An Improved Stereo Camera Pattern Projector for Use in Robotics

A Design Project Report Presented to the School of Electrical and Computer Engineering of Cornell University in Partial Fulfillment of the Requirements for the Degree of Master of Engineering, Electrical and Computer Engineering

Project Completed in Partnership with Carnegie Robotics, LLC.

By:
William Laney
wel53

Cornell ECE Advisor:
Dr. Peter Jessel

Signature: ______________________________

Date: ______________________________

Carnegie Robotics Advisor:
Dr. Omead Amidi

Signature: ______________________________

Date: ______________________________

Contents

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my Carnegie Robotics advisor Dr. Omead Amidi and my Cornell advisor Dr. Peter Jessel for their continuous support of my research, and for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all aspects of research and writing of this project.

I would like to thank everyone at Carnegie Robotics for their support in this project. Without the team their none of this would have been possible.

Finally I would like to thank my family: my parents Judy and Hyrum, my brother Joshua, and my grandfather Peter for supporting me in every aspect of my life.

Abstract

In stereo vision, pattern projectors are used to aid in image matching and to increase the quality of the produced depth data. In this study we pursued multiple avenues for creating improved pattern projectors for steros vision. We attempted two approaches which used light emitting diodes (LEDs). The first approach utilized LED gradual illumination and was done in simulation. The second approach experimentally tested patterns produced by LED GOBO projectors. Separate from LED, technology we prototyped a laser based pattern projector utilizing diffractive optical elements. Finally, we evaluated two commercially available laser based pattern projectors: one making use of diffractive optical elements, and the other vertical-cavity surface-emitting laser (VCSEL) diodes. All work done in this study was conducted with the goal of improving the performance of the Carnegie Robotics MultiSense stereo camera.

Chapter 1

Introduction

Stereo vision systems can be broken into two broad categories based on their illumination strategies. In the first category are passive systems that rely solely on the natural light in the environment. The second category uses active systems, which have a light source for the camera. Active systems can use one of three main approaches. The first is to simply bathe the subject in light. While this allows the camera system to work on scenes without sufficient natural light, it does not inhearently improve the performance of the camera. The next is to project a distinctive pattern of light onto the scene. This solves the low light problem while also improving the performance of the camera system. The reason for this is that the projected pattern increases the texture in the scene artificially which in turn allows for better image matching by the camera [1]. This artificial texture is necessary to get accurate depth data from low texture scenes, e.g. smooth walls. The final active illumination strategy projects specific light patterns onto the scene, and the stereo algorithms exploit knowledge of the exact pattern being projected to estimage depth[2].[1]

When selecting a pattern to project, it is important to consider the texture it will

[1] The phrase "structured light" is commonly used to distinguish between types of active illumination. There are however competing definitions as to if structured light means any sort of system that structures light into a pattern, or only systems in which the algorithm has pre-knowledge of the pattern. To minimize this confusion the authors will avoid the phrase structured light entirely.

introduce into the scene, and thus how it will improve the stereo matching results. The best patterns are pseudo-random dots that cover the entire field of view of the camera [3]. They create high contrast and distinctive patterns for the camera to match. Research has been done into the ideal random dot patterns to make matching as easy as possible while minimizing the chance of mismatches across the image[4]. Aside from optimizing dot patterns, little has been published about the practical issues associated with the design, deployment, and improvement of these systems. Furthermore, little to no research seems to be underway into alternative patterns or techniques that may be used.

This research focuses on improving the second type of active illumination, where a pattern is projected but the camera algorithms have no knowledge it exists. The project is primarily concerned with robotics applications, where the camera system is mobile and may be in operation around people. Special attention has been paid to pattern performance with the Carnegie Robotics MultiSense stereo camera.

The structure of this report is as follows: Section 2 lays out the specific issues we aim to address along with additional background. Section 3 introduces the idea of a gradual illumination LED pattern projector. It then goes on to test the concept through simulation, and finds it to not be a viable strategy. Section 4 describes experiments with a number of different pattern projector technologies. These technologies include both light emitting diode (LED) and laser-based approaches. Finally, Section 5 presents our overall conclusions.

Chapter 2

Issues to be Addressed

Random dot pattern projectors are traditionally made by placing a diffractive optical element in front of a laser source[5] [6]. While this is simple in principle, it introduces many issues for real-world use. Laser diodes can be delicate devices and have highly specific power and temperature constraints. Additionally, all laser systems introduce eye-safety concerns and are governed under IEC 60825 specifications[7].

Due to this and other technical and market reasons, there are few random dot pattern projectors available commercially. The majority are designed for use in industrial inspection and are thus not suitable for more general purpose applications. While general purpose pattern projectors are optimized for size and cost, there is room for tradeoffs in exchange for better performance.

The most straightforward way to improve the performance of a pattern projector is to increase the optical output power. Increased power leads to greater range for the projector and better pattern visibility in scenes with high ambient light. The key tradeoffs for increased power with laser based systems are increased heat, greater complexity in electrical design, and eye safety concerns.

LEDs do not suffer from the same eye safety concerns at higher power as do laser based systems. What makes lasers dangerous to the eye is that they are highly collimated and thus focus all of their power into a very small area. Furthermore, most lasers used in pattern projectors operate in invisble infrared wavelengths, rendering natural reflexes to protect the eye ineffective. LEDs on the other hand are not collimated and thus disperse quickly, distributing their power over a wide area.

There has been research toward making random dot pattern projectors with LEDs, but the results have been poor at distances greater than 3m [1] [8]. This is because the light from the LED diverges quickly, making the dots indistinguishable from each other over anything more than short distances. Furthermore, the dots are normally created using a metal gobo (go before optics) which works by blocking light, and is thus inefficient.

We built on existing LED pattern projection research in two separate ways. We first attempted a new approach to the problem where natural illumination patterns from LEDs are used. We then expanded on the existing GOBO research by using more optically efficient glass GOBOs to create patterns.

In addition to the LED research, we built a laser based diffractive optical element pattern projector. We also evaluated two commercially available pattern projectors with the MultiSense camera. This evaluation led us to research the state-of-the-art of vertical-cavity surface-emitting laser diode array pattern projectors.

Chapter 3

Gradual Illumination LED Pattern Projection

Gradual illumination LED pattern projection is based on the idea that the divergence of light out of an LED could itself be used as a pattern for stereo matching. The LED light has the greatest intensity in-line with the LED and the intensity decreases radially from the center point. This creates a gradual light gradient pattern that could conceivably be used by the stereo matching algorithm. Creating a pattern projector of this type would realize the benefits of using LEDs over lasers, including improved eye safety, higher fault tolerance, lower prices, and wider market availability. Furthermore, LEDs do not suffer from the phenomenon of laser speckle, which causes problems in stereo matching [9] [8]. Despite these potential benefits, we are not aware of any prior research that uses the unmodified output of LEDs, or any sort of gradual illumination strategy, for stereo matching.

The following sections outline our methodology and results in examining a gradual illumination LED pattern projector. We created simulator tools to produce LED light patterns and test them with stereo matching algorithms. We ultimately determined that the LED gradual illumination strategy is not feasible with the Carnegie Robotics MultiSense.

3.1 Simulator Design

There are three steps involved in evaluating possible patterns through simulation. The first is generating the possible patterns. The second is converting these patterns into two images that are equivalent to the output of a stereo camera. The final step is performing image matching on the two stereo images in order to obtain results.

To generate possible patterns, we created a simulator capable of outputting the illumination pattern of LEDs with arbitrary arrangement, field of view, and distance from the projector. We additionally created a simulator that replicated the behavior of a random dot projector. We then implemented a technique to create a pair of stereo images from the simulation. Finally, we implemented a semi-global matching algorithm to produce disparity results. In addition, we received a pre-built simulator of the Carnegie Robotics MultiSense camera that also produced disparity results. These simulators are presented in detail below.

3.1.1 LED Simulation

The behavior of the LED simulator was derived from a combination of information taken from LED datasheets and observations of LED light. LED datasheets generally give the viewing angle of the LED, and a chart showing intensity of the LED versus angle. The viewing angle is defined as the total angle around which the relative intensity of an LED is 50% or greater. After viewing charts of LED intensity versus angle we determined that the fall-off of intensity from the central bright point of most LEDs is roughly exponential in nature. We additionally determined that the light pattern of an LED directed at a vertical surface is circular when the LED is normal to the surface, and becomes elliptical as the LED is angled. This phenomenon is intuitive and easy to observe by simply pointing a flashlight at a wall and moving one's wrist.

Using this information we created a simulator that produced an expected LED light pattern on a vertical surface. The results of the simulator are dependent on the following variables:

- Distance between the LED and the surface (z)
- Coordinates of the LED with respect to the center point of the surface (x).
- Angles of the LED off the plane of the surface (θ)
- The field of view of the LED (ϕ).

Note that x and θ are two dimensional.

To find the intensity of light emitted by the LED at a certain point, P, the simulator first calculates the location of the full intensity point of the LED (3.1). It additionally calculates four more points that represent the vertices of an ellipse where the LED light is at 50% intensity (3.2).

$$P_{Full} = z * \tan(\theta) + x \tag{3.1}$$

$$P_{vertices} = z * \tan(\theta \pm \phi/2) \tag{3.2}$$

The simulator then defines a line between the point P and the central bright point and solves for where this line intersects the 50% intensity ellipse. The intersection point is defined as P_{50}. Using the intersect point and the full intensity point the simulator defines a decay parameter (δ) that describes the exponential decay in the light intensity along this line. (3.3).

$$\delta = \frac{\ln(0.5)}{|P_{50} - P_{Full}|} \tag{3.3}$$

Figure 3.1: Example image output of the LED pattern projector simulator. This image shows the output of an LED 5m from the surface with a 60° field of view. The LED was placed at the origin and not angled.

The simulator uses the decay parameter and the euclidean distance between point P and the full intensity point to determine the light intensity, I, at the point P. (3.4).

$$I = \exp(\delta * |P - P_{Full}|) \tag{3.4}$$

This process is repeated for every point on the surface, and leads to an image of the type in Figure 3.1.

To simulate multiple LEDs, the described process is repeated for each LED and the light intensities from each LED are added together. The intensity values are normalized to be between 0 and 255. This means that no clipping takes place, and the resulting images have perfect exposure. An example of multiple LEDs in a single image can be seen in Figure 3.2

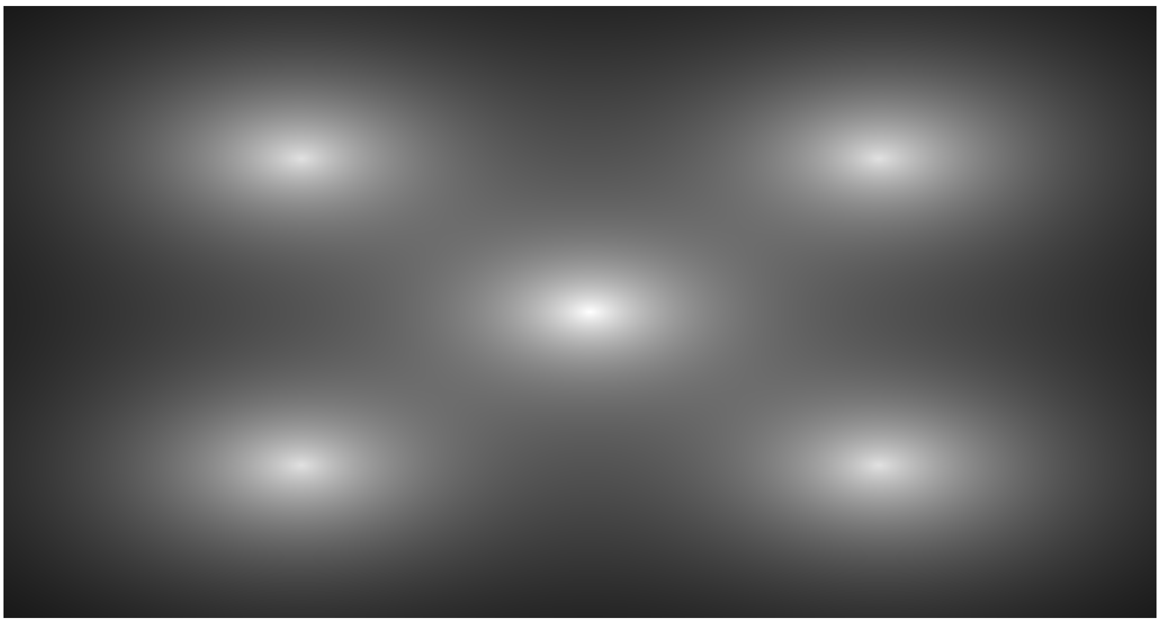

Figure 3.2: Output of the LED pattern projector simulator with five LEDs 5m from the surface. All five of the LEDs have a field of view of 20°. The center LED is placed at the origin and not angled, the four corner LEDs are shifted 0.1m and angled 30° in their respective directions.

3.1.2 Generation of Stereo Pair Images

Stereo matching is based around having two images that have a slight offset from each other. These images constitute the stereo pair. The algorithms used in this paper require the stereo pair to be rectified, meaning that pixels corresponding to the same location between the two images must only be offset in the horizontal direction. The magnitude of this offset is referred to as the disparity.

