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 5

Object Detection and Tracking in Natural Environments

Realistic AUV operation environments, such as rivers, dams and harbours are typically characterised as having a high degree of uncertainty. The use of active sonar in such natural shallow water results in receiving echoes reflected from targets, marine animals and other underwater objects. Furthermore, ambient noise caused by underwater activities in natural water environments, and strong reverberation arising from the bottom backscattering caused by rocks, coral and muddy river bed will also be recorded by the sonar hydrophone. The presence of echoes, ambient noise and strong reverberation limits the performance of target detection using active sonar. In this chapter, a fuzzy logical based algorithm used in the detection of underwater obstacles in natural shallow water environments, is described in detail.

5.1 Detection with *a priori* knowledge

In Chapter 4, a sequence of image processing techniques was described for small obstacle detection in structured environments where the water is relatively calm and clean. For the purpose of comparison, the detector that was developed in Chapter 4 is applied to a natural shallow water environment in this section. However, it cannot provide reliable results without *a priori* knowledge of the environment. Here, *the a priori* knowledge is consisted of **background intensity** and **target size**. In this study, the knowledge is referred to as two parameters: intensity threshold and area threshold.

Ambient noise and boundary reverberation were the two major factors which form background noise against the detector in natural environments. In acoustic images, they display as image components with different colours and sizes (see Figure 5-1). For the Super SeaKing, the luminance of an image is controlled by the parameter ‘Adlow’, which is an intensity threshold used to filter out the background noise and isolate the image components. The higher the value, the less sensitive the sensor will be. Using the SeaNet Pro (survey-data-acquisition and logging-software package), human operators can easily adjust ‘Adlow’ and decide on an appropriate value according to the output images. Object echoes can be identified by human operators through the adjustment. Figure 5-1 illustrates one acoustic image after applying two different values of ‘Adlow’. In the top one, noise can be observed as purple in colour and real object echoes are in green and yellow. The bottom one is relatively cleaner than the top one, although the colour intensity of the object echoes is relatively low (object echoes are displayed in purple). Therefore, most ambient and reverberation noise can be eliminated from the acoustic image using a proper intensity threshold, which is referred to as the value of ‘Adlow’.

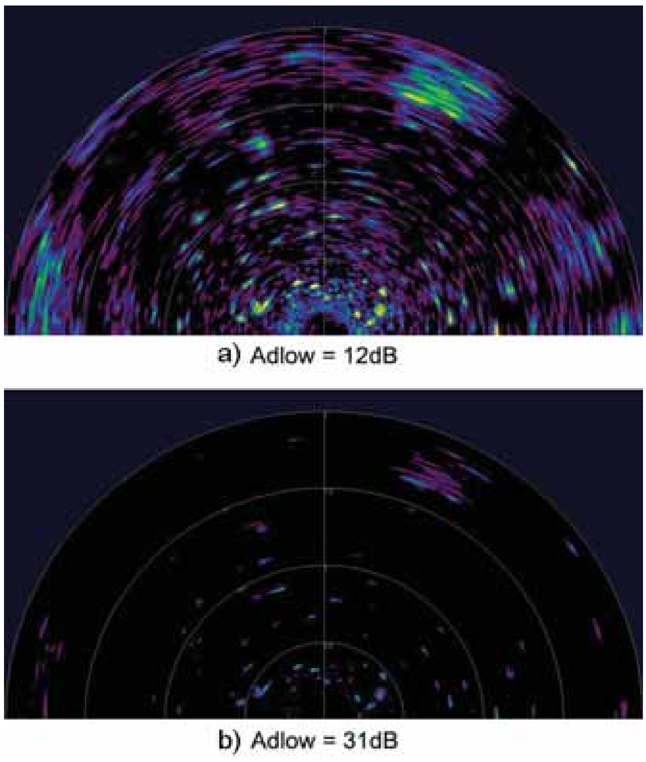

Fig. 5-1: Sonar image thresholded by two different intensity values. a) Adlow = 12dB; b) Adlow = 31dB. A high threshold can filter out most background and retain the object echoes in the image. However, some information about the objects is missing after thresholding.

Another parameter, 'target size', acted as an area threshold that estimated target objects against ambient noises for the morphological detector. Those noises, generated from fish, small rocks and other aquatic creatures, display high colour intensity but their sizes are small in the acoustic image. In Figure **5-1**, some small speckles which are close to the sonar centre can be observed after the intensity threshold is applied. To eliminate those noises, an area threshold which relies on the parameter 'target size' was applied. Based on the above descriptions and algorithms presented in Chapter 4, the target detection contains the following steps.

Intensity thresholding:

Thresholding each single ping by the pre-determined parameter 'Adlow' before the pre-processing stage.

Pre-processing:

Performing the self noise filter and reverberation suppression filter (see section 4.2.1 and 4.2.2).

Median filtering:

Smoothing the raw data with a 3-by-3 neighbourhood median filter.

Acoustic image representation:

Mapping the raw data into Cartesian space.

Otsu binary thresholding:

Thresholding the grey level image into a binary image

Morphologically detecting:

Performing morphologic operations and then obtaining the size of each image component.

Area thresholding:

Evaluating the size of each image component with the 'target size'. Image components whose sizes are larger than the threshold will be marked as target objects and remain in the image.

Figure 5-2 shows the detection result for one acoustic image frame.

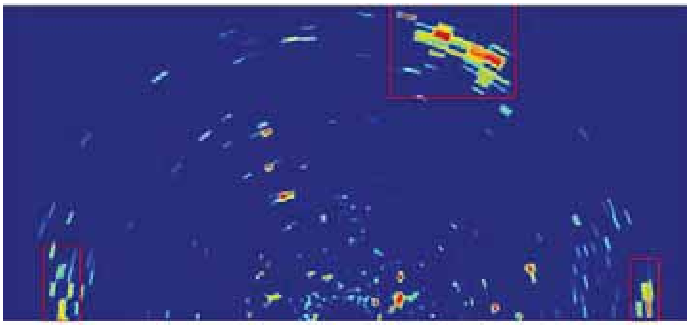

Fig. 5-2: Target objects detection result

The code for this detector is shown in Figure 5-3.

```
%% Image processing techniques for natural environment detection
% input: 'sonar_raw_data'
% output: 'resultImg',Result image with labelled targets
% output: 'decTarget' a class which save the object information, including
% the target numbers, centers, areas and bounding boxes of the targets
function [resultImg, decTarget sonarX, sonarY] = ImageDetector(sonar_raw_data)
% Pre-processing
rawData = read(sonar_raw_data);% read data
thresholdRaw = getinput();% read threshold
% intensity threshold
find(rawData < thresholdRaw) = 0;
% Media filtering
smoothData = media(rawData);
% represent in X-Y plane
acousticImg = transfer(smoothData);
binaryImg = Otsu(acousticImg);
% set the area threshold for morphologic detector. threshold.1: less than 2
% meters, threshold.2: over ten meters
thresholdArea = setthreshold();
oct = STREL(3,'ocatgon'); % a octagon STREL structuring element
dilObjects = dilate(objects, oct); % Dilation operation
numObjects = getNum(dilObjects); % calculate the numbers of the objects
% Evaluating the regions in the image;
if Objects.size <  thresholdArea
Object.intensity = 0
% 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) % save the data into the class

% extract the boundaries of the target by morphological operation
resultImg = boundary(targetsImg);
```

Fig. 5-3: Code for detecting in the natural environment using imaging processing.

5.2 Objects detection without *a priori* knowledge

Using a *priori* knowledge, the aforementioned detector can successfully extract the obstacles from the acoustic image. However, the accuracy of the detector is compromised without such environmental knowledge. Also, it is inefficient to adapt the random underwater conditions to the detector. A technique that does not require specific knowledge is very much required when dealing with the random operational environment of the AUV. On the other hand, the conversion from raw data to acoustic images costs excessive computational time, which could be a problem for real-time performance of AUVs. To improve system efficiency, it is necessary to filter out most noises before the conversion or detection of obstacles directly from the raw data. In this section, the reason for the selection of fuzzy logic is discussed. An analytical study of the ping follows and this helps to understand the developed fuzzy inference system. Finally, the developed fuzzy detector is detailed.

5.2.1 Challenges of automatic detection

In Chapter 2, the most popular methods for the detection process using underwater sonar were reviewed. However, they all have limitations. For instance, Matched Filter may be the simplest way to detect small obstacles, but it has great difficulty coping with coloured reverberation [24]. The thresholding or adaptive thresholding technique may filter out the echoes from objects due to the presence of strong reverberation [28]. Moreover, Adaptive Clutter Filter, which is used to suppress the background clutter, requires information concerning the average highlight region and the shadow region in side scan sonar (SSS) images [47]. Unfortunately, in the present study, the shadow region was always covered by strong reverberation when mechanically scanned image sonar (MSIS) was used in shallow water. The low resolution of the MSIS restricts the use of fractal-based analysis [48], spatial point processing [49] and dual hypothesis [50] techniques all classify objects according to their textures. Furthermore, the success of these three methods depends heavily on large training samples and simplifying the modelling assumption [51].

Static and moving targets detection with MSIS in shallow water is heavily degraded by the presence of reverberation. The crux of automatic detection using MSIS in shallower water is to identify distinct features of object echoes. However, in this study, it is a complex challenge for the following reasons:

- the uncertainty of the underwater environment;
- small target echoes which were surrounded by large water-bottom reverberation;
- insufficient development of mathematic models for non-stationary reverberation.

Since common mathematical methods cannot solve these problems properly and efficiently, a more intelligent and automatic method which uses human-like reasoning is required to meet the aims of this study. The most popular artificial techniques based on physiological metaphors are neural networks and fuzzy logic. Neural networks are renowned for their ability to learn and generalise from example data even when the data is noisy and incomplete [52]. It acquires the knowledge from training and stores the knowledge in a distributive manner throughout the whole neural network structure. Then the neural network responds to a novel situation by using the previously learned knowledge. Neural networks are capable of accommodating poorly modelled, nonlinear dynamical systems [52]. However, the need for a large number of samples for the training procedure is not suitable for this study. Further, the representation of the distributed knowledge in neural networks is difficult to be interpreted by easily understandable concepts. Fuzzy logic was first proposed in 1965 when Lotfi Zadeh published a paper titled 'Fuzzy Sets' in the journal *Information and Control* [53]. It is designed as a mathematical tool for dealing with imprecision and information granularity [54]. The term fuzzy logic has two different senses. In a narrow sense, fuzzy logic *(FLn)* is a logical system – an extension of multivalued logic that is intended to serve as a logic of approximate reasoning. In a wider sense, fuzzy logic *(FLm)* is more or less synonymous with fuzzy set theory; that is, the theory of classes with unsharp boundaries [55]. Lotfi Zadeh remarked: "In almost every case you can build the same product without fuzzy logic, but fuzzy is faster and cheaper [55].” The use of fuzzy logic has following advantages.

- ***It is easy to understand:***

Fuzzy logic is based on natural language for human communication. The mathematical concepts behind fuzzy reasoning are very simple.

- ***It is adaptive to imprecise data:***

Nothing is precise. Fuzzy proposition is a matter of degree which builds understanding into the process rather than tacking it onto the end.

- ***It models nonlinear functions:***

Fuzzy logic relies on heuristic knowledge which is subject to the designer’s experience and interpretation of the system. A fuzzy system is a nonlinear matching of input and output data.

Therefore, fuzzy logic is more suitable than neural networks for this study.

5.2.2 Nature of the ping

Human knowledge is gained from experiencing the world in which we live. People cannot perceive the world exactly when they lack information. Descriptions for human perception are pervaded by uncertain concepts which do not have sharply defined boundaries. In section 5.1, image processing techniques were shown to successfully detect objects with the assistance of a *priori* knowledge. In fact, this knowledge was known by the researcher. A human operator detects a target object by identifying its distinguishable features according to his/her own experience, and so does the fuzzy logic. For instance, the target was a large component in an acoustic image which had high colour intensity. The term 'large' and 'high' are simple evaluating expressions in human mind. Using this heuristic reasoning, a fuzzy system can be built for the purpose of detection. However, this section offers a closer inspection into the nature of the ping. It attempts to identify distinctive features of objects’ echoes for the fuzzy detector.

A complete ping contains the environment information (reflection history of sound waves) along the propagation path of the sound beam emitted from the sonar. It is acquired by sampling the sound into a sequence of discrete Bins (a more detailed description can be found in section 3.3.2). Therefore, a single ping can be considered as a discrete signal and can be written as Equation 5.1.

$$Ping = \{Bin[n]\} \qquad 1 \le n \le \text{scan range} \tag{5.1}$$

Equation 5.1 means that a *Ping* can be represented as a discrete set of *Bin*[n], in which the *nth* number in the sequence is denoted *Bin*[*n*].

