Improvement of automatic man-made object detection in underwater videos by using of navigational information

Isabelle Leonard, Andreas Arnold Bos, Ayman Alfalou, Nicolas Mandelert

IMPROVEMENT OF AUTOMATIC MAN-MADE OBJECT DETECTION IN UNDERWATER VIDEOS BY USING OF NAVIGATIONAL INFORMATION

ISABELLE LEONARD[1], ANDREAS ARNOLD-BOS[2],
AYMAN ALFALOU[1], NICOLAS MANDELERT[2]

Abstract. In order to detect man-made objects in underwater video, we propose and validate a novel method based on background subtraction methods. The other main contribution is the introduction of *a priori* elements deduced from positioning sensors. These elements allow to enhance the visibility of underwater objects thanks to the calculation of the sun position in relation to the vehicle position, the detection with the distance from the object and the post-processing with constraints defined on the vehicle movements. These constraints allow to reject false detections and to better know the position of the detected object. We tested our algorithm on data acquired at sea and show that we improve detection results and decrease false alarm rate, comparing to our former work. Both algorithms have been applied on the same videos. We still have to increase the true detection rate while reducing processing time *i.e.* processing time should be close to video rate.
Keywords: underwater images, mine detection, background subtraction, use of navigational information

1. Introduction

Underwater mines represent an important threat. This threat is generally addressed using a system with four steps: a detection step, a classification step, an identification step, and a neutralization step [1]. Nowadays, the trend is to design autonomous systems (Autonomous underwater vehicles, AUV) to avoid the involvement of clearance divers. Some of these AUVs are equipped with optical video camera besides sonars and positioning sensors. In case of identification mission, detection and guidance are done by sonar. When the vehicle is close to the mine, the video camera is activated. However, video images are affected by the underwater medium. Scattering and absorption cause images with weak contrast: objects are difficult to distinguish on the ocean floor. Besides, real time preprocessing and detection algorithms are necessary to improve identification results and closed-loop vehicle guidance.

In this article, our vehicle is supposed to be able to automatically identify a mine using a video camera. For that, we assume that the vehicle knows approximately

[1]Departement Vision L@bISEN, ISEN, 20 rue Cuirasse Bretagne, 29228 BREST CEDEX 2, FRANCE
(isabelle.leonard, ayman.al-falou)@isen.fr

[2]General Sonar Studies, Thales Underwater Systems, Route de Sainte Anne du Portzic, CS 43814, 29238 BREST CEDEX 3, FRANCE
(Andreas.Arnold-Bos, nicolas.mandelert)@fr.thalesgroup.com

the position of the target object, using sonar navigation data. Thus, first of all, our system has to detect this object and position the mine exactly.

We present a novel method based on background subtraction and an adaptation of our detection method to use navigational information. Videos are analyzed image per image. One of the contributions of this article is the use of navigational information in order to improve true detection rate. Another key point of our method is to consider the temporal aspect, *i.e.* the link between successive images.

First we will detail our problematic and present some performance criteria. Then we will detail our proposed algorithm and our first experimental results. Finally, we will compare the performances of our algorithm with a state-of-art algorithm.

2. Problem statement

This work takes place in the underwater mine detection domain. This underwater medium affects the light used to illuminate the target scene through absorption and scattering phenomena. The visibility range reaches only a few meters. This limits the use of video cameras in underwater applications. Moreover, underwater images have a weak contrast, require preprocessing steps and restraint the efficiency of edge detection methods. Accordingly, we are interested in background subtraction methods less sensitive to underwater medium than edge detection methods.

Furthermore, we must take processing time constraint into account. In fact, our system should be embedded on an AUV. This vehicle guides itself depending on the results obtained with the detection and identification steps. Thus, the processing time should be close to the video recording rate.

Our preliminary tests, using experimental data acquired at sea, showed the good algorithm efficiency.

3. How to measure algorithm performances?

First of all, we have to define some parameters to test the detection algorithm performances. The first parameter is the definition of mine presence zones. The other parameters are used to calculate the detection probabilities. To compute these probabilities, we create for each experimental image an annotation file, containing among others the viewer-object distance and the position of the region of interest located around the target object.

