

---

Theses and Dissertations

---

Spring 2015

# Development of airborne light field photography

Michael Dominick Yocius  
*University of Iowa*

Copyright 2015 Michael Dominick Yocius

This dissertation is available at Iowa Research Online: <http://ir.uiowa.edu/etd/1812>

---

## Recommended Citation

Yocius, Michael Dominick. "Development of airborne light field photography." PhD (Doctor of Philosophy) thesis, University of Iowa, 2015.  
<http://ir.uiowa.edu/etd/1812>.

---

Follow this and additional works at: <http://ir.uiowa.edu/etd>



Part of the [Electrical and Computer Engineering Commons](#)

DEVELOPMENT OF AIRBORNE LIGHT FIELD PHOTOGRAPHY

by

Michael Dominick Yocius

A thesis submitted in partial fulfillment  
of the requirements for the Doctor of  
Philosophy degree in Electrical and Computer Engineering  
in the Graduate College of  
The University of Iowa

May 2015

Thesis Supervisor: Associate Professor Thomas Schnell

Copyright by  
MICHAEL DOMINICK YOCIUS  
2015  
All Rights Reserved

Graduate College  
The University of Iowa  
Iowa City, Iowa

CERTIFICATE OF APPROVAL

---

PH.D. THESIS

---

This is to certify that the Ph.D. thesis of

Michael Dominick Yocius

has been approved by the Examining Committee  
for the thesis requirement for the Doctor of Philosophy  
degree in Electrical and Computer Engineering at the May 2015 graduation.

Thesis Committee: \_\_\_\_\_  
Thomas Schnell, Thesis Supervisor

\_\_\_\_\_  
David R. Andersen

\_\_\_\_\_  
Er-Wei Bai

\_\_\_\_\_  
Thomas F. Boggess

\_\_\_\_\_  
Anton Kruger



To my grandparents, Dominick and Agnes Yocius and Norbert and Dorothy Patek, who instilled the value of education from a very early age.

## ACKNOWLEDGMENTS

I would like to thank Professor Tom Schnell for advising me throughout my time as a graduate student. His lab, the Operator Performance Laboratory (OPL), allowed me to cut my teeth as an engineer. The scope of the different types of projects that I have had the opportunity to be a part of is an experience that I could have only had at OPL. I would like to thank all of the labmates I have worked with at OPL, especially Matt Cover, Nick Lorch, Carl Richey and Joe Engler. Over the years they have helped teach me software development, aviation concepts, systems design and many other fields.

I would also like to thank Professor David Andersen for teaching multiple courses on optics that gave me the background for this project. Thanks for all of the office hours over the years that helped me work through different concepts and discuss outdoor recreation. Professor Er-Wei Bai for teaching signal processing courses in ways that helped me understand concepts that I used in this dissertation. Professor Thomas Boggess for teaching some of the hardest material I learned in college. Professor Anton Kruger for always stopping me in the hallway to talk about what I was currently working on, from solar bike club to this dissertation.

I would like to thank my family for their support throughout this whole process. Thanks for making Iowa City a second home for the Yocius family the last 10 years. With Maggie, Jackie and Danny also being enrolled at the university during parts of that time, my parents have made countless trips to visit. They were all with me throughout this process. I would also like to thank Brittany Borghi for her support during my final push to finish graduate school. I could not have done this without everyone's help.

## ABSTRACT

Light field photography offers a new approach to digitally captured images. These commercially available cameras are able to capture the 4D light field in a single image. This allows for a variety of image processing capabilities that traditional cameras do not offer. For example, the image can be digitally refocused after it is captured and its depth can be estimated. In terms of application, these capabilities could be beneficial on airborne platforms. However, a limitation of currently available light field cameras is that they are not fully functional at medium or long ranges. If these cameras were to capture light fields at longer ranges, they would have a practical application when mounted on low-flying aircrafts. This dissertation takes current light field photography techniques and modifies them so they work better to capture medium-range images. The majority of cameras that capture the 4D light field use a microlens array to modulate the incoming light before it hits the image sensor. Previous work using printed modulation masks garnered the same effect obtained by microlens arrays. This dissertation details the development of a modulation mask that has medium-range applications. A new way of extracting the 4D light field from raw images that uses a digital Fourier transform is presented. This method works for images captured with microlens arrays and printed mask cameras. Two prototype cameras were built and tested to demonstrate some of these concepts. The concepts demonstrated by these cameras could be used in the future designs of light field cameras.

## PUBLIC ABSTRACT

Light field photography offers a new way of looking at digitally captured images. These commercially available cameras allow for image processing capabilities that traditional cameras do not offer. For example, the image can be digitally refocused after it is captured, and its depth can be estimated. In terms of application, these capabilities could be beneficial on airborne platforms. However, a limitation of currently available light field cameras is that they are not functional at medium or long ranges. This dissertation takes current light field photography techniques and modifies them so that the cameras are better able to capture medium-range images. If these cameras were to capture light fields at longer ranges, they would have a practical application when mounted on low-flying aircraft. Two prototype cameras were built and tested to demonstrate some of these concepts. The concepts demonstrated by these cameras could be used in the future designs of light field cameras.

## TABLE OF CONTENTS

LIST OF TABLES .....	viii
list of figures .....	.ix
INTRODUCTION .....	1
Statement of the Problem.....	1
Background.....	1
Objectives .....	7
AIRBORNE PLATFORMS.....	8
THIN LENS AND STEREOSCOPIC CONFIGURATIONS .....	14
LIGHT FIELD IMAGING BACKGROUND .....	19
CAPTURING THE 4D LIGHT FIELD.....	23
A Microlens Array Light Field Camera.....	25
Printed Mask Light Field Camera.....	30
MANUFACTURING OF A PRINTED MASK.....	35
Printing Considerations .....	37
Pixel Density .....	38
Printing Transparencies.....	39
Printing Black and White Negatives .....	40
Printed Mask Prototypes.....	41
TESTING PRINTED MASKS .....	44
Example Analysis of Printed Mask 1 $p = 2$ $s = 5$ Using Fresnel Diffraction.....	48
Analysis of Printed Mask 2: $p = 2$ $s = 10$ .....	53
Analysis of Printed Mask 3: $p = 4$ $s = 5$ .....	56
Analysis of Printed Mask 4: $p = 10$ $s = 20$ .....	57
Analysis of Printed Mask 5: $p = 5$ $s = 20$ .....	62
Potential of Using Fresnel Diffraction to Design New Mask Patterns.....	63
INSTALLATION OF PRINTED MASK.....	65
PROCESSING CAPTURED LIGHT FIELD IMAGES .....	71
Fourier Domain Light Field Extraction.....	71
Constructing the Light Field Equation from Fourier Slices .....	89
Digital Refocusing.....	90
PROCESSING LIGHT FIELD IMAGES FROM PRINTED MASK LIGHT FIELD CAMERA .....	93

Sampling Theory for Imaging Sensors .....	93
Lens Effects on Printed Mask Modulation .....	94
Mask 2 Processing .....	97
Mask 4 Processing .....	110
Summary of Results .....	119
3D DEPTH ESTIMATION .....	120
Stereo Correspondence .....	122
CONCLUSIONS AND FUTURE WORK .....	127
WORKS CITED .....	131
APPENDIX A: PRINTED MASKS .....	134
Printed Mask $p = 2, s = 5$ .....	135
Negative Printed Mask $p = 2, s = 5$ .....	136
Printed Mask $p = 2, s = 10$ .....	137
Negative Printed Mask $p = 2, s = 10$ .....	138
Printed Mask $p = 4, s = 10$ .....	139
Negative Printed Mask $p = 4, s = 10$ .....	140
Printed Mask $p = 10, s = 20$ .....	141
Negative Printed Mask $p = 10, s = 20$ .....	142
Printed Mask $p = 5, s = 20$ .....	143
Negative Printed Mask $p = 5, s = 20$ .....	144