Similarly, the amplitude of the *Bin* or the colour intensity of each *Bin* can be defined mathematically as a function that maps an input sequence *Bin*[*n*] into an output sequence with values *bin*[*n*], say

$$bin[n] = f\{Bin[n]\} \tag{5.2}$$

According to the sound propagation and reflection behaviour underwater, if the $bin[n] = 0$, it indicates that there is no echo or there is no object reflecting sound waves. If the $bin[n] \neq 0$, it means there is an object (or objects) reflecting the sound waves and the echoes are recorded by the hydrophone (see Figure 5-4).

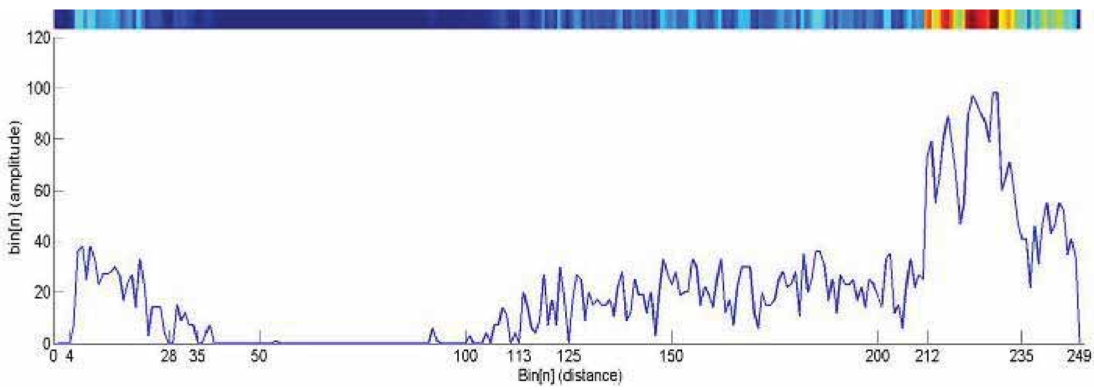

Fig. 5-4: A single ping received in the natural environment.

In Chapter 4, a single ping was partitioned into several sub-returns according to the pulse length (or the distance). Each sub-return represents a specific sound reflection surface. For instance, a sub-return can be observed from *Bin* [125] to *Bin* [249] which means reflections occurred from *Bin* [125] to *Bin* [249]. Among these, the highest colour intensity is acquired at *Bin* [228] (*bin* [228] = 100). Since the colour intensity is relatively higher than other Bins, it can be deduced by a human operator that there are object echoes inside the sub-return. There are two important parameters for a sub-return, which are the pulse length and the highest colour intensity, respectively. Therefore, these two parameters can be used for fuzzy rules to extract target objects against the background.

5.2.3 Fuzzy logical detector

A typical design for a fuzzy inference system (FIS) usually consists of five functional blocks as shown in Figure 5-5. However, a complete ping is firstly pre-processed by an intensity threshold to filter out low level noises. Then the ping is partitioned into several sub-returns according to the pulse length. Each sub-return is the input of the fuzzy inference system and the output is a crisp (non-fuzzy) number in percentage which represents the true probability of obtaining an object echo. In this study, if the output is over 49%, the corresponding sub-return is considered to be object echoes.

NOTE:
This figure is included on page 102 of the print copy of the thesis held in the University of Adelaide Library.

Fig. 5-5: Fuzzy logical detector. (The image of the FIS refers to [56])

5.2.3.1 Pre-processing

As shown in Figure 5-4, targets' echoes are always surrounded by reverberation in shallower water. However, since the aim of pre-processing is to filter out low level noise which is satisfied with the standard normal distribution, a one level intensity threshold is applied to each completed ping. Meanwhile, pre-processing can speed up processing for the whole system because the redundant data are eliminated.

Given a completed ping, the mean (denoted as m) and the standard deviation of the intensity value (denoted as σ) can be easily calculated. According the empirical rule in statistics, about 68% of the values lie within 1 standard deviation ($m \pm\sigma$) of the mean. Hence, the value of the threshold (T) is assigned to the sum of m and σ ($T = m + \sigma$). All the pixels whose colour intensity values are lower than the T will be filtered out. Moreover, the self noise around the sonar sensor will be eliminated in the pre-processing stage (see section 4.2.1). A complete ping after pre-processing is shown in Figure 5-6.

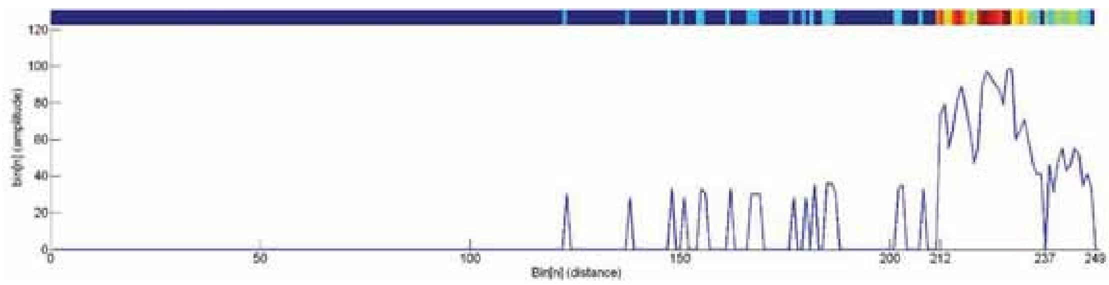

Fig. 5-6: A single ping after the pre-processing phase.

5.2.3.2 Fuzzy sets and membership functions

Human thinking and reasoning frequently involve fuzzy information which can be called fuzzy concepts. Fuzzy sets deal with such fuzzy concepts and expressions for computers. Whereas classical (or crisp) sets can be defined by characteristic functions, a fuzzy set $\tilde{A}$ can be characterized by membership functions by Equation 5.3 [57].

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\} \quad 5.3$$

where X is a collection of objects denoted by x, and $\mu_{\tilde{A}}(x)$ is called the membership function of $\tilde{A}$ (see Figure 5-7). The database in Figure 5-5 defines the membership functions of the fuzzy sets. In this study, the triangular membership functions are used because of their striking simplicity. However, even under some weak assumptions, these specific triangular membership functions immediately comply with the relevant optimization criteria [58].

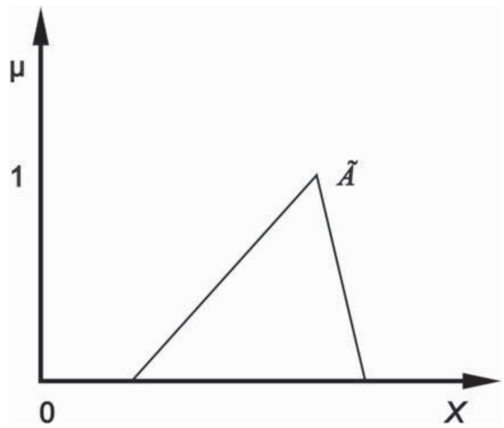

Fig. 5-7: A triangular membership function of fuzzy set $\tilde{A}$.

Basic set-theoretic operations for fuzzy sets operations include the union, intersection, and complement. They are defined differently as follows:

Union:

$$\tilde{C} = \tilde{A} \cup \tilde{B} = \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\} \quad x \in X \quad 5.4$$

Intersection:

$$\tilde{C} = \tilde{A} \cap \tilde{B} = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\} \quad x \in X \quad 5.5$$

Complement:

$$\overline{\tilde{A}} = \mu_{\overline{\tilde{A}}}(x) = 1 - \mu_{A}(x) \quad 5.6$$

5.2.3.3 Fuzzy rules

A fuzzy system is performed using interference rules, which are known as IF-THEN rules. The rule base in Figure 5-5 is described using *IF antecedent*, *THEN consequent* format. In this study, three fuzzy rules are employed to capture the imprecise modes of human reasoning:

- If the maximum amplitude is high then the probability is high;
- If the maximum amplitude is average and pulse length is large then the probability is average;
- If the maximum amplitude is low and pulse length is small then the probability is low.

Where maximum amplitude and pulse length are linguistic variables [59], high and low are linguistic values.

5.2.3.4 Fuzzy inference system

The fuzzy inference system (FIS) built in this study is based on Mamdani's fuzzy model [60]. The fuzzification interference (see Figure 5-5) firstly transforms the crisp inputs (each sub-return) into degrees which match with linguistic values. Then a decision-making unit (see Figure 5-5) performs the inference operations based on fuzzy rules. The fuzzified inputs are taken into the antecedents of the fuzzy rules. If a rule has multiple antecedents, the fuzzy operator (AND, OR) is applied to obtain one evaluation that represents the result of the antecedent for that rule. Finally, the defuzzification interference (see Figure 5-5) transforms the fuzzy results of the inference into a crisp output K, which is most likely to obtain an object echo. If the output is less than 49%, the corresponding sub-return will be classified as noise and then the colour intensity will be assigned to zero, which represents no echo return. If the output is over 49%, the corresponding sub-return will be considered as object echoes and will remain in the ping history. The developed FIS system which consists of two inputs (maximum amplitude and pulse length) and one output (probability) is shown in Figure 5-8.

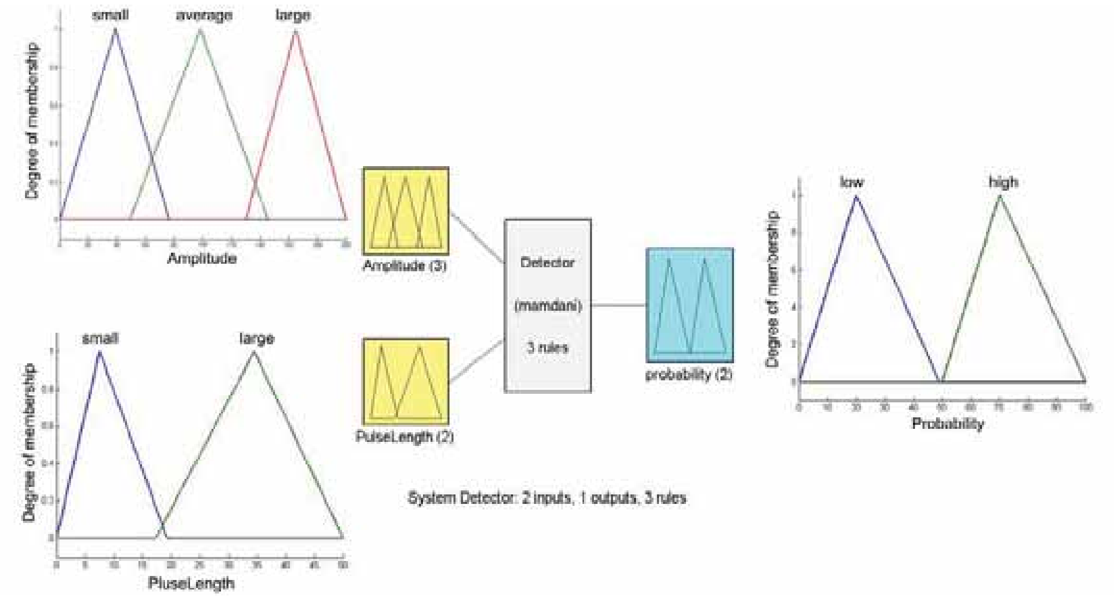

Fig. 5-8: An overview of Mamdani's fuzzy interface method.

For instance, the ping plotted in Figure 5-6 was partitioned into 15 sub-return and the outputs of the FIS are listed in Table 5-1.

Tab. 5-1: Outputs of the echo after defuzzification

echo	1	2	3	4	5	6	7	8	9	10
output	23.0028	23.0028	23.1290	23.0597	23.1290	23.1857	23.0597	23.1857	23.0597	23.0597
echo	11	12	13	14	15	16	17	18	19	N/A
output	23.1290	23.1857	23.1857	50	38.0166	N/A	N/A	N/A	N/A	N/A

According to the fuzzy output, only the 14th sub-return is object echo (or it contains object echoes). Since the bearing of the ping, sonar scan step and scan range are known, the object will be deduced by the following calculations. For instance, an object echo was detected in a sub-return at Bin [212], the bearing of this certain ping is 115th, the sonar scan step is 1.8 degrees and scan range is 200 cm respectively. The azimuth (θ) of the ping is $115 \times 1.8° = 207°$. The range (r) is $212 \times 200/250 = 170.3$ cm, where 250 is the total number of samples in the ping. The detection result is shown in Figure 5-9.