3.1. **Definition of mine presence zones.** In this article we work on underwater videos. We apply our algorithm on the video, *a posteriori*, not during the video acquisition. But in order to guide the underwater vehicle when the algorithm is be embedded on it, we will have to define mine presence zone. We can ask a human operator to select zones where there is an object. But this solution is not a robust one and especially not a repetitive method. Indeed, the presence zone detection differs in a random movement, at the beginning and the end of the zone, depending on the operator and the video. Some people wait for the entire object when others need few millimeters to detect it. Therefore, we are looking for a subjective method. For that, we carefully investigated our videos and chose a distance criterion. We empirically fixed a maximum viewer-object distance to 6 meters. According to visibility and turbidity, this distance is a satisfactory compromise.

3.2. **Detection probabilities.** Our new algorithm is able to detect several objects in an image. For that, we have to distinguish true and false detections. Moreover, we wish to test detection performances of our algorithm and compare them with another one (our former detection algorithm). To solve this problem (distinguish true and false detections), we defined several probabilities, summarized in figure 1.

		Reality	
		Mine	Nothing
Results	Mine	Ptp true positive	Pfp false positive
	Nothing	Pfn false negative	Ptn true negative

Figure 1. Definition of different used probabilities

These probabilities are defined as follows:

- if our algorithm detects an object with correct location, we have a true positive detection, noted Ptp (*cf.* figure 2)

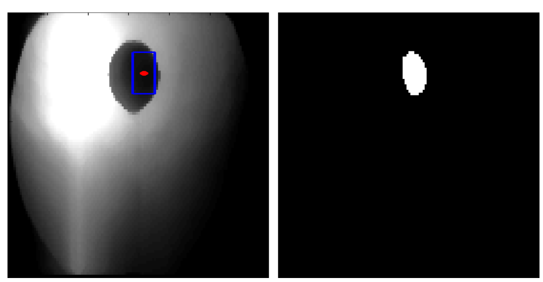
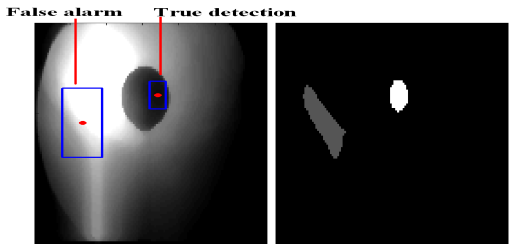
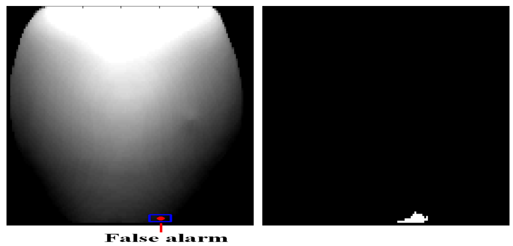


Figure 2. Example of a true positive detection

- if our algorithm detects an object with wrong location (*cf.* figure 3a) or in an empty image (*cf.* figure 3b), we have a false alarm or a false positive detection, noted Pfp (*cf.* figure 3)



(A) Example of a true detection and a false alarm



(B) Example of a false alarm

FIGURE 3. Example of false alarm detection

- if we miss a detection in an image with an object (*cf.* figure 4), we have a false negative detection noted Pfn

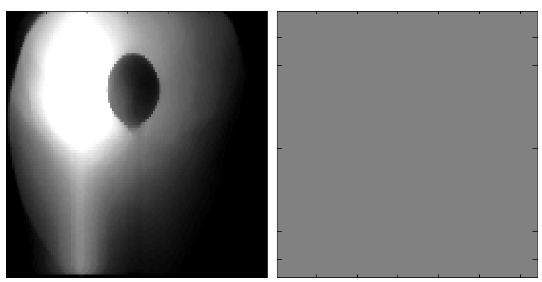


FIGURE 4. Example of a missed detection

- if no object is detected in an empty image (*cf.* figure 5), we have a true negative detection, noted Ptn

