2012-07-31

Topographic Reconstruction Through a Wavy Interface Using a Structure Light Technique

Amin Sarafraz
University of Miami, sarafraz@gmail.com

UNIVERSITY OF MIAMI

TOPOGRAPHIC RECONSTRUCTION THROUGH A WAVY INTERFACE USING A STRUCTURED-LIGHT TECHNIQUE

By

Amin Sarafraz

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2012

UNIVERSITY OF MIAMI

A DISSERTATION submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

TOPOGRAPHIC RECONSTRUCTION THROUGH A WAVY INTERFACE USING A STRUCTURED-LIGHT TECHNIQUE

Amin Sarafraz

Approved:

Brian K. Haus, Ph.D.
Associate Professor of Applied Marine Physics

M. Brian Blake, Ph.D.
Dean of the Graduate School

Antonio Nanni, Ph.D.
Professor of Civil, Architectural and Environmental Engineering

Kenneth J. Voss, Ph.D.
Professor of Physics

James W Giancaspro, Ph.D.
Associate Professor of Civil, Architectural and Environmental Engineering

SARAFRAZ, AMIN (Ph.D., Civil Engineering)
Topographic Reconstruction Through a Wavy Interface Using a Structured Light Technique (August 2012)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Brian K. Haus.
Number of pages in text: (103)

Three dimensional (3D) surface imaging has been an active area of research for decades and has reached a stage of maturity, thanks to digital cameras and ever-increasing computational power. Its utilization in science and technology is rapidly growing due to its potentials in civil engineering, mapping, structural monitoring and inspection, archaeology, and cultural heritage documentation, to name a few. In this research we will focus on development of a 3D surface imaging technique for applications in a special environment in civil engineering, i.e. underwater.

Some of the most important methods of 3D surface imaging are based on the use of structured light. These techniques simplify the problem of 3D reconstruction with the help of controlled illumination and can yield excellent results in practice. They basically consist of acquiring 2D image(s) of the scene while illuminating the scene with spatially varying intensity pattern(s). The nonplanar geometric shape of the surface distorts the projected structured-light pattern as seen from the camera. The principle of structured-light 3D surface imaging techniques is to extract the 3D surface shape based on the information from the distortion of the projected structured light pattern.

In this dissertation, a new structured-light method for underwater 3D surface imaging has been developed which can simultaneously estimate both the geometric shape of the water surface and the geometric shape of underwater objects. The method utilizes a projector outside of water and an underwater camera. It requires only a single image and thus can be applied to dynamic as well as static scenes. The proposed method is specifically applicable to 3D reconstruction for structural monitoring in hazardous areas that are inaccessible to direct inspection. Experimental results show the utility of this method in non-invasive underwater 3D reconstruction applications. The performance of the new method is studied through a sensitivity analysis for different parameters of the suggested method.

DEDICATION

To my parents, Khalil and Mahjan

ACKNOWLEDGMENTS

I would like to express my gratitude and appreciation to my advisor, Dr. Brian Haus for persevering throughout the time it took me to complete this research and write the dissertation.

I am also grateful for all the good-natured support of Dr. Antonio Nanni, the co-chair of my committee and the head of the CAE department.

The members of my dissertation committee, Dr. Kenneth Voss and Dr. James Giancaspro, have generously given their time and expertise to improve my work. I thank them for their constructive discussions and contributions.

My thanks must go also to Lizett M. Bowen and Maria L. Aldana, the secretaries of the CAE department for all their patience and help throughout my studies at the University of Miami.

I would like to acknowledge Mike Rebozo for assisting me to set up all the experiments in the lab. Without his expertise it would have been impossible for me to collect data for my research.

I would also like to acknowledge Kristin Ponge for all the time, assistance and support she gave me during my PhD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1

Introduction

1.1 Preface

Understanding the three-dimensional structure of surrounding objects is an essential component in many civil engineering applications. Examples of these applications include topographic mapping, Non Destructive Testing (NDT), navigation, 3D virtual cities, and Geographical Information Systems (GIS), among others.

Analysis of visual information is a powerful apparatus which enables engineers to tackle the problem of understanding (reconstructing) the three dimensional structure of objects. Humans along with a large number of animals mainly "understand" the three-dimensional structure of the surrounding world through their visual system. It is believed that this task is achieved by integrating a number of visual cues (perspective deformation, stereo, shading, shadows, occluding contours, (de)focus, texture, motion parallax, highlights) in conjunction with prior knowledge about the scene. In many circumstances, even a single image of an object is sufficient for humans to extract a good mental representation of its three-dimensional shape.

Due to the richness of visual information, the problem of reconstructing the three-dimensional structure of objects from a set of two-dimensional images has always been of interest to researchers and professionals in civil engineering [1, 2, 3, 4, 5].

For example, for the purpose of topographic mapping, photogrammetric techniques have been used for decades to derive three-dimensional data from stereo images. Photogrammetric 3D reconstruction is a common practice specially in large-scale mapping projects. Almost all of the national topographic maps are produced from aerial or satellite images. 3D maps serve as the foundation for applications such as 3D virtual cities, GIS, navigation, etc.

Non Destructive Testing (NDT) is another example of an application area where obtaining 3D information from images is very common. NDT has been defined as comprising those test methods used to examine an object, material or system without impairing its future usefulness [1]. NDT embraces a variety of methods, including visual testing, radiography, ultrasonic, liquid penetrant inspection, and thermography, to name a few. Visual testing is the first natural choice for Non Destructive Testing, mainly due to its low cost, minimal equipment requirements, massive amount of information in visual contents and the ability for sampling to be performed while the subject is being used or processed. Visual testing can refer to both human and machine operated visual testing. By applying computer vision techniques such as 3D reconstruction, image segmentation and/or classification, visual inspection can be implemented efficiently, automatically or semi-automatically with certain interactive manipulations by the operator. Once the images are taken, appropriate computer vision techniques can reveal potential defects within the component under investigation.

Optical imaging, as a tool for visual testing, has a number of advantages including acquiring information of a component without direct contact and its capability of instantaneous recording of two-dimensional images of the objects is important for investigating dynamic behaviors of a component.

The limitations to visual testing include detection of only surface discontinuities and high equipment costs of aids for some visual testing. In case of human operated visual testing, other limitations such as poor or variable resolution of the eye, fatigue of the inspector, and distractions will be added.

The above examples shed a light on how and why optical imaging and computer vision technology became a standard aid in civil engineering and play an important role in different applications in this area. In this research we will focus on a special environment in civil engineering, i.e. underwater.

Civil engineers quite often deal with underwater environments such as oceans, rivers, and lakes. For instance, the directions, heights and surface slopes of three dimensional sea waves are highly important for the design of offshore and nearshore structures, as well for the development of an understanding of momentum, mass and heat transfer across the air sea interface [6, 7].

Also, underwater visual inspection is the integral part of monitoring underwater structures or disaster management in oceans or seashores. Some recent examples in this regard include two recent major disasters, i.e. :

- The reactors in Japans nuclear power plant in 2011 were flooded with seawater (an extremely hazardous area that required minimal human exposure).

- British Petroleum oil spill in the golf of Mexico 2010 (5000 feet below the surface of the water in almost absolute darkness). Visual estimation of petroleum outfall was an important component of understanding the extent of the problem.

In both of above-mentioned disasters, Remotely Operated Vehicles (ROV) and/or Autonomous Underwater Vehicles (AUV) equipped with optical cameras were deployed for large scale and long term data collection tasks. Since these systems could be deployed in such hazardous environments without risking human divers, they played an essential role in structural monitoring and disaster management operations.

Furthermore, 3D information can be very helpful in managing underwater environments and plays a crucial role in watershed management and coral reef conservation. According to the Global Coral Reef Monitoring Network in 2008, the world has lost 19% of the original coral cover and 15% of the worlds reefs are under short term threat of major losses in 10 to 20 years. In 2011, the World Resources Institute factored climate change into their predictions to state that more than 60% of the worlds reefs are under immediate threat of direct human pressures and this will rise to 95% by 2050 [8]. This effect is mainly due to damaging activities by humans in watershed areas, combined with over-fishing and global climate change. Good quality 3D maps are key components of GIS technology for monitoring these

precious coastal ecosystems. It is highly important to obtain 3D information with minimum disturbance of these environments and optical imaging can play a key role in collecting accurate 3D information for this purpose.

Another emerging application of optical sensors and vision systems is underwater security and surveillance. A video camera mounted on an underwater vehicle provides lots of information that can be examined for effective path planning and navigations as well as object detection and classification. Also, video cameras are high resolution sensors and are invaluable in situations that require accurate measurements at short ranges. Governments and private security and surveillance agencies are interested in using optical cameras for the purpose of underwater mine detection, harbor patrol, counter-surveillance, and anti-terrorism. They are required to continuously protect ports, offshore oil platforms, coastal energy terminals, high value vessels, naval bases, energy installations, and VIP compounds around the world against underwater intrusion and sabotage. With image-based systems, security personnel can view underwater structures for abnormal changes in conditions or be on the lookout for sabotage or terrorist interference.

Finally, recent advances in photogrammetry and computer vision, mainly in automatic matching and dense map generation, will change archaeological excavations that are often irreversibly destructive. Since underwater sites can never be experienced firsthand by the majority of archaeologists or the general public, it is crucial to provide a faithful and accurate 3D map and detailed documentation of such a site which can provide virtual access to all archaeological data. As

soon as appropriate tools for underwater photography became available, a trend towards obtaining 3D maps of underwater sites using optical imaging also developed. Using optical imaging and photogrammetric techniques, archaeologists are now able to obtain accurate 2D or 3D representations of underwater sites and ship wrecks. Mapping, archaeological surveys, and more generally, surveys dedicated to cultural heritage are an interpretation of the real world and underwater imaging is an intrinsic part of these tasks.

1.2 Different approaches for 3D reconstruction using visual information

Given the capabilities embedded in visual information and the ever-increasing demand for 3D reconstruction from images, it is highly desirable to design a system that would replicate the function of the human visual system. This ideal artificial system would be able to automatically extract the three-dimensional geometry of the objects using only the information contained in a set of 2D pictures of objects. For the purpose of civil engineering, in most cases the 3D reconstructed geometry should even be more accurate than human perceptual understanding of the 3D world. Although, we are still far from such an ideal system, tremendous research efforts have been conducted in this regard. These research efforts resulted in several different solutions. Existing Image-based 3D reconstruction systems can be categorized into two categories: passive and active.

Passive systems use only images acquired in natural (unstructured) light for computing 3D geometry. They employ different cues that contain information

about 3D shape such as stereoscopic disparity, texture, motion parallax, (de)focus, shadows, shading, specularities, occluding contours and other surface discontinuities [4, 9, 10, 11, 12, 13, 14, 15, 3]. Among these cues, stereoscopic disparity reliably gives reasonable accuracy. However, its main drawback is that it can not be used on untextured surfaces.

A single moving camera can substitute for the pair of cameras in stereoscopy. This is the case that commonly happens in conventional photogrammetric mapping using aerial photos. The camera takes two (or more) snapshots of the world from different locations in space at two (or more) different instants in time. The reconstruction procedure is then identical to traditional stereo scenarios: the structure of the world is triangulated using the two camera images used as a stereo pair. The main advantage of these methods is that the camera can be moved freely by the user in the scene and hence is more ergonomic. There exists, however, two major limitations of such an approach [16]: First, in order to perform geometrical triangulation, one should determine the motion disparity between the two camera positions. Second, since the two pictures are taken at different times, the objects and optical transmission properties must remain rigid between those two acquisition times.

In contrast to passive systems, active systems use an external projecting device to illuminate the scene. They employ a combination of imaging devices (one or more cameras) and lighting devices (laser/LCD projectors) calibrated one with respect to each other. The principle of their underlying methods is quite intuitive:

the projector emits light patterns (structured lights) that are reflected by the scene and detected on the image acquired by the camera. The three-dimensional structure of the scene is then computed by geometrical triangulation. Aside from being quite intensive to variations in texture within the scenes, these technologies have the advantage of yielding very good accuracies (errors in reconstruction can be as small as a part in one or two thousand).

A variety of such active systems (structured light techniques) have been developed for applications in the underwater environment (or participating media in general). They range from using thin laser beams [17], to using a number of carefully spaced confocal sources [18], to more sophisticated time-gated [19] and synchronization-gated techniques [20]. An implicit assumption made in these systems is that they are immersed in a clear medium and therefore the light is neither scattered nor absorbed by the medium in which the scene and sources are immersed.

Contrary to the above methods, Narasimhan et.al [21] introduced an analytic image formation model that is based on the physics of single scattering. The model describes the interactions of structured light with the medium and the scene. Using the image formation model for light striping, they developed an algorithm to detect objects and obtain a 3D reconstruction of the scene in the presence of single scattering.

1.3 Objective

In all of the above-mentioned structured light techniques, the assumption is that the projector and the camera are in the same (homogeneous) medium, e.g. both camera and projector are placed underwater or in an open-air environment. However, there are applications that need the projector and the camera to be in different mediums. For example, consider air-sea interface studies that need simultaneous estimation of the water surface and sea floor in shallow waters.

In the context of this dissertation, we mainly focus on the problem of estimating the 3D geometry using active lighting in underwater environments. With that specific goal in mind, we will present a technique for estimating the 3D shape of underwater objects as well as the water surface. Specifically, we seek to develop a structured-light technique that employs a projector above water and a camera located under water to estimate both the water surface elevation and the topography of the sea floor. In this case, the water surface is dynamic and changes instantaneously. Because it must deal with a dynamic surface, the method has to use a single image. For this purpose, we project a unique bi-dimensional pseudorandom color pattern from the projector that encodes positions as spatially distributed codes. In this pseudorandom code, any subset (or window) is unique and defines a specific location. Each code is composed of a 3x3 neighborhood of square color patches that ensures simplicity in the pattern generation and reduces the ambiguity in the color recognition.

An optical camera is used to acquire a single image of the underwater scene illuminated by the structured-light (pseudocode) pattern. The 3D surface shape of the water surface and underwater objects will be estimated based on the information from the distortion of the projected pattern.

We assume that the water is clear and the scattering effect is negligible. Therefore, the method is suitable for pure sea water and/or a clear ocean (see [22] for definition of different water types). If the scattering effect dominates the absorption, the method still can be used but the scattering effect needs to be removed initially from the images.

1.4 Outline

The manuscript is organized in the following chapters.

Chapter 2 reviews major image based techniques for water surface reconstruction. In addition this chapter provides a review of recent advances in non-contact 3D surface measurement techniques based on structured light. It is beyond the scope of this dissertation to cover all possible 3D surface imaging techniques. Instead, we have selected representative techniques in order to understand fundamental technical principles and typical system characteristics.

Chapter 3 describes the proposed system for underwater object reconstruction based on a structured light technique. We introduce the fundamental notation used in this dissertation. This is followed by a description of the camera-projector calibration procedure which is critical to the successful operation of any structured-light 3D surface imaging system. In this chapter we also introduce a new simple

mathematical formalism that enables us to explore and compute geometrical properties of three-dimensional scenes. The main contribution of this work is a new parametrization of underwater object reconstruction with minimum contact with water that leads to a set of useful properties.

Chapter 4 presents the results of several laboratory experiments used to define the ability and the effectiveness of the proposed method in both synthetic and real scenarios. The experiments were conducted with wavy and flat water surfaces. Both planar and non-planar objects were reconstructed to assess the performance of our method.

Finally, in chapter 5 we conclude, and discuss some directions for future developments.

2

Previous Works

Three-dimensional surface reconstruction deals with measurement of the x,y,z coordinates of points on the surface of an object. The result of the measurement may be regarded as a map of the depth z as a function of the position x, y in a Cartesian coordinate system. Applications of 3D surface reconstruction range from traditional mapping, structural monitoring, cultural heritage documentation, quantifying landform change to forensic, medical, biology and many other fields.

Three dimensional surface imaging has been an active area of research for decades and has reached a stage of maturity, thanks to digital cameras and ever-increasing computational power. We will discuss optical imaging techniques that have been developed for both sea surface reconstruction and underwater object 3D reconstruction using structured-light techniques.

2.1 Water surface reconstruction

Under the assumption of statistical stationarity or homogeneity, various temporal or spatial methods have been proposed to recover shape information of water surface waves. These methods are generally applied to time series obtained at a chosen location. Among the methods in common use are: the Direct Fourier Component method [23]; the Maximum Likelihood method [24, 25]; the Maximum Entropy method [26] and a method based on Wavelet transforms [27]. Time series observations generally obtained using Heave-Pitch-Roll buoys or its more recent

successor, the three axis accelerometer buoy, three component current meters, GPS sensors or gauges. All of these techniques are single point observations and can not reveal the instantaneous 2D structure of the surface. Arrays of point gauges or scanning lasers can provide such information but require significant structures that may disturb the surface. An alternative to these approaches is to use optical techniques.

It has long been recognized that optical techniques are useful for the measurement of water surface slopes. Cox et. al. in [28] used the sun as a light source to measure the surface wave spectrum through the specular reflection observed in airborne photographic images. An innovative approach to laboratory measurements was developed by Cox [29]. He illuminated a wedge with varying thickness from the bottom in the laboratory. The light that passed through this wedge was observed by a phototube. The refraction of the light through the surface then caused the intensity of observed light to be proportional to the surface slope. By changing the orientation of the wedge, both along-tank and cross-tank slopes could be observed. Hughes et al. [30] developed a laser point slope technique that used a camera to record the refracted position of a laser beam passing through the surface. A lens was used to eliminate the effects of wave height variation.

Two-dimensional slope measurements were obtained by Keller and Gotwol in [31] and Jähne and Riemer in [32] using a single camera located above the surface and a one dimensional graduated light source located below the surface.The authors in [31] used a light source through a colloidal suspension to generate an exponential

gradient of light intensity in a single axis. A camera was used to record images of this slope component at a rate of 60 Hz. After the light passed through the water surface, it passed through a compensation wedge to linearize the intensity gradient. The authors in [32] used rows of lamps positioned underneath the Delft wind-wave facility. The light shone through an aqueous suspension of latex particles which provided a light intensity gradient in either the along-tank or cross-tank direction depending on which lamps were illuminated. The light gradient was imaged by a camera after passing through the air-water interface. The instantaneous surface slopes in the direction of the light gradient were then inferred from the observed image grey levels, and temporal resolution was achieved by analyzing an image sequence. The advantage of these light gradient methods is that the data analysis is simple and can be performed in real time and the techniques can be readily extended to two spatial dimensions.

Shemdin et al. in [33] used stereo photography to observe the two-dimensional wavenumber spectrum of short gravity waves in the field from a tower. In another approach, Waas et al. [34] made combined slope-height measurements using a Reflective Stereo Slope Gauge. This instrument consists of a pair of point-light sources and video cameras housed in a single package suspended over the surface. Linear polarization filters are placed in front of the lights and cameras so that only specular reflections from the left light are imaged by the right camera, and only specular reflections from the right light are imaged by the left camera. The apparent displacement of specular reflections between the left and right images

(parallax) are proportional to the elevation of the surface facet where the reflection occurred. By computing the parallax for all specular reflections, the position and slope of specular points on the surface are found.

In another approach, Bock and Hara [35], Martinsen and Bock [36], and Lee et al. [37] employed a scanning laser slope gauge in which a scanning laser beam is located below the surface, and a receiver is located above the surface. Then, by recording the instantaneous position of the transmitted and received beams, a two-dimensional slope map of the surface is reconstructed. These instruments have been successfully deployed in the laboratory and field, and are capable of recovering significant statistical information about dynamics of the ocean surface.

To design a system capable of recovering instantaneous joint elevation-slope maps, while not disturbing the air-surface interface, a variety of passive image analysis techniques were investigated. At the air-sea boundary, light rays are reflected back into the atmosphere or refracted into the water volume, and multiple reflections and scattering at the surface can be neglected. At this level, a water surface displays specular reflectance properties. Consequently, the water surface acts like a partially silvered mirror that reflects the sky brightness into the camera. Most image analysis techniques designed to recover shape information from images (feature matching, photometric stereo, and shape-from-shading) assume Lambertian surface reflectance properties. For a Lambertian type surfaces, the apparent brightness of a surface point is strongly dependent on the position of the light source relative to the surface, and weakly dependent on the position of the observer. Fea-

ture matching algorithms take advantage of this reflectance property by assuming that small surface features can be identified in images taken from different positions [15, 9]. Similarly, photometric stereo techniques [15, 38] and shape-from-shading algorithms [11, 10] assume that the pixel brightness depends only on the position of the light source relative to the surface. Several researchers have extended the photometric stereo technique to specular surfaces [12, 39, 13, 40]. These methods, however, require a single camera viewing a stationary surface illuminated by a sequence of extended light sources. Shape-from-specularity techniques [14] have successfully recovered shape information from specular surfaces by analyzing the reflections produced by a small number of point light sources [12, 41, 42, 43]. These methods cannot be extended to natural lighting conditions because it is not possible to isolate an individual specular reflection.

An optical flow technique [44] has been employed to recover the instantaneous shape of a water surface by analyzing the apparent motion of a submerged pattern. This technique is limited, however, to gradually undulating surfaces where the maximum slope does not exceed 5". Because of these restrictions, none of the above methods can be readily adapted to recovering the shape of a moving specular surface (such as water waves) under natural lighting conditions.

Schultz in [2] adapted the multiple image shape-from-shading technique introduced in [45, 11] to recover the fine scale structure of the ocean surface. In this method, known as Specular Surface Stereo, the surface shape is recovered by solving the inverse problem of finding a surface shape (defined by an elevation

and gradient map) that results in a set of synthetic images that closely match a set of observed images. To accomplish this, the algorithm requires an image of the illumination source, three images of the water surface, a few initial elevation estimates, and an image synthesis model that predicts the appearance of an image of the surface.

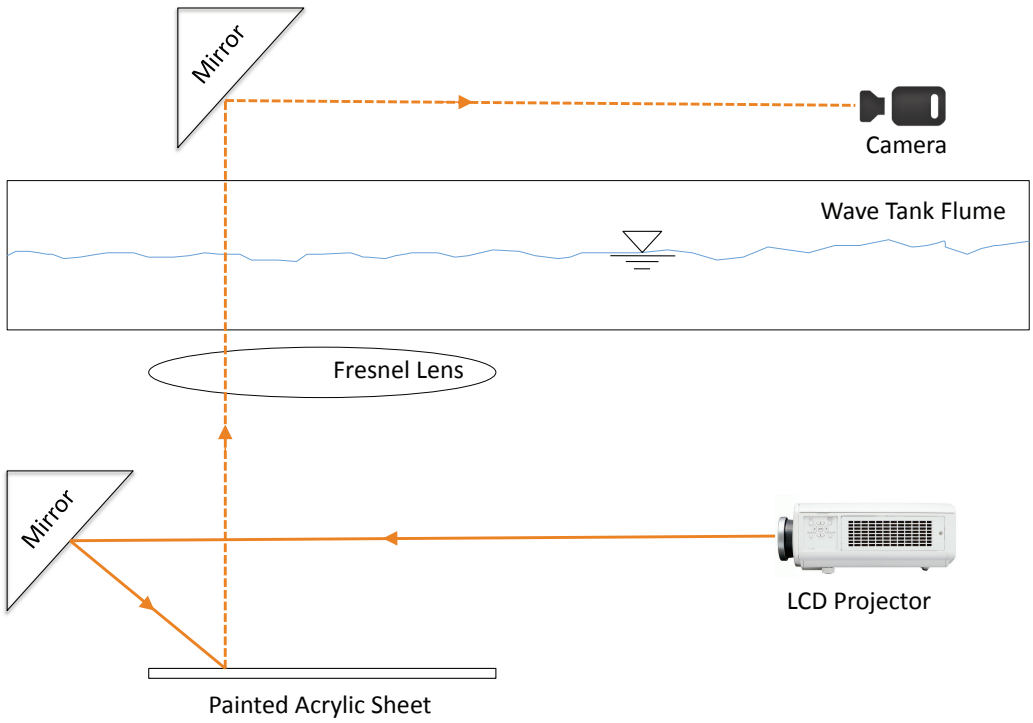


Figure 1: Imaging Slope Gauge setup at the University of Miami's ASIST lab used for the purpose of water surface estimation

Finally, the imaging slope gauge used by Haus et. al. in [46] follows the design of Jähne and Riemer [32], in which an area on the floor beneath a laboratory tank is illuminated with blue and green light having intensity gradients down and

cross-tank respectively. In place of a light box and screen lying beneath the tank, a digital image of known color gradients is projected onto a flat-white painted acrylic sheet below the wave tank. A projector is used to provide carefully controlled gradients. The camera peers down at the image area with a 0 nadir angle. A Fresnel lens is installed just beneath the tank and above the illuminated area alleviating the need for tedious surface elevation measurements. Figure 1 shows the setup used in this method. One advantage of this approach is that the image observed by the camera can be easily analyzed for linearity and range of color gradients. If the observed image does not meet the experimental requirements for the color mask, the projected image can be digitally adjusted to provide desired color gradients and range in the observed mask.

