AUTOMATIC TARGET RECOGNITION FOR THE HUGIN MINE RECONNAISSANCE SYSTEM

Ø Midtgaard Norwegian Defence Research Establishment (FFI), Kjeller, Norway
P E Hagen Norwegian Defence Research Establishment (FFI), Kjeller, Norway

1 INTRODUCTION

In spite of considerable research efforts over the last two decades, Automatic Target Recognition (ATR) systems for high-resolution sonar images have not yet been considered sufficiently reliable for regular operational use. This may be due to unrealistic expectations of achieved performance level, and correspondingly improper use of such systems, but may also signal a continued need for:

- Improved detection and classification algorithms adapting to varying seafloor, target and sonar conditions
- Improved sensor data providing better discrimination between targets and non-targets
- Improved exploitation of all available sensor data through data fusion

FFI seeks to address all three issues within the HUGIN Mine Reconnaissance System (MRS) development. In cooperation with Kongsberg Maritime, FFI has developed an interferometric Synthetic Aperture Sonar (SAS), the HISAS 1030 [1]. This sonar will be the main payload sensor on the new HUGIN 1000-MR Autonomous Underwater Vehicle (AUV), which is designed for Mine Counter Measures (MCM) and Rapid Environmental Assessment (REA) operations [2]. The first vehicle will be delivered to the Royal Norwegian Navy for deployment on an MCM vessel in mid-2007. The HUGIN MRS post-mission analysis (PMA) system seamlessly integrates components from different vendors, including FFI's ATR module and FOCUS toolbox [3] for synthetic aperture processing. This paper presents our strategy for development and integration of the ATR module, together with a description of the initial system components.

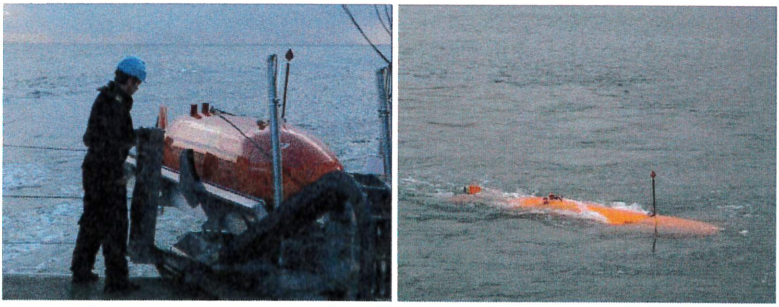

Figure 1. HUGIN 1000 AUV immediately before and after launch during operations in Italy in 2005.

2 DEVELOPMENT AND IMPLEMENTATION

In MCM operations, *time* is usually the most critical resource. The MCM force needs to complete its task and report the result as rapidly as possible. Any delay in the MCM effort will delay other assets

– be it larger military vessels in transit, oil tankers or other critical supplies, or an amphibious landing force.

In AUV based mine hunting, the time to clear an area can broadly be divided into three main parts:

1. AUV survey – gathering data for detection and classification
2. PMA – data processing, detection and classification
3. Identification and disposal

The survey phase is sped up by increasing the area coverage rate (ACR) of the AUV. By replacing a high-frequency side scan sonar (SSS) with a high-resolution SAS, the ACR can be increased by a factor of 5 or more.

The ID/disposal phase is sped up due to the increased classification performance attainable by a SAS [1,4], which will reduce the number of false contacts to be identified.

The PMA phase is probably the most difficult to speed up. Having a higher-resolution sensor that provides a range of different products means that more contacts may have to be analysed, and more time may be spent analysing each contact. A few extra minutes spent analysing a contact can be well spent; the time for identification (using divers, ROVs or one-shot disposal systems) can easily be 30 minutes per contact. However, a well-functioning ATR system can also reduce the time spent in PMA.

If automated detectors and classifiers with sufficient performance are available, operators will not have to go through the entire data set, but may instead concentrate on the automatically detected and classified contacts. This is indeed a long-term goal of the HUGIN MRS development programme. This level of performance does not appear to have been reached by any ATR system in use today, despite a substantial research effort. One reason for this may be limitations of the sensors and algorithms used; another may be the psychology of automated vs. manual target recognition [5].