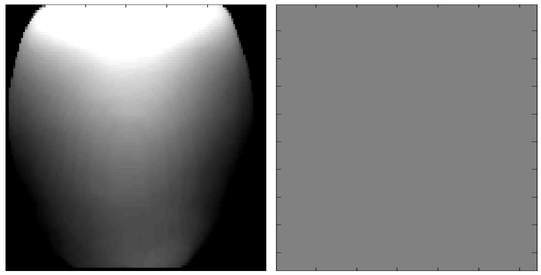


Figure 5. Example of a true negative detection

As we define the mine presence-zone parameter according to viewer-object distance, we present our detection results according to the same viewer object distance too. However, we cannot calculate probabilities for each single distance. For that, we group distances by intervals equal to 0.5 meter: we group all our detection results for a given video with a viewer-object distance from 0 meter to 0.5 meter and so on. Thanks to annotation files, we know the viewer-object distance for each image. Thus we know how many images and how many true objects correspond to each distance interval. For each distance interval, each probability is defined as:

$$\frac{\text{number of detection objects in the considered distance interval corresponding to the probability definition}}{\text{number of all objects in the distance interval}} \tag{1}$$

4. Algorithm

Underwater images have a poor contrast caused by light absorption, which increases with water turbidity. To increase the mine detection rate, we have to preprocess images. We presented our preprocessing in a previous article [2]. To limit the moir effect and the processing time, we resize the images. Then we apply the edge enhancement proposed by Arnold Bos *et al.* [3, 4]. Finally, we use Phong's model [5] to reduce the sun reflection on the sea floor.

According to Phong [5], the received intensity I_r is the combination of the ambient light I_a (a constant), the scattered light I_d and the specular light I_s. In the underwater realm, the absorption balances this combination:

$$I_r = e^{-c.z}(I_a + I_d + I_s) \tag{2}$$

where c represents the absorption coefficient and z the distance between the object and the viewer. The specular intensity depends on the viewer and light source positions. The scattered light depends on the light source position [5]. The equation 2 can also be written as:

$$I_r = e^{-c.z}(I_a + (-\vec{L}\vec{N})I_e + (\vec{R}\vec{O})I_e) \tag{3}$$

where I_e is the emitted intensity, the vector $\vec{L}$ represents the source-object vector, the vector $\vec{N}$ is the vector perpendicular to the object, the vector $\vec{R}$ represents the reflected light and the vector $\vec{O}$ represents the object-viewer vector. These vectors are explained on figure 6.

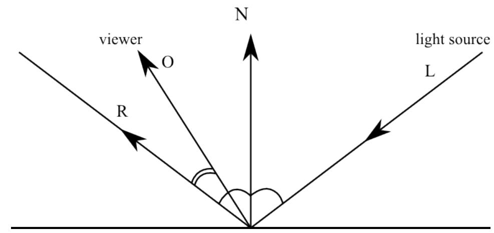


Figure 6. Definition of angles and vectors used to calculate the received intensity

Navigational information contain the AUV position and time of video recording. Thus we know the viewer position and we can calculate the sun position thanks to Reda and Andreas algorithm [6].

On figures 7a and 8a, we present preprocessed images.

In our previous publication [2], we used the phase of the image spectrum to detect mine edges. We obtained good results. However, this method has a very low detection rate especially when we have images with weak contrast, as shown on table 1. The detection probability (table 1, 4th column) is low (below than 35%) while the false alarm rate (table 1, 5th column) is very high (above 45%).

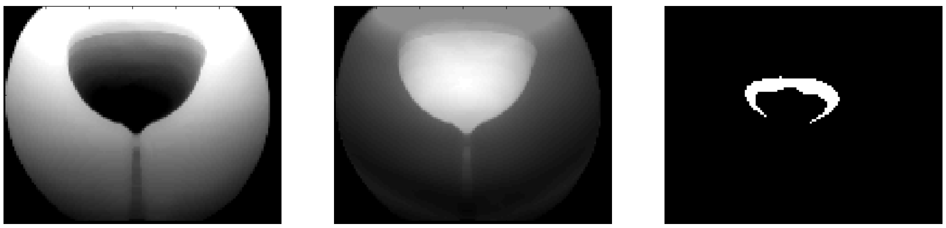
Mine	Number of studied images	Number of images with a mine	Ptp	Pfp	Pfn	Ptn
Manta	25205	18275	18.87%	71.64%	42.34%	14.74%
Cylinder	49251	37564	32.42%	46.72%	20.96%	33.21%
Sphere	11376	7919	31.45%	69.94%	25.95%	7.10%
Other objects	13905	10222	17.92%	71.15%	26.90%	14.35%
Empty videos	17389	0		8.97%		74.90%

