Underwater Object Detection and Tracking

Chapter · February 2020

DOI: 10.1007/978-981-15-0751-9_76

2 authors:

Divya Priyadarshni
Banasthali University

Maheshkumar Kolekar
Indian Institute of Technology Patna

SoCTA (Soft Computing: Theories and Applications) - 2018

Underwater Object Detection and Tracking

Divya Priyadarshni[1], MaheshKumar H. Kolekar[2]

[1]School of Automation, Banasthali Vidyapith,Tonk- 304022, Rajasthan, India
[2]Department of Electrical Engineering, Indian Institute of Technology, Bihta- 801103, Patna, India

Abstract

Underwater imagery is a powerful tool for hydrographic inspection including the bathymetry and aquatic possibilities over the extent of the swath. This paper describes a flexible technique for detecting a specific object from the clump of objects, using the reference imaging and target imaging. It facilitates to track the path followed by the object in the streaming video. The proposed methodology operates as an optimal interface between the ideal and the empirical situations, whether it is for image processing or video processing. The objectives accomplished through this paper are the summation as, irrespective of the resolution and clarity of image or video frame, the algorithm reliance and achieves the target. To this end, a reliable measure and analysis of data on certain conditions have been done which shows the higher precision of the approach, with the detected and predicted locations, as compared to the traditional techniques which often fail at certain points.

Keywords- image processing; underwater imaging; object detection; object tracking

1 Introduction

The hydrographic industry has often been adopting several techniques like Synthetic Aperture Sonar (SAS) and variable resolution surface creation for conduit channel scrutiny projects as the suite for underwater imagery and video capture via Autonomous Underwater Vehicles (AUVs) or Remotely Operated Vehicles (ROVs) [1][2]. Also, tracking underwater objects during active circumstances can be done by signal processing methods such as sector scan sonar, side scan sonar and SAS, where, the acoustic formation is done i.e., echoes are reflected by the target and are analysed by the receiver, in order to detect the object's presence [3][4]. This acoustic analysis involves the approximation of the time-of-flight of the reflected signal in water, therefore requires the absolute value, known as amplitude techniques. However, amplitude techniques possess several drawbacks that limitate the tracking performance [5]. Underwater applications such as object detection, imaging of wrecks and underwater tracking have always been a continuous research area with its wide field variance. Nevertheless, due to interference processes and formation methods, there is always some noise introduced into the area of interest, which reduces the effectiveness of the wide applications of underwater acoustics. Hence, there is limit down in the dexterity of couth interpretation of underwater images, restricts edge separation, image segmentation, target recognition and classification, and it introduces ambiguity in underwater navigability and texture parametric inversion [6]. However, this paper proposes a method where the underwater images are processed for detecting objects from the cluttered scene as shown in Fig 1 (a) and further the trajectory followed by the object in the video stream is tracked as shown in Fig 1 (b).