In HUGIN MRS, a stepwise introduction of ATR capability is planned, roughly as follows:

1. Use clutter density estimators and/or seafloor complexity classifiers to help the operators decide in what areas to focus the PMA effort. Analysis then proceeds fully manually – the operators go through the entire data set using three displays in parallel: a sonar image waterfall display for detection, a target analysis display for classification, and a geo-referenced mosaic and target position display for data association.
2. Run ATR on all sonar data in the background, in parallel with manual analysis. After manual analysis is completed, associate manually detected contacts geographically with ATR detections. Provide the operator with sonar images of the most likely candidates that were not detected manually.
3. As above, but also display ATR detections in the geo-referenced display, with an optional symbol in the waterfall display. Only contacts with confidence above a user selectable threshold, or a given number of contacts per km^2, will be displayed.
4. Replace the manual detection with automated detection and pre-classification in areas automatically deemed suitable for ATR – typically, areas with very low clutter density and low complexity, such as flat, sandy bottom. A simple solution may be to suggest manual detection in all areas (or survey lines) where the ATR contact density is above a given threshold. Classification will proceed based on the list of ATR contacts, typically sorted by detection confidence.
5. As (actual and perceived) ATR performance improves, gradually raise the threshold for when to apply ATR.
6. When ATR performance is sufficiently high, move the ATR processing chain into the AUV. This will facilitate transmission of detections to the host vessel during mission, and in-mission identification using an optical sensor and/or circular SAS.

For each step in the above list, a higher level of performance is required for the ATR system to actually be beneficial for an operator. Level 1 is likely to be very useful with current algorithms; level 2 and 3 are potentially productive, and will at worst be deactivated by the operator in case of algorithm malfunction. The most significant, but also most difficult, leap is from level 3 to 4, as it requires an operator to trust the ATR to not miss any contacts the operator himself would detect. This trust must be built as the operators get familiar with system performance during steps 2 and 3.

The first versions will rely on a human operator to perform the final object classification, and are thus not truly automated systems. At levels 4 and 5, the main benefit will be faster post-mission analysis as the operator does not have to evaluate all sensor data. The requirements on algorithm performance are that a high probability of mine object detection is achieved, as detections missed by the ATR will never be seen by the operator; and that the number of false alarms is manageable for human evaluation within given time constraints. FFI aims to gradually develop an ATR system for operation onboard the HUGIN AUV during MCM missions (level 6). At this level, the ATR system will also reduce the identification and disposal time – as identification of many mine-like contacts can be performed during the first (survey) phase.

3 SENSOR DATA AND PROCESSING

As mentioned above, insufficient sensor data quality is one possible explanation for the inadequate performance of ATR systems today. The most common AUV based MCM sensors today are inexpensive, very high frequency (900 kHz or higher) short-range side scan sonars that provide good resolution but contain numerous response artifacts (saturated targets, strong surface returns, poor SNR, etc); and lower-frequency (400-500 kHz) high-quality side scan sonars with longer range but somewhat poorer resolution. A high-performance SAS, such as the HISAS 1030 on the HUGIN 1000-MR AUV, should provide significantly increased data quality, and thus boost ATR performance potential.

A novel issue with the HUGIN PMA is the close interaction between ATR and SAS processing. The ATR algorithm will not only process the streaming SAS data, it will also use its detection results to selectively request various modes of SAS re-processing for regions of interest. This data will then be ready without delay to an operator performing final object classification at the same regions of interest. The production of this augmented classification data may be prohibitive for the complete mission due to high computational load or it may require input parameters from the detection process.

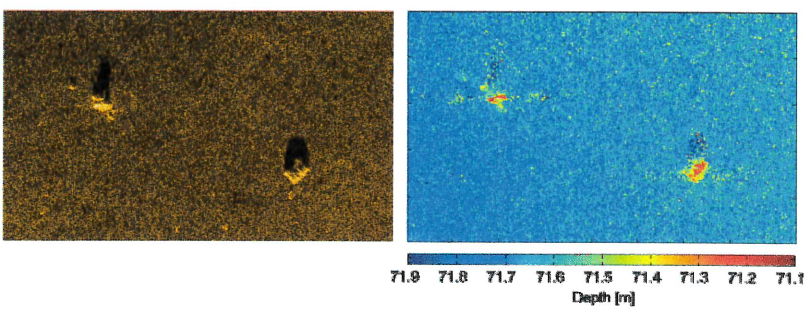

Figure 2. Imagery and bathymetry of rocks and barrel obtained with the HISAS prototype.

