Sonar and Video Perception for an Autonomous Mine Disposal Vehicle

Nicolas Mandelert, Andreas Arnold-Bos

Sonar and Video Perception for an Autonomous Mine Disposal Vehicle

Nicolas Mandelert, Andreas Arnold-Bos

Thales Underwater Systems
Route de Sainte Anne du Portzic
Site Amiral Nomy
CS 43814
29238 Brest Cedex, France

Nicolas.Mandelert@fr.thalesgroup.com

***Abstract* – In the context of modern mine warfare, a solution for future mine disposal consists in using Autonomous Underwater Vehicles (AUVs). These will allow mines to be attacked with limited human intervention. Thales Underwater Systems is currently working on perception algorithms for such an AUV. The vehicle sensors for perception are a sonar and a video camera. Sonar is used at long range for object detection and approach, and for object classification at medium range. The video camera is activated for identification at short range. The vehicle is then guided to the attack position by automatic video tracking. This paper presents the results of the automatic sonar detection, tracking and classification algorithms, and of automatic video identification and tracking algorithms applied on data recorded at sea. It demonstrates how these perception algorithms can provide an autonomous vehicle with high-level information (such as relative positioning and target identification) adequate to find, reach and destroy a mine.**

I. INTRODUCTION

In a mine warfare operation, mine disposal is performed by divers or ROV. To avoid exposition of divers and reduce the time of operation necessary to the disposal of a previously classified bottom or moored target, autonomous mine disposal vehicles (AMDV) will be an efficient system, but only if they are really capable of a fully autonomous mission, from launch to identification and attack.

Thales Underwater Systems is developing a perception module, based on sonar and video sensors, to guide an autonomous vehicle towards a target, and identify it. For this development, sonar and video sensors have been specified and real sonar and video data have been recorded at sea on dummy mines in the condition of a complete AMDV mission. The sonar is a mechanically steered sonar (Tritech Super Seaking). The video camera is a Monacor TVCCD30 black and white analogue camera.

This paper aims at presenting the sensors' specifications, the perception system architecture and the results on real data of the different algorithms that have been implemented for each phase of the vehicle mission.

II. AMDV MISSION CONTEXT

The disposal vehicle is launched after a search operation, which gives a target position and class. Because this operation is conducted at long range, the uncertainty on the target position can be quite high (25 m max. error is our hypothesis), and the target classification is not certain. Thus the first task of the AMDV is to detect and re-locate the target, then to reach it and identify it before attack.

A mechanically steered sonar will be used for detection, re-location, and for approach guidance. At very short range (less than 5m), the video camera will give images for identification, relative positioning of the target and final attack.

The mission from launch to attack is composed of the following phases:

1. Transit to the expected target location
2. Detection of objects inside the uncertainty area
3. Approach of first detected object
4. Elevation estimation of target (if it is a moored mine)
5. Sonar classification
6. Video detection and approach
7. Video identification
8. Video attack guidance

During transit, no sonar or video perception is being used, since the vehicle is guided by acoustic positioning.

The first phase for perception is the detection phase. The vehicle is hovering outside the uncertainty area of the target, so that the sonar is able to scan the whole area, and detect all objects that give an echo in the area. These objects are listed and sorted, so that the first in the list is the most likely target.

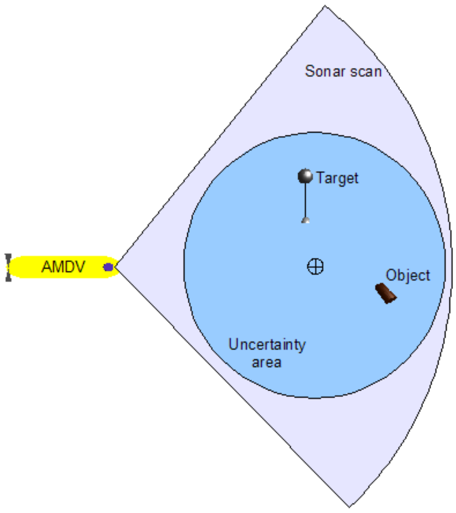

Figure 1: Detection phase description.

Once the list of detected objects is available, the second phase is the approach of the first object in this list. During approach, the sonar is used as a homing sensor, to guide the vehicle towards the contact. The sonar scan is reduced to increase the frequency of measurements. This phase ends when the vehicle is at classification range (5 m) from the target.

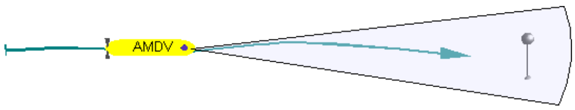

Figure 2: Approach phase description.

