MBES WATER COLUMN DATA PROCESSING FOR SMALL UNDERWATER TARGET RECOGNITION

Isabelle Quidu[a], Michel Legris[a]

[a] Lab-STICC UMR CNRS 6285, ENSTA Bretagne, 2 rue François Verny, 29806 Brest Cedex 9, France

isabelle.quidu@ensta-bretagne.fr
michel.legris@ensta-bretagne.fr

Abstract: *In the context of the mine warfare modernization plan, the French Ministry of Defense is studying the possibility of close object inspection by an autonomous underwater vehicle. To that end, the vehicle is equipped with a multibeam echosounder (MBES), and travels to within 10 meters of the object. With each narrow beam formed by the MBES, a sounding is measured by a detection algorithm that estimates the two-way propagation time between the sonar and the seabed. Unfortunately, manufacturer algorithms are designed for bathymetry estimation and suffer from two drawbacks for object recognition: first, multiple detections per beam are seldom available; second, too many false alarms arise in case of strong specular reflection. In order to achieve better recognition of small objects, in this paper we propose processing the water column data where the 3D shape of the object is visible. This huge volume of data needs to be reduced, however. Data are first enhanced by the use of the bitonic filter, which combines non-linear morphological and linear operators. Then, hysteresis thresholding is applied, allowing multiple detections, as well as false alarm mitigation. This technique is assessed on various data sets collected by the French Defense Procurement Agency (Direction Générale de l'Armement, Naval Techniques section) with three different MBES: R2Sonic2022 (700kHz), EM2040 (400kHz) and MB2250 (2.25MHz). Results show that the set of detected points allows us to complete the final step of recognition based on 3D shape matching.*

Keywords: *MBES water column data, target recognition*

1. INTRODUCTION

For the purpose of underwater mine warfare, the French Defense Procurement Agency, DGA TN (Direction Générale de l'Armement, Naval Techniques section), develops future MCM (Mine CounterMeasure) concepts. In order to mitigate human risks and to increase stealth, underwater mine warfare missions will mainly be achieved by AUVs (Autonomous Underwater Vehicles) deployed from a mothership or USV (Unmanned Surface Vehicle). A piece of these operational systems is the I-AUV or Identification-AUV, able to determine the type of mine. Common MCM procedures require identification to be performed either by divers, or by automated optical imaging. The main identification sensor is the camera but, except for specific optical imaging (like the *TileCam optical camera* in [1]), it is usually insufficient in turbid water. To gather additional information, the use of an acoustic system is proposed. This system is a multibeam echosounder (MBES) carried by the vehicle which travels to within 10 meters of the object. Usually, from each narrow beam formed by the MBES, a sounding is measured by a detection algorithm that estimates the two-way propagation time between the sonar and the seabed. Unfortunately, such manufacturer algorithms are designed for bathymetry estimation and generally provide a single detection per beam. By looking carefully at the water column data where the 3D shape of the object is visible, we sometimes notice that multiple echoes can be relevant per beam. This is even more noteworthy when specular reflection arises, together with sidelobes (due to the transmitter and/or the receiver arrays). In order to achieve better recognition of small objects, we propose in this paper to enhance the water column data (section 2) and to reduce this huge volume of data to a few relevant detections (section 3). Results will be obtained from three different MBES datasets (section 4). We will close with a comparison with the soundings obtained using the manufacturer's algorithm.

2. WATER COLUMN DATA ENHANCEMENT

In image processing, denoising while preserving edges (in terms of contrast and localization) is challenging because noise and edges are both high frequency components. Common methods include a preliminary step of detection of edges in order to only low-pass filter the areas with low gradient values. A famous example is the anisotropic filter [2]. However, the latter is not easy to set parameters for and, because of high gradient values, does not denoise nor connect pixels on the edges. Instead, we prefer the bitonic filter [3]. A bitonic sequence (defined in the context of sorting as an extension of monotonic) is one which increases monotonically (or not at all) to a peak, then decreases monotonically (or not at all), *i.e.* it has at most one local maxima or minima. Finally, noise is anything which is not bitonic over the given range. Notice that only the data rank matters, not the level. As such, a bitonic filter is based on rank filters with a small centile c, in order to preserve local maxima or minima, while rejecting unsignificant impulses. More precisely, a robust opening operation is used, with a small centile c rather than the usual minimum, and $(100 - c)$ instead of the usual maximum. A robust closing operation is similarly computed. In order to eliminate residual noise, a Gaussian filtering of the residue between the morphological (opening or closing) operation output and the original signal is performed (in image processing, this residue is known as the output of a top-hat filter). These two smoothed errors are used as weights to combine the previous opening and closing operators' outputs. This combination preserves mean signal values in the case of a noisy signal. The advantages of the bitonic filtering are twofold: a) contrast enhancement and b) denoising (impulsive as well as white Gaussian noises).

