Indian Journal of Geo-Marine Sciences
Vol. XX(X), XXX 2015, pp. XXX-XXX

Automatic target detection of sonar images using multi-modal threshold and connected component theory

Subhra Kanti Das[1]*, Soma Banerjee[1], Dibyendu Pal[1], Sambhunath Nandy[1], Sankar Nath Shome[1] & Somnath Mukherjee[2]

[1]Robotics & Automation, CSIR-CMERI, Durgapur-713209, India

[2]Computer Vision & Design Engineer, Kritikal Solutions Pvt. Ltd., India

*[E-mail: subhrakanti.das82@gmail.com]

Received 30 September2014; revised 28 November 2014

The aim of this paper is to present a complete progressive development of object detection from underwater acoustic images. Object detection with respect to automatic target detection in underwater autonomous vehicle system is still in a severe problem in context of surveillance and other defense activity. The present work is based on robust method in perspective of segmentation and feature extraction. Underwater acoustic images suffer from typical noise associations and are often of low contrast. In this perspective, a multi-modal thresholding is adopted for automatic segmentation of the images thus obtained and a graph theoretic approach based on connected components is formulated in order to interpret features embedded within the image context. An imaging SONAR is used for carrying out necessary experimental work. The proposed algorithm is executed in comparison with multi-level thresholding and *K*-means clustering. Effectiveness is established in the context of both running time and quality of processed image as well. The latter aspect is determined by a Figure of Merit (FOM) parameter.

[**Keywords:** SONAR, Automatic target detection, thresholding, features-detection, acoustic image]

Introduction

Automatic Target Detection is now becoming a routine process of underwater imaging using SONAR for multipurpose applications which completely belong to a subset of underwater object tracking. In this paper, it has been restricted for acoustic images using SONAR as the important aid for path planning as regards underwater autonomous vehicular systems, which further lies in the effective retrieval of information from SONAR images. Acoustic images as supplied by SONARs are associated with noise typically emanating from multipath and specular reflections especially in shallow waters or confined water bodies like basins and pools. Consequently acoustic images are often distorted and do not however completely embody the objects or features being detected. As a result feature extraction from such noisy images becomes a real issue. Moreover underwater features are often unstructured like abrupt continental rise or dynamic like the moving keel of big ships as well as aquatic animals. In such cases no suitable model based feature extraction or pattern matching can be committed to the available source images in order to identify and determine the position and other information about the detected features.

A bulk of work has been reported in literature with a common characteristic. Most of the available methods center around the rich class of image processing techniques including spatial as well as temporal filtering of acquired SONAR images followed by segmentation at multiple levels and finally identifying the hotspots present in the segmented image, to determine the position as well as size of the detected objects. However, a major aspect in detection of objects lies in the fact that not all such segmented hotspots should be of concern to the vehicle. There often remain small isolate and speckled regions which might not ever be any object of potential collision with the concerned AUV. In the present scope of this paper, an attempt has been made to present a complete progressive development of object detection from underwater acoustic images. The present work is based on robust method in perspective of segmentation and feature extraction. Underwater acoustic images suffer from typical noise associations and are often of low contrast. In this perspective, a multi-modal thresholding is adopted for automatic segmentation of the images thus obtained and a graph theoretic approach based on connected components is formulated in order to interpret features embedded within the image context. An imaging SONAR is used for carrying out necessary experimental work.

An automatic target identification method has been introduced[1]. Emphasis is laid on supervised learning techniques in order to make the system learn about target morphologies and ultimately pattern recognition is applied in order to identify the target from the known template. The work can be extended to both SONAR as well as RADAR images. In Ref. 2[2] the authors have conducted detection of moving objects using sector-scan SONAR with an underlying objective to track observations. Optical flow estimation is used in order to estimate the dynamic state of moving objects. These are utilized in the prediction of future detections which are actually matched with subsequent observations. A tracking tree method has been devised for storing multiple correspondences. However, the effectiveness of the method is restricted to platforms with significant computational capabilities for example ROVs and underwater surveillance systems with human intervention. In Ref. 3[3] a method has been introduced as yet another obstacle avoidance framework for the Advanced ROV Package for Automatic Mobile Investigation of Sediments tool-skid. Segmentation of the acoustic images is carried out in phases with filtering and single value thresholding followed by feature extraction and tracking. Kalman filter has been adopted for estimation of the dynamic characteristics of the objects. While planning algorithms are devised both for static as well as dynamic obstacles, scope for being used in autonomous underwater vehicles with limited computational power is yet not explored. In Ref. 4[4] the authors have achieved 3D information extraction from a sequence of SONAR images. A forward looking sector scanning SONAR has been employed for the purpose, which looks down upon the seafloor. Ultimately 3D information is extracted by comparing the various positions of features in the sequence of 2D images collected by the SONAR. Contrastingly a parameter based segmentation technique for SONAR images has been proposed in the Ref. 5[5]. A Markov Random Field model is proposed for the image and parameters for the model are estimated using Bayesian probability. Segmentation is applied to a structured side scan SONAR image with a prior model for the image. The method is effective in segmenting the image by considering image model, shadow as well as reverberation. A new method has been proposed in Ref. 6[6]as regards the unsupervised method of segmentation for seafloor images acquired by a side scan SONAR. Region-based active contours method has been employed for the purpose with an assumption that, the underlying image consists of two distinctively divided regions. A 3D volume growing segmentation has been proposed for the realistic rendering of seafloor objects in the Ref. 7[7].Selective feature extraction from real-time SONAR images acquired from a COTS multiband SONAR has been proposed in the work summarized in the Ref. 8[8] towards collision avoidance for the Meredith AUV. Extraction of selective features leads to a decisive state wherein potential obstacles are detected. The information thus obtained is subsequently used in a 2D grid based collision avoidance framework. Despite discussion on effectiveness of the overall segmentation, the feature extraction method adopted is not reported in sufficient details. In Ref. 9[9]the authors have carried out underwater SLAM for an AUV using forward looking SONAR. A region aggregation and expansion method has been adopted for feature extraction from SONAR images. However, a fixed threshold intensity value has been regarded for isolation of the features from the background. Symbolic analysis using pattern recognition for identifying seafloor features targeted at mine inspection has been carried out in the Ref. 10[10]. In the referred scope, features have been classified into two distinct categories viz. mine-like and non-mine-like objects. Further work regarding key-point based texture recognition of side scan SONAR images for seafloor has been carried out[11]. Nevertheless an interesting object detection algorithm has been proposed in the Ref. 12[12] wherein the covariance in intensity distribution is considered for the object to be tracked with respect to the undisturbed background. A threshold value is chosen for the ultimate segmentation of the image