To generate the stereo pair of images the simulator takes the following inputs:

- Distance between the centers of the two camera sensors (baseline distance). This offset is only in the horizontal direction, the sensor have no vertical offset
- Field of view of the sensors
- Sensor resolution in pixels
- Distance between the camera and the surface

Using the field of view of each camera sensor and the distance to the surface, the simulator calculates the total area that can be seen by either camera sensor. The LED simulator then produces a light pattern over this area. The light pattern image is split into a left and right image based on what can be seen by the individual camera sensor. This creates the stereo pair of images necessary for image matching.

3.1.3 Semi-global Matching

To evaluate the stereo pair images we implemented a semi-global image matching algorithm based on Heiko Hirschmuller's 2005 landmark paper on the topic, and his 2008 follow up [10] [11]. At a high level, the algorithm works by calculating a cost at each pixel for each possible disparity. The costs are then aggregated along lines radiating out of each pixel to the edges of the image. The disparity for each pixel is

Figure 3.3: Example output from the semi-global matching algorthim.[1]
Left: Left stereo pair image
Center: Right stereo pair image
Right: Resulting disparity image. Lighter pixels represent points that are closer to the camera.

selected as the disparity value that gives the lowest aggregate cost at that pixel. The cost aggregation is weighted to produce smooth results.

In our implementation of the algorithm, the initial pixel costs were calculated using a sum of absolute differences with a five by five kernel. The images were padded with black to allow calculations at the edges. Sixty-four possible disparities were evaluated.

The cost aggregation was done along eight paths from each pixel. This is the minimum number of paths that Hirschmuller says is effective. The path directions were in the horizontal, vertical, and diagonal directions from the pixel. An example of the results produced by the algorithm can be seen in Figure 3.3.

3.1.4 Carnegie Robotics MultiSense Simulator

In addition to the semi-global matching algorithm we wrote, Carnegie Robotics supplied a simulator of their MultiSense camera algorithm. This simulator takes a stereo image pair as an input and outputs a disparity image that is pixel-identical to

[1]Stereo image pair taken from the data set created by Daniel Scharstein, Alexander Vandenberg-Rodes, and Rick Szeliski in conjunction with their 2003 paper [12].

Figure 3.4: Example output from the semi-global matching step of the MultiSense simulator.
Left: Left stereo pair image
Center: Right stereo pair image
Right: Resulting disparity image. Lighter pixels represent points that are closer to the camera.

that produced by the MultiSense camera. The MultiSense camera makes use of a sophisticated semi-global matching algorithm to produce high quality real-time results. This algorithm is accompanied by both pre- and post-processing steps for the images. The semi-global matching algorithm used by the MultiSense has a significantly different implementation than that originally put forward by Hirschmuller. This is because that original algorithm is unable to run in real time on large images [13].

The MultiSense simulator additionally produces intermediate results at different points throughout the algorithms pipeline. For the purposes of this study we used the disparities produced at the semi-global matching step, before any post processing is applied. An example of the results produced by the MultiSense simulator can be seen in Figure 3.4.

The MultiSense simulator outputs 16bit grayscale images. These images can often be difficult to view, with the images appearing very dark. For the purposes of analysis in this report the MultiSense disparity images have their intensity range adjusted so that the images are brighter and detail is easier to visually discern. The settings used are the same for all of the MultiSense test images for easy comparison.

3.2 Simulator Results

We tested various simulated LED patterns with both the semi-global matching algorithm and the MultiSense simulator. We began by using simple patterns with the semi-global matching algorithm that produced perfect disparity results. We then attempted to use the same patterns with the MultiSense simulator. These patterns proved ineffective with the MultiSense simulator. We took these results and attempted to use them to create more complex LED patterns that would work with the MultiSense Simulator. While we were able to improve the MultiSense simulator results, we concluded that gradual illumination with LEDs is not effective with the MultiSense algorithms.

3.3 Analysis Technique

In order to create a baseline for evaluation we built a random dot simulator. It takes in the desired number of dots and places them with maximum intensity on the surface. The dots are placed using a uniform random distribution. While a truly random dot distribution is not mathematically ideal for stereo matching [4] it did yield perfect disparity results in this study. An example of a random dot distribution can be seen in Figure 3.5.

The disparity in 3.5 is the result of perfect matching. The entire image, excluding the left edge, is a uniform shade of gray. This represents every pixel having the same disparity value, and thus being the same distance from the camera. The LED simulator projects light onto a surface that is a uniform distance from the camera so uniform disparity is expected. The left edge is black because matching is not possible on this edge due to the offset between the left and right images.

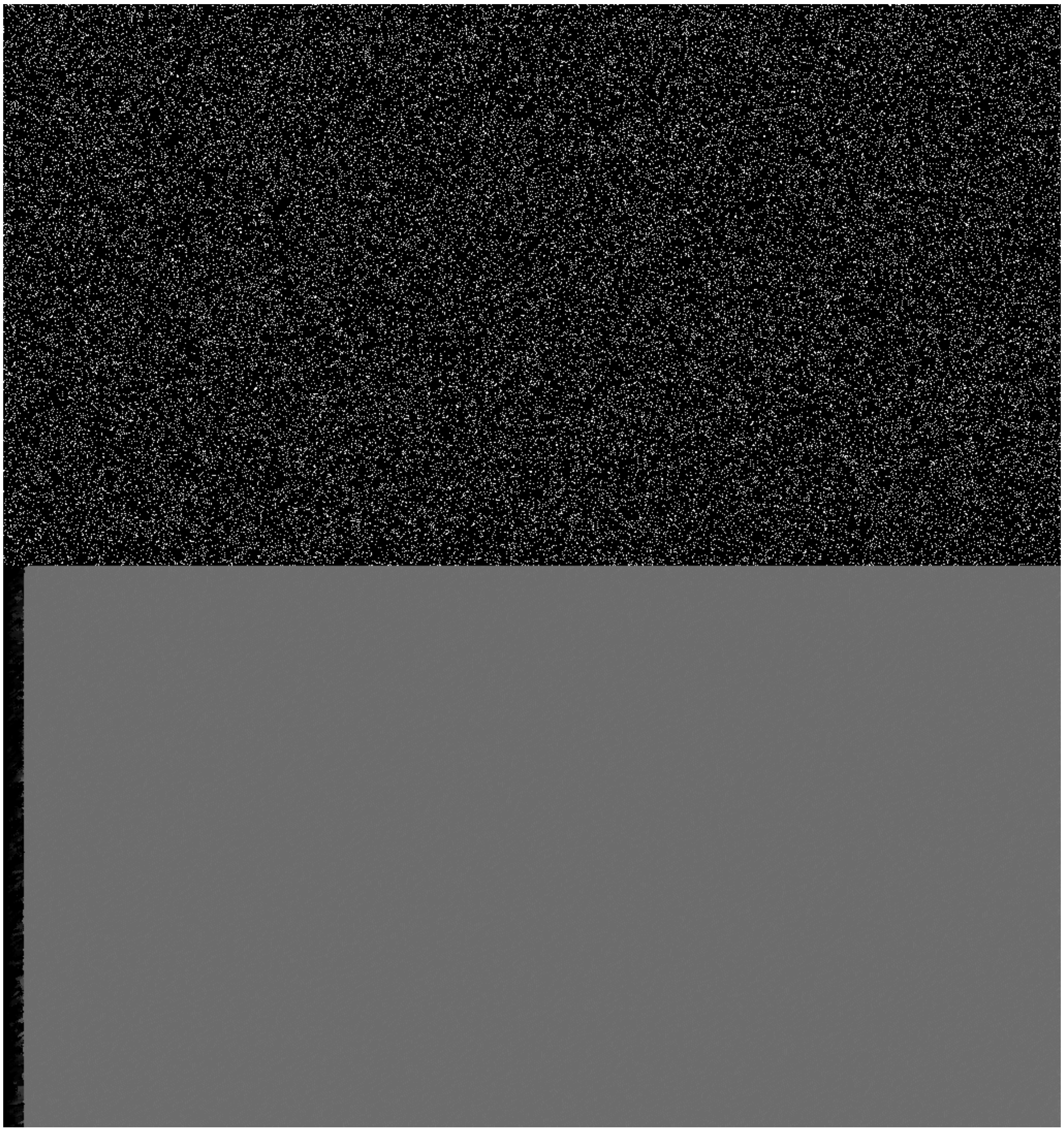

Figure 3.5: Simulated 60,000 dots image (top) and resulting MultiSense simulator disparity (bottom). This represents a perfect disparity image, and what this research is trying to achieve.

For this report analysis of disparity images produced from simulated LED patterns was done through qualitative visual comparison to the disparity in Figure 3.5. Since the LED pattern simulation produces such idealized versions of the images that would be obtained with physical LEDs and a MultiSense an LED pattern can only be declared successful if it produces perfect results. With all the simulated patterns tested the perfection of a resulting disparity was never ambiguous, so there was no need for more in depth quantitative analysis of the images.

3.3.1 Semi-global Matching

The researchers semi-global matching algorithm produces perfect disparity results from the simulated LED images. This comes mathematically from the implementation of the semi-global matching algorithm and the LED simulator. The semi-global matching algorithm calculates initial costs through a sum of absolute differences. Since the LED simulator creates noise free images the absolute difference at the correct disparity is always zero. Additionally the LED simulator represents an image of a uniform vertical wall. This means that the correct results have the same disparity value at all pixels in output, since all points are the same distance from the camera. The semi-global matching aggregates costs and has weights that give lower aggregate costs when adjacent pixels have the same disparity. If a pixel has multiple disparities that gives it an initial cost of zero; the final disparity for that pixel is the one that keeps it at the closest disparity value to its neighbors. These two factors together lead to the algorithm producing correct results for any LED pattern that does not produce perfectly consistent illumination across the image. An example of a correct semi-global matching result can be seen in Figure 3.6.

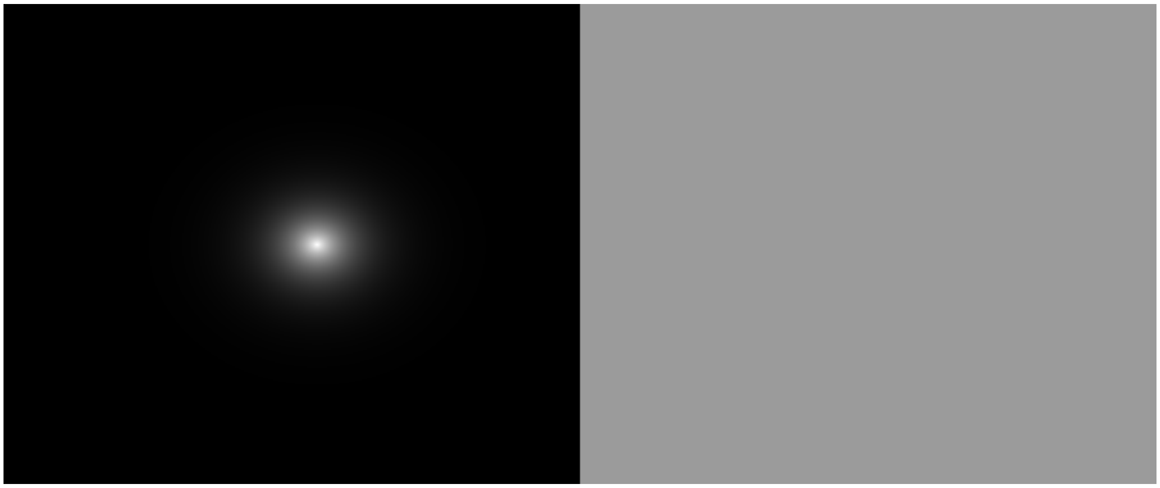

Figure 3.6: Left stereo image of a single LED 1m from the surface with a 10° field of view (left) and resulting SGM disparity (right)

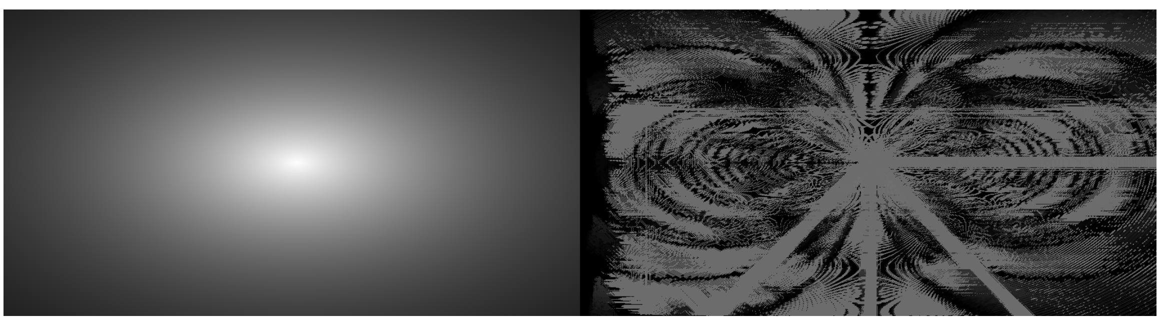

Figure 3.7: Left stereo image of a single LED 5m from the surface with a 60° field of view (left) and resulting MultiSense disparity (right)

3.3.2 Initial Characterization of Gradual Illumination with the MultiSense Simulator

The semi-global matching algorithm we implemented on the MultiSense requires more texture than the one implemented. The simple LED patterns that produced perfect results with the our non-real time SGM algorithm do not produce similar results with the MultiSense real-time SGM algorithm. Example images can be seen in Figures 3.7 and 3.8.

