

Chapter 1

CONCLUSIONS AND FUTURE WORK

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 are used as the primary domain to validate these tools. The two applications that drove the design of these object recognition tools are the subway car and scallop recognition from sea bed images. Additionally, a proof-of concept demonstration of a multi-view approach capable of performing classification without prior segmentation is also proposed. Finally, during the course of this exploration on object recognition techniques, an underwater [Remotely Operated Vehicle \(ROV\)](#), CoopROV, was built to facilitate underwater object recognition applications. The different insights gleaned from development of each of these tools will be discussed in the remainder of this section. This section will conclude with possible ways to extend and improve these tools.

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 descriptor was 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 the information about an object. Additionally, while evaluating eigen-value shape descriptors, it was determined that discretization errors and segmentation errors, often associated with complex shape profiles, can significantly impair the performance of eigen-value based shape descriptors. Hence eigen-value shape descriptors is a useful tool for identifying objects provided we can guarantee that the shape of the object can be described in a discrete domain with minimal error. Furthermore, 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 are very restrictive and render recognition of complex objects like scallop infeasible. The complexity of

the scallop shape lies in its distinguishing crescent profile along with its unique representative texture. Since eigen-value descriptors neither handle complex shape profile like crescents nor textural information, a more powerful approach is warranted. This lead to the development of multi-layered object recognition framework proposed in Chapter ??.

With the advent of robotic vehicle that can generate image datasets ranging millions of images, automated processing of images is becoming an increasingly important problem. 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 datasets. Modular architectures also contains the built-in flexibility to be reconfigured to solve varied object recognition problems. This automated scallop recognition framework was a part of a scallop survey process that utilized [Autonomous Underwater Vehicle \(AUV\)](#) images to study scallop population. This approach offers a hands-off automated alternative to manual enumeration that was also performed as a part of this scallop survey. This four-layered automated scallop recognition framework combined a series of off-the-shelf computer vision tools like visual attention and graphcut, along with custom designed noise filters and machine learning classifiers. This framework was successful in recognizing 60% of scallops in a dataset of over 8000 images. This multi-layered framework is unique in its ability to handle noisy natural images under varied environmental conditions. To improve the performance of the classifier proposed here and also cut the false positive rate, more information about a target object can be helpful. To inject this additional information into the classification process, a multi-view object recognition approach discussed in Chapter ?? is proposed.

The multi-view approach discussed in Chapter ?? is machine learning technique that is designed for binary classification tasks. This multi-view approach is designed as a solution to build a robust object classifier 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. The classifier here uses 13 images of each object specimen captured from different heights. All these 13 images are combined to build a single object representation. The availability of height information provides another feature dimension for the machine learning technique to capitalize on. Another salient character of this approach is the use of histogram-based global feature descriptors that do not require segmentation of object pixels. Since segmentation can be challenging in noisy natural images, the lack of need for segmentation greatly boosts the appeal of this method. This method is evaluated on a combined dataset of 22 specimens comprising Orange and Strawberry collected using a special imaging apparatus in an underwater environment. All the specimens are correctly classified. The exceptional performance in a small dataset of 22 images is an indicator of the potential of this approach to be used as a classification layer in a multi-layered approach like the one described in Chapter ??.

To test an object recognition algorithm designed to operate in underwater images, a submersible robotic vehicle with a camera sensor is often required. Most commercially available robotic vehicles are either expensive or inflexible to customization needed for research experiments. To address this problem, CoopROV, a low cost ROV was designed as a research prototype to evaluate object recognition algorithms. CoopROV carries an Inertial Measurement Unit (IMU), stereo cameras, depth and pressure sensors. The onboard electronics on the robot allow easy inclusion of new sensors to accommodate different research needs. CoopROV also offers a Robot Operating System (ROS) software interface, that allows to easily expand CoopROV's capabilities by using the plethora of existing ROS packages. The low cost and the flexibility to modify the software and hardware of CoopROV, positions it as an ideal tool to support object recognition experiments like the ones discussed in Chapter ?? and Chapter ?? of this dissertation.

The body of work in this dissertation comprises three object recognition frameworks validated on a specific object recognition problem. There are also details of a

ROV that was designed as a test bed to evaluate these algorithms. There are some possible directions to extend these object recognition efforts. For instance the scallop recognition effort in Chapter ?? could benefit from using more descriptive templates that capture the position of the bright crescent along with the dark crescent, that is currently being used. Unlike the dark crescent, or shadow cast due to the strobe light that always appears on top of scallop, the bright crescent can occur at any point on the rim of the scallop. The bright crescent seen here is the bright interior of the scallop that is sometimes visible. To capture such a feature, a series of templates with the bright rim showing at different points in the rim will be required. As for the multi-view object recognition method discussed in Chapter ??, a more expansive dataset could be used to test this algorithms. Another possible improvements of this multi-view approach is generalization to accept images from any series of specified heights instead of a fixed set of heights. Easing the restrictions on different views needed by the multi-view algorithm might make the experiments and data collection effort easier. This approach could also be extended to multi-class classification from the current binary classification. On the front of CoopROV, analysis and redesign of the acrylic frame to improve weight distribution could be helpful. Designing controllers and building a localization system are other avenues where future improvement on the robot can be made.

This dissertation offers object recognition tools designed to work on noisy natural images. Each of the tools developed in this dissertation 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 with minimal customization of the framework. 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. The over reaching goal of this dissertation is to contribute to the object recognition tools available to robots to strengthen their perception abilities. Better scene understanding and perception abilities ultimately allows robotic systems to be autonomous.

BIBLIOGRAPHY