## LIST OF TABLES

Table 1 This table contains variable descriptions of a single-lens imaging system. ....	17
Table 2 This table shows the pixel density of important image formats. ....	39
Table 3 This table displays prototype printed mask design values. ....	42
Table 4 This table depicts $p$ and $q$ locations, in which $G(p,q)$ is a nonzero for Mask 1. ....	51
Table 5 This table details the expected and actual $x$ and $y$ values of the delta functions for a diffraction pattern produced by Mask 1 at $z$ distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters. ....	52
Table 6 This table details the $p$ and $q$ locations where $G(p,q)$ is nonzero for Mask 2. ....	54
Table 7 This table shows the expected and actual $x$ and $y$ values of delta functions for a diffraction pattern produced by Mask 2 at $z$ distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters. ....	55
Table 8 This table details the $p$ and $q$ locations where $G(p,q)$ is nonzero for Mask 4. ....	59
Table 9 This chart depicts the expected and actual $x$ and $y$ values of the delta functions for a diffraction pattern produced by Mask 4 at $z$ distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters. ....	61
Table 10 This table shows the sensor sampling rate for a prototype light field camera. ....	94
Table 11 This table illustrates the modulation frequency locations for Mask 2. ....	98
Table 12 This table contains the modulation frequency locations for Mask 2 when the images are captured with a focal length of 12 mm. ....	104
Table 13 This tables shows the modulation frequency locations for Mask 2 when images are captured with a focal length of 33 mm. ....	106
Table 14 This table contains the modulation frequency locations for Mask 4. ....	111
Table 15 This table shows the printed Mask 4 modulation frequency locations for images captured with a focal length of 13 mm. ....	115
Table 16 This table contains the modulation frequency locations for images modulated by Mask 4. These images were captured with a focal length of 6 mm. ....	119

## LIST OF FIGURES

<p>Figure 1 This is a model for a traditional camera. A main lens focuses multiple light rays at an image sensor. A pupil restricts which light rays hit the sensor. In a traditional camera, the sensor is able to record only the total of all of the rays hitting an image sensor. A light field camera is able to record each of these rays in a 4 dimensional function. Each ray is defined by the location it enters the main lens and the location it hits the image sensor. Two of the dimensions represent the location on the main lens where the ray passes through. The other two dimensions represent the location where the ray hits the image sensor.....</p>	4
<p>Figure 2 This figure illustrates a coordinate system for imaging systems integrated into an airborne platform with a GPS and IMU. This image is from <i>Evaluating of GPS/IMU Supported Aerial Photogrammetry</i> ©2006 IEEE [21].....</p>	11
<p>Figure 3 This figure depicts a thin lens model.....</p>	15
<p>Figure 4 This figure depicts the geometry of a single-lens imaging system. ....</p>	16
<p>Figure 5 This figure demonstrates the angular plenoptic coordinate system for a human eye.....</p>	20
<p>Figure 6 This figure depicts a single lens and image sensor coordinate system.....</p>	21
<p>Figure 7 The Stanford Multi-Camera Array contains fifty two cameras that are closely packed together. This image is from <i>High-Speed Videography Using a Dense Camera Array</i> ©2004 IEEE [33].....</p>	24
<p>Figure 8 This diagram illustrates a microlens array in a single-lens camera for the capturing of light field images.....</p>	26
<p>Figure 9 Each microlens forms circles on the image sensor. Each circle is produced by its parent microlens directly above it. The content of each circle is a view of the main lens from the perspective of its parent microlens. This mapping is shown for two different microlenses.....</p>	27
<p>Figure 10 This figure is an example of a raw image captured with a light field camera that has a microlens array on top of its image sensor. The beak of the bird is enlarged to show the pattern that the microlens array produces at the level of the pixel. The raw image is from Light Field Toolbox of Matlab[35] .....</p>	28
<p>Figure 11 This figure is a raw light field image of a white background.....</p>	29
<p>Figure 12 This figure depicts the modulation mask architecture.....</p>	30
<p>Figure 13 The left image illustrates an enlarged view of a printed cosine mask with four harmonics. The right image is an enlarged view of a single period of the mask.....</p>	31
<p>Figure 14 This figure details the Matlab function for a printed mask. ....</p>	36



Figure 15 This figure is an example of a mask generated by the Matlab function shown in Figure 14. This mask is 500×500 pixels with a minimum peak distance of 20 pixels. ....	37
Figure 16 This figure shows a negative of a printed mask. This is the negative of the mask shown in Figure 15. ....	41
Figure 17 This is an example of a printed mask on 35mm film. ....	43
Figure 18 The following photograph demonstrates how to test a printed mask with a laser pointer. ....	44
Figure 19 This is an enlarged image of the diffraction pattern. It is produced by coherent visible light passing through the printed mask. ....	45
Figure 20 This figure shows the Mask 1( $p = 2$ $s = 5$ ) test pattern at a distance of 0.5842 m down range. ....	53
Figure 21 This figure shows the test pattern produced by Mask 2( $p = 2$ $s = 10$ ) at a distance of 0.5842 m down range. ....	56
Figure 22 This image shows the Mask 3( $p = 4$ $s = 5$ ) test pattern at a distance of 0.5842 m down range. ....	57
Figure 23 This image is of the Mask 4( $p = 10$ $s = 20$ ) test pattern at a distance of 0.5842 m down range. ....	60
Figure 24 This image is of Mask 5's test pattern ( $p = 5$ $s = 20$ ) at a distance of 0.5842 m down range. ....	63
Figure 25 This figure is comprised of images that show the different steps of the instillation of the printed mask for a Canon EOS Rebel Xsi. Subfigure i) shows a fully assembled camera. Subfigure ii) is of photo of a camera with the back panel removed. Subfigure iii) depicts the image sensor's mounting after the removal of the top-layer circuit board. Subfigure iv) is of photo of the removed image sensor's mounting. Subfigure v) is of depicts the image sensor's mounting with the IR filter removed. Subfigure vi) illustrates the placement of a printed mask onto the image sensor. ....	67
Figure 26 These images depict different steps of the installation of the printed mask into a Fujifilm Finepix S5200. Subfigure i) is a view of the fully assembled camera, ii) is a view of the camera with the back panel removed, iii) is a view with the top-layer circuit board removed, and iv) is a view of the camera with the bottom-layer circuit board removed and the image sensor back plate exposed. Subfigure v) is an image of the camera with the back plate removed and with an elevated image sensor. ....	69
Figure 27 This figure is an image captured from a Lytro camera that has been converted to grayscale. The image on the left is the full image. The image on the right is an enlarged section of the calibration target. ....	73
Figure 28 This figure is a color version of the test image collection with the Lytro camera. The image was color balanced and focused with the Lytro software. ....	74