When dealing with 3D matrices, the practical implementation of the bitonic filter carries out 2D filtering channel by channel, which means ping by ping, in our case. Given the expected acoustic signature of the object returned by the MBES (*i.e.* edges of the object in the sonar image in polar coordinates), the structuring element that is used is a diamond whose size in pixels depends on the chosen degree of detail dm (in meters), and on the image resolution (in meters). The shape of the structuring element is a trade-off between the necessary neighborhood consideration for smoothing and the preservation of high frequency components of the edges (like corners). Following the filtering of the polar images, another filtering is performed beam by beam on "images" whose axes are the ping number axis and the range axis. This time, given the fact that data have already been filtered in range, the structuring element is set to only 3 pixels along this axis, while it is again set to a number of pixels tuned by the same chosen degree of detail along the ping number axis.

3. MULTIPLE DETECTION PER BEAM

After water column data enhancement, object edges are now denoised and enhanced. However, the volume of data is still too large for the final step of object recognition. The sonar image has to be processed so as to only keep a few detections per beam. Generally, these detections are automatically provided by the system. The systems used here compute a single detection or sounding, but the built-in algorithms are too sensitive to strong specular reflection, fish, or sidelobe effects (Fig.). This can lead to false alarms and to misrecognition. Unfortunately, these predictable physical effects cannot be removed easily from the image. For example, simply normalizing the polar image in order to mitigate the receiver sidelobe pattern also induces an unwanted enhancement of the transmitter sidelobe pattern in the transversal direction.

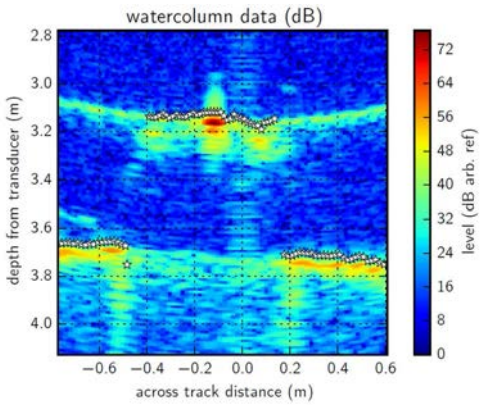


Fig. 1: Soundings measured by the R2Sonic 2020 MBES (700 kHz) on a sphere with strong specular reflection and sidelobes (sonar image in Cartesian coordinates).

In order to offer the possibility of multiple detections per beam with a reasonable false alarm rate, a 2D hysteresis thresholding is performed on each (previously enhanced) polar sonar image. This allows us to separate strong-highlight areas (over the upper threshold) from weak-highlight areas (over the lower threshold, and connected to at least one strong-highlight pixel). The trickiest part of this algorithm is the threshold adjustment. In our case, in order to have a fully automated procedure, the upper threshold is set as the proportion of pixels in the seafloor mask introduced below. This works in most cases, except in case of objects with very weak highlights. The lower threshold is a fraction k of the upper threshold. An example of hysteresis thresholding is showed in Fig. 2, on a polar sonar image where we can see the shape of a Manta mine.

The hysteresis principle is used here in order to suppress isolated weak highlights, as well as to mitigate sidelobe effects. However, this is still insufficient for sonar systems with

classical beamforming for which fairly large areas can be labelled as weak highlight after thresholding, and connected to a strong highlight. In order to discard them, we propose a post-procedure that consists in detecting the false weak areas, then to remove them from the ternary image obtained after hysteresis. The false weak area detection is achieved by using two masks: let us call “weak mask” (*resp.* “strong mask”) the binary image obtained by thresholding the enhanced image with the lower (*resp.* upper) threshold. These masks must first be pre-processed: a morphological smoothing of the weak mask is performed so as to remove or fill segments that do not contain at least np connected pixels along each beam; a morphological opening of the strong mask is performed with the cross-structuring element. The strong mask is then dilated with the same structuring element as for the bitonic filter. Finally, false weak-highlight pixels are pixels of the weak mask that are not marked by the dilated strong mask.