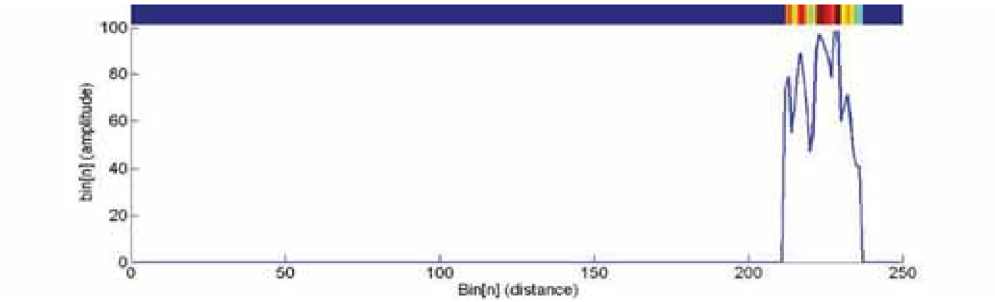

Fig. 5-9: A ping after performing fuzzy detector.

Figure 5-10 illustrates one frame of raw data after the fuzzy detector has been applied. However, since the raw data are hard to interpret, the corresponding acoustic images are also shown.

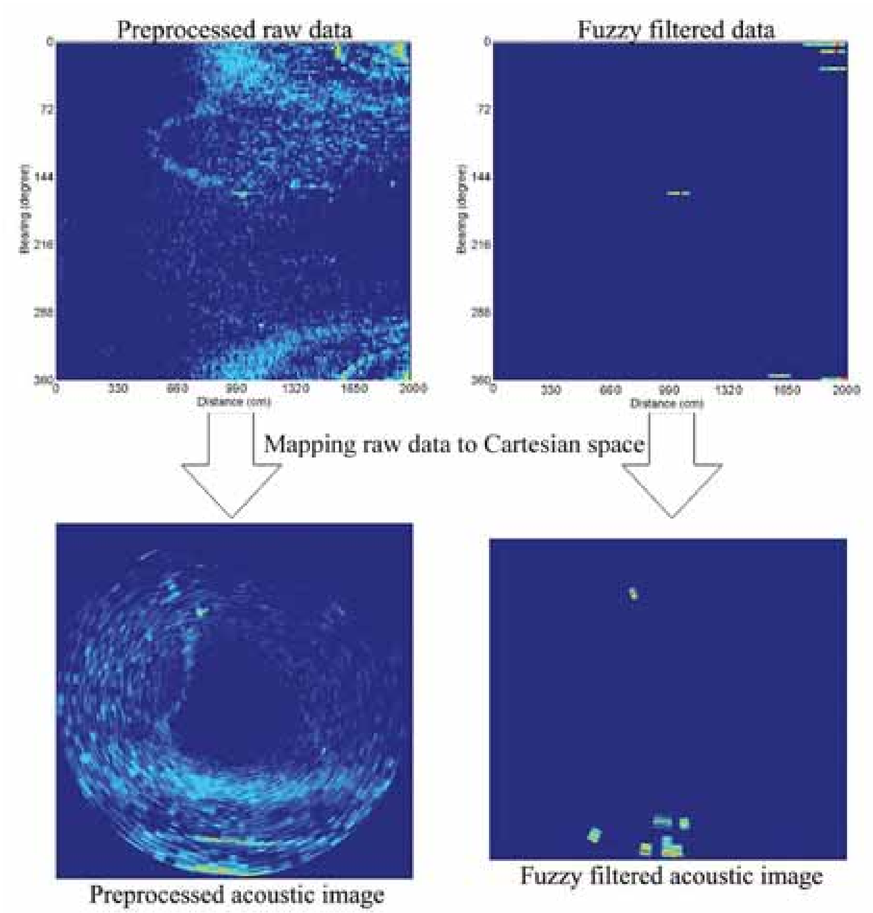

Fig. 5-10: Fuzzy logic filtering result for raw data

The code for the natural environment object detection using fuzzy logic is shown in Figure 5-11.

```
% fuzzy logic decison making and detection system
function [Ping,polarResult] = fuzzyFiltering(ping_beam);
% input: sensor received ping
% output: filtered ping
% output: object information in polar coordinate sysytems
% read each single ping into memory and acquire the parameters
[ping,bearing]  = read(sonar_ping);
binNum = getNum(ping); % the number of bins in single ping.
subIndex = label(ping)   % label the zero in the ping;
n = getNum(subIndex) % return the numbers of sub_ping;
echo = cell(n,1) % construct a class to save echo;
for i  = 1 to n
    echo{i} = segment(ping);
    subPing_marker = mark(ping) % record the sub_ping location in the
ping
    pluseLength = getlength(echo{i});
    maxAmplitude = getmax(echo{i});
    % Evaluating the echo by fizzy logic system
    probability = fuzzy(pluseLength,maxAmplitude)
    if(probability<49) % automatic selection
        echo{i} = zero;
    else
        polarResult.bearing = bearing; % get objects orientation
        polarResult.distance = subPing_marker;
        polarResult.length = pluseLength;
    end
end
% restore the filtered data into ping for displaying
for j = 1 to n
    Ping = restore(echo{j});
end
```

Fig. 5-11: Code for the natural environment detecting using fuzzy logic

5.3 Experimental results

This section presents two experiments in natural environments. The first experiment was carried out in a small artificial pond which corresponds to the proposed acoustic imaging processing method. The second one was performed in the River Torrens, specifically at Elder Park, Torrens Lake, Adelaide.

5.3.1 Small pond

In order to conduct this experiment, a small man-made pond located in Rymill Park, Adelaide, was used. This environment proved to be extremely challenging. Firstly the depth of the water was approximately 70 cm. Due to the depth of the water and the fact that the base of the pond was muddy, it was extremely difficult to obtain suitable target materials which could be suspended at a constant depth. The sonar was fixed to a permanent base and located centrally in the pond and the sensor scan distance was set at two metres. It was decided that the researcher's lower legs were a suitable target and this made the placement of the target at different distances more achievable. Figure 5-12 shows the experiment's environment and location. The detecting and tracking results of 15 continuous frames using *a priori* knowledge are shown in Figure 5-13, Figure 5-14 and Figure 5-15.

NOTE:
This figure is included on page 109 of the print copy of the thesis held in the University of Adelaide Library.

Fig. 5-12: Experiment pond in Rymill Park (image from Google map).

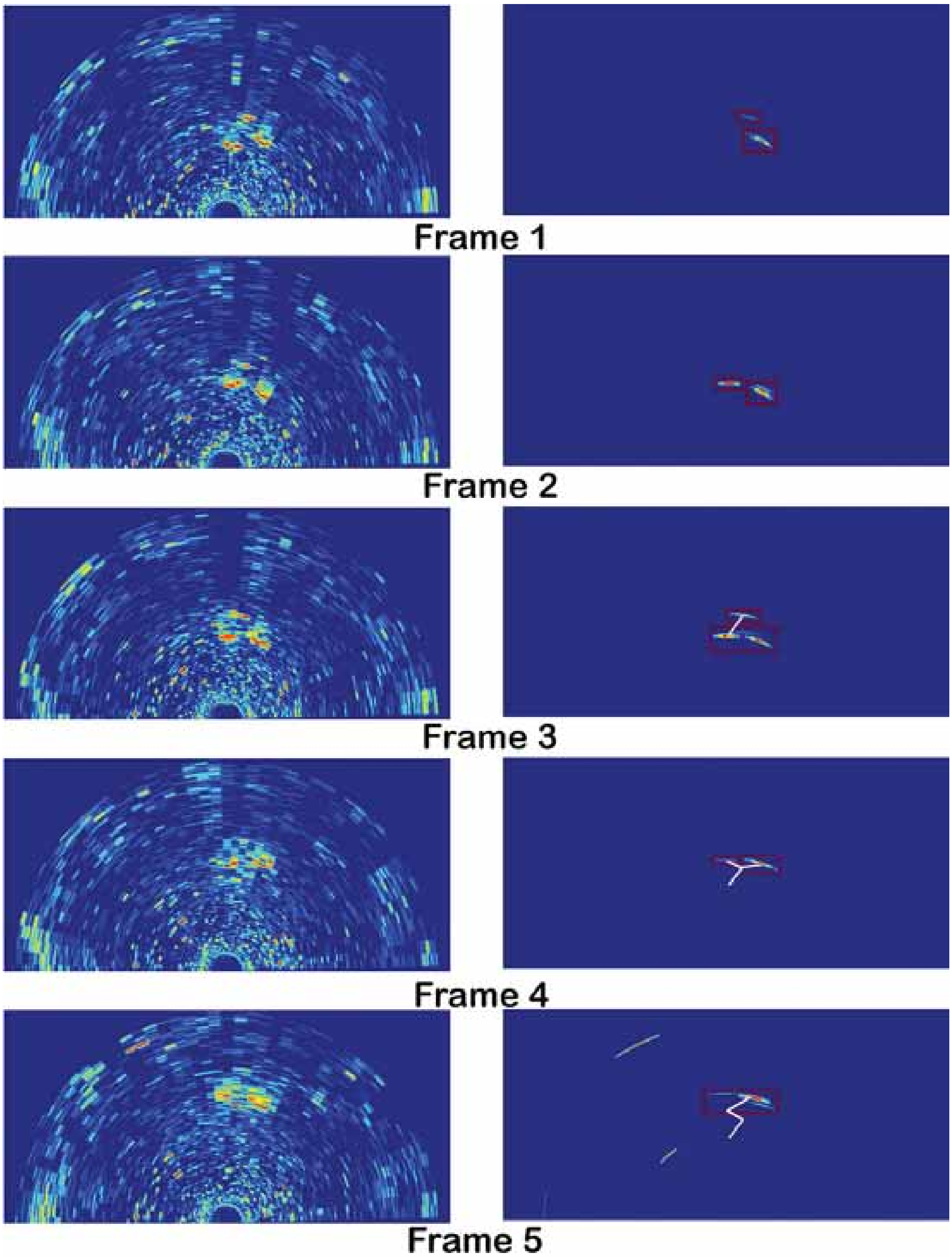

Fig. 5-13: Objects detection and tracking results using image processing techniques. From Frame 1 to Frame 5.

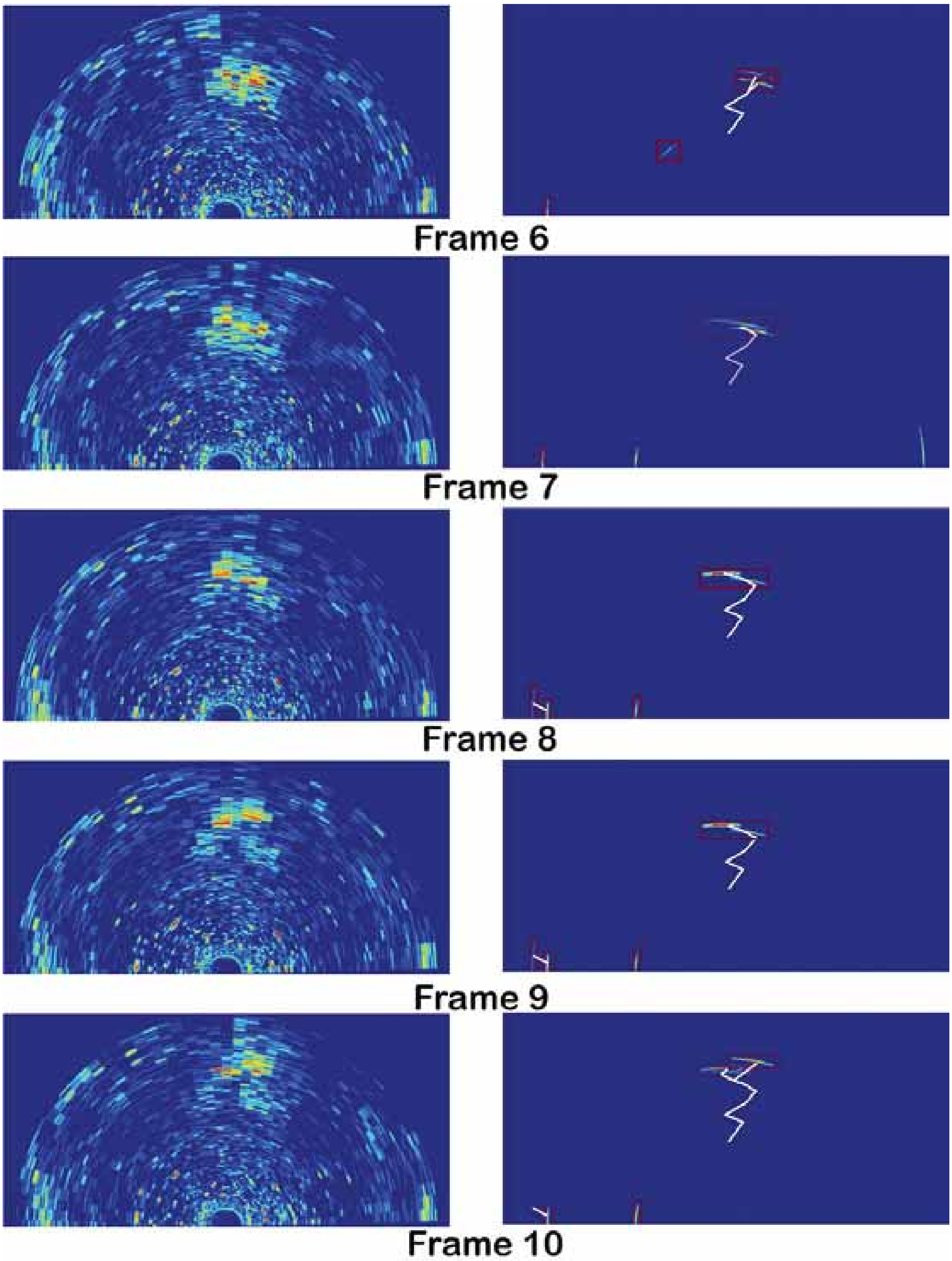

Fig. 5-14: Objects detection and tracking results using image processing techniques. From Frame 6 to Frame 10.