2.2 Object reconstruction using structured-light techniques

Some of the most important methods of 3D surface imaging are based on the use of structured light. These techniques simplify the problem of 3D reconstruction with the help of controlled illumination and can yield excellent results in practice. They basically consist of acquiring 2D image(s) of the scene while illuminating the scene with spatially varying intensity pattern(s). The nonplanar geometric shape of the surface distorts the projected structured-light pattern as seen from the camera. The principle of structured-light 3D surface imaging techniques is to extract the 3D surface shape based on the information from the distortion of the projected structured-light pattern.

A large number of structured light techniques for 3D surface imaging have been developed so far. These techniques can be classified into two categories: those that require multiple images (multi-shot techniques) [47, 48, 49, 50] and those that require only a single image (single-shot techniques) [51, 52, 53, 54, 55] to reconstruct a surface. The single-shot techniques can further be classified into three broad categories: techniques using continuously varying structured-light patterns [51, 56, 52, 57], techniques using 1D encoding schemes (strip indexing) [58, 54, 55], and techniques using 2D encoding schemes (grid indexing) [59, 53, 60].

Generally, multiple-shot techniques may often result in more reliable and a more accurate 3D reconstructed surface. Because of this, if the target is static, the multiple-shot techniques are recommended. However, if the target is moving, using single-shot techniques is required. In our case, since we are dealing with moving objects (dynamic water surface), the single-shot techniques are of more interest.

The key difference between all these techniques is the structure of the illuminant (which could typically be a point, a stripe, multi-point, or multi-stripe) that differentiates a single projected light spot from the acquired image(s). Each technique has its own advantages and disadvantages, depending on the specific applications. Here, we discuss several representative techniques, but by no means all techniques, along that continuum.

The very basic methodology is to employ scanners that sweep across the surface (see Figure 2). A plane of light sweeps across a surface while a camera acquires the stripe reflection and triangulates to the light plane, scanline by scanline.

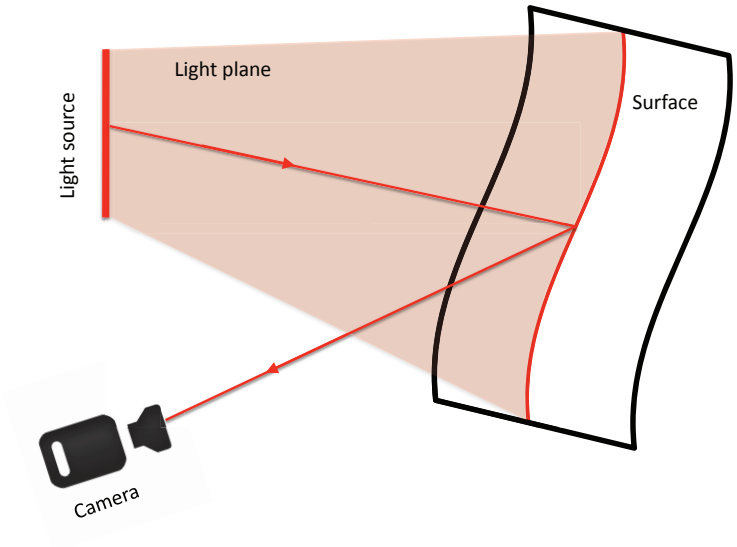


Figure 2: A typical scanner setup. A sheet of light sweeps across a surface while a camera acquires the stripe reflection

Rioux et al. [61] employed a flying spot and linear sensor array. Kanade et al. [62] swept a light plane and recorded the time at which a peak is observed at each sensor pixel. This time is then used to triangulate the sensor line of sight back to the position of the stripe at that time. Curless and Levoy [63] generalize this temporal analysis (calling it spacetime analysis) to other scanner configurations and observe that it substantially increases immunity to shape and reflectance variations which affect purely spatial analyses. While the Kanade method and the Rioux method [64] can achieve high frame rates, both require highly customized hardware.

The binary coding [65, 47, 48, 49, 66] uses black and white stripes to form a sequence of projection patterns, such that each point on the surface of the object

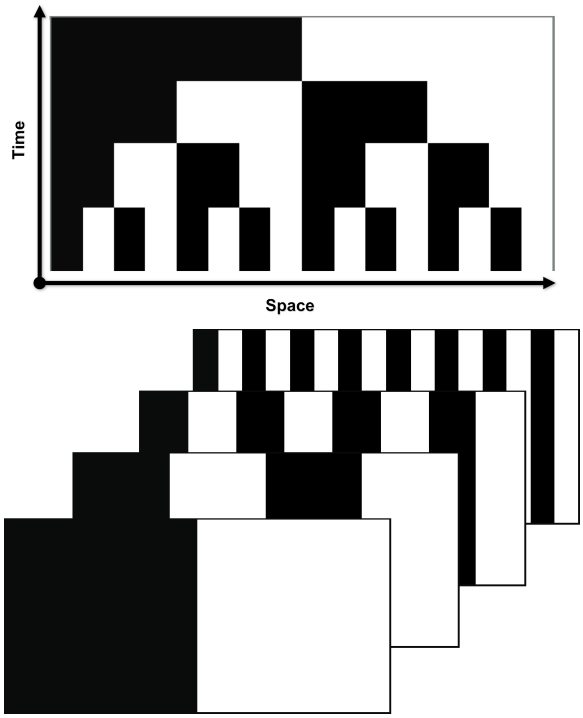


Figure 3: Binary pattern used in multi-shot structured light techniques for 3D reconstruction

possesses a unique binary code that differs from any other codes of different points. In general, N patterns can code 2^N stripes. Figure 3 shows four example patterns which can be used in binary coding method. Once this sequence of patterns is projected onto a static scene, there are 2^N unique areas coded with unique stripes. The 3D coordinates x, y, z can be computed (based on a triangulation principle) for all 2^N points along each horizontal line, thus forming a full frame of the 3D image. Binary coding technique is very reliable and less sensitive to the surface characteristics, since only binary values exist in all pixels. However, to achieve high spatial resolution, a large number of sequential patterns need to be projected. All objects in the scene have to remain static. The entire duration of 3D image acquisition may be longer than a practical 3D application allows for.

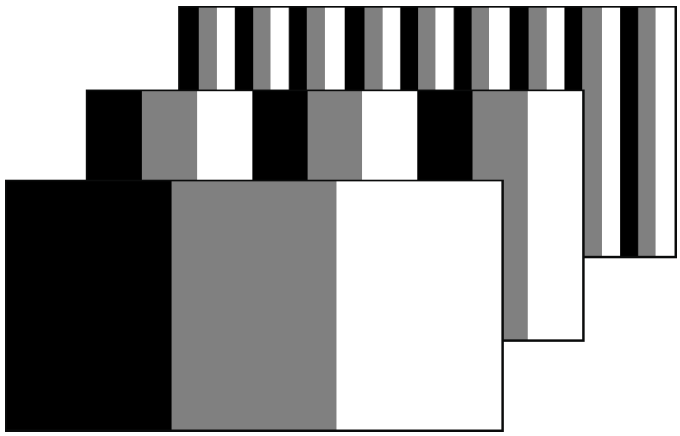


Figure 4: Gray pattern used in multi-shot structured light techniques. It uses several distinct levels of intensity (instead of only two in the Binary pattern), to produce unique coding and reduces number of required projection patterns.

In the direction of using fewer images, gray-level patterns are developed. For example, one can use M distinct levels of intensity (instead of only two in the binary code) to produce unique coding of the projection patterns. Figure 4 illustrates a gray pattern with 3 distinct level of intensity; white, gray, and black. In this case, N patterns can code $M \times N$ stripes. Each stripe code can be visualized as a point in an N-based space, and each dimension has M distinct values [67]. For example, if $N = 3$, and M is 4, then the total number of unique code stripes is $64 = 4^3$. In comparison, for 64 stripes with a binary code, 6 patterns are needed. Caspi et al. [68] reduce the number of images further by using a color generalization of Gray codes. Finally, Hattori and Sato [69] refine the original hierarchical stripe technique by introducing sub-pixel offsets to the final stripe pattern to get finer resolution. Their approach is similar to spacetime analysis and uses $log(N)$ stripes plus m shifted versions of the finest stripes.

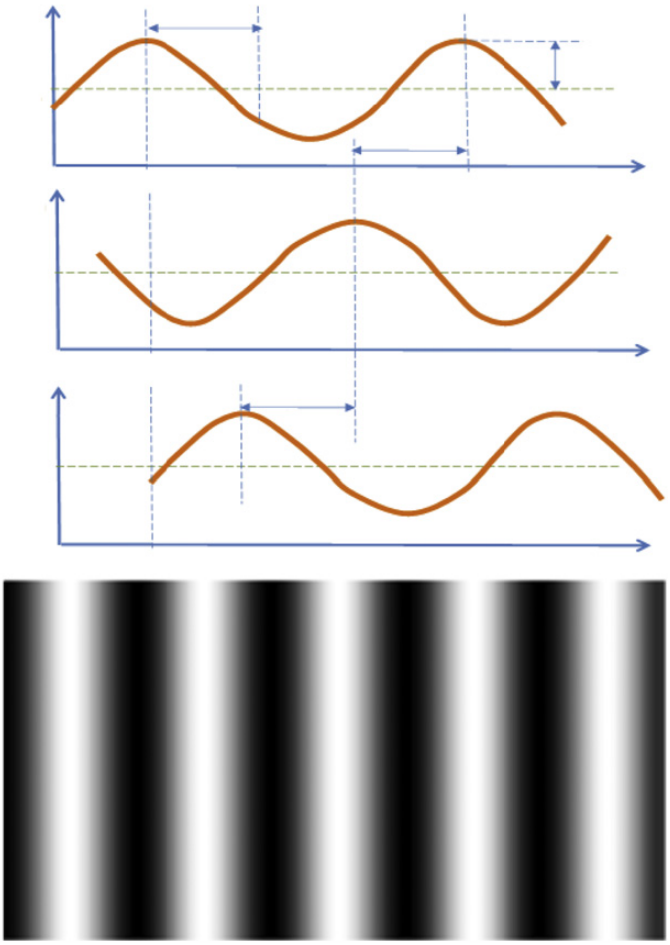


Figure 5: Phase Shift used in multi-shot structured light techniques. A set of sinusoidal patterns are projected. The variation form of these projected fringe patterns are similar, except the phase of the fringe patterns are shifted a constant angle with respect to each other. The minimum number of projections is three.

Another methodology under multiple-shot category is Phase-shift. It is a well-known fringe projection method for 3D surface imaging [50]. A set of sinusoidal patterns is projected onto the object surface. The intensities for each pixel (x, y) of the three projected fringe patterns will be recorded. The phase information $\varphi(x, y)$ can be retrieved (i.e., unwrapped) from the intensities in the three fringe patterns. Phase unwrapping is the process that converts the wrapped phase to the absolute phase. The 3D (x, y, z) coordinates can be calculated based on the

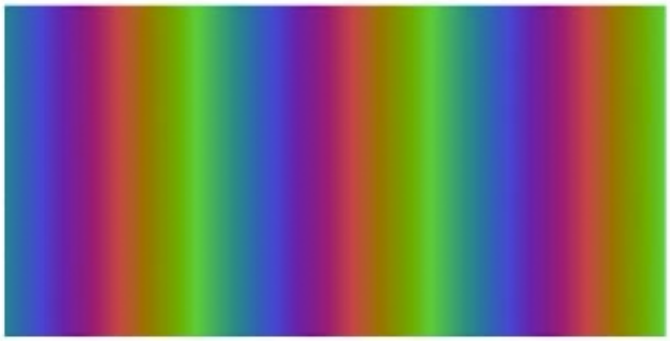


Figure 6: Color Phase Shift used in single-shot structured light techniques

difference between measured phase $\varphi(x, y)$ and the phase value from a reference plane [70]. Figure 5 shows a standard light pattern for the Phase Shift method and Figure 6 shows a color pattern used in an extension of this method.

Hall-Holt and Rusinkiewicz [71] describe a method that consists of projected stripe patterns that vary over time and are based on coding the boundaries between projected stripes. By finding the nearest stripe patterns over time, a unique code can be determined for any stripe at any time. The constraint in this case is that the object should be rigid and move slowly to avoid erroneous temporal correlations.

Major drawbacks of the multiple-shot techniques include their inability to acquire the 3D scenes in dynamic motion. The last set of techniques we describe are those that take advantage of color information or a unique encoding scheme in the projection pattern and require only one acquired image of the object under the color pattern illumination to derive the full frame of the 3D image with (x, y, z) coordinates of each visible point in the scene.

To obtain 3D data by projecting one pattern and a single image, one can use stripe indexing techniques [55],[72],[73]. In order to distinguish one stripe from

others, randomly distributed cuts can be used to identify each stripe. This way, each stripe is divided into many small line segments (see Figure 7). Then, segment matching can be performed between the image and the pattern to obtain 3D data [55]. However, this only applies to a 3D object with a smooth and continuous surface when the pattern distortion due to surface shape is not severe. Otherwise, it may be very difficult to recover the unique segment pattern and will result in false matches. One solution to reduce the number of false matches is to use adjacency relations among segments and segment pairs whose adjacent segments also correspond with each other are extracted and considered to be correct matches. Then, from the resultant matches, the correspondence can be propagated to get an entire 3D reconstructed object.

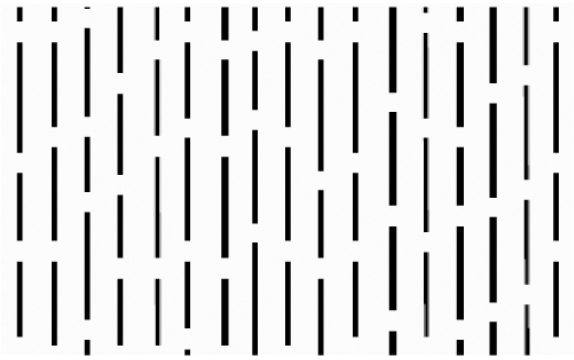


Figure 7: Stripe Indexing using segment pattern used in single-shot structured light techniques

Durdle et al. in [72] introduced another stripe indexing approach that utilizes more than two intensity levels. They arrange the intensity levels of stripes such that any group of stripes has unique intensity pattern within a period of length (See Figure 8). This approach, however, might not be categorized in single shot

techniques since it uses two stereo cameras and a projector for 3D reconstruction.

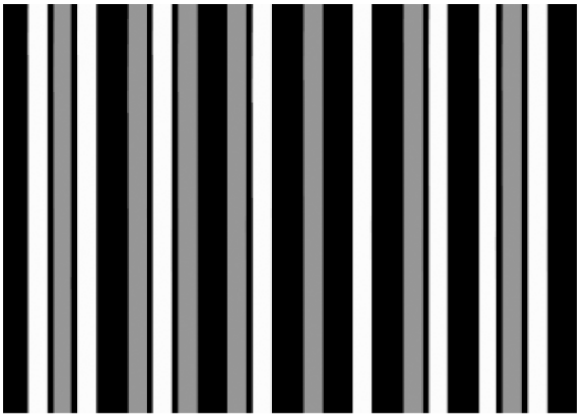


Figure 8: Stripe Indexing repeated gray level used in single-shot structured light techniques

Zhang et al. [73] presented a color structured light pattern for recovering object shape that can be used in both single-shot and multiple-shot approaches. They used De Brujin sequence [74] in order to create unique stripes. Figure 9 shows their stripe pattern and an example image captured by their camera. In the De Brujin sequence, no subsequence is correlated to any other. This unique feature can be used in constructing a stripe pattern sequence that has unique local variation patterns that do not repeat themselves. Such uniqueness makes the pattern decoding an easier task. Unlike the method in [72], they are using one camera. In order to resolve the vertical ambiguity in finding correspondences, they rectify the image and the projected pattern so that the displacements are purely horizontal and therefore, the search space for finding correspondences is one dimensional.

Proesmans et al. [75, 76] demonstrate a scanner which projects a grid pattern onto the scene and matches the observed grid to the projected pattern by

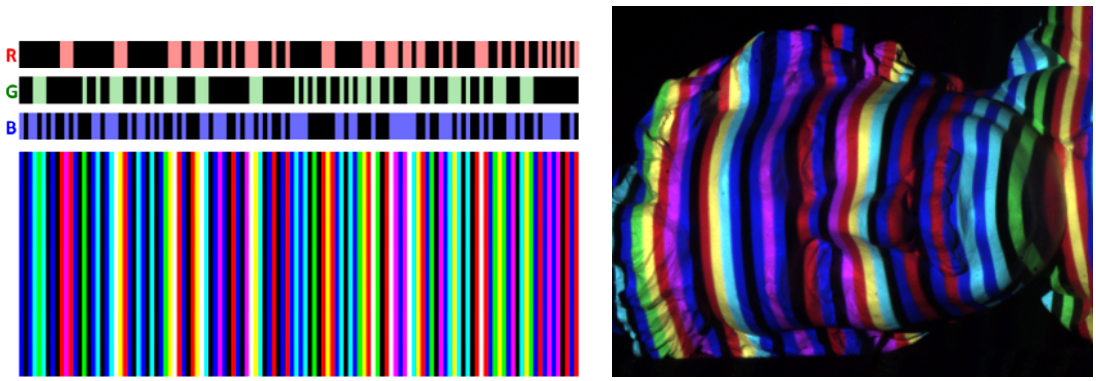


Figure 9: Color Stripe Indexing used in single-shot structured light techniques. (left) an example of the projected stripe pattern, (right) an image captured by the camera.

a global 2D grid optimization algorithm. In this case, the constraint is that the visible portion of the object consist of a single connected component. Boyer and Kak [77]project a color stripe pattern in which each window of spatially adjacent stripes has a unique color intensity configuration. Davies and Nixon [78] propose a color dot pattern with a similar spatial neighborhood property. In both cases, local neighborhoods must exhibit enough spatial coherence to preserve the windows. Lavoie et al. [79] encode pseudo-random grids as bi-dimensional binary arrays (PRBA) composed of sequences in which any subset (or window) is unique and defines a specific location (See Figure 10).

Finally, Chen et al. [80] describe a stereo vision method that also uses projected color stripes, but solves correspondences between edges in the two cameras through dynamic programming, thus obtaining a global optimum. The constraint in this case is surface monotonicity between the cameras. Each of the color pattern methods above has the additional constraint that the surface does not change the reflected color too much, e.g., doesn't cause a color channel to drop out completely.

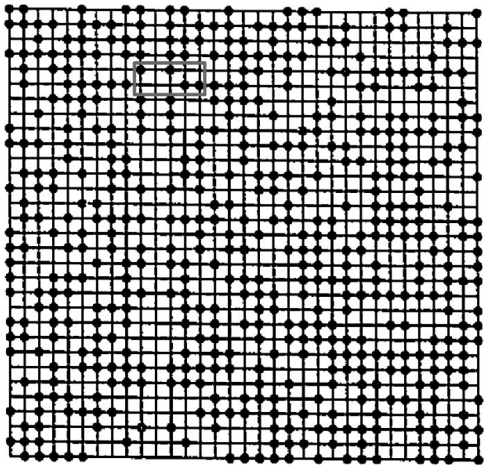


Figure 10: Pseudo Random Binary Array (PRBA) used in single-shot structured light techniques

Figure 11: Rainbow pattern with spatially varying wavelength illumination used in single-shot structured light techniques

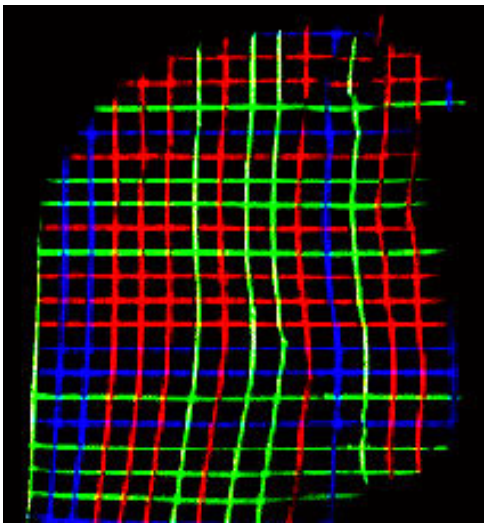


Figure 12: Color coded grid by both vertical and horizontal stripes used in single-shot structured light techniques

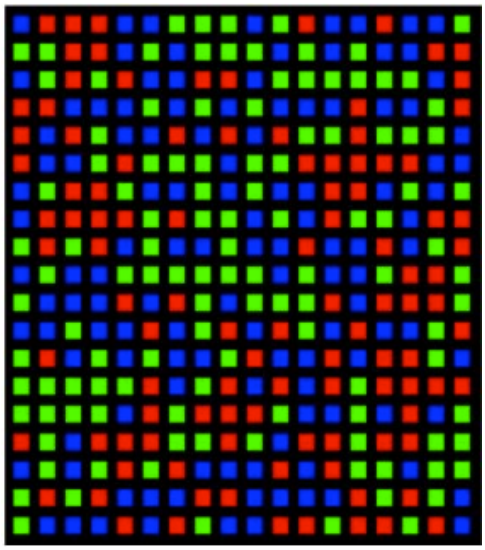


Figure 13: Psuedo Random Code is a 2D array of color-coded dots used in single-shot structured light techniques

Although the single-shot techniques are useful for achieving large spatial resolution, these techniques present their own drawbacks. The limited bandwidth of the projectors provokes integration of intensities over adjacent pixels. In addition, light intensities vary due to the different colors and depths of the measuring surface. Also, the error quantization introduced by the camera adds more difficulties to the color identification. Therefore, the correct identification of every projected intensity or color is not easy to achieve. To overcome these difficulties, in most cases, such techniques use devices that project a unique wavelength for every grey level or color along with cameras with large depth-per-pixel (about 11 bits per pixel) for accurate quantization.

Note that there are many ways to improve the performance of a 3D surface imaging system by combining more than one encoding scheme. Examples of these methods can be found in [81]. Also, due to the complexities in the essence of

image-based 3D reconstruction problem, a large number of techniques have been developed to date and each technique has its own set of advantages and disadvantages. In order to select a technique that can be applied to a specific application, one should ponder the application requirements and choose the method that matches those requirements based on indexes such as accuracy, resolution, speed, cost, and reliability.