Figure 29	Analog frequency, angular frequency and sample frequency mapping .....	75
Figure 30	This figure depicts an unshifted DFT of the test image. ....	77
Figure 31	Sampling window for an unshifted 1D DFT signal.....	78
Figure 32	Sampling window for a shifted 1D DFT function.....	79
Figure 33	This figure illustrates a DFT shift function. ....	80
Figure 34	This image shows the DFT of the target image.....	81
Figure 35	This figure shows an enlarged view of the center of a DFT image.....	83
Figure 36	This figure shows each sub-image that is identified in the DFT image. ....	86
Figure 37	This figure depicts spatial representation of light-field slices shown at their modulation frequency locations. ....	88
Figure 38	This image shows the geometry required for refocusing light field images. ....	90
Figure 39	This figure depicts four images captured from a light-field camera and then refocused at different values of $\alpha$ . The raw light field image used to produce the four images was part of the Light Field Toolbox for Matlab[35]. ....	92
Figure 40	This figure shows the mask and main lens geometry to explain the diffraction pattern spreading. ....	95
Figure 41	This figure shows a Laplacian operator of a DFT image. The original image had a focal length of 12 mm. The modulation mask was Mask 2. The slices of the light field are segmented by the red boxes. ....	100
Figure 42	This figure is a DFT of an image that was captured with a focal length of 12 mm. This image was modulated with Mask 2. The slices of the light field are segmented by the red boxes. ....	101
Figure 43	This figure shows the spatial representation of light field slices arranged according to their modulation frequency locations. This is image was captured by the camera with Mask 2 installed. The camera had a focal length of 12 mm. ....	102
Figure 44	This is a raw image from a test camera with Mask 2 installed. The image was captured with a focal length of 13mm. ....	103
Figure 45	This figure illustrates the Laplacian of a DFT image. The original image had a focal length of 33 mm. The modulation mask that was used was Mask 2. The slices of the light field are segmented by the red boxes. ....	107
Figure 46	This figure depicts a DFT of an image that was captured with a focal length of 33 mm. The modulation mask that was used was Mask 2. The slices of the light field are segmented by the red boxes. ....	108

Figure 47 This figure shows the spatial representation of light-field slices arranged according to their modulation frequency locations. This was captured with a focal length of 33 mm. The modulation mask that was installed was Mask 2. ....	109
Figure 48 This figure shows a Laplacian operator of a DFT image. The original image had a focal length of 13 mm. The modulation mask used was Mask 4. The slices of the light field are segmented by the red boxes. ....	112
Figure 49 This is a DFT of an image that was captured with a focal length of 13 mm. This image was modulated with Mask 4. The slices of the light field are segmented by the red boxes. ....	113
Figure 50 This image shows the spatial representation of light-field slices arranged according to their modulation frequency locations. This is an image that was captured with Mask 4. The focal length was 13 mm. ....	114
Figure 51 This image shows a DFT image after a Laplacian operator has been applied. The original image had a focal length of 6 mm and the modulation mask was Mask 4. The slices of the light field are segmented by the red boxes. ....	116
Figure 52 This is a DFT of an image captured with a focal length of 6 mm. The image was modulated with Mask 4. The slices of the light field are segmented by the red boxes. ....	117
Figure 53 This image shows spatial representation of light-field slices arranged according to their at their modulation frequency locations. This image was captured at focal length of 6 mm by a camera with Mask 4 installed. ....	118
Figure 54 This figure illustrates a single slice of a light-field image. The left side of the image is the main lens. The right side of the image is the imaging sensor. This image was taken from <i>Gradient-Based Depth Estimation from 4D Light Fields</i> © 2004 IEEE [14]. ....	122
Figure 55 This is a raw light field image captured with a Lytro camera. This was used for stereo correspondence test. ....	124
Figure 56 This figure shows light-field slices of a test image for a stereo correspondence test. ....	125
Figure 57 This figure depicts a disparity image for each slice of the light field. The disparity is measured between the center slices and the slices located at the center of each disparity image. ....	126
Figure 58 This figure shows considerations that should be taken when designing light-field cameras. ....	128

# INTRODUCTION

## Statement of the Problem

Light field photography is a new way of looking at digitally captured images that traditional cameras do not offer [1]. Images can be digitally refocus after capture and depth within the images can be estimated [2]. The abilities of a low-cost light field camera could be beneficial in a low-flying airborne platform. A single sensor to capture images and estimate depth is beneficial in an environment where weight constraints are a factor and the demands for more up-to-date aerial images are high.

Commercially available light field cameras have limitations that are not desired in airborne platforms. Their benefits work only with short range image capture and typically have a fixed length lens. This limits the environments that the camera can operate in. Before looking at the feasibility of airborne light field photography, current light field camera designs need to be modified for airborne systems. Existing refocusing and depth estimation methods need to be modified to work with new light field camera designs.

## Background

Virtual globe applications, such as Google Earth and NASA World Wind, have allowed more people to access aerial photographs than ever before [3]. These applications offer an intuitive user interface to view and interact with large-format images. Virtual globe applications are able to merge multiple data streams together in ways that a traditional Geographic Information System (GIS) software cannot. These visualization tools handle the sorting, queuing, and presentation of different data streams. This allows end users to create more applications for the data, such as geographical data visualization, navigation, precision agriculture and urban mapping[4]. With this

visualization infrastructure in place, application-specific image acquisition can be plugged into the virtual globe. An airborne light field camera may be able to provide images with properties that are not currently being offered by other types of single-camera imaging platforms.

The United States Geological Survey (USGS) has topological maps, aerial photographs, and satellite images that are publicly available [5]. In addition to other public and private databases, these datasets constitute the GIS, which the virtual globe applications visualize. Researchers at the Georgia Institute of Technology have overlaid real-time video sources, such as security cameras, into Google Earth [6]. The live video feeds are rectified based on camera location and perspective. The live video is then placed on top of traditional GIS images on the virtual globe, producing a blended and rectified image view.

Another technology, wide-area persistent surveillance (WAPS) [7], uses an airborne camera array to capture a dense collection of aerial images. Multiple images of each imaging area are collected. With information relating to the aircraft's position and orientation, imaging areas can be globally referenced. A light field can be constructed from multiple images that are taken from slightly different perspectives of the same area. From this light field, a 3D scene can be extracted. The light field can be used to create a 3D scene in the same way that stereo images can be used to produce a 3D representation. This reconstructed scene can then be placed into virtual globe applications or other 3D visualizers.

In addition to showing high-resolution aerial images, virtual globe applications visualize 3D terrain data in a manner that allows for the perspective and scale to be

determined by users. With the perspective and scale set, users can overlay images located in the large public and private datasets onto the 3D terrain. Though the images are likely to have different collection dates and resolutions, the virtual globe software is able to effectively merge them to create a unified visualization.

The popularity of using aerial images is very large. The global aerial imaging market was valued at \$970 million in 2013 and is projected to grow at a compounded annual growth rate (CAGR) of 13.4 percent from 2014 – 2020 [8]. Likewise, the technology used to capture these images has improved, and the number and variety of platforms has grown. Early platforms included hot-air balloons, kites, pigeons, rockets, and fixed-wing aircraft. These platforms may have been carrying film cameras that were not stabilized, and they were capturing images that could not be geo-referenced [9]. Fortunately, modern systems have grown to include unmanned fixed-wing aircraft, helicopters, and multi-rotor unmanned aircraft systems. Quadrotors are an inexpensive and accessible consumer product. Additionally, digital camera systems are standard, and gyro-stabilized platforms reduce motion blur. Digital formats like GeoTIFF have also become standard. Images in a GeoTIFF file format contain embedded geographical reference information in the digital image. This affords GIS applications the freedom to utilize multiple platforms for image acquisition [10].

A traditional camera contains a main lens that focuses entering rays at the image sensor. These rays represent a collection of light rays that are traveling in the same direction and have the same color. Typically there is some type of pupil, or iris, that prevents some light rays from hitting the image sensor. What actually gets recorded at the sensor is the total power of all of the rays that are hitting that location. This is shown

in Figure 1. A traditional camera is unable to identify each of the rays that is being summed. The goal of all light field cameras is to be able to record the rays themselves within the camera.

