

Chapter 1

INTRODUCTION

1.1 Outline

Chapter 1: If robotic systems are to decrease the strain on human workforce, they need to be autonomous; perception and data interpretation with in-built robustness to noise are central to autonomy.

Section 0: Intro

Section 1: Robotics systems are beneficial for handling tasks characterized as dull, dirty and dangerous for humans.

Para 1: Robotics systems are beneficial in environments hazardous or inaccessible to humans like war zones, nuclear radiation prone regions, and deep sea.

Para 2: When the reasoning ability and intelligence of a human expert is sometimes deemed necessary to deal with various elements in a potentially inhospitable environment, robots can still be tele-operated by human operators.

Section 2: Unmanned aerial, ground and underwater vehicles mostly rely on human decision making via teleoperation while working in natural environments.

Para 1: Tele-operation allows for a human operator to make decision for a robot in real-time.

Para 2: Though tele-operation solution can combine the durability of a robotic agent and intellect of a human, it has to deal with sometimes unreliable communication links and constant need for human attention.

Para 3: Automated decision making is essential for robots that do not have a human operator.

Section 3: To make informed automated decisions, robotic systems need to sense and interpret the state of their environments.

Para 1: Sensors allow a robotic system to gather data about its environment.

Para 2: Interpreting the state of an environment involves identifying objects and gaining a semantic understanding of different elements in the environment.

Para 3: Knowing the state of the environment is essential to make *informed* automated decisions.

Section 4: The capability of a robot to image and label objects in its surroundings, or in other words do object recognition, is more challenging in unstructured natural environments than in man-made environments.

Para 1: Robots in the past have been specialized to do tasks in well defined environments like assembly lines where environmental variables like visibility and lighting can be controlled.

Para 2: In natural environments, despite the its unstructured nature and unpredictable variations in environmental conditions, a robot needs to understand and parse the environment into labeled objects.

Para 3: To work in natural environments, the object recognition capabilities of robotic system needs to be robust to noise and variations in environmental conditions.

Section 5: Noise in sensor data obtained in natural environments limits the amount of useful information that can be extracted for the object recognition process.

Subsection 1: The sources of noise associated with an application need to be carefully analyzed before designing appropriate filters to mitigate the noise.

Para 1: Noise can enter sensor data through several forms, for example the noise in a camera sensor can be sub-divided as: noise inherent to

the sensor, noise due to the setup associated with data collection and noise introduced by the environment.

Para 2: To understand the noise inherent to a sensor, as an example, the various noise sources in a camera sensor can be analyzed.

Para 3: Noise can be introduced by the data-collection setup, in form of movement or vibration of the sensing apparatus.

Para 4: Environmental conditions can themselves be sources of noise, in the context of an underwater camera, this can be in form of poor lighting, light absorption by water or scattering from suspended particles in water.

Para 5: A method that is designed to work in natural environments, requires dedicated effort to filter noise to get useful information from sensor data.

Subsection 2: Once the noise sources are analyzed, dedicated filters, some specialized for the sensing apparatus and others specialized for the sensing environment need to be designed as a part of the object recognition procedure.

Para 1: Filters that deal with a specific sensing apparatus are required to work with the noise generated by the sensor.

Para 2: Specialized filters are required to deal with the noise from the environment conditions that robot will be exposed to.

Para 3: Analyzing and filtering noise is an important component of object recognition systems that work in natural environments.

Section 6: Existing object recognition methods do not generally translate from one application domain to another.

Para 1: Since object recognition is a challenging problem, researchers often specialize their object recognition modules to work in specific controlled environments to optimize performance.

Para 2: Object recognition methods designed for specific controlled environments do not fare well with wide variations in environmental conditions.

Para 3: Porting an object recognition method specialized for a specific environment condition to work on an entirely different application with drastically different environmental conditions can be a laborious process and might no longer offer the previous levels of performance.

Section 7: To work in natural environments, the object recognition process should be robust and sufficiently general to deal with various kinds of objects.

Para 1: To work in natural environments with unpredictable conditions, the object recognition techniques should be robust to noise and variations in environmental conditions.