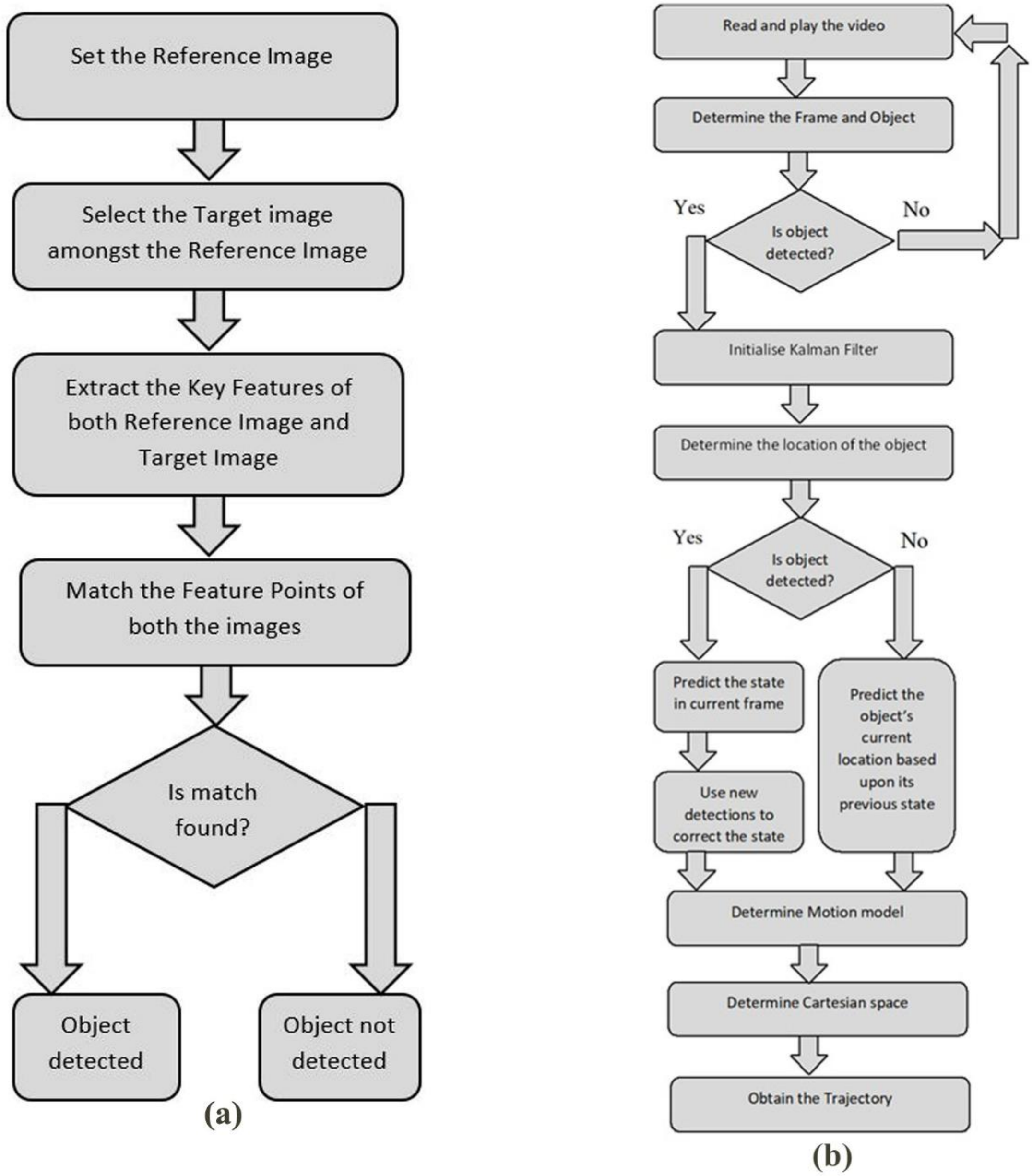

Fig. 1(a). Flowchart for Object Detection **(b)** Flowchart for Object Tracking

2 Object Detection

There have been ample researchers and marine biologists who counted the number of objects in the present frame or identify them in the streaming videos [7]. To that end, object detection is the need. The main hindering met was to make the correct preference of the object detection algorithm, as it is the one, which influences the performance of the object detecting and counting systems [8][9]. The main optimization problem found is the processing time and data acquired i.e., the algorithm's pliability is restrained at the points where application of filter and parameter sequence needs to be followed up, reflecting the inadequacy of these algorithms due to the large processing time required [2]. Therefore, a moving average algorithm is explored, based on removing a reference image exhibiting the background, from the current input image. A specific object's detection can be done, based on finding point relations between the reference and the target image. There are several advantages of comparing Target image and the Reference image in spite of detecting them randomly as, it can detect objects despite a scale change or in-plane angular rotation. It is also manifest to trivial out-of-plane rotation and occlusion. This method works best for objects that exhibit non-repeating texture patterns, which give rise to unique feature matches and the algorithm flowchart is depicted in Fig 1(a).