The increased along-track resolution provided by SAS processing is expected to be a vital prerequisite for reliable mine-object classification based on response shape details. The authors believe, however, that MCM sensor development should comprise more than just improved image resolution. Intelligent fusion of diverse types of data e.g. multi-aspect, multi-frequency sonar and bathymetry data may increase ATR robustness. The HUGIN ATR system will utilise both sonar imagery and bathymetry for enhanced performance (Figure 2).

4 ATR PROCESSING SCHEME

4.1 Overview

The ATR processing chain for SAS data is displayed in Figure 3. The blue boxes show ATR modules, while the grey boxes show FOCUS and interactive data analysis modules. The streaming sonar imagery and bathymetry can be produced by the FOCUS toolbox at typical resolution of 3-8 cm for the complete mission. Following some basic pre-filtering such as normalisation and denoising, the data is fed into two modules for detection of proud and short-tethered mines, respectively. Each detected contact is assigned a confidence value indicating its degree of mine likeness and its position in global coordinates is calculated.

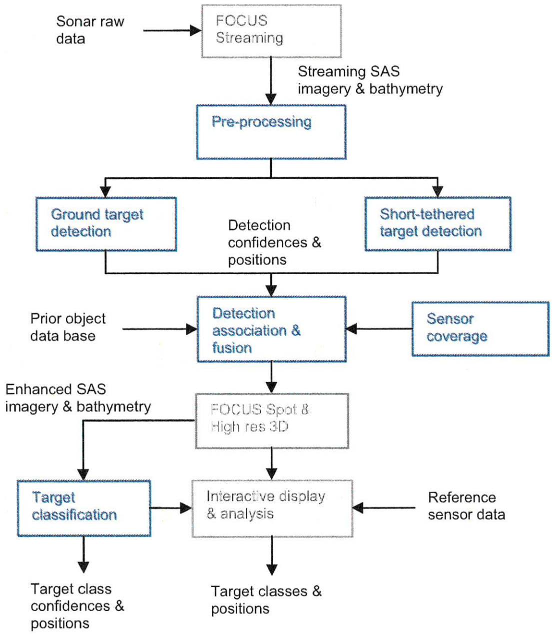

Figure 3. Main components and data flow in ATR processing string for HUGIN MRS.

Detections from different sensor views and algorithms are fused to produce a single confidence value for all detections associated the same seafloor object. The total sensor coverage is also estimated and used to identify negative (missing) detections, which contradict the presence of a mine-like object at the given seafloor position.

FOCUS modules are then invoked to produce enhanced classification data for small regions (typically 15m x 15m) encompassing the fused detection clusters that have high mine-likeness score. These data may include full resolution SAS imagery and bathymetry, multi-aspect imagery and Fixed Focus Shadow Enhancement (FFSE) [6]. For post-mission analysis, the final object classification will be performed by a human operator. FFI is however investigating methods for automated classification that can be useful operator support or a required part of an in-mission ATR system implemented on board the AUV. Reference sensor data from earlier missions may be imported for detailed change detection. The final results are classes and global positions of the detected targets.

4.2 Data Normalisation

The purpose of data normalisation is to produce an image with a more consistent background level throughout the image by removing large-scale response variations. These variations are typically caused by varying seafloor reflectivity or imperfect signal gain compensation in sonar images and sloping terrain in bathymetry data. Normalisation enhances small, local discontinuities such as mine object responses. It is thus used for pre-processing to improve the performance of succeeding detection algorithms both for sonar imagery and bathymetry data. Normalisation along individual image range columns have been reported [7], but in our work a 2D technique has been developed.

First, a low-pass version of the data is produced e.g. by convolution with a mean or median mask whose length and width are smaller than the input parameters Δx and Δy, respectively. Following this an estimate of local response level at image position (i,j) is obtained as the median of the lowpass values in the nine locations shown in Figure 4, consisting of the pixel (i,j) itself and the eight-neighbour pixels displaced as ($i\pm\Delta x$, $j\pm\Delta y$). Sensor altitude data is used to exclude water column pixels from the background response estimate. The parameters Δx and Δy determine the maximum dimensions of response deviations that are preserved during normalisation. The normalised image response for pixel (i,j) is finally obtained by dividing or subtracting this local background response from the original pixel value, depending on whether a multiplicative (e.g. sonar magnitude) or additive (e.g. bathymetry) image model is appropriate.