Para 2: Multi-layered object recognition frameworks with dedicated modules for dealing with different sub-problems like filtering and hypothesis testing allow easy customization of layers based on application requirements.

Para 3: Layered Object recognition frameworks designed for natural environments with inbuilt robustness to variations in environmental conditions can be easily ported to a multitude of applications.

Section 8: Object recognition is central to building robots capable of automated and informed decision making, and thus achieving full autonomy; especially while operating in noisy natural environments.

Para 1: Object recognition capability allows a robot to interpret its environment.

Para 2: Informed decision making comprises using information about the state of the environment to making decisions.

Para 3: Noisy natural environments pose a challenge to object recognition and hence dealing with noise is critical for a robotic system the depend on object recognition to make automated decisions.

Para 4: Robust object recognition capabilities translates to autonomous decision making, this combined with autonomous actuation results in full robot autonomy.

1.2 Benefits of Robotic Systems

Robotic systems are useful for tasks that come under the purview of the “ $3D$ ’s”—dirty, dull and dangerous. Currently, manufacturing and assembly lines, where tasks are of repetitive nature, are places where robotic systems are ubiquitous. “Dirty” tasks like traversing tunnels or sewers are being slowly delegated to tele-operated robots. The other avenue where robotic systems are being promoted relates to activities deemed dangerous for humans, like deep sea exploration, operation in radiation, biologically, or chemically contaminated zones, or conflict areas. With enabling new advances in the field of robotics, robots are beginning to take over new roles to complement and assist humans in various ways.

Though advances in robotic systems have enabled robots to outperform humans in certain complicated tasks that are usually considered to require intelligence, such as competing in games like Go [1] and Chess [2], robotic systems still lack the ability to reason and operate autonomously in most unpredictable real world environments. In such cases, a human expert is deemed necessary to make decisions for the robot and guide the machine to accomplish its objectives. Another domain where tele-operation is prevalent is in operations over inhospitable environments. Some prime examples of such remote operation in dangerous environments include the rescue and repair effort at radiation affected zones in Fukushima Daichii nuclear power plant [3], or combat operations via remotely operated weapons like Packbot [4] and Predator [5] that have been deployed at war-zones in Iraq and Afghanistan. Though tele-operation appears to offer a solution, a more scalable alternative is aiming for full autonomy of robotic systems to minimize the strain and demands on human workforce.

1.3 Tele-operation

Tele-operation offers an avenue for humans to leverage the benefits posed by a robotic system to perform activities that are *dirty* or *dangerous* without being physically present on the site of operation. During tele-operation, the human operator typically operates the robot from a safe distance away from the point of activity of the robot. This safe distance could range from a few meters to over a few hundred kilometers, based on the nature of the activity. For instance, in case of a tunnel or sewer inspection, the operator can stay a few meters outside the structure and drive a robot that is traversing inside. On the other hand, in case of weaponized drones like the predator [5], the operator typically controls the drone from a safe remote command center, several hundred kilometers from the point of action of the drone. Tele-operation allows the operator to process the sensory information collected by the robot and make *informed decisions* for the robot. The robot then translates an *informed decision* into a sequence of actions. The tele-operation solution allows a human operator to safely accomplish an objective without being in discomfort or risk.

Tele-operation brings together durability of a robotic agent and human intelligence to provide a powerful solution to solve problems ranging from deep sea exploration to safe operation in combat zones. Despite its attractiveness, tele-operation still requires the constant attention and involvement of a human operator. Additionally, since the human operator is not present, a reliable communication channel is essential to remotely control the robot. If the communication links are unreliable, or subject to willful sabotage by an enemy in a war-zone, the robotic system can get disconnected from the human operator. Such communication interruptions could possibly result in mission failure. The constant need for a human operator and the need for reliable communication channels are some of the shortcomings that limit the applicability of tele-operated systems.

If the objective is to move towards decreasing the strain on human workforce, it is imperative to decrease the reliance of robots on humans for operation. This emphasizes the need for automated decision-making in robotic systems. Furthermore,

building automated, decision-making capable robotic systems obviates the need for reliable communication channels. Thus improving the cognitive capabilities to enable automated decision-making is essential to build robust self-sufficient robotic systems.