2.1 Deciding the Target Image amongst the Reference Image

The image captured may consist of several objects and the foreground and background together might be cluttered, so, in order to detect the specified object, a target image from the reference image is chosen, Fig 2 (a) shows target image and Fig 3 (a) shows reference image [10]. Let I and O be the input image and additional background other than the object to be detected, respectively. Then the target image in the n^{th} frame, Tn is given by:

$$Tn = I - O \tag{1}$$

2.2 Feature Extraction

The key feature points in both the images are detected and the strongest 10 points of the target image and the reference image are visualised respectively, Fig 2 (b) and 3 (b). The feature extraction uses the Detect Surf algorithm. However, many other methods for the same can be considered.

2.3 Matching the Feature Points

Once the key features of both the images are decided, the contour inlier and outlier matching of the images using their descriptors is done via Geometric Transform function. This sub loop relates the transformation of the matched points between target and reference, and allows locating the target object in the reference scene, Fig 4. If the match among the above feature points occurs, that implies, the object detection is successful and the algorithm plots the desired object out of the reference scene, Fig 5. However, the algorithm fails to detect any object if the target image does not match the reference image.

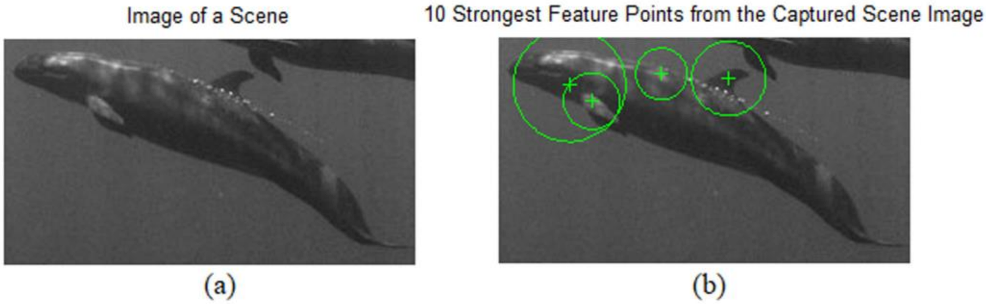

Fig. 2. (a) Target image; (b) Key features of target image

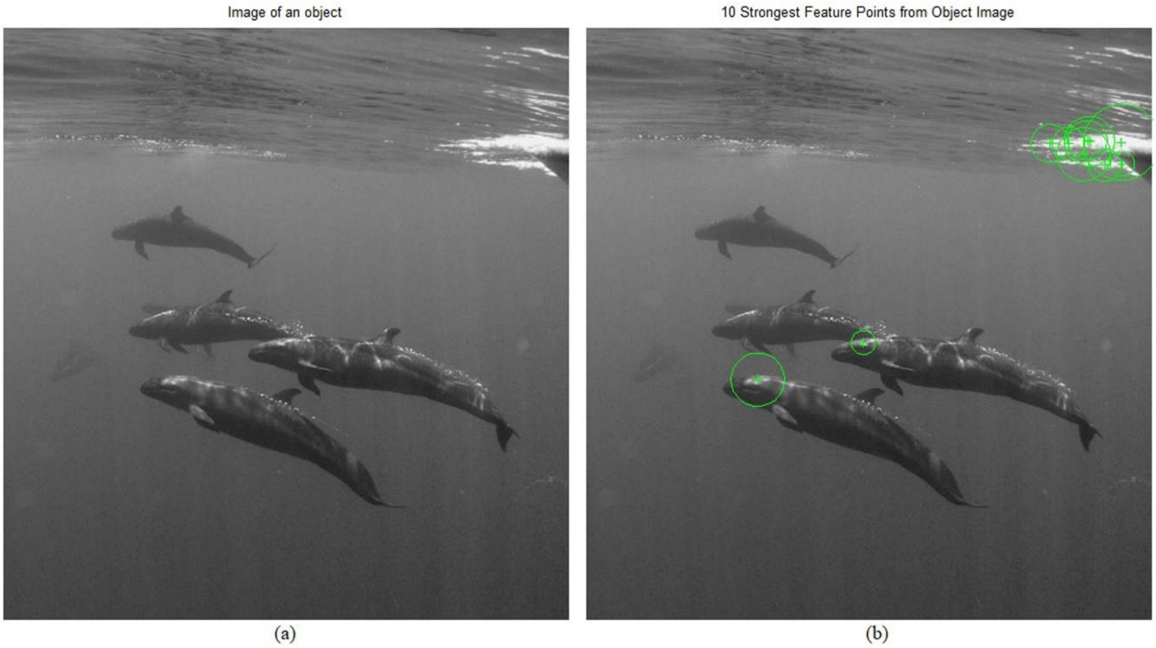

Fig. 3. (a) Reference image; (b) Key features of reference image

3 Object Tracking

Tracking an object is the advanced version of detecting an object, however, they are not same [8][11]. Object detection means locating the object in a single frame whereas, object tracking means detecting the object in more than one frames i.e., locating the object(s) in a video stream. The general framework for tracking single object under dynamic conditions include optimisation of noise dilation and occlusion methods[12] that can be a robotic vision based heuristic tracking system[13] or extracting features followed by Kalman filter[14]. Here, the proposed methodology is formularised in three steps i.e., Detection, Prediction and Data Association and the algorithm can be given in the form of a flowchart in Fig 1 (b).