Table 1. Results obtained with the method proposed in [2]

To improve these rates, we are interested in other kinds of methods. Edge detection methods are not always the most efficient methods when the contrast is limited. Region subtraction and especially background subtraction methods can solve this problem. This is not a new method in the underwater domain [7]. Moreover, these methods have demonstrated their good detection performances in other domains, *e.g.* Edgindton *et al.* [7] proposed and validated a new system based on these methods to detect animals. Thus, we adapted our algorithm based on these methods to detect correctly most of the true objects: we increase the true detection rate. Based on the background subtraction methods, we decided to use several images to create correctly a background image. In fact, we observe that the detection is more robust when we create the background image using mean images instead of using only one image to create the background. For that, we select the first hundred images of the video (where there is no object). We preprocessed these images and create the background image by averaging the images.

We decided to compare preprocessed images (*cf.* figures 7a and 8a) to preprocessed images with background subtraction (*cf.* figures 7b and 8b). Then we compared both images and looked for corresponding high intensity and very low intensity zones. Thresholds have been fixed empirically. Figures 7c and 8c show the results of the different steps.

Figures 7a and 8a present the preprocessed images. The images with background subtraction are shown in figures 7b and 8b. The comparison results are on figure 7c and 8c. On these examples, mines are clearly visible. On figure 7c, we obtain only the spherical mine. However, on figure 8c, the mine is truly detected but other zones are also presented, corresponding to difference between the sea floor and the images used to create the background, at the beginning of a given video. Thus there are some false alarm detections (*cf.* figure 8c).

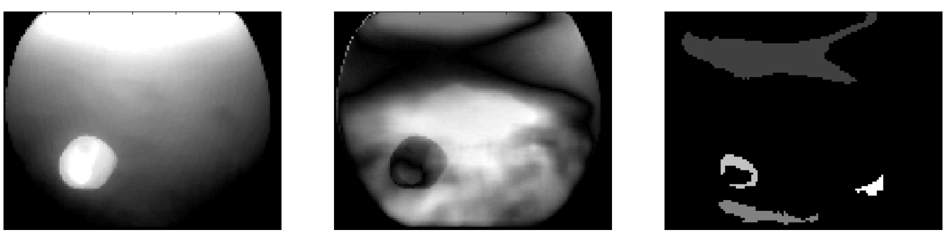


(a) Preprocessed image

(b) Preprocessed image with background subtraction

(c) Result of the image comparison

FIGURE 7. Result of our algorithm on spherical mine image



(a) Preprocessed image

(b) Preprocessed image with background subtraction

(c) Result of the image comparison

FIGURE 8. Result of our algorithm on Manta mine image

The results obtained with this method are presented on table 2. Detection probability (table 2, 4th column) is higher than detection probability previously obtained (*cf.* table 1), above 30%. The false alarm rate (table 2, 5th column) decreases. The true negative detection rate (Ptn, table 2, last column) is higher than 90% for mine Manta and spherical mine. The different results presented on table 2 show a good improvement compared to table 1, but the developed method is not the best one since all the probabilities are not close to the optimal probabilities.

It can be noted that we also improved the processing time. Our algorithm has been run using Matlab, version R2007a, using an Intel Core 2 Quad CPU cadenced at 2.66Ghz. The processing time is 0.08s per image, which is close to the video rate.

Mine	Number of studied images	Number of images with a mine	Ptp	Pfp	Pfn	Ptn
Manta	25205	18275	32.86%	20.83%	26.19%	96.61%
Cylinder	49251	37564	49.41%	62.37%	16.48%	31.61%
Sphere	11376	7919	43.23%	2.10%	56.41%	99.40%
Other objects	13905	10222	46.82%	71.94%	34.72%	13.96%
Empty videos	17389	0		4.13%		95.87%

TABLE 2. Results obtained with our proposed method

Further analysis of our results shows that detection can be changed between two consecutive images, when neither the underwater vehicle moves nor the mine. So we thought that we could improve the detection probabilities thanks to use of a special constraint application. The constraints we apply come from navigational information concerning the vehicle movement data: position (x, y and z) and orientation (tilt, roll and head). Mines are not supposed to move. The vehicle has no abrupt movements. Moreover, when the mine has been detected in the center of the considered image, it is not likely to disappear. Consequently, detection stops can be avoided by navigational information and detection position analysis.

