Automatic Underwater Multiple Objects Detection and Tracking Using Sonar Imaging

Shi Zhao

Submitted in Fulfilment of the requirements for the degree of
Master of Engineering Science

School of Mechanical Engineering
The University of Adelaide
Adelaide, South Australia

CHAPTER 4

Objects Detection and Tracking in Structured Environments

The information extracted from sonar measurements is crucial for AUV navigation. Detecting underwater objects will not only assist the AUV to avoid oncoming obstacles but also solve localisation and mapping problems. The underwater circumstances often present different measurement issues to sonar. Such problems, like ambient noise and reverberation found in near proximity of sonar, form the background which may cause false alarms to the detector. Simulating the AUV operational environments and testing the sonar sensor in artificially structured habitats will open the door for future applications in complex natural surroundings. In this chapter, a set of algorithms are exploited in order to recognise static and dynamic targets from acoustic images, obtained in a man-made test tank.

4.1 Introduction

The raw data replied from the Super SeaKing is in the form of a bilateral matrix. Returns are aligned vertically according to the bearing. Each row represents the transducer orientation of a return and each column represents range measurement results. The raw data is hard to interpret by human operators unless it is transformed into Cartesian coordinate systems. An acoustic image in x-y plane will represent the real sonar operational environment directly. In this study, a complete sonar scan sector is described as one scan period and the data reply of this scan period is described as one frame of raw data. The transferred image in the Cartesian coordinate systems is described as a frame of acoustic image. Figure 4-1 shows one frame of raw data and a corresponding acoustic image in the Cartesian coordinate system. Generally, for the displaying purpose, sonar

raw data and acoustic images are all displayed in pseudo colour instead of grey colour in this thesis.

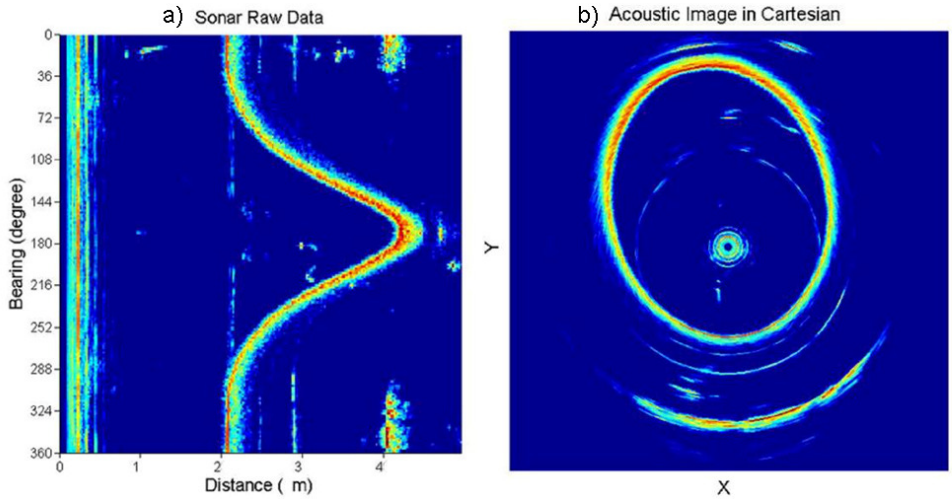


Fig. 4-1: a) Sonar raw data. The size of the raw data is 200×248 pixels, which means the scan step was 1.8° and 200 pings were emitted from the transducer. Furthermore, 248 discrete bins were obtained in each return. b). Same data presented in 2D x-y plane.

An obstacles detection approach includes signal processing and several image processing tasks, which begins with the pre-processing filter and ends with the feature extraction operation. The processing will start once the raw data are acquired in digital form. The procedure can be illustrated by a hierarchical image pyramid, as seen in Figure 4-2. At the very lowest level, raw data are considered as discrete digital signals. Each single return is pre-processed to filter out the sensor self noise and reverberation energy. The next level up is the acoustic image. Raw data is reproduced in Cartesian coordinate systems. The discrete signals are then converted into individual pixels. Consequently, some image processing techniques are introduced to partition the image into several sub regions, which contain a set of similar pixels. These regions are the potential targets. Threshold analysis is then performed to detect the real targets over the background. Edges of the real targets will be extracted after the detection. A mathematic model of the operational environment is finally generated and the sonar sensor can be finally localized according to the mapping results. A high level representation of the

image will result in more reduction in the amount of data. All the operations and image representation in Figure 4-2 will be explored later in this chapter.

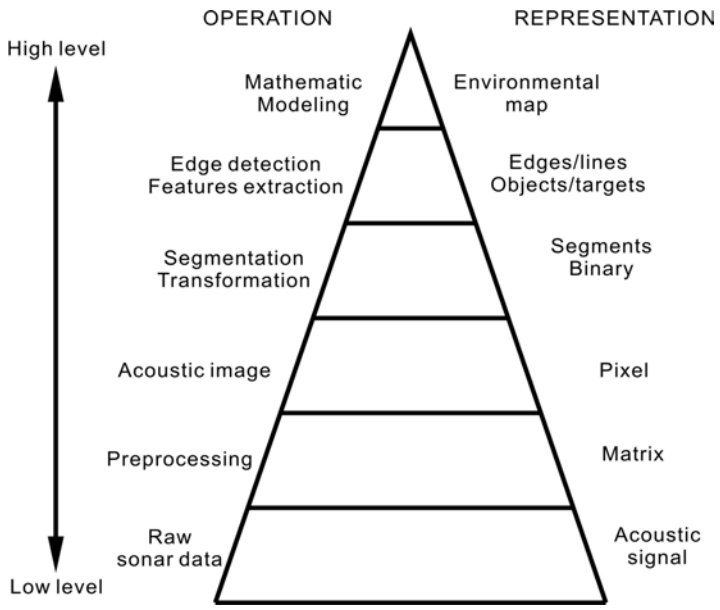


Fig. 4-2: The hierarchical image pyramid.

4.2 Pre-processing the sonar raw data

In structured shallow water conditions, ambient noise is no longer the main component of the background and can be negated. On the contrary, reverberation energy displayed on the acoustic image will result in false alarms and bias to the tacking results. False alarms often occur when small targets are observed by the sonar sensor because their features are similar to the reverberation. Since discriminating reverberation from real targets is relatively hard in an acoustic image, a filter is designed in this research to filter out the reverberation energy in pre-processing procedures. The pre-processing mainly focuses on eliminating the sensor self noise and reverberation energy which may interfere with the detection of real obstacles and adding artificial signatures for further sonar localisation.

4.2.1 Sonar self noise filter

Besides the noise from electronic and mechanical systems inside sonar sensor, another principal factor that causes self-noise is the emission of the sound waves from the projector. When projecting the waveform for each return, the hydrophone will capture the sound waves first rather than the echoes from targets. In a single return, a high amplitude segment can be observed in the first several Bins (see Figure 4-3). When reproducing the acoustic image in a Cartesian coordinate system, these false phantom returns will orbit the sensor in a circular motion (see Figure 4-4). From the Super SeaKing hardware specification, the minimal range resolution is 0.5 meters. Therefore, all the segments within this range are considered to be noise and will be eliminated. Their intensity values will be replaced by zero, which represents no echo return.

4.2.2 Reverberation suppression filter

When sound is transmitted underwater, scattering and reflection occur until the vibration energy is absorbed by the water. Reverberation represents the energy which is reflected back to the sound source. Likewise, target echoes represent a special case of reverberation [30]. Unwanted reverberation is most responsible for creating the interfering background in structured water environments, because of the absence of the ambient noise.

4.2.2.1 Reverberation in shallow water environments

There are three classes of reverberation to be considered in sonar performance: volume reverberation, surface reverberation and bottom reverberation. In most artificial structured environments, where the water depth is below 30 meters, surface reverberation and bottom reverberation are dominant and volume reverberation is seldom dominant [30]. Surface reverberation and bottom reverberation may be analytically considered

together as boundary reverberation, since a two-dimensional distribution of scatterers is involved [33].

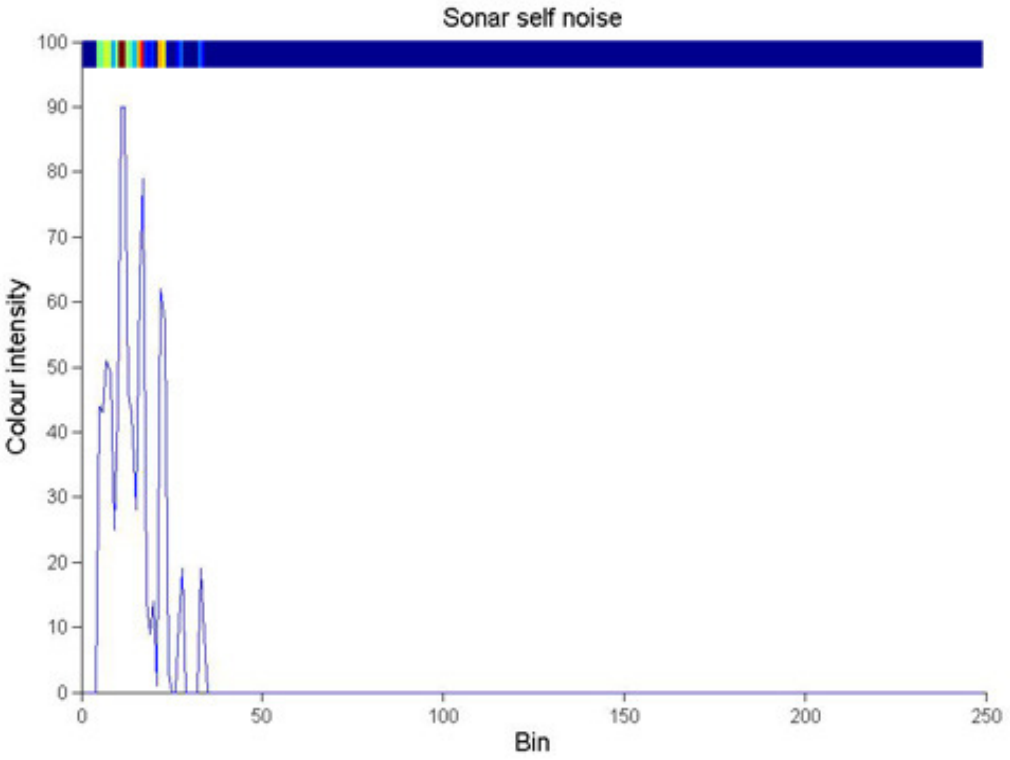


Fig. **4-3**: Sonar self noise in a single return. The beam on the top illustrates the colour intensity of the return .

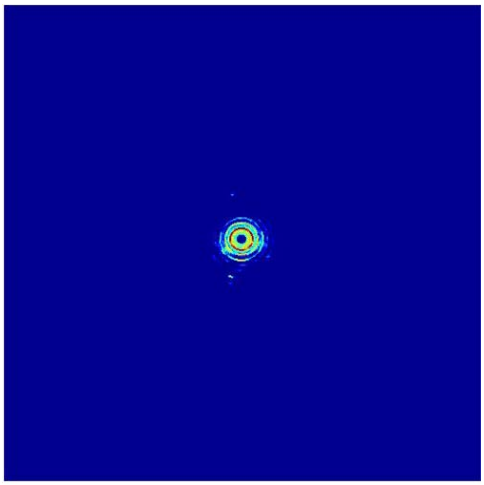


Fig. **4-4**: Sonar system noise displayed in an acoustic image.

Assuming all the sound waves in a structured environment are originally emitted from the sonar sensor, the actual beam produced by the sonar will finally determine the reverberation strength and position in an acoustic image. Figure 4-5 shows an ideal two-dimensional beam pattern which is generated by Super SeaKing. The sound wave is incidental at grazing angle α/2 upon the boundary between the water surface and the air, where α is the vertical beamwidth of the sonar beam. If the sonar depth is given, the distance between the sonar centre and the surface reverberation can be calculated. Since the acoustic path from the sonar is relatively straight, a right triangle is formed. The calculation used to determine the range of the surface reverberation is as follows: the range (R) of the surface reverberation is equal to the product of the depth of the sensor (H) and cotangentα/2 (referred to Equation 4.1).

$$R = H\cot(\alpha/2) \qquad 4.1$$

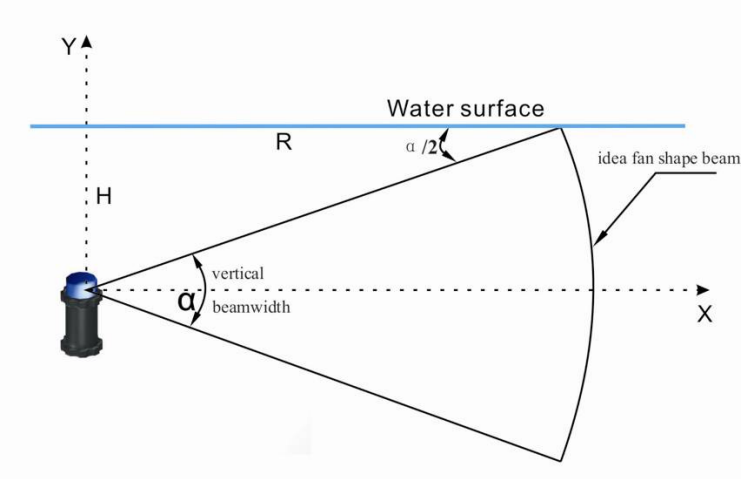


Fig. 4-5: An ideal beam pattern and calculation of surface reverberation.

Nevertheless, all sound emitting devices will transmit side lobes at the side of the main beam. Side lobes represent the undesired propagation directions of the sonar beam. Varieties of techniques in the transducer design and composition try to suppress the side-lobes as much as possible; these lobes however always exist. Figure 4-6 simulates an actual beam pattern with side lobes. In most cases the contribution due to the side lobe of the transducer can be neglected. However, in shallow water environments, reflected

energy from side lobes can be easily picked up by the sonar receiver. Although the precise mathematical pattern of the Super SeaKing beam is not available, Equation **4.1** can be used to roughly distinguish the main lobe reverberation from side lobe reverberation. Given the depth of the sensor and idea vertical beamwidth, the range of reverberation caused by main lobe can be calculated by Equation **4.1**. Since the range of side lobe reverberation will always smaller than the one which is caused by main lobe.

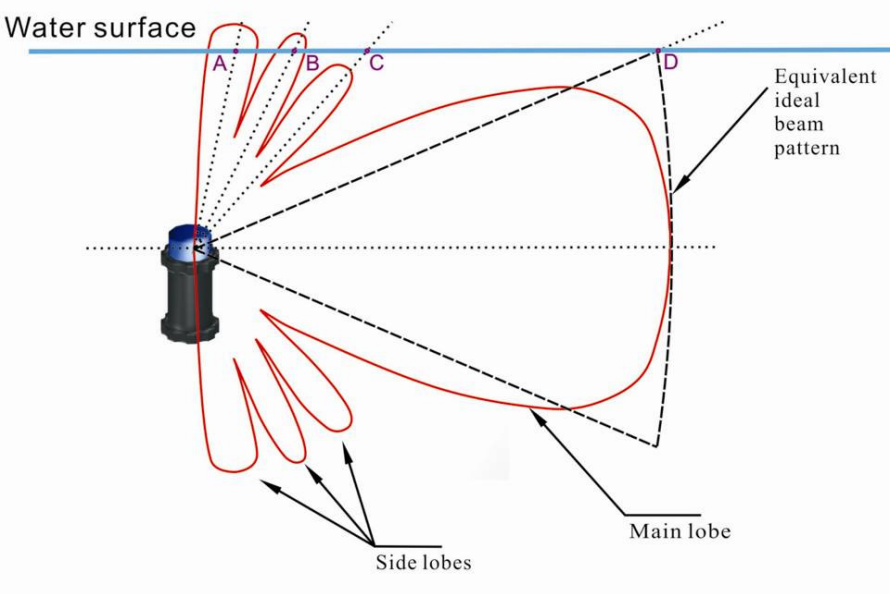


Fig. **4-6**: Side lobes beside the sonar main lobe. Reverberation will also receive by the sonar around point A,B,C and D.

4.2.2.2 Boundary reverberation in acoustic images

In artificial structured environments, since the boundaries of the water surface and bottom are relative smooth, nearly all the sound energy is reflected in the specular direction and sonar cannot capture it directly. Theoretically, the reverberation from a single sound beam is relative lower than a real object. However, it is still very important to filter out reverberation from sonar images properly for the following two reasons:

- Sonar scanning strategy

Super SeaKing rotates the sound beam with a small steering angle increment. An acoustic image is generated by stitching small scanned sections together. When the sonar is static inside the water, boundary reverberation will occur at the same range but different azimuths. Corresponding to the acoustic image, small reverberation will connect together and display as a regular geometric arc or a ring surrounded the sonar centre (depending on the sonar scan sector). Figure 4-7 shows an example image generated in an elliptic water tank when sensor was lowered two meters deep underwater water. In this image, we can only observe the boundary reverberation from the side-lobe because the main-lobe strikes the walls first (based on the calculated result from Equation 4.1).