The paper is organized as follows: Section II introduces the methodology adopted in segmenting acoustic images and detecting potential features. Section III discusses the results as obtained from acoustic images collected by physical SONAR. The proposed algorithm is executed in comparison with multi-level thresholding and *K*-means clustering. Effectiveness is established in the context of both running time and quality of processed image as well. The latter aspect is determined by a Figure of Merit (FoM) parameter. The proposed method is shown to outperform the other reference methods in both the aspects. Conclusion is presented in section IV.

Materials and Methods

The present section provides an overview for region-wise multimodal thresholding and connected component based feature extraction from SONAR images. The stage wise activities may be summarized as: filtering, thresholding and feature extraction using a connected component approach drawing inspiration from graph theory. Fig. 1 is illustrative of the control flow in the algorithm adopted for extraction of information from acoustic images.

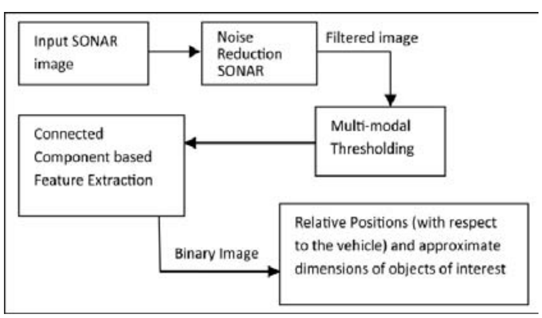

Fig. 1–Schematic outline of the overall algorithm proposed for the effective retrieval of information from SONAR images.

Segmentation is necessarily carried out with the basic objective in identifying the areas of interest present in the image. The usual segmentation procedure for SONAR images involves denoising the image followed by thresholding for conversion to binary image. SONAR images are attributed with acoustic noise typically arising from backscatter as well as specular reflections (which may usually happen as the acoustic echo signal suffers from multiple reflections). Consequently, the image needs to be filtered prior to any information extraction process. In the present context, median filtering is used for its potentiality in preserving the details of the image. Moreover, median filter is an effective method that can suppress isolated noise without blurring sharp edges. The set of processes involved in the proposed framework may well be summarized as follows:

Filtering

A 5×5 median filter has been adopted for noise reduction the SONAR image. The filter works effectively in maintaining the details of the image. There are several approaches for more robustness in these sections to produce a best linear and nonlinear classical noise reduction filtering process. In Ref. 13[13] an improved sigma filter have been proposed for speckle noise reduction that has been further improved in the latter work[14] where anisotropic diffusion have been incorporated for the best result estimation.

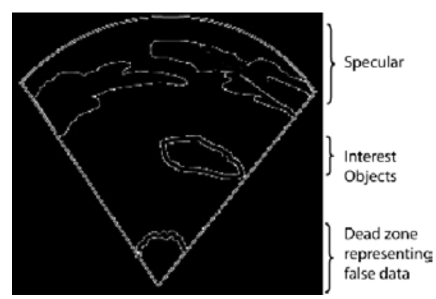

Fig. 2–a typical morphological classification of SONAR images consisting of relevant and irrelevant information; specular reflections usually appear towards the edges of the scan; dead zone represents the minimum scan range below which the SONAR fails to detect objects; the dead zone usually is of high intensity due to multiple interference of acoustic pressure signals at the source end.

Thresholding

A single, fixed threshold, generally gives results which are highly dependent on the background level. An automatic bimodal histogram based threshold technique have been incorporated[15] based on improve Otsu method. But there are no such interpretations whether foreground object is much closer to the back ground. In the present context, a region-wise thresholding method has been formulated using multiple threshold intensities chosen from a multi-modal histogram of the filtered image at a particular sampling instant.