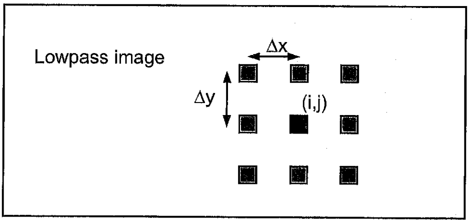

Figure 4. Nine-pixels median mask for estimation of local background response for pixel (i,j).

4.3 Ground Object Detection

A simple, yet effective, method for detecting mine-like, proud bottom objects in sonar images is deformable match filters [7]. They typically consist of a highlight (echo) mask followed by a shadow mask at a variable distance in the positive range direction. The filter is convolved with the sonar image, calculating the average response within the highlight mask at each pixel. The shadow mask response is calculated for all positions within a preset distance interval behind the highlight mask, and the minimum value is selected as shadow output for the current filter position.

Different combinations of the echo and shadow mask responses yield match filters with different characteristics. Three filter examples are:

$$mf_1 = (echo - bg) + \alpha(bg - shadow) \tag{1}$$

$$mf_2 = echo/(shadow + \delta) \tag{2}$$
$$mf_3 = \max(echo - bg, 0) * \max(bg - shadow, 0) \tag{3}$$

Where $echo$ and $shadow$ are the average response values within the highlight and shadow mask, respectively. bg is the value of a featureless seafloor (bg=1 with our normalisation method), α >1 is a weight factor to increase contribution from shadow responses and δ is an infinitesimal value introduced to avoid division by zero.

The outputs of the three filters are visualized in Figure 5. The weighted sum filter in (1) produces large output values for strong echoes, even when shadow response is close to the background level. The output from the ratio filter in (2) increases dramatically as shadow response approaches zero. The product filter in (3) requires both small shadow and high echo values to produce large output values.

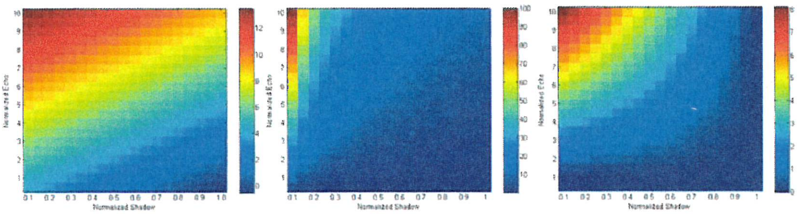

Figure 5. Filter outputs with typical normalized echo and shadow responses for match filter no 1 with α=5 (left), no 2 (centre) and no 3 (right).

The size and shape of the echo and shadow masks should ideally match a given mine type and orientation perfectly. However, to avoid convolving the input image with a large number of different match filters, a generic filter with echo and shadow masks that are narrow in the along-track direction has been chosen. The shadow mask's across-track length increases linearly with range.

To minimise image edge induced artifacts, sensor data is filtered in large blocks, each covering a large seafloor region. As this region may contain varied seafloor conditions, a filtered image threshold based on global image statistics may be unsuitable, and thresholds are thus calculated from local neighbourhood statistics.

4.4 Short-Tethered Object Detection

Due to the varying aspect angles along the synthetic aperture used for building the SAS image, shadow responses from tethered objects are partly filled in with seafloor backscatter. These shadows are thus typically weak and indistinct. Figure 6 presents a SAS streaming image of an acoustic release transponder tethered 3-4 m above the seafloor at around 55 m range. The data was obtained with the HISAS prototype and processed with the FOCUS toolbox to an image resolution of 3x4cm.The strongest echo corresponds to the transponder floatation jacket, while the anchor shows up at slightly larger slant-range as a highlight with a shadow right behind. The transponder shadow is hard to recognise, but can be restored through FFSE processing [6] which also slightly defocuses the surrounding seafloor. Features like lengths of object and tether can then be estimated from the shadow and echo responses. FFSE is performed for a given input range, and thus requires that the tethered object is already detected by another method.

The fact that shadows of tethered objects move along the seafloor when sonar aspect-angle is varied, can however be utilised in a detection algorithm identifying this feature in multi-aspect images. For an interferometric sensor such as HISAS 1030, detection can also be based on bathymetric data. FFI is currently investigating these and other methods for short-tethered object detection in SAS data.