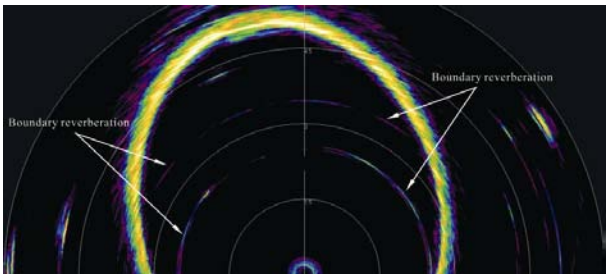


Fig. 4-7: Boundary reverberation caused by side-lobe in sonar image.

- Small targets

Target objects introduced in this research are in centimetre-scales. Their sizes are much smaller than the normal sonar targets, such as underwater mine or a school of fish. Most targets can only be seen by a single scan. Further, the insonfication of the target is small and sound back toward the sonar is a process of scattering instead of reflection. The real geometric shape of the tangent surface cannot be acquired accurately by the sonar. Targets, no matter spheres, cylinder or cube, in acoustic images may be displayed similar to rectangles. In addition, the appearance (shape and colour intensity) of the object will vary from scan to scan. Therefore, comparing with the boundary reverberation, targets are more like random noise instead of real objects underwater. Figure 4-8 shows a sonar image with three ball shape objects, which was generated in the same water tank as Figure 4-7.

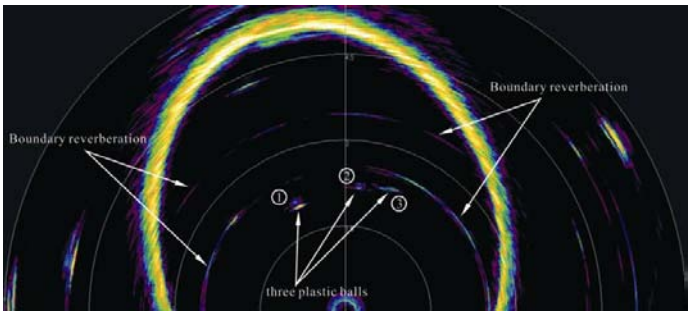


Fig. 4-8: Sonar image with three rigid spheres and boundary reverberation. Sphere 1 and sphere 2 are roughly in square shape in the image, while sphere 3 is just a small curve

4.2.2.3 Features of reverberation and target echo

To successfully filter out the boundary reverberation, it is necessary to seek the distinct characteristics between the real targets and reverberation. Therefore, a rigid sphere (18mm in diameter) was located at the position where the boundary reverberation occurred (see Figure 4-9).

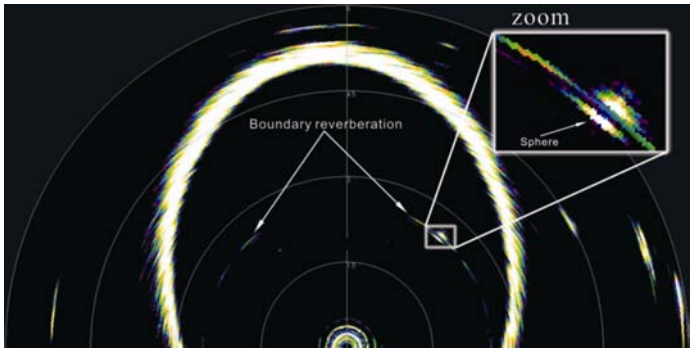


Fig. 4-9: Image when object echoes and reverberation are observed at the same range.

Since the sonar image is formed by returned sound beams, to compare the difference, the best way is to return to the original ping. In a single return, the distance between each bin can be obtained according to the amount of the bins and sonar scan range, which is equal to the scan range divided by the number of the bins. Consequently,

given the range of the reverberation (or the sphere), the corresponding Bin can be calculated by Equation **4.2**, where *Bin* is the Bin index number, *R* is the range of reverberation in meters (or the sphere) *S* is the sonar scan range in meters, and *N* is the amount of bins in single beam.

$$Bin = RN/S \qquad 4.2$$

The rest of section compares two aspects: signal strength (colour intensity) and pulse length (energy span). The signal strength refers to the returned power from an object which is detected by the hydrophone. The pulse length represents the signal duration which is reflected by the specific ensonified surface of an object. Figure **4-10** illustrates a sequence of sonar returns which have scan step size of 3.6 degree. Assuming the first return starts at bearing 180° and after seven steps, the bearing of the eighth return is 205.2°. Reverberation energy can be observed at three adjacent returns (205.2°, 208.8° and 212.4°). The acoustic signals then encounter the sphere at 216.0 degrees. Figure **4-11** plots the above seven returns by colour intensity versus the distance. In this case, sonar scan range is 6 meters. Measured by the SeaNet Pro, the reverberation occurred at about 2.5 meters; 489 bins were recorded in each single return. Based on Equation **4.2**, reverberation (or the sphere) will be sampled around Bin 203.

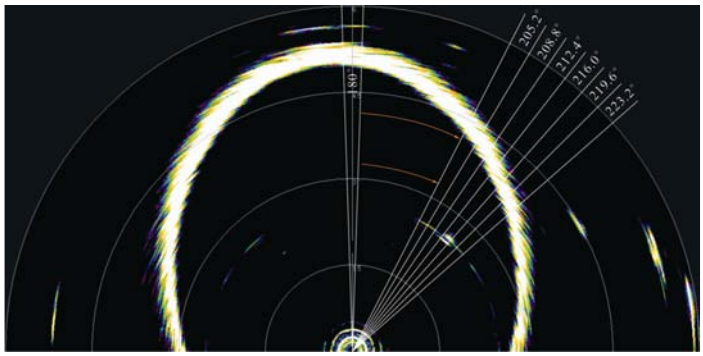


Fig. **4-10**: A sequence of scanned pings in a sonar image.

A closer inspection of boundary reverberation and target echoes is shown in Figure **4-11**, where the X axis represents the distance and Y axis represents the colour

intensity. The echoes around Bin 200 are zoomed inside the black rectangle for comparison purpose.

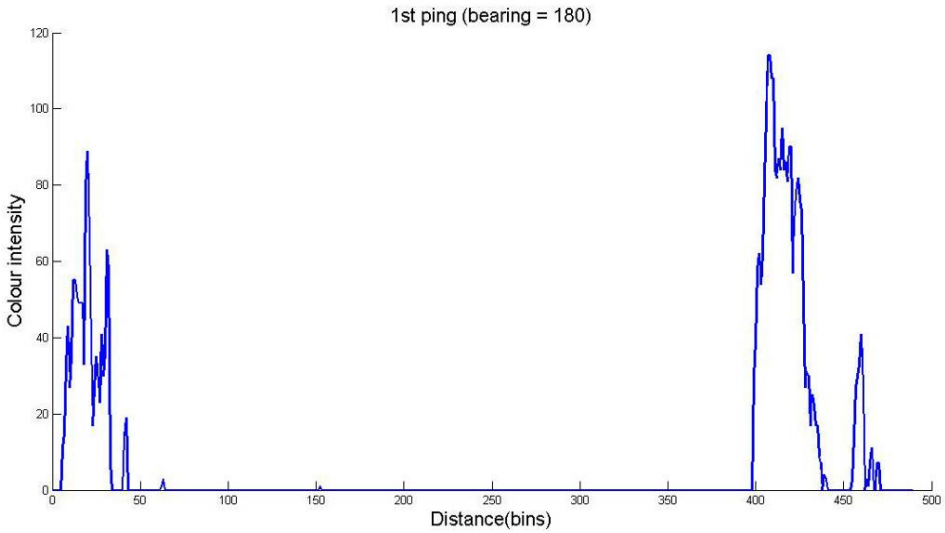


Fig. **4-11 a**: 1st Return. In the first return, sensor self noise can be observed from Bin 1 to Bin 58. The echo from the tank boundaries roughly starts with Bin 400 and till to Bin 440. No reverberation energy is detected. If we partition the return into several sub-returns by the pulse length, there are seven sub-returns in the history of 1st return.

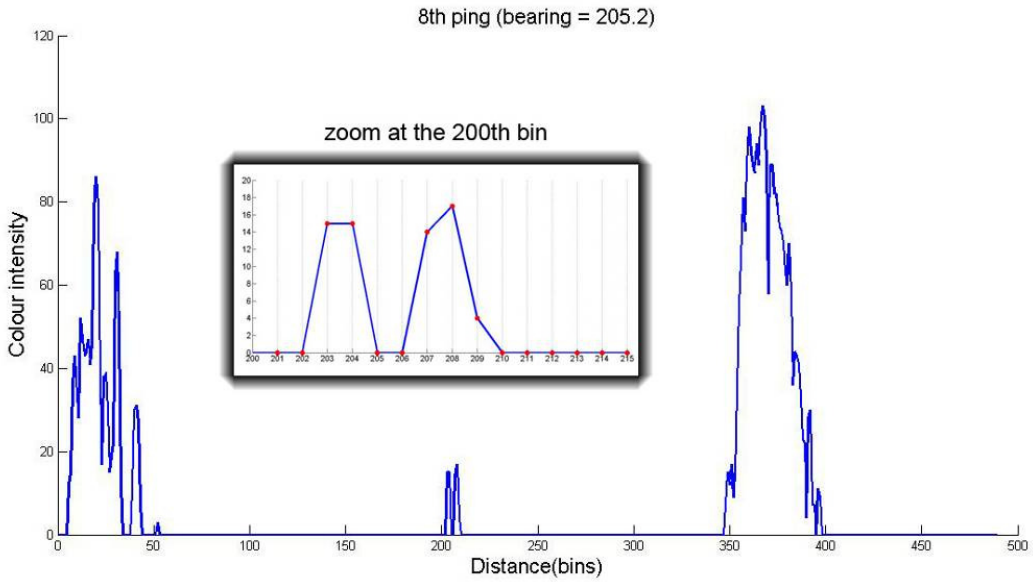


Fig. **4-11 b**: 8th Return where the transducer bearing is 205.2°. Reverberation energy is first received at Bin 203. From the zoomed image, we can see clearly that reverberation was separated into two sub-returns. The pulse length (or distance) of those two sub-returns is two bins and three bins separately.

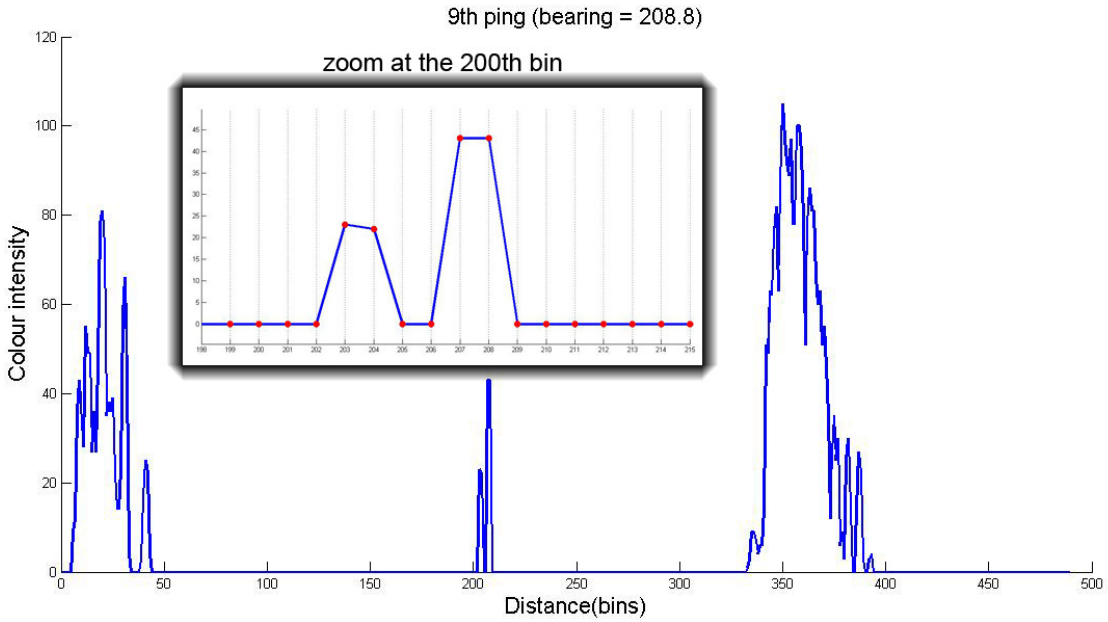


Fig. **4-11 c**: 9th Return where the transducer bearing is 208.6°. Similar reverberation energy is received from Bin 203. The pulse length of those two sub-returns is all two bins.

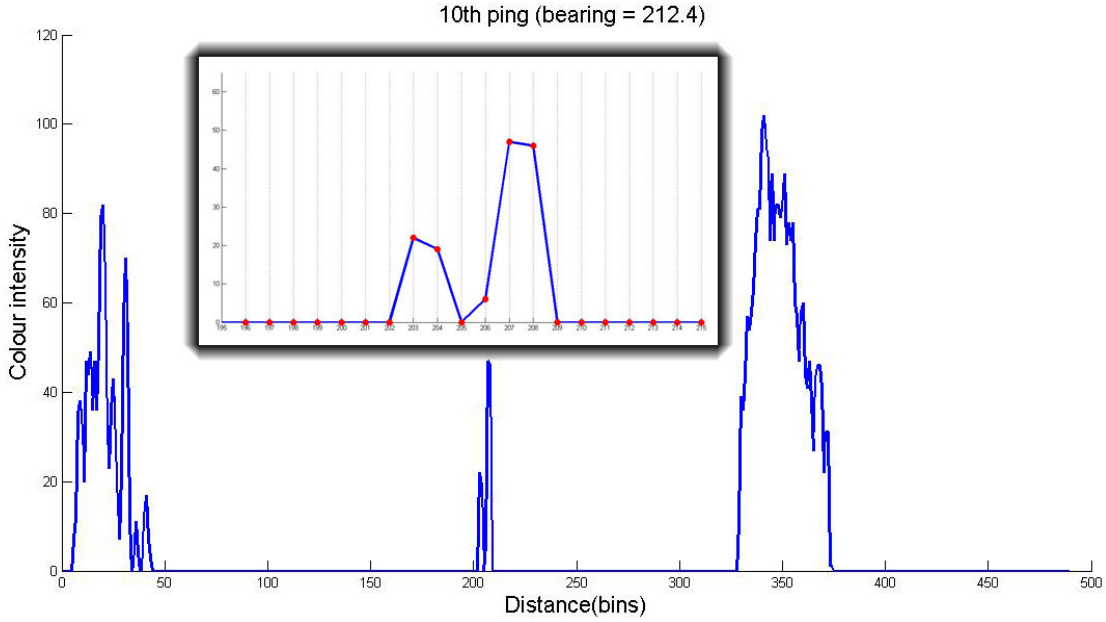


Fig. **4-11 d**: 10th Return where the transducer bearing is 212.4°. Reverberation energy is first received at Bin 203. The pulse length of those two sub-returns is two bins and three bins separately.

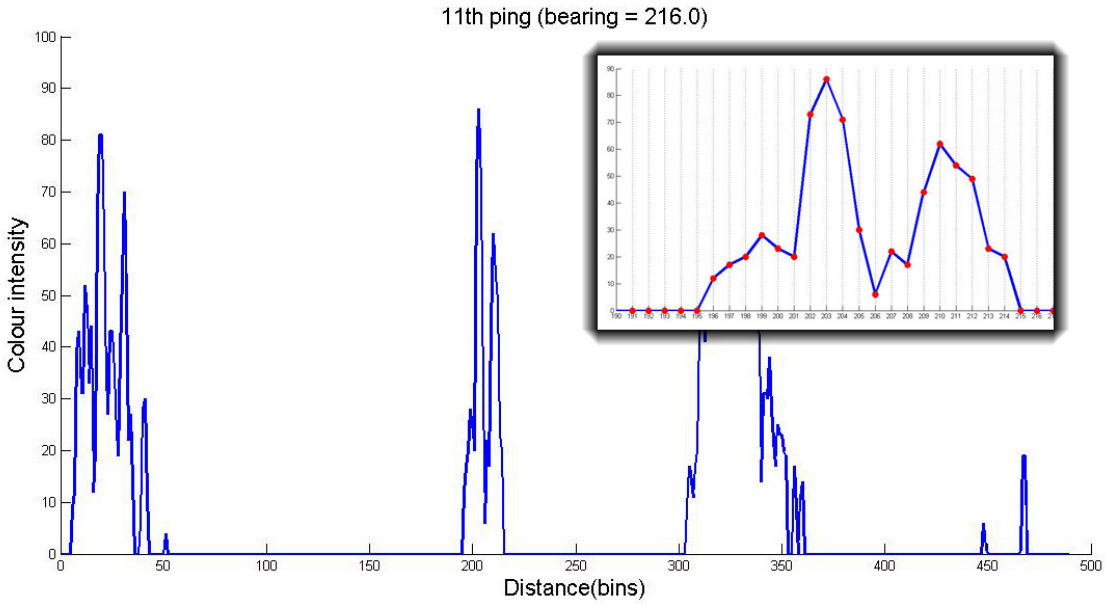


Fig. **4-11 e**: 11th Return where the transducer bearing is 216°. Echo reflected from the target is firstly received at Bin 196 and the energy span of the target is 18 bins. The maximum intensity value at Bin 203 is observably higher than the intensity value of reverberation. However, the colour intensity values of other bins are closed to the reverberation.