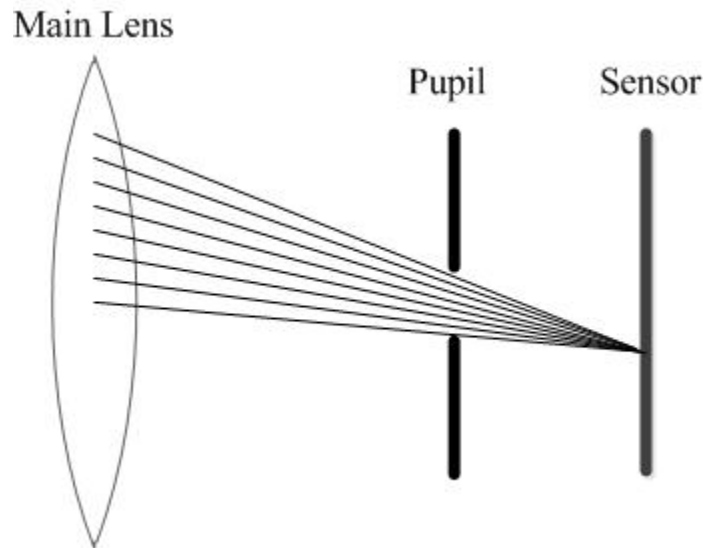


Figure 1 This is a model for a traditional camera. A main lens focuses multiple light rays at an image sensor. A pupil restricts which light rays hit the sensor. In a traditional camera, the sensor is able to record only the total of all of the rays hitting an image sensor. A light field camera is able to record each of these rays in a 4 dimensional function. Each ray is defined by the location it enters the main lens and the location it hits the image sensor. Two of the dimensions represent the location on the main lens where the ray passes through. The other two dimensions represent the location where the ray hits the image sensor.

The popularity of light-field photography has increased in recent years because of the development of hand-held plenoptic cameras [11]. A plenoptic camera is equipped with a microlens array or printed mask that is placed directly over the camera's imaging sensor. This filters the light that passes through the camera. Based on the properties of the main lens and the filtering device, light-field rays can be extracted from

measurements on the imaging sensor. Each ray is described by its 2D intersection with the main lens and its 2D intersection with the image sensor. The two measurements produce a 4D ray. These rays make up the 4D light field, which can then be processed using integral imaging to create a 2D image [12]. Plenoptic cameras are referred to as light field cameras because they capture the light field. For the rest of this paper they will be referred to as light field cameras.

A 4D light field image contains more information than a traditional 2D image. Stereo images can be extracted from a 4D light field image, which allows them to be used for 3D visualizations and other computer vision applications [1]. The complete 4D light field is not needed to create a 2D image. It is possible to make two virtual eyes with a single-lens camera, which allows the user to manipulate the images produced in a manner similar to those produced by a binocular system. Using ray tracing, it is then possible to determine how the light rays within the 4D light field hit each virtual eye. The virtual eye will then produce stereo images. It is possible to digitally refocus a light field image at different depths of field [2]. This is done by changing the focal plane of the image with a Fourier slice transform [2].

The ability to virtually change the focal plane of an imaging system post-capture offers functionality that traditional imaging systems do not. Because the image focus does not need to be selected at capture time, there are fewer requirements for real-time focusing and calibration. Typically the iris of the camera is fixed and the exposure time is set based on the brightness of the scene. Imaging systems that are focused at infinity do not require adaptive focusing, and many airborne systems are focused in this way. A camera that is focused at infinity will produce out-of-focus images when those images are



taken from an altitude lower than the focal distance. Focal distance is a function of image sensor size, pupil size, and focal length of the main lens. For small systems like a DSLR, a typical range for focal distance can be between 0 and 25 feet [13]. A light field camera allows for correction of images that were not properly focused at capture time.

Additionally, different applications are able to change the focal depth based on the user's needs. The focus of the captured image can be dictated by the end user and not by the capturing system.

The depth of objects within an image can be estimated in a 4D light field image [14]. The trajectory of a light ray passing through a camera can be used to interpolate the origin of the beam in reference to the camera. By taking the gradient of the sampled light field, light rays are traced as they pass through the camera. The information relating to the path a light ray takes through the camera allows for the interpolation of the direction and the origin of the ray. A collection of ray origins and the point at which they intersect with the main lens allows the depth of an object to be estimated.

Capturing the 4D light field from a single captured image in an airborne platform is made simpler by one device. That device is the single-lens light field camera. A single-lens camera containing fixed-lens optics simplifies the complexity of hardware common in traditional multi-camera setups used to capture 4D light fields. In addition, light field cameras have the potential to be a low-cost option, especially when compared to traditional systems. Traditional systems typically incur additional costs because the multiple-camera setup requires mountings to hold the cameras rigidly together.

Information contained in the 4D light field can help solve some of the traditional issues that are present in photographic GIS applications. Airborne images collected with

a light field camera could offer low-cost collection solutions for GIS applications. With low-cost collection, GIS applications could collect images of a target area multiple times a year. Multiple image sets of the same area over time may be useful for natural resource-based industries such as agriculture or forestry. The combination of light field images and inertial measurements has the potential to produce a novel architecture for simplifying the creation of overlays in virtual globe software applications. These overlays would be made of images that are focused on the ground plane. They would be able to capture low-resolution 3D terrain data without the use of additional sensors or Digital Terrain Elevation Data (DTED). Comparing the differences between the collected terrain data and the DTED is a potentially useful application of this technology.

### Objectives

The primary objective of this dissertation is to implement a system that can capture light field images in an airborne platform by identifying the limitations of currently available cameras and attempting to correct them in a prototype camera. The new design needs to have a main lens with an adjustable focal length. This focal length could be tuned to the targeted flying altitude of an airborne platform. Existing methods for refocusing and depth estimation need to be validated with the new camera design. If successful, these design concepts could be used to develop small, light-weight and low-cost light field cameras that could be used on a targeted airborne platform. The images produced could be used in end-user applications, such virtual globes.

## AIRBORNE PLATFORMS

The development of an airborne imaging system depends on a variety of requirements. These requirements are based on the flying characteristics of the airborne platform and the characteristics of the application. For the development of an airborne light field camera in this dissertation, the targeted airborne platform is a low-flying vehicle. Low-flying vehicles have lower operating costs than other airborne platforms. The low cost of operations make this technology feasible for a broad range of applications. Creating imaging systems for high-altitude vehicles is a fairly mature field [9]. Additionally, the benefits of the proposed light field camera design will not work when the area being imaged is focused at infinity, which the case for high-altitude imaging systems. The described algorithms for the extraction of the 4D light field will not work without the use of extremely large lenses, which are not practical in that application.

Rotorcraft are good platforms for the proposed airborne light field camera because they are able to fly low, move slowly, and provide a stable infrastructure for imaging systems. Additionally, by choosing a rotorcraft platform, it will not be necessary to develop an algorithm for correcting extreme motion blur. Algorithms for the detection and correction of motion blur for non-stabilized imaging systems have already been well-studied and established [15].

Early prototypes of imaging systems can be tested on specially equipped experimental helicopters, but one goal of this camera is to create a system that can be scaled down to operate on an unmanned aerial system (UAS) such as a quadrotor. A small UAS offers a low-cost, low-maintenance, and easy-to-control aerial platform [16].

If the proposed camera design had the ability to refocus images, estimate depths, and create 3D scenes, it would be worth scaling down for use on small, unmanned systems. The low cost of a quadrotor system makes it more accessible for civilians. A low-cost airborne platform and imaging system with these capabilities would be useful in private, small-scale GIS applications such as agriculture, law enforcement, and private surveying. If the imaging system was able to perform in real-time, it would be useful for the visual navigation of airborne platforms. Though there are several advantages to a UAS, it is currently more practical to deploy a manned rotorcraft for many users. Operation of an unmanned asset for commercial application is not currently legal in many states and research applications require a Certificate of Authorization (COA). For this reason, a manned helicopter can act as a surrogate for a UAS until the FAA has finalized the small UAS rule [17].

