

## Chapter 1

### CONCLUSIONS AND OUTLOOK

#### 1.1 Outline

This work can be used for robotic image analysis in natural environments.

Section 0: Intro

Section 1: This work offers a multi-layered object recognition tool that can work with noisy low-resolution natural images.

Section 2: The object recognition tool is tested on underwater image datasets in the context of an automated scallop survey application.

Section 3: A novel distance based object descriptors is developed to enhance object detection accuracy.

Section 4: This distance based object descriptor is validated through image data from a self-developed underwater robotic platform CoopROV.

Section 5: This work offers tools for natural image processing and object recognition which have been validated through experiments and also tested on natural image datasets.

Section 6: This work can be extended in the future to produce more advanced tools for natural image labeling and object recognition.

Section 7: This work can be extended to recognize more than one category of objects and be used in a multitude of object recognition domains.

Section 8: The current CoopROV design can be improved by adding controllers that complement the object identification tasks.

The work in this dissertation offers robotic image analysis tools designed to operate in noisy natural environments. Underwater environments have been the primary domain used to validate these tools. The two applications that drove the design of these object recognition tools are the automated submerged subway car recognition from artificial reef sites, and scallop recognition problems. Additionally, a multi-view object classifier, that can operate in noisy images without segmentation, is also proposed. Finally, a low-cost [Remotely Operated Vehicle \(ROV\)](#), named CoopROV, was built to facilitate underwater experiments. The different insights gleaned from development of each of these tools are discussed below.

Chapter ?? showcases the application of eigen-value based shape descriptors to the subway car recognition problem. Eigen-value based shape identification is a tool that can be used to identify simple geometric shapes. By reducing subway car recognition to rectangle matching, eigen-value based shape descriptors were successful in recognizing subway cars from seabed images. Though shape matching can aid object recognition, other cues like texture can often play an important role in distinguishing an object. Since eigen-value shape descriptors only use shape information, they are not suited for applications where non-shape cues are more relevant in encoding object information. Additionally, while evaluating eigen-value shape descriptors, it was determined that discretization errors and segmentation errors can significantly impair the performance of the descriptors. In other words, eigen-value shape descriptors are a useful tool for identifying objects, provided we can guarantee that the shape of the object can be represented with minimal discretization error. For this method to work, the objects we are interested in should exhibit a shape profile that is significantly different from all other objects in the background. These requirements may be restrictive, and could render recognition tasks involving complex shapes problematic.

A challenging problem where the validity of such assumptions is in question is scallop recognition. The visual cues presented by a scallop include its distinguishing crescent profile, along with its texture. Since eigen-value descriptors can neither handle complex shape profiles like crescents, nor do they capture textural information, a more

refined approach is warranted. This led to the development of multi-layered object recognition framework proposed in Chapter ??.

Marine biologists and oceanographers often depend on large natural image datasets to study benthic phenomena, similar to surveying marine organisms. With the availability of robotic imaging vehicles, the generation of datasets ranging millions of images are becoming increasingly common. Automated image processing tools are necessary to deal with such large image datasets. Most existing techniques are built on restrictive assumptions that do not generalize to noisy natural image datasets. The multi-layered object recognition approach described in Chapter ?? addresses this problem by providing researchers with a modular architecture that can scale to large natural image datasets. Modular architectures have the flexibility to be reconfigured, in order to solve different object recognition problems. The particular multi-layered framework was tasked with automated scallop recognition from images taken by an [Autonomous Underwater Vehicle \(AUV\)](#), as a means to assess scallop populations. This framework was successful in recognizing 60–70% of scallops in a dataset of over 8000 images. The uniqueness of this approach lies in its ability to handle noisy natural images under varied environmental conditions. To improve the performance of the classifier proposed in Chapter ??, and thereby reduce false positives, more information about a target object can be helpful. To inject this additional information into the classification process, a multi-view object recognition algorithm was formulated and discussed in Chapter ??.

The multi-view approach discussed in Chapter ?? is a machine learning technique designed for binary classification tasks. The objective here is to formulate a robust object classifier that can perform acceptably even in case of noisy image data. Since noisy single-views of an object might often not contain sufficient information to unambiguously recognize it, information from multiple views are combined here to build a machine learning classifier. In this approach, the information from 13 images of each object specimen, captured from different heights, is encoded into a single object model. Additionally, the use of histogram-based global feature descriptors here obviates segmentation of object pixels. Since segmentation can be challenging in noisy

natural images, this feature is attractive. This method is evaluated on a combined dataset of 22 specimens comprising oranges and strawberries. All the specimens are correctly classified. Despite the small dataset (22 specimens), the exceptional performance of this classifier in noisy underwater data is encouraging. To decrease false positives, an approach like this could be used as a classification layer in a multi-layered approach, like the one described in Chapter ??.

To remotely collect data in underwater environments, a submersible robotic vehicle is typically required. Most commercially available robotic vehicles are either expensive or difficult to customize for research needs. To address this problem, CoopROV, a low-cost ROV was designed as a research prototype to support experimentation in underwater settings. CoopROV carries an Inertial Measurement Unit (IMU), stereo cameras, depth and pressure sensors. The onboard electronics on the robot allow customization and inclusion of new sensors. CoopROV also offers a Robot Operating System (ROS) software interface that allows easy access to plethora of existing software tools. The low-cost, and innate flexibility to modify the software and hardware, makes CoopROV an ideal platform to support underwater experiments.

There are some possible directions to extend the object recognition tools proposed in this dissertation. For instance, the scallop recognition approach of Chapter ?? could benefit from using more descriptive templates. In Chapter ??, we see that the appearance of a scallop is characterized by two crescents near the scallop rim: one bright and one dark. Currently, only the position of a dark crescent is captured by the templates used. Using templates that capture additional information, like the position of the bright crescent could be interesting. As for the multi-view object recognition method discussed in Chapter ??, a more expansive dataset could be used to test this algorithm. Another possible improvement is to generalize this multi-view algorithm to accept images from any series of heights instead of a fixed set of heights. Relaxing the restrictions on the different views needed by the multi-view algorithm could make the experiments and data collection effort easier. This multi-view approach could also be expanded to multi-class classification for dealing with more than 2 object classes.

The CoopROV frame could be redesigned to improve the trim (weight distribution) of the vehicle. Designing controllers and building a localization system are other avenues where future improvements on CoopROV is possible.

The main contribution of this dissertation are object recognition tools designed to work on noisy natural images. Each of these tools have been validated on different underwater applications. The developments here are intended to provide automated solutions to researchers dealing with large natural image datasets. The availability of modular reconfigurable architectures to allow researchers to build custom solutions to solve their object recognition needs is the prime objective. The techniques proposed here are first steps in providing tools capable of handling noisy natural images. Further exploration on this domain is necessary to build more robust architectures that can recognize multiple classes of objects. Some examples of object recognition applications, like subway car and scallop recognition, are not all-encompassing but provide an initial thrust in this domain. Further work to build extensive tools required by the marine research community is warranted. In a broader scheme, this dissertation contributes in strengthening the perception capabilities of robots. Ultimately, better scene understanding and perception are vital building blocks that support robot autonomy.

## BIBLIOGRAPHY