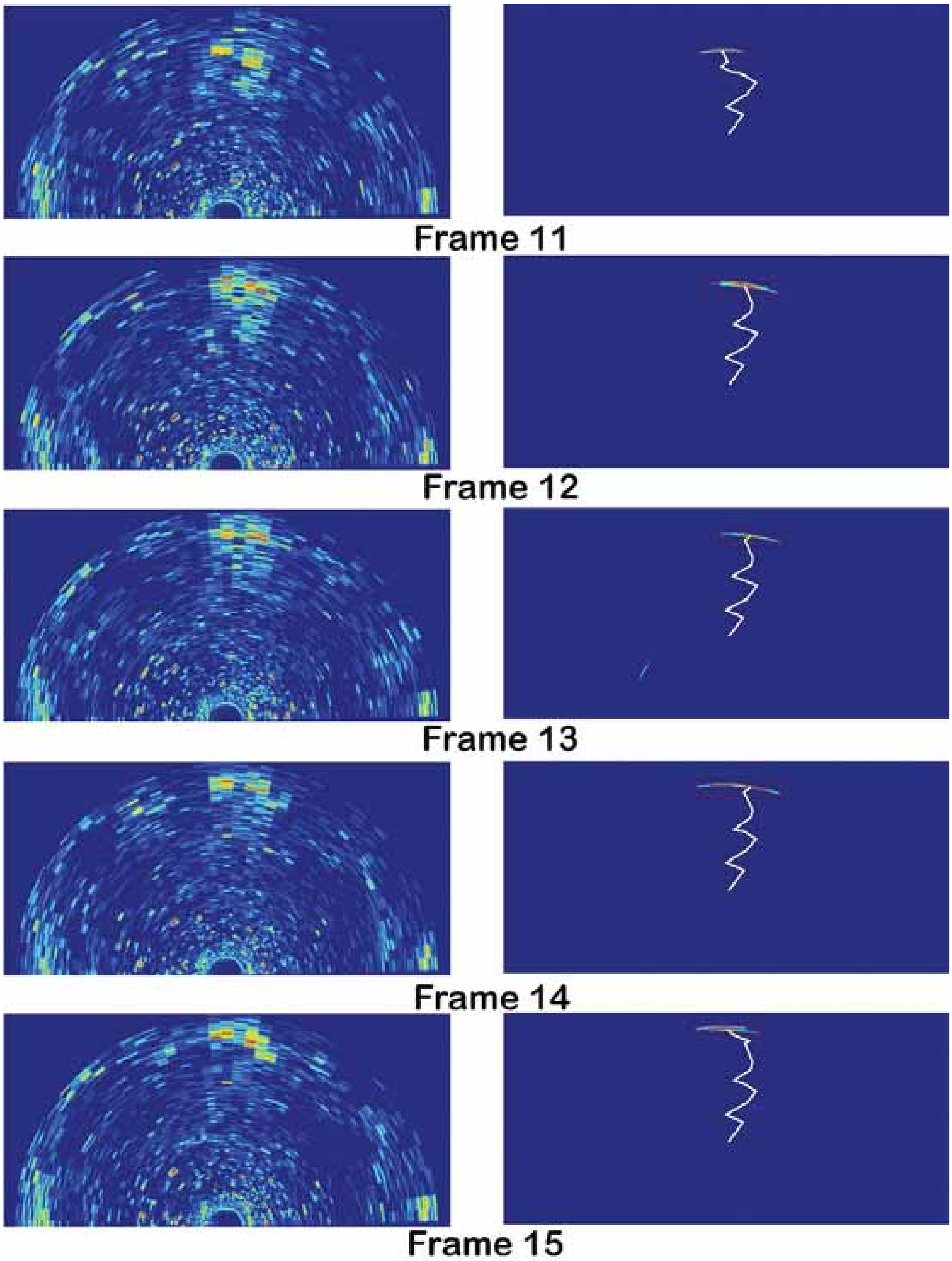

Fig. 5-15: Objects detection and tracking results using image processing techniques. From Frame 11 to Frame 15.

Figure 5-16 illustrates the object moving trajectory using NN tracker, which was described in Chapter 4.

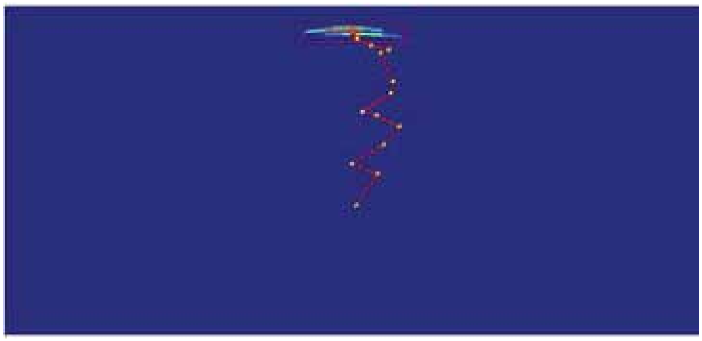

Fig. 5-16: Moving trajectory of human leg in the pond

With the assistance of *a priori* knowledge of the environment, the detector developed by image processing techniques provided a reliable result. However, it did not adapt to the uncertainty caused by the underwater environments. When the environment changes significantly, the detector leads to poor results.

5.3.2 River Torrens

For the second experiment, an area located near the River Torrens and adjacent to the Elder Park in central Adelaide, was selected. This area is a naturally occurring river bed with a man-made retaining wall located on the southern bank of the river and a man-made retaining wall with naturally occurring vegetation located on the northern bank (see Figure 5-17). The principal purpose for conducting this experiment was to test the reliability of the Fuzzy Logic decision-making system in natural environmental conditions.

For the initial testing and calibration of the sonar equipment it was decided to fix the sensor to the end of the boat ramp that is used to embark and disembark passengers, who cruise the River Torrens on the popular Popeye river boat. At this point the River Torrens is approximately 75 metres wide and varies in depth between 2-3 metres. On the left of the boat ramp, approximately 20 paddle boats are moored in a fixed line.

NOTE:
This figure is included on page 114
of the print copy of the thesis held in
the University of Adelaide Library.

Fig. 5-17: Experiment's River Torrens environment (image from Google map).

Figure 5-18 and Figure 5-19 illustrate the results of this test when the sonar sensor was mounted under the ramp. The original sonar image in Figure 5-18 was generated at 670 kHz with a scan step 0.9 degrees and the sonar scan range was set to 10 meters. According to sound beam pattern discussed in Chapter 3, it was inferred that the sound waves would travel in the water for about 5 meters (horizontal distance) before hitting the river bottom. Giving the vertical beam width (40 degrees), the water depth was calculated at approximately 1.82 m. The muddy and rough river bottom reflected the sound waves, and reverberation energy represented as a dark blue colour in the sonar image. It should be noted that in order to attach the sonar sensor to the ramp, it was necessary to invert the sensor, and as a result the images projected appeared in opposite order, that is to say left appears right and right appears left.

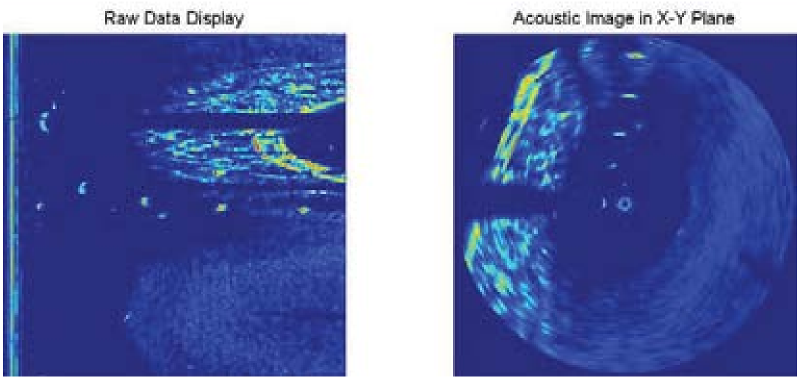

Fig. 5-18: Sonar raw data and acoustic image obtained from the River Torrens

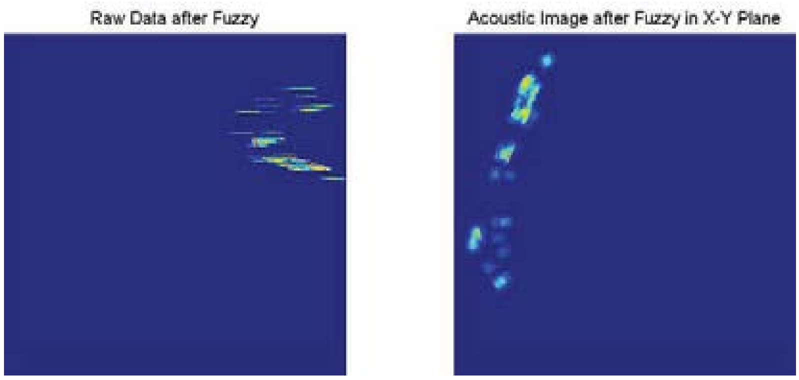

Fig. 5-19: Sonar raw data and acoustic image after applying fuzzy filter

After the initial testing, it was evident that both the location and equipment were suitable to complete the experiment. The sensor head was fixed to the front left-hand side of the paddle boat at approximately 1.4 metres depth, the boat being constructed primarily of fibreglass. In order to truly represent the AUV operational scenario, the experiment was conducted whilst the paddle boat was moving at approximately 1m per second. To conduct the second phase of the experiment, a total of 3 paddle boats were used: Paddle boat A housed the sonar while Paddle boats B and C each had specific objects that were submersed and secured at the same depths as the sonar sensor. Two types of objects were used (see Figure 5-20), the first being a metal cuboid object measuring 18cm × 40cm × 40.7cm (width × length × height) and the second a metal cylinder tube object measuring 20.5cm × 0.2cm × 50cm (diameter × thick × height). Both objects were attached to buoys and controlled by the operators on the boats (see Figure 5-21). Initially, all three paddle boats were driven to the middle of the river. Boat A was centralised and boats B and C were placed in orbit around Boat A at approximately 5 to 8 metres. During the sampling, Boats B and C continued to orbit Boat A and at predetermined stages were individually removed from the area being scanned. Boat A also travelled along the river during the experiment.

Another very important factor was the very clear detection of the river bank while the boat was in motion. In the acoustic images, very noticeable interferences came mainly from the river bottom. Since no other navigation sensor was presented in the experiment, the speed of boat A was controlled under 1m/s. The main purpose of velocity control is to reduce image distortion caused by the motion and to test the NN tracking algorithm.

Fig. 5-20: Target objects for the experiment

Fig. 5-21: Sampling in the middle of the River Torrens

Figure **5-22** to Figure **5-27** show the data flow with thirty image frames and the detection results of the two moving objects. Figure **5-28** shows the tracking results using NNSF, however, the accuracy of trajectories of the target objects are not verified because the accurate motion of the sonar sensor was unavailable.