Airborne images become more useful once they are georeferenced.

Georeferencing is the process of finding geographical reference points within an image. With georeferencing, airborne and ground images can be scaled, rotated, and projected. This allows them to be referenced with and compared to other images taken in the same area [18]. Georeferencing is required for images in virtual globe applications. With the low cost of global positioning systems (GPS) and digital imaging sensors, the amount of georeferenced images is steadily increasing [19]. Because these databases contain georeferenced images from the same locations with the same perspectives and orientations, differences in the imaging areas can then be quickly observed.

One way to find a reference location for an aircraft in relation to a global coordinate system is to use a GPS and an inertial navigation system (INS) that employ

inertial measurement units (IMU). This allows for positions to be estimated on the basis of accelerations and angular rates. For applications that involve projecting images into georeferenced planes, global Cartesian coordinate systems, such as the Earth-Centered, Earth-Fixed (ECEF) coordinate system, are convenient. The ECEF system is a Cartesian coordinate system for which the origin of the axes are at the center of the earth. The z-axis points north and the x-axis intersects with the equator at zero degrees longitude. Measurements along these axes are in meters. Approximations for conversions back and forth from the ECEF system to geodetic coordinate systems exist and are fairly accurate [20]. At different points, both coordinate systems are convenient and it is useful to be able to switch between them when necessary.

A GPS and IMU system can return ECEF coordinates. Additionally, a GPS and IMU system can return either the rotational values of the GPS's antenna or the calculated centroid of the aircraft. Using this coordinate system, the ECEF coordinate and rotational values for the camera can be determined. Figure 2 depicts an aircraft with the coordinate system for a GPS antenna and an image sensor. The imaging sensor's coordinate system is centered at the camera's perspective origin.

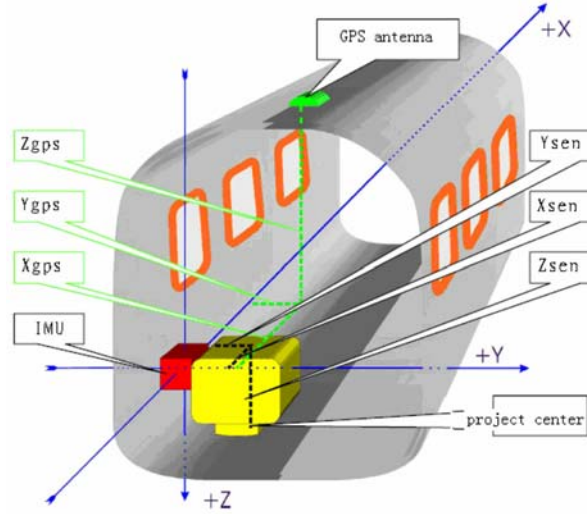


Figure 2 This figure illustrates a coordinate system for imaging systems integrated into an airborne platform with a GPS and IMU. This image is from *Evaluating of GPS/IMU Supported Aerial Photogrammetry* ©2006 IEEE [21].

The transform between these coordinate systems is adding the offsets between the origins of the two coordinate systems and applying a rotation matrix [21]. Equation 1 shows one way of applying this transform, where  $(X_A, Y_A, Z_A)$  is the ECEF coordinate calculated by the GPS and IMU and  $R$  is a rotation matrix to rotate the axis into the same orientation as the imaging sensor. After the rotation, the ECEF value is shifted linearly to the center of the image sensor by offset  $(X_s, Y_s, Z_s)$ . This offset is the difference between the GPS and IMU coordinate origin and the coordinate origin of the imaging sensor. The resulting value is an ECEF value of the camera's perspective origin.

$$\begin{bmatrix} X_I \\ Y_I \\ Z_I \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} + R \begin{bmatrix} X_A \\ Y_A \\ Z_A \end{bmatrix}$$

1

With the ECEF coordinate and the rotations of the imaging sensor's center, it is possible to reproject the image. This enables the image to be georeferenced. Once

georeferenced, the image can be visualized in a virtual globe or in different GIS applications. Because many aerial images are captured without a direct overhead perspective, they are prone to lens distortion and need to be reprojected to be placed on the same plane as a Nadar map. This reprojection can be performed by using a projective transform on the image [22]. The projective transform is a  $3 \times 3$  matrix that maps pixels from the captured image to locations on a flat map plane. The coefficients of this transform can be approximated by the geometry of the inertial measurements provided by the GPS and the IMU.

A GPS and an IMU can only be used to approximate the coefficients of the projective transform matrix because a GPS and an IMU are not perfectly precise. However, finding matching ground control points between pairs of images that cover the same area can further refine the approximations of the coefficients [23]. Matching ground control points can be manually tagged or detected by an automated process. The shift invariant feature transform (SIFT) has been shown to be robust in image stitching [24]. Examples of this include the matching of ground control points. Generally, an image will produce many SIFT features that can be used to match control points. However, SIFT features can produce some points that are false positives. Using algorithms, like the Random Sample Consensus (RANSAC), coefficients for the projective transform matrix can be found by regressing through large numbers of possible matching ground control points [25].

For a robust solution, a GPS and an IMU can be used to acquire approximations for the coefficients of the projective transform matrix. Additionally, an algorithm can be used to match ground control points to further refine the solution. A higher-dimensional

SIFT transform may be able to identify features in a 4D light field image that can be used for making these approximations. There might be other features contained in a 4D light field image that can be leveraged for robust image stitching.

The above section outlines aspects of the targeted airborne platform for a light field camera. This type of imaging system is targeting low-flying, stable aircraft that are able to provide inertial measurements to be used for georeferencing. Additional image stitching refinements can be delivered from this platform. With this desired platform specified, the next section will explain techniques for capturing light field images.



## THIN LENS AND STEREOSCOPIIC CONFIGURATIONS

Some knowledge of basic optical systems is necessary to understand how light field cameras work, including an understanding of thin lens and elementary optics. Depth estimation also requires knowledge of thin-lens equations and trigonometric identities. Stereoscopic imaging properties of light field cameras will be covered. These fields are relevant to 3D imaging and depth estimation.

A thin lens is defined as a lens with a negligible thickness across the optical axis, as compared to the focal length. If a light ray enters a thin lens at a coordinate  $(x,y)$  on a lens surface, it will leave the lens approximately at the same coordinate on the other side [26]. With this assumption in place, it is possible to find the focal length using the lens maker's equation. The lens maker's equation is shown in Equation 2 [27]. In this equation,  $R_1$  and  $R_2$  are the radiuses of the curvatures of the two lens fronts and  $n$  is the reflective index of the lens.

$$\frac{1}{F} = (n - 1) \left( \frac{1}{R_1} + \frac{1}{R_2} \right)$$

2

With the focal length of the lens defined by the lens maker's equation, it is possible look at a single-lens configuration, as shown in Figure 3. The lens has a focal length of  $F$ . An object of interest is on the left side of the lens, and a real image of that object is on the right side of the lens. The real image is an inverted image of the original object. The location of the real image is determined by Equation 3. The focal length of the lens and the distance between the object and the lens determines where the real image is located.

