

# Chapter 1

## UNDERWATER ANIMAL RECOGNITION

### 1.1 Outline

Previous efforts to recognize underwater organisms in their natural environments exist; these methods are constrained to specific type of organisms and do not translate to other application domains.

Section 0: Intro

Section 1: Attempts to count moving sea-animals like fish using stationary cameras is performed through background subtraction.

Para 1: One way to detect animals in natural environments is through background subtraction; in this method a snapshot of the environment is taken and any change to this snapshot is considered to be evidence of a moving animal.

Para 2: Detected animals are further classified into different species based on their appearance characteristics.

Para 3: Background subtraction is only applicable in scenarios where the background is static and the camera is stationary.

Section 2: Background subtraction is inherently challenging when dealing with sedentary animals like scallops that require mobile robotic platforms for image acquisition.

Para 1: Sedentary animals do not move and hence cannot be detected from their motion as in background subtraction, in this case the imaging platform is required to move to capture images.

Para 2: Movement of the camera results in changing background and hence generating a static model for the environment becomes problematic.

Para 3: Lack of a background model and lack of motion based segmentation makes background subtraction in applicable in the context of sedentary animals like scallops.

Section 3: Previous scallop recognition efforts involve artificial scallop beds with controlled imaging conditions or natural image databases with few tens of images.

Para 1: Proof of concept studies for validating scallop recognition algorithms are often tested on artificial scallop beds with controlled conditions that do not represent the variability of environmental conditions observed in natural images.

Para 2: Previous scallop recognition efforts on natural image datasets are limited to small set of few tens of images which do represent the ability of the algorithms to generalize on large databases containing thousands to millions of images.

Section 4: Currently existing methods do not offer effective frameworks for counting stationary animals like scallops from noisy low-resolution underwater image datasets.

Para 1: Currently existing methods are primarily focused on detecting moving animals like fish through background subtraction.

Para 2: Existing methods that work on sedentary animals are primarily trained to operate on controlled environments and lack the robustness to deal with the wide gamut of conditions exhibited by natural environments.

Para 3: The methods that work on natural images are restricted to select set of images bringing into question their ability to generalize over large datasets that exhibit wide variability and substantial amount of noise.

Para 4: There is no existing method that can count stationary animals like scallops from noisy low-resolution underwater images.

Section 5: A general object recognition tool, easily customizable to recognize any type of target, needs to be developed for automated processing of large natural image datasets.

Para 1: With the availability of robotic platforms that can collect staggering amounts of data in form of millions of images, there is a need for data-processing tools that can automate the processing of natural image datasets.

Para 2: Marine scientists and government agencies involved in wildlife surveys require tools that can automatically process imagery data to determine population statistics of animal species.

Para 3: A generalized object recognition tool that can sift through large datasets and produce population counts from natural image datasets need to be developed.

In natural settings, living organism often tend to blend into their environments to evade detection via camouflage. Webster’s thesis work [21] provides a detailed exposition on the visual camouflage mechanisms adopted by animals to blend into their background. Under such circumstances of camouflage, there are very limited visual cues that can be used to identify animals. Even in the presence of visual cues, the task of identifying animals from natural scenes is shown to be a cognitively challenging and complex task [22].

There is an extensive body of existing literature related to identifying underwater animals. Some of this work is specialized for identifying different types of animals under specific environmental conditions. In general, the existing methods can be broadly divided into methods devised for identifying mobile organisms and methods for stationary organisms. The former category is useful in dealing with a wide range of sea organisms like the varied species of fish which swim through water. The latter less

studied category is identifying sedentary underwater animals like scallops and corals. Both these categories present their own set of challenges. In the rest of this chapter we will visit the techniques relevant to moving animals and show how they are different from the methods employed for stationary animals. We will follow this by providing an overview of the existing literature on recognising stationary animals with special emphasis on methods developed for identifying scallops. The shortcomings of the existing methods in recognising stationary underwater animals is addressed through a multi-layered method discussed in Chapter ?? of this thesis.

## 1.2 Methods for Recognition of Moving Underwater Organisms

Recognising and counting mobile marine life like fish [18, 2, 23] and studies in aquaculture [24] has been previously attempted. The recurring theme in these efforts involve the use of stationary cameras to detect the presence of moving species provided the background can be described by a prior model. This technique of assuming a known background and using changes in the background as an evidence for the presence of a moving object entering the field of view of a sensor is called background subtraction. In the marine species identification case, any changes to the background is assumed to be caused by a moving marine organism. The pixels in the image that deviate from the background model can be labelled as the pixels belonging to the organism.