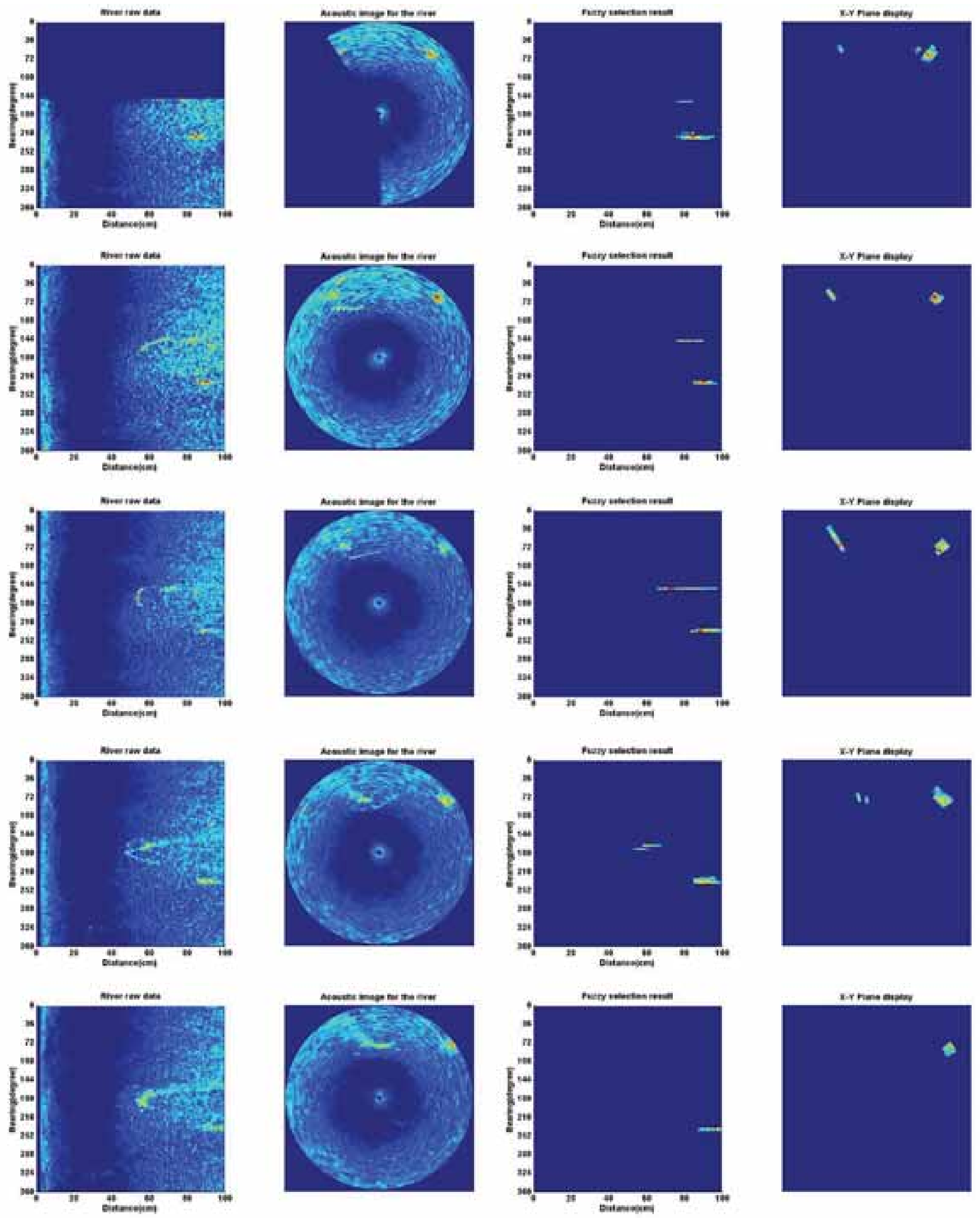

Fig. 5-22: Two moving target objects detection results.

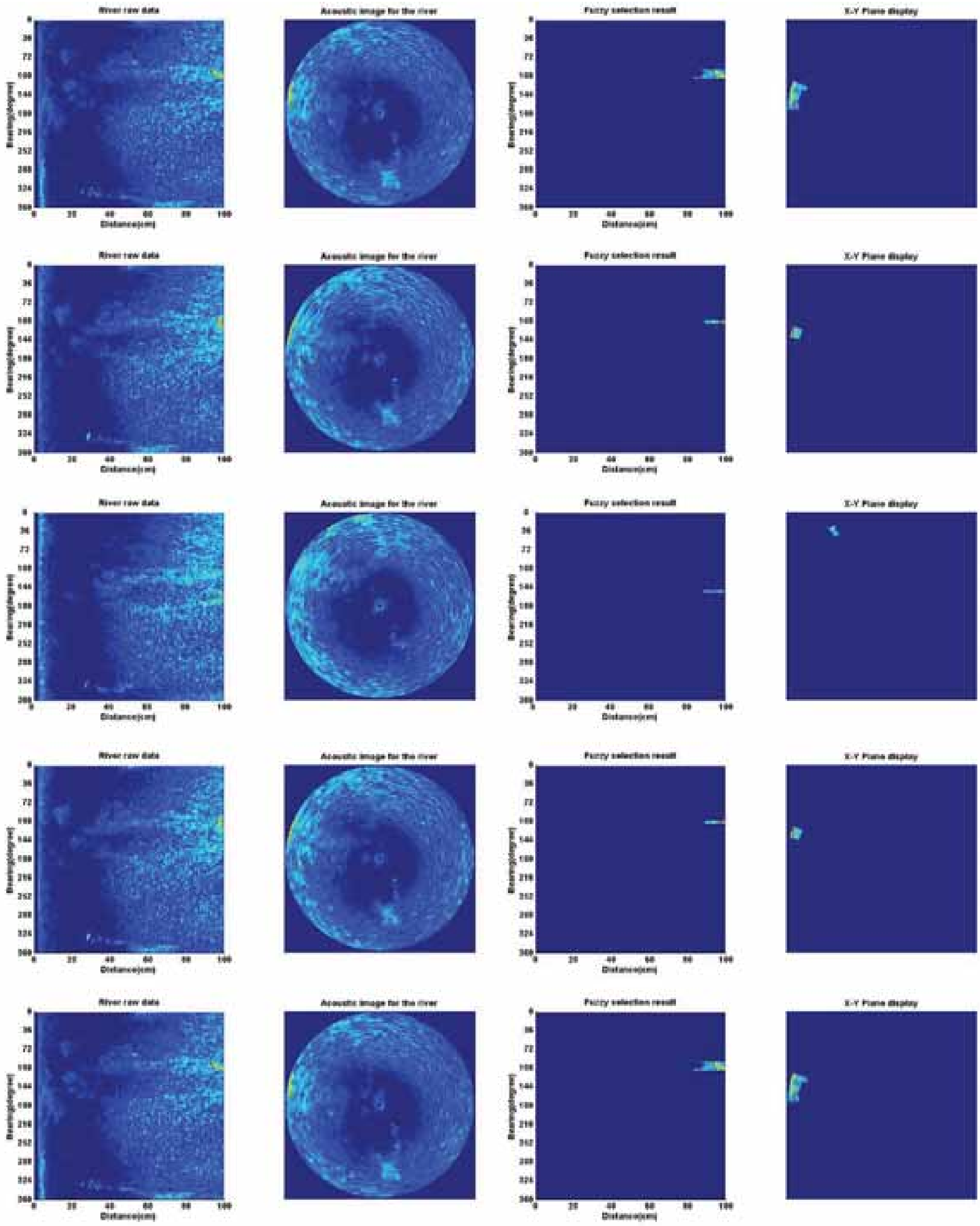

Fig. 5-23: Two moving target objects detection results.

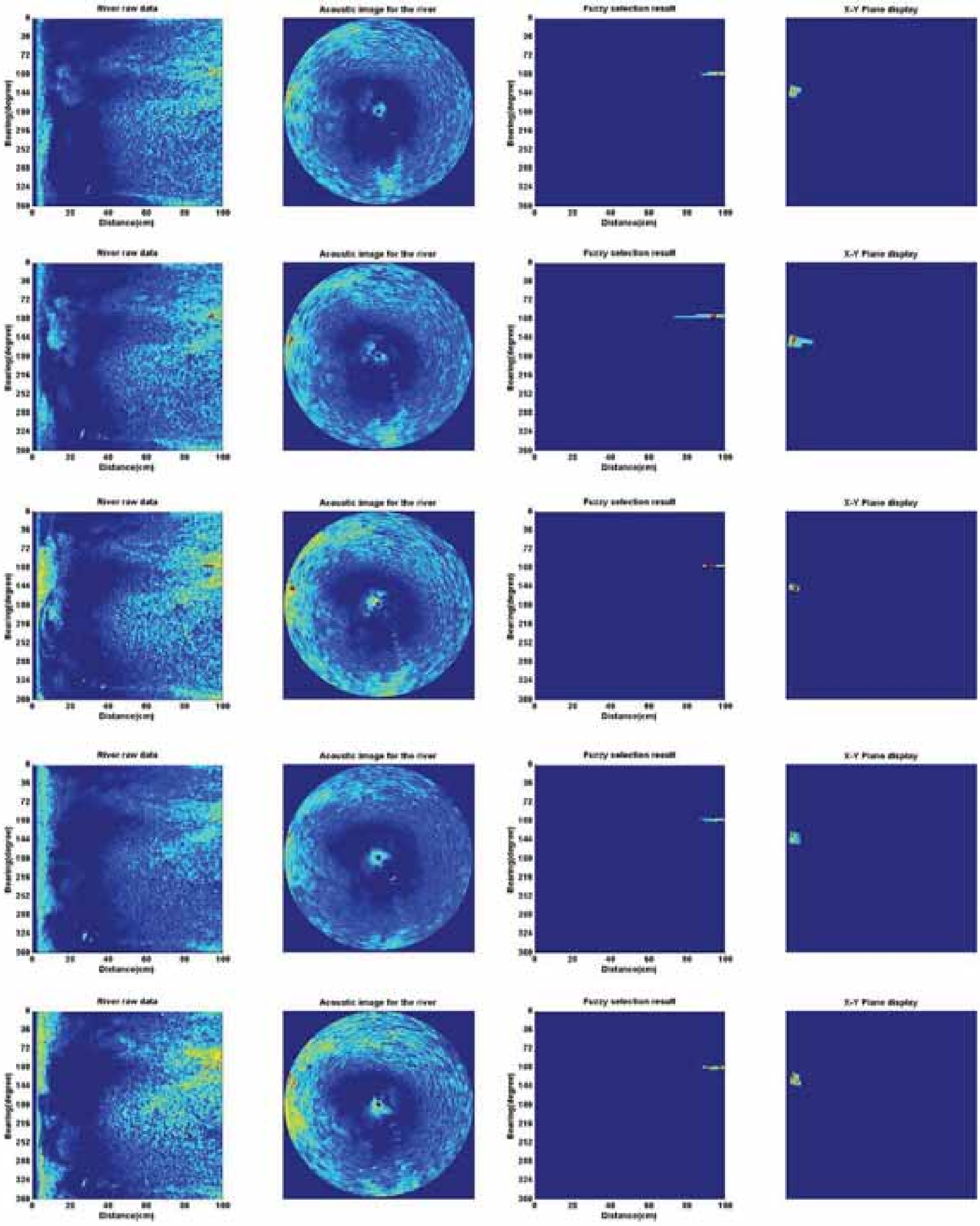

Fig. 5-24 : Two moving target objects detection results.