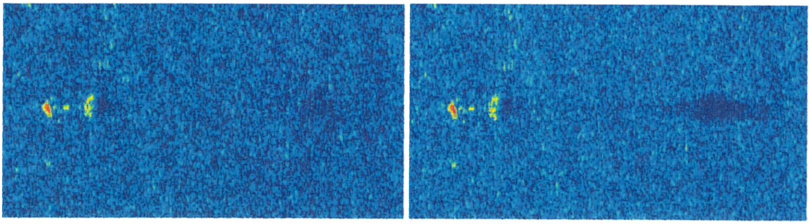

Figure 6. Left: Streaming SAS image of short-tethered acoustic transponder with blurred shadow. Right: Transponder shadow restored through fixed focus shadow enhancement. Dynamic range shown is 40 dB.

4.5 Detection Fusion

Several clues regarding the possible presence of a mine-like object may be found during MCM mission evaluation. The evidence consists of detection results for a given seafloor position from one or more of the following sources:

- Multiple sensor views due to survey line overlap or cross-hatching
- Multi-sensor data (e.g. SAS and nadir gap filling sensor)
- Multiple datasets from single sensor (e.g. image and bathymetry from interferometric sonar)
- Multi-mission data (e.g. change detection)
- Multiple ATR algorithms run on same data set

FFI has developed a detection fusion mechanism for all these cases using Dempster-Shafer theory of evidence and creating the probability mass functions from detection performance estimates [8,9].

Each detection supports the hypothesis *M: there is a mine-like object at the given position*. The complementary hypothesis M^C*: there is **not** a mine-like object in the given position*, is supported by negative detections. A negative (missing) detection occurs when no detection is called at an imaged seafloor position and a detection was made at this global position in sensor data from another mission track (or a detection was called by another ATR algorithm, sensor, etc). After associating all positive and negative detections corresponding to the same seafloor position, a combined belief value for hypothesis M is calculated. A detection with low confidence value adds less support to hypothesis M than a detection with high value. A missing detection at extreme sensor range where the probability of detection is low, will similarly add only a small amount of support to hypothesis M^C.

If available, a prior data base with position coordinates of known non-mine, mine-like bottom-objects (NOMBOs) can be imported as certain non-mines in the fusion mechanism, thus eliminating associated detections in the current mission.

5 CONCLUSIONS

FFI has developed an initial version of an ATR system for detection of mine-like ground objects in high-resolution sonar images. This will be delivered to the Royal Norwegian Navy as part of the HUGIN 1000-MR system in 2007. Previous sonar image ATR systems, developed by different institutions, appear to have failed to meet the expected performance in varied environments. Recognising this, our ambitions for the first version are limited to supervised processing in benign seafloor conditions. FFI will continue its development efforts to improve sensor data quality, ATR algorithms and data fusion. Based on operational experience and new development results, improved capabilities will be added in later versions and degree of system automation gradually increased.

6 REFERENCES

1. P E Hagen, R E Hansen and B Langli: Interferometric synthetic aperture sonar for the HUGIN 1000-MR AUV. Proc. UDT Pacific 2006 Conf. (2006).
2. P E Hagen, N Størkersen and K Vestgård: The HUGIN 1000 military AUV system. Proc. UDT Europe 2004 (2004).
3. R E Hansen, T O Sæbø, H J Callow, P E Hagen and E Hammerstad: Synthetic aperture sonar processing for the HUGIN AUV. MTS/IEEE Oceans '05 Europe (2005).
4. P E Hagen, R E Hansen and Ø Midtgaard: SAS and side scan sonar systems compared: Experimental results from HUGIN AUVs, Proc. Underwater Acoustics Measurements 2007, Crete, Greece (2007).
5. R T Kessel: Apparent reliability. Defence R&D Canada Atlantic Tech. Report 2005-155. Dartmouth, NS, Canada (2005).
6. R E Hansen, J Groen and H J Callow: Image enhancement in synthetic aperture sonar. Int. Conf. on Detection and Classification of Underwater Targets. Proc. of the Institute of Acoustics, Vol. 29, Pt 6. (2007).
7. G J Dobeck: Fusing sonar images for mine detection and classification. Proc. SPIE '99, Vol. 3710, pp 602-614. (1999).
8. Ø Midtgaard: HUGIN MRS - object detection fusion. Tech. Report FFI/Notat-2006/01558. Kjeller, Norway (2006).
9. V Myers and Ø Midtgaard: Fusion of contacts in synthetic aperture sonar imagery using performance estimates. Int Conf. on Detection and Classification of Underwater Targets. Proc. of the Institute of Acoustics, Vol. 29, Pt 6. (2007).