Once a marine organism is detected through background subtraction, then other computer vision or machine learning techniques can be used to classify the organism into a specific species based on its visible characteristics. This classification task can be achieved through conventional machine learning approaches. For instance the salmon species classification algorithm developed by Williams. et. al. [23] uses active contours to model the shape of the fish before comparing these contours to known salmon species. However if the pixels corresponding to the organism are corrupted by high levels of noise, some specialized technique that is robust to noise might be required.

Background subtraction requires the distribution of background pixels to be represented as a mathematical model. In an underwater setting, such a background model can only be obtained if the camera is stationary and the background model represents regions which are stationary and do not change over time. The opportunity to employ the background subtraction based techniques arises in underwater environments with stationary fixtures designed to study a specific underwater location. In instances where such stationary arrangement of cameras are not available, background subtraction is not applicable.

### **1.3 Methods for Recognition of Sedentary Underwater Organisms**

Since sedentary marine animals like scallops do not move, a mobile robotic platform is required to traverse the underwater terrain to image and recognize sedentary marine animals. Extending background subtraction to work with mobile robotic platforms is challenging as motion of the platform causes changes in its background and hence generating a model for the background to perform background subtraction is problematic. This makes the task of detecting sedentary organisms more challenging than moving organisms. The lack of background model then prompts the need for a foreground model. If a foreground is available, the task of detecting an organism can be realised as a search for pixels satisfying the foreground model in the image.

Detecting an organism also involves segmenting all pixels of the organism before one can classify the organism into a known category. The motion-based segmentation of marine animals that results from background subtraction is no longer possible in case of sedentary organisms. Furthermore, the task of segmentation can be challenging in noisy images with weak edges as the boundary pixels of the foreground object cannot be easily distinguished from the background pixels.

The lack of background model and motion based segmentation makes background subtraction infeasible. This leads to the need for techniques that depend on foreground models and use of other features to detect and segment organisms from the background. This task gets complicated if the organism does not present significant

visual cues, as in the case of creatures exhibiting camouflage, that make it distinctive from the background. High levels of noise or unpredictable environmental variable could also significantly impede any animal recognition mechanism.

## 1.4 Scallop Recognition

Previous efforts to detect animals like plankton [13, 19], clam [7] and a range of other benthic megafauna [17] exist. Most methods here are specialized for a specific species or only tested in controlled environments, and in some cases even require specialized apparatus like in the plankton recognition studies [13, 19]. On the otherhand, the proposal by Schoening [17] combines a series of automated tools like specialized color correction, segmentation and classification modules along with some level of manual expert support for identification of several marine organisms like sea anemones and sponges from natural image datasets.

There are several aspects that make the Scallop recognition problem challenging. Scallops, especially when viewed in low resolution, do not provide features that would clearly distinguish them from their natural environment. This presents a major challenge in designing an automated identification process based on visual data. To compound this problem, visual data collected from the species' natural habitat contain a significant amount of speckle noise. Some scallops are also partially or almost completely covered by sediment, obscuring the scallop shell features. A highly robust detection mechanism is required to overcome these impediments.

There is a range of previously developed methods specialized for scallop recognition [1, 3, 5, 4, 6, 10, 12, 11] that operate on different assumptions either with regards to the environmental conditions or quality of data. Existing approaches to automated scallop counting in artificial environments [5, 4] employ a detection mechanism based on intricate distinguishing features like fluted patterns in scallop shells and exposed shell rim of scallops, respectively. Imaging these intricate scallop shell features might be possible in artificial scallop beds with stationary cameras and minimal sensor noise, but this level of detail is difficult to obtain from images of scallops in their natural

environment. A major factor that contributes to this loss in detail is the poor image resolution obtained when the image of the target is captured several meters away from it. Overcoming this problem by operating an underwater vehicle too close to the ocean floor will adversely impact the image footprint (i.e. area covered by an image) and increase the risk of damaging the vehicle.