During approach, if the target is a moored mine, the elevation estimation phase is activated. It consists in changing the pitch of the vehicle from –20° to +20°, and measuring the echo strength, in order to estimate the contact elevation, based on the sonar's vertical aperture. The vehicle goes then to the estimated immersion and continues approach.

After approach, sonar classification is performed. The vehicle is stabilized in front of the contact, and contact characteristics are extracted from the sonar images.

Then the video detection phase begins. The sonar guidance is still used until the contact is detected and localized in the video images. As soon as the video is able to detect the object, the vehicle is guided by the video.

The identification is performed at about 2 m range from the contact. At this distance, the object is big enough and it is still entirely visible in the camera field of view.

Once the object has been identified as the target, the attack phase begins. The vehicle goes to attack position. This position is defined so that the shaped charge of the vehicle is pointed to the center of the target's explosive. During this phase, the video attack tracking provides the vehicle management with the relative target position and attitude.

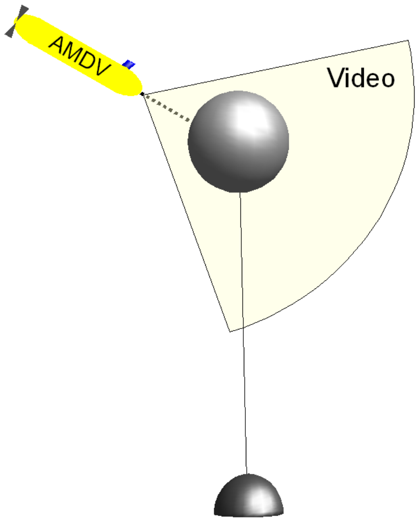

Figure 3: Attack phase.

In case the selected object is not identified as the target, the vehicle goes to the next contact in the detection list, and to do that, a perception re-localisation is performed before a new approach. The re-localisation phase is identical to the detection phase, except that the uncertainty of contact position is smaller than the initial target position uncertainty.

III. PERCEPTION ALGORITHMS

The principle of the perception module is to be a high level sensor, that is able to give absolute and relative positioning data to the vehicle's mission management system, and to perform video-based identification. It receives sonar and video data from the sensors, and mission data from the mission management of the vehicle. For guidance phases, the input of this module is always a contact position (absolute or relative, depending on the phase) and an associated uncertainty. The output is a measured position and uncertainty.

The perception module is composed of seven main algorithms:

1) Sonar detection is in charge of building the list of contacts inside the target uncertainty area, and to sort it on a likelihood criterion. The position of each detected echo is given in absolute Cartesian coordinates.
2) Sonar approach gets the last relative position of current contact in the sonar sector. If the contact is found around its estimated position, its new bearing and range is sent back to the vehicle management.
3) Elevation estimation computes the relative elevation of a moored mine using variation of vehicle pitch to find the direction of the echo maximum

level. The estimated elevation gives the mission management a better estimation of the mine position in the water column, and the vehicle altitude can be set to this estimated altitude.

4) Sonar classification extracts the length, width and height of the contact from the sonar image. It computes the echo's and shadow's shape and calculates the object's dimensions from it.
5) The video detection and localisation task is in charge of deciding if the object is visible in the video images, and to give mission control its relative position. Once the object is detected, its position is calculated and video is used to continue the approach.
6) The video identification task decides if the object is the target or not. The decision is based on a database of known mines, composed of spheres, cylinders and Manta (truncated cones) mines. If the identification is positive, the identification process gives the relative orientation of the target to the mission management.
7) The last task is the final attack guidance. A target point is set on the video image, and this point is tracked during the attack phase, until the firing position is reached. The video tracking is able to keep the attack point position even if the mine is very close and not entirely visible.

During the sonar phases, the sonar setting are calculated by the perception module to optimise the sonar range, direction and scan width according to the current phase and relative target position.

IV. AUTOMATIC PROCESSING RESULTS

Each task of the perception module has been tested on real data recorded at sea, with the specified sonar and video camera. The conditions of acquisition were as close as possible to the conditions of a real AMDV mission, except that the vehicle was manually driven. As the data was recorded, there was no interaction between vehicle navigation and automatic perception.

The data have been collected in La Ciotat (southern France), on a sandy seabed. Several dummy mines have been deployed on the area. Two tethered spheres (at a 5 m and 17 m altitude), one cylinder, one Manta and two other objects have been put in a square of 15 m width. The following results show these objects, in sonar and video data.