1.4 Components of Automated Informed Decision-Making

The first component of informed decision-making is gathering the information needed to make a decision. For a robotic system, information about the environment and the objective it needs to accomplish, dictate the decisions it makes. Collecting sensory data, is the prime and often the only mechanism, available for a robot to learn about its environment. Based on the type of sensors, different measurements describing the state of the environment are available. These raw sensor data need to be processed to get actionable information that enable a robot to reason about the state of its environment. Then decision may be seen as extrapolation of the reasoning process to choose an action from the set of actions available to a system. If the entire sequence, from sensory data aggregation to choosing an action based on processed information, is performed by a robotic system without any external intervention, then the due process can be characterized as automated, informed, decision-making.

Transforming sensory data into usable information involves parsing the data to gain a semantic understanding about the state of an environment. The basic stage in understanding a scene¹ is separation and labeling of different objects present. A more refined version of this understanding is deciphering the relationship between the objects in a scene. Such a detailed semantic understanding allows a robot to infer the ability of the scene elements to actively participate in the robot's planned actions.

The ability of a robot to detect and label objects in its environment is an important step towards gaining semantic understanding of its environment. Since semantic understanding of a scene dictates the ability of a robotic system to make informed decisions, object recognition capability is central to automated informed decision-making.

¹ Scene here is defined as the state of an environment at a given point in time.

1.5 Automated Decision-Making in Natural Environment

Building robotic systems that are fully autonomous, with an ability to make informed decisions based on their observations under all possible environmental conditions, is an open problem. The existing solutions typically apply to certain specialized domains only. Before further discussion on this topic, it is important to understand that natural environments pose significant challenges that compound the problems faced by robotic systems operating in other regimes. Section 1.5.1 states the definition of natural environment in the context of this dissertation and, Section 1.5.2 goes into the details on the challenges faced by robotic systems in natural environments.

1.5.1 Natural Environment

The word *Natural Environment* in this dissertation refers to uncontrolled and unstructured real world environments, with scenes containing predominantly naturally occurring objects; the robot has no direct control over any environment parameter. (However we assume that the robot can use accessories, like artificial light sources, to influence the state of the environment and thereby enhance sensor measurements.) Natural environments are typically challenging to model due to the unpredictability in the environment parameters and unknown noise parameters. Deep sea or forest landscapes with vegetation are some prime examples of such natural environments.

1.5.2 Robotic Systems in Natural Environments

In the past, robots have been specialized to do tasks in well defined environments, like assembly lines, where environmental variables (e.g. visibility and lighting) are strictly controlled. If such environment variables are known in advance, a mathematical model that maps robot actions to finite known environment states can be built. In such cases, based on the state of the environment observed, a robot action that complies with a predefined database of logical rules is chosen. Systems that operated on this principle, sometimes referred to as *Expert Systems* [6], came into existence in 1980's. Such systems fail in scenarios where the environment cannot be perfectly

modeled. Despite the advances in field of robotics since then, robotic systems still face challenges when dealing with unpredictable environments similar to those in unstructured natural settings found in forests or deep sea.

The task of understanding natural scenes and recognizing objects like animals from them is shown to be a cognitively challenging task even for humans [7]. The unstructured nature and unpredictability associated with natural environment makes the task of building models to capture natural scenes challenging. Using global features to characterize the nature of a scene is, in general, possible [8]. However, these models can only offer a high level understanding of the scene. For instance, the high level understanding translates to differentiating between widely different scenes like a beach, foliage or a busy city street. It does not inform about separating the different components or objects in those scenes. Most object recognition techniques depend on features like edges or texture to identify such objects. Such feature-based approaches are difficult to implement in natural scenes where the edges of objects can be difficult to distinguish. The measurement noise in such natural scenes further complicates the object recognition task. Thus, for a robotic system to be successful in making automated informed decisions in natural environments, it has to possess robust object recognition capabilities that can handle noise and variations in environmental conditions.