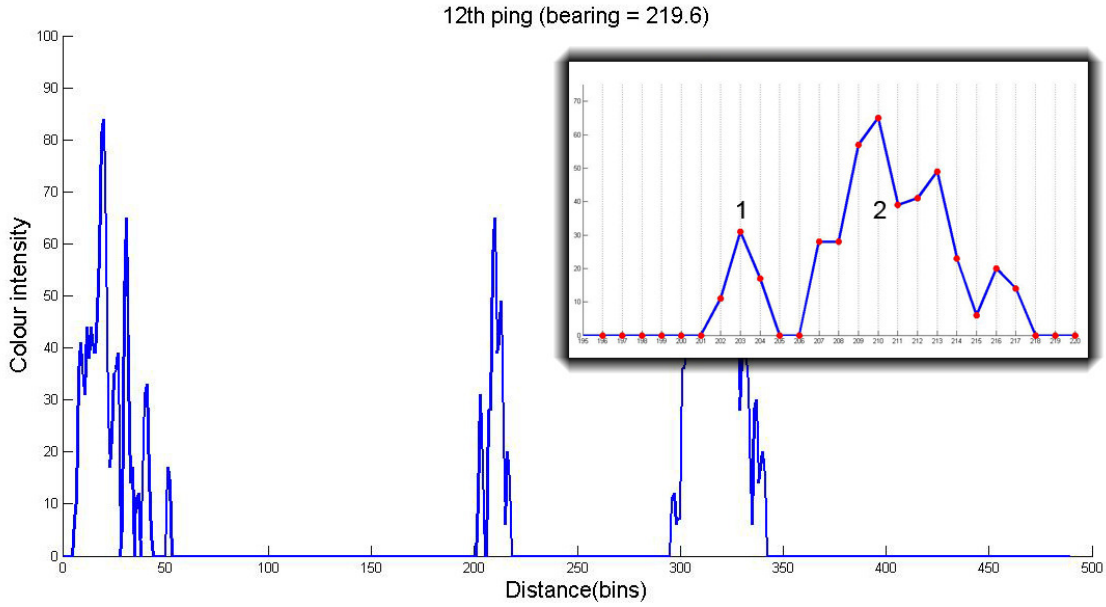


Fig. **4-11 f**: 12th Return where the transducer bearing is 219.6°. Both the boundary reverberation and target echo can be observed in this return. The energy span for the reverberation is three bins. On the contrary, the energy span for the target is eleven bins.

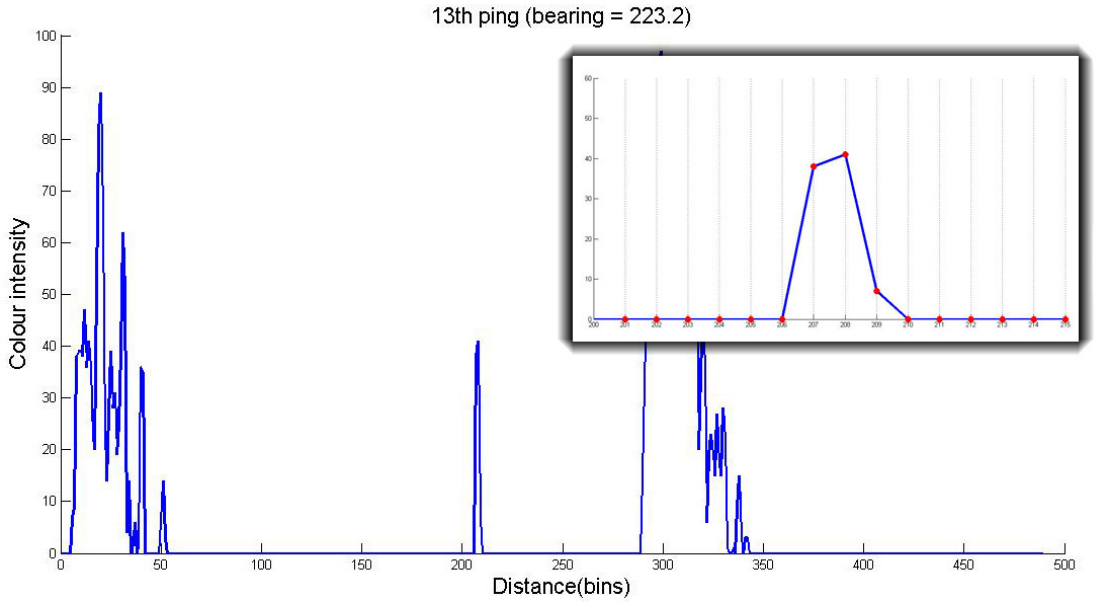


Fig. **4-11 g**: 13th Return where the transducer bearing is 223.2°. With the advance of the scanning procedure, reverberation is received at Bin 207 again.

Figure **4-8** till to Figure **4-11** demonstrate that the colour intensity of the most reverberation segments is less than that of the object. However, the outside contour of the sphere has similar colour intensity as the reverberation energy (see Figure **4-11 f**, zoomed area). In this example, the maximum reverberation energy was observed at 10^{th} return with colour intensity valued of 50. If using an intensity threshold to filter out the reverberation, a part of the echo from the target will be eliminated at the same time. Such intensity thresholding operation excludes useful information and may result in missing detections. Alternatively, in Figure **4-11**, the reverberation energy detected by the sensor only spreads less than four bins but the echo from the sphere extends over ten bins. The abovementioned examples in this section were acquired when sonar scan range was 6 meters. A single return was partitioned into 485 bins, which means that each bin represented about 0.012 meters. That is to say, the sound wave may travel at least 0.132 meters (11 bins) to across the sphere. (The number of bins may change according to different sonar scan step. It is depending on the micro-controller inside the sonar sensor. However, whether it is automatically adjusted by the micro-controller or it is a default value by the sonar system is unknown, because no information is given form the Tri-tech.)

The reason for this phenomenon may because of the sound reflected strategy. When sound waves strike a smooth surface, for instance still water, specular reflections occur. Most energy reflects at same angle with respect to the incident angle. Some of sound waves may also transmit through the water surface. The returned energy detected by the sensor hydrophone is relatively small. Meanwhile, the fan shape beam results in a small size of reflected boundary at the water surface with respect to the cross-section of the sphere. Because the pulse length represents the distance that sound wave travelled, the distance threshold can successfully filter out the boundary reverberation energy and retain the echo from the target. Therefore, the original design of reverberation suppression filter is to partition each single return into several sub-returns by pulse length and then calculate the energy span (distance) of each sub-return. If the energy span is less than a certain value, this sub-return will be classified as boundary reverberation and their colour intensity values will be assigned to zero.

4.2.2.4 Multi-range approach for Reverberation Suppression Filter

Normally, the propagation of acoustic signals underwater is much complex than the above examples. Multiple reflections, Doppler shift and boundary roughness will affect the acoustic imaging and another factor needed to be considered is the reverberation caused from side-lobes. The energy span of reverberation and targets may be very close to each other. A simple distance threshold may work perfectly in one environment, but provide poor results in a new operational environment. Furthermore, the aim of target detection is to provide inputs for the AUV to avoid obstacles. The region close to the AUV is more important than others, which are further away for reasons concerning avoidance procedure. All the potential objects which may cause damage within this region need to be detected. Also, to maintain a proper safe distance between obstacles and the AUV, information should be obtained sufficiently and rapidly. According to the above two considerations, a multi-range approach is developed to meet the requirements. In this multi-range approach, a single return is partitioned into a number of sections. The criterion used to distinguish the reverberation and target is

different within different sections. The similar approach will be applied to evaluating the real targets from background noise in the later session as well. The principle of the region partition is determined by two parameters:

- AUV cruising speed versus the sonar scan period

The cruising speed of the AUV is designed to 0.5m/s. A complete 360 degree scan will take about 4 seconds (scan range less than 50 meters). Therefore, a two meters section is a vital range for AUV safety. Any obstacles within this range need to be detected with high priority.

- Sonar scan range

Sonar scan range is another factor that needs to be considered not only because it will affect the scan period but also because the resolution of an acoustic image will decrease with the scan range increase. The same target will be displayed in different sizes depending on the sonar scan range, since each bin will represent a different distance. For this project, the maximum sonar scan range is 20 meters.

In practical performance, a single return is divided into three sections. The first section will start from zero till 2 meters. The rest section will divide according to the proportion of 1:2. For instance, if the sonar scan range is 20 meters, the first section is 0-2 meters and the second one is 2-8 meters and the last one is 8-20 meters.

4.2.2.5 Reverberation Suppression Filter design

- First section

Because reverberation within this region is caused by side-lobes, sub-returns which have energy span less than 3 bins were considered as reverberation and eliminated straightaway. The rest sub-returns will remain unaltered for the further process.

- Second section

Sub-returns which have energy span amounting to less than 5 bins were eliminated immediately. Sub-returns which have energy spanning more than 8 bins

remained unaltered. Sub-return between 5 to 8 bins will be processed by an average filter (denoted as Equation **4.3**). This average filter can suppress the reverberation to some extent and retain useful information for further detection.

$$Y(k) = \frac{1}{N}\sum_{k=1}^{N} X(k) \qquad \textbf{4.3}$$

where *X(k)* is the input and *Y(k)* is the output. *N* is the number of bins. *k* is the index number of the sub-return.

- Third section

Sub-returns which have an energy span less than 8 bins were eliminated immediately. Sub-returns which have an energy span larger than 10 bins remained unaltered (potential echo from the object). Returns which have energy spans between 8 and 10 bins will be processed by the same average filter as Equation **4.3**.

Another important issue needs to be addressed is the situation when objects overlap section boundaries. The partition action may breaking up the object (into sections) and may result in fragmentation of genuine target echoes. It seems that such operation could potentially be individually misclassified as reverberations. However, the probability of this event is quite low. Also the misclassification will not lead to the misdetection, because the objects may not scan by a single return. For instance, the 18mm-radius sphere was scanned by the sonar in two continued beams (216 degree and 219.6 degree) and the number of bins is 18 and 11 separately. Even if the sphere overlaps the third section boundaries, at least one sub-return is remained after applying the multiple-range reverberation filter.

Figure **4-12** illustrates the result of a single return after the self noise filter and reverberation suppression filter is applied. The sonar scan range was 6 meters and 249 bins were obtained in each return. Therefore, each bin represented about 0.024 meters. The code of developed reverberation suppression filter is shown in Figure **4-13**.

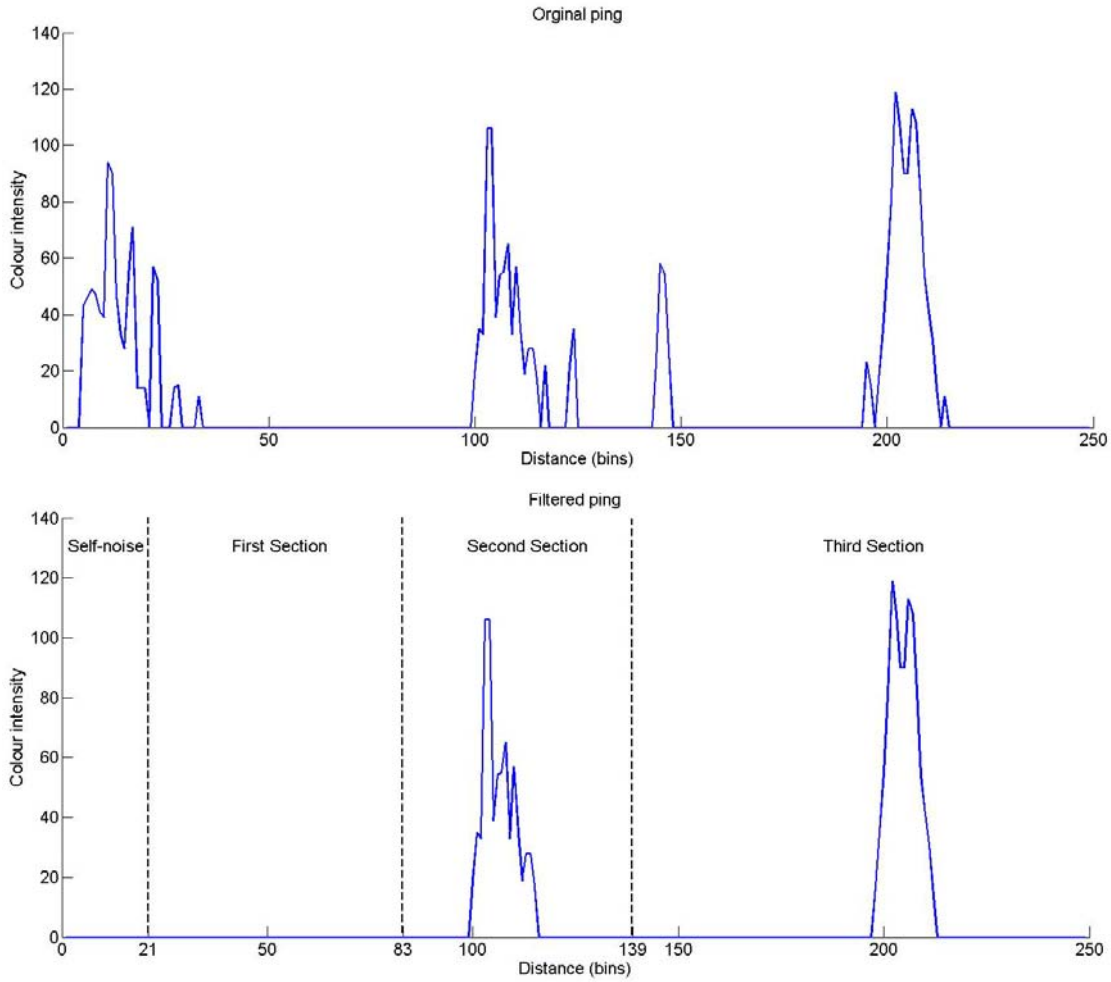


Fig. **4-12**: A single return after filtering out self noise and reverberation.

```
%% multi-range reverberation suppression filter
% pre-processing the sonar single return.
function [Return] = reverberationFilter(sonar_return)
% Initialization:
% read each single return into memory and acquire the parameters
return = read(sonar_ping)
binNum = getNum(return) % the number of bins in single return.
range = getRange(sonar_configure) % sonar scan range.
rangeMark_1 = round(2/range*binNum); % range segmentation
rangeMark_2 = round((range-2)/3/range*binNum);

% acquire sub-returns
subIndex = label(return)   % label the zero in the return;
n = getNum(subIndex) % return the numbers of sub_return;
subReturn = cell(n,1) % construct a class to save sub_return;
for i  = 1 to n
subReturn{i} = segment(return);
subReturn _marker = mark(return) % record the sub_return location in
the return
pluseLength = getlength(subReturn{i});

% Multi-range approach for suppression filter
 switch  subReturn _marker
        case  subReturn _marker <= rangeMark_1
            if(pluseLength<3) % automatic selection
                 subReturn = zero;
            end
        case rangeMark_1 <  subReturn _marker <= rangeMark_2
            if(pluseLength<5)
                subReturn{i} = zero;
            else
                if (5<pluseLength<8)
                    subReturn{i}=avragefilter() % average filter
                end
            end
        case rangeMark_2 <  subReturn _marker <= range
            if(pluseLength<8)
                subReturn{i} = zero;
            else
                if (8<pluseLength<10)
                    subReturn{i}=avragefilter() % average filter
                      end
            end
    end
end

% restore the filtered data into return
for j = 1 to n
    Ping = restore(subReturn{j});
end
```

Fig. 4-13: Automated reverberation suppression filter.

4.3 Acoustic image representation in Cartesian

The sensorial raw data are arranged in a polar coordinate system, since each point on the two-dimensional plane is determined by an angle and a distance respectively (see Figure 4-1 a)). Since polar representation is difficult to interpret by a human operator, one frame raw data will transform to Cartesian coordinates after pre-processing stage. The conversion from a polar system to a Cartesian system is not complex, because the conversion trigonometric equation is fairly straightforward. However, the main problem with this conversion is that the acoustic image in a Cartesian system is not completed. For instance, given a particular return, the bearing of the return, the distance measurements along the return and the colour intensity in the return can be obtained. Then, the position of the return in Cartesian coordinates can be calculated. It is a straight line in the x-y space. If we assign the colour intensity into the Cartesian coordinates, we will generate a series of lines intersecting with the straight line. To fill the gap between those lines, a better approach is to assume that each ping is not a simple line of measurement but a sector that occupies the gaps and then assigned the colour intensity accordingly (see Figure 4-14). An example for reproduced acoustic image can be found in Figure 4-1 b).