We tested eight patterns created with fewer than 10 LEDs arranged in geometri-

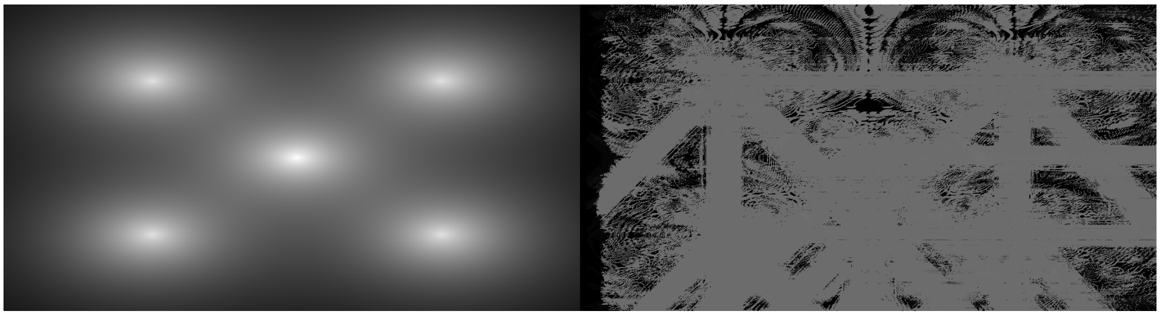

Figure 3.8: Left stereo image of five LEDs 5m from the surface, all with a field of view of 20° (left) and resulting MultiSense disparity (right)

cally uniform patterns. None of the tested patterns led to acceptable disparity results. From this we concluded that more LEDs arranged in more complicated patterns would be necessary to produce acceptable results.

Through experimentation, we found that the MultiSense does not produce symmetric disparities with symmetric input images. We also found that the results were improved with a greater number of LEDs. We further found that results degraded with distance. In general increasing distance has a square relationship with disparity degradation with the simple patterns.

3.3.3 Development of Complex Patterns for the MultiSense Simulator

Using the results found in 3.3.2 we developed patterns spedificly tuned to help the MultiSense. We began by slightly increasing the number of LEDs used. These LEDs were positioned such that we would illuminate the portions of the image identified to have the worst matching. Specifically we identified that the top of images had worse matching than the bottom, with the top left being the worst position. These observations were made by examining the disparity images and identifying the area of the image with the most black (mismatched) pixels. An example of a pattern made to target the top of an image can be seen in Figure 3.9.

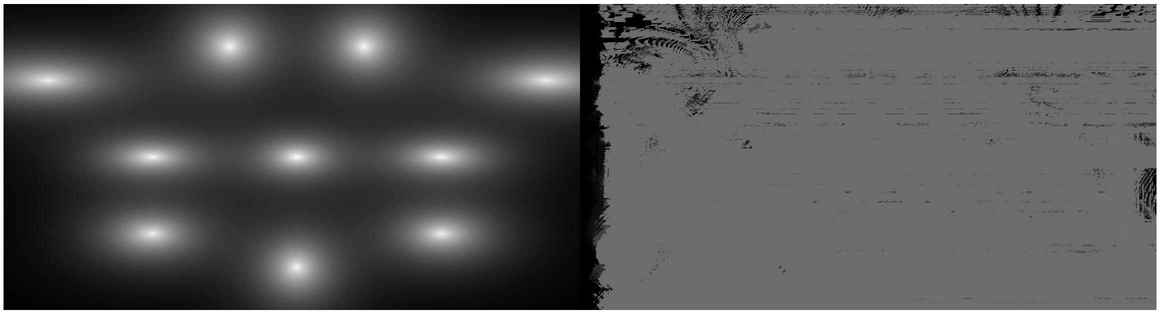

Figure 3.9: Left stereo image of a more complicated pattern of 10 LEDs 5m from the surface (left), and resulting MultiSense disparity (right). All LEDs have a field of view of 10°.

This technique did produce improved results; however clear mismatches were still visible. We tested eleven patterns similar to the one shown, varying the LED placement, LED field of view, and number of LEDs. None of these patterns performed significantly better than the one shown in the example. In general, the top and far right edges of the image produced poor results. Additionally small spots of mismatch were visible in the center of the images.

Finally we moved away from attempting to minimize the number of LEDs used in a pattern and designed patterns with up to thirty-six LEDs. We also tested different LED fields of view with these denser patterns. With more LEDs, smaller fields of view are necessary to keep the LEDs distinct. The smallest field of view tested was 6° as this was the smallest field of view of an LED that we could find readily available on the market.

Through this iterative process we found that increasing the number of LEDs only increased the quality of disparity results up to a point. The relative improvement from each LED decreased as the number of LEDs increased. Furthermore, at around thirty LEDs the overall quality of the images began to decrease. This is because the target was starting to become fully illuminated with a large number of LEDs. We

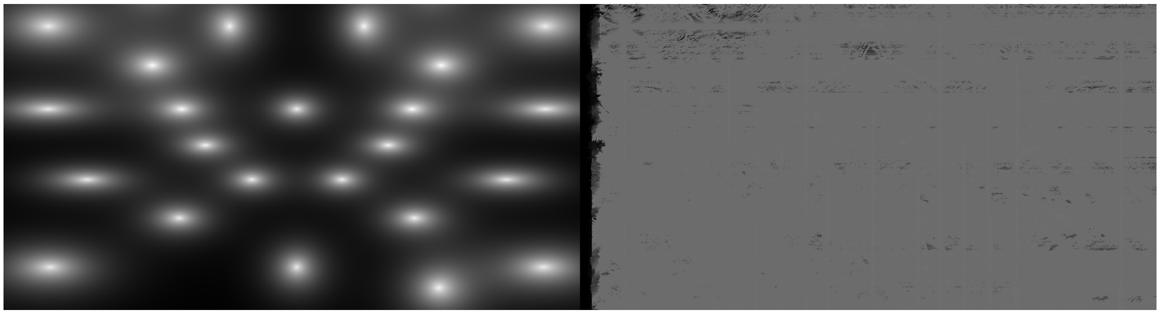

Figure 3.10: Left stereo image of a pattern of 25 LEDs 5m from the surface (left), and resulting MultiSense disparity (right). All LEDs have a field of view of 6°.

also found that LEDs had to be carefully placed in order to improve disparity results. A poorly placed LED would either have no effect or would decrease the quality of disparity results. Ultimately, the pattern we found to produce the best disparity results consisted of twenty-five LEDs. The pattern can be seen in Figure 3.10.

While at first glance the resulting disparity in Figure 3.10 may appear good, it is marred by small areas of mismatched pixels. As laid out in Section 3.3 this is unacceptable. From this we can conclude that the gradual elimination strategy is not viable with the MultiSense camera.

Increasing the number of LEDs or changing the arrangement did improve the disparity results. While decreasing the LED field of view and increasing the number of LEDs may improve the results, this would take the research out of the realm of commonly available LEDs. Furthermore, twenty-five LEDs carefully positioned and angled on the face of a camera is not feasible to manufacture. Increasing the number of LEDs would further exacerbate this problem.

3.4 Further Research

No further research is warranted into using LED gradual illumination strategies with the MultiSense camera. The success of these strategies with a simple offline

SGM algorithm does open the possibility that LED gradual illumination may be successful with different real-time SGM implementations. It is, however, unlikely that LED gradual illumination will outperform laser-based random dot projectors, and the technology is relegated to areas where lasers are not feasible. Furthermore, the simulations presented portray a best-case scenario. No noise is considered, the surface does not absorb light, and the projector is always pointed directly at the surface. All of these are confounding variables that would decrease the matching accuracies in real world applications.

The area where LED gradual illumination is most promising is as an alternative to uniform illumination. Some stereo vision systems do not use a pattern projector, but do contain a light source to illuminate scenes if ambient lighting is not present. These systems could be improved by using a light source that creates nearly uniform illumination, much like the twenty to thirty LED projector presented.

Chapter 4

Experimentation With the MultiSense Stereo Camera

Due to the poor performance of the gradual illumination LED pattern projector in simulation, the decision was reached to abandon that line of research. We began work testing various pattern projection technologies with the goal of identifying possible alternatives to the current MultiSense pattern projection. In the following sections we outline four sets of experiments conducted towards this goal. The first set of experiments was done with the existing MultiSense pattern projection to establish a baseline of performance. The second set evaluated patterns created with diffractive optical elements. The third set used a commercially available projector manufactured by Engage Photonics. The final set of experiments was done using a go between optics (GOBO) projector that was manufactured for use in theater and architectural lighting applications.

For these experiments, Carnegie Robotics supplied a MultiSense S21B stereo camera [14]. The camera was configured with standard lenses that had a focal length of 4.8mm, giving the camera a field of view of 115° x 68°. A picture of the camera can be seen in Figure 4.1. It outputs grey scale images that were 1024 x 544 pixels in size. The camera was capable of outputting unrectified images from the left and right image sensors, rectified images from the left and right image sensors, and the

Figure 4.1: A Carnegie Robotics MultiSense S21B, image courtesy of Carnegie Robotics, LLC. [15]

disparity results from the rectified images. Rectification is the process of aligning images so that we are only offset from each other in one dimension.

For the purpose of this research we recorded the left and right rectified images from the camera. We used these images with the MultiSense simulator (described in Section 3.1.4) to obtain disparity results. This was done so that we could obtain disparities produced at the semi-global matching step. The post processing done by the MultiSense algorithm after the semi-global matching step in part mitigates the problems introduced by low texture in an image. Due to this the disparity results at the semi-global matching step give more direct information about the effects of texture on the quality of image matching than the final disparities.

In the following experiments disparities were examined with a similar qualitative approach to the one laid out in Section 3.3. Section 4.1 presents results from the current MultiSense pattern projector, and these results can be used as a baseline for comparison. Additionally, due to the large field of view of the MultiSense S21B the test target and projector pattern rarely filled the entire field of view of the camera. In these cases only disparities from the area of the image that included both the test target and the pattern were considered.

This experimentation ultimately showed that the projector currently used by Carnegie Robotics performs well and none of the other technologies are well suited to replace it. The GOBO projector, however, did show good results and is a promising area of further research. It currently suffers from practical issues surrounding its size and field of illumination.

4.1 Multisense Pattern Projector

The Carnegie Robotics MultiSense projector makes use of four pattern projector integrated chips (ICs) and is mounted on the front of the MultiSense. Four ICs are needed to fully illuminate the field of view of the camera. The ICs themselves are based on vertical-cavity surface-emitting laser (VCSEL) technology. Each IC produces approximately 5,700 dots and has an optical power of 250mW. The projector operates at a 850nm light wavelength. This wavelength is near infrared and invisible to the naked eye.

4.1.1 Results

We took readings with the MultiSense and MultiSense projector in two different arrangements. In the first arrangement, the camera and projector were located 1.56m from the test target; in the second, the setup was located 6.7m from the test target. For both of these arrangements all overhead lighting was turned off and an effort was made to eliminate sources of ambient light in the area.

Figure 4.2 shows the 1.56m test setup. Figure 4.3 shows the resulting images and disparities from this setup without any pattern projection. Figure 4.4 shows

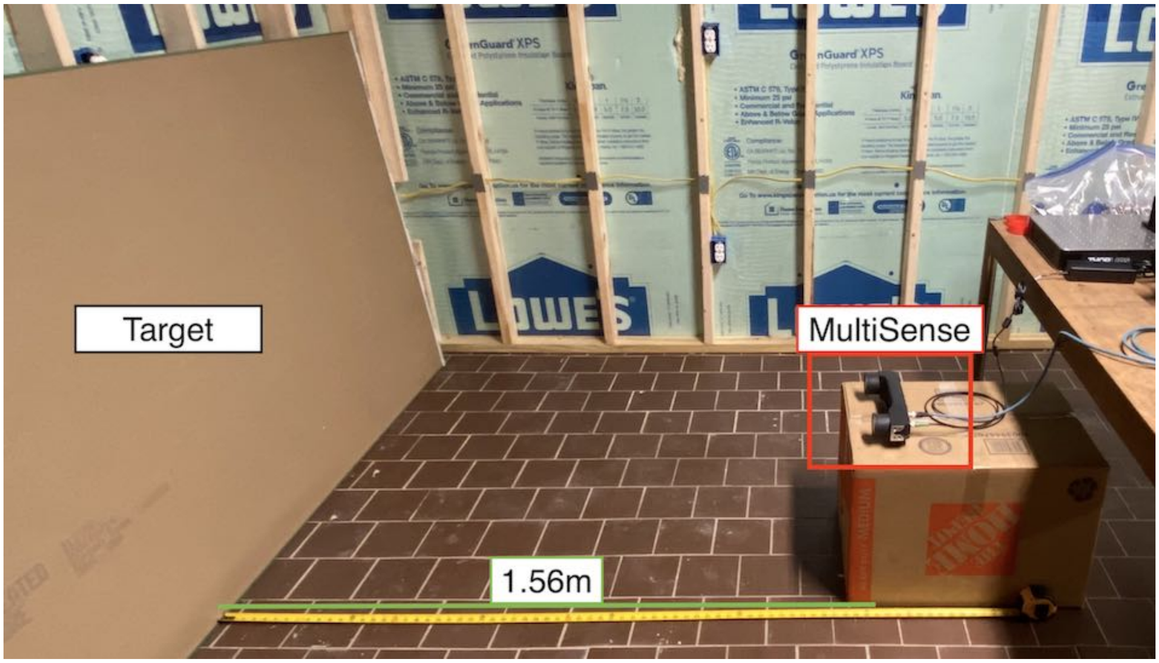

Figure 4.2: Experimental Setup with MultiSense 1.56m from the test target.

the results with the MultiSense projector. In this setup the MultiSense projector produces disparity results over the area of the test surface that are on par with the ideal disparity results shown in figure 3.5.