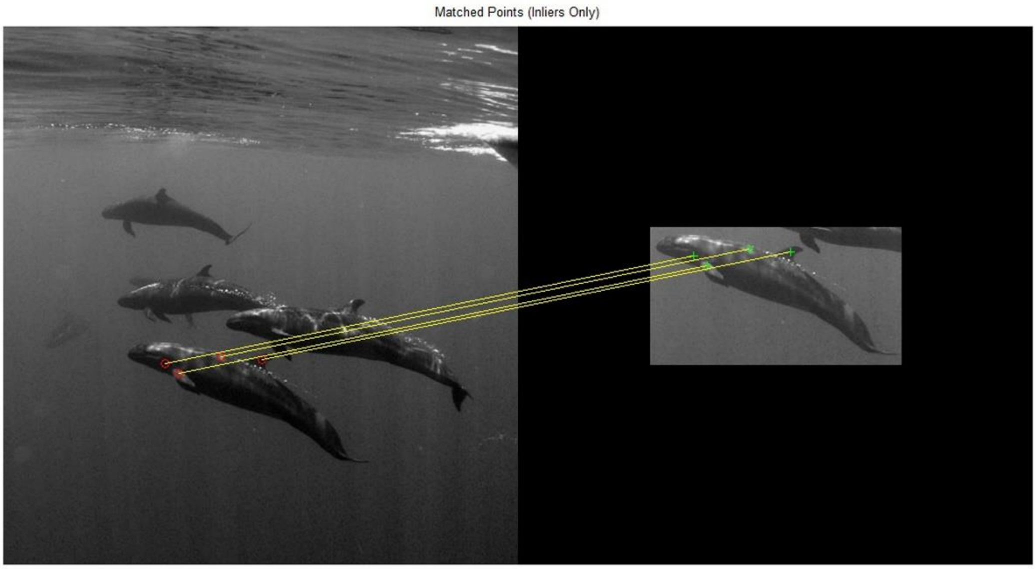

Fig. 4. Matching the key features

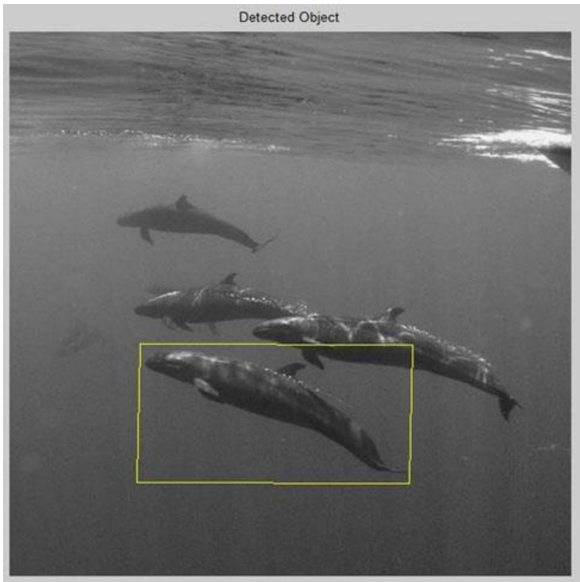

Fig. 5. Object detected

3.1 Detection

The detecting algorithm depends upon the target type and camera type. Here, we use the camera type as stationary to track the target in the streaming video. Target type can be a clearly visible single target among the group of randomly moving fishes, AUVs, ROVs, etc. in any underwater video which might include all of these objects too [12]. This step performs the subtraction of the background using Gaussian Mixture Models, which sets the object properties as the called function argument and hence, detects the objects of interest. The function used is foreground detector.

3.2 Prediction

We use correct and predict methods to eliminate noise present in the tracking system. Predicting also includes the estimation of the object's location in case of any sudden disappearance of the object from the frame. This step includes the initialisation of the Kalman filter and the assumption of the probability of finding the object will be more near its previous detected location in the upcoming frame. This method is especially effective for high frame rates and uniform motion of objects. However, the Kalman filter uses a mathematical model of constant acceleration for more précised and refined results. Simultaneously, the deviation from the motion model and detection errors are also taken into account.

3.3 Data Association