1.6 Noise in Sensor Measurements

Sensors are designed to measure a physical quantity. This physical quantity is captured by the sensor in form of a signal. The measurement reported by the sensor almost always does not exactly match the real value of the physical quantity. This error in the measured value is considered as measurement noise. The ubiquitous presence of noise in almost all sensor measurements has spurred a field of study in the name of *Filtering* [9]. Filtering deals with the estimation of signals from noisy data. If the source and nature of noise that affects measurements of a sensor can be somehow modeled, then specialized filters can be designed to recover the underlying signal from noisy sensor measurements. The challenge here is in identifying the noise sources and

modeling them. The sources of noise in sensor data can be manifold. An exhaustive discussion on sources of noise and methods to model them is beyond the scope of this dissertation. In the remainder of this section, we will focus on specific sources of noise that affect imagery data collected by robots operating in natural environments (Section 1.6.1) and some filters that can help recover the underlying signal in these cases (Section 1.6.2).

1.6.1 Noise Sources

The sources of noise associated with a particular application need to be carefully analyzed before designing appropriate filters to mitigate this noise. In the specific case of imaging applications that employ a camera sensor in natural environments, the various noise sources can be broadly divided into (i) noise inherent to the sensor, (ii) noise associated with data collection setup, and (iii) noise introduced by the environment. For instance, in the case of an underwater imaging application using an [Autonomous Underwater Vehicle \(AUV\)](#), the noise associated with the imaging setup constitutes noise introduced due to vibration in the motion of the [AUV](#). Additionally, the environmental parameters associated with deep sea environments could introduce noise which could further degrade the sensor measurements. Further details about the different noise sources of this type is discussed in the rest of this section.

In an imaging application, some noise is inherent to the camera sensor. Studies [10] have analyzed the various constituents of noise in a [Charge-Coupled Device \(CCD\)](#) video-camera. According to these studies, the sources of noise in a camera can further be divided into three categories, namely: illumination-independent noise, illumination-dependent noise, and digital processing noise. After further analysis, this work suggests that there are four major noise sources that contribute to the camera sensor noise. Out of these four sources, *readout noise* and *fixed pattern noise* are constant illumination-independent noise sources. One of the factors that determine the value of readout noise and fixed pattern noise is the temperature of the sensor. The remaining two primary contributing noise sources, *photo-shot noise* and *photo-response non-uniformity noise*

are illumination-dependent noise sources. Its important to note that, in illumination dependent noise sources the intensity of noise is a function of the intensity value of the pixel. Modeling all the noise sources for different camera sensors poses a challenge in itself.

Another source of noise is the physical setup which holds the camera. This setup could be an [AUV](#) or just a stationary imaging rig. Such physical structures that hold the camera can introduce noise in form of vibration or motion of the structure during imaging. Such motion could lead to out-of-focus or blurred pictures. There could also be cases where the physical setup can produce artifacts like shadows that could affect the sensor measurements. The few examples mentioned here are not all encompassing. Each imaging setup needs to be studied to identify the possible sources of noise it can introduce into sensor measurements.

Environmental conditions are one of the chief sources of noise. There is a key difference between noise due to environmental conditions and other sources like sensor noise or the imaging setup noise: the former is relatively more difficult to model compared to the latter due to wide range of variables associated with environment. These environmental variables are fairly limited in controlled environments, but in case of natural environments the number of variables are far higher. Furthermore, the ability to predict or model the environmental variables in natural environments is far more challenging compared to controlled environments. To understand some of the variables associated with natural environments, let us consider the case of imaging experiments in underwater environments. Some of the variables at play here are non-uniform low illumination, absorption of light by water, and scattering of light by different particles in water. Each of these parameters can have considerable impact on the quality of the image acquired by a camera. Modeling these environment variables and accounting for the noise induced by them in the sensor measurements is critical to extract useful information from the sensor data.

Any object recognition method has to account for the different noise sources

that can affect sensor measurements. Especially in case of natural environments, environmental conditions typically act as major noise sources. In any case, it is important to take appropriate steps to mitigate the effect of noise in the sensor measurements before useful information can be extracted. Even in the absence of accurate models of different noise sources, it is important to attempt to model the noise seen in the sensor data and devise filters to extract the underlying signal from recorded sensor measurements.

