Joint sonar and video sensing for a fire-and-forget underwater mine disposal munition

Nicolas Mandelert, Andreas Arnold-Bos

Joint sonar and video sensing for a fire-and-forget underwater mine disposal munition

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* – Today, autonomous, fire-and-forget munitions launched from the surface are under study for their potential use in future underwater mine warfare. Without any human intervention, these would navigate towards a designated object, identify it, and destroy it if necessary. Since the weapon would self-destruct in the process, low-cost, off-the-shelf sensors are needed. However, a simple, mechanically steered sonar cannot alone allow for proper automatic target identification and terminal guidance. As an alternative, Thales Underwater Systems is proposing a dual sonar/video sensing approach. Sonar is used for long to medium-range object detection and approach, and for preliminary object classification. Short-range approach, identification and attack are done with a standard video camera. This paper demonstrates the results of our perception algorithms on real data acquired at sea. It shows how the developed perception algorithms can provide the munitions mission management system a significant ability to track and destroy a mine autonomously.**

I. INTRODUCTION

The current trend in underwater mine warfare involves Remotely-Operated Vehicles (ROVs) launched by a surface ship to inspect and eventually destroy mine-like objects previously detected by sonar. However ROVs require the dedicated attention of human operators plus unwieldy fiber optics to relay information to the surface.

The logical next step is thus to develop an Autonomous Mine Disposal Vehicle (AMDV), a kind of fire-and-forget weapon that would be launched on the mine-like echo. The AMDV would then automatically transit to the area of operations, survey the potential targets in the area, and if a mine is found, destroy it. Mine/non mine discrimination is insufficient, since the attack strategy varies with the mine type: the shaped charge must explode at a specific point with a specific angle to be efficient. In this context, human intervention is supposed to be as limited as possible, except for an optional confirmation before attack, which would involve the acoustic transmission of a single image by an acoustic link.

Thales Underwater Systems is developing a complete perception module (PM) for such an AMDV. Since the vehicle is disposable, the sensors must as simple as possible and available off-the-shelf. In this case, there is a mechanically steered Tritech Super Seaking sonar, a Monacor TVCCD30 black and white analogue camera and a single laser beam illuminating the scene for optical range determination. Others inputs to the PM are navigation data (estimated absolute and relative position and angles) and contact data coming from an external, third party mission management module. The outputs are the same as the inputs but refreshed by the sensors.

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. The real data was acquired at sea during trials in October 2007 of an AMDV prototype.

II. MISSION OPERATIONAL PHASES

The disposal vehicle is launched after a survey operation, which yields a target position and class (moored or bottom object) as well as the position of surrounding contacts on the area of operations (AO). Because this operation is conducted from a long range, typically 500 m, the uncertainty on the target position can be quite high (25 m max. is our hypothesis), and the target classification is not certain. Before launch, the AMDV is fed a map containing the contacts and the target. The first task of the AMDV is to transit from the ship to the AO. This is done without using PM, which only begins to be used once near the AO. Figure 1 illustrates the different operational phases happening on the area. We now detail the purpose of each different operational phase and software modules associated to it.