3

Proposed System

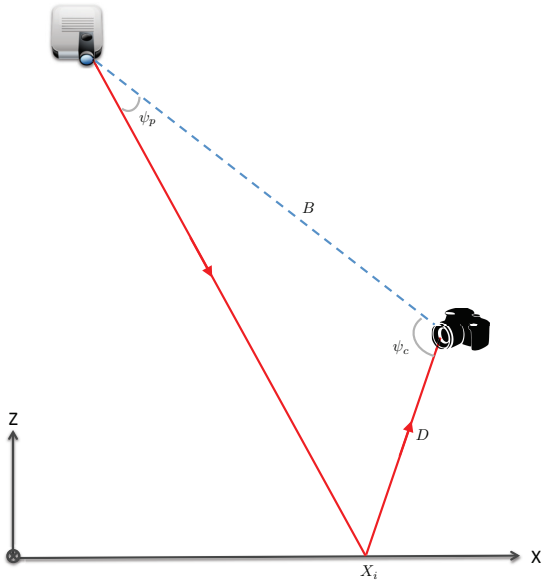
3.1 Problem statement

Consider a typical structured-light system that both camera and the projector are in the same medium, see Figure 14(a). A spatially varying structured illumination is generated by a projector or a light source modulated by a spatial light modulator. An imaging sensor (a video camera, for example) is used to acquire a 2D image of the scene under the structured-light illumination. If the scene is a planar surface without any 3D surface variation, the pattern shown in the acquired image is similar to that of the projected structured-light pattern. However, when the surface in the scene is nonplanar, the geometric shape of the surface distorts the projected structured-light pattern as seen from the camera. The principle of structured-light 3D surface imaging techniques is to extract the 3D surface shape based on the information from the distortion of the projected structured-light pattern.

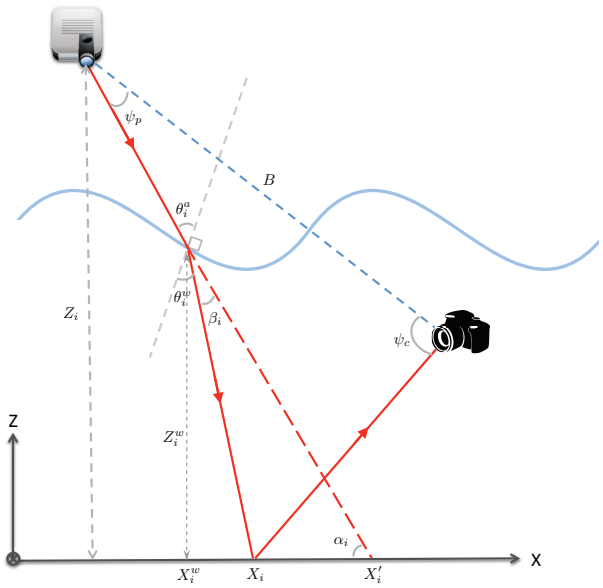
As shown in Figure 14(a) , the geometric relationship between an imaging sensor, a structured-light projector, and an object surface point can be expressed by the triangulation principle as

$$D = B\frac{sin(\psi_p)}{sin(\psi_c + \psi_p)} \tag{1}$$

Now consider the scene in Figure 14(b) where the projector is located above water and the camera is under water. In this case, the distortion of the pattern is



(a)



(b)

Figure 14: Illustration of a ray in a projector-camera system: (a) both camera and the projector are in the same medium, (b) camera and the projector are in different medium, in this particular case the projector is above water and the camera is underwater.

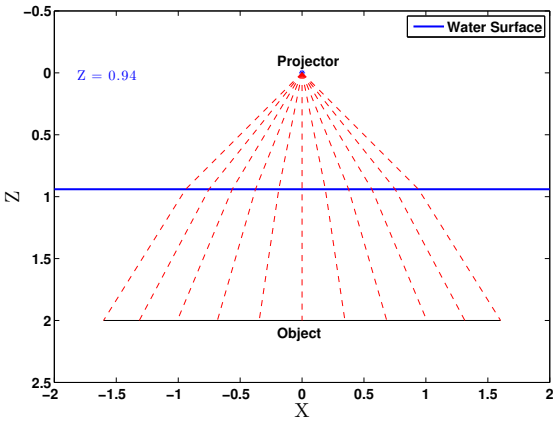
due to both the geometric shape of the water surface and the geometric shape of the object surface. To recover the object surface, one should consider distortions from the water surface.

To demonstrate the effect of water surface shape on the projected pattern, consider the scene in Figure 15 where a projector illuminates an underwater flat object while its line of sight is perpendicular to the object's surface. These plots show the effect of refraction of the projected rays when they enter the water. Figure 15(a), (b), and (c) illustrate the behavior of the rays against a flat, sloped, and sinusoidal water surfaces, respectively. As can be seen, each ray has a different refraction behavior based on its incidence angle. When a light ray passes through the water surface, the Snell's law can describe the relationship between the angles of incidence and refraction of that ray, i.e.

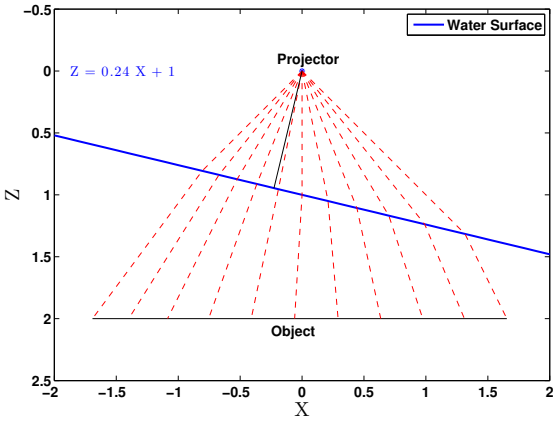
$$n_w \sin(\theta_i^w) = n_a \sin(\theta_i^a) \tag{2}$$

where n_a represents the refractive index of the air and n_w is the refractive index of the water. Also, θ_i^a and θ_i^w respectively represent the incidence angle and the refraction angle of the i^{th} ray.

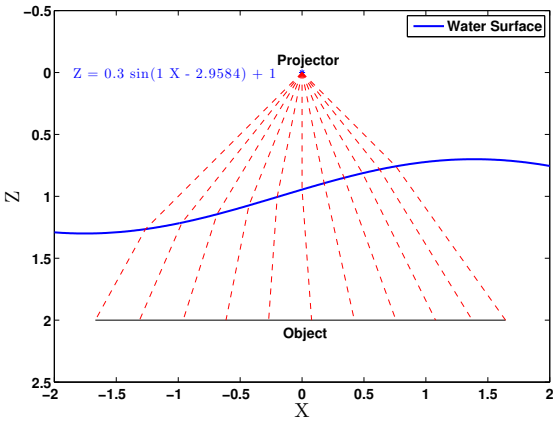
Because of the refraction effect, a unique projected pattern will have different shapes, even on a flat object at the same distance from the projector, depending on the water surface shape. If one can estimate the water surface, then it is possible to predict the path of each projected ray and account for the refraction effect on the distortion of the projected pattern on the underwater object. We will discuss how to estimate water surface in section 3.4.



(a)



(b)



(c)

Figure 15: Illustration of the ray refraction after hitting water surface with a (a) flat surface with Height =0.94(m), (b) flat surface with Slope=0.24 and Intercept 1, (c) Sinusoidal water surface

In order to estimate the geometric shape of the water surface and underwater objects, we first need to calibrate the camera-projector system. In the next section, we will explain the details of the camera-projector (system) calibration.

3.2 System calibration

Calibration is an essential part of a camera-projector system which involves intrinsic calibration of both camera and the projector along with extrinsic calibration of camera -projector. The accuracy of calibration step plays a critical role in measurement accuracy of camera-projector system. Camera calibration is the first step in the system calibration which can be done by taking several images of a checkerboard pattern at different angles.

3.2.1 Camera calibration

Camera calibration is a well-known problem in computer vision. It establishes the relationship between a pixel on a 2D image (in camera coordinates) and a straight line in 3D space (world coordinates) along which the object point is located, taking lens distortion into consideration. Usually, a pinhole camera model and a set of intrinsic parameters are used to characterize the relationships. Several approaches are proposed and their accompanying toolboxes are available [82, 83, 84]. The most common approach is proposed by Zhang [83] and implemented by Jean-Yves Bouguet [85]. It requires images at several angles and distances of a known calibration object. A planar checkerboard pattern is used as the calibration object because it is very simple to produce, can be printed with a standard printer, and has distinctive corners that are easy to detect. An example

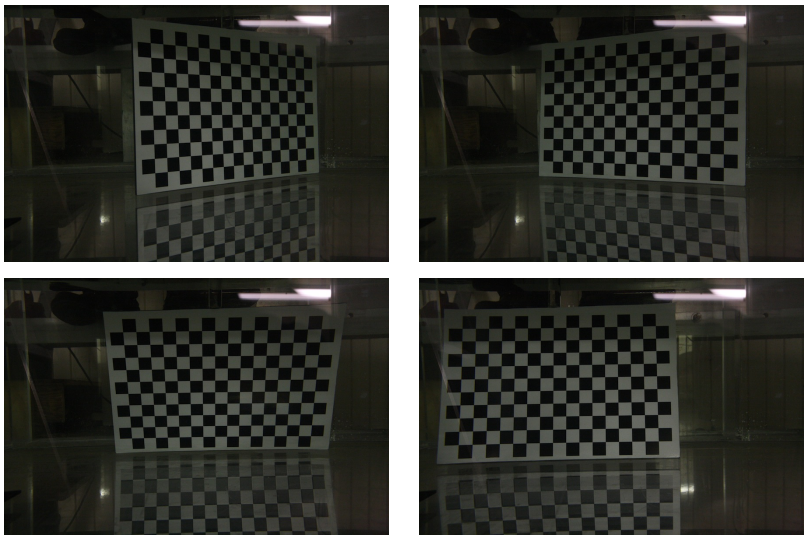


Figure 16: Samples of camera calibration images

image involving such a pattern is shown in Figure 16. From the images of the calibration pattern 2D to 3D correspondences are constructed.

The intrinsic parameters in this approach are:

Focal length: this approach returns two focal lengths that are expressed in units of horizontal and vertical pixels. Both focal lengths are usually very similar. Their ratio, often called "aspect ratio", is different from 1 if the pixels in the CCD array are not square. Therefore, the camera model naturally handles non-square pixels.

Principal point: the intersection of optical axis and the image plane.

Skew coefficient: This coefficient encodes the angle between the x and y sensor axes. Consequently, pixels are even allowed to be non-rectangular. Some authors refer to that type of model as "affine distortion" model.

Distortions: the image distortion coefficients. The geometric relationship

modeled in the Eq.(1) accounts for images which are not influenced by lens distortions. In practice lens distortions are inevitable and must be corrected. The distortion model used in this approach was first introduced by Brown [5] and is called the "Plumb Bob" model. For more details, refer to [5].

Assume the plane of the planar calibration board in world coordinates to be $Z = 0$, then each point on the calibration board becomes $M = [X, Y, 0, 1]^{\mathsf{T}}$. Therefore, an object point M and its image point m are related by a homographic matrix H [83],[81]:

$$m \sim K[r1, r2, r3, -Rt][X, Y, 0, 1]^{\mathsf{T}} \tag{3}$$

$$m \sim H[X, Y, 1]^{\mathsf{T}} \tag{4}$$

where $H = [h_1, h_2, h_3] = K[r_1, r_2, -Rt]$ is a 3×3 matrix defined up to scale, and r_1, r_2, r_3 are 3×1 column vectors of the rotation matrix. Noticing that for rotation matrix R the column vector r_1, r_2, r_3 is orthonormal, we therefore have

$$h_1^{\mathsf{T}}(KK^{\mathsf{T}})^{-1}h_2^{\mathsf{T}} = 0 \tag{5}$$

$$h_1^{\mathsf{T}}(KK^{\mathsf{T}})^{-1}h_1^{\mathsf{T}} = h_2^{\mathsf{T}}(KK^{\mathsf{T}})^{-1}h_2^{\mathsf{T}} \tag{6}$$

Each homography can provide two constraints on the intrinsic parameters. As $K^{-\mathsf{T}}K^{-1}$ in the Eq.(6) is a symmetry matrix, it can be defined with a 6D vector:

$$A - K^{-\mathsf{T}}K^{-1} = \begin{pmatrix} A_1 & A_2 & A_4 \\ A_2 & A_3 & A_5 \\ A_4 & A_5 & A_6 \end{pmatrix} \tag{7}$$

Let the ith column vector of H be $h_i = [h_{i1}, h_{i2}, h_{i3}]$. Then, we have

$$h_i^{\mathsf{T}} A h_j = v_{ij} a \tag{8}$$

where

$$v_{ij} = \begin{pmatrix} h_{i1}h_{j1} \\ h_{i1}h_{j2} + h_{i2}h_{j1} \\ h_{i2}h_{j2} \\ h_{i3}h_{j1} + h_{i1}h_{j3} \\ h_{i3}h_{j2} + h_{i2}h_{j3} \\ h_{i3}h_{j3} \end{pmatrix}$$

and

$a = [A_1, A_2, A_3, A_4, A_5, A_6]$

The two constraints can then be rewritten as a homogeneous equation. In order to solve a, at least three images from different viewpoints are needed. In practice, more images are used to reduce the effect of noise, and a least-squares error solution is obtained with singular value decomposition. Finally, the result can be optimized to minimize the reprojection error by minimizing the energy function below:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} ||m_{ij} - \bar{m}(K, R, t, M_j)||^2 \tag{9}$$

3.2.2 Camera-Projector calibration

The only difference between the optical model of a projector and a camera is their direction of projection. Thus, a projector can be considered as an inverse camera. It can be calibrated using camera calibration algorithms if the correspondence between projector and the world coordinate system is established. However,

the inverse model makes the problem of relating a pixel on a 2D image (in projector coordinates) with a straight line in 3D space (world coordinates) difficult, as we cannot tell where a given point in 3D space will be projected in the inverse camera coordinates.

To overcome this difficulty, a pre-calibrated camera and a calibration plane (in this case checkerboard) can assist. If a calibrated camera observes a checkerboard of known size on a plane, the resulting correspondences can be used to solve for the equation of the plane in world space. Now given any point in the camera that falls onto this plane, the 3D world position of the point can be determined through ray-plane intersection. Therefore, if the projector emits a spatial code that the camera can locate, this point can be translated into the corresponding 3D world point. For instance, if the projector emits a checkerboard pattern, the corners of the pattern can serve as spatial codes yielding projector to world space correspondences. Figure 18 shows the projected checkerboard along with the physical checkerboard at different orientations. Then the projector can be calibrated using the Zhang based calibration method that was used to calibrate the camera. This method is straightforward in theory and relatively easy to implement. However, the calibration accuracy of these methods depends heavily on the accuracy of the pre-calibration of the camera.

Both image capture and projection is generally described by a standard pinhole camera model with intrinsic parameters including focal length, principle point, pixel skew factor, and pixel size; and extrinsic parameters including rotation

and translation from a world coordinate system to a camera or projector coordinate system. We assume that we have a calibrated camera. The projector calibration can be summarized as follows:

Step 1: Recover calibration plane in camera coordinate system

Step 2: Project a checkerboard on calibration board, capture an image of the checkerboard and detect its corners

Step 3: Determine 3D coordinates of each projected corner by intersecting its corresponding ray from the camera and the recovered calibration plane

Step 4: Calibrate the projector using the correspondences between the 2D points of the image that is projected and the 3D projected points.

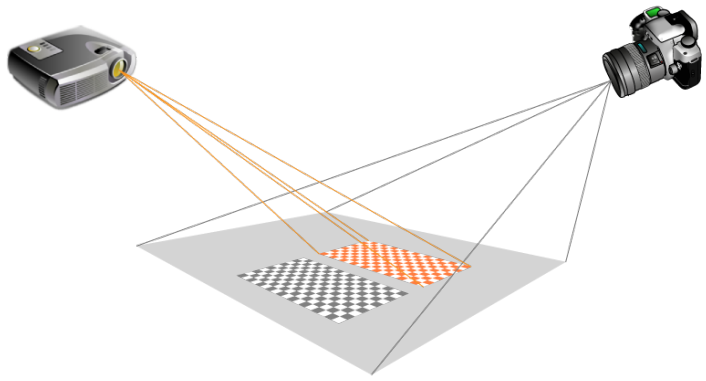


Figure 17: Schematic view of the projector, camera and the calibration plane. By changing the orientation of the calibration plane and taking images of the plane by a pre-calibrated camera, the intrinsic parameters of the projector and the translation and the rotation matrices of the camera with respect to projector can be estimated.

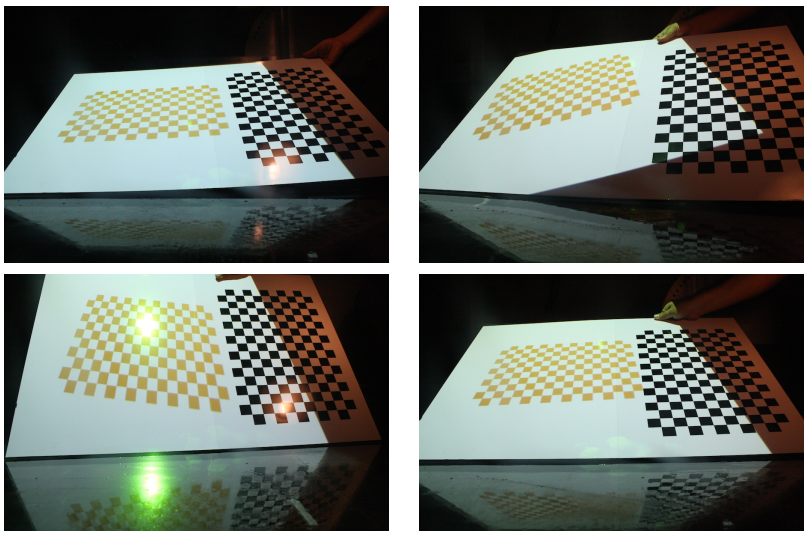


Figure 18: Samples of projected checkerboard along with physical checkerboard at different orientations used for projector calibration

3.3 Pseudocode pattern

Since in this application (at least) water surface is dynamic, we have to use a structure for the illuminant that enables us to localize all light spots in a single-shot. We adopt the random pseudocode structure proposed in [53, 86]. They introduced a brute force algorithm to generate an array that preserves the uniqueness of subwindows, but it may not exhaust all possible subwindow patterns. The method is relatively intuitive to implement in computer algorithms. For example, Figure 19 (left) shows a 6×6 array with subwindow size of 3×3 using three code words (R, G, B). The computing procedure is as follows: first fill the upper left corner of the 6×6 array with a randomly chosen pattern. Then, add a three-element column on the right with random code word. The uniqueness of the subwindow is verified before adding such a column. Keep adding the columns until all columns are filled with random code words and subwindow uniqueness is

verified. Similarly, add random rows in the downward direction from the initial subwindow position. Afterwards, add new random code words along the diagonal direction. Repeat these procedures until all dots are filled with colors. Again, this computational procedure may not guarantee the generation of a pseudo-random array for all the array sizes and code words, but good results have been achieved for many cases. Figure 19 (right) shows an example of a pseudo-random array with 50×50 dimensions.

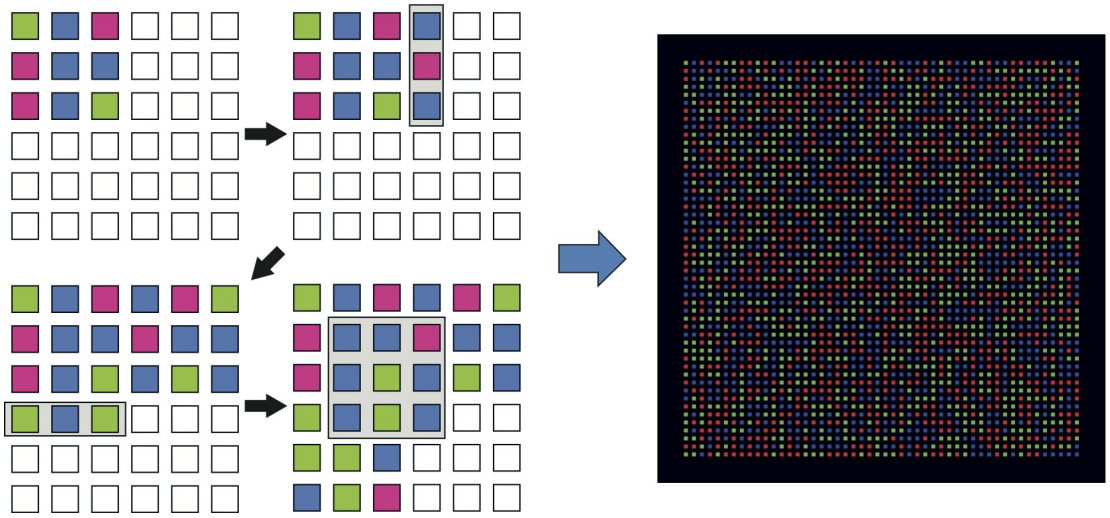


Figure 19: Generating pseudocode: (left) step by step pseudocode generation (right) example of a generated 50×50 pseudocode

In all our experiments we are using three colors to create the pseudocodes but there is no restriction on the number of colors that could be used in order to create such pseudocodes. However, increasing the number of colors could potentially increases the noise sensitivity when measuring colorful scenes.

3.4 Water surface estimation

Consider the scene in Figure 14(b). Without loss of generality, we consider the water surface along the XZ plane. A projector is located above water and a camera under water. We assume that both camera and the projector are calibrated, so the transformation from camera frame to projector frame is established. The water surface (blue line) is a dynamic surface. A light ray (red line) emitted from the projector passes through the water surface and refracts based on Snell's law. Then it reflects from the object surface towards the camera. We assume that the water surface has a parametric model, e.g. it can be represented by a polynomial.

Using a calibrated camera, the rotation and translation matrices of any plane can be obtained. Consider a plane with an attached checkerboard (calibration grid) and concentrate on the camera reference frame attached to that grid. Figure 20 shows the reference frame (O, X, Y, Z) attached to that calibration grid.

Let P be a point with coordinate vector $\mathbf{X}_p = (X_p, Y_p, Z_p)$ in the grid reference frame shown in Figure 20. Let $\mathbf{X}_p^c = (X_p^c, Y_p^c, Z_p^c)$ be the coordinate vector of P in the camera reference frame. Then $\mathbf{X}_p$ and $\mathbf{X}_p^c$ are related to each other through the following rigid motion equation:

$$\mathbf{X}_p^c = R^c \mathbf{X}_p + T^c \tag{10}$$

In particular, the translation vector T^c is the coordinate vector of the origin of the grid pattern (O) in the camera reference frame, and the third column of the matrix R^c is the surface normal vector of the plane containing the planar grid in the camera reference frame.