Figure 4.5 shows the 6.7m test setup. Figure 4.6 shows the results of using the Multisense projector at this range. The projector does not have the optical power to illuminate the full scene at this range, leading to poor matching at the test target.

4.2 Diffractive Optical Element Pattern Projection

We prototyped a pattern projector using a diffractive optical element (DOE) and a laser source. DOEs have finally engraved microstructures that reshape an incoming laser beam into an arbitrary pattern [16]. They have a history of use in stereo camera pattern projection applications [5] [6]. Our goal with the prototype

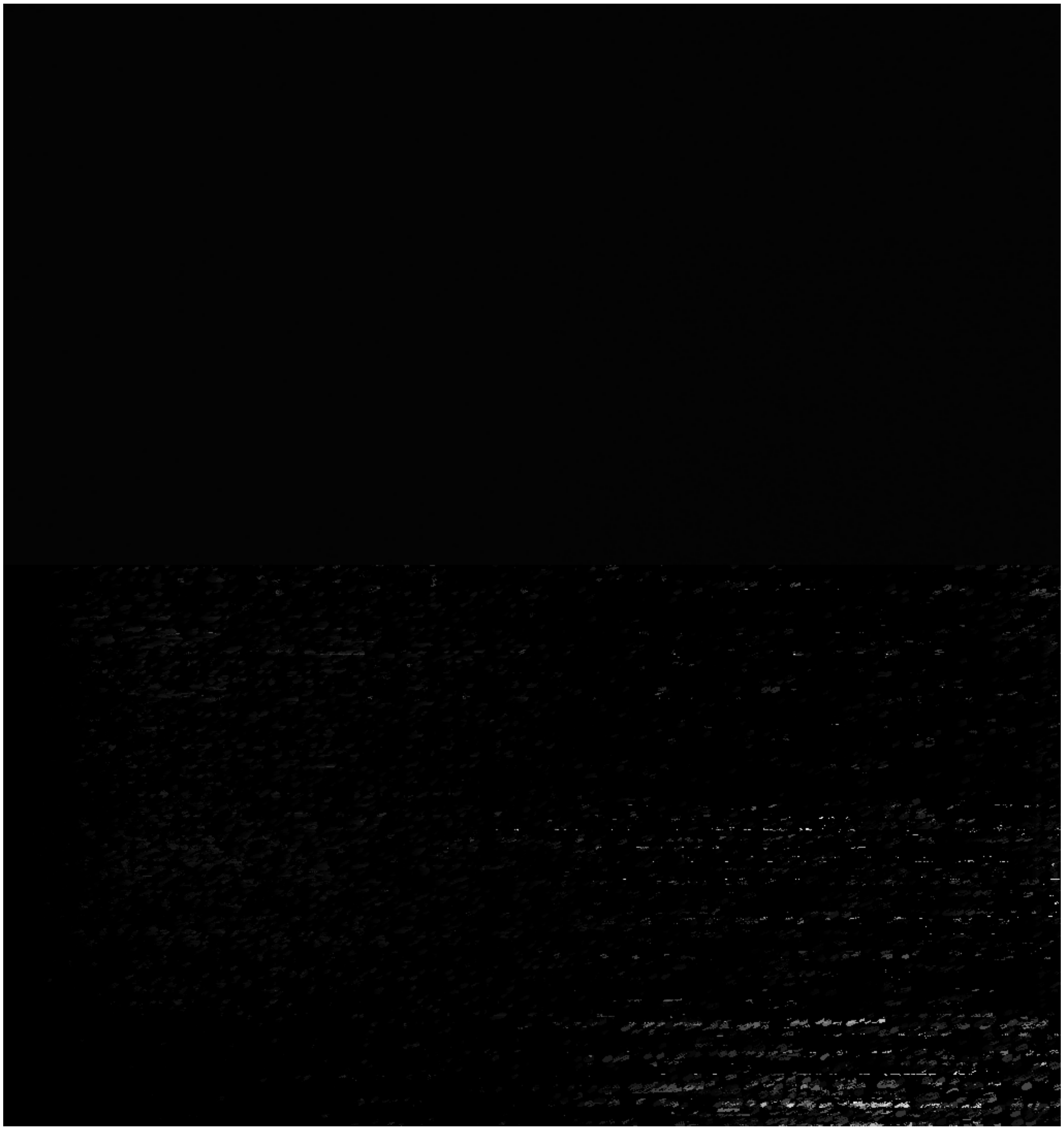

Figure 4.3: MultiSense image with no projector and no overhead light. With no ambient light the image appears all black (top). The resulting disparity (bottom) has almost no correct matching.

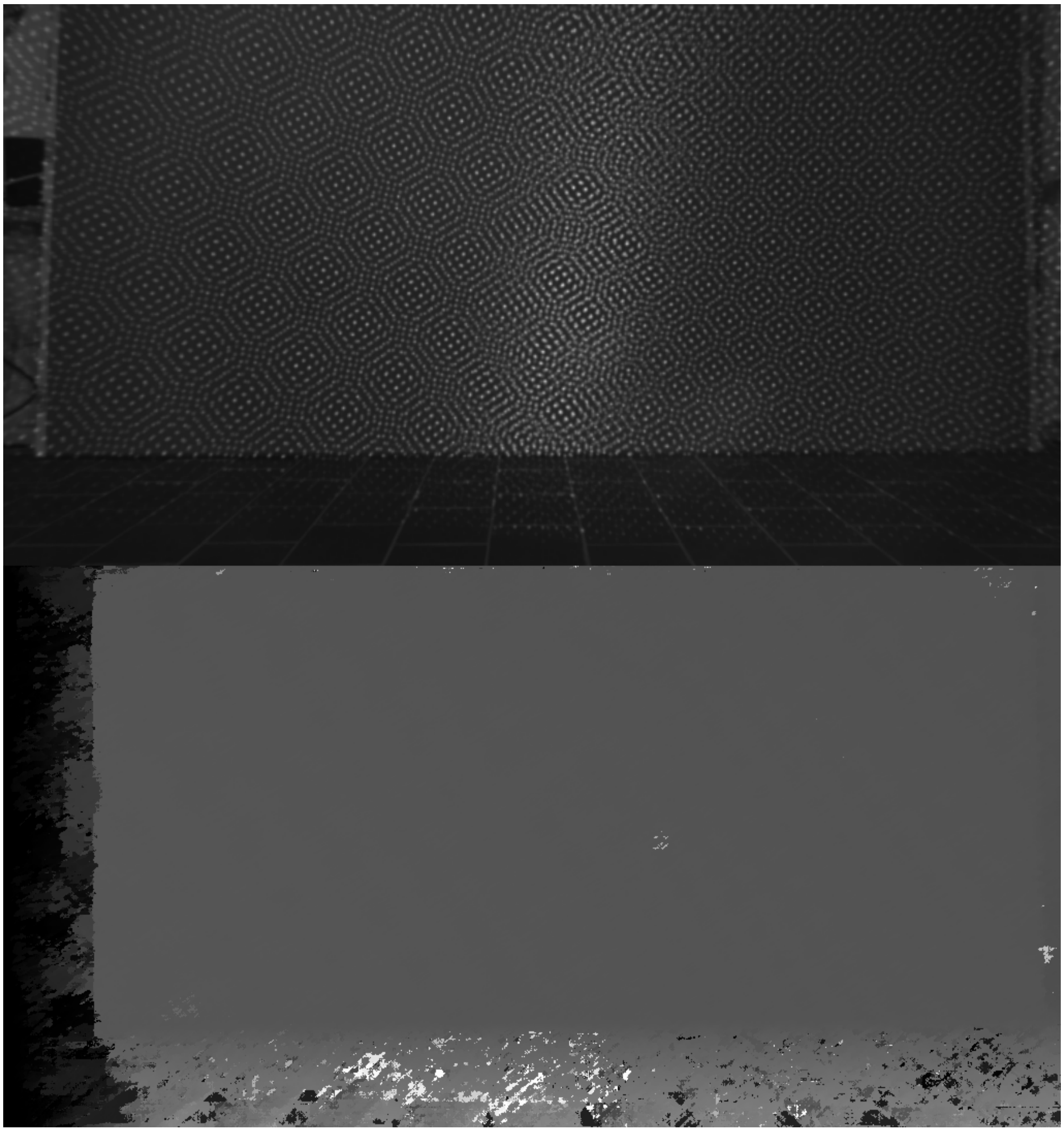

Figure 4.4: Pattern produced by MultiSense pattern projector at 1.56m from the target (top) and resulting disparity (bottom)

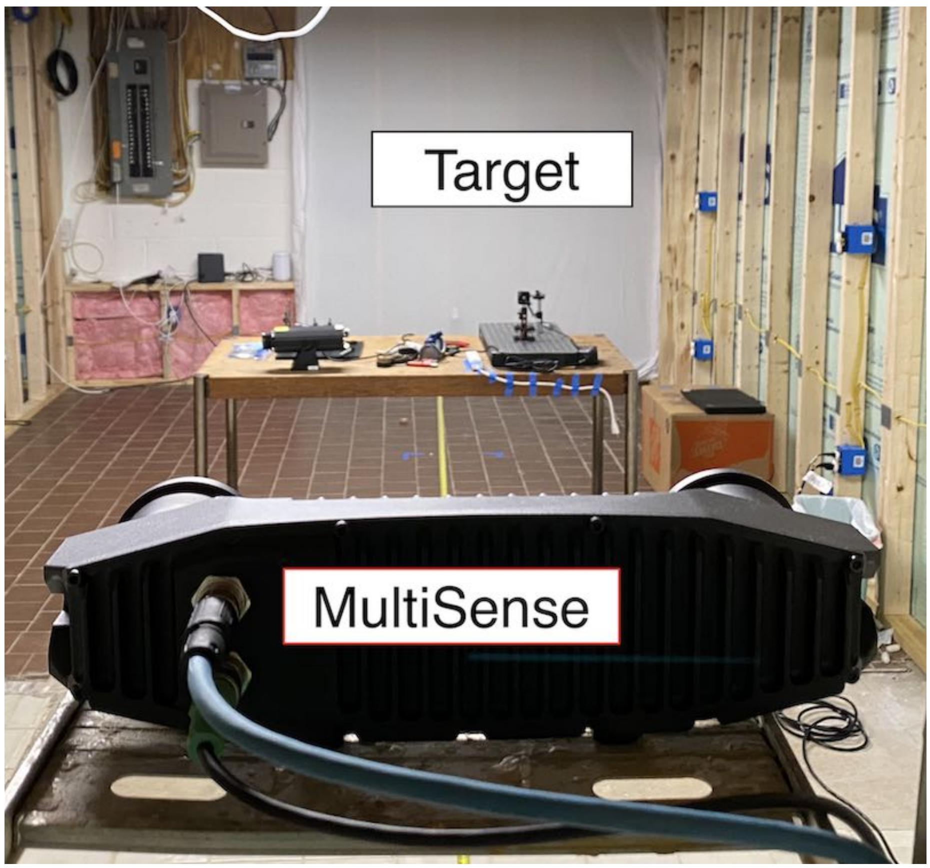

Figure 4.5: Experimental Setup with the MultiSense 6.7m from the test target.