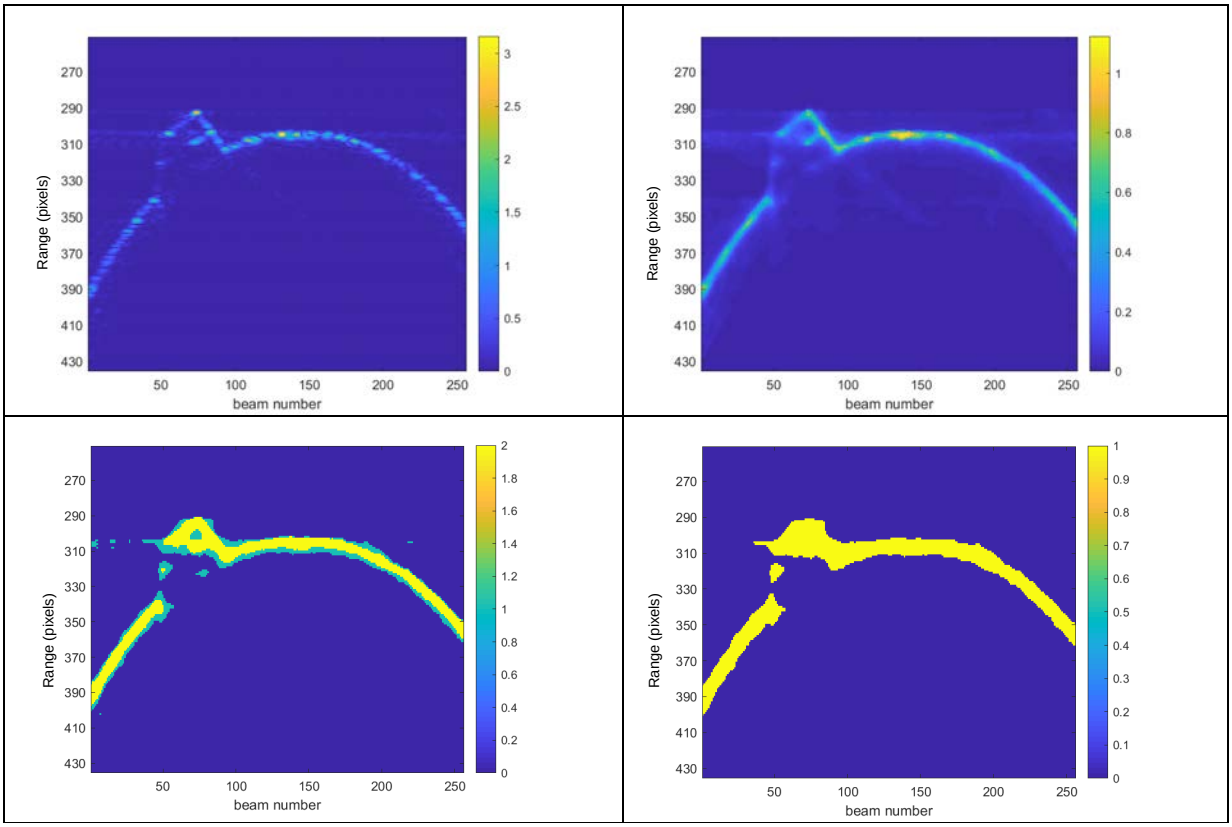


Fig. 2: EM2040 MBES (400kHz): polar image of a Manta mine (a); enhanced image (b); ternary image (c); hysteresis image (d)

The resulting image can now be processed with the following edge thinning procedure. Practically, the ternary image is processed column by column (or beam by beam). Each column vector is composed of values ranging from 0 (no edge) to 2 (strong-highlight pixel). A decision occurs on each pixel segment different from zero. In order to distinguish object detections from seabed detections, a seafloor mask is computed by averaging polar sonar images over the sequence and performing a thresholding at a centile that guaranties a full mask of the seafloor (that means, without hole for a given beam). If there is not a null intersection between this segment and the seafloor mask, the detected point is located at the mean index of the segment, and its value set to 0.5. In case of a null intersection, if the segment is composed of weak highlights, the detected point is located at the mean index of the segment, otherwise it is located at the index of the pixel of maximum value in the enhanced sonar image, and its value set to 1. An example of sidelobe effect mitigation and edge thinning is given on Fig. 3.