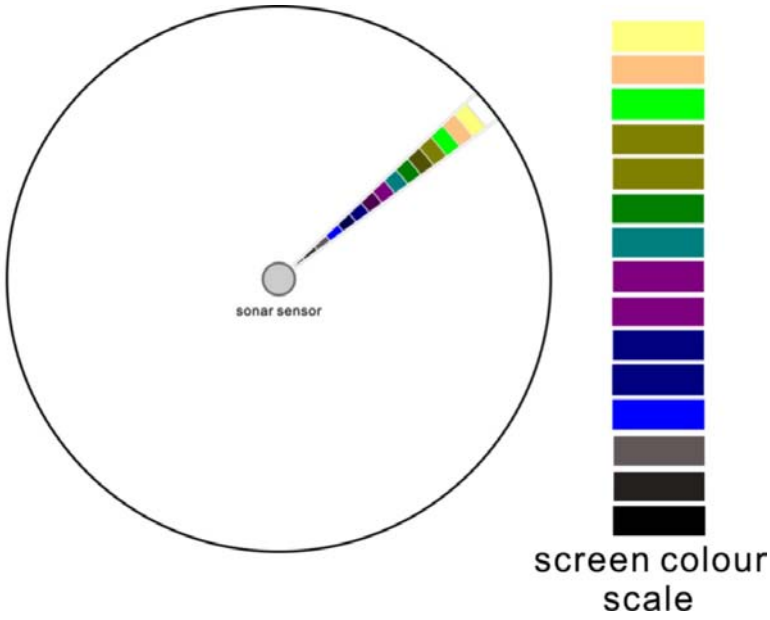


Fig. 4-14: Mapping the raw data to a Cartesian system.

This approach successfully maps the sensorial measurements from a polar system to a Cartesian system. Also, the conversion of each single point only changes the spatial coordinates and remains colour intensity unaltered. The location of the sonar sensor can also be acquired after the conversion. However, it changes the resolution of an acoustic image. From the Figure 4-14, it can be identified that the image resolution will decrease when the scan range increases.

4.4 Binary thresholding

Binary images are the simplest images for processing because there are two possible intensity values for each pixel. Numerically, the two values are usually denoted 0 and 1. The silhouette of an object can be depicted clearly in a binary image. Target objects can be easily distinguished from the background after an acoustic image is quantized to a binary image. Therefore, to isolate the targets from the background, a simple and effective solution is through a binary image. Thresholding technique will convert a grey scale image into a binary image by turning all pixels below a certain threshold to zero and all pixels about that threshold to one. This approach assumes that target and background pixels can be segmented by grey-level values. The thresholding step can critically affect the successful classification of the target from noise background. Hence, the selection of an adequate threshold (grey-level) for extracting targets from the background represents another important issue in this research.

Figure 4-15 illustrates a grey scale acoustic image and its corresponding histogram image. As shown in Figure 4-15, objects in the image commonly have higher intensities than the background. The grey-level histogram of an acoustic image has a deep and sharp valley between the two peaks which represent objects and background. Since there are only two dominant modes in the histogram, one global threshold is enough for partitioning the acoustic image.

Sezgin and Sankur categorized thresholding techniques into six classes: histogram shape-based methods, clustering-based methods, entropy-based methods, object attribute-based methods, the spatial methods and local methods [34]. Among them, clustering-

based method is a processing whereby a data set is replaced by clusters, which are collections of data points that belong together [35]. The Otsu method [36] is another name for the clustering-based global thresholding method. It separates two clusters (black and white pixels) using an exhaustive search method of the threshold that minimizes the weighted within-class variance. The optimal threshold can be selected automatically and stably for a grey scale image with bimodal histogram. It is easy to perform because only the zeroth and the first cumulative moments of the grey level histogram are utilized.

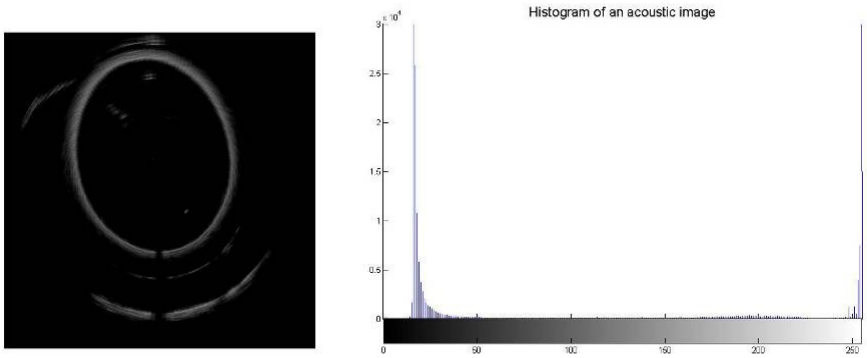


Fig. **4-15**: Grey scale acoustic image and its histogram.

The following deduction of the Otsu threshold is adapted from [35]. Assume the luminance levels of all the pixels in an acoustic image are in the range [1, 2,…L], where the L is the maximum luminance value in the image. The number of pixels at level j is denoted by n_j and the total pixels of the image is N, and $N = n_1 + n_2 + \ldots n_j$. If the normalized grey level histogram is considered as a probability distribution, the probability density function of the grey level histogram is given by Equation **4.4**:

$$P_j = n_j / N \quad \text{where } P_j \geq 0 \text{ and } \sum_{j=1}^{L} P_j = 1 \qquad \textbf{4.4}$$

The foreground object becomes a set of pixels with luminance values above the threshold T, while the background contains the rest pixels which luminance values are less than T. The foreground (objects) and background are expressed as $p_f(j)$, $T+1 \leq j \leq L$ and $p_b(j)$, $0 \leq j \leq T$ respectively. The mean of the acoustic image, background and foreground are given by Equation **4.5**, respectively:

$$\mu = \sum_{j=1}^{L} j \cdot P_j$$

$$\mu_b(T) = \sum_{j=1}^{T} j \cdot P_j \qquad \textbf{4.5}$$

$$\mu_f(T) = \sum_{j=T+1}^{L} j \cdot P_j = \mu - \mu_b(T)$$

Then the probabilities of class occurrence are given by Equation **4.6**:

$$w_b = \sum_{j=1}^{T} P_j = w(T)$$

$$w_f = \sum_{j=T+1}^{L} p_j = 1 - w(T) \qquad \textbf{4.6}$$

where w_b is the probabilities of the background class occurrence which is equal with the probabilities occurrence $w(T)$, and w_f is the probabilities of the foreground (objects) class occurrence

Class mean levels are denoted as Equation **4.7**:

$$\mu_b = \mu_b(T) / w(T)$$

$$\mu_f = \left[\mu - \mu_b(T)\right] / \left[1 - w(T)\right] \qquad \textbf{4.7}$$

where w_b and μ_b are the zeroth and the first-order cumulative moments of the histogram at *Tth* level, respectively.

Hence, the total mean level of the original acoustic image μ can be rewritten to

$$\mu = w_b \mu_b + w_f \mu_f \qquad \textbf{4.8}$$

the class variance is then given by Equation **4.9**:

$$\sigma^2(T) = w_b (\mu - \mu_b)^2 + w_f (\mu - \mu_f)^2 = w_b w_f (\mu_b - \mu_f)^2 \qquad \textbf{4.9}$$

substituting Equation **4.8** and Equation **4.7** into Equation **4.9**:

$$\sigma^2(T) = \frac{\left[\mu w(T) - \mu_b(T)\right]^2}{w(T)\left[1 - w(T)\right]} \qquad \textbf{4.10}$$

the optimal threshold T is equivalently maximizes $\sigma^2(T)$ and is given by Equation **4.11**:

$$T = \arg \max_{1 \le T \le L} \left[\sigma^2 (T) \right] \qquad \textbf{4.11}$$

Figure **4-16** illustrates the binary image after Otsu threshold which is referred to the previously displayed grey level image in Figure **4-15**.

Fig. **4-16**: Binary image after Otsu thresholding.

4.5 Mathematical morphology detector

To detect the presence of targets underwater and avoid oncoming obstacles, it is necessary to extract them from binary images. After the pre-processing and binary thresholding, multiple sound reflections outside the test tank appear to be the major issues which impede the detection. If the structured environment can be identified first, those multiple reflections can be eliminated by logical operations. In addition, due to the Doppler shift, dynamic objects appeared in the acoustic image are frequently become small fragmentations. On the one hand, if those fragmentations are labelled as separate targets, it will increase the computational burden and produce erroneous tracking results. On the other hand, discarding those fragmentations will result in the loss of useful information for navigation. Mathematical morphological approaches aim to maintain the

stability of geometric characteristics in an image and analyse the geometric structures which are inherent from the image. It is a powerful tool in extracting and describing image components. Therefore, a target detector is developed based on morphological technique.

4.5.1 Preliminaries of set theory in digital image

The rationale mathematical morphology is to describe an image component as a set. The standard set notations used to describe image operations are listed in Table 4-1.

Tab. 4-1: Basic set operation in mathematical morphology

Notation	Description
Z^2	The Cartesian production of a set of integers Z
A	the image or image components
A^C	the compliment of the image (inverse)
$A \cup B$	the union of image A and B
$A \cap B$	the intersection of image A and B
$A\text{-}B = A \cap B^C$	the difference between A and B (the pixels in A but not in B)

Apart from the above set operations, mathematical morphology technique extends to another two special set operations for image processing, which are called reflection and translation [37].

The reflection of set A, denoted $\hat{A}$, is defined as

$$\hat{A} = \{-x \mid x \in A\} \tag{4.12}$$

The translation of a set A by point $z = (z_1, z_2)$ denote as A_z, is defined as

$$A_z = \{w + x \mid x \in A\} \tag{4.13}$$

where w is the displacement vector of point z

The reflection for a two dimensional set A is to rotate every point by 180 degree with respect to the origin coordinate space. The translation of A by a vector w results by

translating every point of A along the vector w. Figure 4-17 shows an example of a translation and a reflection operation. Assume that A is an image component, which contains seven pixels and w is a vector of (3, 3). The reflection will rotate A over 180 degree by its origin (2, 5). The translation will shift every pixel by (3, 3).

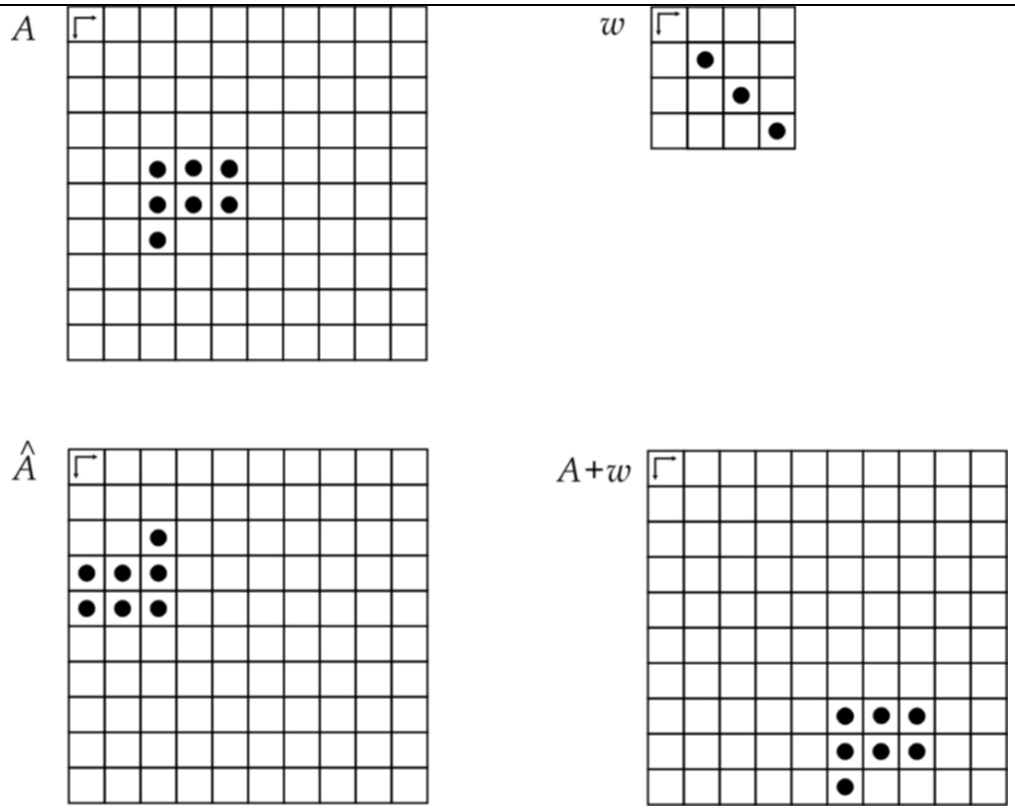


Fig. 4-17: Translation and reflection operation in an image.

4.5.2 Structuring elements

In mathematical morphology, a structuring element is a set used to probe or interact with a given image. The shape and size of a structuring element control the specific operation manner and result. Selecting the structure elements depends on the shapes of image components. Since most target objects in this study are of simple geometric shapes and forms, the structuring elements used in this study are shown in the

following images (Their sizes may change during the performance to differentiate image objects or features).

1. Disk-shaped structuring element (see Figure **4-18**).

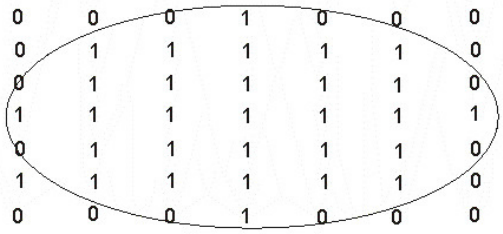


Fig. **4-18**: A disk element (Radius = 3)， where R specifies the distance from the centre of the structuring element to the boundary.

2. Octagonal structuring element (see Figure **4-19**)

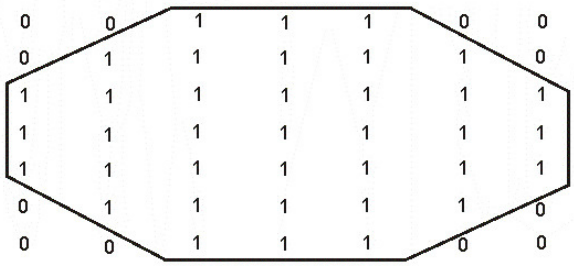


Fig. **4-19**: An octagonal element (Radius = 3)

3. Square structuring element (see Figure **4-20**)

1	1	1
1	1	1
1	1	1
1	1	1

Fig. **4-20**: A square structuring element with width (*w*) is 3 pixels

4.5.3 Binary morphology operation

In a binary image, there are only two Boolean values for each pixel: '0' or '1'. In this study, the foreground set (pixels), which comprise all objects in an image is referred to as white colour ('1'). The rest of the complement sets are called the background, which are in black colour ('0'). In this study, four important morphological operations are employed: dilation, erosion, open and close. Based on these four operations, the practical uses of morphology in this study include: region grow and boundary extraction. The rest of this section details those morphological operations.

4.5.3.1 Dilation and erosion operation

Typically, in mathematical morphology, A always represents the object of interest (image component) in an image and B is a structuring set, called a structuring element. Dilation and erosion are fundamental to morphological image processing. The dilation and erosion are dual with each other when the negation of a formulation employing the first operator is equal to that formulation employing the second operator on the negated variables [38]. The dilation of a binary image A by B, denoted by $A \oplus B$, is defined as:

$$A \oplus B = \{c \mid c = a + b, \ \textit{for some } a \in A \textit{ and } b \in B\} = \bigcup_{b \in B} A_b \qquad \textbf{4.14}$$

The dilation operation will combine all the points in the two sets by translation A with respect to each element of B. Figure 4-21 illustrates a dilation operation. After dilation operation, objects in the binary image are expanded and small holes are filled. Broken segments are therefore joined together and the contour of the object is smoothed. The erosion of a binary image A by B, denoted by $A \Theta B$, is defined as:

$$A \Theta B = \{x \mid x - b \in A \textit{ for every } b \in B\} = \bigcap_{b \in B} (A)_{-b} \qquad \textbf{4.15}$$

The erosion operation will combine two sets using vector subtraction of elements in set B. Dilation can be represented as a union translated by structure element B, whereas

erosion can be represented as an intersection of the negative translated by structure element *B*. The erosion will shrink the object as well as remove small objects and disconnect objects gradually in the image (see Figure 4-22).