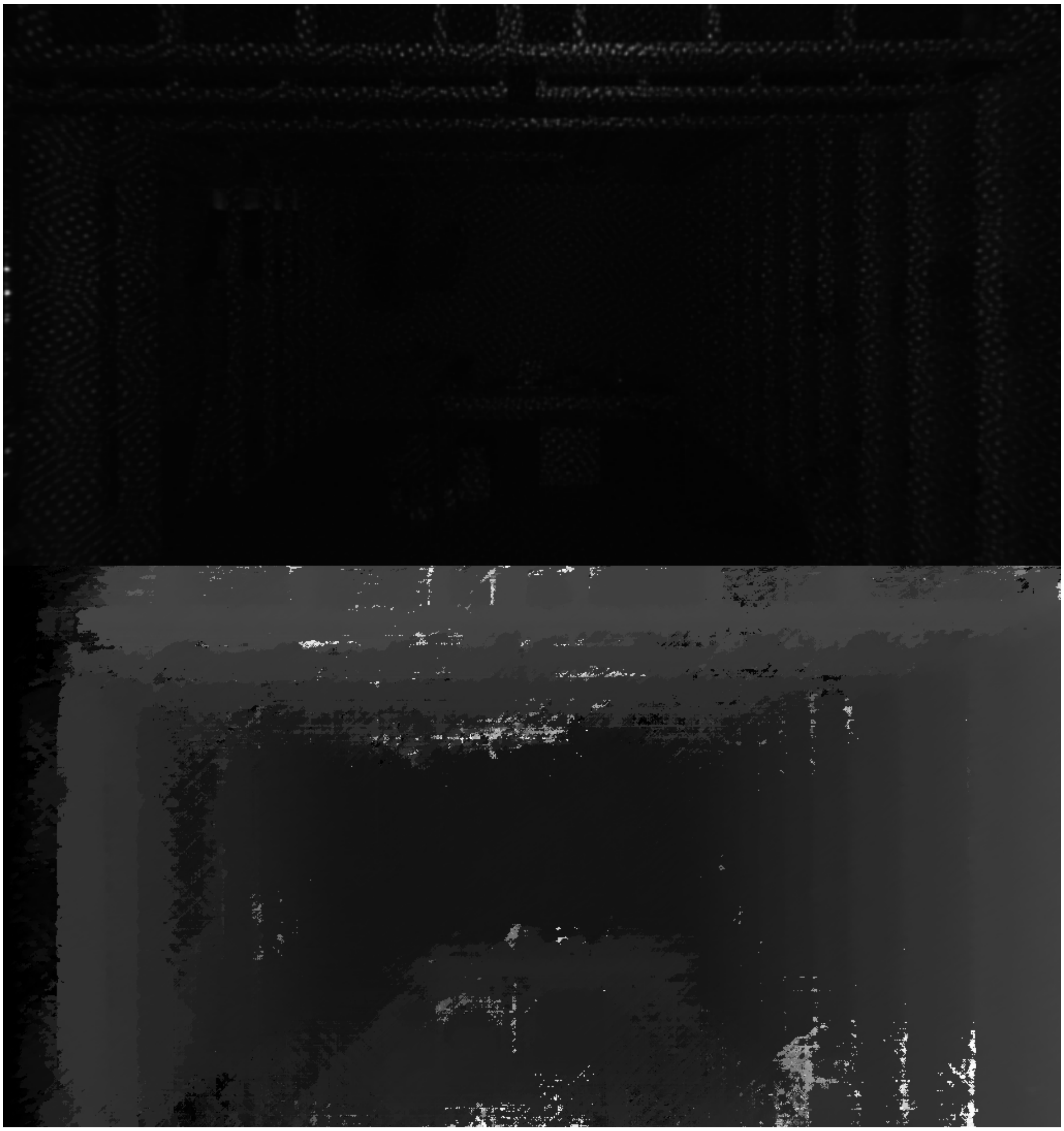

Figure 4.6: Pattern produced by MultiSense pattern projector at 6.7m from the target(top) and resulting disparity (bottom)

projector was to evaluate the feasibility and practicality of a DOE based pattern projector manufactured by Carnegie Robotics. We found that while it is possible to build such a system, trade-offs would have to be made in both size and eye safety for it to outperform the current MultiSense pattern projector.

4.2.1 Setup

For our prototype projector we used a HoloEye DE-R 335 DOE [17] and a ThorLabs CPS650F laser diode module [18]. The DE-R 335 creates a pseudo-random 33,000 dot pattern and is designed to work with 630-660nm (red) light. The CPS650F module is an integrated laser diode, driving circuit, and collimating optics. It has an optical output power of 4.5mW at 650nm, and adjustable focus.

It is ultimately desirable to have the pattern projector function in the near infrared range so the resulting pattern is not visible. We conducted these DOE experiments using visible light for convenience and eye safety while experimenting.

For our setup we mounted the DE-R 335 and CPS650F on an optical breadboard approximately 2m from the test target. We adjusted the laser such that the beam was in focus on the target. We placed the DOE 20mm in front of the laser source, and rotated the DOE so that it produced a level rectangular pattern. Images of the optical setup can be seen in Figure 4.7, and the resulting pattern can be seen in Figure 4.8.

4.2.2 Results

We had difficulty imaging the DOE pattern with the MultiSense camera. The pattern produced by our setup was not bright enough for the camera. Using the MultiSenses default exposure time of 0.03s the pattern did not overcome the noise floor of the camera. At the Multisense maximum exposure time of 1.0s the pattern

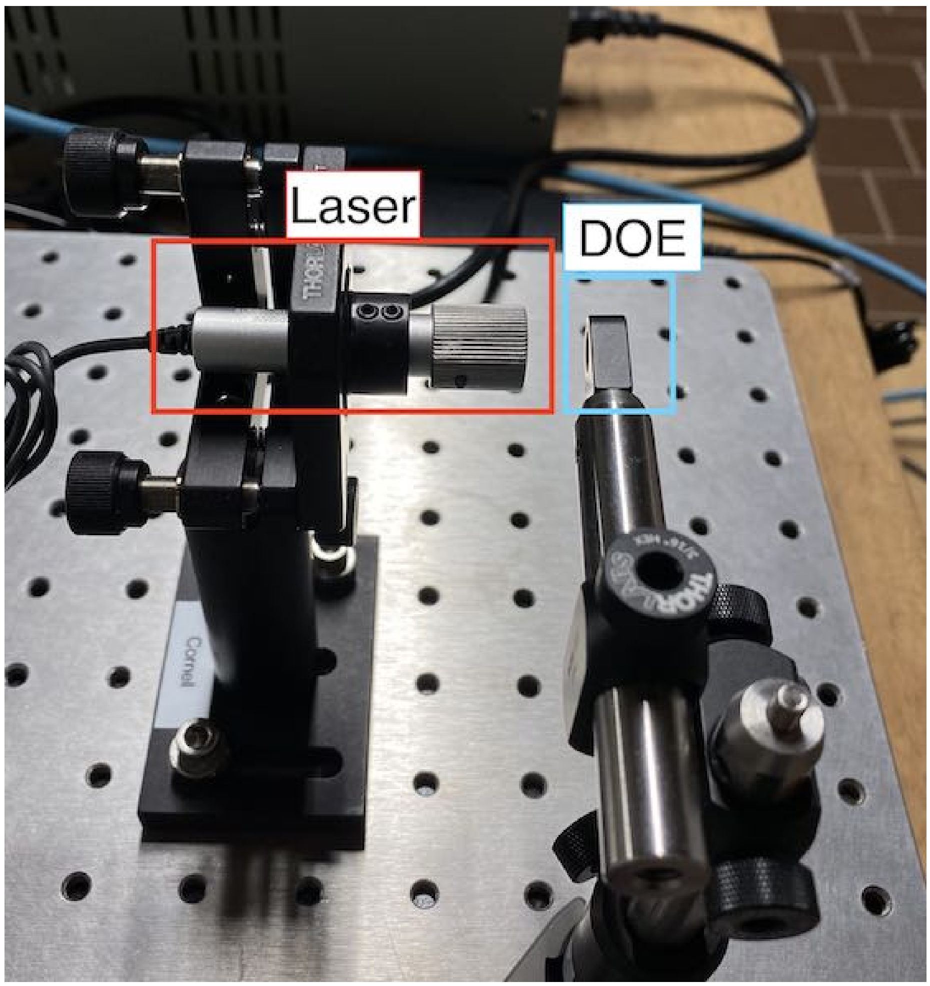

Figure 4.7: CPS650F laser and HoloEye DE-R 335 DOE setup

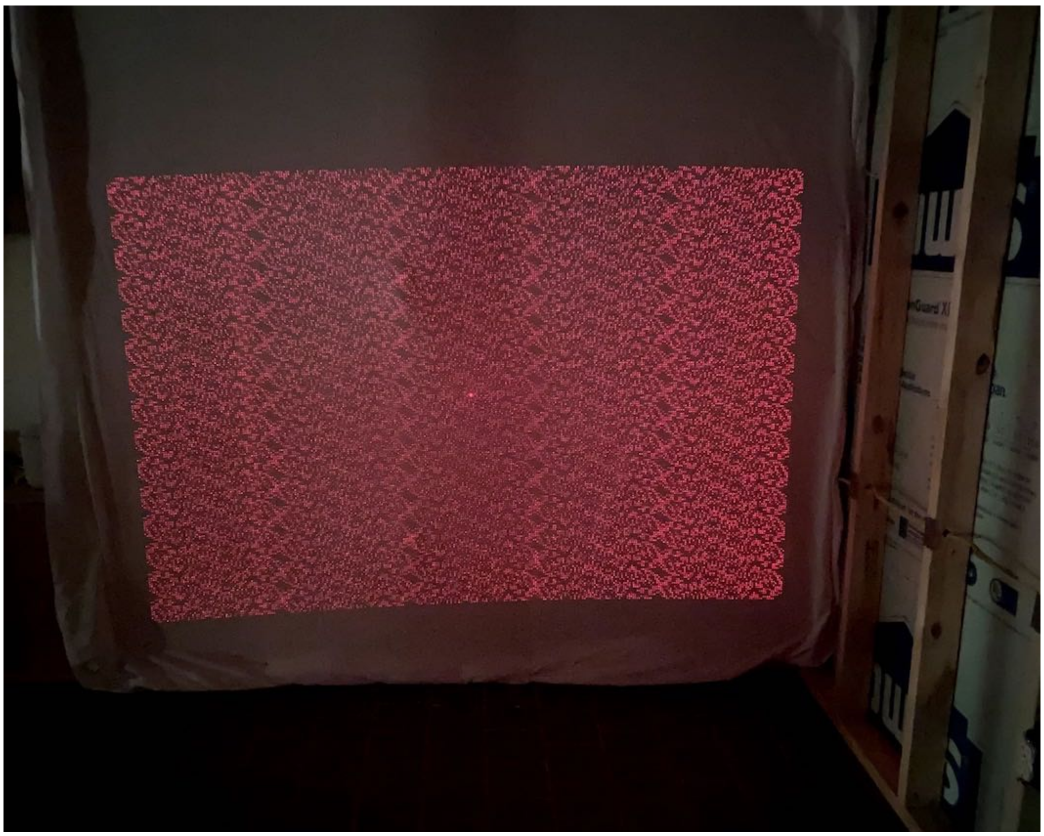

Figure 4.8: Pattern from HoloEye DE-R 335 DOE

did overcome the noise floor, but not enough to lead to good stereo matching. This can be seen in Figure 4.9.

To create a simulacrum of using the DE-R 335 with a more powerful laser source we used the image from Figure 4.8. This image was taken on an iPhone 11 Pro with a 0.25s exposure and has a resolution of 4032 x 3024 pixels. We cropped the image such that we obtained two images of the left center portion of the original image. These two images were offset from each other by sixty pixels in the X direction and both had the same resolution as the multisense images. We then used these images as inputs to the multisense simulator. The left input image and the resulting disparity can be seen in Figure 4.10. The resulting disparity image is perfect. This points to pattern characteristics of the DE-R 335 being a good choice for use with the MultiSense. These results need to be recreated using a higher powered laser and the MultiSense before a firm conclusion can be reached.

The need for a higher powered laser points to a problem with the discrete DOE and laser source approach. The DOE greatly disperses the optical power of the incoming laser beam, making the resulting pattern less dangerous to eyes than the initial laser beam. The danger of the initial beam is however still present. If the DOE was to be removed or fail in some way, it would be possible for people to be exposed to the full power of the laser. This risk is taken into account in eye safety laser ratings. Furthermore, the DOE has a prominent zero order spot. This spot can be seen in Figure 4.8 as the bright point in the center of the pattern. The zero order spot is the attenuated incoming beam, and is always the brightest point in the pattern. As the power of the laser source increases the zero order intensity increases proportionally. This zero order point sets the eye safety limit, not the intensity of the entire pattern. As the optical power of the laser source increases one must insure that the zero order spot remains within eye safety limits.