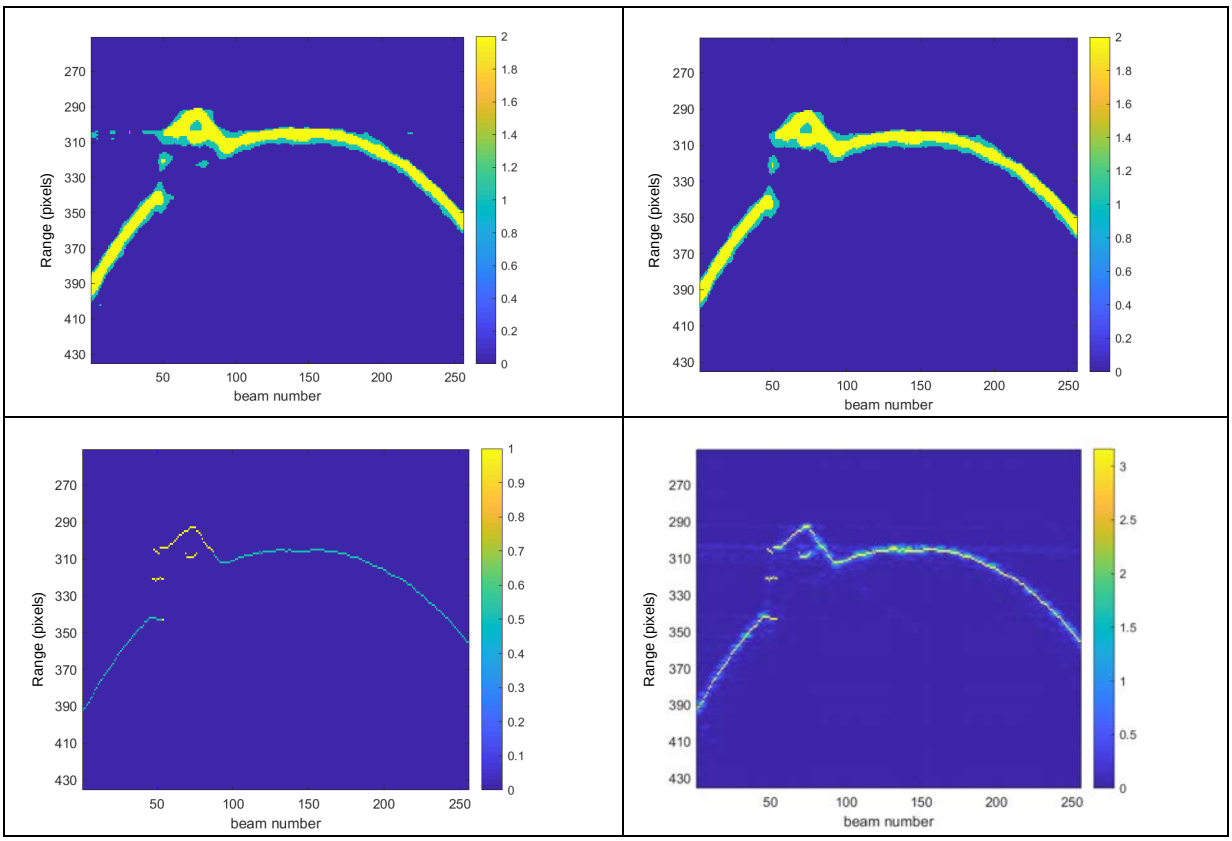


Fig. 3: EM2040 MBES (400kHz): ternary image (a); after post-processing (b); after edge thinning (c); superimposition of detections and polar image (d)

4. RESULTS

a) Dataset description

Since 2014, in preparation for this study, several MBES (specifications given in Table 1) were evaluated:

- BlueView MB2250-45 carried by the Daurade AUV and property of DGA TN (trials in 2014),
- R2Sonic 2022 rented by ENSTA Bretagne and mounted on a survey boat (trials in 2014),
- Kongsberg2040C lent by the National Hydrographic Service « SHOM » and mounted on a survey boat (trials in 2016)

An example of polar image obtained with each sensor is given Fig. 4.