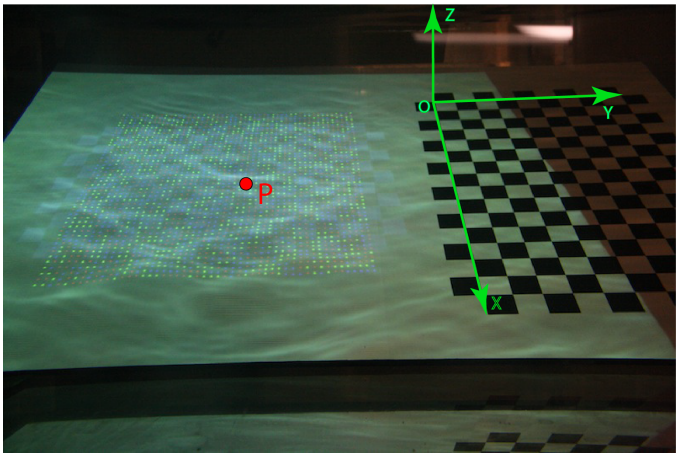


Figure 20: Object coordinate system. Using a calibrated camera and a checkerboard attached to the white board, the coordinates of all dots of the psuedorandom code can be calculated. The Z component of all these dots are equal to zero.

The translation vector T^c and the rotation matrix R^c of the plane can be recovered in the camera reference frame using corresponding points. For this purpose, one can use the corners of the checkerboard. These corners can be easily determined in the image plane. Once the coordinates of a point is determined in the image plane, it can easily be expressed in the camera reference frame using the intrinsic camera parameters. On the other hand, each corner on the calibration board becomes $\mathbf{X}_p = (X_p, Y_p, 0)$. By plugging corresponding $\mathbf{X}_p$'s and $\mathbf{X}_p^c$'s into Eq. 10, the T^c and R^c will be estimated. Once the transformation from the plane coordinates system to camera reference frame is established, the coordinates of any point on the plane can be determined in camera reference frame.

Using the rotation matrix and translation vector between camera and the projector obtained from camera-projector calibration, one can transform object points from the camera reference frame to the projector reference frame. The coordinates of the pseudocode in the projector image plane can easily be expressed

in the projector reference frame using the intrinsic projector parameters obtained from the camera-projector calibration.

For a moment, suppose that we know the parametric representation of water surface function in the projector reference frame; i.e. $Z^w = F(X^w, \mathbf{a})$ where $\mathbf{a} = \{a_0, a_1, ..., a_n\}$ are parameters of the function. Let us represent the intersection of ray i and the water surface as (X_i^w, Z_i^w). See Figure 21 for illustration of these parameters. This intersection can be determined by finding a point that satisfies both the water surface function and the line that represents the ray from the projector:

$$Z_i^w = Z_{proj} + \frac{(Z_i' - Z_{proj})(X_i^w - X_{proj})}{(X_i' - X_{proj})} \tag{11}$$

Where (X_{proj}, Z_{proj}) are the projector coordinates. According to Figure 21, one can write X_i as:

$$X_i = X_{proj} + \Delta X_i^a + \Delta X_i^w \tag{12}$$

where

$$\Delta X_i^a = \pm|(Z_i^w - Z_{proj}) \cot(\alpha_i)| \tag{13}$$

and

$$\Delta X_i^w = \pm|(Z_i^w - Z_i) \cot(\alpha_i + \beta_i)| \tag{14}$$

In above Eq.(13) and Eq.(14), α_i can be easily estimated using the intrinsic parameters of the projector which have been already estimated in the calibration step, i.e. $\tan(\alpha_i) = f_{proj}/x_i^p$. Also, β_i is the difference between the incident angle

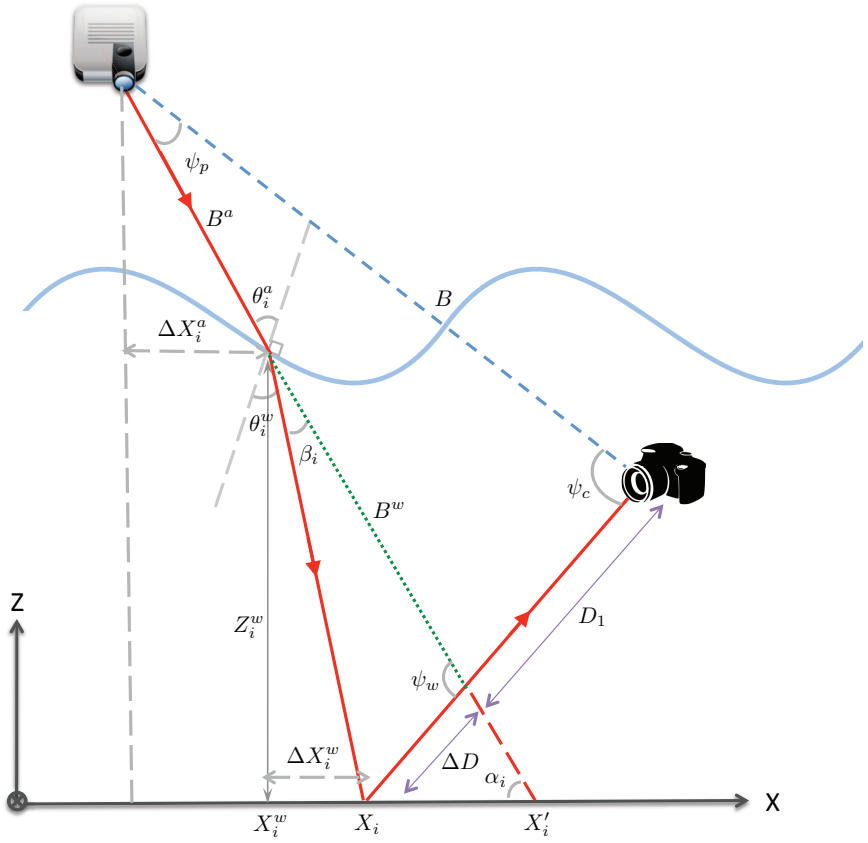


Figure 21: Two dimensional path of a light ray emitted from a projector located above the water and captured by an underwater camera. The ray refracts after hitting the water surface based on its incidence angle θ_i^a.

θ_i^a and the refractive angle θ_i^w:

$$\beta_i = \theta_i^a - \theta_i^w \tag{15}$$

The incident angle for each ray can be calculated by the following equation:

$$\theta_i^a = \frac{\pi}{2} - (\alpha_i - \gamma_i) \tag{16}$$

where γ_i is the slope of the water surface at the intersection of the ray i with the

water surface. If the water surface is known, it can be calculated by:

$$\tan(\gamma_i) = \frac{dF(X_i^w, \mathbf{a})}{dX_i} \tag{17}$$

once the incident angle θ_i^a is determined, the refractive angle θ_i^w can be calculated using Snell's law:

$$\theta_i^w = \sin^{-1}(\frac{n_a}{n_w}\sin(\theta_i^a)) \tag{18}$$

where n_a is the refractive index of the air and n_w is the refractive index of the water.

In conclusion, if the shape of the water surface is known, one can trace each ray and easily predict the coordinates of that ray in the projector coordinate frame. Now, the problem of water surface estimation becomes an optimization problem. In other words, we need to find a set of water surface parameters $\mathbf{a}$ that minimizes the distance between estimated coordinates of rays and their observed coordinates:

$$\underset{\mathbf{a}}{\text{minimize}}\ J(\mathbf{a}) \tag{19}$$

where $J(\mathbf{a})$ is the cost function and is equal to

$$J(\mathbf{a}) = \frac{1}{2n}\sum_{i=0}^{n}(X_i - X_i^p)^2 \tag{20}$$

In Eq.(19) and Eq.(20), n is the number of rays that will participate in water surface estimation. Note that n should not be less than the number of parameters that describe the water surface. We used the exhaustive search algorithm as a

global optimization technique to solve this problem. Other optimization techniques can be used to find the water surface parameters as long as they do not entrap in local minima.

3.5 Object reconstruction

Once the water surface parameters are determined, the path of each ray passing through the water surface can be predicted. When both the camera and the projector are in the same medium, in order to compute object coordinates, only one distance, D, needs to be determined (Figure 14(a)). This can be done using Eq.(1). However, in our proposed system, because of refraction of the light rays, this equation is not valid for computing the distance between the camera and the object. In the presence of the refracted light rays, the distance between the camera and the object, can be computed by adding or subtracting ΔD from D as shown in Figure 21.

$$D = B\frac{sin(\psi_p)}{sin(\psi_c + \psi_p)} \pm \Delta D \tag{21}$$

in Eq.(21), if the ray refracts towards the camera, use $-\Delta D$, and if it refracts aways from the camera, use $+\Delta D$.

$$\Delta D = B^w \frac{sin(\beta_i)}{sin(\beta_i + \psi_w)} \tag{22}$$

where

$$\psi_w = \begin{cases} \psi_c + \psi_p & \text{if ray refracts away from the camera} \\ \pi - \psi_c - \psi_p & \text{if ray refracts towards the camera} \end{cases} \tag{23}$$

B^w is the distance between two intersections: (1) the intersection of the ray from projector and the water surface and (2) the intersection of the undeviated ray from projector and the reflected ray from object towards the camera. By the water surface function obtained using the procedure in Section 3.4, one can easily obtain its intersection with the ray projected from the projector. Also, using D calculated by Eq.(1), the coordinates of the camera, and the direction of the ray reflected from the object, the second intersection can be obtained.

4

Evaluation

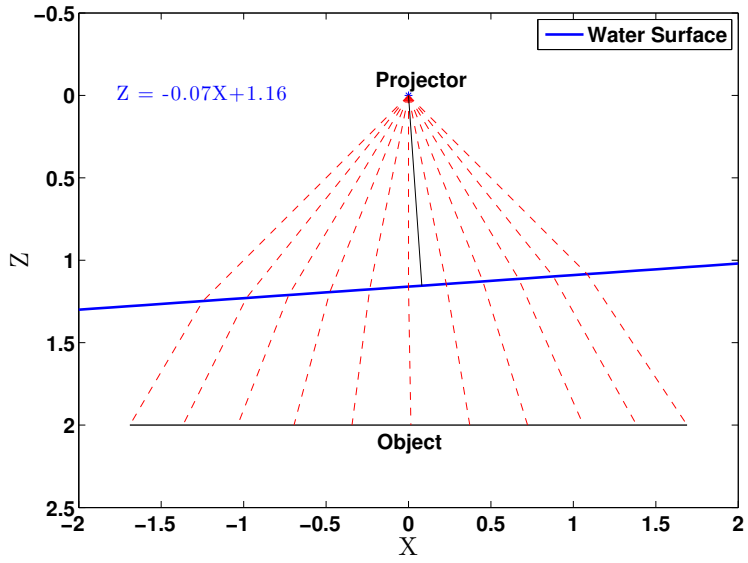
4.1 Experiments and results

Several experiments were conducted to justify the ability of the proposed structured light method to estimate the water surface and the underwater object shape. We present the results for both synthetic and real data. A cylindrical object with a flat top and a spherical object were used for 3D reconstruction. These examples demonstrate that the method works for objects with different shapes. The experiments were conducted in the presence of both flat and wavy water surfaces. We represent the flat water surface with a linear model and the wavy water surface with a sinusoidal model. Other parametric models could be used if they are more appropriate for a particular condition of the water surface.

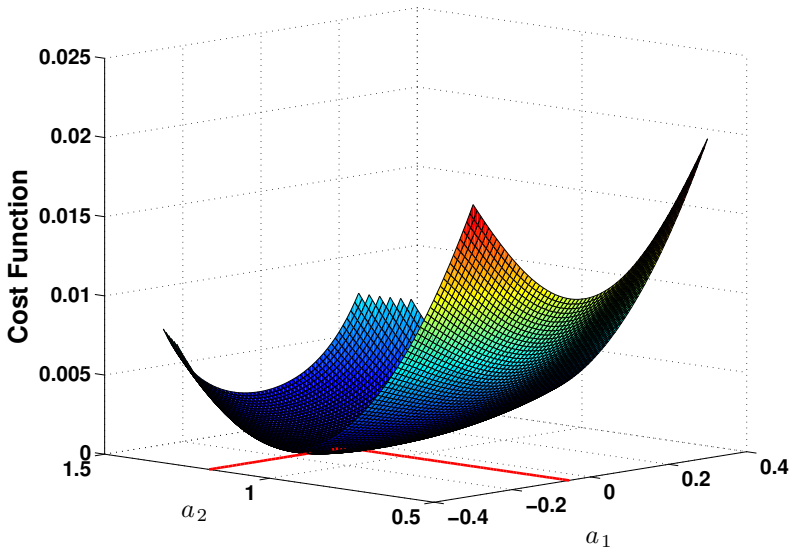
4.1.1 Synthetic data

To evaluate the correctness and performance of the implemented method, we created synthetic data for both linear and sinusoidal water surfaces (see Figure 15 (b),(c)). The world coordinate origin is assumed to be at the projector center. All internal calibration parameters for the projector are known and were set to $(x_c^{Proj}, y_c^{Proj}) = (0, 0)$, $f_{Proj} = 100$pixel, and no lens distortion for simplicity. With known water surface parameters, we were able to calculate (X, Z) pairs for each ray using Eq.(11) to Eq.(14) .

In the first synthetic dataset, it is assumed that the water surface has a linear function, i.e. $Z_i^w = a_1 X_i^w + a_2$. In this dataset, a_1 and a_2 were set to -0.07



(a)



(b)

Figure 22: (a)Synthetic flat water surface of the form $Z_i^w = a_1 X_i^w + a_2$ and the rays from the projector. The a_1 and a_2 were set to -0.07 and 1.16 (b) Corresponding cost function.

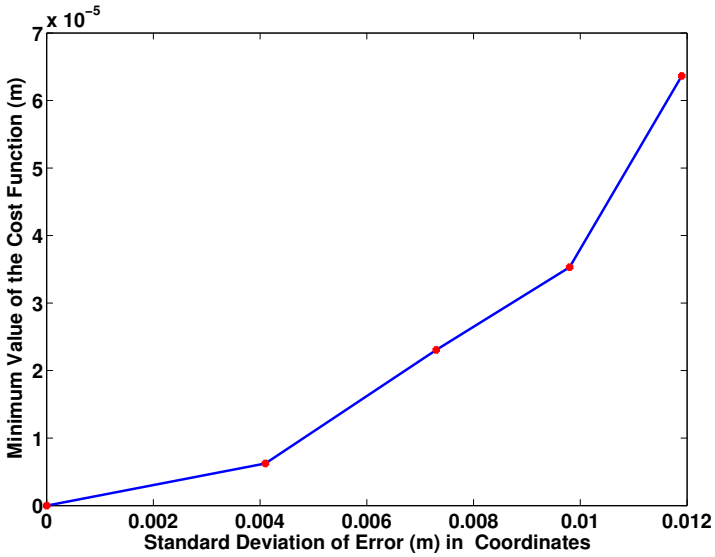


Figure 23: Minimum value of the cost function for estimating a flat water surface parameters while adding different levels of noise to the actual coordinates of 11 rays shown in Figure 22(a). As the level of noise in the input data increases, the minimum value of the cost function increases.

and 1.16 respectively. In other words, it is assumed that the projector is located 1.16(m) above a water surface with slope -0.07. The (X, Z) pairs for 11 rays were calculated using Eq.(11) to Eq.(14) . In addition, random noise was added to (X, Z) coordinates with a relatively high standard deviation $\delta = 0.0113$ (m). That is, after calculating the coordinates of each point in the camera coordinate system and transferring them to the projector coordinate system, the (X, Z) pairs have error with $\delta = 0.0113$ (m). Figure 22(b) shows the cost function over the range of a_1 and a_2 for this dataset with increment 0.001 for both parameters. The minimum value of the cost function was $4.7357e-05$ (m) that happens at $a_1 = -0.054$ and $a_2 = 1.183$. This shows even in the presence of relatively high amount of noise, there is only 23 millimeter error in the estimated height and 0.016 error in the slope which is a good estimate for the water surface parameters.

We added different levels of noise to the coordinates of those 11 rays and repeated the above procedure to estimate the parameters of the flat water surface. Figure 23 plots the minimum value of the cost function versus standard deviation of errors in the input coordinates. As it is expected, the higher the level of noise in coordinates, the higher the value of minimum in the cost function.

In the second synthetic dataset, we use a sinusoidal function for the water surface, i.e. $Z_i^w = A\sin(kX_i^w - \phi) + B$. In this case, there are four parameters that need to be estimated and they were set to: $A = 0.1(\text{m})$, $k = 5(m^-1)$, $\phi = \frac{-\pi}{2}$, and $B = 1$. These parameters represent a setup consisting of a wavy water surface with amplitude 0.1(m), wavenumber $5(m^{-1})$, phase $\frac{-\pi}{2}$ and a projector located 1 meter above the average water surface. As in the previous case, the (X, Z) pairs for 11 rays were calculated using Eq.(11) to Eq.(14) . No random noise was added to the coordinates in this dataset. The cost function has four dimensions. Figure 24 shows profiles of the cost function for all four parameters. In each profile, 3 parameters were fixed and set to the values corresponding to the minimum of the four dimensional cost function. Since the signal was noiseless, the estimated value of the parameters are exactly the same as their actual values. Note that for k the minimum happens for two values, i.e. -5 and $+5$. Plugging either of these two values into the water surface equation does not affect the shape of the water surface and it will be identical for both values. In other words, only the positive values for wavenumber could be assessed.

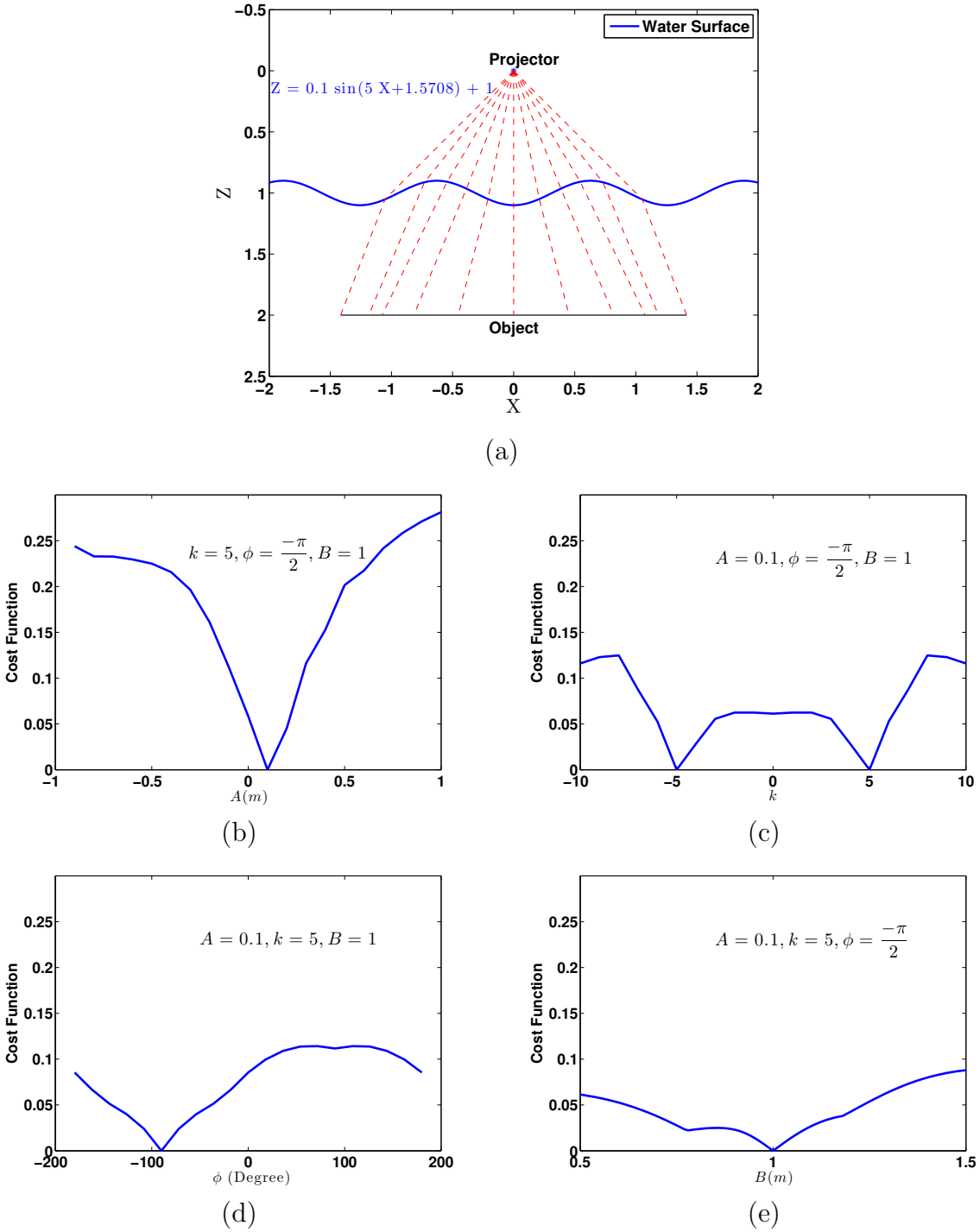


Figure 24: Cost function profiles for a synthetic Sinusoidal water surface of the form $Z_i^w = 0.1\sin(5X_i^w - \frac{-\pi}{2}) + 1$. Cost function profiles for a synthetic sinusoidal water surface of the form $Z_i^w = A\sin(kX_i^w - \phi) + B$ where $A = 0.1, k = 5, \phi = \frac{-\pi}{2}$, and $B = 1$. Each profile represents a cross section of the four dimensional cost function along one parameter: (a)Amplitude A, (b)Wavenumber k, (c)Phase ϕ, (d)Distance between the projector and the average water surface B.

4.1.2 Real data

All experimental data were acquired in the Air-Sea Interaction Saltwater Tank (ASIST) at the Rosenstiel School of Marine and Atmospheric Science, University of Miami. ASIST is a state-of-the-art wind-wave tank with fully transparent acrylic walls and its primary purpose is to study various physical processes at the air-sea interface. The working section of ASIST is 15(m) long, 1(m) wide and 1(m) high. For simulation of realistic ocean conditions, it is equipped with a mechanical wave generator (wave frequencies 0.25Hz-3Hz, amplitudes 0-0.1(m)), a wind generator (wind speed in the tank's centerline 0-30 (ms^{-1}), a current generator (current speed 0-0.5 (ms^{-1})), and a water temperature control (5-40 C).

The mechanical wave generator in ASIST is a flat vertical panel, moved horizontally by a hydraulic motor according to a pre-programmed trajectory. During all experiments with wavy water surface, the shape of the motor trajectory was exclusively a sine wave with various frequencies and amplitudes.

A Canon EOS Digital Rebel XS camera was used to capture images of the distorted pseudorandom code. The camera resolution was set to 3888×2592 pixels. An NEC's MT1065 projector was used to project the calibration checkerboard and the pseudorandom code. This projector is rated at 800:1 contrast and utilizes 3 LCDs equipped with micro lenses to increase light output. The resolution of the projector was set to its native resolution which is 1024×768 pixels (XGA). Also, the resolution of the output images from the computer was set to 1024×768 to avoid resampling of the pixels by the projector.

As a first step in 3D reconstruction, both camera and the projector were calibrated using the method presented in Section 3.2. The projector was installed above the water surface looking downward at a height of 1.24(m). The camera was located next to the transparent side wall of the tank. The camera-projector calibration was done while the tank was empty. Then we filled the tank with fresh water and calibrated the camera as though it was underwater. The calibration results for our setup are summarized in Table 1. The ratio of the focal length of the wet camera (filled tank) to the dry camera (empty tank) was 1.3369 which is close to the ratio of the refractive index of the water to the refractive index of the air.