This step involves the tracing the trajectory of the object in motion based on the prediction history of the detected locations in the multiple frames with respect to time. Therefore, the data obtained is plotted on a bipartite graph, which represents both the detection and prediction for each scale point as shown in Fig 6 (a) and (b) [9]. Data association also considers the fact, that random new objects can appear in the visual field, or, the specific object being tracked can leave the field of view. Hence, in any particular frame of video, there might be a need to create some number of new tracks and discard some number of existing tracks, based upon the clear presence of a single target. In addition to the matched presence of the target, the unmatched detections cannot be tracked at the same time i.e., if the scene has multiple entry points, undetected random objects in that particular region are considered noise and the only single entry point is tracked. This can turn out to be one of the future work.

4 Results:

In the graph 6(a) and (b)-
X-axis= Time
40 divisions= 8 sec
1 division= 0.2 sec
5 divisions= 1 unit= 1 sec

Y-axis= Location
50 divisions= 12.7 cm
1 division= 0.254 cm
5 divisions= 1 unit= 1.27 cm

Formula used

$$\textit{Accuracy} = |\textit{detected location} - \textit{predicted location}| \quad (2)$$

where, accuracy is defined as the closeness of the reading to the true value

Precision= |max{(average value- maximum value), (average value- minimum value)}| *(3)*
where, precision is defined as the degree by which a reading can be produced again and again

Analysis of the accuracy and cases with the above calculations is shown in Table 1 and 2.

Note: '+' in the graph represents the actual detected location and '○' in the graph represents predicted location.