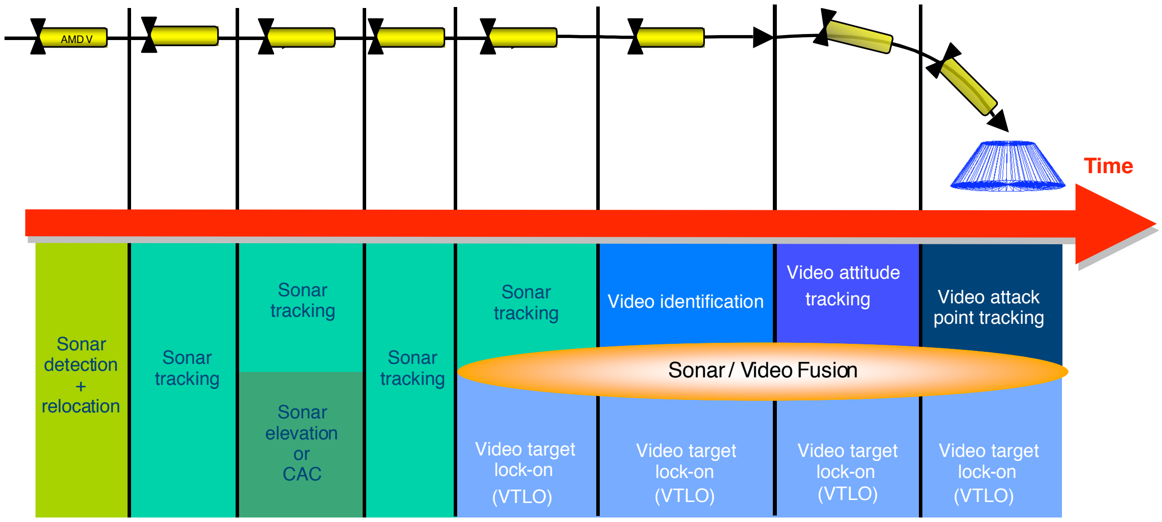

Figure 1: operational phases for AMDV. Sonar modules are in green and video modules in blue.

A. *Sonar-Only Phases*

1) Sonar detection This is done at the border of the AO with a long range, wide scan sonar setting (see figure 2). The absolute position of the contacts found by sonar, is matched to the position of the contacts found in the map; this allows for some degree of drift correction. The contact closest to the theoretical position of the mine-like echo is then promoted to target status.

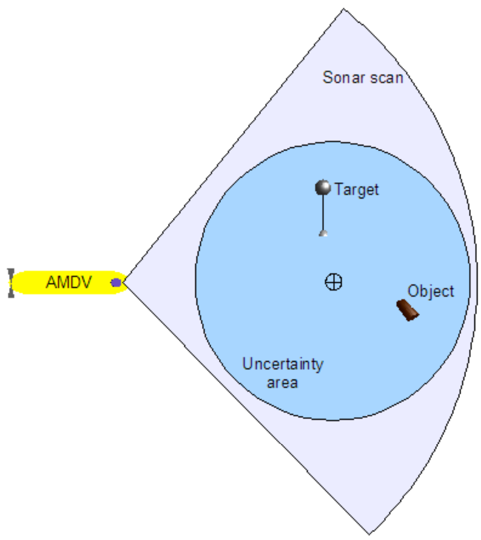

Figure 2: detection phase

2) Sonar tracking During approach, the sonar is used as a homing sensor, to guide the vehicle towards the contact. The sonar sector is narrowed to increase the frequency of measurements. This phase ends when the vehicle is at classification range (5 m) from the target.

3) Sonar Target Elevation Estimation If the target is a moored mine, the elevation estimation phase is activated at about 20 m from the mine. 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.

4) Sonar Computer-Aided Classification After approach, sonar classification is performed. The vehicle is stabilized in front of the contact, and contact dimensions are extracted from features in the sonar images (echo and shadow size). This information can be used as a complement to information retrieved in the video images.

B. *Sonar To Video Transition*

When the vehicle is too close to the object, it is impossible to accurately track it by using sonar only, since the echo becomes too large and too hard to segment. The sonar guidance is still used until the contact is detected and localized in the video images. This is done by the Video Target Lock-On (VTLO) module, which is also in charge of determining the distance to the target by using the laser beam (see §III.E). As soon as VTLO sees the object, the vehicle is guided by video. It is up to a Fusion module to decide which source is used at a given time to provide information to the mission management module.

C. *Video-Only Phases*

1) Video Identification 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. The Video Target Lock-On module is still being used in that phase since the vehicle must be servoed to keep the target in the image while identification runs.

2) Video Attitude Tracking (VAT) During this phase, the video attack tracking provides the vehicle management with the relative target position and attitude; the results are merged with VTLO results in a Fusion module. This serves two purposes. First, when the object has been identified as the target, the vehicle must move to a precise attack position so that the shaped charge is pointed at the most vulnerable point of the mine. Also, in this phase, an image can optionally be sent to a surface operator via an acoustic link for identification confirmation. This can take a couple dozen of seconds and VAT/VTLO is used for station-keeping in the meantime.

3) Video Attack Point Tracking When the vehicle is in position, it must advance until it hits the target. In this phase VTLO provides an accurate estimation of the distance thanks to the laser. The impact point is continuously tracked in the image to guide the vehicle. When the vehicle is close enough, it explodes.

If 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. AUTOMATIC PROCESSING DESCRIPTION AND RESULTS

Each task of the perception module was tested on real data recorded at sea in a campaign, which took place in October 2007 at La Ciotat (southern France). The conditions of acquisition were as close as possible to the conditions of a real AMDV mission, and the vehicle itself was designed to the final specifications, except that the vehicle was manually driven. As the data was recorded, there was no interaction between vehicle navigation and automatic perception. Several dummy mines have been deployed on the area, mostly made of a sandy seabed. 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 paragraphs gives more details on the algorithms we implemented and shows results we obtained with the trials data.