b) Parameter setting

In the implementation of the bitonic filter, the diamond structuring element for morphological operations, with centile $c = 10\%$, is of size $l_i \times l_j$ and the rectangular window for Gaussian filtering is of size $2\sigma_i \times 2\sigma_j$ where $\sigma_x = \frac{l_x}{3}$, $x = i, j$. For each sensor, l_i (*resp.*

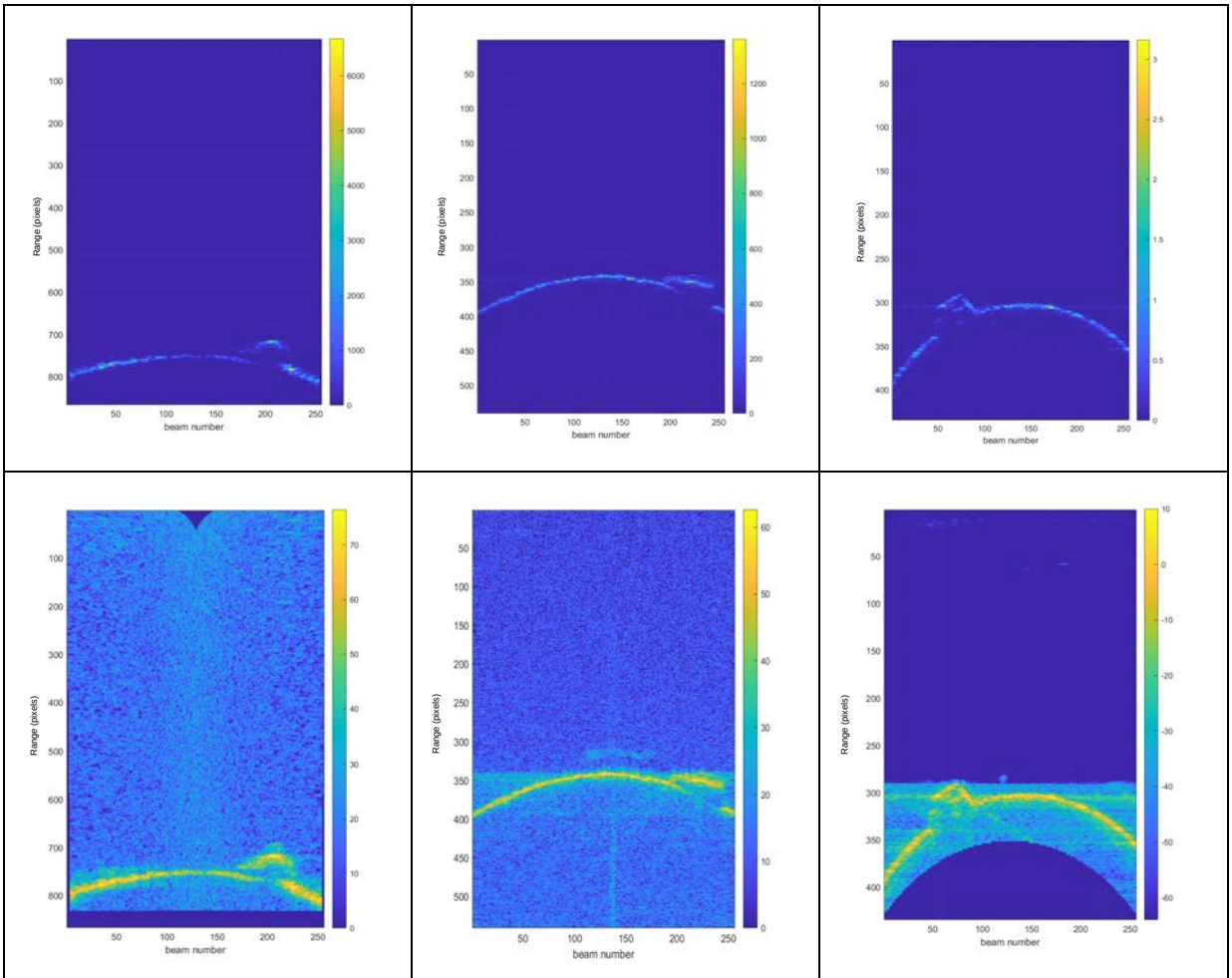
l_j) is fixed by considering the pixel size r_i (*resp.* r_j) at the shortest range to the seabed and the chosen degree of detail dm :

$$l_x = int\left(\frac{dm}{r_x}\right), \quad x = i, j \text{ where the } int(\ ) \text{ function gives back the closest integer.}$$

Whatever the sensor, $dm = 0.15m$.