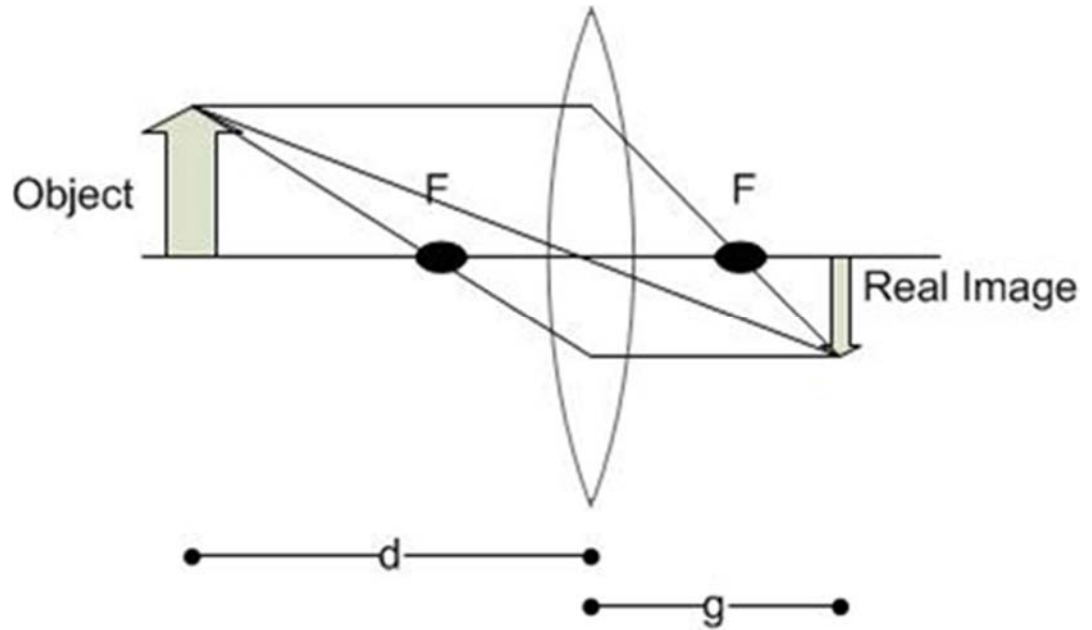


Figure 3 This figure depicts a thin lens model.

$$\frac{1}{F} = \frac{1}{d} + \frac{1}{g}$$

3

The thin lens model does not include a fixed imaging sensor. Figure 4 is a thin lens model with a fixed imaging sensor plane included, adding several distances of interest to the model. Table 1 defines these distances. In the setup, there is an object located at a distance of  $d$  from the main lens. This object creates a real image at a distance of  $g$  from the lens. The sensor plane is placed in front of the real image at a distance of  $f$  from the lens. With the sensor plane not located at the real image, light rays that normally converge at the real image will create a displacement on the sensor plane at a height of  $h$ . These same light rays will create a displacement on the lens at a height of  $v$ .

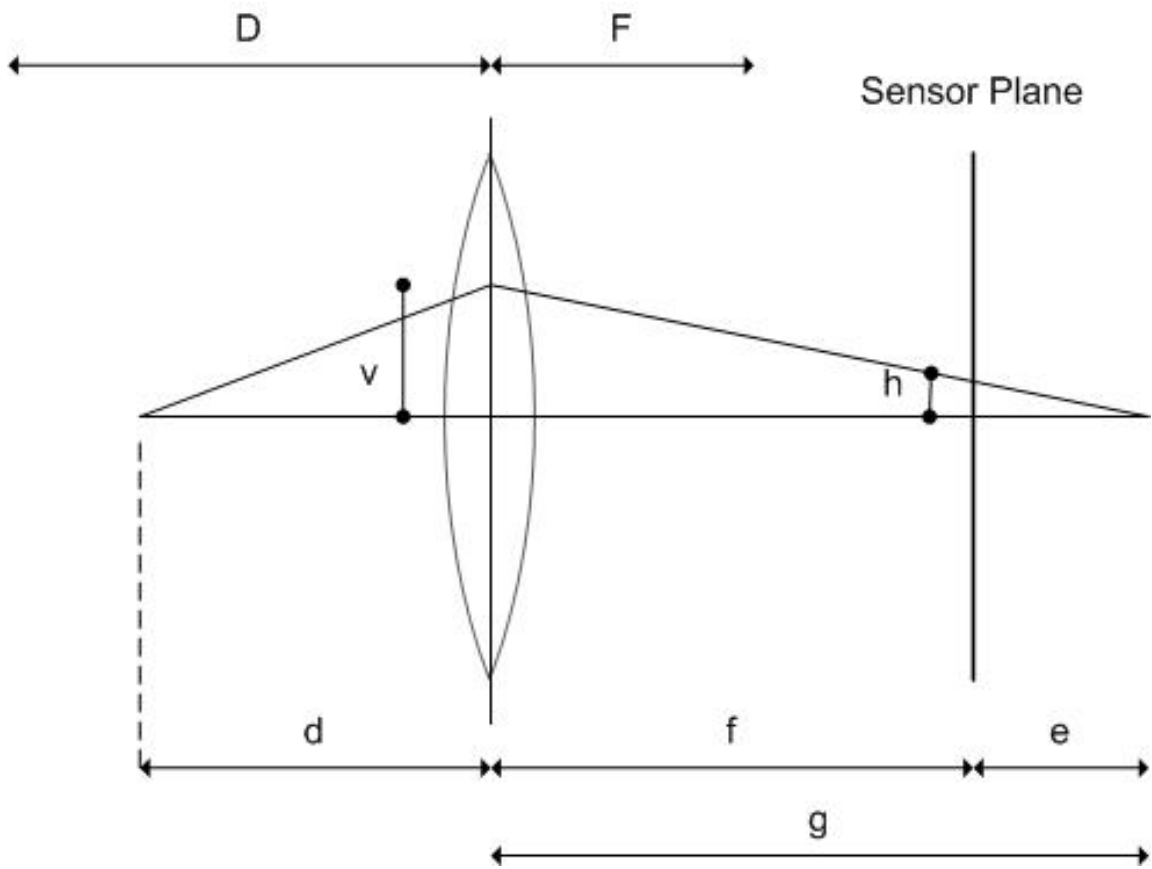


Figure 4 This figure depicts the geometry of a single-lens imaging system.

Table 1 This table contains variable descriptions of a single-lens imaging system.

<b>Variable</b>	<b>Description</b>
$F$	Focal length of lens
$f$	Distance from lens to sensor plane
$D$	Distance from sensor plane to conjugate plane
$d$	Distance from object to lens
$g$	Distance from conjugate focus to object
$e$	Distance from conjugate focal point to sensor plane
$v$	Displacement of aperture
$h$	Displacement of object on sensor

Knowing the location of the real image is beneficial because it is the point at which the image is in focus. That point is called the conjugate focus of the object. When the displacement heights of  $h$  and  $v$  are known, similar triangles can be used to find the location of the virtual image [1]. Equation 4 uses similar triangles and the two displacement heights to determine the distance between the conjugate focus and the object ( $g$ ).

$$\frac{1}{g} = \frac{1}{f} \left(1 - \frac{h}{v}\right)$$

4

By substituting Equation 3 into Equation 4, it is possible to find the distance to the original object [1]. The substitution is shown in Equation 5, in which  $d$  is the distance to the original object. This equation will be used later for depth estimations with a single-lens camera. Normally, traditional cameras are unable to measure the displacement distances of  $h$  and  $v$ . However, this is possible with a light field camera.

$$\frac{1}{d} = \frac{1}{F} - \frac{1}{f} \left(1 - \frac{h}{v}\right)$$

5

The assumptions that can be made from thin lens models can be transferred and applied to many aspects of light field camera design. Thick lenses or complex main lens setups can introduce spherical aberrations and other undesired distortions. There are well-established computer vision techniques for correcting distortions with the use of calibration images [28]. Simple camera models and calibration images make it possible to detect distortions and correct them in images collected in the future [29]. When spherical aberrations or other distortions become an issue with light field images, several of these established techniques can be used to correct the images in a manner that allows them to fit into an established camera model.