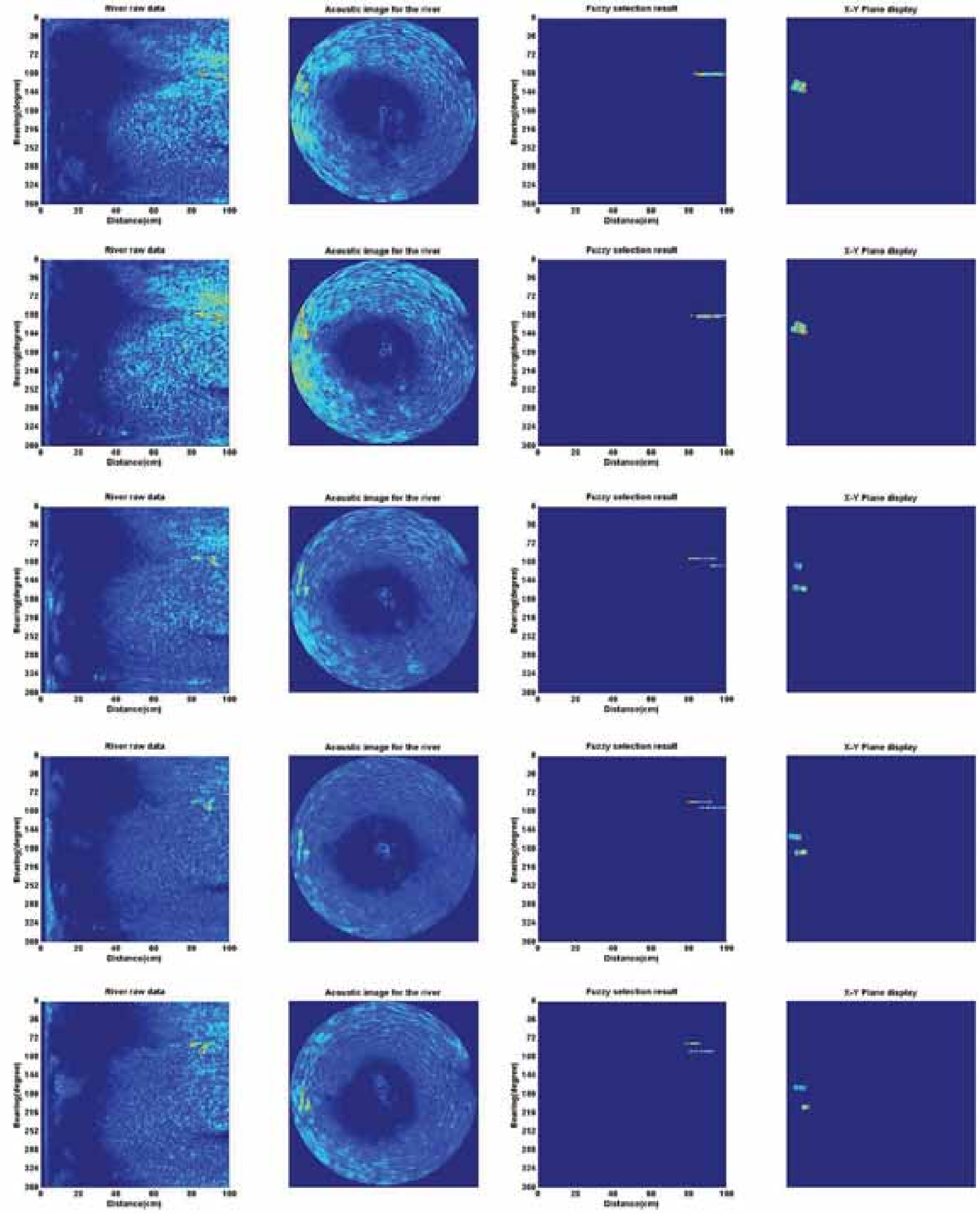

Fig. 5-25: Two moving target objects detection results.

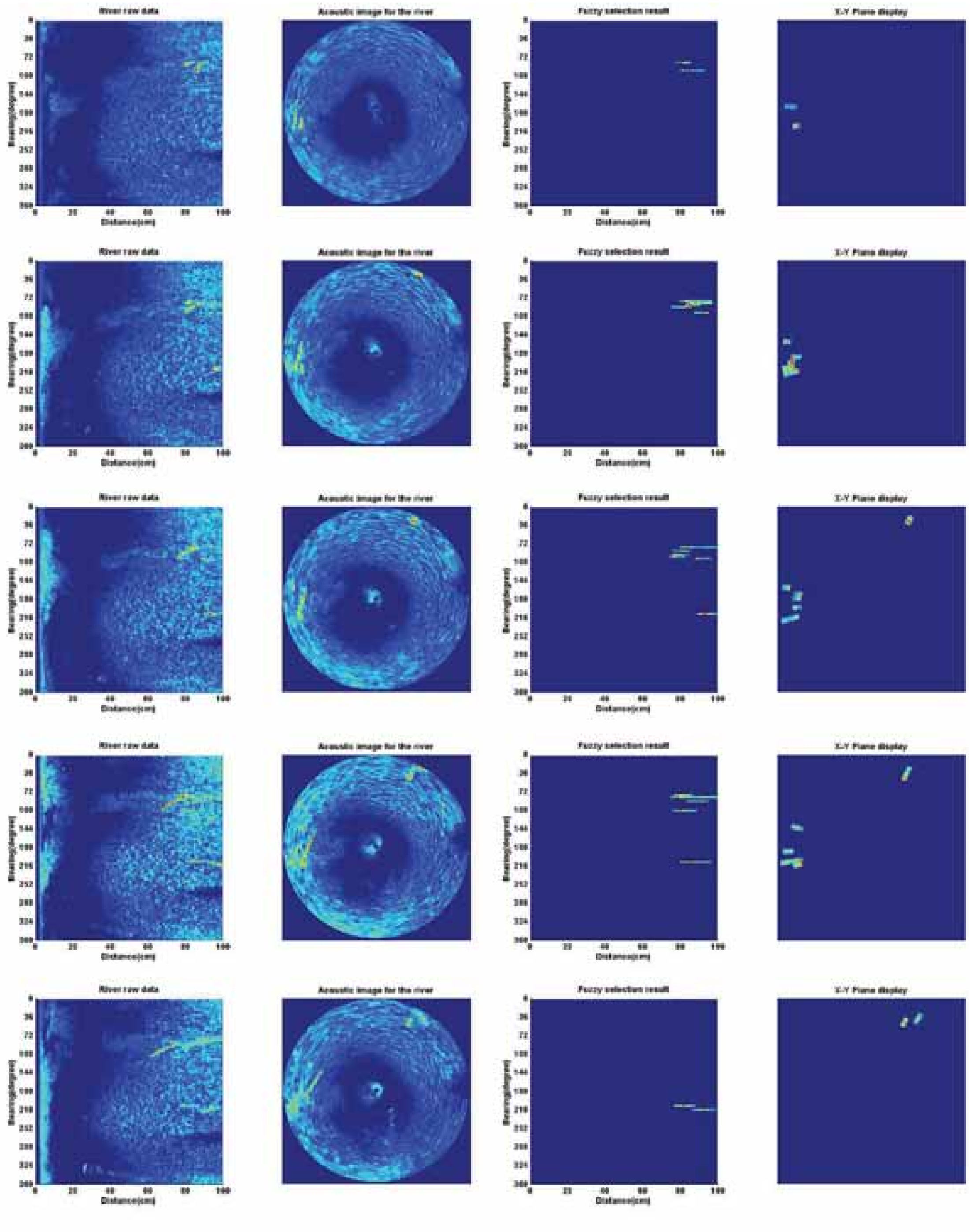

Fig. 5-26 : Two moving target objects detection results.

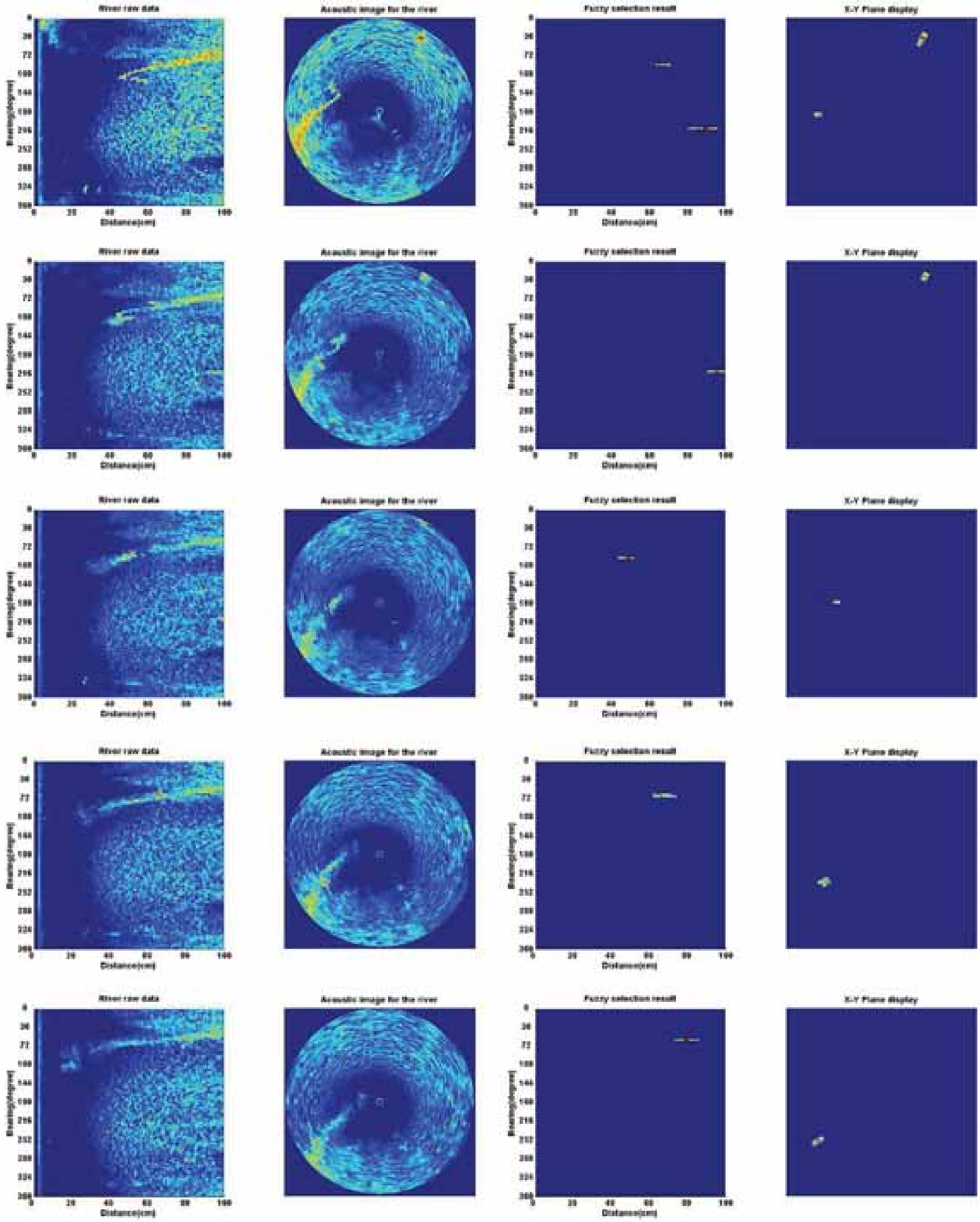

Fig. 5-27: Two moving target objects detection results.

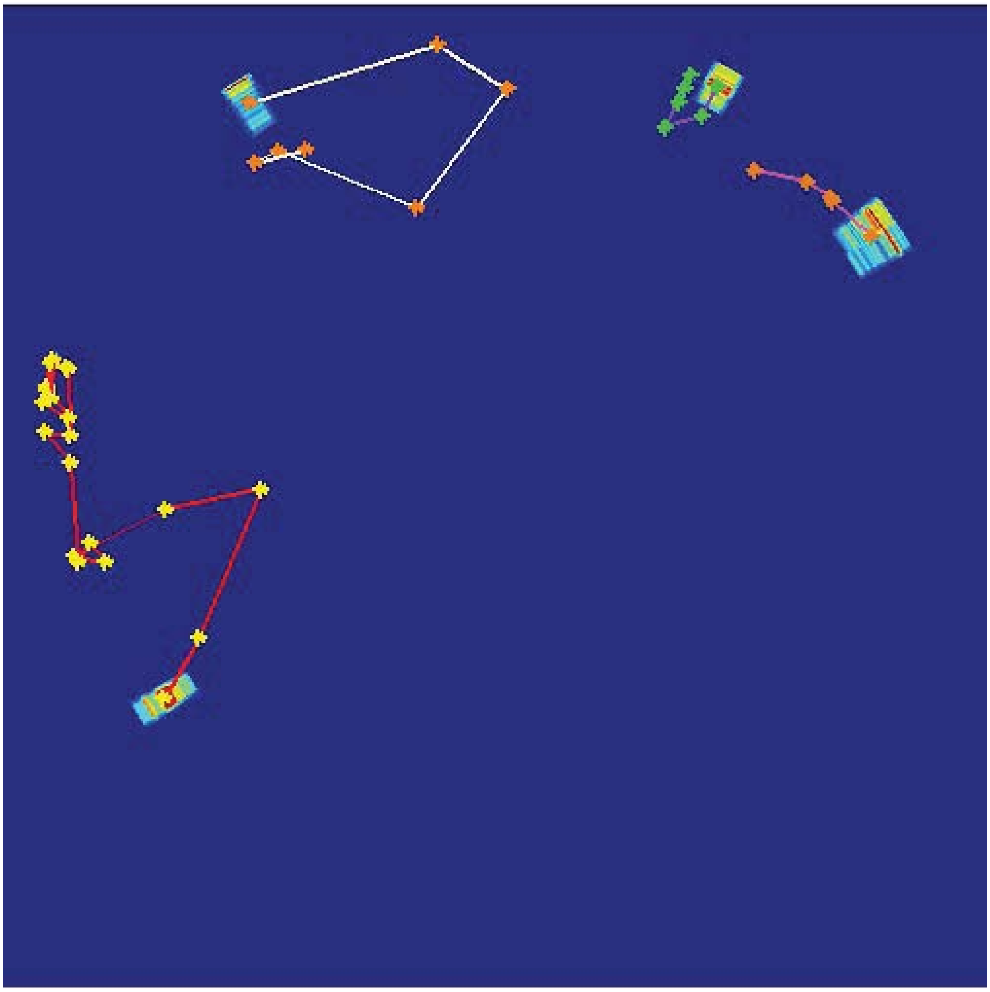

Fig. 5-28: Objects trajectories in the whole sonar scan. Because the target objects appeared out of sonar scan rage, four different trajectories were recorded.