	Frequency (kHz)	*Bandwidth (kHz)*	*Beamwidth*	*Swath sector*	*Range sampling rate (m)*	*Azimuth sampling rate*	*Distance travelled between pings (m)*
BlueView MB2250-45	2250	-	$1° \times 1°$	45°	0.006	0.16°	0.046
R2Sonic 2022	700	60	$0.6° \times 0.6°$	60°	0.011	0.23°	0.023
Kongsberg 2040C	400	60 max	$1° \times 1°$	60°	0.013	0.26°	0.019

Table 1: MBES specifications



BlueView 2250-45 (Rockan mine) Sonar altitude = 4.65m	R2Sonic 2022 (Rockan mine) Sonar altitude = 3.75m	Kongsberg 2040C (Manta mine) Sonar altitude = 4.0m

Fig. 4: Polar images from the different sensors (Up: linear scale; Down: dB scale)

Concerning the hysteresis thresholding, $k = 0.5$, except for MB2250 where $k = 0.6$, which leads to a higher value for the low threshold because of the high resolution of this sensor that can more easily induce false alarms. Morphological smoothing is performed with $np = 3$ pixels, except for MB2250, for which $np = 5$ pixels.

Last, Teledyne BlueView uses a blazed acoustic array to form and steer beams: the image is generated by time-frequency processing. Consequently, a given beam corresponds to a given frequency and then to a given resolution. Unfortunately, this impacts the backscattering level, which fluctuates along the beams. To avoid this, a preprocessing which consists in normalizing each row of the polar image is performed for this particular dataset. An example is given Fig. 5.

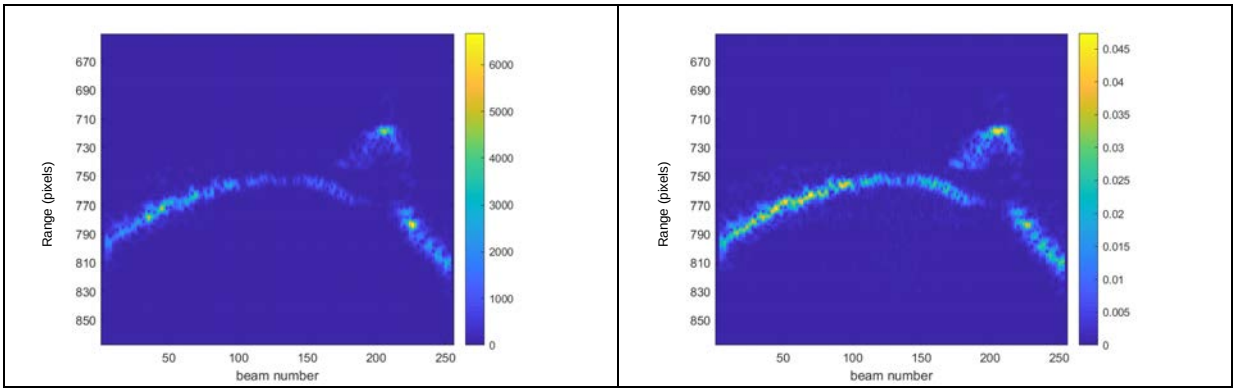


Fig. 5: MB2250 polar image normalization (left: without; right: with)

c) 3D detection views

As expected, compared to soundings measured by manufacturer detection algorithms, our algorithm perceives more details like multiple detections per beam and better describes objects which usually return strong specular reflection. An example is given in Fig. 6 for the MB2250 MBES with a cylinder.