A. *Detection*

The detection task is applied on several successive scans on the area. First, a detection is performed on each scan, then recurring contacts over several scans are associated, giving more confidence on the detected echoes. Then the list of contacts is created, and matched with known objects in the area. As mine hunting has been performed on this area before, it is useful to get a list of known contacts as an input and to try to match this list with the detected contacts. By doing that, if the association is good, the chance to get the right contact in the first place is increased.

The detection result is displayed on the figure below. The sonar scan is shown in green-yellow, and the six known objects are displayed at their true position. The target is the moored mine at position (0,0) and the uncertainty area is the green circle around it. Each detected contact is represented in red (with its uncertainty area), and a confidence level is written in white close to it. An unique ID for each contact is written also next to the confidence level.

This figure shows that the algorithm detects all known objects, and some other echoes mainly due to noise or true unknown objects. The contact with the highest confidence is the moored mine. Thus the detection will be successful and leads to approach of the target in the first attempt.

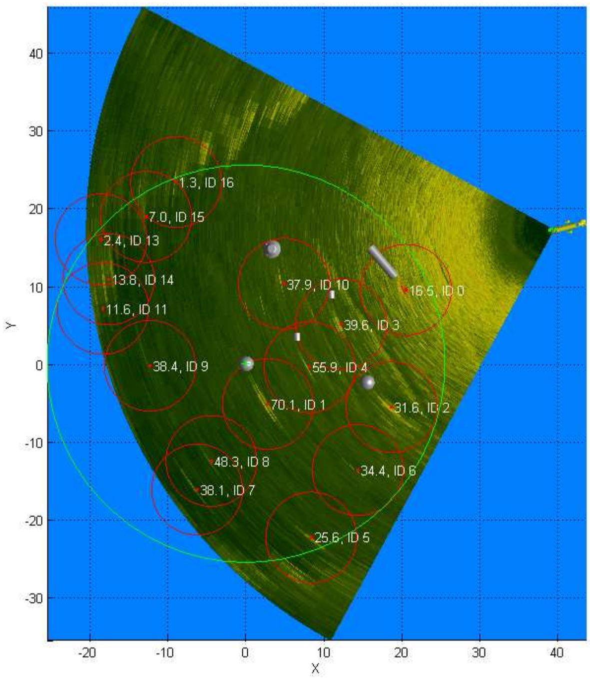

Figure 4: Detection phase results .

B. *Approach*

During approach, the estimated position of the current contact is received by the perception, which sends back an updated measured position, in range and bearing. If the echo is not detected, the uncertainty of the contact position is increased, thus making the search area of the contact bigger. The approach will then be successful even if contact is lost for a limited time.

The following figure shows on the left one sonar scan with the current contact position in red, and the result on a full sequence of the measured range and bearing (in red), compared to the true position (manually annotated) in green. On this example, the tracking is successful from a 30 m range to about 8 m, even if it failed on some sonar scans.

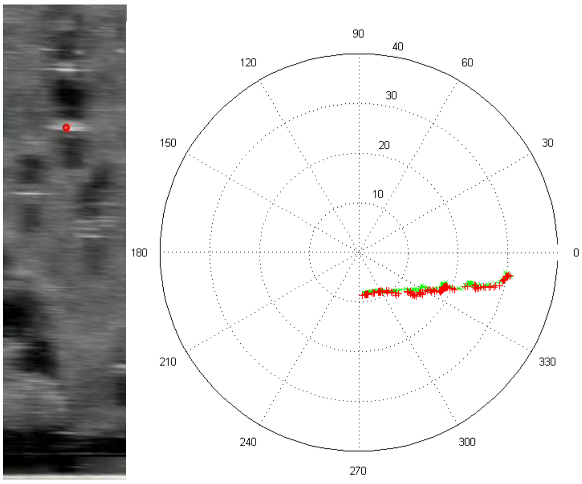

Figure 5: Sonar approach results

C. Elevation estimation

Elevation estimation is performed only if the target is a moored mine. During approach, at about 20 m range, the vehicle is stopped and changes its pitch from –20° to +20°. During this phase, the approach task continues to update the contact position and the echo level is measured in the sonar image for different vehicle pitches. When enough data have been gathered, the elevation of the contact is estimated by fitting the theoretical sonar vertical lobe to the measured levels. The following figure shows the measured and theoretical curves as a function of the vehicle pitch, and the estimated and true elevations for a tethered sphere at an altitude of 17 m.