Feature Extraction

The binary image obtained after threshold consists of few scattered white pixels, on a black background. Considering the general morphological classification of a typical SONAR image as illustrated in Fig. 2, the objects of interest can be effectively determined by clustering the white pixels among themselves and heuristically selecting large clusters to be objects of interest. The process of feature extraction proposed in the present context starts as a pixel aggregation technique and draws inspiration from graph theory in that pixel aggregates are treated as graph nodes and patches are considered to be connected components in the graph. Thus the identification of objects relies on the effective determination of all such connected components of the graph representative image.

The subsequent sub-sections elaborate on the proposed approach towards retrieving ambient information from SONAR images.

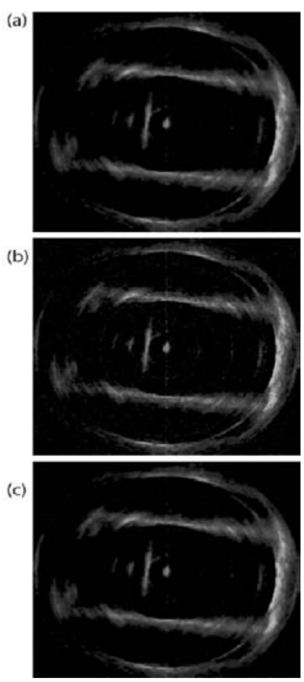

Fig. 3—(a) Original grayscale image; (b) Filtered image with 3×3 window; (c) filtered image with 5×5 window.

Noise Reduction of SONAR images

A typical median filter [$n \times n$], has been used for this purpose. Median filter is a spatial technique, which involves replacing the intensity value of the current pixel with the median value of intensities bore by the neighborhood pixels. The choice of neighborhood is critical to the performance of the filter in terms of noise elimination. For the present work, a 5×5 window has been selected.

$$g(x,y) = median[f(x,y)] \tag{1}$$

Linear classical filter[16]can be used for getting better and robust result in noise reduction where filtering process is depends on some specific constraint that can help where the smoothness is required and where not. It has been followed by Eq. 2 as:

$$g(x,y) = \lfloor f1(x,y) + k(f(x,y) - f1(x,y)) \rfloor \tag{2}$$

where $f1(x,y)$ is the real image, $f1(x,y)$ is the locally estimated averaging image, $g(x,y)$ is the filtered image and k is adaptive filtering coefficient.

When $k \rightarrow 1$, the filter image retains all the features and image pixels and when $k \rightarrow 0$ the filtered image will be the averaging of the original intensities.

Fig. 3 comparatively exhibits the quality of images obtained after filtering with 3×3 as well as 5×5 window respectively. It may be well observed that the details of the filtered image having a 3×3 neighborhood are quite comparable to that obtained with a 5×5 neighborhood.

Multimodal Threshold of Filtered image

Subsequently after the image has been filtered we are left with the purpose of segmentation, with distinct identification of hotspots present in the image in contrast to the background. Thresholding results in the generation of a binary image consisting of the hotspots delineated from the background in that the image consists of only two gray levels. In the present context, global thresholding [$n \times n$] is carried out using the spatial properties (i.e. pixel locations) of the image alone. This is because acoustic images are almost uniformly illuminated and hence the thresholding is not dependant on the properties imparted by a neighborhood to the pixels.

An adaptive thresholding technique dependant on the histogram has been adopted in comparison to the direct thresholding method. However, the histogram obtained for SONAR images can highly be multimodal which means having more than one particular gray level with the highest relative frequency of pixels. This leads to the immediate inference that appropriate thresholding cannot be done by considering a single gray level for the entire image which may lead to loss of details. Instead considering local threshold can provide better separation of the objects from the background. The proposition can be realized with a simulated histogram as shown in Fig. 4.

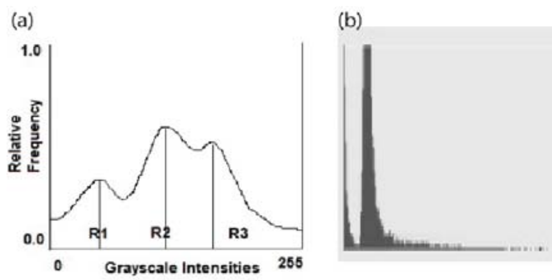

Fig. 4–(a) a typical multimodal histogram providing basis for selection of local thresholds over different regions in the grayscale space. R1, R2 and R3 are the grayscale regions demarcated by the peak intensities represented by solid lines on the histogram; (b) actual histogram obtained for the filtered SONAR image.

Fig. 4 illustrates a multimodal histogram having multiple intensities in the image with relatively high frequencies in different regions of the grayscale space i.e. 0-255. Depending upon the values of the Peak Intensities corresponding grayscale regions are demarcated. Regions are in between two peaks. The regions thus obtained are mapped onto the spatial intensity distribution of the image. Subsequently a pixel is transformed with a threshold using the intensity (*left peak* + *right peak*)/2 corresponding to the grayscale region to which the particular pixel belongs. As a result, the final binary image does not exhibit loss of significant details.