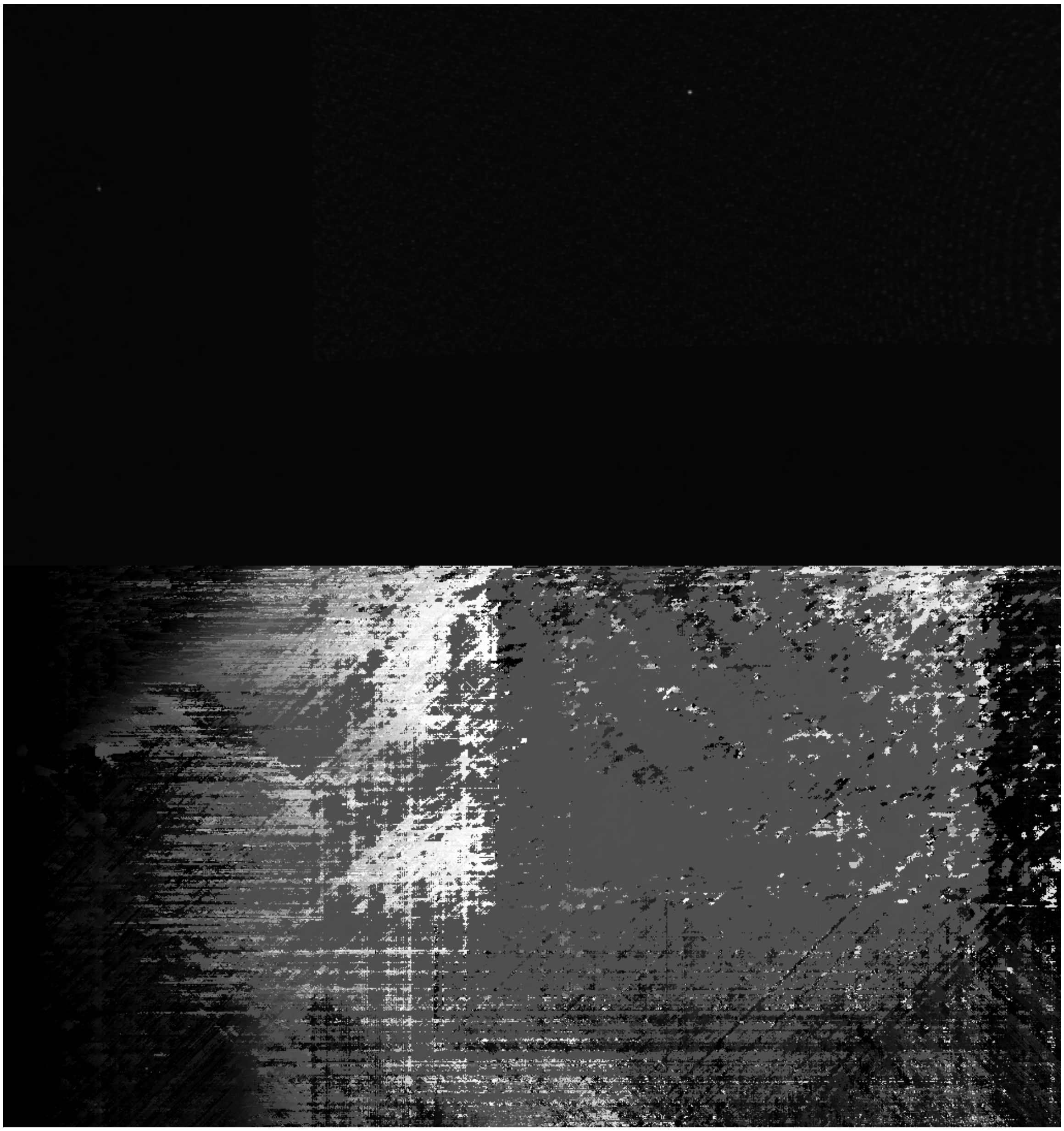

Figure 4.9: MultiSense DOE pattern image (top) and resulting disparity (bottom). The only part of the pattern that is easily visible is the zero order sport near the top right.

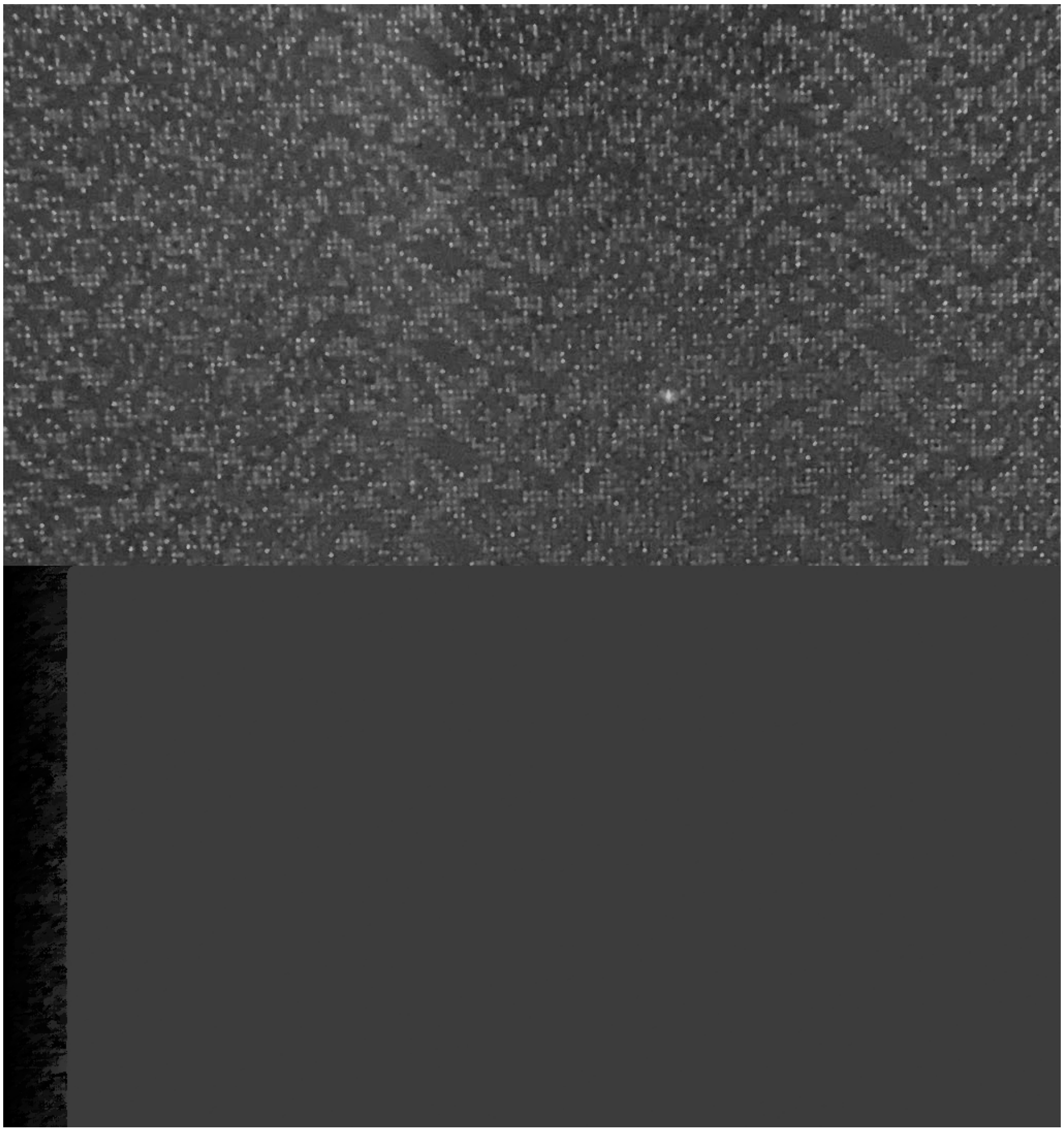

Figure 4.10: Cropped iPhone 11 Pro DOE pattern image (top) and resulting disparity (bottom)

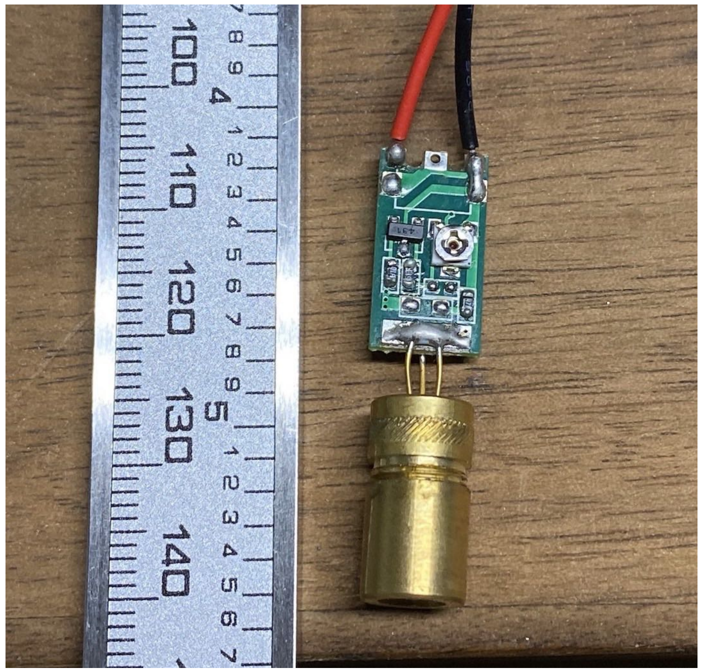

Figure 4.11: Engage Photonics CEE850 10k pattern projector

4.3 Engage Photonics Pattern Projector

We purchased a CEE50-10k pattern projector from Engage Photonics. This pattern projector operated at 850nm wavelength and produced approximately 10,000 dots with a 59° x 46° field of illumination. The projector was in a TO-can package and included a driving circuit. A picture of the projector and circuit can be seen in Figure 4.11.

The pattern produced by the projector and the resulting disparity can be seen in Figure 4.12. The pattern shows a prominent zero order spot leading us to assume that the projector works off of very similar principles to those discussed in section 4.2. The main differences between this projector and the prototype we built in that section was that the Engage Projector had the laser source and DOE integrated into a single package. It additionally had a much more powerful laser source.

The disparity results achieved with this projector are excellent. However the size and eye safety concerns from the more powerful laser diode do not make it a good general replacement to the existing MultiSense projector. The Engage Projector would be a good choice for applications that require higher optical output power and can tolerate eye safety concerns.

4.4 GOBO Pattern Projection

Previous attempts to use go between optics (GOBOs) for stereo matching have been ineffective [8]. GOBOs traditionally are pieces of metal with patterns cut in them. GOBO projectors function by shining light at the GOBO and then projecting the light that passes through the cut out portions of the metal pattern. In previous research, attempts were made to use this technology to produce random dot patterns. These projectors did not function well at range because the dots would get large and eventually combine producing uniform illumination.

In addition to metal GOBOs, which create a pattern by simply blocking light, there are glass GOBOs that use engraved glass to produce patterns consisting of bright areas and shadows. These GOBOs are more optically efficient than metal GOBOs because they do not fully block light. This does mean that we can-not produce patterns with discrete dark and bright areas which is needed for a random dot pattern. We can, however, produce complicated textured patterns that are suitable for stereo matching.

Since GOBO projectors do not use lasers and are often based on LED light sources, we have many of the benefits discussed in Section 2. Additionally the efficacy of the types of patterns produced by glass GOBOs for stereo matching is not well under-stood. We purchased an ECO Spot LED10 GOBO projector [19] and three glass GOBOs with different patterns. We found that all the patterns tested produced good stereo matching results with the MultiSense. We additionally tested one of the pat-

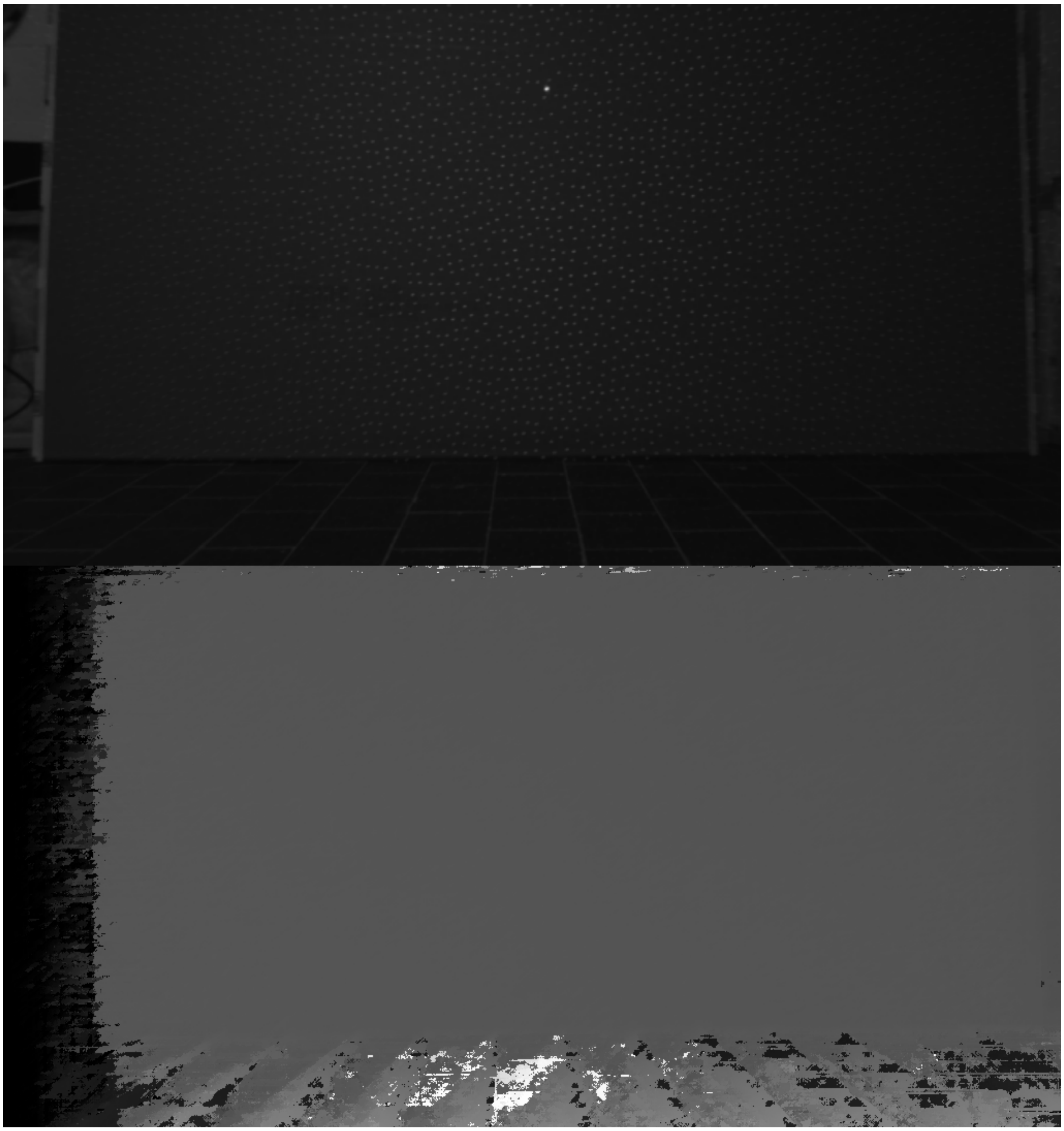

Figure 4.12: Engage Photonics CEE850 10k pattern (top) and resulting disparity (bottom)