Note that, it is assumed that the projector coordinate system coincides with the world coordinate system and thus all the extrinsic parameters for the projector are equal to zero. Also, for the robustness of the calibration, the principal point of the projector was set to the center of the projector. Fifteen images of the checkerboard at different angles were used to calibrate the camera. Also, nineteen images of the projected checkerboard on the actual checkerboard at different angles were used to calibrate the camera and the projector. All the image and the pattern coordinates used in the following experiments, were corrected for the lens distortion before applying the equations.

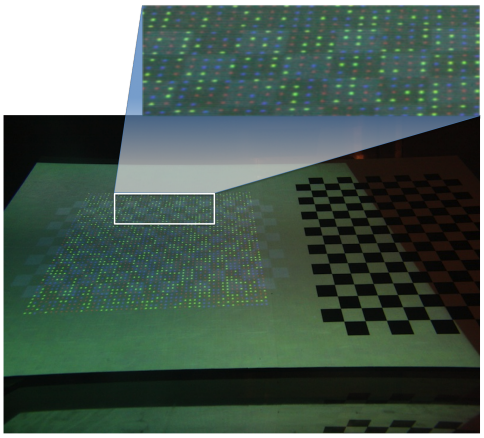
In the first experiment, a flat board was used as an object. A checkerboard was attached to this flat board so that the coordinates of every spot could be estimated using a calibrated camera as was described in Section 3.4. This board

	Symbol	Camera	Projector
Intrinsic Parameters			
Principal Point (Pixels)	x_c	1968.6	511.5
Principel Point (Pixels)	y_c	1545.3	383.5
Focal Length (Pixels)	f_x	4424.8	1802.1
Focal Length (Pixels)	f_y	4451.2	1835.6
Skewness Parameter	s	0	0
Pixel Size Ration	$\lambda_x = \frac{f_x}{f_y}$	0.9941	0.9817
Lens Distortion Parameters			
	k_1	0.0436	0.2406
	k_2	0.3627	-0.4427
	k_3	0.0261	0.0515
	k_4	-0.0021	0.0187
	k_5	0	0
Extrinsic Parameters			
Rotation (Degree)	θ_X	82.73	0
Rotation (Degree)	θ_Y	-7.01	0
Rotation (Degree)	θ_Z	4.92	0
Location (m)	X_c	0.1590	0
Location (m)	Y_c	0.7185	0
Location (m)	Z_c	1.3966	0

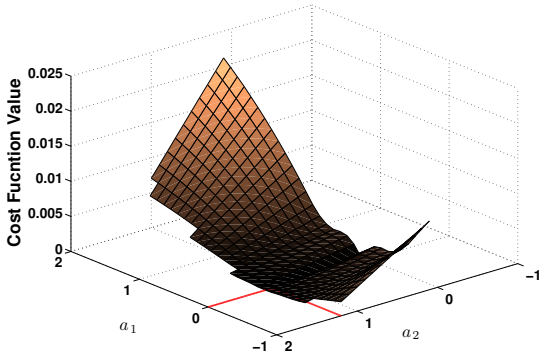
Table 1: System calibration results: Intrinsic and Extrinsic parameters for both camera and the projector

was placed in the tank with a slope such that it could be seen by camera while the projector could illuminate the pseudocode on top of it. In this experiment, the water surface was flat and no wave was generated. As can be seen in the acquired image of the pseudocode in Figure 25(b), the water surface did not impose any distortion, except scaling, in the pseudocode. This means that if we use a simple linear function $F(X_i^w, \mathbf{a}) = a_1 X_i^w + a_2$ to represent the water surface, we should expect the slope a_1 to be zero and a_2 to be equal to the distance between the projector and the water surface. Figure 25(a) shows the cost function for different combinations of a_1 and a_2. The increments for a_1 and a_2 were equal to 0.1. The dark areas show lower values and the bright areas show the higher values of the cost function. Where there is no intersection between predicted refracted rays and the object (board) within a defined limit, the value of the cost function was set to infinity. As can be seen for different a_2's, wherever $a_1 = 0$ the minimum of the evaluation function occurs. The minimum value of all pairs of a_1 and a_2 occurs at $a_1 = 0$ and $a_2 = 1.2$ which is indicated by the intersection of two lines on Figure 25(a). Therefor, the best representation for the water surface function would be $F(X_i^w, \mathbf{a}) = 0X + 1.2$.

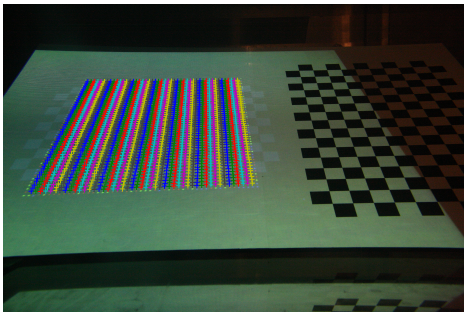
In the second experiment, we took images of the projected pseudocode approximately every 10 minutes while the tank was filling up with a constant rate. Figure 26(a) shows 6 samples of captured images. The water surface in all these images was flat and no wave was generated. With increasing height of the water surface the pseudocode pattern shrinks and this effect can be seen in the



(a)

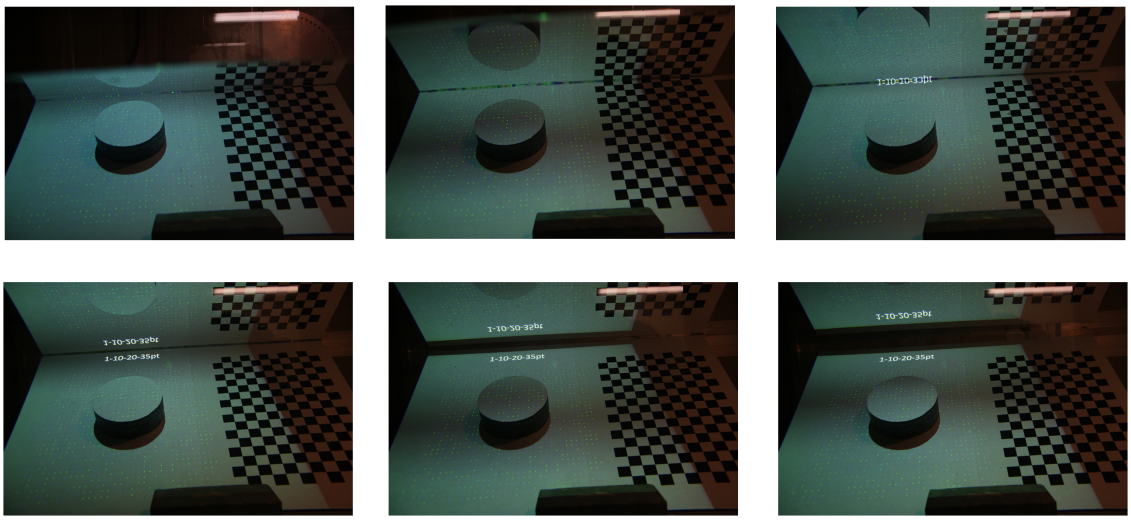


(b)

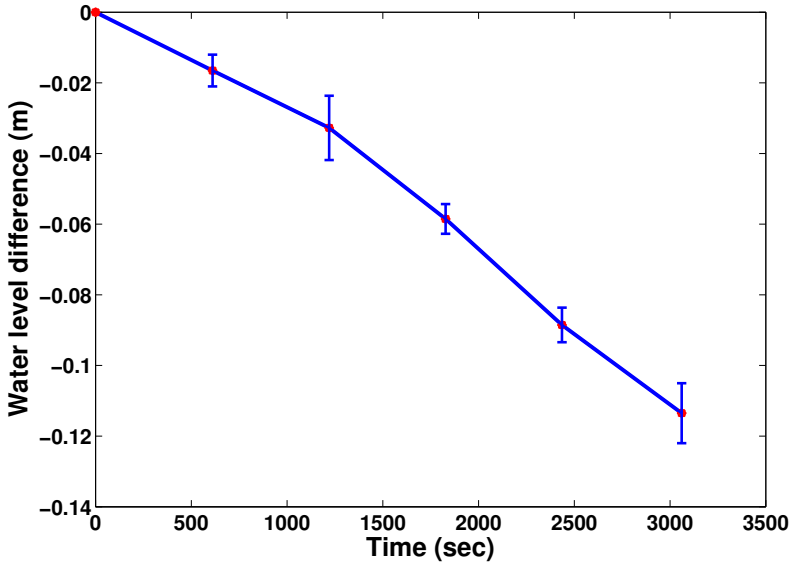


(c)

Figure 25: (a) Image of the pseudocode pattern on a flat board under a flat water surface (no wave), (b) Corresponding evaluation function for $F(X_i^w, \mathbf{a}) = ax + b$ for different pairs of a and b with the increment of 0.1, (c)Extracted pseudocode pattern.



(a)



(b)

Figure 26: (a) Estimated water level differences in specific time intervals while filling up the tank with fresh water at a constant rate, (b) corresponding images of projected pseudocodes.

consecutive images in Figure 26(a). Note that the camera and the projector are fixed and the distance between the projector and the water surface decreases as the tank is being filled up. Then the water surface for each image is estimated, assuming the water surface follows a flat surface model, i.e. $F(X_i^w, \mathbf{a}) = a_1 X_i^w + a_2$ as in the previous experiment. The estimated a_1 for all cases were close to zero. Figure 26(b) shows the estimated height difference for the water surface versus time. As expected, the plot has a linear pattern and was consistent with the constant rate of filling up the tank.

In the next experiment, we generated a wavy water surface using the mechanical wave generator. It was set to generate a water surface wave using an input signal of frequency 2 (Hz), maximum voltage 1 (volt) and minimum voltage 0.5 (volt). The wave was sinusoidal and it is assumed to be modeled as $F(X_i^w, \mathbf{a}) = A\sin(kX - \phi) + B$. In this model, A is amplitude, k is wavenumber, ϕ is shift, and B represents the distance between the projector and the mean water surface. The mechanical wave maker settings results in a wave with $k = 16$ (m^{-1}) and $A = 2.5$ (cm). Figure 27(a) shows the image of the pseudocode captured under the wavy water surface. As highlighted in the image, even under the effect of high curvature waves, the pseudocode pattern is detectable and can be extracted. Although it might not be possible to extract all the projected rays, it is possible to extract a large enough number of rays in order to recover the water surface parameters. In this experiment, the sinusoidal model has four parameters and at least four rays are necessary to trigger the water surface estimation.

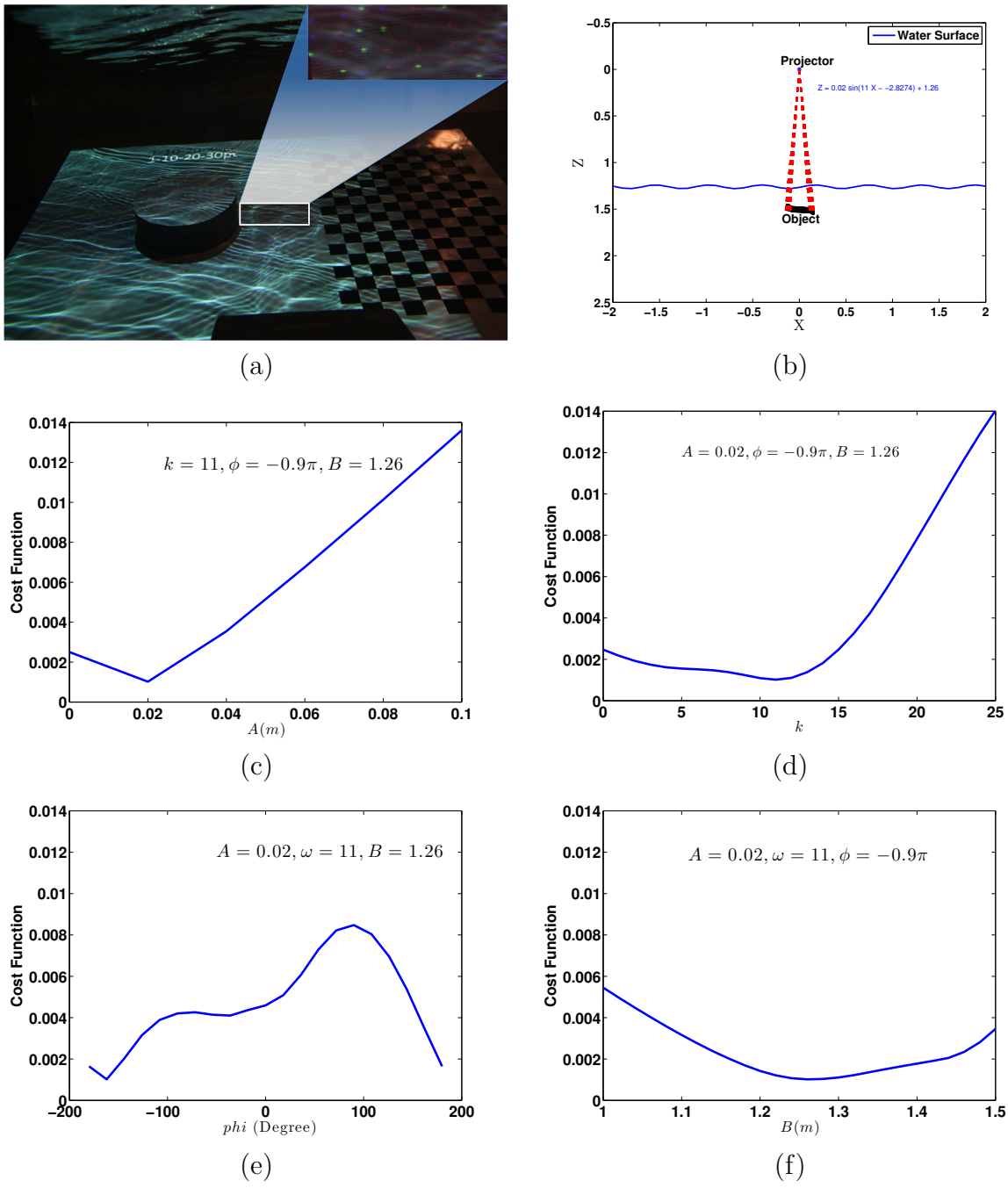


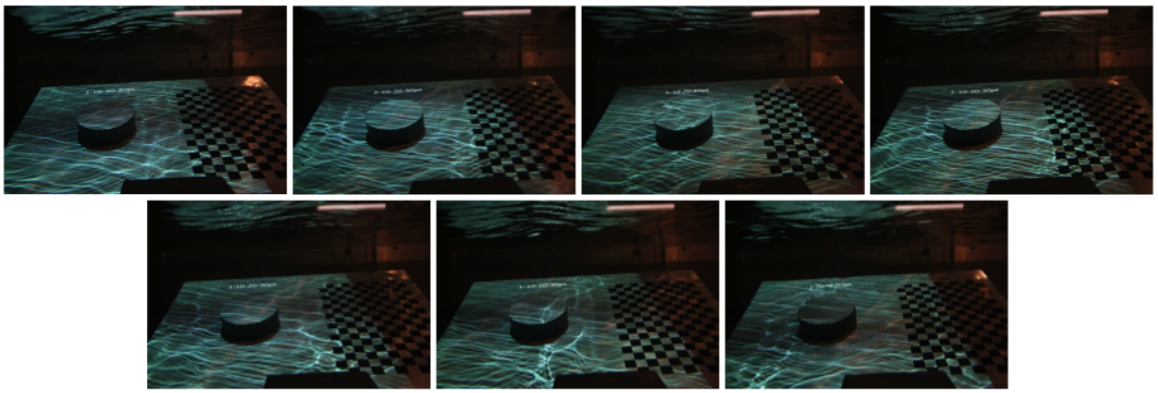
Figure 27: (a) An image used for water surface reconstruction, the highlighted section shows that even under the effect of high curvature waves, white flickers, the pseudocode can easily be detected, (b) Reconstructed sinusoidal water surface, the object, and the projector, (c)-(f) Cost function profiles for a wavy water surface modeled as a sinusoidal function of the form $Z_i^w = A\sin(kX_i^w - \phi) + B$. Each profile represents a cross section of the four dimensional cost function along one parameter: (c)Amplitude A, (d)Wavenumber k, (e)Phase ϕ, (f)Distance between the projector and the average water surface B.

We calculated the cost function for the entire span of the possible solutions to find the best set of four parameters (A, k, ϕ, B) within defined limits for each parameter; $A \in [0\ 0.1]$(m) with 0.02(m) increment step, $k \in [0\ 25]$ with 1 increment step, $\phi \in [-\pi\ \pi]$ with $\frac{\pi}{10}$ increment step, and $B \in [1\ 1.5]$(m) with 0.02 (m) increment step. The minimum of the cost function happens at $A = 0.02(m)$, $k = 11(m^{-}1)$, $\phi = -2.8274$, and $B = 1.26(m)$. Figure 27(c)-(f) show the profiles of the cost function for all four parameters at the minimum. The parameter B which represents the distance between the projector and the average height of the water surface, is very close to our measurements, i.e. 1.24 (m). In our setup, since the area covered by the entire pseudocode pattern does not contain several wavenumbers, the difference between the cost function for $k \in (0, 16)$ is small (see Figure 27 (d)). However, the cost function drastically increases for higher values of k. Therefore, in this experiment, by plugging values less than 16 for k into the water surface model, the difference between the resulting shapes of the water surface in the region covered by the entire pseudocode would not be distinct. If a better estimate in the final value of k is needed, the pseudocode has to cover a large enough area to at least encompass one full wavelength. Figure 27(b) demonstrates the position of the projector, object, and the estimated water surface with respect to each other.

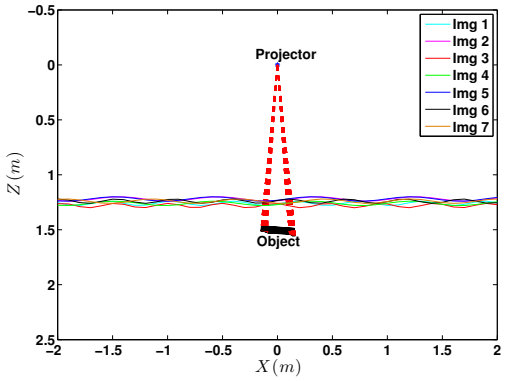
Assuming a sinusoidal model, the water surface parameters were calculated for seven consecutive images. The images were taken at 5 second intervals while the wave generator was creating a wavy water surface with the same characteristics.

Figure 28(a) shows all seven consecutive images. Ideally estimated A, B, and k should be the same for all these images while the shift parameter ϕ might be different for each image. As in the previous experiment, the cost function for different combinations of parameters were calculated and the set of parameters corresponding to the minimum value of the cost function were picked as the final solution. Table 2 summarizes the final results for all these images. As expected, the parameter A is equal for all seven images. Note that we used a 2 (cm) step size for A in calculating the cost function and if a finer step size were used, the estimated values would be slightly different but within an acceptable range. Also, the distance between the projector and the average water surface, B, for all images are within the expected range. As previously mentioned, since the pseudocode pattern does not contain a full wavelength of the generated wave, the value of k does not affect drastically the shape of the final water surface, i.e. there is a subtle difference between the water surfaces obtained by plugging different numbers of $k < 16$ into the sinusoidal water surface model. For $k >= 16$ the water surface model represents one (or more than one) full wavelength within the borders of pseudocode pattern and the effect of k would be more visible. Given that, in all images we obtained a good estimate of k. The value for shift parameter ϕ as expected is different for each image. Figure 28(b)-(c) shows all the reconstructed water surfaces using the parameters estimated from seven images.

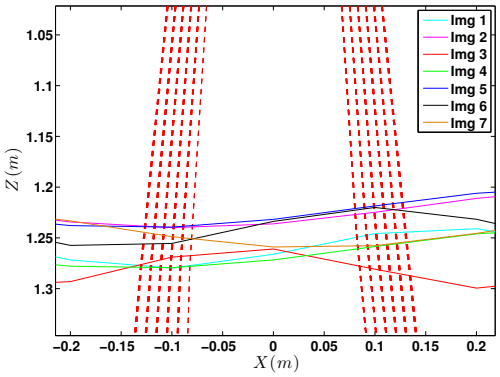
A better way to assess the accuracy of the proposed method is to reconstruct an object with known geometric shape and compare the reconstructed shape



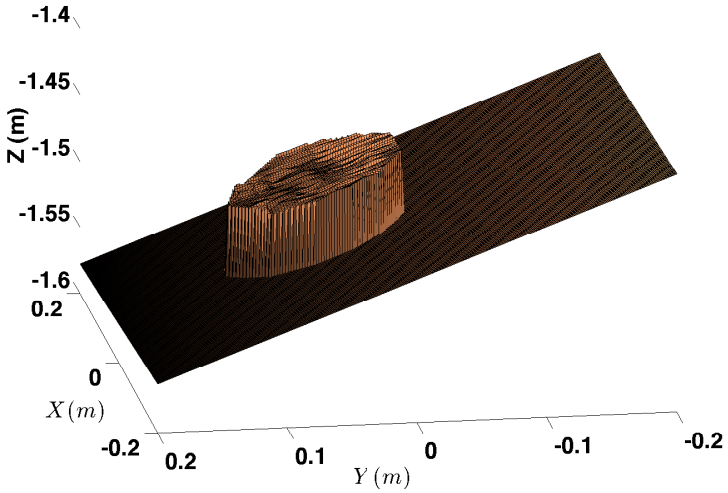
(a)



(b)



(c)



(d)

Figure 28: (a)Seven consecutive images of a flat board and cylindrical object at the presence of wavy water surface taken by an underwater camera, (b) Reconstructed sinusoidal water surfaces for each image with respect to the location of the projector and the object, (c) Reconstructed water surfaces are highlighted for better visual representation, (d) Reconstructed cylindrical object attached to the board using combined data from all seven images.

Image Number	1	2	3	4	5	6	7
A	0.02	0.02	0.02	0.02	0.02	0.02	0.02
k	11	7	13	7	7	12	8
ϕ	-2.83	-2.20	1.26	-2.51	-2.51	2.83	-1.26
B	1.26	1.22	1.28	1.26	1.22	1.24	1.24
Avg. Error (m)	0.002	0.0023	0.0019	0.0016	0.0018	0.0023	0.0013

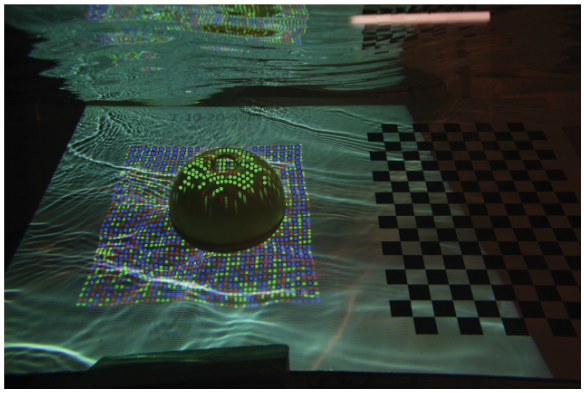
Table 2: Estimated water surface parameters for seven images shown in Figure 28(a) using a sinusoidal model $Z = A\sin(kX - \phi) + B$ for the water surface.

with the original one. For this purpose, we used a cylindrical object with known thickness (0.053 meter) and attached it to a board. Several images of this object were taken while a wavy water surface was generated (see Figure 28(a)). The water surface was estimated for each image in the previous experiment and the results are summarized in Table 2. The increment steps for all parameters of the sinusoidal model are the same as in the previous experiment. Then, using the estimated water surface parameters and Eq.(21) and Eq.(22), the three dimensional coordinates of the projected pattern on both the cylindrical object and the board were calculated. A plane was passed through all the points on the board. Using the 3D coordinates of the points on the cylindrical object and the equation of the plane, the distance from each point on top of the cylinder to the plane was calculated. This distance for all the points should be the same and equal to the thickness of the cylinder. Cumulatively, in seven images, 166 points were extracted from the pattern on top of the cylindrical object. The distances for points to the board were calculated. The average error in the distance for these points was zero and the standard deviation of error was 0.0023 meter. Higher accuracy could be

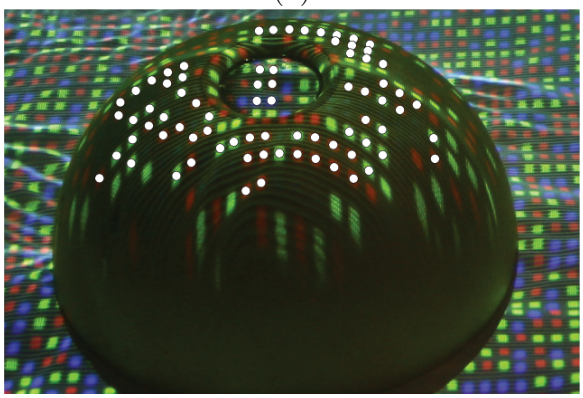
achieved if the rays were thinner and sharper. However, in the presence of waves, the thiner rays could easily disappear and are more difficult to detect in the images. Figure 28(d) shows the reconstructed geometric shape of the board and the attached cylindrical object using all the points obtained from the seven images. Note that we can reconstruct the geometric shape of the cylindrical object with a single image, however, in this experiment we used the single shot method in a multi-shot scheme to obtain more points on top of the object.