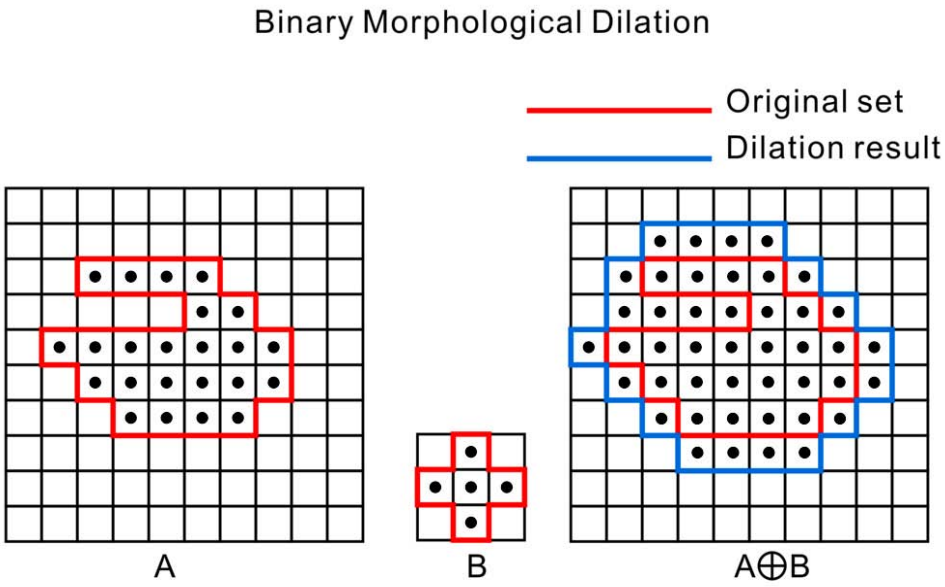


Fig. 4-21: Dilation operation

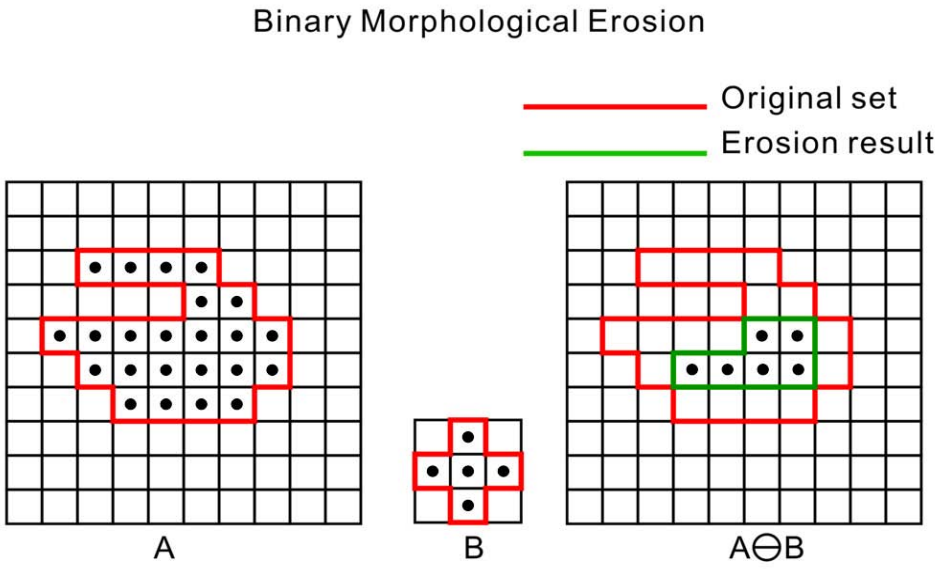


Fig. 4-22: Erosion operation

4.5.3.2 Open and close operation

Dilation and erosion are usually present in pairs. Erosion followed by dilation is an opening operation and dilation followed by erosion is a closing operation. The opening operation can be geometrically considered as 'rolling' the structuring element along the inner boundary of an image A. It therefore will smooth the image. In comparison, a closing operation is 'rolling' B on the outer boundary of an image A. It can fill the holes and remove the gap in an image. The definitions of morphological open and close are given by Equation **4.16** and Equation **4.17**, respectively:

Opening operation, denoted by $A \circ B$, is defined as

$$A \circ B = (A \Theta B) \oplus B \qquad \textbf{4.16}$$

Closing operation, denoted by $A \bullet B$, is defined as

$$A \bullet B = (A \oplus B) \Theta B \qquad \textbf{4.17}$$

4.5.3.3 Region growing (Hole filling)

In a binary image, a hole is a background region which is surrounded by a connected border of the foreground object. The region grow algorithm for filling the hole is based on dilation, set complementation, and intersection. Let A denote a set with the connected boundary and enclosing a hole. B is the symmetric structuring element. Giving a point x_0 in the hole, the Boolean value of this point is assigned to 1 and is located as the initial point. A set X_0 (the same as the A) is constructed with only one element (Boolean value: 1) at point x_0. The filling procedure can be described as Equation **4.18** [39]:

$$X_k = (X_{k-1} \oplus B) \bigcap A^c \qquad k = 1, 2, 3, \cdots \qquad \textbf{4.18}$$

where k is the number of the background points. The procedure terminates at iteration step k if $X_k = X_{k-1}$.

The region grow result R is given by union the X_k and A.

$$R = X_k \bigcup A \qquad 4.19$$

The region grow procedure is shown in Figure 4-23 below.

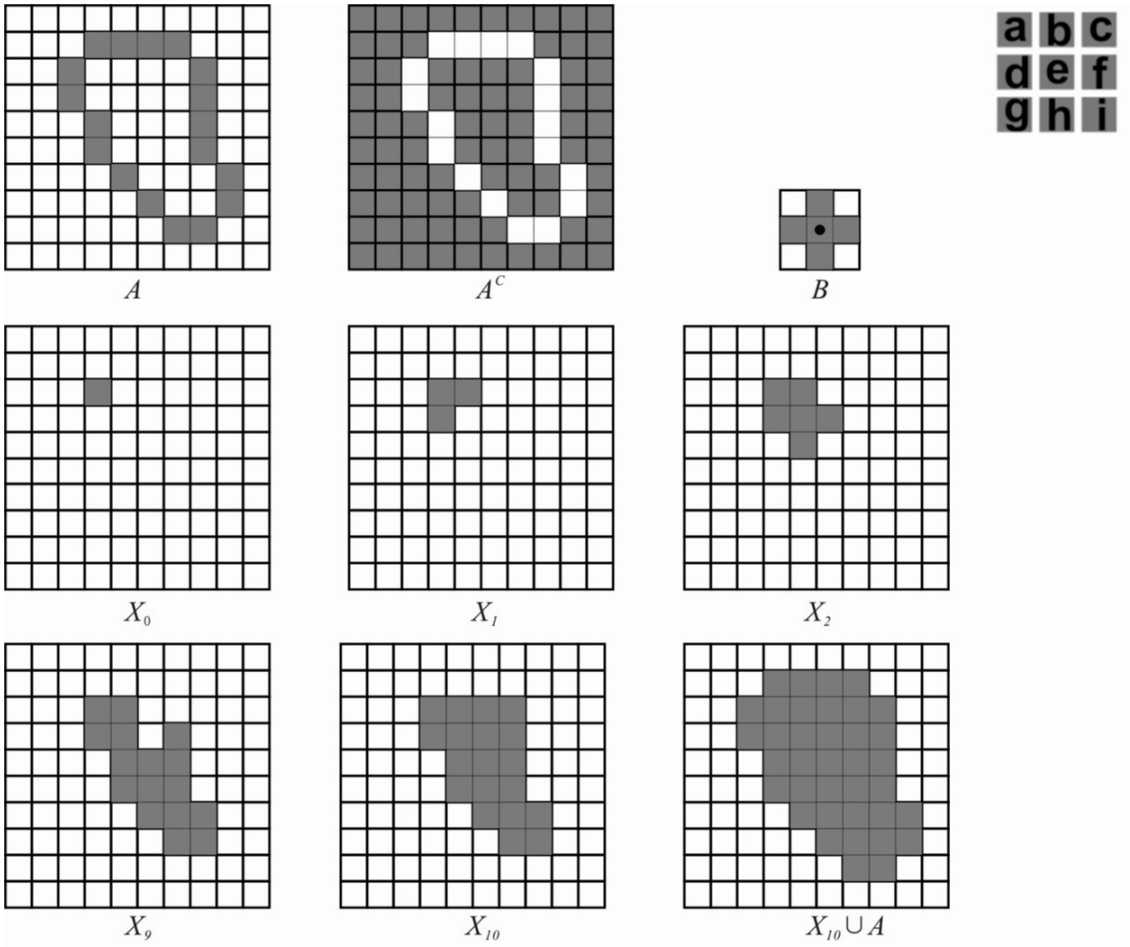


Fig. 4-23: Region growing to fill the hole in the binary image. a) Set *A*. b) Complement of *A*. c) structuring element *B*. d) Initial point in set *A*. e)-h) Different steps for growing. i) Final result. The matrix on the top right corner indicates the number of the image.

4.5.3.4 Boundary extraction

If the boundary of a set *A* is denoted as *Ψ(A)*, it can be obtained by Equation 4.20:

$$\psi(A) = A - (A \Theta B) \qquad 4.20$$

which means the A is firstly eroded by B and then performs the set difference between A and its erosion result. Figure 4-24 show the mechanics of the boundary extraction.

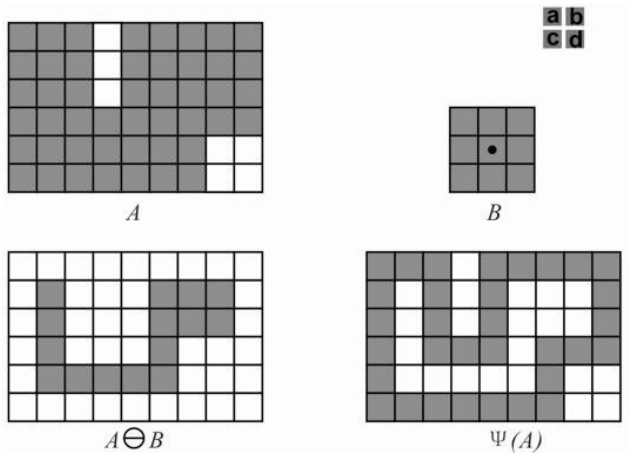


Fig. 4-24: Boundary extraction procedure. a) Set A b) square structuring element c) erosion operation d) boundary of A.

4.5.4 Environment detector for acoustic image

When the sonar raw data is available the detector will perform in the acoustic image which was generated after the pre-processing stage. The detection of the structured environment assumes that the environment itself is the largest region in the binary image. Foreground objects (regions) are firstly dilated by a 4-STREL-octagonal structuring element (refer to Figure 4-19 while R = 4). This operation will bridge the gaps between regions and smooth the contour of individual regions. Secondly, regions in the binary image are ranked and labelled by the area (the actual number of pixels in the region). Thirdly, except for the maximum region, other regions are classified as background and their intensity values are assigned to zero. Fourthly, the binary image is eroded by a 4-STRE-disk structuring element (refer to Figure 4-18 while R = 4) to shrink the

boundaries of the operational environment. Meanwhile, the inner boundaries of the environment need to be extracted for mapping and localisation. This operation can be addressed using Boolean operations. Followed by the Canny edge detection algorithm [40], a set of points which contains the edge of the operational environment can finally be obtained. The Canny method uses a Gaussian filter with a specified deviation to reduce noise and then applies two thresholds automatically to the gradient: a high threshold for low edge sensitivity and a low threshold for high edge sensitivity. The edge is linked by incorporating the weak pixels which are 8-connected to the strong pixels. The code of developed environment detector is shown in Figure **4-25**.

4.5.5 Targets detector for acoustic images

Similarly, thanks to the Boolean operations, only the objects underwater are left in the binary image. A 6-STREL-disk structuring element (refer to Figure **4-18** while R = 6) firstly dilated the foreground objects and then eroded the dilation result (morphological close). The morphological close operation can fuse the narrow breaks between the small fragments, eliminate small holes inside the regions and fill gaps in the contour. After the close operation, foreground objects are aligned by their area size. As discussed in the previous section, the detector partitions acoustic images into three regions which all cover the full horizontal view of the sonar sensor. For instance, given the scan range is 20 meters. The first region originates at the centre of the sensor with a radius of 2 meters. The second region starts at 2 meters and ends at 8 meters. The third region starts at 2 meters and ends at 20 meters. A multiple-range-threshold is generated for automated target selection. The foreground object, where the distance is less than 2 meters and the area over 15 pixels which is about 0.008 m^2 in physical dimension, is marked as targets. In the same vein, if the distance is over 2 meters, the threshold will be 35 pixels (0.02 m^2 in physical dimension). The detector will generate a new binary image which contains the targets. The boundaries of these targets can be extracted by morphological operation (referred to as Equation **4.20**). The code of developed target detector is shown in Figure **4-26**.

```
%% acquire the operational environment
function[edge,tankBW,tankFill]=environmentDetector(binary_image)
% 'binary_image' is the binary image in the memory
% 'edge' is the points set of the tank edge
% 'tankBW' is the silhouette of the environment
% 'tankFill' is the grown image

oct = STREL(4,'octagon');% a octagon STREL structuring
element
dilBW = dilate(BW, oct);% Dilation operatoin
% calculate the numbers of the region
regionNum = getNum(dilBW);

% calculate the area of the region and store in a cell
array
for i = 1 to regionNum
    regionArea.{i} = area(dilBW)
end
tankNum = max(regionArea.{i});% find the largest region
% find the rest regions label and assign the pixels to
background
for i = 1 to regionNum
    while i != regionNum % leave the maximum region only
        dilBW(find(dilBW==n))=0;
    end
end

tempBW = dilBW;% acquire the environment
disk = STREL(4,'disk');% a disk STREL structuring element
tankBW = erode(tempBW,disk); % erosion to thin the boundary
% silhouette image of the environment
%% Boolean operations for extracting the inner boundaries

tankFill = regoingrow(tankBW);% fill the holes by
foreground value '1'
% acquire the inner boundaries

Boundaries = not(tankBW|(not(tankBW|tankFill)));

edge = cannyedge(Boundaries);% Canny edge detection
```

Fig. 4-25: Enclosed environment detector.

```
%% target detector
function [resultImg,decTarget]=targetDetector(BW,tankBW,tankFill)
% 'BW' is the orginal binary image in the memory
% 'resultImg' Result image with labelled targets
% 'decTarget' a class which save the object information,
including the target numbers, centres, areas and bounding
boxes of the targets

% obtained the environment and objects
tankObjects = BW&tankFill;
% eliminating the wall and leave the objects inside.
objects = tankObjects&(not(tankBW))
oct = STREL(3,'ocatgon');
dilObjects = dilate(objects, oct); % Dilation operation
numObjects = getNum(dilObjects); % calculate the numbers of
the objects

for i = 1, i < numObjects
    % Get area of the dilaition objects
    area(i) = getArea(dilObjects);
   center(i) = getCenter(dilObjects);
    % calculate the distance between the sonar and objects
    distance(i) = getDistance(center(i), sonarX,sonarY);
    if (objectArea(i) < 35)&&(distance > 2 meters) % Multi-
range thresholding
        dilObjects(i)=0;
    else
        if(objectArea(i) < 15)&&(distance < 2 meters)
            dilObjects(i)=0;
        end
    end
    i++;
end

% generate an image whose size same as the binary images
but only contains the targets
targetsImg = getTarget(dilObjects);
numTargets = getNum(targetsImg);
decTarget.num = numTargets;
decTarget = getResult(targetsImg)
% extract the boundaries of the target
resultImg = boundary(targetsImg);
```

Fig. 4-26: Target detector.

4.6 Target tracking using nearest neighbour

Tracking is the process to estimate the state of a target over a time interval of interest from the measurements. The crux of the multi-target tracking is to associate a set of measurements with the same target, which is also called data association problem [41]. However, due to the uncertainty associated with the measurements, such as missing detections, false alarms and interfering targets or other countermeasures, data association become a serious problem. In this study, a nearest neighbour standard filter (NNSF) [40] was adopted to address the data association problem. The selection of nearest neighbour filter is mainly based on the following reasons:

The kinematic constraint of targets and the AUV. Both dynamic obstacles and the AUV introduced in this study are moving at low speed and their movements tend to be more like a linear behaviour. Obstacles will not suddenly appear or disappear during the sonar scanning.

The elaborate detector. Since the uncertain measurements are mainly related to the origin of the detection procedure, a sophisticated detector will simplify the tracking problem. The developed detector in this study considered all the factors which cause false alarms and missing detection. It has been approved to be very reliable and accurate to real sensorial data which were collected in an elliptical test tank (see section 4.8).

Hardware limitation and system requirement. Acoustic signal can propagate a longer range than optical signal but the resolution of an acoustic image is relatively lower than an optical image. The sonar sensor adopted in the research is good at detecting objects instead of classification them. Also, the use of a stereo camera on the AUV will provide more information of the target and the fusion of two sensorial data will lead to a new tracking strategy.

In the NNSF, the validated measurement nearest to the predicted measurement is used for updating the state of the target [40]. The distance was chosen to determine which measurement is closest in the study. Since the target objects were separated from the acoustic image, their spatial information was therefore obtained. Once targets are detected, their contours are extracted and NNSF is performed to track their centroids

across the acoustic image. However, the problem of choosing the nearest neighbour is the incorrect measurement. It means that, with some probability, the NNSF will 'believe' that incorrect measurements are correct. This is because the filter-calculated error covariance matrix does not account for the possibility of processing an incorrect measurement and it lead to loss of track [40]. The result is a target trajectory across the acoustic image (see Figure **4-38**).

