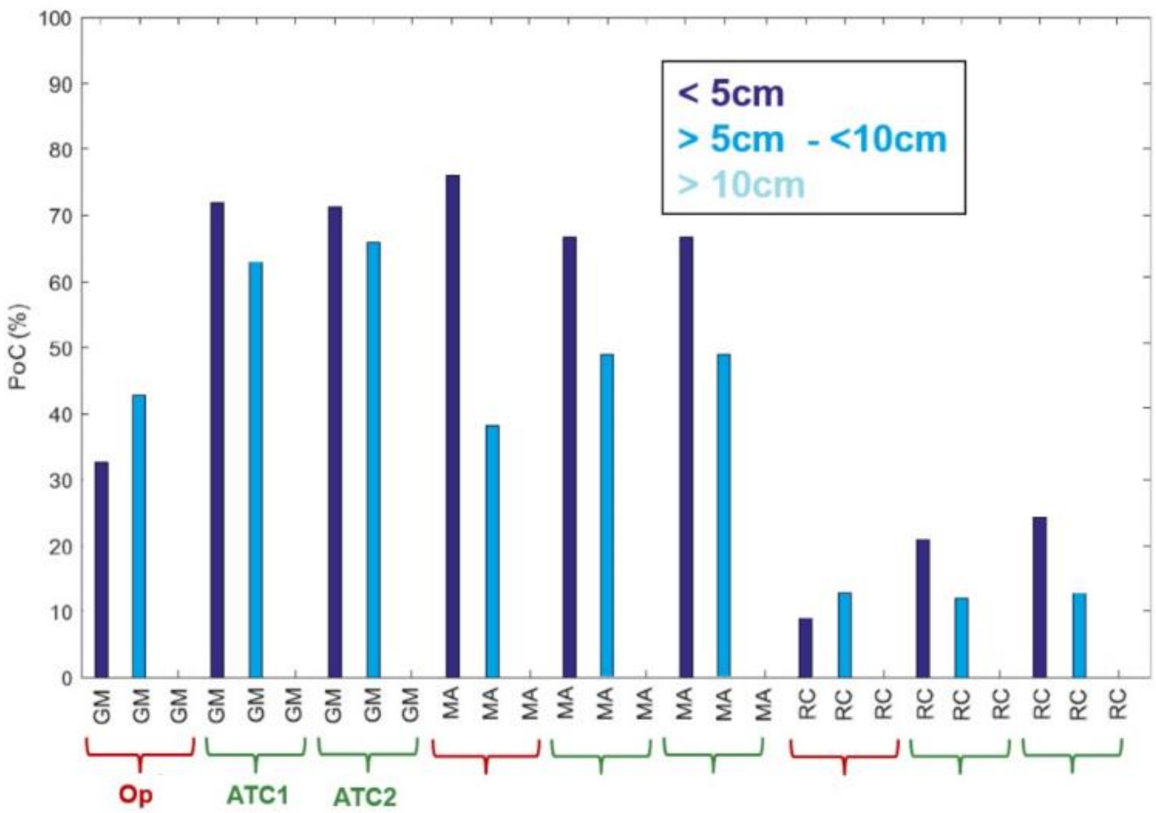

Fig. 5: Probability of classification for operators (BE Op) and algorithms (ATC1, ATC2) as a function of mine type and image resolution.

6. ACKNOWLEDGMENT

The authors would like to thank CMRE STO for giving us access to part of the MUSCLE database from the Autonomous Undersea Surveillance group in order to train our algorithms and for allowing us to use their lacunarity software. The authors would like to thank also the human experts involved in the experiment as well as the crews of RV Belgica [12] and the different Belgian Navy vessels involved in data collection for their time and efforts in contributing to this study.

REFERENCES

[1] J. Stack, "Automation for underwater mine recognition: Current trends and future strategy," in SPIE Defense, Security, and Sensing, vol. 8017, 2011.
[2] D. Williams, M. Couillard, and S. Dugelay, "On Human Perception and Automatic Target Recognition: Strategies for Human-Computer Cooperation", in International Conference on Pattern Recognition, 2014.
[3] Lopera, O., A. Borghgraef and E. Mersch; The Special Case of Sea Mines, in Mine Action, Beumier C, ed. InTech 2017, pp 251-279.
[4] J. Isaacs, "Sonar Automatic Target Recognition for UXO Remediation," Naval Surface Warfare Center PCD, Proceedings of the Institute of Acoustics, Vol. 36, 2014.
[5] O. Lopera and S. Papili, "Sea-floor characterization related to object-detection performance of sonar systems: a case study", in Oceans 2017, Aberdeen, UK, 2017
[6] O. Lopera, "Human in the loop for autonomous underwater threat recognition", in Oceans 2018, Charleston SC, 2018
[7] D. Williams, "Fast Unsupervised Seafloor Characterization in Sonar Imagery Using Lacunarity", IEEE transactions on Geoscience and Remote Sensing, Vol. 53, No. 11, pp. 6022-6034, 2015.
[8] O. Lopera and S. Papili, "An introductory study of the impact of environmental parameters in the performances of imaging sonar systems", in Oceans 2016, Monterey, CA, 2016.
[9] L. Sha, and L. Nolte, "Effects of environmental uncertainties on sonar detection performance prediction," The Journal of the Acoustical Society of America, vol. 117, no. 4, pp. 1942–1953, 2005.
[10] D. Williams and E. Fakiris, "Exploiting environmental information for improved underwater target classification in sonar imagery," IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 10, pp. 6284–6297, 2014.
[11] X. Lurton, and G. Lamarche, "Backscatter measurements by seafloor-mapping sonars: Guidelines and Recommendations". Geohab, Eastsound, WA, USA. 200pp. 2015.
[12] http://www.howbigisbelgica.be