Furthermore, existing work on scallop detection [1, 3] in their natural environment is limited to small datasets (often less than 100 images). The framework proposed in [3] employs a sliding window approach to targettedly search for the presence of a scallop. The large number of overlapping windows that need to be processed per image raises scalability concerns if this method were to operate on a large dataset containing millions of images. Additionally, the few tens of images natural image used as a test set raises questions about the generalizibility of this method and its ability to function under varied environmental conditions. The work by Dawkins [1] is more detailed in its treatment of the natural environmental conditions spanning the scallop habitat. The images used here are collected using a towed camera system with minimal noise which greatly enhances the performance of the machine learning and computer vision algorithms. Despite the elaborate imaging setup designed to minimize noise, the results shown are derived only from a few tens to hundreds of images. It is not clear if the machine learning methods proposed in [1] can extend to noisy image data captured by [Autonomous Underwater Vehicle \(AUV\)](#)s. From these studies alone, it is not clear if such methods can be used effectively in cases of large datasets comprising several thousand seabed images. An interesting example of machine-learning methods applied to the problem of scallop detection [6] utilizes the concept of Bottom-Up Visual Attention (BUVA). The approach is promising but it does not use any ground truth for validation.

The work in [10, 12, 11] offers a multi-layered object recognition framework validated on a natural image dataset for scallop recognition application. The main emphasis here is to develop a technique that can work on low quality noisy sensor data collected using [AUV](#)s. The other attribute is to build a scalable architecture that can

operate on large image datasets in the order of thousands to millions of images and can be generalized for recognizing other marine organisms. A detailed account of this multi-layered approach is discussed in Chapter ??.

## 1.5 Summary of Existing Methods

Existing methods on recognizing moving underwater organisms like fish depend on background subtraction. Though background subtraction offers a convenient way to detect the presence of an animal along with motion based segmentation to isolate the pixels of the organism, the need for a well defined background model limits its applicability. A well defined background model is only obtainable only in scenarios where the background is unchanging or if the evolution of the background can be captured by a mathematical model. Such modeling is possible if the imaging setup is stationary and the [Field of View \(FOV\)](#) of the predominantly captures a static background. In such a case any moving object can be identified and isolated from the rest of the background. This is only possible when the imaging system is expressly looking for moving objects. In case of stationary organism, the imaging systems has to move to capture data and identify organisms from the images. In this case, the background distribution is very challenging to model if not impossible.

In order to recognize sedentary organisms, a mobile imaging system capturing images is required. Stationary camera systems have been used to identify scallops from artificial beds [5, 4]. However in these cases, the region to be imaged was small enough to fit inside the [FOV](#) of a network of stationary cameras. The environment in this case comprised an artificial scallop bed of known substrate along with full control of parameters like illumination. In natural environments the magnitude of the region to be imaged is sufficiently large to make it prohibitively expensive to use a large array of stationary cameras. Mobile sensors offer a cost effective and practical alternative to this problem. The use of a mobile camera system however exposes the system to a gamut of environmental conditions. The bifurcation of the gathered data into foreground and background without full knowledge of the changing environmental factors poses a



technical challenge. The methods developed for identifying organisms under artificial controlled conditions are not transferrable systems that operate over natural image datasets.

A few automated scallop recognition systems that have been validated on natural image datasets [1, 3, 6, 10, 12, 11]. Even those that are validated on natural image datasets often limited to small test sets of less than a few hundred images. The sizeable cost and effort required to collect natural image data explains the reason for the limited amount of work in this domain. Furthermore, developing an automated system to recognize organisms from natural images also entails a manual annotation effort to obtain a generalizable representation of a species that spans different variations of its habitat. Such a labelled dataset is intended to serve as a knowledge base for a machine learning system developed to recognize an organism. The work by Fearn [6] discusses a relative comparison of a multitude of image processing techniques and computer vision techniques applied to detect scallops from artificial beds. However, there is no explanation the magnitude of the dataset, the annotation process or a clear quantification of the performance of these techniques with regards to ground truth. The work by Dawkins [1] offers a combination of several tools helpful for scallop identification. However, most of the results shown are derived from a small dataset of less than a hundred images. The scallop recognition approach in [10, 12, 11] is designed to deal with noisy AUV images that are characterized by high levels of speckle noise, uneven illumination and low contrast. This is a multi-layered architecture that has been validated over a dataset containing a few thousand images. A more detailed comparison of the work in [1] against [11] along with working details of [11] can be found in ??.