A more formal description of the adaptive multimodal histogram based thresholding approach may be provided as the following sequence:

1. Finding out region-wise threshold intensities corresponding to high frequencies in the overall histogram:

$$P = \left\{ p_i \mid \forall p_i, hist(p_i) > \frac{hist(p_i - 1) + hist(p_i + 1)}{2} \right\} \tag{3}$$

2. Determining regions in the gray level space (0-255) on the basis of the peak frequencies:

$$R = \left\{ r_i \mid \forall r_i, r_i \in \left\{ \left\lfloor \frac{P(i+1) - P(i)}{2} \right\rfloor, \left\lfloor \frac{P(i) - P(i-1)}{2} \right\rfloor \right\} \right\} \tag{4}$$

3. Applying region-wise threshold gray levels for the entire image:

$$b(x, y) = 0, \left(g(x, y) \in \{r_i\} \wedge g(x, y) \geq P(i) \right) \tag{5}$$

$$b(x, y) = 1, \left(g(x, y) \in \{r_i\} \wedge g(x, y) < P(i) \right) \tag{6}$$

A pseudo-code for the proposed algorithm is presented in Table I.

Connected component approach for feature detection

Thresholding generates a binary image having only two intensities 0 and 255. This ultimately leaves us with the major problem of identifying objects of interest and approximating their relative positions in the image. The binary image thus obtained from grayscale-regioning threshold, is tainted with discrete scattered white pixels against a black background. Such isolated pixels may be termed as "grains" in the present context. Contrastingly large number of pixels which almost remain adhered to each other may be termed as "patches". At this point it may be clearly understood that patches contribute more to the identification of objects of interest. Therefore, the identification of features may be carried out with the elimination of grains and preservation of patches. Subsequently identification of the significant patches can contribute to the detection of potential objects of interest. A connected component approach has been devised in this connection in order to identify features within the SONAR image. Connected component is a graph theoretic entity useful in understanding to what degree a given graph is connected. This means that with a notion of connected components within a graph, we can compute the node reachability. Every graph in itself is a connected component and that every graph has at least one connected component, i.e. itself.

Considering the image to be a graph, pixels can be assumed to be nodes. However, in the present context of segmentation a pixel shall be referred

Table 1 Pseudo Code for Thresholding with Grayscale-Regioning on Multimodal Histogram

```
Steps of the Proposed Algorithm :-
Input:  SONAR image f(x,y)
Apply Median Filtering on the image
              g(x,y)= median ( f(x,y) )
Hist[0:255] ← Compute the image histogram.
peakIntensities [0:127] ← array for storing the modes of the histogram
j:= 0
peakIntensities[0:127] := 0
FOR i := 1 : 254
    IF (hist[i]>hist[i-1] &&hist[i]>hist[i+1])
            peakIntensities[j++] := i
            peakIntensities[j] := 255;
END IF
FOR (all segments of histogram ranging from peakIntensities[i] :
    peakIntensities[i+1])
            threshold_value:= floor ((peakIntensities[i]+peakIntensities[i+1])/2)
            FOR (all pixels in the same segment)
                        IF (threshold_value<pixel_value)
pixel_value := peakintensity(i+1)
                        ELSE
                                                pixel_value := peakintensity(i)
                        END IF
    END FOR             END FOR END FOR
```

Table 1I Pseudo Code for Connected Component Approach for Feature Extraction

Create an array pointLocations[] with (Image_Height X Image_Width)/(window_size X window_size) elements. Assuming each index of the pointLocations[] array contains the co-ordinates of a point. Each point represents a bucket, consisting of windowSize X windowSize no of pixels.

```
Create an array bucket[] with (window_size X window_size) elements.
FOR i = 0 to Image_Height
   FOR j = 0 to Image_Width
      IF i>= window_size/2 or j >= window_size/2 or i< (Image_Height-(window_size/2)) or
 j < (Image_Width-(window_size/2)) THEN
        Y :=0
        count :=0
        FOR l = -(window_size/2) to +window_size/2
           FOR k = -(window_size/2) to +window_size/2
              IF finalArr[i+l][j+k] = 255 THEN
                 count := count+1
              END IF
           END FOR
        END FOR
        IF count >= (window_size X window_size)-2 THEN
           FOR l = -window_size/2 to +window_size/2
             FOR k = -window_size/2 to   +window_size/2
finalArr[i+l][j+k] = 255
              END FOR
           END FOR
           Insert an entry in pointLocations[] array for the current bucket & mark this bucket as unvisited.
        END IF
```

```
    END IF
  END FOR
END FOR
Create an array white_areas[], whose each elements is an array of adjacent buckets.
K := 0
FOR each element(bucket) in the array pointLocations[]
  IF this bucket is unvisited THEN
    Create an array area[] which contains all buckets which are adjacent.
    Add this bucket to the array area[] & mark this bucket as visited
    Add this array area[] to the array white_areas[]
    Call procedure trace(bucket, area[],  pointLocations)
  END IF
END FOR
```

Table III—Specifications for MiniKing SONAR of Tritech make

Parameters	Particulars
Operating Frequency	675 KHz
Beam-width, vertical	400^0
Beam-width, horizontal	3.0^0
Range	100 meters (300 feet)
Souce Level	210 dB per 1 μPa@1 m
System Bendwidth	35 KHz
Scan Sectors	360^0 continuous or 180^0 forward, left or right

Table IV—Average Running Time of Proposed Method, *C*-Means and Multilevel thresholding based segmentation approach (on same machine configuration)

Proposed method (sec)	*C*-Means (sec)	Multi-level Thresholding (sec)
0.043	0.217	0.098

Table V—FOM values for the proposed and referred methods

Tested Process	FOM value
K-means Clustering	0.2618
K-means Threshold	0.1928
Mpth	0.5223
Proposed	0.5688

Table VI—Relative positions and area of occupancies for 4 objects as identified by the target detection algorithm

Objects	Relative Positions (x,y)	Relative Area of Occupancy
A	(21,73)	0.0160
B	(83,-16)	0.0319
C	(19,-93)	0.0156
D	(-30,2)	0.0002

only to white pixels. This is due to the fact that white pixels solely constitute our objects of interest within the image. A pixel is said to have an edge to another pixel if the pixel is within its neighborhood and has the same grayscale intensity The necessity for the adjacent pixel to be in the neighborhood of a pixel lies in the fact that, pixels, which are distant from each other spatially despite having the same intensity, do not offer any relevance to the formation of a patch.