4.7 Map building for an elliptical environment

The inner edge of the environment has been obtained after performing the edge detector. It is a set of discrete points in the X-Y plane. Since no other navigation sensors are introduced in this research, a non-linear curve fitting technique is adopted to generate the algebraic map. Using this algebraic map, the accuracy of the sensor observation can be measured and the position of the sonar sensor can be identified finally. Curve fitting is the process of finding a mathematical formula for approximating a set of tabular data values [42]. Given a set of data points: (x_1, y_1), (x_2, y_2),…,(x_n, y_n) where x is independent variable and y is dependent variable, the quantity y_i is dependent on one or more variable $x_1, x_2,\ldots,x_n$

$$y_i = f(x_i) \qquad i = 1 \text{ to } n \tag{4.21}$$

Suppose $u=g(x)$ is a mathematical function which can express the independent variables, say, Equation **4.22**

$$u_i = g(x_i) \tag{4.22}$$

The deviation (error) d from each data point is given by:

$$d_i = y_i - u_i \qquad i = 1 \text{ to } n \tag{4.23}$$

If all the errors are sufficiently small, then the function g is the desirable algebraic assumption for the data set. The two major issues for this problem are [41]:

A mathematical form for the function g.

An appropriate evaluation for the parameters of the function g.

The first issue is relatively simple, since the test tank is a structured environment with an elliptical shape. The evaluation method used in this study is the least square which Gauss discovered in 1795. The least square assumes that the best-fit curve has the minimal sum of the deviations squared from a given set (x_1, y_1), (x_2, y_2),…,(x_n, y_n). The criterion of fitting is given by Equation **4.24**:

$$D = \sum_{i=1}^{N} d_i^2 = \sum_{i=1}^{N} (y_i - u_i)^2 = \text{minimum} \qquad \textbf{4.24}$$

where N is the total number of the data set and D is also called algebraic distance. Ellipses in two-dimensions can be represented algebraically by Equation **4.25**

$$g(x) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \qquad (1) \qquad \textbf{4.25}$$

where *a, b, c, d, e* and *f* are the ellipse coefficients, and with the constraint b^2-$4ac$<0.

Let $\chi = \left[x^2 \; xy \; y^2 \; x \; y \; 1\right]^T$ and $\alpha = \left[a \; b \; c \; d \; e \; f\right]^T$, the general quadratic curve equation (4.25) can be written as an implicit second order polynomial [43]:

$$g(\alpha, \chi) = \alpha^T \chi = ax^2 + bxy + cy^2 + dx + ey + f = 0 \qquad \textbf{4.26}$$

The algebraic distance D is then given by Equation **4.27** [42]:

$$D = \sum_{i=1}^{N} g(\alpha; \chi)^2 \qquad \textbf{4.27}$$

Bookstein recommended the minimum D can be solved by rank-deficient eigenvalue system [44]:

$$A^T A\alpha = \lambda C\alpha \qquad \textbf{4.28}$$

where $A = [x_1, x_2, \ldots, x_n]^T$ is called the designed matrix and C is the quadratic constraint matrix on the coefficients *a, b, c, d, e* and *f*.

Bookstein's constraint is $\|\alpha\|^2 = 1$ while Gander *et al.* [45] impose $\alpha + C = 1$.

For purposes of comparison, these two approaches were implemented in this study. The outputs of the ellipse fitting are: major axis length, minor axis length, orientation and the centre of the ellipse. The comparison result is shown in Table 4-2. This result indicates that there is only a tiny difference between these two ellipse-fitting approaches. In this task, Ganders' approach will perform approximately 2 times faster than Bookstein's. Gander *et al*.'s approach is therefore chosen to fit the data set.

Tab. 4-2: Ellipse coefficients output

Method	Center (x)	Center (y)	Major axis (pixels)	Minor axis (pixels)
Bookstein	243.686	199.540	153.3503	122.0485
Gander	243.686	199.540	153.3451	122.0521

4.8 Experimental results

The experiments presented in this section are executed under laboratory conditions to test the target detection algorithm. The tests took place in the Maritime Operation Division (MOD), Defence Science and Technology Organisation (DSTO), Adelaide, Australia test facility. The final experiments correspond to the real data obtained in an elliptic cross section water tank. The data sets acquired by the sensor are used to test the developed reverberation filter, morphological detector and mapping algorithm.

Four hard plastic buoy balls were used as target objects because their echo strengths were relatively independent of the orientation. The red ones are 14cm in radius and yellow ones are 8.8cm in radius, respectively (see Figure 4-28). Weights secured the buoys to the bottom of the test arena and ropes were fixed to the top of the spheres in order to control the motion. The depth of the targets was set at 3.0m below the water's surface. The sonar was secured by the claw and lowered to a depth of 2.5 metres. The test platform was mounted on a bridge above the test tank.

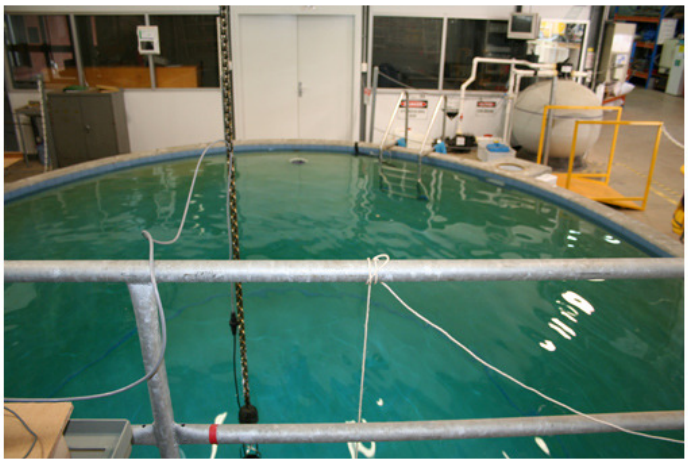


Fig. 4-27: Test tank

4.8.1 Static targets detection

To test the target detection algorithm in various scenarios, experiments were arranged in an orthogonal array according to the Taguchi Methods [46]. The engineering system was divided into input factors, control factors, noise factors and output factors. The output factor or response will affect the input factors, control factors and noise factors. Although the methods are normally used in industrial experiments to study the effect of several control factors, they can also examine the capability of the targets detection algorithm, because the orthogonal array provides a set of well balanced (minimum) experiments which will predict the best results. In this study the sonar target detection system is divided into four factors (see Table 4-3):

Tab. 4-3: Sonar parameters classification

Factors	Description
Input factor	scan range
Control factors	Gain; Adlow; Dynamic range; Signal frequency; Scan step
Noise factors	Boundary reverberation; Sensor self noise; Multiple reflection
Output factor	Detected targets

The orthogonal array designed for the experiment is $L_{16}(4^5)$, meaning there are five control factors and each control factor contains four levels. Sixteen groups of data were collected in this experiment. In the experiments, the signal frequency and scan step only have two variations, because the sensor only has two working frequencies and the scan period has two scan steps (0.45° and 0.9°) is too long to meet the AUV navigation requirements. The experiment deployment is shown in Figure 4-28.

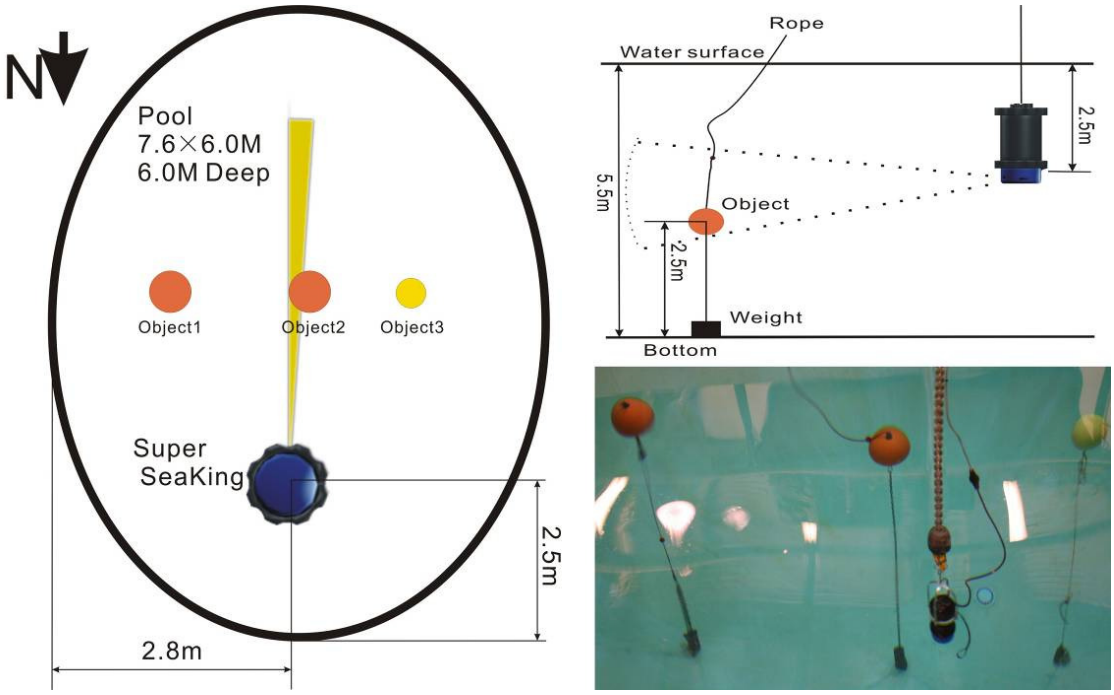


Fig. 4-28: Experimental set-up

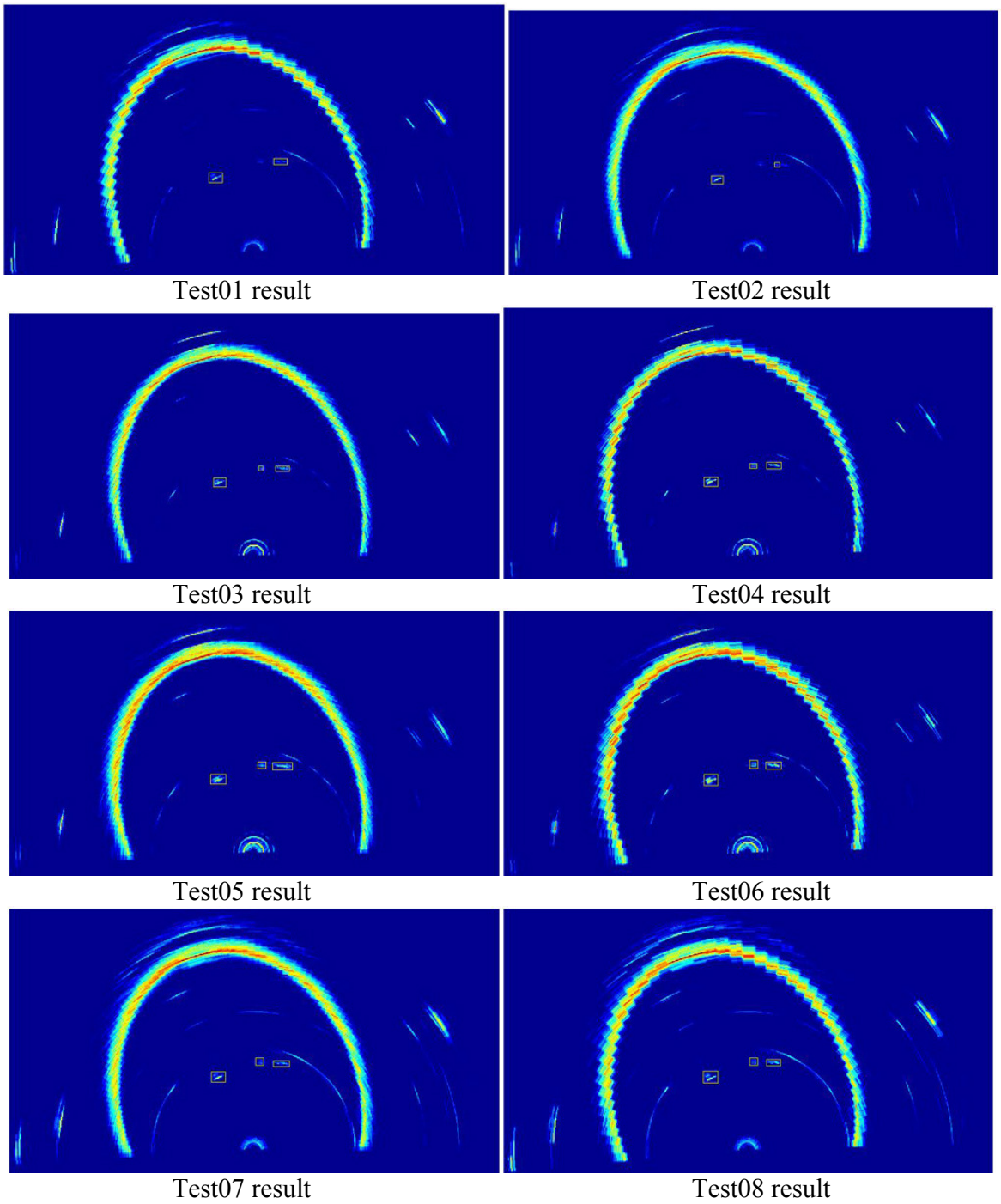
The designed orthogonal array and the final results are shown in Table 4-4 and Table 4-5, respectively. Targets detection results are shown from Figure 4-29 to Figure 4-30, where the targets are highlighted by the bounding boxes and overlapped on the original image. The miss detection represents that a target is displayed on the acoustic image but it is missed by the detector. The false alarm represents that the detector classifies a non-target echo as a real target. When the 'Gain' was set to zero, the developed detector marked the yellow sphere as background noise and miss detection occurred (see Figure 4-29). Also, since the detector cannot detect the target in three continued sonar scan, only three image frames were taken.

Tab. 4-4: Orthogonal experiment arrangement

Factors	Gain (%)	Adlow(dB)	Dynamic Range (dB)	Frequency(KHz)	Scan Step(degree)
Test 01	0	5	10	300	3.6
Test 02	0	10	15	300	1.8
Test 03	0	15	25	670	3.6
Test 04	0	20	35	670	1.8
Test 05	10	5	15	670	1.8
Test 06	10	10	10	670	3.6
Test 07	10	15	35	300	1.8
Test 08	10	20	25	300	3.6
Test 09	15	5	25	670	1.8
Test 10	15	10	35	670	3.6
Test 11	15	15	10	300	1.8
Test 12	15	20	15	300	3.6
Test 13	20	5	35	300	3.6
Test 14	20	10	25	300	1.8
Test 15	20	15	15	670	3.6
Test 16	20	20	10	670	1.8

Tab. 4-5: Detection results

	No. of image frames	Detected targets	Missing detections	False alarms
Test 01	4	2	4	0
Test 02	3	2	3	0
Test 03	3	3	2	0
Test 04	3	3	0	0
Test 05	5	3	0	0
Test 06	5	3	0	0
Test 07	5	3	0	0
Test 08	5	3	0	0
Test 09	5	3	0	0
Test 10	5	3	0	0
Test 11	5	3	0	0
Test 12	5	3	0	0
Test 13	5	3	0	0
Test 14	5	3	0	0
Test 15	5	3	0	0
Test 16	5	3	0	0



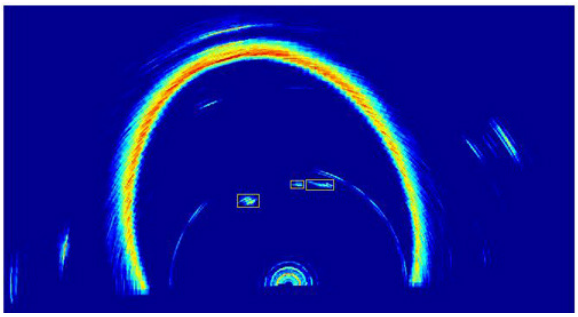
Test01 result Test02 result

Test03 result Test04 result

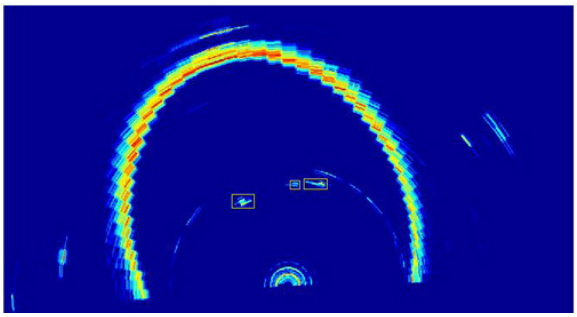
Test05 result Test06 result

Test07 result Test08 result

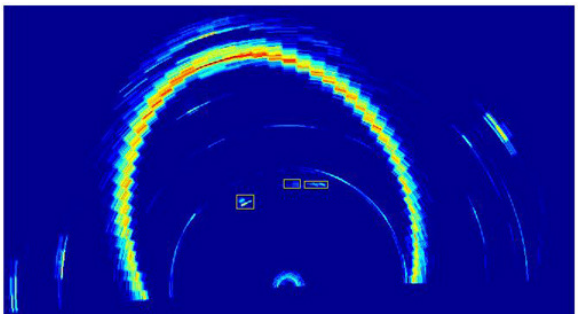
Fig. 4-29: Static targets detection results for the first eight tests.