A. Detection

The detection task is applied on several successive scans on the area. Detection is performed on each scan, and recurring contacts over several scans are associated to eliminate spurious echoes and derive a confidence criterion. Then a list of contacts is created, and matched with the 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.

One particular detection result is displayed at figure 3. 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). 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 number 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 at the first attempt.

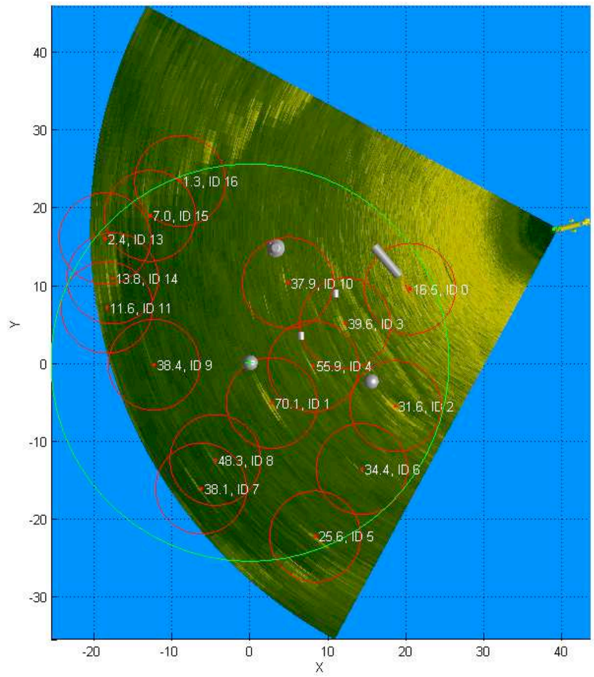

Figure 3: Sample detection phase sonar data with results superimposed on it.

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.

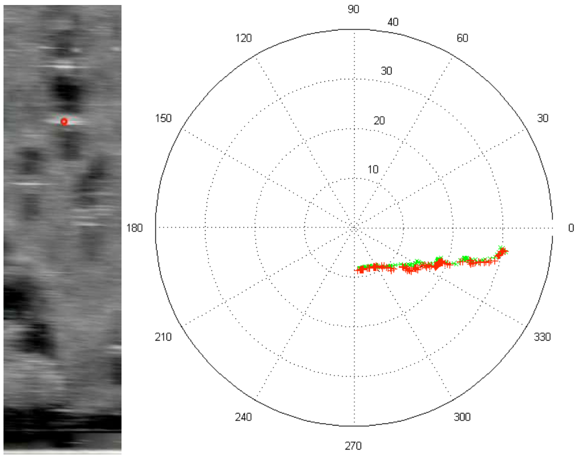

Figure 4: Sonar approach results

Figure 4 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 a manually annotated "true" position 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.

C. Elevation estimation

Elevation estimation is performed only if the target is a moored mine. During approach, about 20 m from the target, 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, fitting the theoretical sonar vertical lobe to the measured levels gives the target elevation. 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.

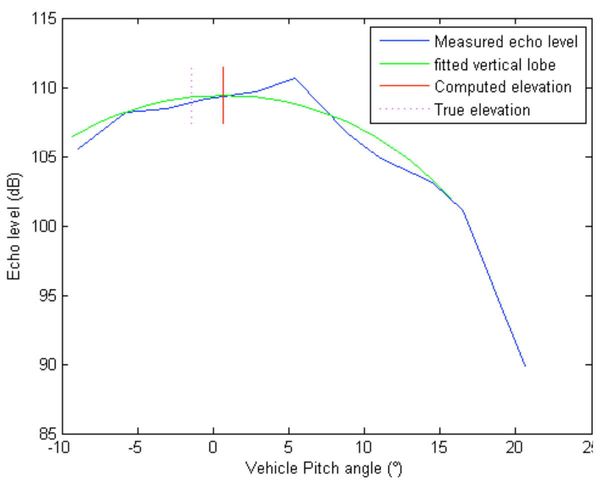

Figure 6: One elevation estimation result done at 20 m from the mine.

D. Classification

Classification is performed at a range of about 5 m, so the objects resolution is high enough (about 13 cm at a 5 m 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 the echo, by 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's length for the height of the object, and the echo width and length for the object's width and length.