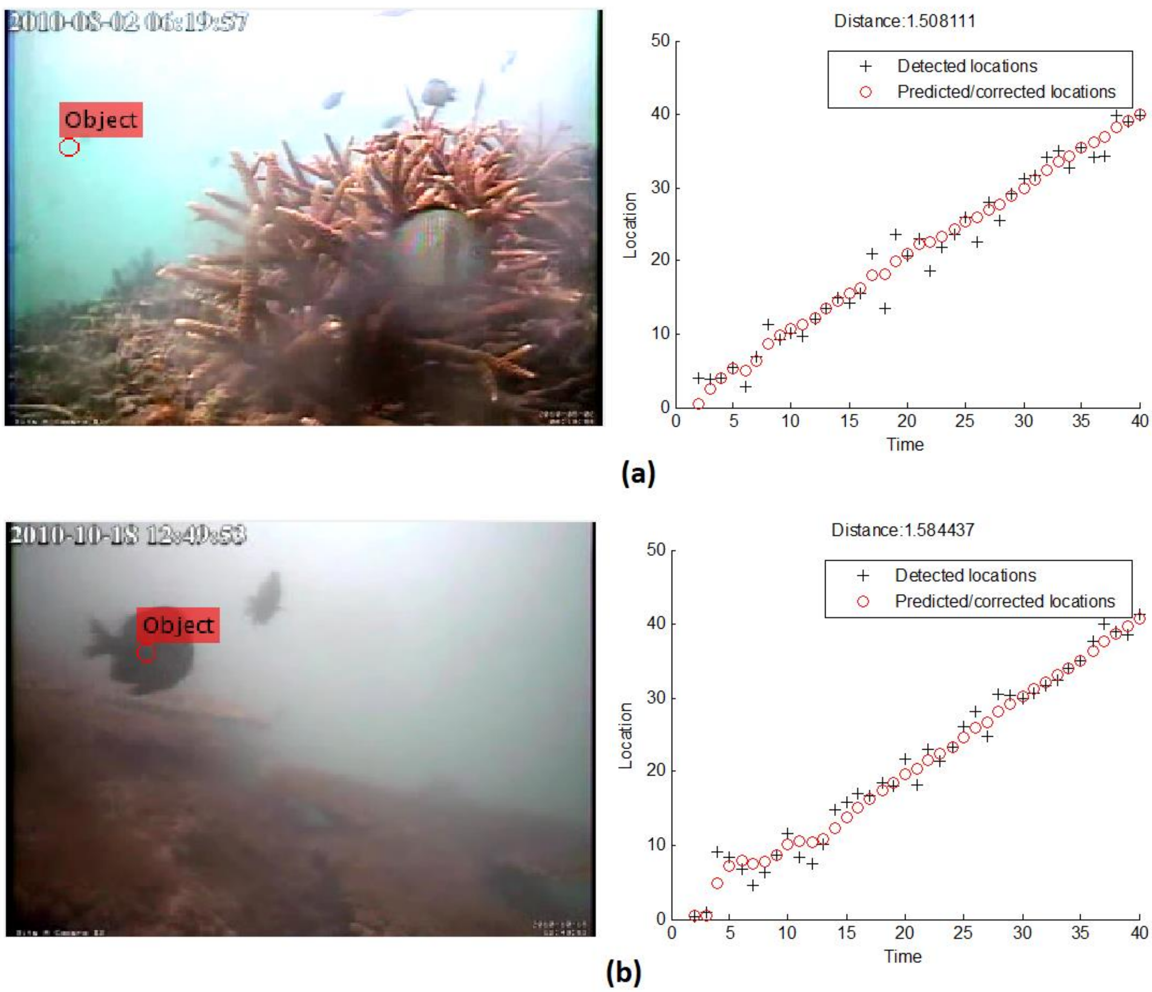

Fig. 6. Object Tracking (a) Video1 streaming and graph; (b) Video2 streaming and graph

Table 1. Analysis of Object Detection images

	Change in inclination	**Change in background**	**Change of frame**
Image1	0° (Detected)	Detected	Not Detected
Image2	30° (Detected)	Detected	Not Detected
Image3	90° (Detected)	Detected	Not Detected
Image4	180° (Detected)	Detected	Not Detected
Image5	270° (Detected)	Detected	Not Detected

Table 2. Analysis of Object Tracking videos

	Object Size	**Video Resolution**	**Accuracy(cm)**	**Precision(cm)**
Video1	Very small	Low	0.53, 1.40, 1.24, 0.48, 0.93	0.80
Video2	Very small	Very Low	0.38,0.38, 0.66, 0.40, 0.44	0.78
Video3	Large	Low	0.46, 0.39, 0.49, 0.53, 0.40	0.82
Video4	Medium	High	0.61, 0.60, 0.41, 0.42, 0.42	0.92
Video5	Small	Medium	0.38, 0.38, 1.42, 0.91, 1.26	0.84

5 Limitations

- Algorithm cannot read RGB image. Image necessarily needs to be a gray image for detecting the object.
- Algorithm doesn't accept change of frame i.e., if the same object with slightest of inclination is captured at 't1' and 't2' in frame 'f1' and 'f2' respectively, the object cannot be detected in f2 with the image at t1 or in f1 with image at t2.
- The algorithm does not provide 100% accuracy. The predictions always possess experimental and environmental errors; however, it is free from gross errors.