As the sonar vertical aperture is 20°, the resolution of this method is not very high (several degrees), but the estimate is sufficient to continue at the correct altitude until video detection, as a 3° error gives only a 1 m altitude error at a 20 m range.

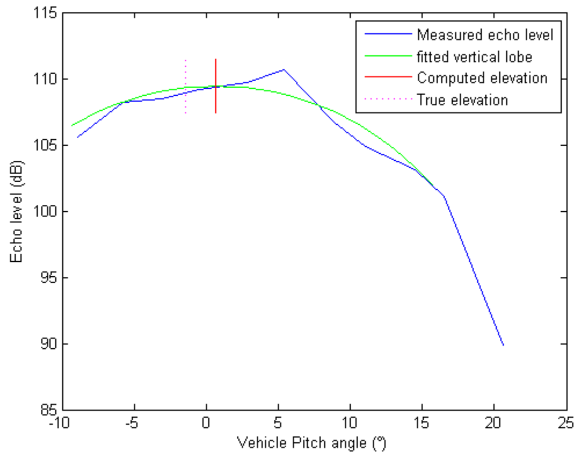

Figure 6: Elevation estimation result.

D. Classification

Classification is performed at a range of about 5 m, to take advantage of a good resolution on the object (about 13cm at 5m range). The vehicle altitude is about 1 m to optimize the grazing angle for shadow detection.

The echo shape is extracted on the image, and shadow is extracted behind echo, using a snake algorithm (as described in ref. [1]).

The two images below give the result of this segmentation on a cylinder image and a Manta. The extracted echo is in red and the shadow in green. The dimensions of the object are computed using the shadow length for the height of the object, and the echo width and length for the object's width and length.

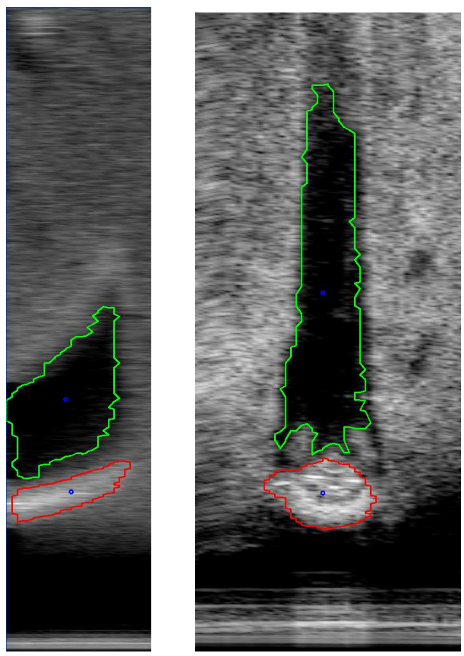

Figure 7: Classification results on a cylinder and a Manta mine.

E. *Video identification*

The identification is an automatic processing of the video images, that allows deciding if the current contact is a known mine or not. There are three known mine types in the database used for identification: "Manta", "tethered sphere" and "cylinder".

Identification is based on a pattern matching algorithm [2], applied after some image enhancement to reduce the effect of water absorption [3]. The known mines are represented in a database by their contours, as illustrated below.

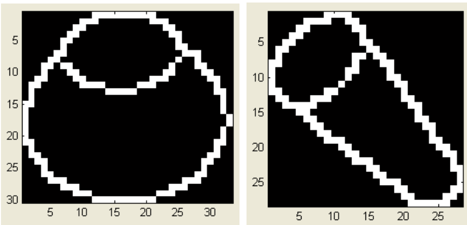

Figure 8: Example of known mine shapes.

The identification process tries to match these shapes with contours from the video images. When the matching is good enough, the decision can be taken, and the relative orientation of the mine is given by the best matching shape orientation.

To increase the confidence in the decision, the coherence of the identification is checked along a video sequence. If the best matching shapes jump from one image to the next, then the identification is not good. The best result is obtained when the target is seen from several angles, thus for good performances of video identification, the vehicle has to change its pitch during this phase.

The following images show in green the best shape identified on the sequence. The Manta and the cylinder are correctly identified, and their relative positions are given by the shape orientation and position in the image.

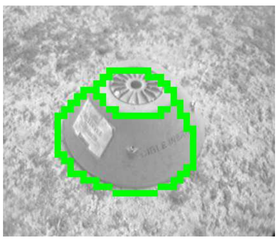

Figure 9: Identification result on a Manta.

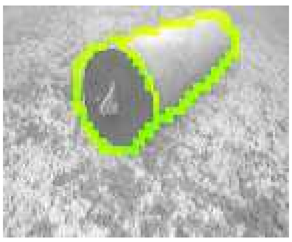

Figure 10: Identification result on a cylinder.