## 1.6 Motivation for a Generalized Automated Object Recognition Tool

Understanding the parameters that affect the habitat of underwater organisms is of interest to marine biologists and government officials charged with regulating a multi-million dollar fishing industry. Dedicated marine surveys are needed to obtain population assessments. One traditional scallop survey method, still in use today, are

dredge-based surveys. Dredge-based surveys have been extensively used for scallop population density assessment [14]. The process involves dredging part of the ocean floor, and manually counting the animals of interest found in the collected material. In addition to being invasive and detrimental to the creatures habitat [9], these methods have accuracy limitations and can only generalize population numbers up to a certain extent. There is a need for non-invasive and accurate survey alternatives.

The availability of a range of robotic systems in form of towed camera and Autonomous Underwater Vehicle (auv) systems offer possibilities for such non-invasive alternatives. Optical imaging surveys using underwater robotic platforms provide higher data densities. The large volume of image data (in the order of thousands to millions of images) can be both a blessing and a curse. On one hand, it provides a detailed picture of the species habitat; on the other it generates a need for extensive manpower and time to process the data. While improvements in robotic platform and image acquisition systems have enhanced our capabilities to observe and monitor the habitat of a species, we still lack the required arsenal of data processing tools. This need motivates the development of automated tools to analyze benthic imagery data containing scallops.

One of the earliest video based surveys of scallops [16] reports that it took from 4 to 10 hours of tedious manual analysis in order to review and process one hour of collected seabed imagery. The report suggests that an automated computer technique for processing of the benthic images would be a great leap forward; to this time, however, no such system is available. There is anecdotal evidence of in-house development efforts by the HabCam group [8] towards an automated system but as yet no such system has emerged to the community of researchers and managers. A recent manual count of our AUV-based imagery dataset indicated that it took an hour to process 2080 images, whereas expanding the analysis to include all benthic macro-organisms reduced the rate down to 600 images/hr [20]. Another manual counting effort [15] reports a processing time of 1 to 10 hours per person to process each image tow transect (the exact image number per tow was not reported). The same report

indicates that the processing time was reduced to 12 hours per tow by subsampling 1% of the images.

Future benthic studies will be geared towards increasing data densities with the help of robotic optical surveys. It is clear that the large datasets in the order of millions of images generated by these surveys will impose a huge strain on researchers if the images are to be processed manually. This strongly suggests the need for automated tools that can process underwater image datasets. With the motivation to reducing human effort, Schoening [17] has proposed a range of tools that can be generalized to organisms like sea-anemones. With an additional requirement of being able to work with low-resolution noisy underwater images, the work in [10, 12, 11] proposes a generalized multi-layered framework that can be used to detect and count underwater organisms. This method has been evaluated on a scallop population assessment effort on dataset containing over 8000 images, the details of which can be found in ??.

## Acronyms

**AUV** Autonomous Underwater Vehicle. [7](#), [9](#)

**FOV** Field of View. [8](#)

## BIBLIOGRAPHY

- [1] Matthew Dawkins, Charles Stewart, Scott Gallager, and Amber York. Automatic scallop detection in benthic environments. In *IEEE Workshop on Applications of Computer Vision*, pages 160–167, 2013.
- [2] Duane R. Edgington, Danelle E. Cline, Daniel Davis, Ishbel Kerkez, and JÃ©rÃ©mie Mariette. Detecting, tracking and classifying animals in underwater video. In *Oceans’06 MTS/IEEE-Boston Conference and Exhibition*, pages 1–5. IEEE, 2006.
- [3] Einar Óli Guðmundsson. Detecting scallops in images from an auv. Master’s thesis, University of Iceland, 2012.
- [4] Koichiro Enomoto, TODA Masashi, and Yasuhiro Kuwahara. Extraction method of scallop area in gravel seabed images for fishery investigation. *IEICE Transactions on Information and Systems*, 93(7):1754–1760, 2010.
- [5] Koichiro Enomoto, Masashi Toda, and Yasuhiro Kuwahara. Scallop detection from sand-seabed images for fishery investigation. In *2nd International Congress on Image and Signal Processing*, pages 1–5. IEEE, 2009.
- [6] R. Fearn, R. Williams, M. Cameron-Jones, J. Harrington, and J. Semmens. Automated intelligent abundance analysis of scallop survey video footage. *AI 2007: Advances in Artificial Intelligence*, pages 549–558, 2007.
- [7] AL Forrest, ME Wittmann, V. Schmidt, NA Raineault, A. Hamilton, W. Pike, SG Schladow, JE Reuter, BE Laval, and AC Trembanis. Quantitative assessment of invasive species in lacustrine environments through benthic imagery analysis. *Limnology and Oceanography: Methods*, 10:65–74, 2012.
- [8] S.M. Gallager, H. Singh, S. Tiwari, J. Howland, P. Rago, W. Overholtz, R. Taylor, and N. Vine. High resolution underwater imaging and image processing for identifying essential fish habitat. In D.A. Somerton and C.T. Glentdill, editors, *Report of the National Marine Fisheries Service Workshop on Underwater Video analysis*, NOAA Technical Memorandum NMFS-F/SPO-68, pages 44–54. 2005.
- [9] SR Jenkins, BD Beukers-Stewart, and AR Brand. Impact of scallop dredging on benthic megafauna: a comparison of damage levels in captured and non-captured organisms. *Marine Ecology Progress Series*, 215:297–301, 2001.