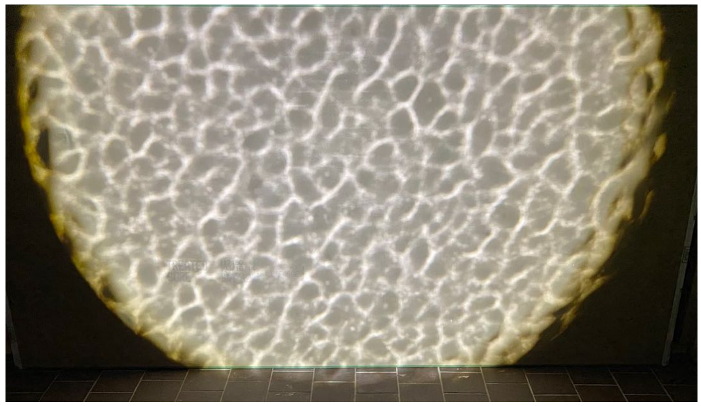

Figure 4.13: Pattern produced by glass GOBO "pebble" pattern

terns to find the maximum range in which we achieved good results. We found this range to be roughly 5m, but were unable to determine if this range limit was set by the optical power of our projector or by the characteristics of the pattern being projected.

4.4.1 Pattern Selection

We selected three separate GOBO patterns to evaluate. They can be seen in Figures 4.13, 4.14, and 4.15. The selection of patterns was difficult because GOBOs are designed and manufactured to be aesthetically pleasing or to reproduce natural patterns. These three GOBOs were selected because the marketing images led us to believe we would produce good texture across the whole pattern. The GOBO in Figure 4.15 was produced using a slightly different technology than the other two tested. It was made by putting a treatment on top of the glass while the other two were made through engraving a pattern into the glass itself.

For the evaluation of the patterns, a similar setup to the one shown in Figure 4.2 was used. The MultiSense was placed 1.56m from the test target and the GOBO projector was placed 2.16m from the test target and 0.38m above the MultiSense. The projector was placed behind and above the MultiSense so that the resulting pattern

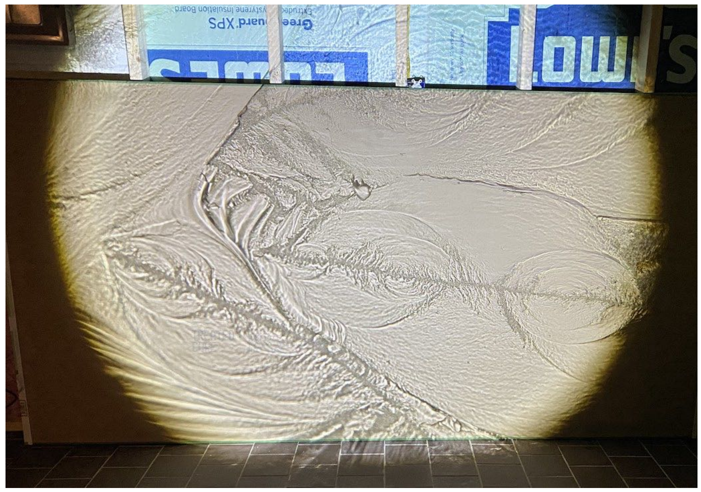

Figure 4.14: Pattern produced by glass GOBO “plume” pattern

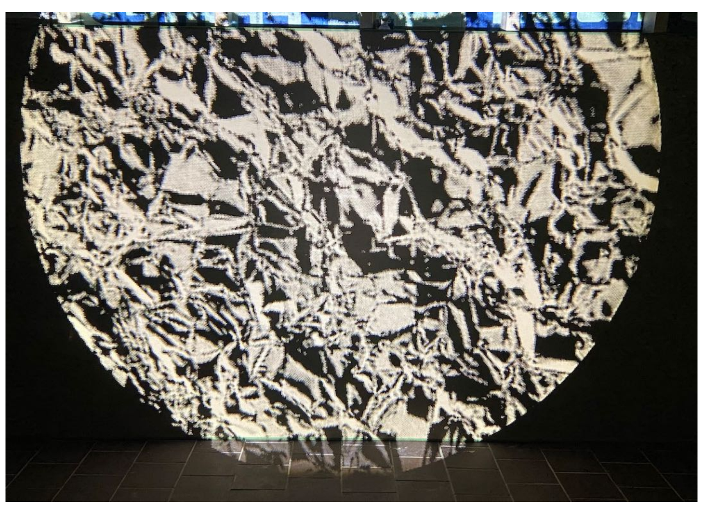

Figure 4.15: Pattern produced by glass GOBO “crumble” pattern

would fill as much of the camera field of view as possible.

All three patterns produced highly quality disparity results. The MultiSense image and resulting disparity of the Figure 4.13 pattern are shown in Figure 4.16 for reference. This performance was a surprising result. Both the patterns in 4.14 and 4.15 have areas with lower texture. We expected this to cause spots of poor matching but this did not occur. We additionally compared the GOBO pattern disparities to the results of using the projector to illuminate the test target with no GOBO This produced uniform illumination on the test target and resulted in significantly worse disparity images than those achieved with the GOBO patterns.

4.4.2 Effective Range

To determine if the glass GOBO patterns suffered from the same range dependency seen in other stereo vision GOBO pattern projectors we tested the pattern in Figure 4.13 at a variety of ranges. In these experiments, we left the MultiSense positioned 1.56m from the test target and varied the distance of the projector to the test target from 2.16m to 5.20m.

The quality of disparity results started to show degradation at approximately 4m (Figure 4.17) and were completely unusable by 5.2m (Figure 4.18). At 5.2m the pattern brightness as seen in the input image was quite low. It was unclear if the failure to match with the projector at this range was due to the pattern brightness or the patterns characteristics. To answer this question an additional study would be needed. All images in this study were taken with the default MultiSense exposure time of 0.03s. Testing should be done to see if increasing the image exposure time or using a more powerful projector could lead to acceptable disparity results at the 5m and greater range.

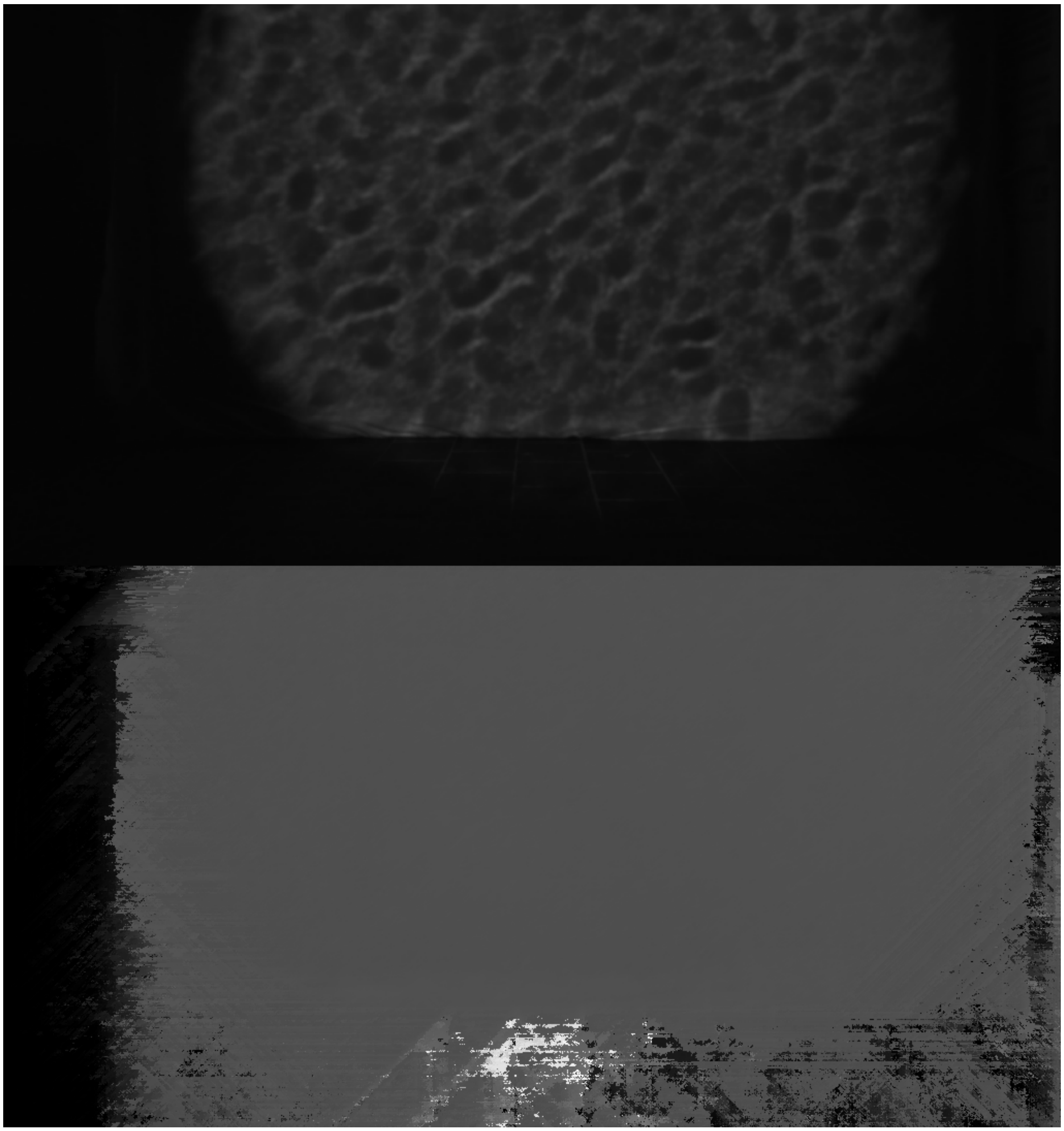

Figure 4.16: GOBO projector 2.16m from the test target (top) and and resulting disparity (bottom)