These results need a comparison with other results obtained with a state-of-art algorithm. We chose an algorithm developed by Cybernetix and Thales [8]. This algorithm preprocesses the image and segments it to obtain contour images. More details are provided in [8]. The results obtained with state-of-art algorithm are presented in table 3.

Mine	Number of studied images	Number of images with a mine	Ptp	Pfp	Pfn	Ptn
Manta	25205	18275	19.78%	3.54%	70.51%	99.02%
Cylinder	49251	37564	42.91%	10.68%	40.42%	84.20%
Sphere	11376	7919	36.05%	3.83%	54.04%	99.64%
Other objects	13905	10222	18.30%	9.63%	73.02%	82.55%
Empty videos	17389	0		8.57%		91.43%

TABLE 3. Results obtained with a state-of-art method

We particularly worked on videos containing spherical mines. Consequently, our detection probability (Ptp, 4th column in table 2) and true negative detection probability (Ptn, last column of table 2) are higher than reference algorithm probabilities (table 3, columns 4 and 7) and false alarm rate and false non detection rate are lower. Our algorithm is more competitive on empty videos. For objects on the sea floor, our background image is not well optimized but we reach a false non detection rate lower than the reference algorithm. On the videos used in this work, we detect a lot of false objects but these false detections can be filtered out during the identification step.

5. Conclusion

In this article, we present a novel method based on background subtraction, on comparison and on the use of navigational information. Our algorithm works in three steps. First we preprocessed our images. Besides classical preprocessing using only information present on the image, we use Phong's model and the sun position to limit the light effects. Then we detect objects. To do this we use the background subtraction algorithm. Navigation information are necessary to learn background when the distance is sufficiently high to be sure of the absence of the object (Object position has been indicated by the sonar detection). Finally, we increase the detection rate and decrease false alarm rate with post-processing. The use of navigational information and fusion of it with detection results form a novel result improvement method. Knowing the vehicle's motion, we can define constraints on the detected object in images.

We tested our algorithm on data acquired at sea and show that we improve detection results and decrease false alarm rate, comparing to our former works. All algorithms have been applied on the same videos. Detection results obtained with the presented algorithm are higher than those obtained with the algorithm based on the phase of the image spectrum and with the reference algorithm. False alarms on empty videos are reduced. False negative probabilities have been decreased between the proposed algorithm and the reference algorithm. In a word, we detect more underwater mines than the reference algorithm.

There still are some improvements to bring on the background image, especially when objects are on the sea floor. Then further work will consist in identifying the detected objects. This second step will help to reduce the false alarm rate.

References

[1] Ocean Studies Board National Research Council, *Oceanography and mine warfare*. National Academy Press, 2000.

[2] I. Leonard, A. Arnold-Bos, and A. Alfalou, "Interest of correlation-based automatic target recognition in underwater optical images: theoretical justification and first results," in *Proc. SPIE*, vol. 7678, 2010.

[3] A. Arnold-Bos, J. Malkasse, and G. Kervern, "Towards a model-free denoising of underwater optical images," in *Proceedings of the IEEE conference on Ocean (Europe)*, 2005.

[4] A. Arnold-Bos, J.-P. Malkasse, and G. Kervern, "A pre-processing framework for automatic underwater images denoising," in *Proceedings of the European Conference on Propagation and Systems*, 2005.

[5] B. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, vol. 18, pp. 311–317, June 1975.

[6] I. Reda and A. Andreas, "Solar position algorithm for solar radiation applications," *Solar Energy*, vol. 76, pp. 577–589, 2004.

[7] D. Edgindton, D. Cline, D. Davis, I. Kerkez, and J. Mariette, "Detection, tracking, and classifying animals in underwater video," in *MTS/IEEE Oceans 2006 Conference Proceedings*, 2006.

[8] N. Mandelert and A. Arnold-Bos, “Joint sonar and video sensing for a fire-and-forget underwater mine disposal munition,” in *Proc. 3rd conference on Maritime Systems and Technology*, 2008.