1.6.2 Noise Filtering

The sensor noise model is often a part of the sensor specification provided by the manufacturer. Thus the person integrating the sensor into a custom application, can sometimes avoid modeling the sensor noise. However, the noise induced by the setup is unique to each system and hence needs to be analyzed to identify the characteristics of noise. Once a model is identified, the subsequent filter design involves designing a filter that can recover the signal from recorded sensor measurements.

The task of filtering noise induced by the environment is more complex relative to that of sensor or setup. The reason for this is the difficulty in modeling environmental variables. In the case of underwater imaging applications, understanding the physics of light propagation, that includes phenomena like absorption and scattering, is required. Some proposed models [11, 12] try to capture light propagation in water and also suggest filter models that can correct illumination and color in noisy underwater images. A more detailed survey by Jaffe [13] describes different existing filter models available to deal with noise in underwater images. In most cases, modeling all the environment parameters is intractable. Hence identifying the dominant noise source or simply identifying the nature of noise affecting the images are possible alternatives. For instance, in scallop recognition [14], the AUV images used were corrupted by speckle noise, hence a median filter [15] was used to minimize noise without elaborate modeling of noise sources.

Since extracting useful information from sensor data is essential for object recognition, noise filters constitute a core component of object recognition systems. The first stage in noise filtering involves identifying noise sources and modeling them. Accurate modeling of noise sources is often challenging, hence most systems use simplified models that only capture dominant noise sources. Since there are several filtering mechanisms available, filter design often involves human intuition to pick appropriate filters that can handle the noise pattern seen. Ultimately, noise filtering allows distilling useful information for informed decision-making.

1.7 Object Recognition in Natural Environments

For an object recognition framework to be effective in natural environments, it should be robust to noise and variations in environmental conditions. To avoid dealing with uncertainties in environmental conditions, researchers tend to specialize their object recognition technique to work in certain controlled conditions. In some instances even the sensing apparatus used can be explicitly designed to detect a certain organism as in the case of plankton recognitions systems [16, 17]. Such techniques, that are highly specialized to solve a specific problem, often do not transfer to other application domains. An example of object recognition in controlled conditions is the scallop recognition system designed to work in artificial scallop beds [18, 19]. In this method, a series of stationary cameras under known conditions are used to perform object recognition. The object recognition systems built on the assumption of controlled environmental conditions, generally do not translate well to natural environments that exhibit wide variation in environmental conditions.

Building specialized object recognition frameworks that translate between different application domains is challenging. At the same time, it is equally challenging to build a generalized object recognition method that can accommodate wide variations in environmental conditions. One approach to this problem is the use of *multi-layered* object recognition frameworks. Multi-layered object recognition frameworks with dedicated modules for dealing with different sub-problems, like filtering and hypothesis

testing, allow easy customization of layers based on application requirements. Thus multi-layered frameworks allow researchers to configure specialized layers that are custom designed for their object recognition problem. The scallop recognition framework [14, 20] offers one such multi-layered framework. Another interesting example is the set of underwater object recognition tools [21] used for recognizing organisms like sea-anemones.

In conclusion, multi-layered object recognition frameworks offer a flexible solution to automated object recognition problems. Customizing a multi-layered object recognition framework to work with a specific problem amounts to picking the right layers that would be effective in solving the problem. The layers used can be adaptations of existing computer vision or machine learning tools. For cases where there is no such off-the-shelf solution, specialized layers can be engineered to fit the problem. A researcher can thus build a multi-layered framework with existing tools in tandem with custom designed solutions to solve specialized object recognition applications.

1.8 Object Recognition as a path to Robot Autonomy

Object recognition capabilities allow a robotic system to identify and label the different components in its environment. Such understanding of individual objects in the robot’s surrounding is required to build a semantic model that captures the relationship between the different components in an environment. Cognitive reasoning relies on such semantic models to determine the best possible action for a robot given its knowledge about the state of the environment. This is why object recognition capabilities are essential to building robotic systems capable of automated informed decision-making.