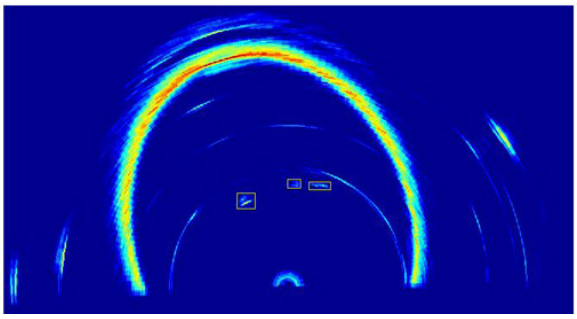
Test09 result



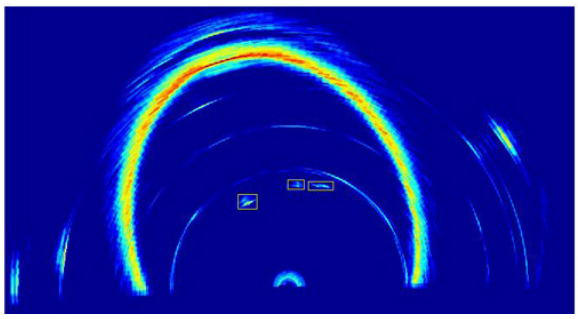
Test10 result



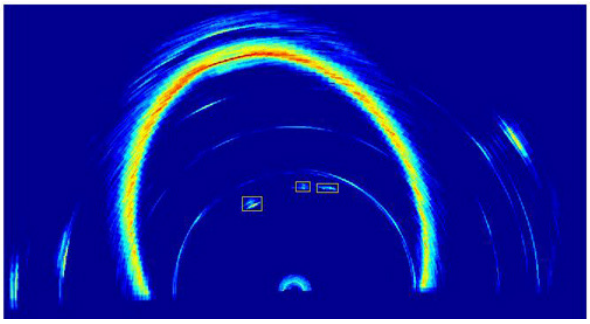
Test11 result



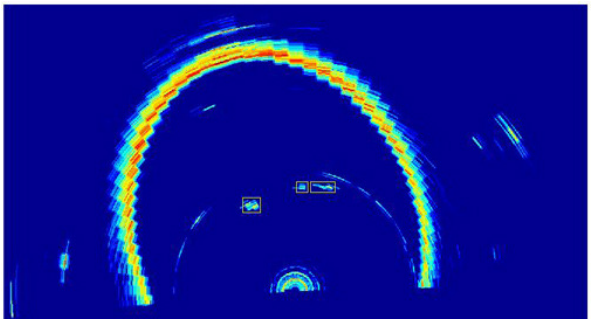
Test12 result



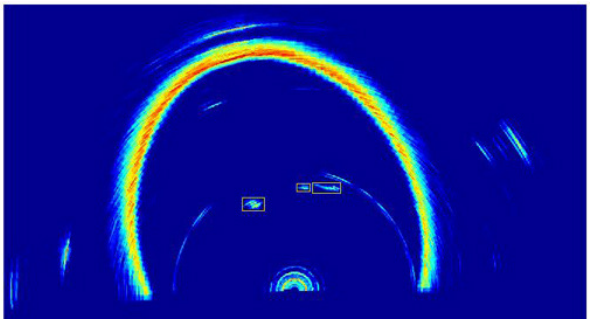
Test13 result



Test14 result



Test15 result



Test16 result

Fig. **4-30**: Static targets detection results for the last eight tests.

4.8.2 Mapping results

To test the mapping algorithm, the sensor was secured by the claw and suspended at a depth of 2.5m in three different positions, while the scan range was set at 6 meters. At each position, over 20 images were taken. During sampling, sensor control factors, the number of targets and their positions were varied.

In the first position, Super Sea-King sensor was operated at 670 kHz and step size was 3.6 degrees. Two spheres were used as dynamic obstacles in the test tank. The dimension of the test tank measured by the SeaNet Pro software was 735.7cm × 581.5cm (Major axis × Minor axis). The procedure of the edge detection is illustrated in Figure 4-31. One of the final mapping results is shown in Figure 4-32 and Table 4-6 shows the coefficients of the elliptic test tank in eight continued image frames.

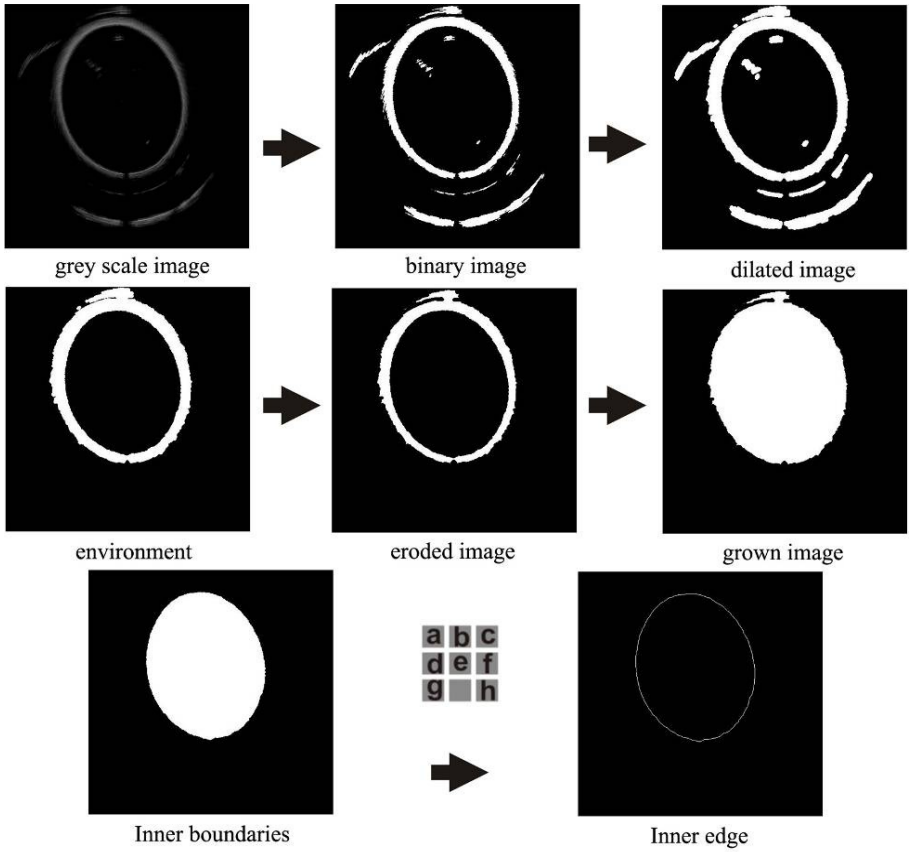


Fig. 4-31: Test tank detection procedure.

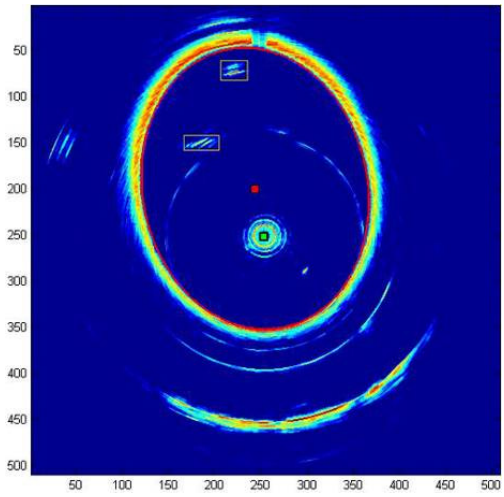


Fig. **4-32**: Elliptical test tank mapping result at the first sonar position. The result is overlapped on an original acoustic image for comparison. The red ellipse indicates the algebraic map. Sensor position is highlighted by the square in green and the obstacle is depicted inside the yellow bounding box.

Tab. **4-6**: Coefficients of the test tank and sonar position

No.	SC	TC	MaL	MiL	PO
1	(257, 252)	(243.59, 199.62)	723.04	575.49	-1.76
2	(257, 252)	(243.59, 199.62)	723.52	574.87	-1.76
3	(257, 252)	(243.70, 199.66)	723.74	575.22	-1.76
4	(257, 252)	(243.63, 199.57)	724.03	574.99	-1.76
5	(257, 252)	(243.46, 199.71)	723.64	574.92	-1.77
6	(257, 252)	(243.49, 199.50)	723.52	575.31	-1.76
7	(257, 252)	(243.45, 199.61)	724.07	574.74	-1.77
8	(257, 252)	(243.43, 199.56)	722.98	575.18	-1.77

No. – number of frames

SC – centre of the sonar sensor (refers to a pixel coordinates [39])

TC – centre of the test tank (refers to a pixel coordinates [39])

MAL – Major axis length of the test tank in centimetres

MIL – Minor axis length of the test tank in centimetres

PO – Orientation of the elliptic pool in radian

In the second position, Super SeaKing was operated at 300 kHz and step size was 1.8 degrees. One 56mm sphere was used as a static obstacle in the test tank. The dimension of the test tank measured by the SeaNet is 730.7cm × 579.5cm (Major axis × Minor axis). Figure 4-33 shows the mapping results and Table 4-7 summarises the coefficients of the elliptical tank in eight continued image frames.

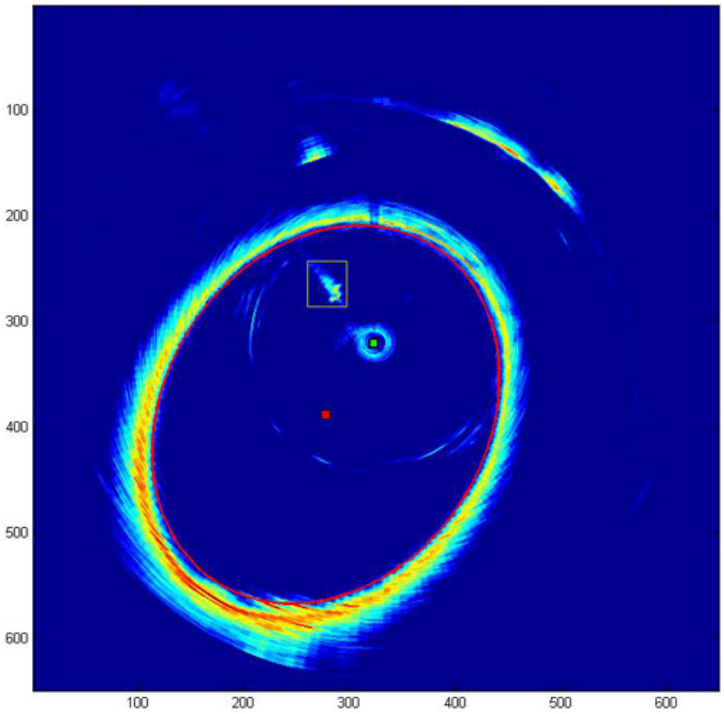


Fig. 4-33: Elliptical test tank mapping result at the second sonar position.

Tab. 4-7: Coefficients of the test tank and sonar position

No.	SC	TC	MaL	MiL	PO
1	(321, 324)	(278.09, 386.42)	710.43	578.73	-0.92
2	(321, 324)	(278.03, 387.89)	714.26	578.62	-0.98
3	(321, 324)	(277.68, 387.89)	715.59	580.47	-0.96
4	(321, 324)	(278.14, 387.68)	711.08	578.88	-0.95
5	(321, 324)	(277.69, 388.21)	712.08	579.73	-0.97
6	(321, 324)	(278.00, 387.67)	710.12	579.95	-0.95
7	(321, 324)	(277.84, 387.69)	710.86	579.32	-0.96
8	(321, 324)	(277.81, 388.06)	711.23	579.32	-0.97

In the third position, Super SeaKing was operated at 670 kHz and step size was 1.8 degrees. Sonar scan sector is 180 degrees only. Three spheres were used as static obstacles in the tank. The dimension of the test tank measured by the SeaNet Pro is 734.7cm × 572.74cm (Major axis × Minor axis). Figure 4-34 shows the results and Table 4-8 outlines the coefficients of the elliptical tank in eight continued image frames.

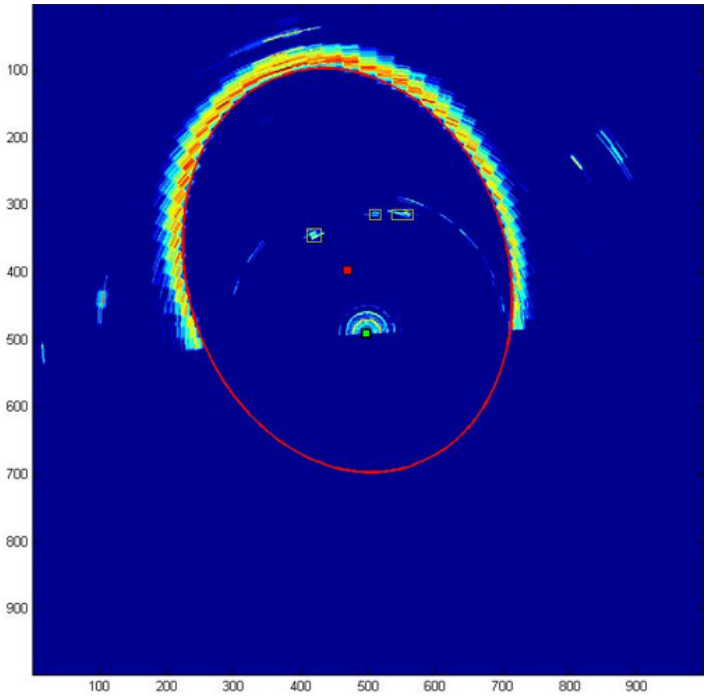


Fig. 4-34: Elliptical test tank mapping result at the third sonar position when sonar scan sector was 180 degrees only.

Tab. 4-8: Coefficients of the test tank and sonar position

No.	SC	TC	MaL	MiL	PO
1	(498, 491)	(469.06, 397.15)	733.87	572.27	-1.87
2	(498, 491)	(468.61, 398.07)	735.82	573.33	-1.87
3	(498, 491)	(469.19, 396.67)	733.01	572.02	-1.87
4	(498, 491)	(468.49, 398.27)	736.58	573.39	-1.87
5	(498, 491)	(469.46, 397.73)	735.70	573.08	-1.87
6	(498, 491)	(469.04, 397.61)	734.94	572.56	-1.87
7	(498, 491)	(468.99, 397.45)	734.14	572.73	-1.87
8	(498, 491)	(469.21, 397.06)	733.82	572.54	-1.87

4.8.3 Dynamic targets detection and tracking results

To test the detection and tracking algorithm, four different trials were conducted in the elliptic tank. The first stage has one static target and one dynamic target. The targets are two buoy balls (8.8cm in radius). The static target was located right behind the sensor and dynamic target was roughly started at 135 degrees (according to sonar bearing) and moved towards the static target (see Figure 4-35). A complete 360 degree scan took about 4 seconds when the sonar scan step was 3.6 degree and scan range was 6 meters. The detected trajectory is shown in Figure 4-36. It is important to understand that as the sensor was hanging upside down, objects in acoustic images will be viewed from left to right.

Fig. 4-35: Targets in the tank. The black line is the trajectory of the dynamic target. (The path is generated from a sequence of digital images by connecting the gravity centres together.)

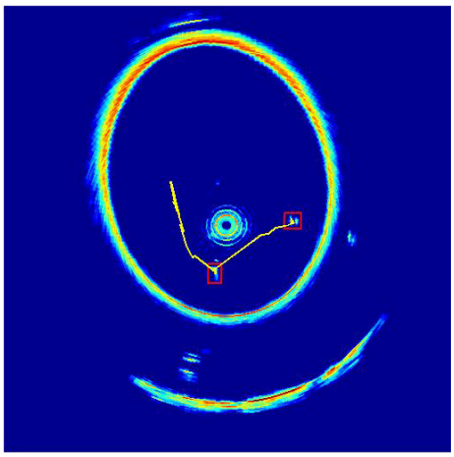


Fig. 4-36: One dynamic and one static targets detection and tracking result. The trajectory of the dynamic target is depicted by the yellow line.

The second scenario is similar to the first one but the starting point was different. The trajectory is shown in Figure 4-37.

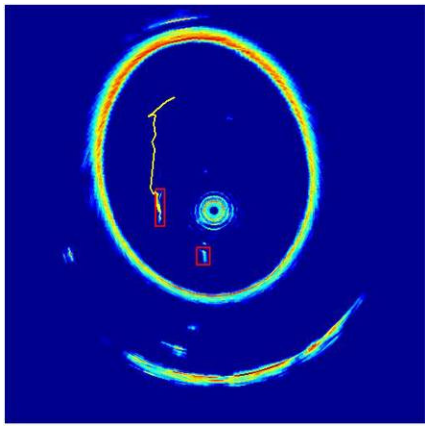


Fig. 4-37: One dynamic and one static targets detection and tracking result. The shape of the dynamic target is shown as a long and narrow rectangle because the underwater rope was also detected by the sensor.