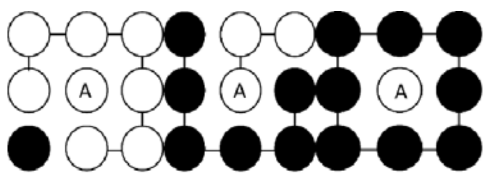

Fig. 5–(left) pixel A is orphan; (middle) pixel A is a pendant; (right) pixel A is strongly connected.

However, pixels not connected to any other pixels in their neighborhood are termed orphans. Pixels which remain connected to only a few of their neighborhood pixels are termed pendants. Hence orphans and pendants both constitute the bulk of grains, which need to be eliminated. On the contrary, pixels connected to a majority of the pixels in its neighborhood are strongly connected. Fig. 5 illustrates an orphan, pendant as well as a strongly connected pixel A. Logically a cluster of strongly connected pixels forms a single connected component in the graph which is homologous to a patch. Structurally it can be conceived that the overall method of feature extraction by the connected component approach needs to be carried out in two passes: (1) identifying "grains" and patches with elimination of "grains", (2) connecting "patches". The first pass involves identifying orphans as well as pendants followed by their elimination. This is practically achieved by inverting the intensity of such pixels to 0. The first pass of the connected component approach is thus described as follows:

Elimination of grains and identification of patches

1. In order to segment an object from a binary image, a 3×3 neighborhood is considered and is termed a *bucket*. A *bucket* is checked for being a 7-pixel strongly connected cluster. This ensures that the *bucket* consists of a "patch" and not a set of isolated "grains" in the form of *orphans* or *pendants*. As a result the bucket can be a candidate for being part of some feature of interest.
2. Following the heuristic as defined in the previous step all the background pixels of a bucket are transformed into white. This is particularly because a major portion of the bucket represents some or all part of a feature, which means that the small portion of background can very easily be ignored in order to finally come with a regular shape of the identified features.

A point of note is that the different features present in the image after thresholding are shape-regularized in the form of small rectangles, granularity of which is a bucket. With this we are left with the task of connecting isolated patches identified in the first pass, in order to form the larger areas of interest. This leads to description of the second pass:

Establishing connections between patches

In the previous step we divided the entire image into patches consisting of number of white *buckets*. In the present step, we have to connect or link the adjacent buckets to get a bigger white patch. Following steps are carried out in order to establish the connectedness of the patches.

1. Each white bucket is marked as "visited" or "not visited".
2. Initially all buckets are marked as "not visited".
3. For each bucket marked "visited" proceed to the next step.
4. If the current bucket is "not visited" then consider all the eight neighbors of that bucket.
5. If any of the neighborhood buckets is adjacent add or link that bucket to the present one and make that bucket as "visited".
6. Repeat all the steps for the newly discovered bucket.

After the second pass, all the white buckets which are adjacent are linked or connected and produce patches which are no more isolated. Table II formally expresses a pseudo-code for the proposed feature extraction method.

Results

For a performance based analysis of the proposed method, an experimental setup was prepared. The targeted platform remained in the form of a lab-scale 2-DOF unmanned surface vehicle as shown in Fig. 6. Equipped with low-cost MEMs Inertial Sensors and a Doppler sensor, the system was driven by a pair of Brushless DC (BLDC) thrusters responsible for the surge and steer motions, thereby imparting to it 2 Degrees of Freedom (DOF), which is sufficient for an effective underwater scanning using SONAR. The SONAR used is mechanical scanning SONAR with single beam from Tritech. The specifications for the SONAR are presented in Table III.

Fig. 6─Image of the test vehicle - SWAN (Surface Water Autonomous Navigator).

Practically, underwater acoustic images are often of low contrast and hence need to be enhanced in order to make the image suitable for effective segmentation. In the present scope of work a Contrast Limited Adaptive Histogram Equalization (CLAHE) method has been exercised upon the filtered SONAR images for subsequent use in the segmentation processes. Forward looking sonar images pixels are mostly 0 or of lower intensity. As a result, normal histogram equalization is less effective for a SONAR image. But CLAHE has one drawback that it over enhances the noise. For CLAHE we have chosen tile size 2×2 and clip limit as 0.01. For adaptive histogram equalization we have subdivided the whole image into 2×2 tiles and then normalized histogram is applied. Clip limit of 0.05 leads to over contrast enhancement. In the last step adjacent tiles are made to conjoin through interpolation. Region-based interpolation is done bi-linearly. The border pixels are treated specially and interpolated linearly. The resulting images after ordinary histogram equalization and CLAHE are illustrated through Fig. 7.