- [10] Prasanna Kannappan and Herbert G. Tanner. Automated detection of scallops in their natural environment. In *21st Mediterranean Conference on Control & Automation, 2013*, pages 1350–1355. IEEE, 2013.
- [11] Prasanna Kannappan, Herbert G. Tanner, Arthur C. Trembanis, and Justin H. Walker. Machine learning for detecting scallops in auv benthic images: Targeting false positives. *Computer Vision and Pattern Recognition in Environmental Informatics*, page 22, 2015.
- [12] Prasanna Kannappan, Justin H. Walker, Art Trembanis, and Herbert G. Tanner. Identifying sea scallops from benthic camera images. *Limnology and Oceanography: Methods*, 12(10):680–693, 2014.
- [13] Catherine McGavigan. A quantitative method for sampling littoral zooplankton in lakes: The active tube. *Limnology and Oceanography: Methods*, 10:289–295, 2012.
- [14] National Marine Fisheries Service Northeast Fisheries Science Center (NEFSC). 50th northeast regional stock assessment workshop (50th saw) assessment report. Technical Report 10-17, p.844, US Dept Commerce, Northeast Fisheries Science Center, 2010.
- [15] L. Oremland, D. Hart, L. Jacobson, S. Gallager, A. York, R. Taylor, and N. Vine. Sea scallop surveys in the 21st century: Could advanced optical technologies ultimately replace the dredge-based survey? Presentation made to the NOAA Office of Science and Technology, October 2008.
- [16] Gregg E. Rosenkranz, Scott M. Gallager, Richard W. Shepard, and Mark Blakeslee. Development of a high-speed, megapixel benthic imaging system for coastal fisheries research in alaska. *Fisheries Research*, 92(2):340–344, 2008.
- [17] Timm Schoening. Automated detection in benthic images for megafauna classification and marine resource exploration: supervised and unsupervised methods for classification and regression tasks in benthic images with efficient integration of expert knowledge. 2015.
- [18] Concetto Spampinato, Yun-Heh Chen-Burger, Gayathri Nadarajan, and Robert B. Fisher. Detecting, tracking and counting fish in low quality unconstrained underwater videos. In *3rd International Conference on Computer Vision Theory and Applications*, pages 514–519. Citeseer, 2008.
- [19] Claus Peter Stelzer. Automated system for sampling, counting, and biological analysis of rotifer populations. *Limnology and oceanography: Methods*, 7:856, 2009.
- [20] J.H. Walker. Abundance and size of the sea scallop population in the mid-atlantic bight. Master’s thesis, University of Delaware, 2013.

- [21] Richard James Webster. *Animal camouflage: disentangling disruptive coloration from background matching*, 2013.
- [22] Felix A. Wichmann, Jan Drewes, Pedro Rosas, and Karl R. Gegenfurtner. Animal detection in natural scenes: critical features revisited. *Journal of Vision*, 10(4):6–6, 2010.
- [23] RN Williams, TJ Lambert, AF Kelsall, and T. Pauly. Detecting marine animals in underwater video: Let’s start with salmon. In *Americas Conference on Information Systems*, volume 1, pages 1482–1490, 2006.
- [24] Boaz Zion. The use of computer vision technologies in aquaculturea review. *Computers and Electronics in Agriculture*, 88:125–132, 2012.