## LIGHT FIELD IMAGING BACKGROUND

Imaging systems, such as traditional cameras, attempt to approximate the optics of the human eye. Film and digital sensors approximate some functions of the retina. The pupil restricts light entering the eye similar to a camera iris restricting light entering the camera. Light hitting the retina is processed to extract motion, color, orientation, binocular disparity and other visual effects. Adelson and Bergen expressed the light field entering the eye as a multivariable function [30]. From this multivariable representation, some of the processing completed by the retina and the brain can be derived. A light field camera attempts to mimic this multivariable representation.

Figure 5 shows a coordinate system for an approximated human eye. Any point on the retina can be expressed in terms of its 3D Cartesian coordinate  $(V_x, V_y, V_z)$ . Light rays entering the pupil are expressed by angles  $(\theta, \phi)$  with respect to an optical axis. The optical axis is parallel to the  $V_z$  axis and intersects the imaging plane at the selected location  $(V_x, V_y)$ . With this notation, any light ray entering the eye that hits the retina can be expressed.

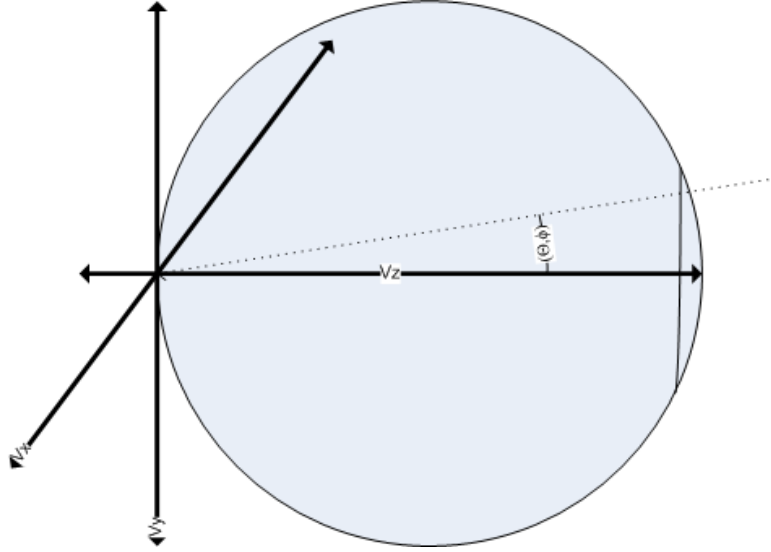


Figure 5 This figure demonstrates the angular plenoptic coordinate system for a human eye.

One form of the plenoptic function is expressed in Equation 6. All of the rays of the light field can be expressed by the wavelength of the ray ( $\lambda$ ), the location at which the ray intersects the imaging plane ( $V_x, V_y, V_z$ ), the angles off of the optical axis ( $\theta, \phi$ ) and the current time  $t$ . This seven-dimensional function can be difficult to use, but it does fully express the light entering the human eye.

$$P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$$

6

For the purpose of a light-field camera, the plenoptic function can be reduced to a 4D function.. The imaging plane is a 2D plane with discrete pixels ( $x,y$ ). For a single lens camera, such as a pinhole camera, all light rays can be expressed by the intersection with the imaging plane and the intersection with the main lens ( $u,v$ ). For camera

geometries that are more complicated, a reduction into this form may be needed.

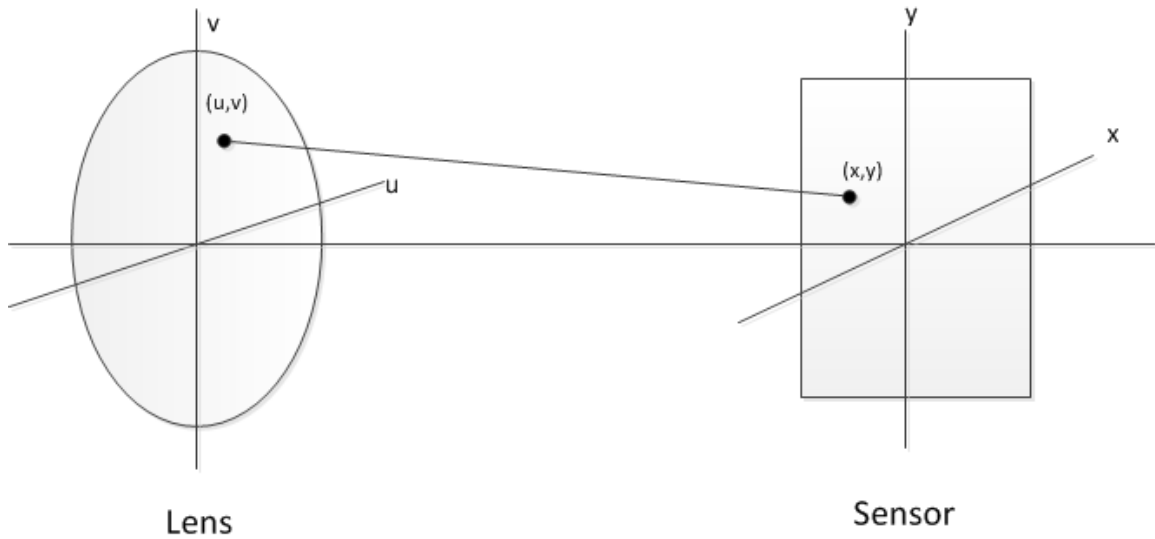


Figure 6 This figure depicts a single lens and image sensor coordinate system.

Equation 6 describes a light field as it enters the human eye. However, it is an over-specification for the functionality of a traditional camera. Most traditional digital imaging sensors, such as a CCD focal plane sensor, are not able to pick up individual frequencies of light. Rather, they work by recording the power within each band of color. Additionally, because each captured image is a discrete moment in time, the time variable can be removed. Using these two factors, Levin reduced the light field into a 4D space for camera applications [31]. The 4D light field can be expressed in the form shown in Equation 7.

$$L = L(x, y, u, v) \tag{7}$$

The 4D light field has also been called the Lumigraph. The Lumigraph has been used in computer graphics and real images. Gortler discussed the construction of a



Lumigraph, the use of it in 3D imaging, and the use of it to reconstruct 2D images with different perspectives [32]. It has been used in some computer vision applications that are creating a 3D virtual environment. Equation 8 expresses a method for reconstructing a standard 2D image from the 4D light field.

$$E(x, y) = \iint L(x, y, u, v) du dv$$

8

## CAPTURING THE 4D LIGHT FIELD

Calculations of the 4D light field have many uses, including the creation of computer graphics and the display of 3D graphics with moving perspectives. There are a number of different approaches to acquiring a 4D light field image. The light field can be produced from computer renderings or from a collection of digital images [12]. There are a few ways of collecting images that be used to record the light field. Some involve multiple cameras, others look at a video from a single camera, and some use diffractive optics[2]. Large arrays of synchronized cameras directed at the same area can capture one form of the light field. An example of this technique is the Stanford Multi-Camera Array, which consists of fifty two cameras packed closely together and is able to approximate a single center of projection [33]. The array is shown in Figure 7. This type of array captures multiple images of the same imaging area and each camera captures a portion in the light field.