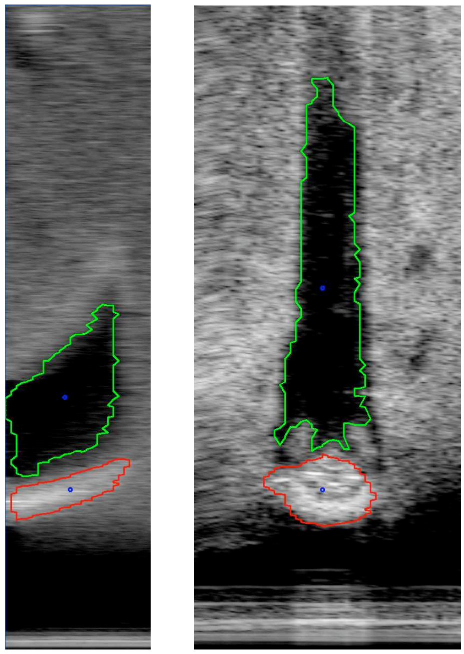
Figure 5: Classification results on a Cylinder and a Manta mine.

E. Video target lock-on and laser range-finding

This module was developed by Cybernétix in cooperation with Thales. Its purpose is to detect whether an object is present in the image, and to locate it. The image is pre-processed and segmented to extract contours. Only contours sufficiently matching simple primitives (lines, ellipses) are retained. The centroid of these primitives gives the target location in the image. The target's image bounding box can also be estimated.

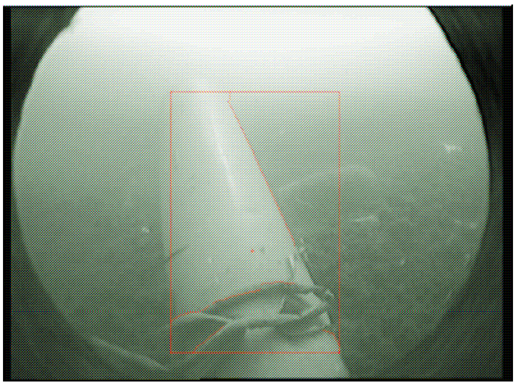

Figure 7: Video target lock-on on a cylinder with bounding box and detected primitives in red.

The module also computes the range to the target by using a laser beam pointed towards the scene. This range is primarily used when very close to the mine, especially during the attack phase. In Figure 8 we present results obtained on real data. The measures are compared to annotations made with a mine model projected in the image. The distance to the geometrical center of the model, can be considered as a pseudo ground truth. The laser does not hit the center of the mine but its surface, so the roughly the radius of the mine must be added to the laser distance to find the pseudo "true" distance. Our conclusion is that the laser distance becomes more or less reliable at about 2.5 meters from the mine but not before.

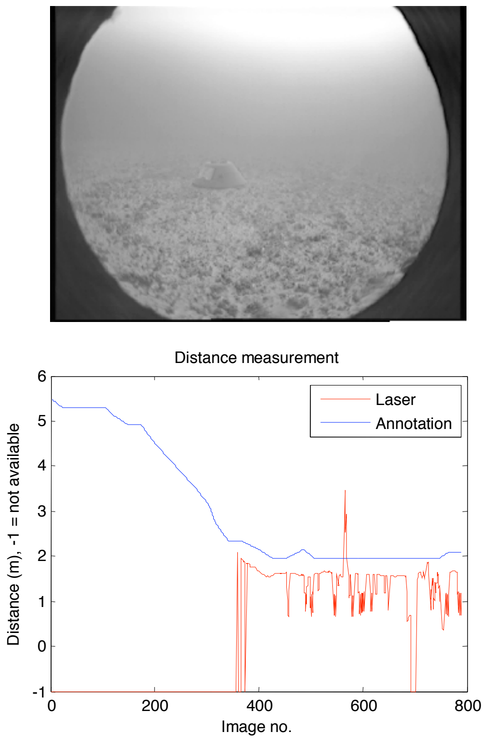

Figure 8: raw laser range measures *vs*. manual annotation for a Manta mine. Value –1 means " no measure".

F. Video identification and attitude tracking

Those two modules use the same method, which is based on a fast pattern-matching algorithm. The patterns are defined as the mine contours. They are labeled in range and attitude: once the best pattern is known, the mine shape, but also the relative mine/vehicle attitude are known too. The consistence of the identification is checked along a video sequence. If the best matching shapes do not fit the vehicle's displacement, then the identification is bad. This allows for deriving a quality criterion. By taking the mine class with the best quality (which must also be above a certain threshold), it is possible to identify the best mine shape. In attitude tracking, the mine class is known and only patterns from that mine class are used to refresh the mission management module with the relative position of the mine.