F. *Video attack guidance*

For video attack guidance, an attack point is designated on the image, and this point has to be tracked during the attack phase, until the vehicle reaches its firing position. An algorithm based on optical flow has been used for this task [4]. This algorithm automatically extracts good features to track in the image, and estimates their new positions in the next image. The following images show the attack point (in red), and the tracked points (in blue) in an attack sequence on the cylinder. Even when the mine becomes bigger than the image, the tracker is able to re-initialize good features to track, and the estimation of the attack point displacement is still estimated correctly. The final position on the mine surface of the attack point is very close to the initial position. This algorithm is very robust and quick (50 ms).

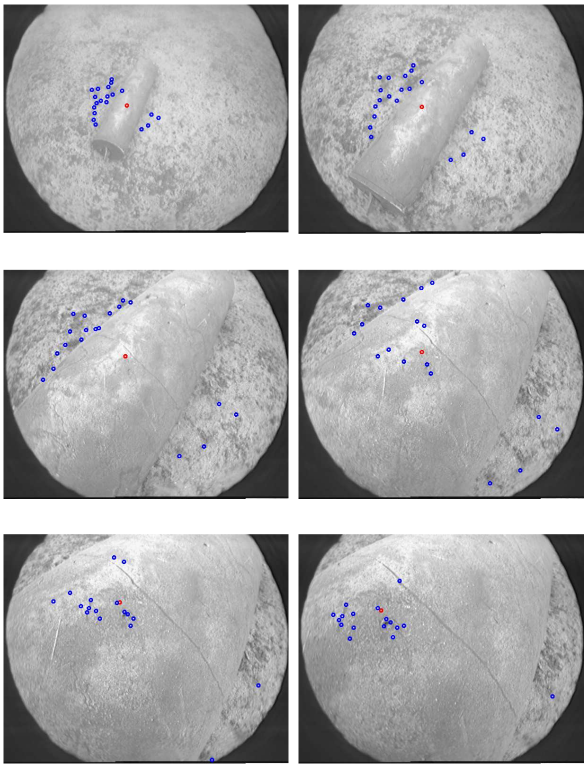

Figure 11: Attack point tracking sequence.

G. *Processing time of the perception*

The perception module is currently being embedded in a demonstrator. All algorithms implemented have been developed and tested to allow real time processing on a maximum of 3 boards with Intel® Core™ Duo processors. The most difficult phase in terms of processing load is the video detection, when both sonar and video processing are working simultaneously.

The processing needs can be dispatched as follows. The sonar detection is not critical as the sector acquisition time is of several seconds. The approach sonar processing needs a maximum frequency of 10 Hz (corresponding to the time of acquisition of a 20° sonar sector at 10 m range).

The sector acquisition for classification takes about 1 s (100° sector, with 0.9° sampling at 10 m range), and classification can be performed in less than 1 s. The video detection and localization is performed at about 10 Hz. The identification task is performed at about 2 Hz and this can be increased if the database is limited to a single object. The video tracking phase is quicker than 10 Hz. This is more than sufficient for the vehicle guidance.

V. CONCLUSION

The perception module presented in this paper is a complete set of sonar and video algorithms adapted to detection, homing and identification of underwater mines for an autonomous mine disposal vehicle. The algorithms have been implemented and successfully tested on recorded real data.

This module is now ready for integration and tests at sea. It will be embedded in a vehicle, and tested at sea in autonomous mode.

VI. AKNOWLEDGMENTS

This perception module for an autonomous mine disposal vehicle has been developped within a project funded by Norwegian FFI and French DGA.

REFERENCES

[1] Franck Fohanno, Nicolas Mandelert, Franck Florin, "Automatic Processing for AUV", Proceedings of OCEANS'06 Asia.

[2] C. Olson, D. Huttenlocher, *"Automatic Target Recognition by Matching Oriented Edge Pixels"*, IEEE Transactions on image processing, Vol 6, n°1, January 1997

[3] Andreas Arnold-Bos, Jean-Philippe Malkasse, Gilles Kervern, *« Towards a model-free denoising of underwater optical images »*, Proceedings of the IEEE conference on Ocean (Europe), Brest, France, 20-23 juin 2005

[4] J-Y.Bouguet, "Pyramidal Implementation of the Lucas-Kanade Feature Tracker". OpenCV Documentation, Microprocessor Research Lab, Intel Corporation,1999.