The task of object recognition and understanding the state of the environment becomes challenging in the presence of noise. Sensor measurements that are supposed to measure the state of the environment could be corrupted due to the presence of noise. In order to avoid being misled by bad measurements, a robotic system needs to employ noise filters to extract the underlying signal from sensor measurements.

Noise filters become a core component of the perception system that aggregates sensor measurements to gain knowledge about the state of the environment. Building accurate models of the surroundings despite the presence of noise is key to the success of a robotic system that is intended to operate autonomously in natural environments.

In conclusion, object recognition capabilities, with robustness to noise, are required for building systems that operate in natural environments. Object recognition, coupled with an ability to model and reason about the state of an environment, allows a robotic system to make automated informed decisions. Automated informed decision-making, in conjunction with autonomous actuation, paves the way for robot autonomy. Autonomous systems can act independently without any human support in a range of environments. Such systems can ease the strain on human work force by taking up tasks that are physically challenging or risky for a human to perform.

1.9 Approach Overview

Given that object recognition in noisy environments is critical for robot autonomy, this dissertation proposes a series of techniques designed to address the object recognition problem in noisy environments. We choose underwater object recognition as the primary domain to conduct the experimental validation for this approach, since operation in natural environment is typically characterized by noisy sensor measurements. The rest of this section offers a brief overview of the different object recognition techniques developed as a part of this overall approach, along with insights on the particular problems they address.

Recognizing submerged subway cars from seabed images was the first application that motivated the development of the need for an object recognition technique capable of operating in noisy seabed images. Since recognizing subway cars can be reduced to detecting rectangles, eigen-value based shape descriptors with in-built [Rotation, Scaling and Translation \(RST\)](#) invariance, were tested to solve this specific object recognition problem. Eigen-value based shape descriptors were successful in detecting simple shapes like rectangles from images. During the course of the study, several

shortcomings of eigen-value shape descriptors were exposed: poor performance in noisy domains, and inability to identify objects characterized by complex shape profiles given constraints on image discretization. Chapter ?? of this dissertation offers a detailed treatment of the eigen value shape descriptors, and the results of the subway car recognition problem they were tested on.

When the scope of the object recognition problem was expanded from simple shapes like rectangles to underwater organisms characterized by complex profiles, eigen-value shape descriptors were no longer a viable choice. The primary application driver for the design of the multi-layered object recognition system in Chapter ?? was the need for a scallop recognition system in image-based benthic survey efforts capable of automated scallop enumeration. The multi-layered object recognition system thus designed was capable of recognizing objects from noisy natural images. This method detected 60-70% of scallops, present in a dataset of over 8000 images. One important point here is that the method was explicitly designed to deal with noise introduced by non-uniform lighting, low resolution, and large levels of speckle noise, that previous scallop recognition efforts were not equipped to handle. Chapter ?? elaborates on the multi-layered object recognition framework developed to meet this challenge.

Despite this multi-layered object recognition framework proposed in Chapter ??, being able to detect objects from noisy images, there were several instances of false positives in the ensuing detections. This prompted the need for a classification technique that can recognize objects with smaller percentage of false positives. To accomplish this we merged information from multiple views of an object before determining its identity. The details involved in this multi-view object recognition approach are discussed in Chapter ??.

During the development of these object recognition algorithms, the need for a low-cost research platform that can be used as a test bed to evaluate different object recognition algorithms was realized. In response to this, CoopROV, a low cost underwater [Remotely Operated Vehicle \(ROV\)](#) was developed. CoopROV's specifications and design procedure is discussed in Chapter ??.

1.10 Dissertation Organization

The rest of this dissertation is organized as follows. Chapter ?? discusses the eigen-value shape descriptors and the subway car detection problem they were evaluated on. The multi-layered object recognition technique developed to support an automated scallop survey effort is discussed Chapter ?. This is followed by Chapter ?, that describes a multi-view object recognition technique that allows to combine information from multiple views of a target object, and decreases the number of false positives seen in the multi-layered object recognition technique. Chapter ? provides details on the underwater ROV designed as a research prototype to test object recognition algorithms. Finally, Chapter ? highlights the insights gleaned during the development of the different object recognition techniques along with possible directions that can be explored in the future.