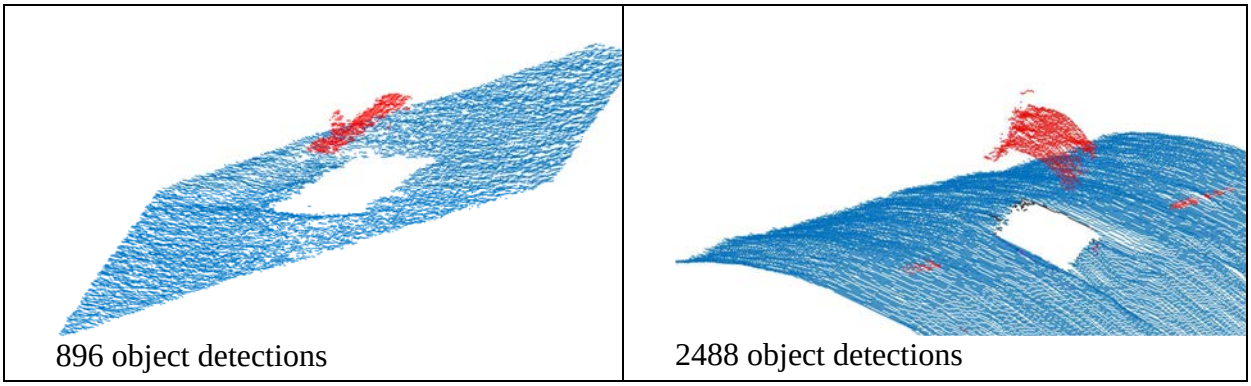


*Fig. 6: MB2250 MBES - Cylinder - Manufacturer detections (Left) compared to our detections (Right) - red (*resp. *blue): first object (*resp. *seafloor) detection per beam; magenta (*resp. *black): second object (*resp. *seafloor) detection per beam.*

The EM2040C MBES manufacturer detection algorithm performs generally better than ours but, again, multiple detections per beam can be useful to help recognizing the object and also, we notice that our algorithm can detect very small objects, as shown in Fig. 7.

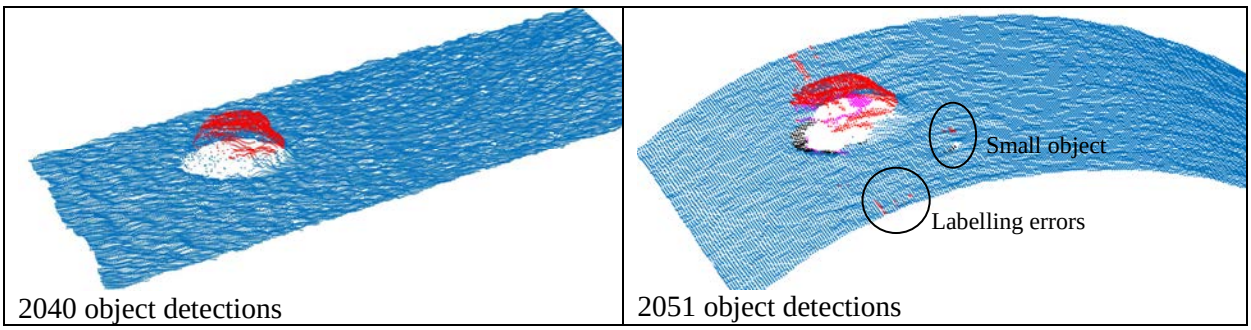


*Fig. 7: EM2040C MBES - Manta mine- Manufacturer detections (Left) compared to our detections (Right) - red (*resp. *blue): first object (*resp. *seafloor) detection per beam; magenta (*resp. *black): second object (*resp. *seafloor) detection per beam.*

5. CONCLUSION

In this paper, an algorithm has been proposed to enhance and then reduce the volume of data gathered by three different MBES. The first contribution is the introduction of the bitonic filter that has proven very successful in our context. The second contribution is the object contour detection by hysteresis thresholding together with morphological operations, followed by an edge thinning procedure. Results showed some improvements compared with other manufacturer algorithms: details can be recovered and multiple detections are available per beam. The next step is to find the best use of these results for recognizing objects for which a 3D model is available.

6. ACKNOWLEDGEMENTS

The authors would like to thank the Direction Générale de l'Armement/Techniques Navales for funding this study.

REFERENCES

[1] **T. R. Krogstad, M. S. Wiig and O. Midtgaard**, Autonomous object identification in mine countermeasure missions, In *OCEANS'11 MTS/IEEE KONA 2011*, 2011.

[2] **P. Perona and J. Malik,** Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 629–639, July 1990.

[3] **G. Treece**, The bitonic filter: linear filtering in an edge-preserving morphological framework, *IEEE Transactions on Image Process*ing, vol. 25, no 11, pp. 5199-5211, 2016.

[4] **S. T. Bachelor, R. L. Thompson and J. L. Seawall**, Acoustic Imaging with Time-Frequency Beamforming, Blazed Arrays, and Other Array Designs, *US Patent* number US-4, 022,377, July 12, 2004.