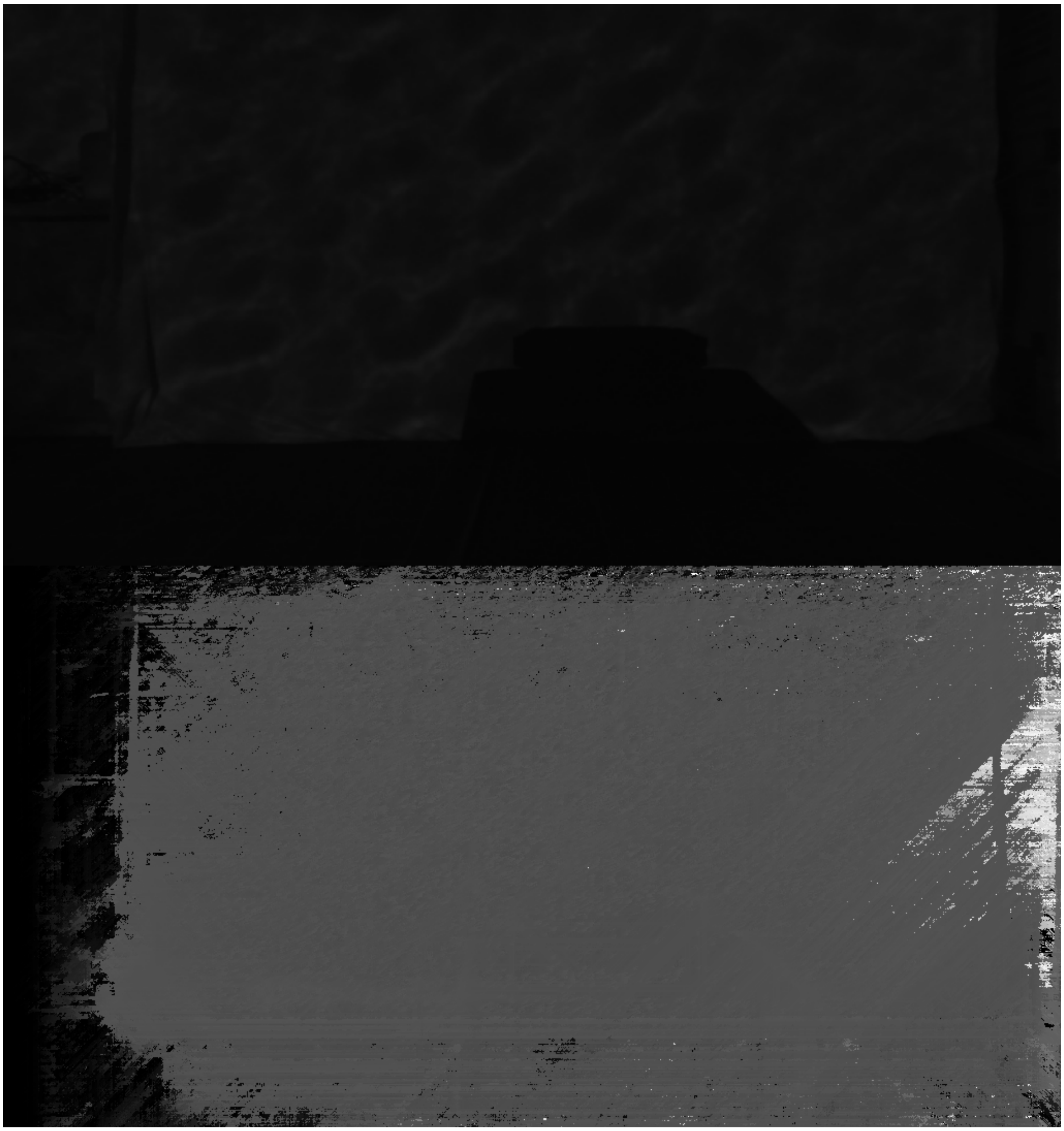

Figure 4.17: GOBO projector 3.99m from the test target (top) and and resulting disparity (bottom)

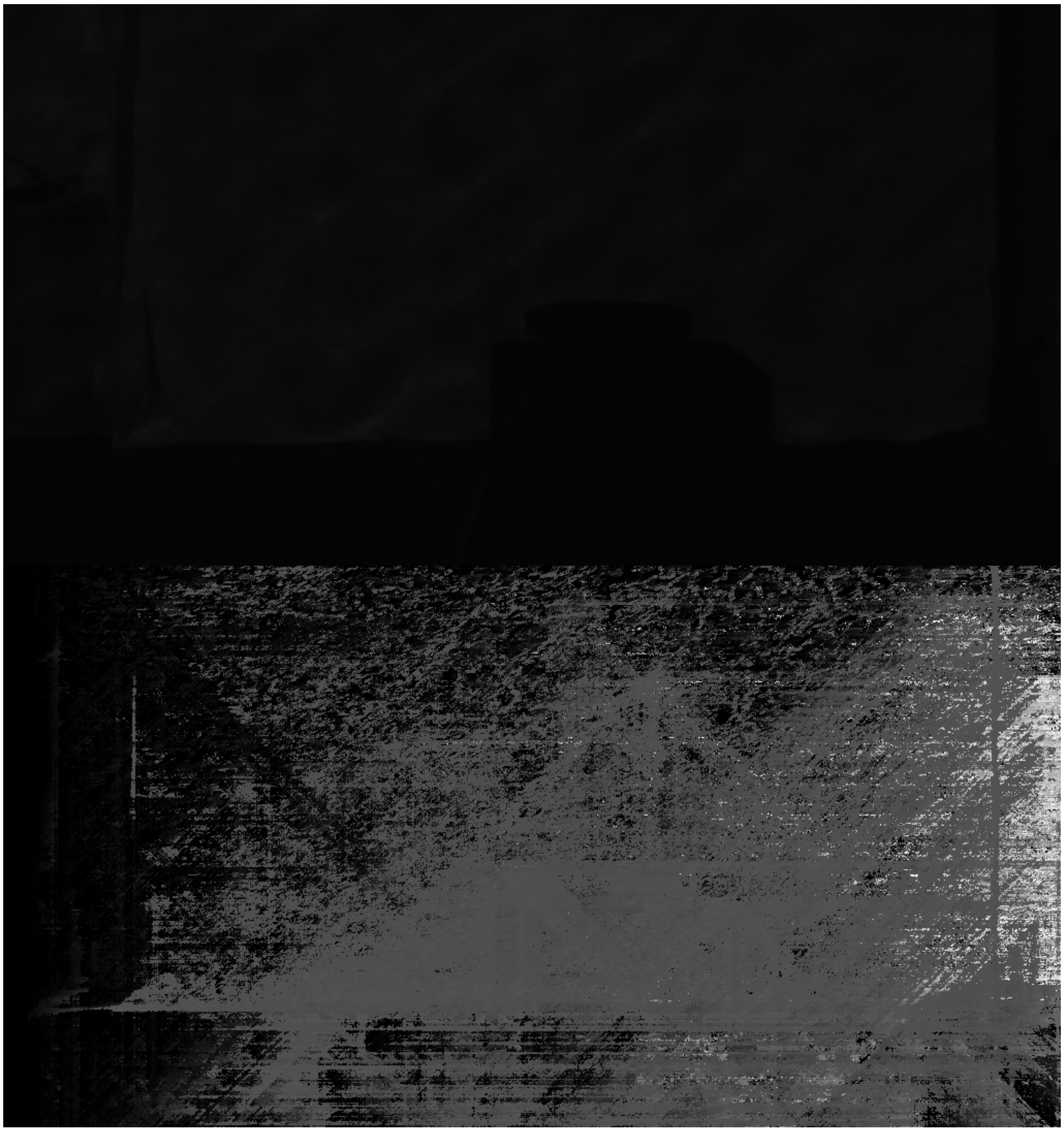

Figure 4.18: GOBO projector 5.20 from the test target (top) and and resulting disparity (bottom)

4.5 Further Research

None of the technologies tested show potential to replace the MultiSense pattern projector in the short term. While more work can be done with the DOEs to produce patterns that are recognizable to the MultiSense, it is not clear that is a worthwhile line of investigation. As the power of the laser diode increases, the difficulty to drive the diode and safety concerns do as well. The ultimate refinement of the DOE approach would most likely produce something similar to the Engage Projector. While that projector does produce high quality disparity results, it does so with questionable eye safety, a large package size, and significant heat dispersion.

The GOBO projector is the most promising line of research. It produced high quality results and can make use of the significant LED benefits discussed in section 2. Furthermore, the patterns produced by the projector are unlike anything currently used by a stereo camera. Much research and refinement is still needed before a GOBO projector can be deployed with a MultiSense. First, the size of the projectors must be reduced. Since GOBO projectors are primarily used in entertainment and architectural lighting situations the size of the projector is not a focus. This means that the smallest projectors on the market are roughly the size of the MultiSense camera itself. Finally, since GOBO projectors produce patterns made to be seen by humans they all function in the visible light range. While we attempted to prototype a GOBO projector that functioned in the near infrared range, we were not able to get it to a point where it could produce results. More work needs to be done in this area to test the technology with larger wavelength light.

4.6 Overview of VCSEL Array Technology

A natural question that arises from the results in this section is how do the pattern projection ICs used in the MultiSense pattern projector work? They use laser diode technologies, but do not have a zero order spot characteristic of DOE projectors. Additionally they are physically quite small.

These characteristics can be achieved by using an array of vertical-cavity surface-emitting lasers (VCSELs). Individual VCSEL emitters have a small maximum output power , but many VCSEL emitters can be arranged on a single silicon die with spacing of a few microns [20]. This means a random dot projector can be built that uses a large number of discrete low powered laser sources instead of splitting one high powered laser source. The challenge with building a projector like this is separating the different beams from each other to create a large field of illumination. Existing micro-optics technologie is capable of this [21].

A VCSEL array is most likely the approach used by the pattern projector ICs in the current multisense projector. It is also most likely the technology used in projectors on other stereo cameras such as the Intel RealSense which have similar pattern chartestics [22]. A VCSEL array projector requires highly specialized manufacturing capabilities to produce. A Silicon wafer must be fabricated, diced, and packaged. Additionally high quality micro-optics must be manufactured and integrated with the VCSEL array.

Chapter 5

Conclusion

Overall our research found that the pattern projector currently in use on the MultiSense camera is a good general purpose option. We explored two possible types of LED projectors. We found that the gradual illumination LED strategy is ineffective and that the use of GOBOs is promising but requires significant further research. We also experimented with diffractive optical element-based alternatives to the current MultiSense projector. This showed that while it is possible to both build and buy projectors based on diffractive optical elements, they are accompanied by tradeoffs that make them worse options than the current projector in use. We found that the best option for stereo vision pattern projectors is using VCSEL arrays with micro-optics. This technique leads to eye safe projectors with good optical output power and small physical footprints. These projectors, however, cannot be built with discrete components. This means that they must be manufactured by companies with IC expertise and capabilities. It seems the best option for pattern projection is to buy a projector from a company that specializes in integrated electro-optical components.

Bibliography

[1] K. Konolige, "Projected texture stereo," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 148–155. DOI: 10.1109/ROBOT.2010.5509796.

[2] W. Jang, C. Je, Y. Seo, and S. W. Lee, "Structured-light stereo: Comparative analysis and integration of structured-light and active stereo for measuring dynamic shape," *Optics and Lasers in Engineering*, vol. 51, no. 11, pp. 1255 –1264, 2013, ISSN: 0143-8166. DOI: https://doi.org/10.1016/j.optlaseng.2013.05.001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0143816613001528.

[3] J. Jiang, J. Cheng, and H. Zhao, "Stereo matching based on random speckle projection for dynamic 3D sensing," in *2012 11th International Conference on Machine Learning and Applications*, vol. 1, Dec. 2012, pp. 191–196. DOI: 10.1109/ICMLA.2012.40.

[4] J. Lim, "Optimized projection pattern supplementing stereo systems," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 2823–2829. DOI: 10.1109/ROBOT.2009.5152786.

[5] S. Thuries, A. Gillet, and B. Puybras, *Us 10,247,547 b2*, United States Patent, Apr. 2019.

[6] J.-c. Jeong, H. Shin, J. Chang, E.-G. Lim, S. M. Choi, K.-J. Yoon, and J.-i. Cho, "High-quality stereo depth map generation using infrared pattern projection," *ETRI Journal*, vol. 35, no. 6, pp. 1011–1020, 2013. DOI: 10.4218/etrij.13.2013.0052. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.4218/etrij.13.2013.0052. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.4218/etrij.13.2013.0052.

[7] *IEC 60825-1: Safety of laser products*, International Electrotechnical Commission, Aug. 2001.

[8] A. Grunnet-Jepsen, J. N. Sweetser, P. Winer, A. Takagi, and J. Woodfill, "Projectors for Intel(R) RealSense(TM) depth cameras d4xx," Intel, Tech. Rep., Mar. 2019.

[9] L. Keselman, J. I. Woodfill, A. Grunnet-Jepsen, and A. Bhowmik, "Intel(R) RealSense(TM) stereoscopic depth cameras," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jul. 2017, pp. 1267–1276. DOI: `10.1109/CVPRW.2017.167`.

[10] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, San Diego, CA, USA, Jun. 2005, pp. 807–814.

[11] ——, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328–341, 2008.

[12] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, Jun. 2003, pp. 195–202.

[13] I. Haller and S. Nedevschi, "GPU optimization of the SGM stereo algorithm," in *Proceedings of the 2010 IEEE 6th International Conference on Intelligent Computer Communication and Processing*, 2010, pp. 197–202.

[14] *MultiSense stereo brochure*, Carnegie Robotics, LLC, 4501 Hatfield Street Pittsburgh, PA 15201, Mar. 2019.

[15] *S21B whitebackground*. [Online]. Available: `https://carnegierobotics.com/multisense-s21b`.

[16] *DOE diffractive optical elements*, 3rd ed., HOLOEYE Photonics AG, Volmerster. 1 12489 Berlin, Germany.

[17] *Holoeye DE-R 335 DOE*, 3.11, HOLOEYE Photonics AG, Volmerster. 1 12489 Berlin, Germany.

[18] *CPS650F spec sheet*, Rev A, ThorLabs, 56 Sparta Avenue Newton, New Jersey 07860 United States, Nov. 2014.

[19] *ECO spot LED10 user manual*, Globus New Media LLC, 2019.

[20] R. Paschotta, *VCSEL arrays*. [Online]. Available: `https://www.rp-photonics.com/vcsel_arrays.html`.

[21] A. Grunnet-Jepsenand, K. Swaminathan, and L. M. Keselman, *Us 20170186166A1*, United States Patent, Jun. 2017.

[22] *Intel(R) RealSense(TM) D400 series product family datasheet*, 7th ed., Intel, Oct. 2019.