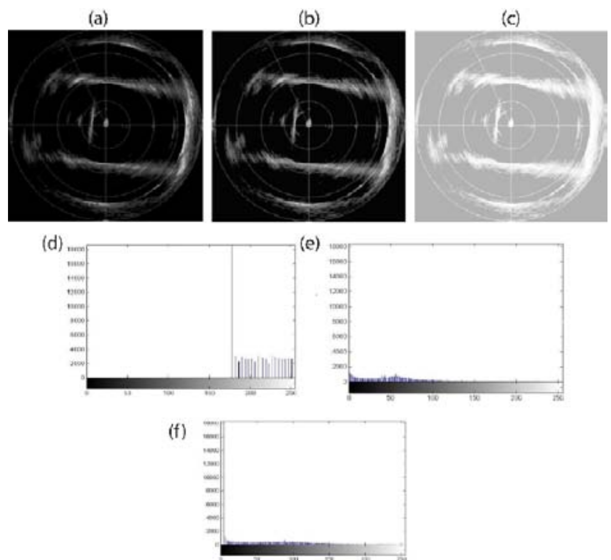

Fig. 7─(a) and (d) are Original image and its corresponding histogram; (b) and (e) are images after ordinary histogram equalization and its corresponding histogram; (c) and (f) are images after application of CLAHE and its histogram.

The proposed method is compared with conventional approach involving multi-level thresholding as well as *K*-means clustering[17]. The algorithmic implications of the reference methods are briefly discussed at the first hand and results are presented subsequently. With *K*-means clustering, the initial clustering begins with random mean assignments over k bins. A pixel is included in a cluster if the Euclidean distance of the pixel from the cluster mean is minimal among all the means. Cluster means are recalculated at the end of iteration. This process continues until there is no change in cluster means. An algorithmic presentation is as follows:

Step 1: Make initial guesses for the means $m_1, m_2, \ldots, m_k$
Step 2: Until there are no changes in any mean
Step 3: Sample x in cluster if$\| x - mi \|$ is the minimum of all k.
Step 4: For i from 1 to k
Step 5: Replace m_i with the mean of all of the samples for cluster i
End for
End until

Multilevel thresholding based segmentation[18] is a good solution where the histogram has more than two dominant peaks. In this method more than two thresholds are needed to partition the histogram. Multilevel thresholding with two thresholds $T1$ and $T2$ classifies a pixel $I(x,y)$ to an object class in the following way:

If $T1<I(x,y)<T2$
Then $I(x,y)$ belongs to one object class.
Else if $I(x,y)>T2$
Then it belongs to another object class
Else if $I(x,y)<T1$
Then belongs to a third object class

An algorithmic interpretation of the multi-level thresholding method is as follows:

Step 1: Repeat steps 2-6, $n/2$ - 1 times; and n is the number of thresholds.
Step 2: Range $R=[a\ b]$; initial value for a=0 and b=255.
Step 3: Find out mean μ and standard deviation σ of all the pixels in range R.
Step 4: Sub-ranges' boundaries $T1$ and $T2$ are calculated as $T1 = \mu\text{-}k1.\ \sigma$ and
$T2 = \mu+k2, \sigma$ and $k1$ and $k2$ are free parameters.
Step 5: Pixels with intensity values in the interval $[a;\ T1]$ and $[T2;\ b]$ are set to the respective weighted means of their values.
Step 6: $a = T1 + 1$; $b = T2 - 1$.
Step 7: Finally repeats Step 5 with $T1 = \mu$ and with $T2 = \mu+1$

Fig. 8 shows a comparison between conventional thresholding methods as well as segment-wise thresholding done on the same SONAR image. It may be clearly observed that the former involves loss of vital details from the image whereas the segment-wise multi-modal adaptive thresholding leads to a better preservation of details as present in the filtered image.

The segment-wise technique actually performs local thresholding over regions in an image whereby, the hotspots as well as background can preserve their own variations as a completely separate entity altogether from the rest of the image. The running time efficiency of the proposed method is depicted in Table IV. The proposed method seems to computationally outperform the other two methods in consideration.

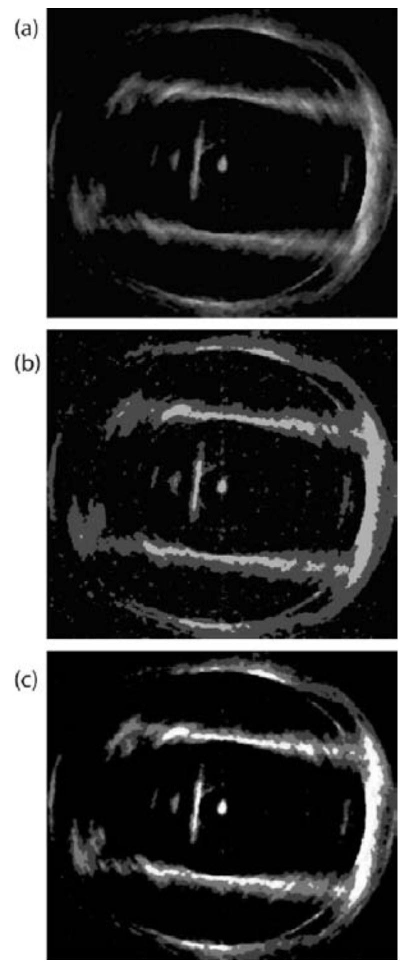

Fig. 8—SONAR image segmented by *C*-Means Clustering (a), Multilevel Thresholding (b) and Proposed Method (c) respectively.