In order to show that this method works for non-planar objects, we cut a spherical object (a round buoy) in half and attached it to the board. The radius of the buoy was 83 millimeter. We projected the pseudocode pattern on this object while the wave generator was making a wavy water surface (See Figure 29(a)). Again we calculated the water surface parameters using the projected rays on the board and eventually the coordinates of projected rays on this object. The white dots on the object shown in Figure 29(b) illustrate the extracted rays on the surface of the spherical object. Then we fit a sphere with 83 millimeter radius to these points to measure the accuracy of the estimated coordinates. The mean of the residuals was close to zero (2.5667e-04 meter) and their standard deviation was equal to 0.0034 meter. The reconstructed object, the extracted points on the object and the plate are shown in Figure 29(c).

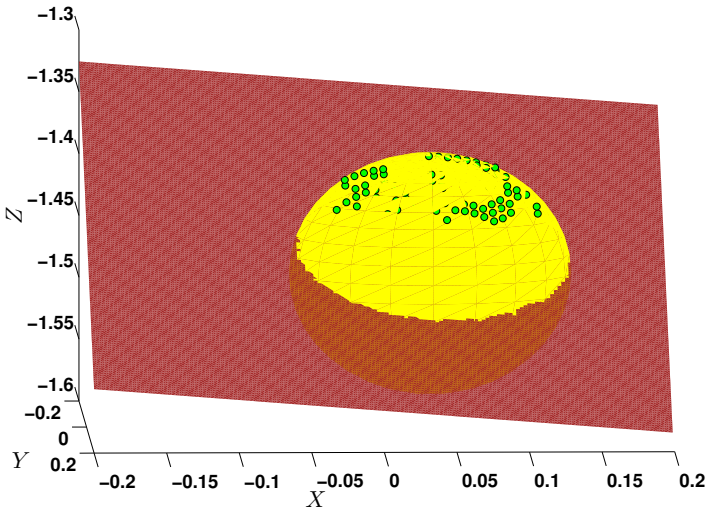
Note that even in the presence of high curvature waves, we were still able to locate the projected rays and obtain an estimate of the water surface which was good enough for our purpose of recovering the 3D shape of underwater objects.



(a)



(b)



(c)

Figure 29: (a)A spherical object attached to a board illuminated by a pseudocode pattern emitted from a projector located above water, (b)White dots depict extracted pseudocode pattern on top of the object, (c)Reconstructed spherical object attached to the board using data from one single image. Note: bigger pseudocodes are easier to extract under a wavy water surface.

4.2 Sensitivity Analysis

It is important to understand the behavior of the proposed method in the presence of error in the input parameters. Inevitably there would be error in the water surface estimation and in the system calibration. The question is how these errors will affect the accuracy of the final reconstructed object. Sensitivity analysis helps us answer this question by measuring changes in output values while changing the values of input parameters. It determines which parameters are the most important and most likely to affect system behavior and/or predictions of the model. Following a sensitivity analysis, values of critical parameters can be refined while parameters that have little effect can be simplified or ignored.

There are many common ways to do sensitivity analysis. A partial derivative can be a sensitivity function for a system with well-defined equations. Statistical techniques can also be used when the model is too complex and the inputs are described with probability distributions.

The sensitivity functions can be categorized into analytic and empirical functions [87]. Analytic sensitivity functions are used when the system under study is well defined and mathematically well behaved (e.g. continuous derivatives). These generally take the form of partial derivatives. They are convenient and have an advantage in that the sensitivity of a system to a given parameter is given as a function of all other parameters and can be plotted as functions of these variables.

Empirical sensitivity functions are often just point evaluations of a system's sensitivity to a given parameter(s) when other parameters are at known, fixed values. They are generally determined by observing the changes in output of a computer simulation as model parameters are varied from run to run. Their advantage is that they are often simpler than their analytic counterparts or can even be determined for an un-modeled physical system [87].

Since our model is mathematically well behaved and taking derivatives is feasible, we will use analytic sensitivity functions to evaluate the effects of error in the input parameters on the model's output. Specifically, we will assess the effect of error in height and slope of the estimated water surface as well as the error in the camera-projector baseline as the representative for the system calibration.

The following types of analytic sensitivity functions are common. They are all based on finding the partial derivative of a mathematical system model with respect to some parameter [87].

- **Absolute**

  The absolute-sensitivity of the function F to variations in the parameter α is

  $$S_{\alpha}^{F} = \frac{\partial F}{\partial \alpha} \mid_{NOP} \tag{24}$$

  It should be evaluated at the normal operating point (NOP) where all the parameters have their normal values. The function F must be differentiable with respect to α and may also be a function of other parameters. Absolute-sensitivity functions are useful for calculating output errors due to parameter

variations and for assessing the times at which a parameter has its greatest or least effect.

- **Relative**

  The relative-sensitivity of the function F to variations in the parameter α is

$$\overline{S}_{\alpha}^{F} = \frac{\%\text{Change in } F}{\%\text{Change in } \alpha} = \frac{\Delta F/F}{\Delta\alpha/\alpha} = \frac{\partial F}{\partial \alpha} \mid_{NOP} \frac{\alpha_0}{F_0} \tag{25}$$

  Similar to the absolute sensitivity, the relative sensitivity should be evaluated at the normal operating point (NOP). Relative sensitivity functions are used to compare the effects of different parameters because they are dimensionless, normalized functions. The lack of dimensions in these functions allow easy comparison of parameters' changes on model outputs even though the parameters may describe widely varying aspects of the model and have different units. As it is shown in Eq.(25), the relative sensitivity functions are formed by multiplying the absolute sensitivity function by the normal value of the parameter and dividing by the normal value of the function.

- **Semi-relative**

  The semi-relative sensitivity of the function F to variations in the parameter α is

$$\widetilde{S}_{\alpha}^{F} = \frac{\partial F}{\partial \alpha} \mid_{NOP} \alpha_0 \tag{26}$$

  Again, the semi-relative sensitivity function should be evaluated at the normal operating point (NOP). As can be seen by the definition, semi-relative sensitivity functions are scaled version of absolute sensitivity functions and

therefore, will have the same shape as absolute-sensitivity functions. Notice that, when a parameter has a value of zero, its semi-relative sensitivity function is also zero.

We will use both absolute-sensitivity and relative-sensitivity functions as the criterions for our assessments.

4.2.1 Sensitivity of the height of water surface

As it is shown in Figure 21 in the previous chapter and formulated in Eq.(12), the X coordinate of a point is the sum of corresponding displacement in air and water, i.e. ΔX_i^a and ΔX_i^w (These formulations can be similarly applied to Y coordinate as well). Error in the height of the water surface will affect both of these displacements. The absolute sensitivity function of X to the height of water surface Z_i^w has the following form

$$S_{Z_i^w}^{X_i} = \frac{\partial X_i}{\partial Z_i^w} = \frac{\partial \Delta X_i^a}{\partial Z_i^w} + \frac{\partial \Delta X_i^w}{\partial Z_i^w} = \cot(\alpha_i) + \cot(\alpha_i + \beta_i) \tag{27}$$

As can be seen, the absolute sensitivity function for the height of water surface is a function of two angles β_i and α_i (See Figure 21 for illustration of these two angles) and is not a function of height itself. Figure 30 depicts $S_{Z_i^w}^{X_i}$ versus different β_i and α_i. As expected, the error in height does not have any effect on the ray at the nadir of the projector if the incidence angle is zero. In this case the effect of error in estimated height of water surface on both ΔX_i^a and ΔX_i^w is equal to zero. Remember that β_i is a function of incidence angle θ_i^a and if the incidence angle is zero, the β_i will also be zero. In fact, there is a (dark-blue) curve in Figure

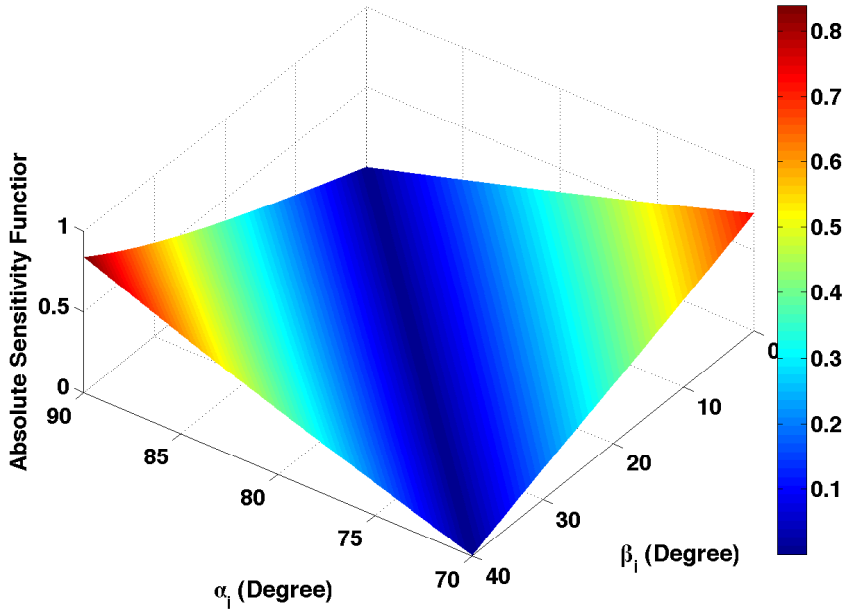


Figure 30: Absolute sensitivity function for reconstructed X (or Y) coordinates to the height of water surface. The function is plotted for different values of β_i and α_i.

30 which represents pairs of β_i and α_i where the effect of error in height is zero. Essentially, for these pairs, the error in ΔX_i^a and ΔX_i^w due to change in height will cancel out each other and their addition is equal to zero. This curve will be distinct for each water type and will change base on the refractive index of the water.

The relative sensitivity function for the height of water surface can be obtained by multiplying the absolute sensitivity function by the height of water surface Z_i^w and dividing by X_i, i.e.

$$\overline{S}_{Z_i^w}^{X_i} = S_{Z_i^w}^{X_i} \frac{Z_i^w}{X_i} = S_{Z_i^w}^{X_i} \frac{Z_i^w}{(Z_i^w - Z_{proj}) \cot(\alpha_i) + (Z_i - Z_i^w) \cot(\alpha_i + \beta_i)} \tag{28}$$

As can be seen, the denominator X_i has a similar form as absolute sensitivity function (see Eq.(27) for comparison). By looking at the absolute sensitivity func-

tion in Eq.(27), we realize that it has two components: Above-water component $\cot(\alpha_i)$ and Underwater component $\cot(\alpha_i + \beta_i)$. The X_i is the weighted sum of these two components and the weights are the distances in the Z direction that the ray travels from the projector to the object above water $(Z_i^w - Z_{proj})$ and underwater $(Z_i - Z_i^w)$. In other words, in this case, the relative sensitivity function $\overline{S}_{Z_i^w}^{X_i}$ is a scaled version of the absolute sensitivity function $S_{Z_i^w}^{X_i}$, with scale being (Z_i^w), divided by the weighted sum of the two components of the absolute sensitivity function.

Figure 31(a) shows the relative sensitivity function $\overline{S}_{Z_i^w}^{X_i}$ for our experimental setup. In this case, the projector is located at 1.24 (m) above the water surface and the center of the object is at the depth equal to 16 (cm). In addition, we plotted the relative sensitivity function for two more cases. Figure 31(b) shows the relative sensitivity for the same parameters as in our experimental setup except the object is assumed to be located 5 meters below water (for more deeper objects). Also, in Figure 31(c), the only parameter that is different from the parameters in our setup is the height of the water surface. As can be seen in Eq.(28) and illustrated in Figure31(a)-(c) , when the denominator is zero (or close to zero), the relative sensitivity of the reconstructed X coordinate to the height of the water surface is not valid. Also, note that the pairs of β_i and α_i with non-valid relative sensitivity are different in each setup. In these cases, it is recommended to use absolute sensitivity as a criterion for sensitivity analysis.

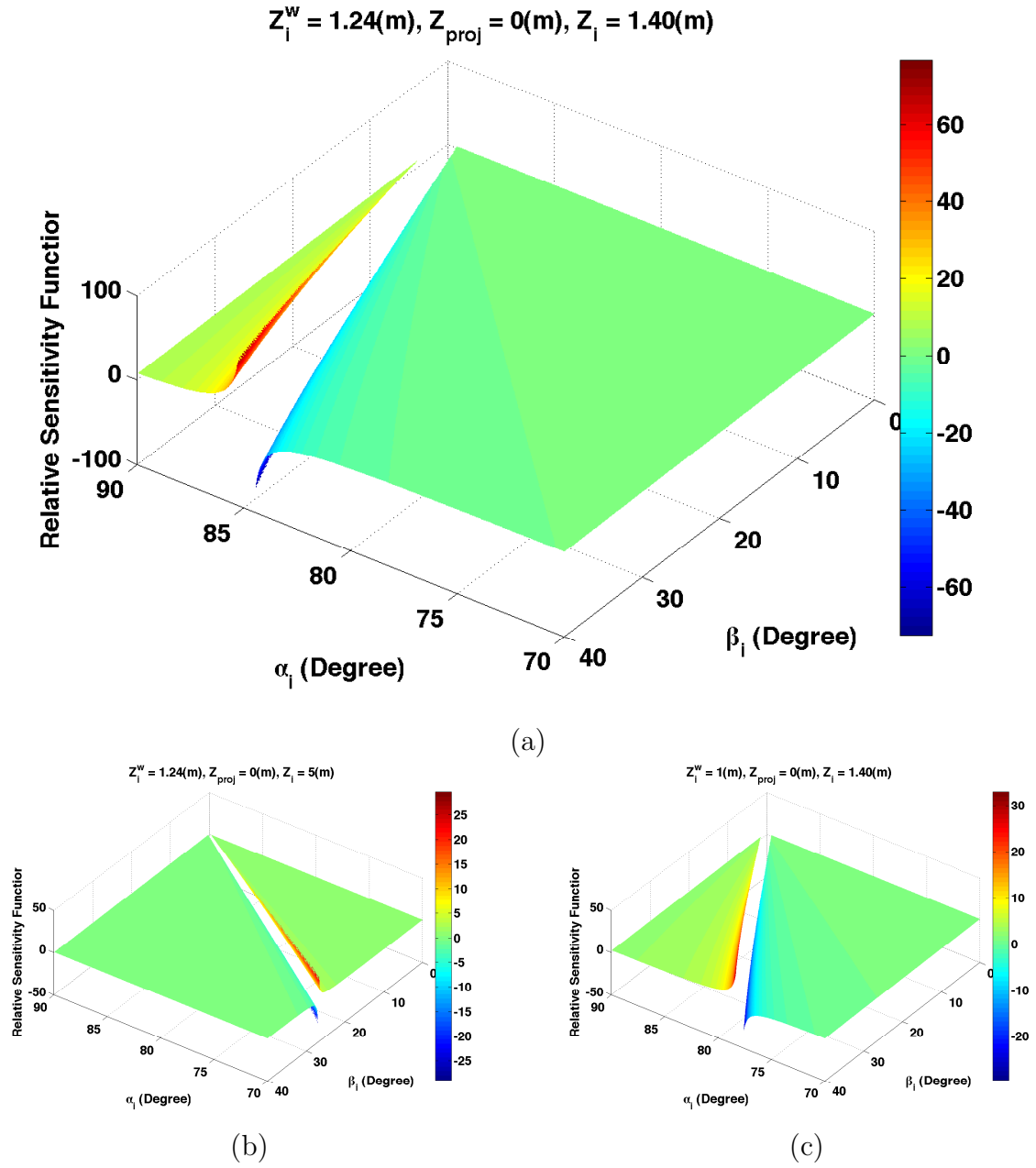


Figure 31: Relative sensitivity function for the reconstructed X or Y coordinates to the height of water surface. The function is plotted for different values of β_i and α_i. Plot of relative sensitivity for the (a) same parameters as in our experimental setup, (b) same parameters as in our experimental setup except the object is at the depth equal to 5 meters, (c) same parameters as in our experimental setup except the projector is closer to the water surface.

4.2.2 Sensitivity of the slope of water surface

The second input parameter of interest is the slope of the water surface. Using the chain rule for computing the derivative of the X with respect to the slope of the water surface, the absolute sensitivity function will be:

$$S_m^{X_i} = \frac{\partial X_i}{\partial m} = \frac{\partial X_i}{\partial \beta_i}\frac{\partial \beta_i}{\partial \theta_i^a}\frac{\partial \theta_i^a}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial m} \tag{29}$$

where

$$\frac{\partial X_i}{\partial \beta_i} = \frac{\partial \Delta X_i^w}{\partial \beta_i} = -(Z_i^w - Z_i)\csc^2(\alpha_i + \beta_i) \tag{30}$$

$$\frac{\partial \beta_i}{\partial \theta_i^a} = 1 - \frac{n_a \cos(\theta_i^a)}{n_w\sqrt{1 - \frac{n_a^2 \sin^2(\theta_i^a)}{n_w^2}}} \tag{31}$$

$$\frac{\partial \theta_i^a}{\partial \gamma_i} = 1 \tag{32}$$

$$\frac{\partial \gamma_i}{\partial m} = \frac{1}{m^2 + 1} \tag{33}$$

In above equations, m is the slope of the water surface and the rest of the parameters are shown in Figure 21. As can be seen, the absolute sensitivity is a function of slope. The absolute sensitivity function for different values of slope and α_i, using parameters of our experimental setting, is shown in Figure 32(a). This function will also have a different shape if the object is at different depth. The deeper the object, the more severe would be the effect of change in slope. This effect is demonstrated in Figure 32(b)-(c). Assuming the object is located 2 meters below the projector and the depth of the object changes from 1 meter to 0.5 meter

by changing only the level of water, one can see that the absolute sensitivity to the slope is higher at 1 meter depth compared to 0.5 meter for all the slopes and α_i's. Also, contrary to the height of water surface parameter, at the nadir of the projector, the change in the slope has its maximum effect on the coordinate X.

The relative sensitivity function for the reconstructed X (or Y) coordinates to the slope of water surface can be obtained using Eq.(25)

$$\overline{S}_m^{X_i} = S_m^{X_i} \frac{m}{(Z_i^w - Z_{proj})\cot(\alpha_i) + (Z_i - Z_i^w)\cot(\alpha_i + \beta_i)} \tag{34}$$

As can be seen, the denominators in Eq.(34) and Eq.(28) are the same. Similarly, the relative sensitivity function to the slope is not valid when the denominator is zero (or close to zero). Figure 33 (a) shows the relative sensitivity using parameters of our experimental setting for different slopes and α_i's. In addition, as in absolute sensitivity to slope, we plotted the relative sensitivity to slope for two hypothetical settings; one with an object at 1 meter depth (Figure 33(b)) and one with an object at 0.5 meter depth (Figure 33(c)). For both of these settings, the object is 2 meters away from the projector and only the height of water surface changes. Note that, as in relative sensitivity to the height of water surface, the pairs of parameters (m, α_i) with non-valid relative sensitivity to the slope of water surface are different in each setup. The reason is that due to the change in the height of water surface or the height of object, the denominator in both sensitivities becomes zero (or close to zero) for different pairs of parameters.