6 Future Work

If there is a cluster of the object in the frame, the algorithm could be improvised to count the number of objects for the exact evaluation. The detected objects could be classified to specific category i.e. if a certain frame consists of several objects like fish, rocks, plants, then irrespective of the exact type or location of the fish, all the fishes in the frame should be detected. To improve object classification accuracy probabilistic classifiers such as Bayesian Belief Network [15], Hidden Markov Model [16] based classifier can be tested. Since background detection of the object also contributes significantly, we will use efficient background modelling techniques in the future [18].

References

1. Juan Carballini, Fernanda Viana, Using Synthetic Aperture Sonar as an Effective Tool for Pipeline Inspection Survey Projects. RIO Acoustics, 978-1-4673-7019-6/15 (2015).
2. Niranjan Anand Hule, Underwater Image Processing For Object Detection. International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 2(2015).
3. Øyvind Ødegårda, Roy E. Hansenc, Hanumant Singhe, Thijs J. Maarleveld,: Archaeological use of Synthetic Aperture Sonar on deepwater wrecksites in Skagerrak. Journal of Archaeological Science 89,1e13 (2018).
4. G.L. Foresti and S. Gentili, A Vison Based System for Object Detection In Underwater Images. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 14, pp. 167-188,No. 2 (2000).
5. Angela Digulescu, Ion Candel, Cornel Ioana, Diana Bucur, Teodor Petrut,:Underwater Object Tracking Using Time Frequency Signatures of Acoustic Signals. OCEANS '14 MTS/IEEE Taipei – Ocean Regeneration,Taipei, Taiwan. pp.978-1-4799-3646-5/14, (2014).
6. Ziqi Song, Hongyu Bian, Adam Zielinski, Underwater Terrain-Aided Navigation Based on Multibeam Bathymetric Sonar Images. J. Marine Sci. Appl. 14: 425-433 DOI: 10.1007/s11804-015-1334(2015).
7. Hemavathy, Dr. Shobha, Object Detection and Tracking under Static and Dynamic environment: A Review. International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 10, October (2013).
8. Spampinato, C, Chen-Burger, Y-H, Nadarajan, G & Fisher, Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos. in Proc. 3rd Int. Conf. on Computer Vision Theory and Applications (VISAPP). vol. 2, pp. 514-519,R (2008).
9. Kapil Gupta, M H Kolekar, Unsupervised Graph Based Video Object Extraction, 21st American Conference on Information Systems, Puerto Rico, USA, (2015).
10. Data set is acquired from 'Fish for knowledge', http://f4k.dieei.unict
11. Enric Galceran, Vladimir Djapic, Marc Carreras, David P. Williams, A Real-time Underwater Object Detection Algorithm for Multi-beam Forward Looking Sonar. Galceran2012ARU.
12. Chua, K. & Arshad, M. R. / Robotics Vision-based Heuristic Reasoning for Underwater Target Tracking and Navigation, pp. 245 - 250, International Journal of Advanced Robotic Systems, Volume 2, (2005).
13. Shaorong Xie, Jinbo Chen, Jun Luo, Pu Xie and Wenbin Tang, "Detection and Tracking of Underwater Object Based on Forward-Scan Sonar". Realistic Dynamic Posture Prediction of Humanoid Robot: Manual Lifting Task Simulation, pp.341-347, (2012).
14. Maheshkumar H. Kolekar, Bayesian Belief Network Based Broadcast Sports Video Indexing, Int. Springer Journal of Multimedia Tools and Applications,54:27-54, (2011).
15. M. H. Kolekar and S. Sengupta, Bayesian network-based customized highlight generation for broadcast soccer videos, IEEE Transactions on Broadcasting, vol. 61, no. 2, pp. 195–209, (2015).
16. M H Kolekar, S Sengupta, Hidden markov model based video indexing with discrete cosine transform as a likelihood function, Proceedings of the India Annual Conference, 157-159, (2004).
17. M H Kolekar, Intelligent Video Surveillance Systems: An Algorithmic Approach, CRC press book, 2018.