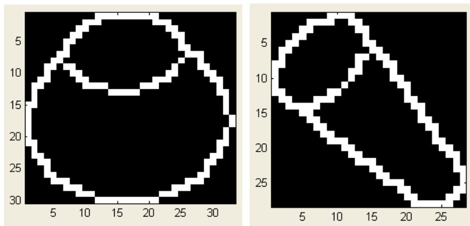

Figure 9: Some mine shapes in the database.

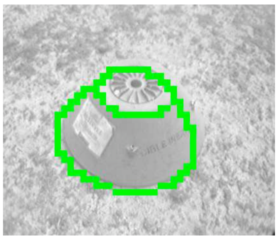

Figure 10: Identification result on a Manta.

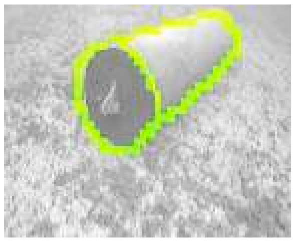

Figure 11: Identification result on a cylinder.

G. Video final attack guidance

Once the vehicle has the correct relative attitude to the mine, it must travel forward to hit the mine while remaining servoed to the final impact point in range, pitch and elevation. Laser range finding is primarily used for range determination. Pitch and elevation is determined by an algorithm based on optical flow [3]. This algorithm automatically extracts tracker points in the image, and estimates their new positions in the next image. The position of the trackers allow for finding the position of the attack point. The following images show the attack point (in red), and the trackers (in blue) in an attack sequence on the cylinder. Even when the mine becomes bigger than the image, the module is able to re-initialize trackers elsewhere, and the estimation of the attack point displacement is still correct. The final position on the mine surface of the attack point is very close to the initial position. This algorithm is very robust and fast (50 ms).

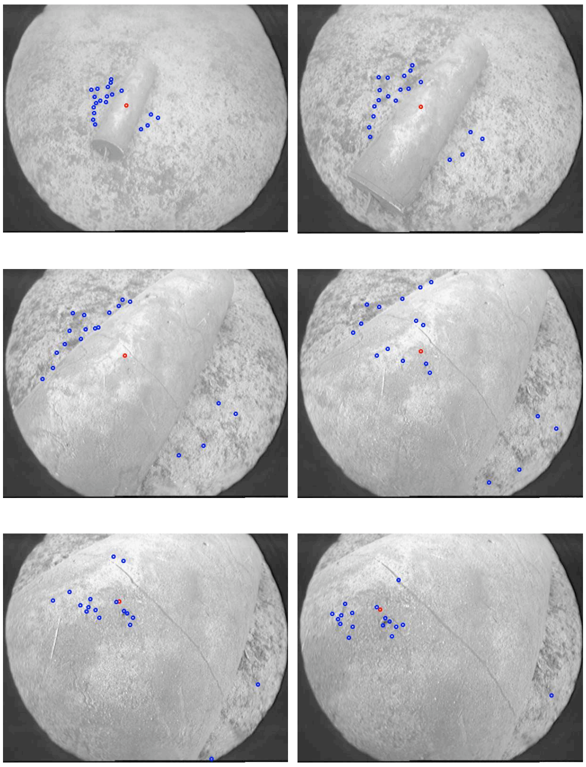

Figure 12: Attack point tracking sequence.

H. Processing time of the perception

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

The processing needs can be summarized as follows. Sonar detection is not critical as the sector acquisition time is of several seconds. Sonar processing during approach must be done in no more than 0.1 s (this corresponds to the duration of the acquisition of a 20° sonar sector at a 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 5 Hz. Identification is performed at about 2 Hz and attitude tracking is faster since the database is limited to a single type of object, that is, the one that has been identified. The optical flow algorithm itself is quicker than 20 Hz but image acquisition limits the processing frequency to about 7 Hz. This is more than enough for the vehicle guidance.

IV. 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. The final closed-loop system (PM+mission management module+guidance) already gave satisfying results in a virtual environment where the vehicle and sensors were simulated as accurately as possible. Final sea trials with the autonomous mode activated are now expected to take place by end 2008.

V. 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, 2006.

[2] 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 June 2005

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