Figure 5-29 to Figure 5-32 show the detection results for a section of natural river bank and one moving object near the bank in ten sonar image frames. However, since the fuzzy detector is performed for each single ping only, it is inadequate for describing such a large object (river bank). Therefore, the image processing techniques presented in Chapter 4 were implemented. Detection results are shown in Figure 5-31 and Figure 5-32.

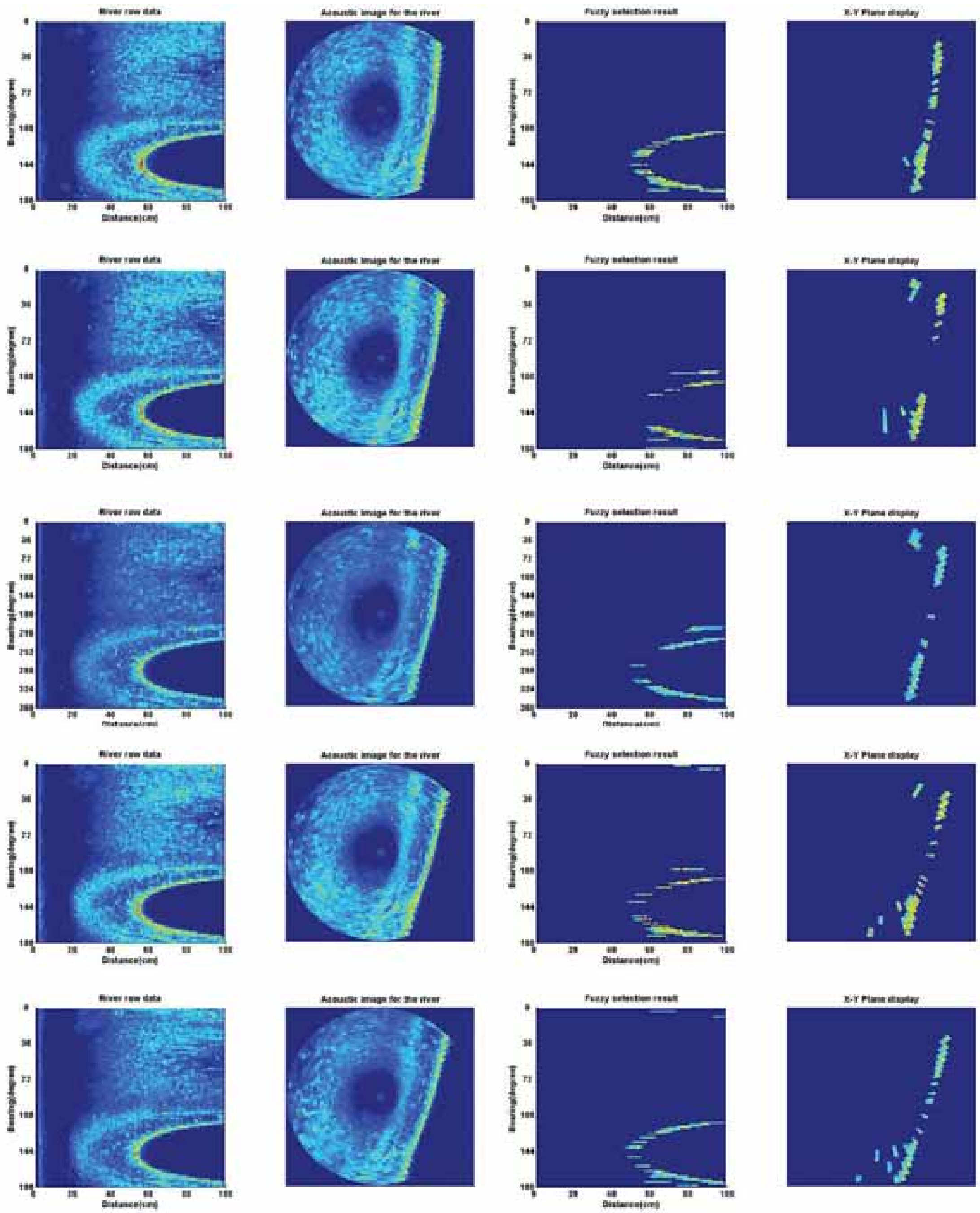

Fig. 5-29: River bank detection results.

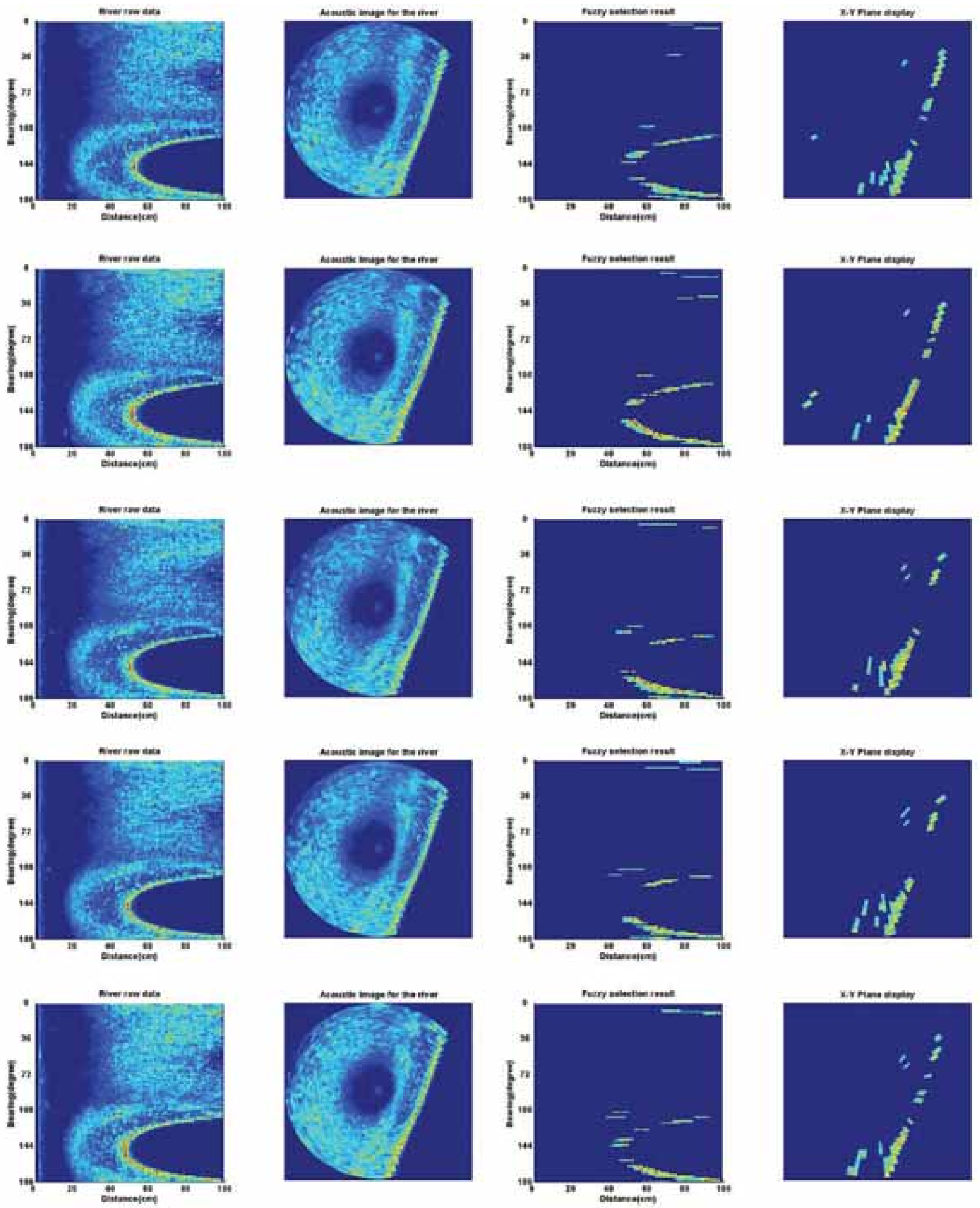

Fig. 5-30 : River bank detection results.

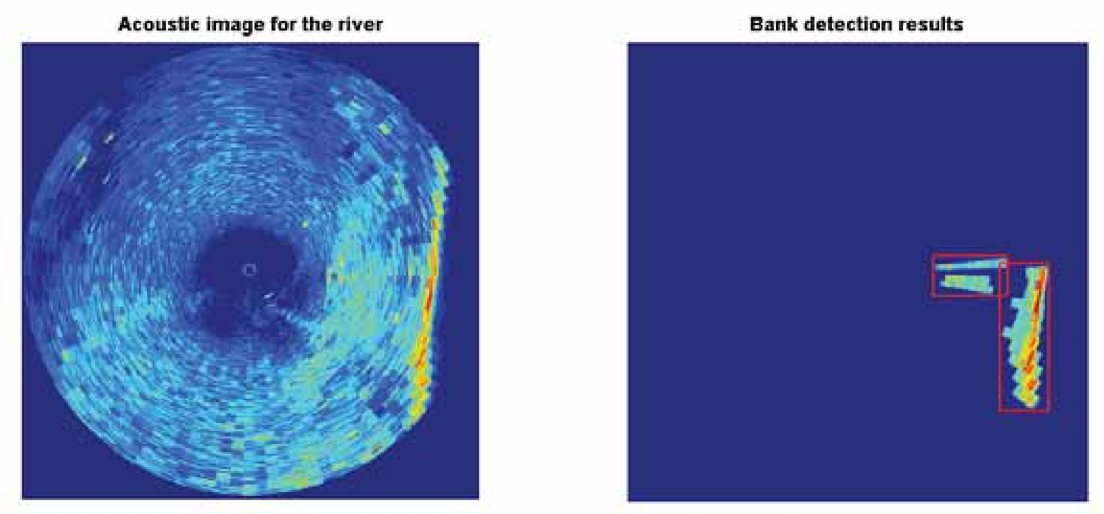

Fig. 5-31 : River bank detection results

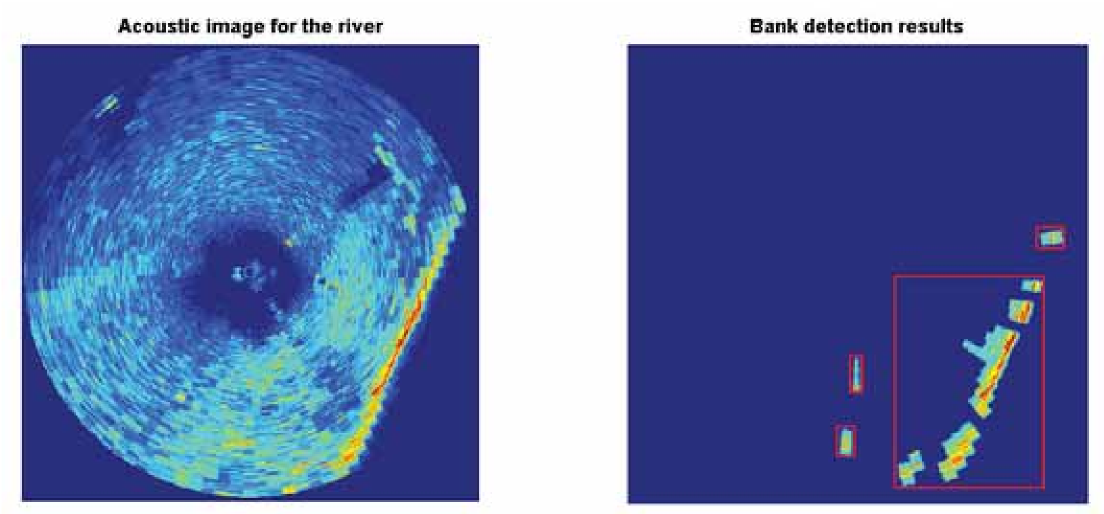

Fig. 5-32 : River bank detection results.