Feature extraction on the acoustic image is illustrated through Fig. 9, with the image after thresholding is being considered for the first pass of extraction. Feature thickening is done in order to demarcate the objects of interest embedded within the image.

Taking note of the quality of processed image, the performance of edge preservation over the speckle reduction of a noisy (speckled) image is estimated using Figure of Merit (FOM) parameter[19]. The parameter may be defined as follows:

$$FOM = \frac{1}{\max\{\breve{N}, N_{ideal}\}} \sum_{i=0}^{N} \frac{1}{1 + d_i^2 \alpha} \quad (7)$$

where, $\breve{N}$ and N_{ideal} are the number of detected and ideal edge pixels, respectively, d_i is the Euclidean distance between the i^{th} detected edge pixel and the nearest ideal edge pixel, and α is a constant typically set to 1/9.The total result can be analyzed by the figure of merit analysis which has been described in Table V. From the table it can be realized that the proposed technique preserves the maximum edge after completing the process. Table VI is a summary of the relative locations and sizes of objects thus identified, which is further illustrated in Fig. 10.

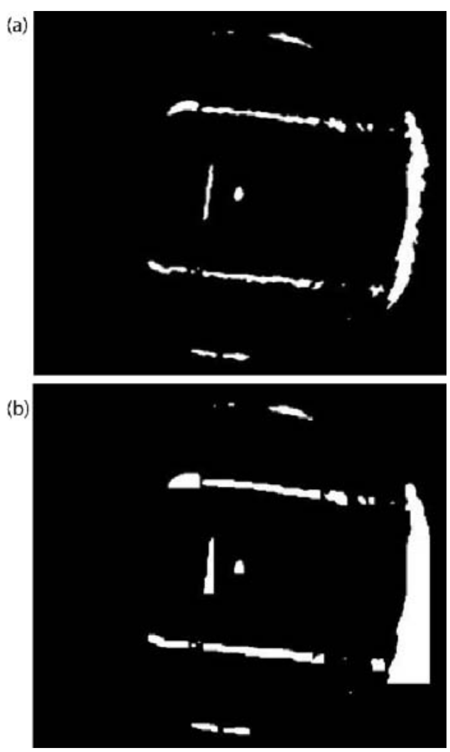

Fig. 9‒After removal of small patches (a) and after feature thickening approach (b).

Discussion

The connected component approach used in feature extraction is drawn in the landscape of graph theoretic algorithms used widely in processing of optical images. According to a survey of graph based image segmentation algorithms provided in[20-22] cut sets have been frequently used. It has been used in conjunction with iterative approaches, fuzzy methods, Euler graphs as well as spanning trees. However, content of information embedded within acoustic images is sparse in comparison to optical images, wherein both the foreground as well as background is quite detailed. Therefore segmentation in case of optical images is relatively non-trivial in comparison to acoustic images. Unlike conventionally used graph methods in segmentation, the proposed connected component algorithm essentially prunes non-requisite information which further simplifies the task of target identification. However, the present algorithm of segmentation and feature extraction is constrained by the fact that the corresponding acoustic image must be having a multi-modal histogram. This leaves us with the scope to explore further in improvising the present algorithm in context of greater adaptability.

Although underwater cameras are commercially available, vision guided navigation is not very popular with AUVs. This is primarily because such a methodology is largely dependent on the turbidity conditions of the surrounding waters. However, SONARs are typically used in collision detection for an AUV. They are mounted on the AUVs in various preferred orientations. Different SONARs ranging from sector scanning to imaging including multi-beam as well as forward looking have been used in this regard.

Conclusion

Consequently, several techniques have been adopted in determining static as well as dynamic objects with online detection as well as offline processing wherein; range and bearing of objects are estimated from reconstructed images instead of taking the actual SONAR images as input.

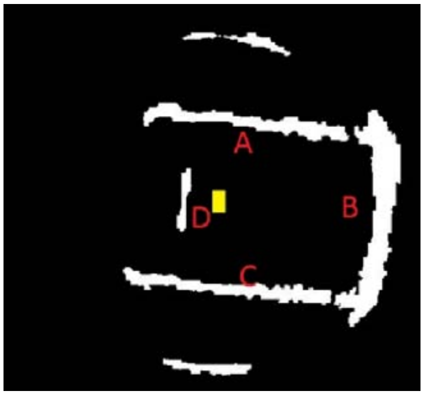

Fig. 10–Relative positions of objects A, B, C and D identified around the vehicle marked in yellow.

In the present scope of this paper a particular contribution is made in the field of acoustic image segmentation using region wise adaptive multi-modal thresholding technique, followed by a graph theoretic approach to feature extraction. Pixel connectedness is used as a criterion in order to come out with proper identification of objects embedded within the acoustic image. The proposed method is tested for validity by taking into consideration images from physical SONAR. The new segmentation method exhibits low running time and better processing quality in comparison to conventional methods like multi-level thresholding and *K*-means clustering.