Acronyms

AUV Autonomous Underwater Vehicle. [10–12](#)

CCD Charge-Coupled Device. [10](#)

ROV Remotely Operated Vehicle. [16](#), [17](#)

RST Rotation, Scaling and Translation. [15](#)

BIBLIOGRAPHY

- [1] D. Silver, A. Huang, CJ Maddison, A. Guez, L. Sifre, GVD Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, and M. Lanctot. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
- [2] M. Campbell, AJ Hoane, and F. Hsu. Deep blue. *Artificial Intelligence*, 134(1):57–83, 2002.
- [3] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, and M. Fukushima. Emergency response to the nuclear accident at the fukushima daiichi nuclear power plants using mobile rescue robots. *Journal of Field Robotics*, 30(1):44–63, 2013.
- [4] BM Yamauchi. Packbot: a versatile platform for military robotics. In *Defense and Security*, pages 228–237. International Society for Optics and Photonics, 2004.
- [5] Predator MQ-1B. <http://www.af.mil/AboutUs/FactSheets/Display/tabid/224/Article/104469/mq-1b-predator.aspx>. Accessed: 2016-10-20.
- [6] SJ Russell, P. Norvig, JF Canny, JM Malik, and DD Edwards. *Artificial intelligence: a modern approach*, volume 2. Prentice hall Upper Saddle River, 2003.
- [7] FA Wichmann, J. Drewes, P. Rosas, and KR Gegenfurtner. Animal detection in natural scenes: critical features revisited. *Journal of Vision*, 10(4):6–6, 2010.
- [8] Aude Oliva and Antonio Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3):145–175, 2001.
- [9] SS Haykin. *Adaptive filter theory*. Pearson Education India, 2008.
- [10] K. Irie, AE McKinnon, K. Unsworth, and IM Woodhead. A technique for evaluation of ccd video-camera noise. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(2):280–284, 2008.
- [11] R. Garcia, T. Nicosevici, and X. Cuf. On the way to solve lighting problems in underwater imaging. In *OCEANS’02 MTS/IEEE*, volume 2, pages 1018–1024. IEEE, 2002.

- [12] J. Ahln, E. Bengtsson, and D. Sundgren. Evaluation of underwater spectral data for colour correction applications. In *Proceedings of the 5th WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing*, pages 321–326. World Scientific and Engineering Academy and Society (WSEAS), 2006.
- [13] JS Jaffe. Underwater optical imaging: the past, the present, and the prospects. *IEEE Journal of Oceanic Engineering*, 40(3):683–700, 2015.
- [14] P. Kannappan, JH Walker, AC Trembanis, and HG Tanner. Identifying sea scallops from benthic camera images. *Limnology and Oceanography: Methods*, 12(10):680–693, 2014.
- [15] GIMP despeckle plugin. http://oldhome.schmorp.de/marc/pdb/plugin_despeckle.html. Accessed: 2016-10-23.
- [16] C. McGavigan. A quantitative method for sampling littoral zooplankton in lakes: The active tube. *Limnology and Oceanography: Methods*, 10:289–295, 2012.
- [17] CP Stelzer. Automated system for sampling, counting, and biological analysis of rotifer populations. *Limnology and Oceanography: Methods*, 7:856, 2009.
- [18] K. Enomoto, M. 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.
- [19] K. Enomoto, M. Toda, and Y. Kuwahara. Extraction method of scallop area in gravel seabed images for fishery investigation. *IEICE Transactions on Information and Systems*, 93(7):1754–1760, 2010.
- [20] P. Kannappan, HG Tanner, AC Trembanis, and JH Walker. Machine learning for detecting scallops in AUV benthic images: Targeting false positives. *Computer Vision and Pattern Recognition in Environmental Informatics*, pages 22–40, 2015.
- [21] T. 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*. PhD thesis, Universität Bielefeld, 2015.