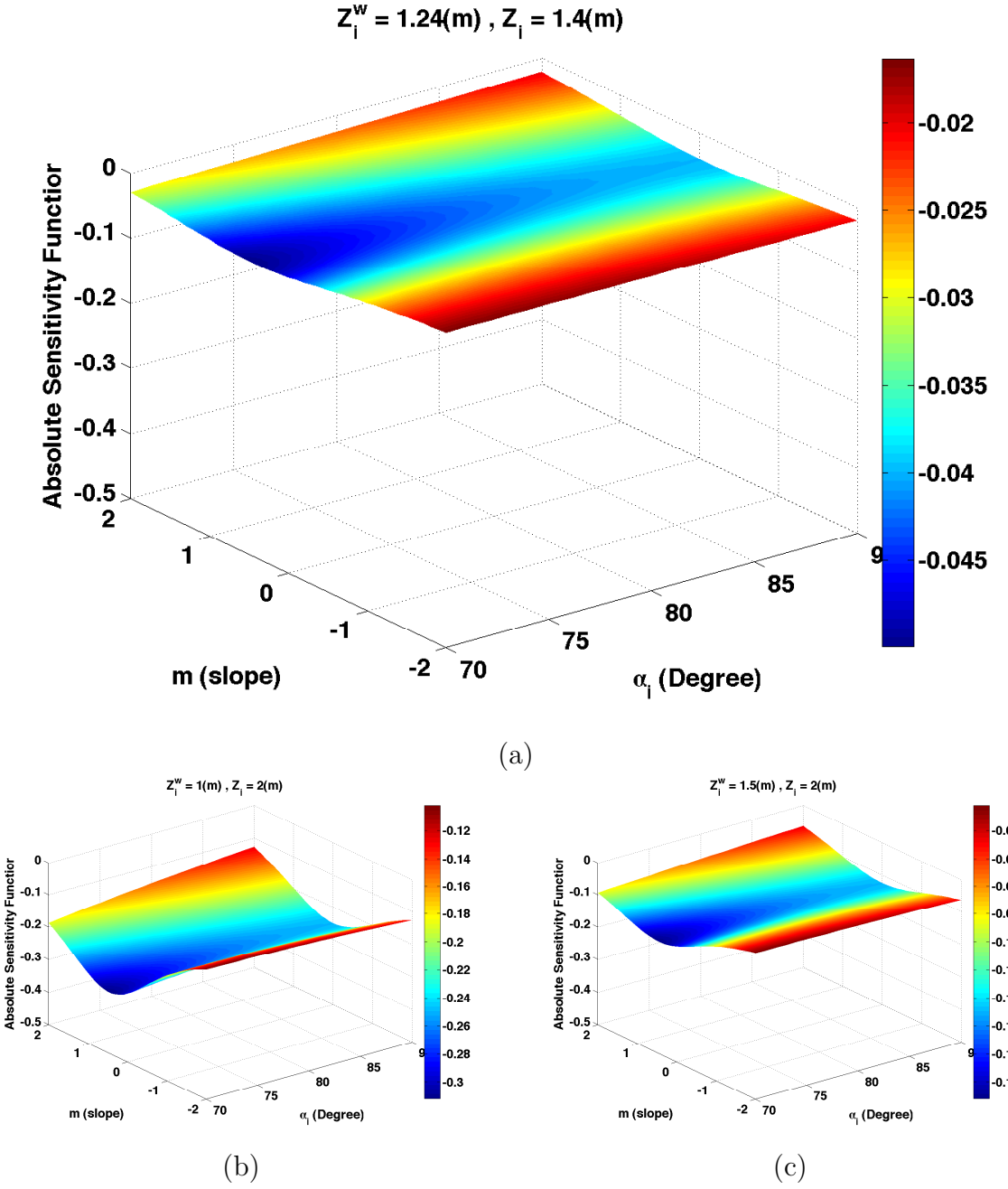
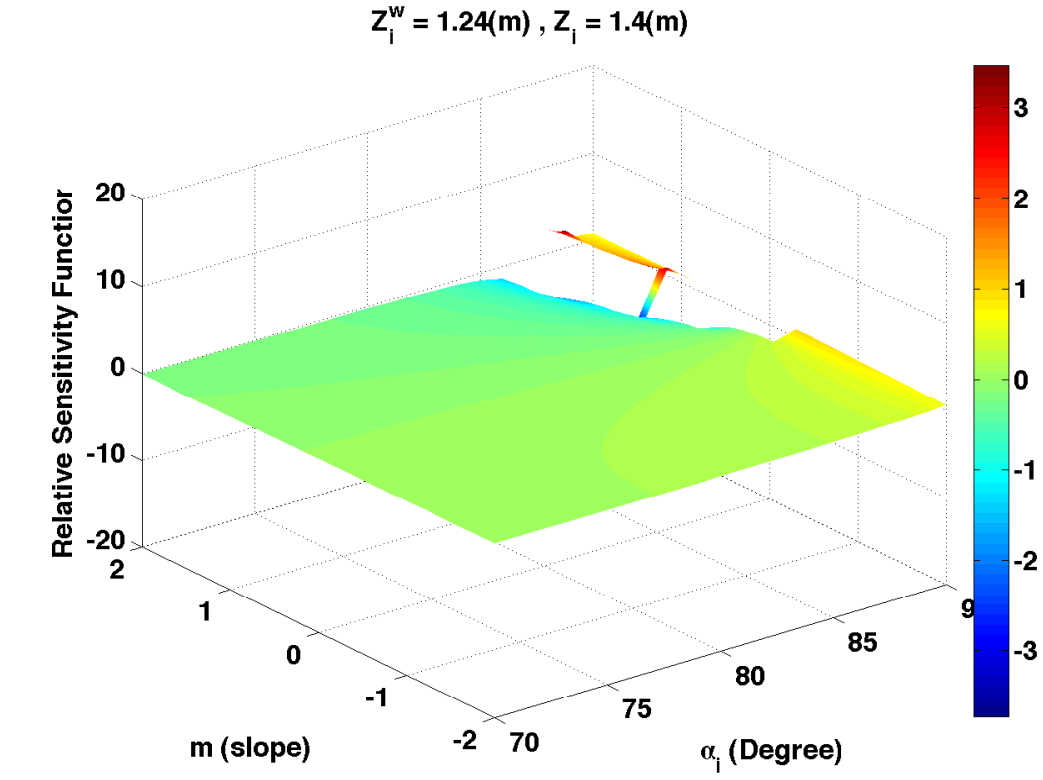
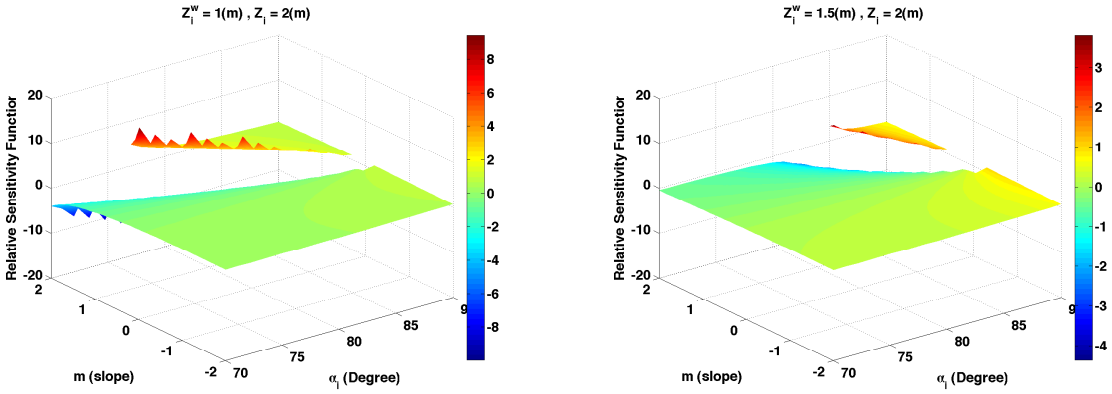


Figure 32: Absolute sensitivity function for the reconstructed X (or Y) coordinates to the slope of water surface. The sensitivity function is a function of slope itself. Absolute sensitivity for (a) our experimental setup $Z_i^w = 1.24(m)$ and $Z_i = 1.4(m)$, (b) hypothetical setup, $Z_i^w = 1(m)$ and deeper object $Z_i = 2(m)$, (c) hypothetical setup, $Z_i^w = 1.5(m)$ and $Z_i = 2(m)$.



(a)



(b) (c)

Figure 33: Relative sensitivity function for the reconstructed X (or Y) coordinates to the slope of water surface. Relative sensitivity for (a) our experimental setup $Z_i^w = 1.24(m)$ and $Z_i = 1.4(m)$, (b) hypothetical setup, $Z_i^w = 1(m)$ and deeper object $Z_i = 2(m)$, (c) hypothetical setup, $Z_i^w = 1.5(m)$ and $Z_i = 2(m)$.

4.2.3 Sensitivity of the baseline

The last parameter for sensitivity analysis is the baseline between the camera and the projector B. Unlike the height and slope of water surface, this parameter is estimated in the system calibration step. We assume there is no other calibration error except error in the baseline estimation. We are investigating the effect of this error on the distance between the object and the camera (D).

According to Figure 21, B^w is equal to

$$B^w = B\frac{\sin(\psi_c)}{\sin(\psi_c + \psi_p)} - B^a \tag{35}$$

where B^a is the distance a ray travels from the projector to the water surface in the air. The absolute sensitivity of the distance between the object and the camera D to the camera-projector baseline B is the partial derivative of D with respect to B and can be obtained using

$$S_B^D = \frac{\partial D}{\partial B} = \frac{\partial D_1}{\partial B} + \frac{\partial \Delta D}{\partial B} \tag{36}$$

where

$$\frac{\partial D_1}{\partial B} = \frac{\sin(\psi_p)}{\sin(\psi_c + \psi_p)} \tag{37}$$

and

$$\frac{\partial \Delta D}{\partial B} = \frac{\sin(\psi_c)}{\sin(\psi_c + \psi_p)}\frac{\sin(\beta_i)}{\sin(\beta_i + \psi_w)} \tag{38}$$

Notice that the absolute sensitivity of distance D to the baseline is only a function of angular parameters ψ_c, ψ_p, and β_i. As shown in Figure 21, the

angle ψ_w can be obtained from calibration parameters ψ_c and ψ_p. Figure 34 shows the absolute sensitivity function for different ψ_c and ψ_p at a specific β_i. The overall shape of the absolute sensitivity function is similar for different β_i's. For higher difference between angles of incidence and refraction, i.e. β_i, the absolute sensitivity is higher. The absolute sensitivity also increases as the ψ_c and ψ_p increase. If the line of sight of the camera and the line of sight of the projector are parallel, both ψ_c and ψ_p will be 90 degrees for objects at infinity.

The relative sensitivity function for the distance between the object and the camera D to the camera-projector baseline B can be obtained by:

$$\overline{S}_B^D = S_B^D \frac{B}{D} \tag{39}$$

where distance D can be calculated by

$$D = B\frac{\sin(\psi_p)}{\sin(\psi_c + \psi_p)} + (B\frac{\sin(\psi_c)}{\sin(\psi_c + \psi_p)} - B^a)\frac{\sin(\beta_i)}{\sin(\beta_i + \psi_w)} \tag{40}$$

Figure 35 illustrates the relative sensitivity function for different β_i for our experimental setup where camera-projector baseline is 1.58 (m) and $B^a = 1.3$(m). As expected from Eq.(40) and illustrated in Figure 35(b), the relative sensitivity for rays with no refraction is equal to 1 for all pairs of ψ_c and ψ_p. By comparing Figure 35(a)-(c), one can see that the relative sensitivity of distance D to baseline B is higher for rays with higher refraction.

Now that we have calculated the relative sensitivity to different parameters, we can easily compare their effects. For example, consider the following relative

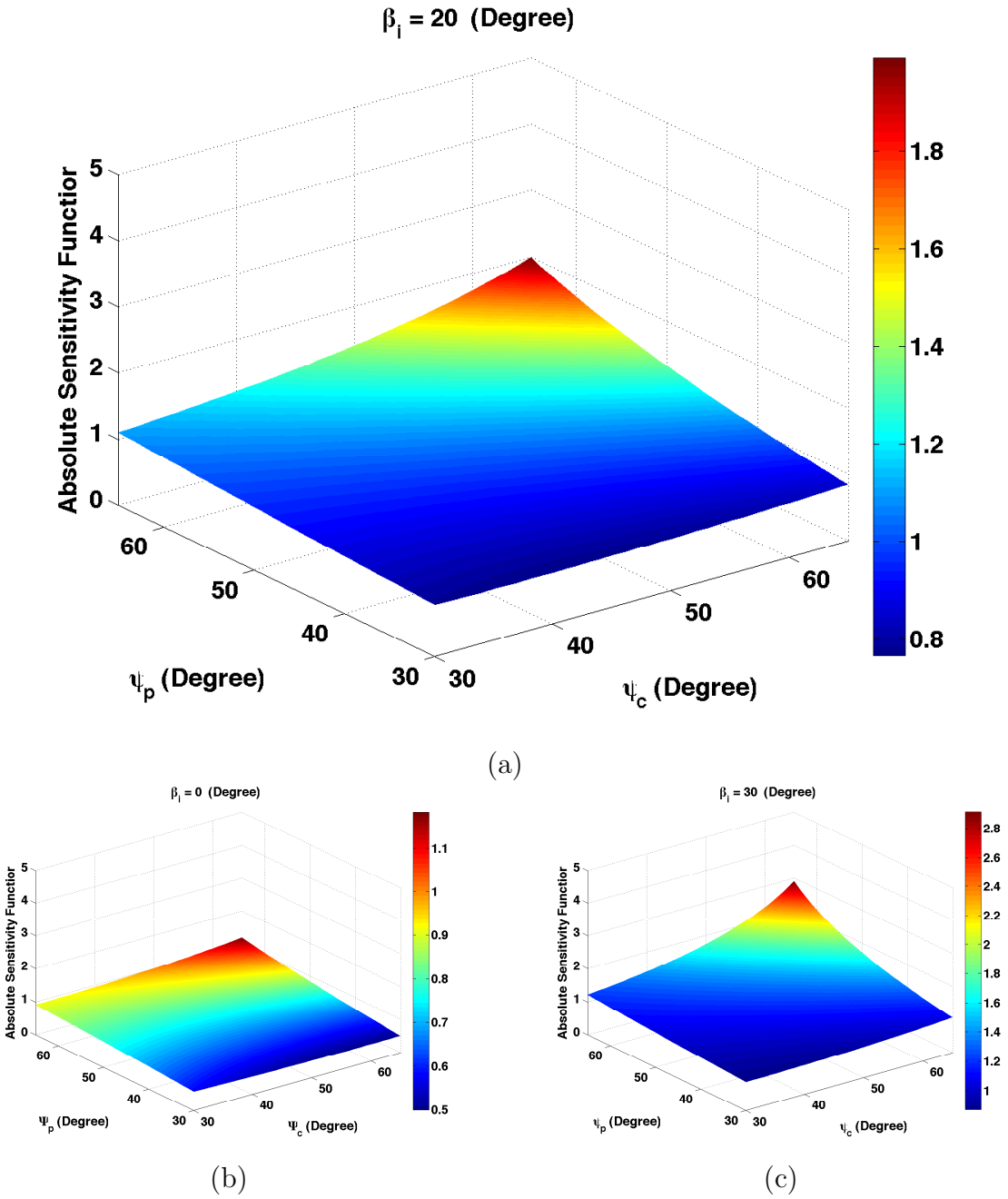
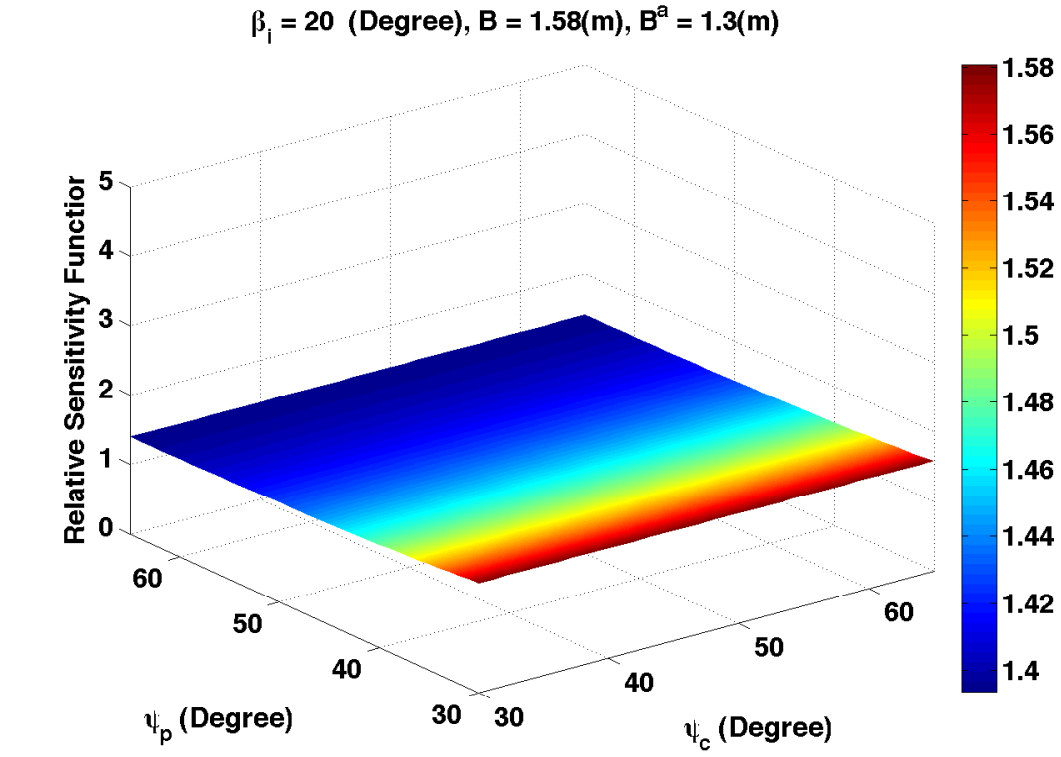
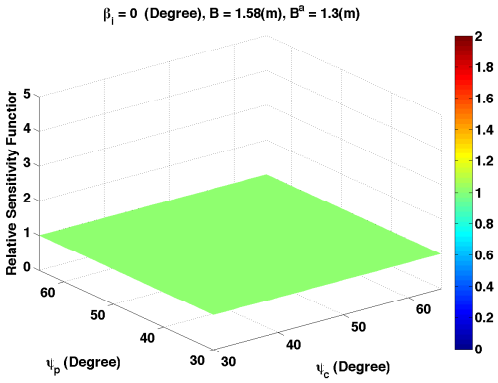


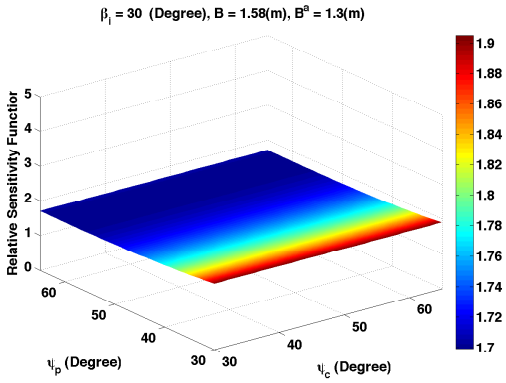
Figure 34: Absolute sensitivity function for estimated distance between the camera and the object D to the baseline between the camera and the projector B. Absolute sensitivity for (a) $\beta_i = 20$ degree, (b) $\beta_i = 0$ degree (with no refraction), (c) $\beta_i = 30$ degree (with extreme refraction). β_i is the difference between angles of incidence and refraction. The absolute sensitivity function changes for different values of ψ_c, ψ_p.



(a)



(b)



(c)

Figure 35: Relative sensitivity function for estimated distance between the camera and the object D to the baseline between the camera and the projector B. Relative sensitivity for (a) $\beta_i = 20$ degree, (b) $\beta_i = 0$ degree (with no refraction), (c) $\beta_i = 30$ degree (with extreme refraction). β_i is the difference between angles of incidence and refraction. The relative sensitivity function changes for different values of ψ_c, ψ_p.

sensitivities; $\overline{S}_{Z_i^w}^{X_i}$: the X coordinates to the height of water surface in our experiment in Figure 31(a), $\overline{S}_m^{X_i}$: the X coordinates to the slope of water surface in our experiment in Figure 33(a), and $\overline{S}_B^D$: the object-camera distance to the camera-projector baseline in Figure 35(b). The $\overline{S}_{Z_i^w}^{X_i}$ for $\beta_i = 17.25$ (degree) and $\alpha_i = 72.72$ (degree) is equal to the $\overline{S}_m^{X_i}$ for slope $m = -0.2$ and $\alpha_i = 90$ (degree). Also, the effects of the height and the slope of water surface on X coordinates with these parameter values are comparable to the effect of baseline on the camera-object distance for rays with no refraction. The value of all these relative sensitivities at their respective parameters are equal to one.

5

Conclusions and Future Work

5.1 Conclusions

We have presented a method for simultaneous reconstruction of the water surface and underwater structures using a structured-light technique. The novelty of our work is in the formulation of the reconstruction problem using an optical camera and a projector while the camera is underwater and the projector is outside water. Compare to other underwater 3D reconstruction methods which require both camera and the projector to be underwater, in our method only one of the two components, i.e. the camera, needs to be underwater. The method uses a checkerboard and a standard calibration procedure to estimate the intrinsic calibration parameters of both camera and the projector as well as extrinsic calibration parameters between the camera and the projector. The intrinsic parameters are composed of focal length, principal point, skew coefficient and lens distortions. The extrinsic parameters include rotation and translation vectors that establish the coordinate transformation from the 3D projector coordinate system to the 3D camera coordinate system. In this method, the projector, while outside of the water, projects a unique random pseudocode color pattern onto the underwater objects. Each subwindow in this pseudocode is unique and therefore can be identified without any ambiguity. Unlike conventional structured-light methods where the distortion of the projected pattern is solely due to the geometric shape of the object surface, in this method, the geometric shape of the water surface and

the object surface contribute to the distortion of the projected pattern. In order to decompose the pattern distortions by the water surface and the object surface, the method estimates the water surface shape and considers its effect on the final distortion of the pattern. When the water surface is known, the refraction of each ray can be calculated and using the path of the projected rays in the water, the underwater object will be reconstructed.

This approach can be a basis for a wide range of underwater applications in civil engineering such as offshore and nearshore mapping, monitoring underwater structures and pipelines, GIS, and coral reef monitoring. In addition, it has applications in other fields such as archaeology, inspection of boats and shipwrecks, underwater security and surveillance, and marine biology research. It is specifically useful for 3D reconstruction in hazardous environments where minimal immersion in one of the media is expedient. We hope this manuscript serves as a source for constructing a practical system for underwater surface reconstruction and depth scanning especially in shallow waters. As the system only utilizes affordable off-the-shelf components, surface reconstruction in such environments is possible by anyone who can afford a projector and a camera.

Although we can not claim our method for estimating water surface surpass the previous methods of water surface estimation (due to the lack of common ground truth data for comparison), we can say that as a byproduct, the method gives a good and rough estimate of the water surface for underwater object reconstruction purposes. We should mention that surprisingly one of the main deficien-

cies of most previous methods for measuring sea waves is the lack of a mechanism to check the accuracy of the final reconstructed water surface. In this thesis, we measured the accuracy of the reconstructed water surface by reconstructing waves in a controlled lab environment as well as reconstructing underwater objects with known shapes using estimated water surface parameters. Although, because of the space limitation in the lab, we were not able to test the larger areas that contain several wavenumbers, the simulated data shows the potential of this method in recovering the water surface.

Since our proposed technique uses an external light source for 3D reconstruction, it falls within the category of active methods for 3D reconstruction. Generally, existing image based techniques for reconstructing the 3D shape of object surfaces can be classified into two main categories: Passive and Active. Passive techniques rely only on the information embodied in the images acquired by cameras under natural uncontrolled lighting while active techniques are based on projecting patterns in the scene using a projector, and infer 3D structure from the way the pattern deform on the objects.

Generally, passive techniques are less invasive than active techniques and provide very accurate, high resolution and detailed data. However, the main drawback of passive techniques such as stereoscopy, is that they depend on the presence of texture in the scene and suffer from a lack of reliable feature points or from unreliable matching between the views. Indeed, such techniques do not work on texture-less objects. Consider for example a surface with uniform color where no

salient feature (patches, points, lines, or curves) can be selected and tracked reliably on the images. In this case, it is impossible to reconstruct the 3D structure of the surface using passive techniques without adding texture. In fact, one solution for reconstruction of such surfaces is to physically add texture to the scene by pasting landmarks on objects. However, this method is not a desirable solution in many applications including but not limited to non destructive testing and coral reef monitoring.

As an alternative, active techniques perform exceptionally well in areas where no features can be readily detected. They are alternative solutions to accommodate the absence of texture in the scene and ensure higher reliability on matches. Beyond time-of-flight solutions that are suitable over larger distances, or active laser scanners that provides very high resolution but require expensive devices, structured lighting remains an affordable but robust strategy. The artificial texture generated by the projector produces sufficient salient image features to allow dense 3D reconstruction. In addition, these techniques yield very good accuracies in the final 3D reconstructed object. As mentioned before, our method lies within the active category and therefore inherits the common limitations and advantages of active techniques.

Another advantage of our single-shot method compare to the stereoscopy, as one of the main passive techniques, is that since the projector constantly sends a unique pattern and the pattern does not change over time, there is no need to synchronize the camera and the projector. Contrary to our method, in stereoscopy

or any multi-camera setup, all cameras must be synchronized otherwise they can not be applied to dynamic scenes.

The advantage compared with multiple-shot structured light techniques, which require multiple patterns to be projected, is that our method permits moving surfaces to be measured. However, since the codification must be condensed in a unique pattern, the spatial resolution is lower. Moreover, local smoothness of the measuring surface is assumed in order to correctly decode the pixel neighborhoods. Since this local smoothness is not always accomplished, errors in the decoding stage can arise producing errors in the 3D measurements.

In our method, when projecting dots (color patches), their mass centers must be located and if a dot appears only partially in the image, it would be hard to accurately locate its mass center. In addition, the size of projected dots has direct impact on the final spatial resolution of reconstructed object; the smaller the dots, the better the resolution. However, decreasing the size of color dots can complicate the process of identification of these dots. One way of overcoming this issue is to use square dots and extract the corners of color patches instead of their mass centers as we used in our last experiment (see the yellow buoy experiment).