Acknowledgements

The present work has been carried out as part of the 12th FYP Network Program of CSIR. The authors would like to express their gratitude to CSIR for approving the project on underwater robotic systems, and to the entire Robotics & Automation group for extending their support in away as to take this work to completion. Special regards may be conveyed to Chandan Har, Pratik Saha and Jyotirmoy Karmakar for preparing necessary test bed in order to carry out requisite experimental work.

References

1. Fisher M.H. and Ritchings R.T., Attributed relational tree approach to signal classification, *IEE Proc. Radar, Sonar Navig.,* 141(6) (1994): 319-324.
2. David M. Lane., Chantler M.J., and Dai Dongyong., Robust Tracking of Multiple Objects in Sector-Scan Sonar Image Sequences Using Optical Flow Motion Estimation, *IEEE J. Ocean. Eng*. 23(1) (1998): 31-46.
3. Maria Lianantonakis and Yvan R. Petillot., Sidescan Sonar Segmentation Using Texture Descriptors and Active Contours, *IEEE J. Ocean. Eng*. 32(3) (2007): 744-752.
4. Andrea Truccoand and Simone Curletto, Extraction of 3-D Information From Sonar Image Sequences, *IEEE Trans. Syst. Man. Cybern. B Cybern.* 33(4) (2003): 687-699.
5. Yao Bin, Li Hai-Sen, Zhou Tian and Sun Sheng-he, Segmentation of complex objects' sonar images using parameter-fixed MRF model, *J. Marine Sci. Appl.* 5(4) (2006): 42-47.
6. Stephen G. Johnson and Michael A. Deaet, The Application of Automated Recognition Techniques to Side-Scan Sonar Imagery, *IEEE J. Ocean. Eng*. 19(1) (1994): 138-144.
7. Palmese and Andrea Trucco, From 3-D Sonar Images to Augmented Reality Models for Objects Buried on the Seafloor, *IEEE Trans. Instrum. Meas.* 57(4) (2008): 820-828.
8. Ken Teo, Kai Wei Ong and Hoe Chee Lai, Obstacle Detection, Avoidance and Anti Collision for MEREDITH AUV, *In Proc. Of OCEANS 2009, MTS/IEEE Biloxi, MS,* (26-29 Oct. 2009).
9. Tiedong Zhang, Wenjing Zeng and Lei Wan, Underwater Simultaneous Localization and Mapping Based on Forward-looking Sonar, *J. Marine Sci. Appl.* (2011) 10: 371-376 DOI: 10.1007/s11804-011-1082-1.
10. Kushal Mukherjee, S Gupta, A. Ray and S. Phoha, Symbolic Analysis of Sonar Data for Underwater Target Detection, *IEEE J. Ocean. Eng*. 36(2) (2011): 219-230.
11. Huu-Giao Nguyen, Ronan Fablet, Axel Ehrhold and Jean-Marc Boucher, Keypoint-Based Analysis of Sonar Images: Application to Seabed Recognition, *IEEE Trans. Geosci. Rem. Sens*. 50(4) (2012): 1171-1184.
12. Zacharia, P.U. and Nataraja, G.D., Fishery and biology of threadfin bream, *Nemipterusmesoprion* from Mangalore-Malpe., *Indian J. Fish.,* 50(1) (2003): 1-10.
13. Del Rio Vera J., Cotras E., Groen J. and Evans B., Automatic Target Recognition in Synthetic Aperture Sonar Images Based on Geometrical Feature Extraction, *Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing, Volume* (2009), Article ID 109438.
14. Punam K. Saha and Jayaram K. Udupa, Scale-Based Diffusive Image Filtering Preserving Boundary Sharpness and Fine Structures, *IEEE Trans. Med. Imag.* 20(11) (2001): 1140-1155.
15. Mengxing Huang, Wenjiao Yu and Donghai Zhu, An Improved Image Segmentation Algorithm Based on Otsu Method, *In Proc. of SPIE*, 6625, 66250I, (2008)
16. Jong-Sen Lee, Jen-Hung Wen, Ainsworth T. L., Kun-Shan Chen and Chen A. J., Improved Sigma Filter for Speckle Filtering of SAR Imagery, *IEEE Trans. Geosci. Rem. Sens*. 47(1) (2009): 202-213.
17. Richard O. Duda, Peter E. Hart, David G. Stork, Pattern Classification, Second Edition, (Wiley-Interscience Publication), 2000.
18. Arora S., Acharya J., Verma A. and Prasanta K. Panigrahi, Multilevel thresholding for image segmentation through a fast statistical recursive algorithm, *Pattern Recogn. Lett*., 29 (2008): 119-125.
19. Pratt W. K., Digital Image Processing. (Wiley, New York)

1977.

20. Basavaprasad B. and Ravindra S. Hegadi, A Survey on Traditional and Graph Theoretical Technique for Image Segmentation, *International Journal of Computer Applications,* Proceedings on National Conference on Recent Advances in Information Technology, (2014): 38-46.
21. Santle K., Camilus and Govindan V.K., A Review on Graph Based Segmentation, *I.J. Image, Graphics and Signal Processing*, 5(2012): 1-13.
22. Faliu Yi and Inkyu Moon, Image Segmentation: A Survey of Graph-cut Methods, *In proc. Of International Conference on Systems and Informatics* (ICSAI 2012), (2012): 1936–1941.