The results prove that the fuzzy logic based detector is able to continuously identify underwater objects. Four missing detections and two false alarms were reported in the above experimental data. Missing detection occurred when target objects were nearly out of the sonar scan range. However, since the sonar sensor operated in a natural river, it was hard to tell that false alarms were caused by the system itself or they were real object echoes from underwater creatures.

5.4 Discussion and concluding remarks

This chapter presented a fuzzy logic-based estimation process for underwater objects detection in natural environments. Although the image processing techniques described in the previous chapter can provide acceptable results with a *priori* knowledge, much useful information regarding the objects had to be discarded. In comparison, the fuzzy logical detector has proven to be robust and reliable according to experimental results. Over 115 image frames were taken in the shallow natural river, and only five false alarms and six missing detections were reported. The fuzzy logic detector achieved satisfactory results in coping with the nonlinear reverberation noise. Also, since this method no longer needs presetting of the sonar control parameters or transferring of raw data into a Cartesian plane, it improves the system processing speed. The locations of small objects can be detected directly from the raw data through a polar coordinate system. For some large objects, such as a river bank, the detected results can be reinforced by the image processing algorithms presented in Chapter 4. This hybrid detector will enhance and benefit the real time applications for AUV navigation. Since the distortion of the acoustic image is small when the boat moves at low speed, the nearest neighbour (NN) tracking algorithms can also provide viable, assessable and reliable results. However, it is relatively weak when objects move out of the sonar scan range momentarily and then are rescanned by the sonar again. A new trajectory will be generated when objects are re-detected by the sonar in such situation.

The independent detector using fuzzy logic presented in this chapter performed within expectation in a natural river. Objects scanned by the sonar can be identified from the sensorial raw data directly. This facility reduces the computational time, thus helping the real time performance for the AUV's navigation.

CHAPTER 6

Conclusions and Future work

This chapter concludes the work presented in this thesis. Firstly, it reviews the contents presented in each chapter and summaries the developed algorithms to a complete system for the AUV. Then the contributions of this research are described. Finally, future possible research directions are outlined.

6.1 Summary

Navigation is of paramount importance for a truly autonomous vehicle. Detecting and tracking potential obstacles which may threaten underwater vehicles is considered in this research. In Chapter 2, a brief overview of the literature and recent approaches were presented. The applications focused on ESIS and MSIS were investigated. The review finally led to the objectives of this study. Chapter 3 concerned the principle mechanisms of underwater sound and sonar imaging. The criteria of sensor selection were also described. It concluded a simple experiment, which was designed for examining the ability of the mechanically scanned imaging sonar (MSIS) and providing a reference data for the future target object selection. Chapters 4 and 5 presented two different approaches according to two distinctive underwater domains and proved the validity of the approaches with real sensor data. Chapter 4 firstly described the reverberation suppression filter for confined environments based on the understanding of sound propagation in such structured environments. It secondly investigated the detection and tracking algorithms. The multiple range approach successfully and accurately detected small static and moving objects, which alleviates the system requirements for the tracking. Meanwhile, the trajectories of the target objects were recorded by nearest neighbour tracking method. Furthermore, the enclose environment was also detected and a mathematical map was built by least square curve fitting. Finally, designed experiments carried out in an elliptic test tank in chapter 4 demonstrated a conclusive result of the detector developed by image processing techniques. Chapter 5 provided a more intelligent approach to filter out the coloured reverberation in very shallow water (about 2 to 3 meters). Firstly, it described a detection method according to *a priori* knowledge using image processing techniques. Secondly, it explained the reason for the selection of fuzzy logic for automatic detection in natural shallow water. Thirdly, a close inspection of the single ping was taken and fuzzy logic was applied to the detector. The fuzzy logic detector in Chapter 5 originated from the reverberation filter. However, it proved to be more powerful and robust than its predecessor. Finally, in a similar vein, experimental data collected in the River Torrens validated the fuzzy approach. The developed navigation system can be explained by the flow chart in Figure **6-1**.

6.2 Contributions

Reverberation and ambient noise are two major issues which interfere with measurements of active sonar in a very shallow water environment. The similarity between small obstacles and weak reverberation in structured environments causes false alarms and miss detections which may lower the efficiency or threatens the safety of the AUV in an exploration. On the other hand, natural shallow water environments are always filled with rocks and coral on the muddy bottom. The bottom reverberation from these surfaces is much stronger than small object echoes. It is a challenge problem to filter out the non stationary and coloured reverberation noise and remain the object information unaltered. Different and stand out from other approaches in the literature, the thesis examined object echoes in a natural way with a mechanically scan image sonar (MSIS) and provided a series of solutions for automatically detecting small obstacles. Furthermore, to simulate the dangerous explosive ordnances, obstacles introduced in the study were all in the centimetre scales. The selection of such small size obstacles and the algorithm for detecting such small size obstacles using a low resolution sonar senor are very rare throughout the literature survey.

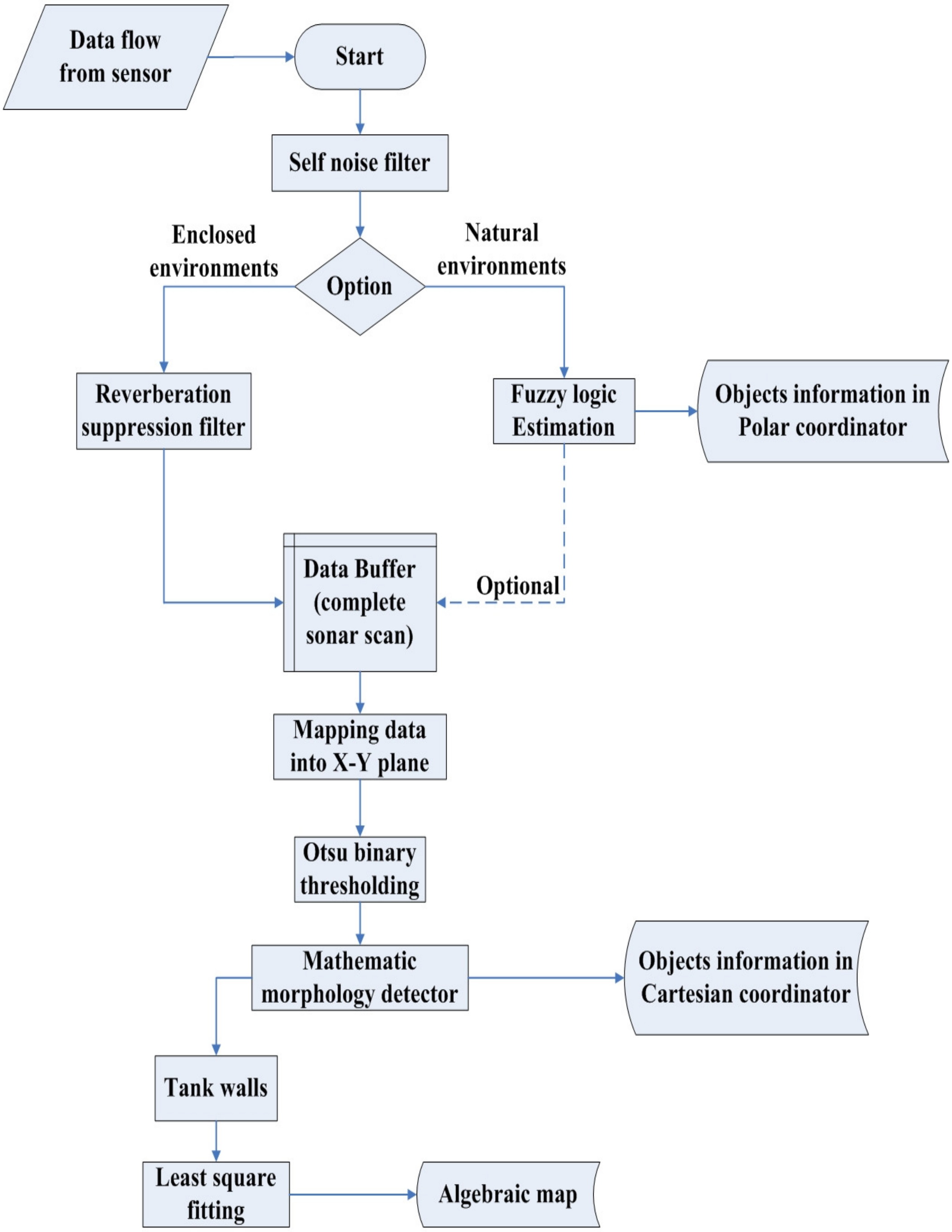

Fig. 6-1 : Flow chart for object detection using imaging sonar.

The work presented in this thesis achieved the proposed goal for the AUV's navigation contributes the following achievements:

***Reverberation suppression filter*:** Even weak reverberation energy picked up by the sonar hydrophone will affect the detection of the small underwater objects. The designed filter reserves the useful information and eliminates interferences for further processing.

***Fuzzy logic detector*:** Although it is inherited from the reverberation suppression filter, the fuzzy logic estimation performs robustly and reliably in dealing with the strong reverberation in very shallow water environments. While the sonar sensor operates in very shallow water, the noisy background will encompass the real obstacles. The use of traditional image processing techniques may exclude valuable information for navigation. Thanks to the fuzzy logical technique, the estimation process appraises the echoes and identifies the real object in each single ping. Instead of reproducing acoustic image in a Cartesian system, it can fully support the navigation and obstacles avoidance missions independently.

***Real-time processing*:** Various approaches were published for automatic detection using imaging sonar. However, to process the sensorial data in real time, a common method is to extract point features of object echoes from acoustic image by identifying the peak intensity values. The actual size and shape of the obstacles are always missing using this method. This study dedicated to an architecture of real-time processing, the developed algorithms were strived to concise and effective. The developed detector for confined environment has proved to run in real time. The sizes, shapes and the distances information of obstacles can be updated in every 4 seconds (determined by sonar scan period). This result was achieved using a P4 2.4GHz, 1GB RAM computer when sonar scan range was 6 meters.

***Feature extraction for pre-known structured environment*:** The method for extracting the features from the acoustic image is another contribution of this study. The presented method successfully processes the continued data flow from the sonar sensor with a high level of accuracy. It provides opportunities for the simplification of the object tracking algorithm. Moreover, the inner walls of the enclosed surroundings are also extracted. The results directly lead to the final mathematical mapping algorithm.

6.3 Further research

In this dissertation, a mechanically scanned imaging sonar data processing system is interpreted. It raises several interesting extensions of the system and algorithms in both theoretical and practical performances.

***Sensor Data fusion*:** The imaging sonar system acts like human eyes to perceive the surrounding objects for the vehicle. To build a true autonomous underwater vehicle, it should cooperate with other sensory systems, such as DVL (Doppler Velocity Log) and gyro to estimate a vehicle's pose and underwater acoustic modem for communication. The observations from different sensor obtained at different times contribute to the perception of the external environment for the AUV. Data fusion is an estimation process which combines the observations into a coherent description of the environment. The motion of the vehicle will distort the acoustic image when the sensor-scanning speed is relatively slow in comparison to the speed of the vehicle. Such distortion may lead to poor estimation results of the objects. In such situations, the displacements and rotations (pose of the vehicle) can be incorporated to correct any distorted acoustic images. Another extension for the data fusion is the development of the localization system. The current algorithm is sufficient while the vehicle is static or moves at low speed (less than 1.5 m/s). Although the proposed localization algorithm has shown promising results, more work needs to be done to successfully localize the vehicle in real time.

Moving objects tracking: The traditional nearest neighbourhood tracking is a simple but compelling algorithm. It exhibits adequate reliability in the discrete-time data flow while the detection results are accurate. However, the application of the nearest neighbour in an environment where spurious detections occur frequently will lead to poor results. This is mainly because one important step is missing: associating the current measurement with the previous one. That is to say, the measurement used for tracking might have originated from a source different with the target of interest.

Intelligent control system: Intelligent control system is the brain of the autonomous vehicle. Measurements from the sensors will integrate and process the system to command and control the vehicle's motion. Estimations from the sonar image processing system are mainly used to avoid obstacles in front of the vehicle. The algorithm for collision avoidance and an efficient navigation system is another important issue for future development.

AUV operational scenarios: The whole project serves the purpose of the underwater vehicle protecting ports and being used for harbour surveillance missions. The improved and newly developed algorithms need to be tested in the ocean, where new problems may be generated, for example the presence of moving boats in a harbour. The wake caused by the engine will be detected by the sonar sensor.

Enhanced ability to gain obstacle information in real time and to achieve guided AUV navigation contributes much of value to the field of autonomous underwater investigation.