The depth per pixel used in the pseudocode pattern is also an important parameter. The noisier the application environment, the smaller the number of colors used should be. When increasing the number of colors, differentiating the colors becomes more difficult. Therefore, the three color combinations used in our experiments are the most robust against noise. Note that two colors could not cre-

ate a large enough unique number of subwindows in our pseudocode pattern. Also, choosing colors with higher contrast among each other and against the underwater object will increase the chance of their detection in the images. In most cases, the read colors does not usually completely match with the projected ones due to the non-linearities of the light spectrum of the projector, the spectral response of the cameras, and the non-uniform albedo of the measuring surface. In order to overcome this problem, a full colorimetric calibration procedure might need to be considered which can be done using color calibration targets. These targets feature different swatches with various shades of colors that can be used to establish the transformation between the recorded colors by the camera and the original ones.

We demonstrated the accuracy of this new method with a number of experimental results in both synthetic and real data collected in a controlled lab environment. In addition, we characterized the performance of the new method through a sensitivity analysis for different parameters of our suggested model.

Note that our method, as most structured-light techniques, only produces partial 3D views of the scene and in order to retrieve a complete 3D model of the scene, it is necessary to merge several 3D views together. If there is sufficient overlap between the partial 3D views, they can be fit together to build a single, complete shape model. This process involves solving two important problems; registration and integration of multiple partial 3D views. In each given direction, only the front part of the object is visible and the other sides are not visible or are partially visible. Either the object or the sensor has to be moved to view the other

sides of the object. The registration with respect to a common coordinate system involves finding the transformation parameters for transforming the different reference coordinate systems to the common coordinate system. After registration the partial models will need to be integrated together to obtain a single complete 3D model.

5.2 Future work

There are many further directions beyond this work. One avenue is to extend this method to the case where both camera and the projector are in the air while the object is underwater. This would be the ideal case for non destructive testing and non invasive measurement of underwater objects since neither the camera nor the projector touch the water. Similar to the presented method in this thesis, at first the water surface needs to be estimated followed by underwater object reconstruction. The same camera-projector calibration method can be utilized to estimate both intrinsic and extrinsic parameters of the camera and the projector. Also, patterns used in single shot techniques should be used to be able to handle the dynamic nature of water surface. In this case, for each ray emitted from the projector, there would be two possible recorded rays in the camera; one reflected from the water surface and one reflected from the underwater object. The latter refracts while entering and exiting water. The reflected rays from water surface can be used in standard structured light formulation to estimate the water surface. Once the water surface is estimated, using camera-projector calibration parameters, one can determine the points where the emitted ray from projector

enters into and exits from water. These two points along with the point where the ray hits the underwater object construct a triangle. By solving for this triangle, one can calculate the object coordinates.

$$D = B^w \frac{sin(\psi_p^w)}{sin(\psi_c^w + \psi_p^w)} \tag{41}$$

The main challenge in this setup is to differentiate between the reflected rays from water surface and reflected rays from underwater object. Since polarization cameras can take simultaneous images of the water surface and underwater objects, they look promising for this purpose.

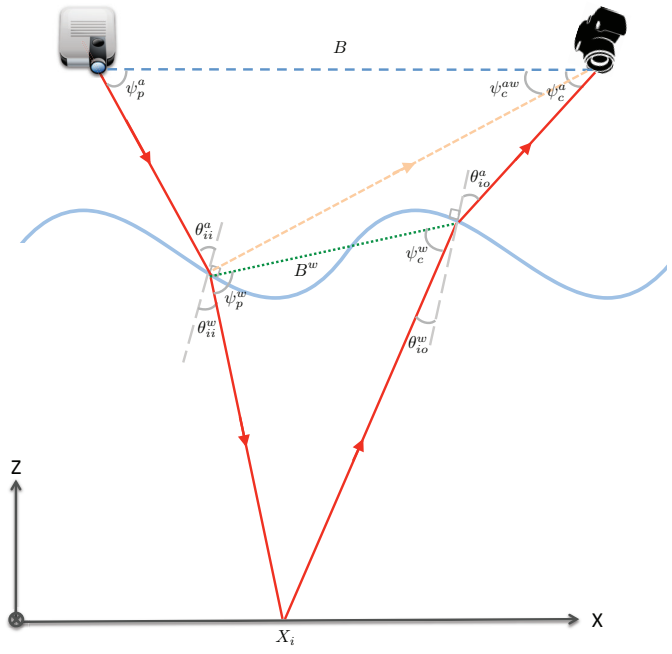


Figure 36: Extension of the proposed method to the case where the camera and the projector are outside water and the object is underwater.

Another avenue is to explore the combination of this method with underwater image enhancement techniques in order to make it more suitable for envi-

ronments with dominant scattering effects. Our method is designed to work in pure sea water and clean ocean and it implicitly assumes that the light is neither scattered nor absorbed by the water. This assumption, however, is violated in ocean environments with dominant scattering effects. An optical image taken in such a medium will be affected by significant loss of contrast and attenuation of brightness. Further, this degradation increases exponentially with distance, making it hard to capture images of the projected pattern. Thus, it is critical to take into account the effects of scattering while applying the proposed system in such media.

Notice that our single-shot technique can always be translated to a multiple-shot technique by expressing the color pattern over a sequence of patterns in order to achieve larger spatial resolution. One can change the size of the color dots in a hierarchical manner in a sequence of shots and achieve a higher spatial resolution. Alternatively, one can progressively shift the pattern horizontally and vertically to increase the number of 3D reconstructed points. As such, advantage is taken of the same uniquely coded feature points several times.

If the underwater object is static, the dynamic water surface can serve as a tool to achieve higher spatial sampling resolution. In fact, in our experiment for reconstructing the cylindrical object, the dynamic water surface helped us to sweep the pseudocode pattern across the static object surface without even changing the size of the color dots. In this way, each dot is not always projected on the same point on the object. In addition, in each shot, different points might be extracted

that accumulatively contribute to the final 3D reconstructed object.

Instead of using one camera for imaging in our setup, it is possible to use two (or more) cameras along with one (or more) projector(s) in a multi-view approach.This way, we can eliminate the need for calibration between the projector and the cameras. This solution gives access to focusing, zooming and brightness settings of the projector to adapt to various operating conditions and depth without influencing the calibration of the acquisition system. This method can also help to alleviate the deficiencies of passive techniques (such as stereo imaging) in dealing with texture-less scenes.

Our method assumes that the water surface model in X and Y directions are independent, so the model parameters in each direction can be estimated separately. This model can be extended to the a more general case where X and Y coordinates are not independent. This way, the 3D shape of the water surface will be estimated using X and Y coordinates of points at the same time.

While projector-camera systems provide high resolution images and detailed information about underwater objects, their visibility is limited to distances of no more than tens of meters even in clear water. As the turbidity of the water increases, their visibility decrease substantially which results in their ineffectiveness in many harbor waters and other highly turbid environments. Acoustic signals, however, propagate in turbid water with little degradation. Thus, sonar imaging becomes a natural choice for imaging in such environment. Unless a very expensive sonar is used, bathymetry obtained using underwater sonar has lower cloud point

density in comparison with the one obtained by an inexpensive optical camera. It is expected that fusion of structured light and sonar images can provide enhanced 3D reconstruction performance as opposed to the single utilization of each sensing modality, and would generally extend the operational range.

LIST OF REFERENCES

[1] M. W. Allgaier, R. E. Cameron, and P. O. Moore, *The Nondestructive Testing Handbook, Third edition: Volume 9, Visual Testing.* American Society for Nondestructive Testing, 2010.

[2] H. Schultz, "Shape reconstruction from multiple images of the ocean surface," *Photogrammetric Engineering and Remote Sensing*, vol. 62, pp. 93–99, 1996.

[3] F. Moffitt and E. Mikhail, *Photogrammetry*, ser. Series in Civil Engineering. Harper & Row, 1980.

[4] J. McGlone, E. Mikhail, J. Bethel, and R. Mullen, *Manual of photogrammetry.* American Society for Photogrammetry and Remote Sensing, 2004.

[5] D. C. Brown, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.

[6] S. Thorpe, "Dynamical processes of transfer at the sea surface," *Progress In Oceanography*, vol. 35, no. 4, pp. 315 – 352, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/007966119580002B

[7] W. K. Melville, "The role of surface-wave breaking in air sea interaction," *Annual Reviews of Fluid Mechanics*, vol. 28, pp. 279–321, 1996.

[8] C. Wilkinson and J. Brodie, *Catchment management and coral reef conservation*, M. Ahmed, C. K. Chong, and H. Cesar, Eds. Global Coral Reef Monitoring Network and Reef and Rainforest Research Centre, 2011, vol. 70. [Online]. Available: http://www.worldfishcenter.org/Pubs/coral_reef/coral-reef.htm

[9] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, no. 4262, pp. 283–287, 1976. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/968482

[10] B. Horn and M. Brooks, *Shape from Shading.* Cambridge, Massachusetts: MIT Press, 1989.

[11] K. Hartt and M. J. Carlotto, "A method for shape-from-shading using multiple images acquired under different viewing and lighting conditions," in *IEEE Conference on Computer Vision and Pattern Recognition (2008)*, 1989, pp. 53–60. [Online]. Available: http://ieeexplore.ieee.org:80/xpls/abs_all.jsp?isNumber=1531&prod=CNF&arnumber=37828&arSt=+53&ared=+60&arNumber=37828

[12] A. Blake and G. Brelstaff, "Geometry from specularities," in *International Conference on Computer Vision*, 1988, pp. 394–403. [Online]. Available: http://research.microsoft.com/en-us/um/people/ablake/papers/ablake/Blake_Brelstaff_ICCV88.pdf

[13] K. Ikeuchi, "Determining surface orientations of specular surfaces by using the photometric stereo method." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 3, no. 6, pp. 661–669, 1981. [Online]. Available: http://www.cvl.iis.u-tokyo.ac.jp/papers/all/446.pdf

[14] G. Healey and T. O. Binford, "Local shape from specularity," *Computer Vision Graphics and Image Processing*, vol. 42, no. 1, pp. 62–86, 1988. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0734189X88901430

[15] B. Horn, *Robot Vision*. Cambridge, Massachusetts: MIT Press, 1986.

[16] J.-Y. Bouguet, "Visual methods for three-dimensional modeling," Ph.D. dissertation, California Institute of Technology, Pasadena, CA, USA, 1999, aAI9941097.

[17] J. S. Jaffe, J. McLean, M. P. Strand, and K. D. Moore, "Underwater optical imaging: Status and prospects," Scripps Institution of Oceanography, La Jolla, Technical Report, 2002.

[18] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic aperture confocal imaging," *ACM Transactions on Graphics (TOG)*, vol. 23, p. 825834, 2004.

[19] P. Naulleau and D. Dilworth, "Motion-resolved imaging of moving objects embedded within scattering media by the use of time-gated speckle analysis," *Applied Optics*, vol. 35, pp. 5251–5257, 1996.

[20] G. R. Fournier, D. Bonnier, L. J. Forand, and P. W. Pace, "Range-gated underwater laser imaging system," *Opt. Eng*, vol. 32, p. 21852190, 1993.

[21] S. Narasimhan, S. Nayar, B. Sun, and S. Koppal, "Structured light in scattering media," in *IEEE International Conference on Computer Vision (ICCV)*, vol. 1, October 2005, pp. 420–427.

[22] C. D. Mobley, *Light and water: radiative transfer in natural waters*. Academic Press, 1994.

[23] M. Longuet-Higgins, D. Cartwright, and N. Smith, "Observations of the directional spectrum of sea waves using the motions of a floating buoy," in *Ocean Wave Spectra*. Englewood Cliffs: Prentice-Hall, Inc., 1963, pp. 111–136.

[24] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408 – 1418, aug. 1969.

[25] M. Isobe, K. Kondo, and K. Horikawa, "Extension of mlm for estimating directional wave spectrum," in *Symposium on Description and Modeling of Directional Seas*, Copenhagen, 1984, pp. 1–15.

[26] A. Lygre and H. E. Krogstad, "Maximum entropy estimation of the directional distribution in ocean wave spectra," *Journal of Physical Oceanography*, vol. 16, no. 12, p. 20522060, 1986. [Online]. Available: http://ams.allenpress.com/amsonline/?request=get-abstract&issn=1520-0485&volume=016&issue=12&page=2052

[27] M. A. Donelan, W. M. Drennan, and A. K. Magnusson, "Nonstationary analysis of the directional properties of propagating waves," *Journal of Oceanography*, vol. 26, no. 9, pp. 1901–1914, 1996. [Online]. Available: http://ams.allenpress.com/amsonline/?request=get-abstract&issn=1520-0485&volume=026&issue=09&page=1901

[28] C. Cox and W. Munk, "Statistics of the sea surface derived from sun glitter," *Journal of Marine Research*, vol. 13, pp. 198–227, 1954.

[29] C. S. Cox, "Measurements of slopes of high-frequency wind waves," *Journal of Marine Research*, vol. 16, no. 3, pp. 199–225, 1958.

[30] B. Hughes, H. Grant, and R. Chappell, "A fast response surface-wave slope meter and measured wind-wave moments," *Deep Sea Research*, vol. 24, no. 12, pp. 1211–1223, 1977. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0146629177905240

[31] W. C. Keller and B. L. Gotwols, "Two-dimensional optical measurement of wave slope," *Applied Optics*, vol. 22, no. 22, pp. 3476–3478, Nov 1983.

[32] B. Jähne and K. S. Riemer, "Two-dimensional wave number spectra of small-scale water surface waves," *Journal of Geophysical Research*, vol. 95(C7), pp. 11 531–11 546, 1990.

[33] O. H. Shemdin, H. M. Tran, and S. C. Wu, "Directional measurement of short ocean waves with stereophotography," *journal of Geophysical Research*, vol. 93, no. C11, pp. 13 891–13 901, 1988.

[34] S. Waas and B. Jaehne, "Combined slope-height measurements of short wind waves: first results from field and laboratory measurements," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, L. Estep, Ed., vol. 1749, dec 1992, pp. 295–306.

[35] E. J. Bock and T. Hara, "Optical measurements of ripples using a scanning-laser slope gauge: Part ii - data analysis and interpretation from a laboratory wave tank," in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 1749, 1992, pp. 272–282.

[36] R. J. Martinsen and E. J. Bock, "Optical measurements of ripples using a scanning-laser slope gauge: Part i–instrumentation and preliminary results," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, L. Estep, Ed., vol. 1749, dec 1992, pp. 258–271.

[37] P. H. Lee, J. D. Barter, K. L. Beach, C. L. Hindman, B. M. Lake, H. Rungaldier, J. C. Schatzman, J. C. Shelton, R. N. Wagner, A. B. Williams, and et al., "Recent advances in ocean-surface characterization by a scanning-laser slope gauge," in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 1749, 1992, pp. 234–244.

[38] R. J. Woodham, "Photometric stereo: A reflectance map technique for determining surface orientation from image intensity," in *Proceeding of SPIE*, vol. 155, 1978, pp. 136–143.

[39] E. Coleman and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Computer Graphics and Image Processing*, vol. 18, no. 4, pp. 309–328, 1982. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0146664X82900016

[40] H. D. Tagare and R. J. P. Defigueiredo, "A theory of photometric stereo for a class of diffuse non-lambertian surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 2, pp. 133–152, 1991. [Online]. Available: http://www.cs.virginia.edu/~mjh7v/bib/Tagare91.pdf

[41] A. Blake, "Specular stereo," in *Proceedings of the 9th international joint conference on Artificial intelligence - Volume 2*, ser. IJCAI'85. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1985, pp. 973–976. [Online]. Available: http://dl.acm.org/citation.cfm?id=1623611.1623671

[42] A. C. Sanderson, L. E. Weiss, and S. K. Nayar, "Structured highlight inspection of specular surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 1, pp. 44–55, 1988. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=3866

[43] A. Zisserman, P. Giblin, and A. Blake, "The information available to a moving observer from specularities," *Image and Vision Computing*, vol. 7, no. 1, pp. 38–42, 1989. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0262885689900188

[44] H. Murase, "Surface shape reconstruction of a nonrigid transparent object using refraction and motion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 1045–1052, 1992.

[45] J. Thomas, W. Kober, and F. Leberl, "Multiple image sar shape-from-shading," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 1, pp. 51–59, 1991.

[46] B. K. Haus, M. A. Donelan, and W. J. Plant, "Optical measurements of the modulation of short wind waves by long waves and their effect on observed c-band radar reflectivity," *Geophysical Research Abstracts, European Geophysical Society*, vol. 5, p. 11426, 2003.

[47] I. Ishii, K. Yamamoto, K. Doi, and T. Tsuji, "High-speed 3d image acquisition using coded structured light projection," in *IEEE RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 925–930.

[48] K. Sato and S. Inokuchi, "Range-imaging system utilizing nematic liquid crystal mask," in *IEEE International Conference on Computer Vision (ICCV)*, 1987, pp. 657–661.

[49] R. Valkenburg and A. McIvor, "Accurate 3d measurement using a structured light system," *Image and Vision Computing*, vol. 16, pp. 99–110, 1996.

[50] P. S. Huang and S. Zhang, "Fast three-step phase-shifting algorithm," *Appl. Opt.*, vol. 45, no. 21, pp. 5086–5091, Jul 2006.

[51] Z. J. Geng, "Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems," *Optical Engineering*, vol. 35, pp. 376–383, February 1996.

[52] J. Geng, "3d surface profile imaging method and apparatus using single spectral light condition," US Patent 6 556 706, April 29, 2003.

[53] P. Payeur and D. Desjardins, "Structured light stereoscopic imaging with dynamic pseudo-random patterns," *Image Analysis and Recognition, Lecture Notes in Computer Science*, vol. 5627, pp. 687–696, 2009.

[54] S. Fernandez, J. Salvi, and T. Pribanic, "Absolute phase mapping for one-shot dense pattern projection," in *PROCAMS10*, 2010, pp. 64–71.

[55] M. Maruyama and S. Abe, "Range sensing by projecting multiple slits with random cuts," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 647–651, June 1993.

[56] C. Heike, K. Upson, E. Stuhaug, and S. Weinberg, "3d digital stereophotogrammetry: a practical guide to facial image acquisition." *Head Face Med*, vol. 6, 2010.

[57] J. Geng, "Method and apparatus for 3d imaging using light pattern having multiple sub-patterns," US Patent 6 700 669, March 02, 2004.

[58] K. L. Boyer and A. C. Kak, "Color-encoded structured light for rapid active ranging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 14–28, January 1987. [Online]. Available: http://dl.acm.org/citation.cfm?id=28748.28750

[59] H. Morita, K. Yajima, and S. Sakata, "Reconstruction of surfaces of 3-d objects by m-array pattern projection method," in *Second International Conference on Computer Vision*, 1988, pp. 468–473.

[60] A. Ulusoy, F. Calakli, and G. Taubin, "One-shot scanning using de bruijn spaced grids," in *3DIM09*, 2009, pp. 1786–1792.

[61] M. Rioux, G. Bechthold, D. Taylor, and M. Duggan, "Design of a large depth of view three-dimensional camera for robot vision," *Optical Engineering*, vol. 26, Dec 1987.

[62] T. Kanade, A. Gruss, and L. Carley, "A very fast vlsi rangefinder," in *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*, apr 1991, pp. 1322 –1329 vol.2.

[63] B. Curless and M. Levoy, "Better optical triangulation through spacetime analysis," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*, jun 1995, pp. 987–994.

[64] J. A. Beraldin, M. Rioux, F. Blais, L. Cournoyer, and J. R. Domey, "Registered intensity and range imaging at 10 mega-samples per second," *Optical Engineering*, vol. 31, 1992.

[65] K. Sato and S. Inokuchi, "Three-dimensional surface measurement by space encoding range image," *Journal of Robotic Systems*, vol. 2, p. 2739, 1985.

[66] J. L. Posdamer and M. D. Altschuler, "Surface measurement by space-encoded projected beam systems," *Computer Graphics and Image Processing*, vol. 18, no. 1, pp. 1–17, 1982.

[67] S. Inokuchi and Matsuda, "Range-imaging for 3-d object recognition," in *International Conference on Pattern Recognition (International Association for Pattern Recognition)*, 1984, pp. 806–808.

[68] D. Caspi, N. Kiryati, and J. Shamir, "Range imaging with adaptive color structured light," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 5, pp. 470 –480, may 1998.

[69] K. Hattori and Y. Sato, "Accurate rangefinder with laser pattern shifting," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, vol. 3, aug 1996, pp. 849 –853 vol.3.

[70] H. Sagan, *Space Filling Curves*. Springer,, 1994.

[71] O. Hall-Holt and S. Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects," in *Eighth International Conference on Computer Vision (ICCV)*, jul 2001.

[72] N. G. Durdle, J. Thayyoor, and V. J. Raso, "An improved structured light technique for surface reconstruction of the human trunk," in *EEE Canadian Conference on Electrical and Computer Engineering*, vol. 2, 1998, pp. 874–877.

[73] L. Zhang, B. Curless, and S. M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," in *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, June 2002, pp. 24–36.

[74] F. J. MacWilliams and N. J. A. Sloane, "Pseudo-random sequences and arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1729, 1976. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1454680

[75] M. Proesmans, L. van Gool, and A. Oosterlinck, "Active acquisition of 3d shape for moving objects," in *Image Processing, 1996. Proceedings., International Conference on*, vol. 3, sep 1996, pp. 647 –650 vol.3.

[76] M. Proesmans, L. Van Gool, and A. Oosterlinck, "One-shot active 3d shape acquisition," in *Proceedings of the 13th International Conference on Pattern Recognition*, vol. 3, aug 1996, pp. 336 –340 vol.3.

[77] K. L. Boyer and A. C. Kak, "Color-encoded structured light for rapid active ranging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 14 –28, January 1987.

[78] C. Davies and M. Nixon, "A hough transform for detecting the location and orientation of three-dimensional surfaces via color encoded spots," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 28, no. 1, pp. 90 –95, feb 1998.

[79] P. Lavoie, D. Ionescu, and E. M. Petriu, "3d object model recovery from 2d images using structured light," *Ieee Transactions On Instrumentation And Measurement*, vol. 53, no. 2, pp. 437–443, 2004. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1284876

[80] C. Chen, "Range data acquisition using color structured lighting and stereo vision," *Image and Vision Computing*, vol. 15, no. 6, pp. 445–456, 1997.

[81] J. Geng, "Structured-light 3d surface imaging: a tutorial," *Advances in Optics and Photonics*, vol. 3, no. 2, pp. 128–160, Jun 2011.

[82] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323 –344, august 1987.

[83] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 1, 1999, pp. 666 –673 vol.1.

[84] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, jun 1997, pp. 1106 –1112.

[85] J.-Y. Bouguet, "Camera calibration toolbox for matlab." [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/

[86] D. Desjardins and P. Payeur, "Dense stereo range sensing with marching pseudo-random patterns," in *Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 216–226.

[87] E. D. Smith, F. Szidarovszky, W. J. Karnavas, and A. T. Bahill, "Sensitivity analysis , a powerful system validation technique," *The Open Cybernetics and Systemics Journal*, vol. 2, no. 1, pp. 39–56, 2008. [Online]. Available: http://www.bentham-open.org/pages/content.php?TOCSJ/2008/00000002/00000001/39TOCSJ.SGM