In the third scenario, one 8.8cm buoy-ball was located at the back of the sensor. Two 14cm buoy-balls were used as dynamic targets. Forty-four image frames were recorded inside the tank. The final trajectories are shown in Figure 4-38 and the detection and tracking results are illustrated by a sequence of acoustic images in Figure 4-39.

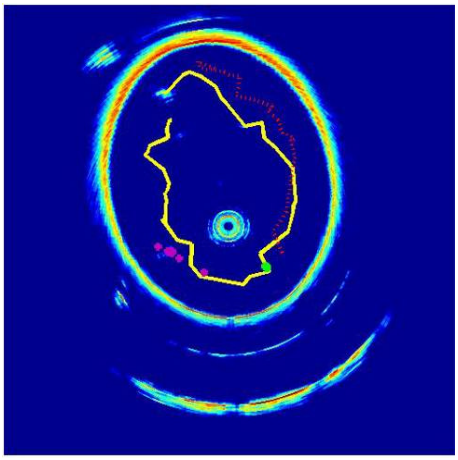


Fig. 4-38: The trajectories of the targets in the whole scan. The green dot illustrates the static targets. Due to the missing detection, one trajectory of the dynamic target is not connected and is shown by red and pink colour.

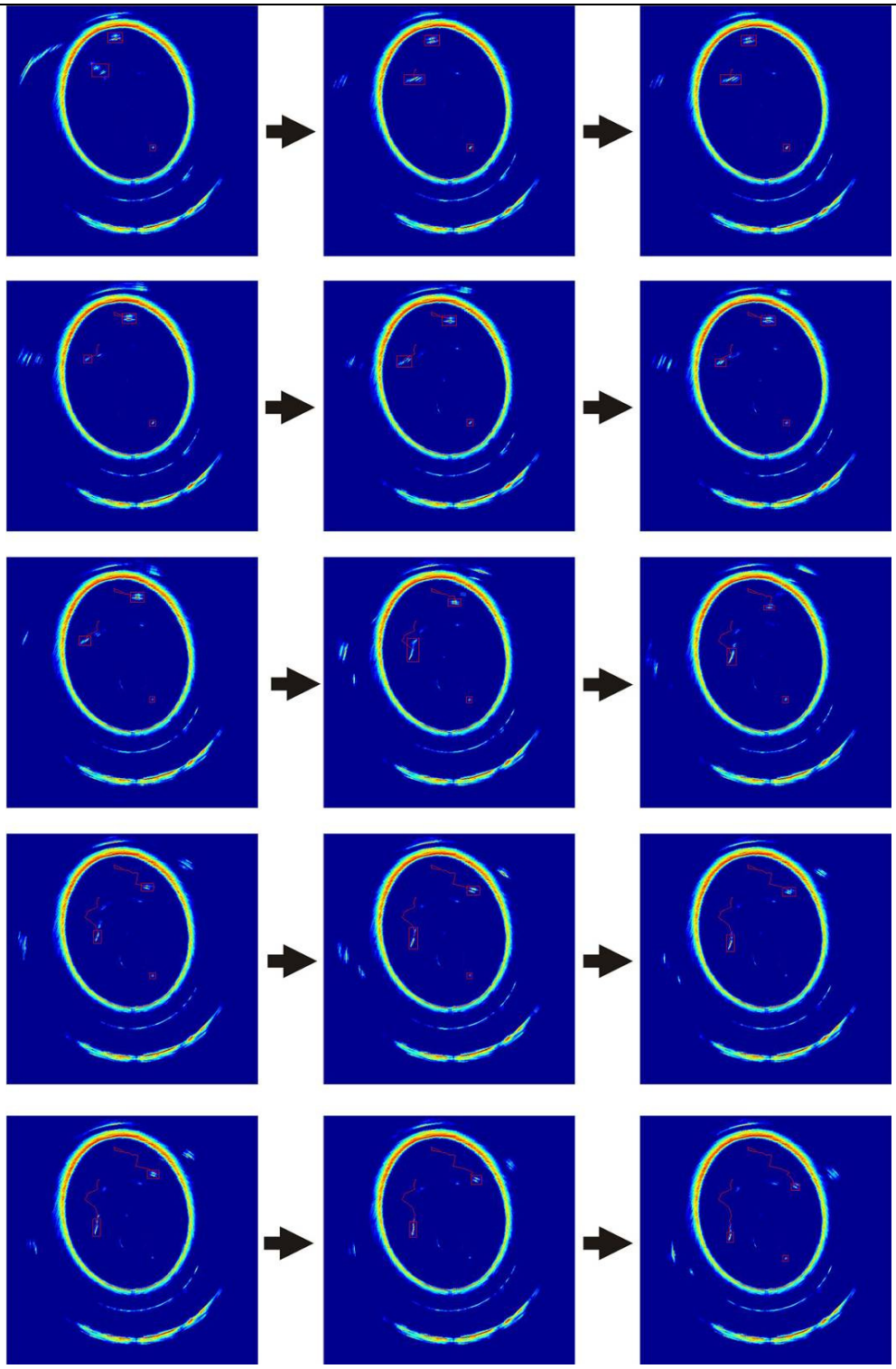


Fig. **4-39**: Sequence of sonar scans for the targets trajectories.

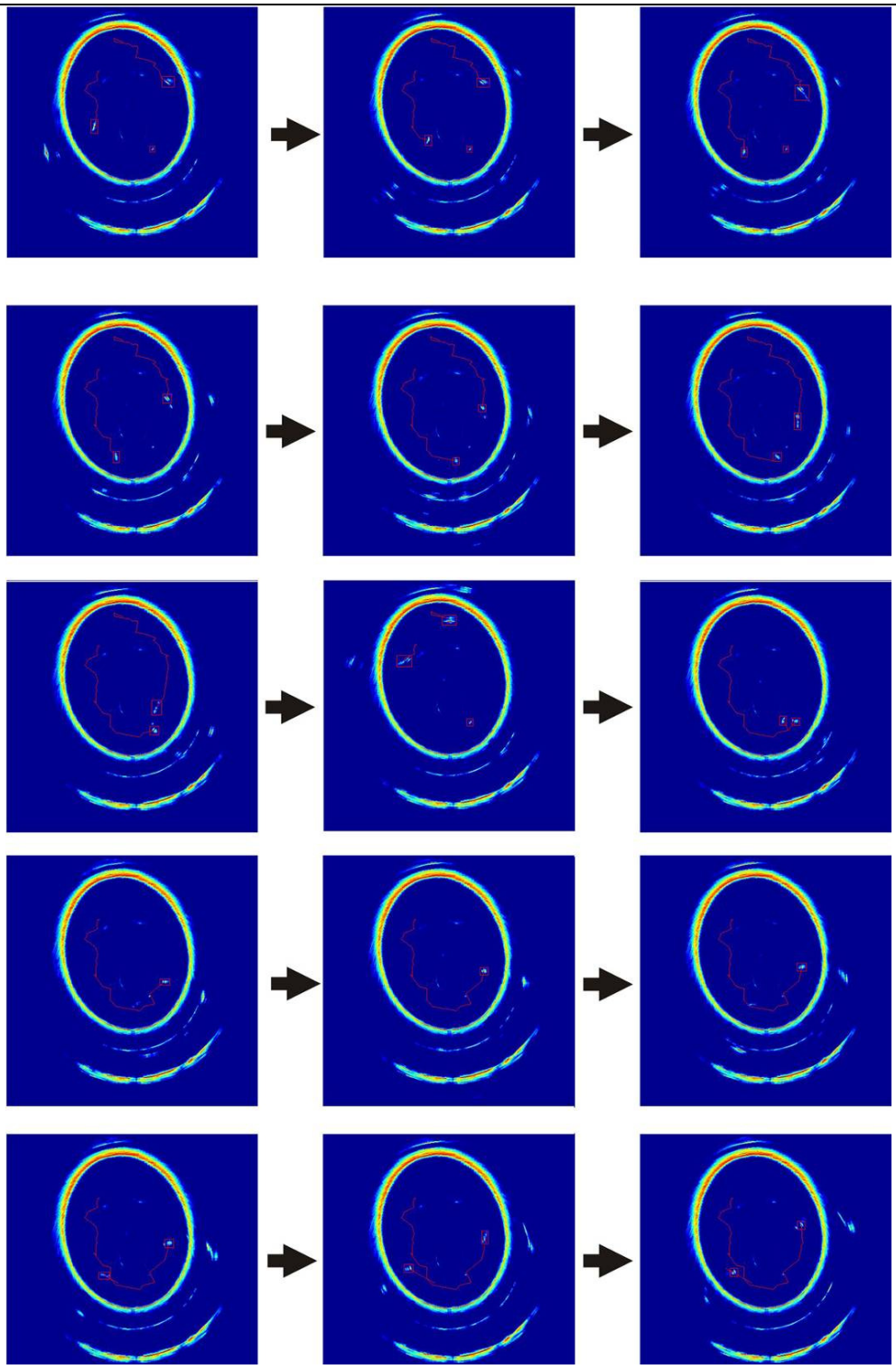


Fig. 4-39: Sequence of sonar scans for the targets trajectories.

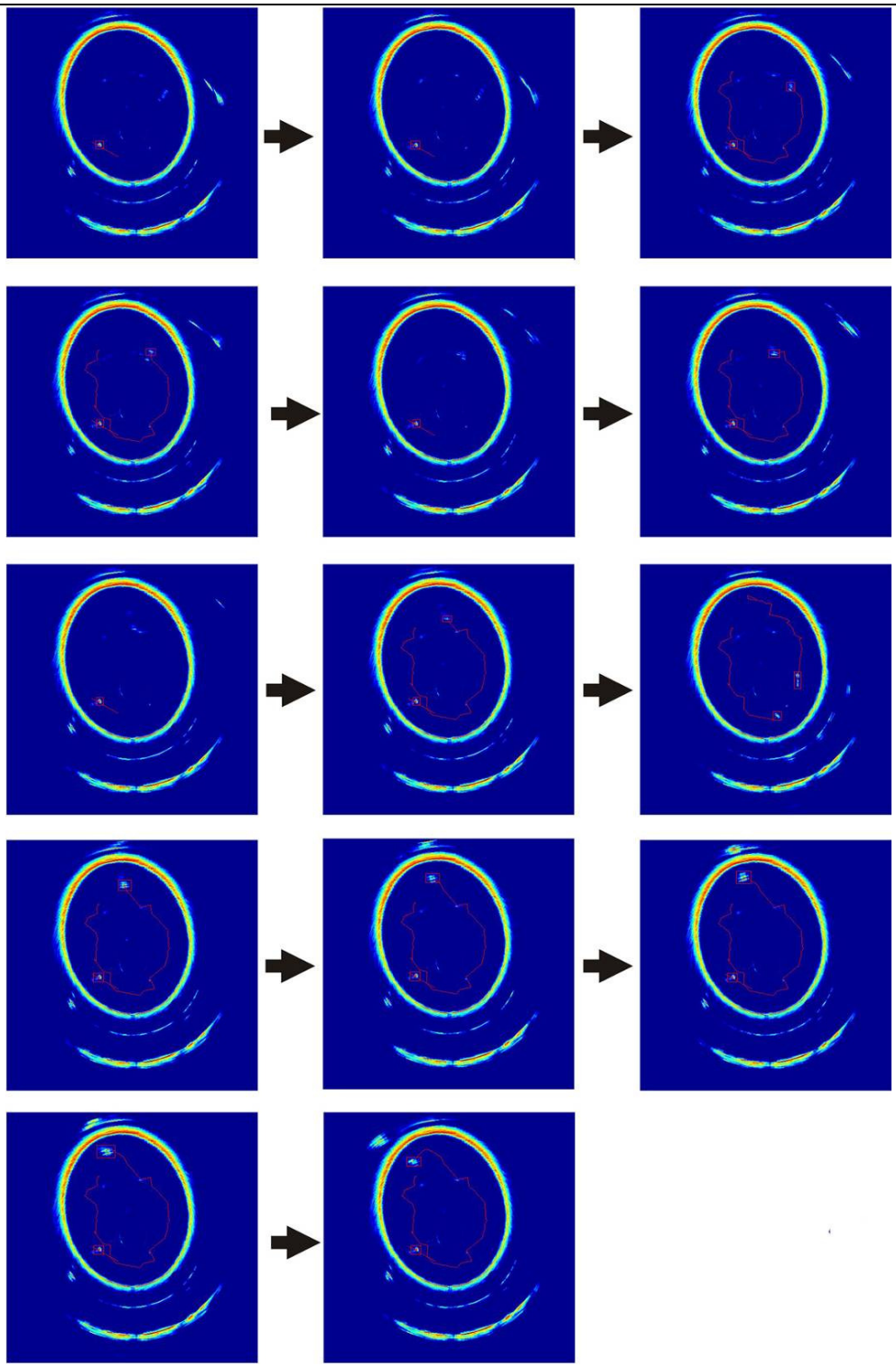


Fig. 4-39: Sequence of sonar scans for the targets trajectories.

In the fourth scenario, one 14cm buoy-ball started at the top of the test tank and was moved in a vertical motion towards the sonar sensor, while another one centrally orbited the sensor. Figure 4-40 indicates the trajectories of the two moving objects.

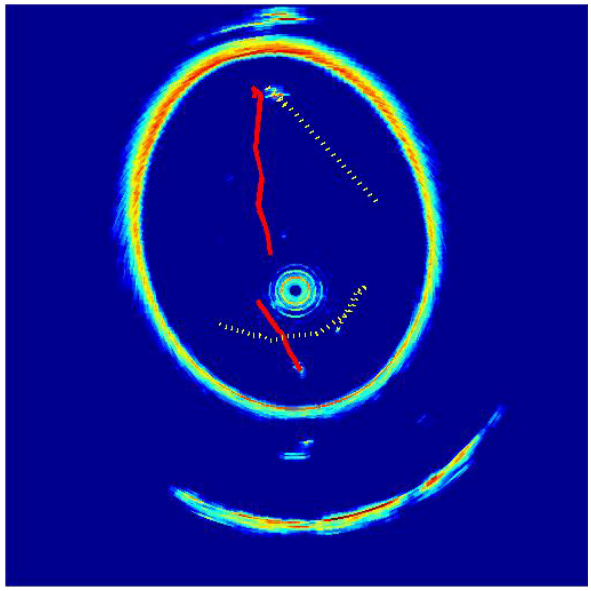


Fig. 4-40: Trajectories of the two moving objects inside the test tank.

4.9 Discussion and concluding remarks

The experimental results strongly support the methodology for detecting small suspended underwater objects. A purpose-designed filter was successful in managing sensor noise and the suppression of reverberation. It also retained useful information for further detection. Over three hundred acoustic image frames were taken in the test tank during different situations. The false alarm (marking the reverberation or other noise as the object) and missing detection (miss the real target) was less than 5%. Only the very strong reflections from the bottom or from the surface would occasionally cause false alarms. The missing detections mainly occurred when the strength of acoustic signals (gain of the sonar sensor) was low or the objects were adjacent to the tank walls, where the dilation operation connected the gaps between the objects and boundaries.

The simple and efficient detection algorithm exhibits a high level of accuracy and assists the object tracking algorithm. Subsequently the mathematical map, which is generated by extracting the features from the acoustic image, shows a very positive result. The detection tracking and mapping results support the AUV's navigation and obstacle avoidance strategies. However, the dimensions of the elliptic test tank are obtained using three different methods: the design dimensions (original dimensions), the SeaNet Pro measurements and the least-square fitting results. Since the original dimension is the standard for comparison against the majority of measurements, errors arise from the current sonar beam shape model and the geometrical shape (elliptic paraboloid) of the test tank. A-fan-shape beam has a 3.6 degree horizontal beamwidth and 40 degree vertical beamwidth. As the beam hits the boundaries of the test tank, specular reflections occur at the vertical tangent surface. To plot the horizontal sonar images, the tangent surface is compressed into one point. Such an acoustic imaging strategy causes the blurred boundaries in the image and therefore results in measurement errors. Moreover, in this study, the real boundary of the test tank in each single beam is assumed to be at the place where the first specular reflections occurred. Hence, the silhouette image of the test tank will affect the final ellipse fitting results. In Figure **4-33**, the inner boundaries acquired by the sonar have several bumps. Therefore, the measured dimensions of the test tank are smaller than the others. Nevertheless, these measurements can be considered as accurate for the purpose of AUV's navigation inside the test tank and also more accurate measurements can be obtained from different image frames.

The self-developed multiple range reverberation suppression filter successfully removed the reverberation noise and retained small object echoes. The morphological detector demonstrated satisfying results in a test tank. The environmental information extracted from the sensorial image will assist the AUV navigation within such confined environments.