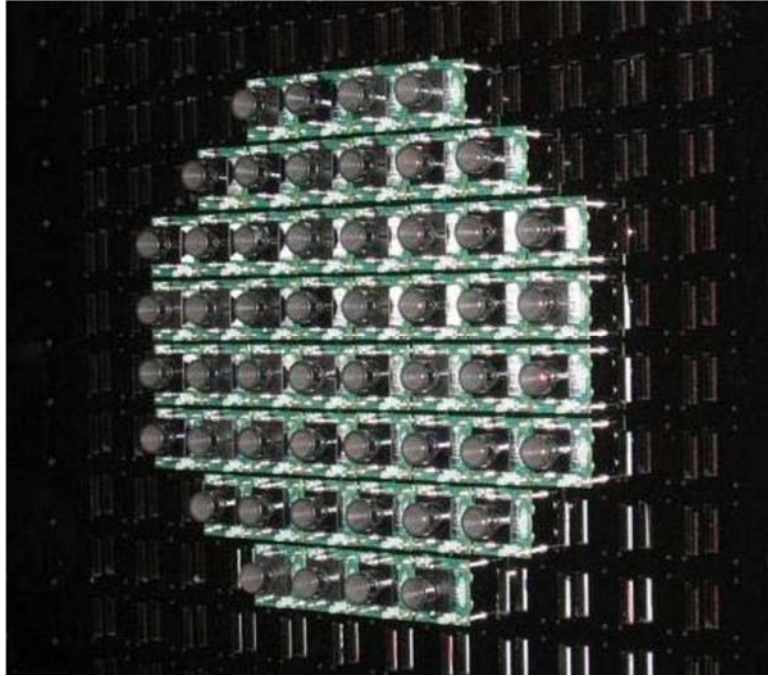


Figure 7 The Stanford Multi-Camera Array contains fifty two cameras that are closely packed together. This image is from *High-Speed Videography Using a Dense Camera Array* ©2004 IEEE [33].

Regarding the operational feasibility of lightweight platforms like micro UAVs, weight of imaging systems is one of the biggest concerns when it comes to feasibility of operation. Another issue with camera arrays is that a global knowledge of each camera's position and orientation in relationship to one another is important for calibrations. In high motion platforms that may have constant vibration and acceleration forces being applied, keeping constant position and orientation between cameras might not be feasible. For this reason, the remainder of this section is focused on single camera systems that can be used for light field capturing.

## A Microlens Array Light Field Camera

One way to record the light field with a single imaging sensor is to use a microlens array. Some of the basic principles of this camera design were outlined by Edward H. Adelson's paper on the subject, which is titled *Single Lens Stereo with a Plenoptic Camera* [1]. Refinements to this research were detailed by Ren Ng in his dissertation on the subject [2]. Ren Ng commercialized this camera design with his company Lytro [34]. Lytro's camera is currently the most popular, commercially available model for light field photography. These cameras are able to capture the 4D light field, but do not record images that are at the same resolution of the image sensor. Each pixel is used to record a light ray, and multiple light rays are required to form a pixel that is produced by a traditional camera. This causes the final image recorded with a light field camera to have a smaller resolution than a traditional camera with the same image sensor.

The basic configuration of a microlens array light field camera is shown in Figure 8. The microlens array is placed on top of the imaging sensor. By doing this, the following two properties can be implied:

- 1) All light that passes through a pixel must pass through its parent microlens.
- 2) All light that passes through a pixel must pass through the pixel's conjugate square on the main lens [11].

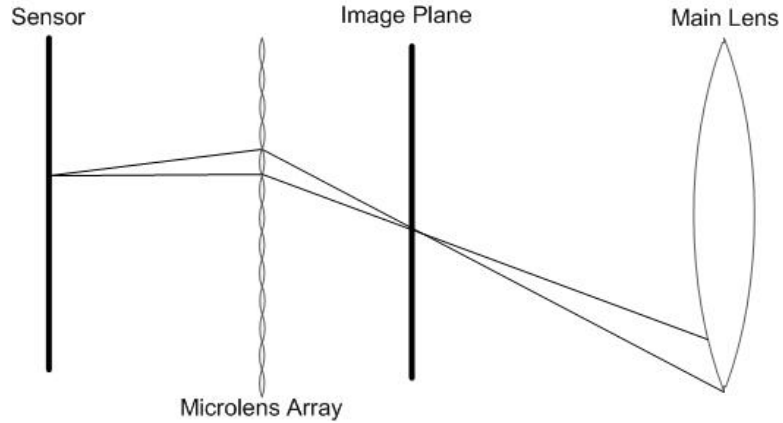


Figure 8 This diagram illustrates a microlens array in a single-lens camera for the capturing of light field images.

These properties cause circles to form on the image sensor. Each circle can be thought of as a very small camera that is taking a photo of the entire main lens. Each of the circles has a slightly different perspective looking at the main lens. This property is shown in Figure 9. The result is similar to having multiple cameras pointed at the same area but on a much smaller scale. Similarly the main lens of a traditional camera, the microlenses' optical properties affect how each circle is formed. Having the microlens placed in a location so that the whole main lens falls into the field of view of the microlens is very important. Standard lens effects, like spherical aberrations, determine the shape, coloring and brightness of each of these circles.

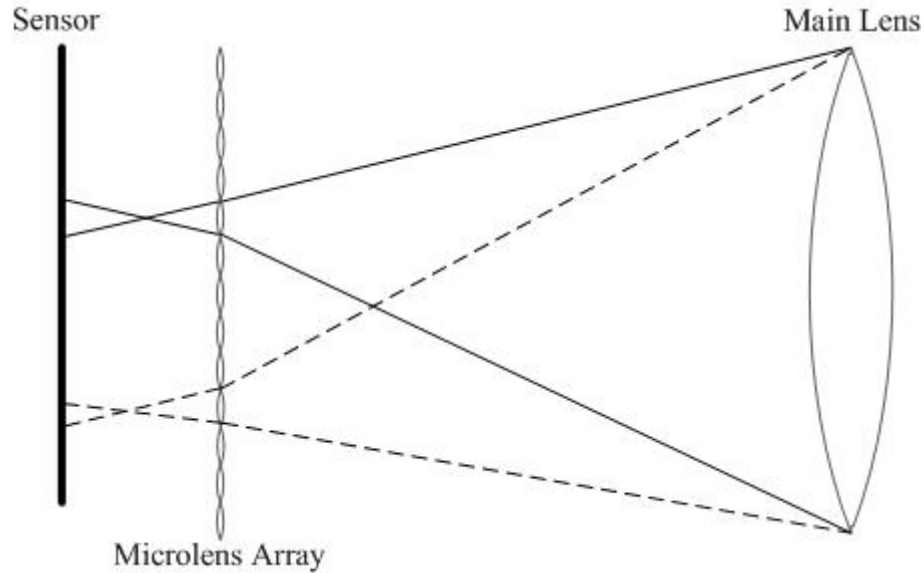


Figure 9 Each microlens forms circles on the image sensor. Each circle is produced by its parent microlens directly above it. The content of each circle is a view of the main lens from the perspective of its parent microlens. This mapping is shown for two different microlenses.

These properties assume the ideal microlens and pixel grid alignment. A raw image captured by a microlens light field camera is similar in appearance to an image captured by a standard digital single-lens reflex (DSLR) camera. A sample of a raw image is shown in Figure 10. The original image is one of the sample images in the Matlab Light Field Toolbox [35]. When portions of the image are enlarged, the effects of the microlens become more apparent.

Each microlens forms a circular image on the sensor, and all of the pixels within each circle are considered a sub-image. Therefore, all light that hits a pixel within a sub-image passed through the same microlens. The location of each microlens within the array provides the value  $(x,y)$  in the light field equation described in Equation 7. The microlens array acts as a virtual version of the image sensor shown in Figure 6. Each

microlens is considered a pixel within the virtual sensor. This can all be inferred from the first property of the microlens light field camera. Each pixel in a sub-image corresponds to a given sub-aperture of the main lens. The location of the sub-aperture within the main lens provides the location  $(u,v)$  in the light field equation.

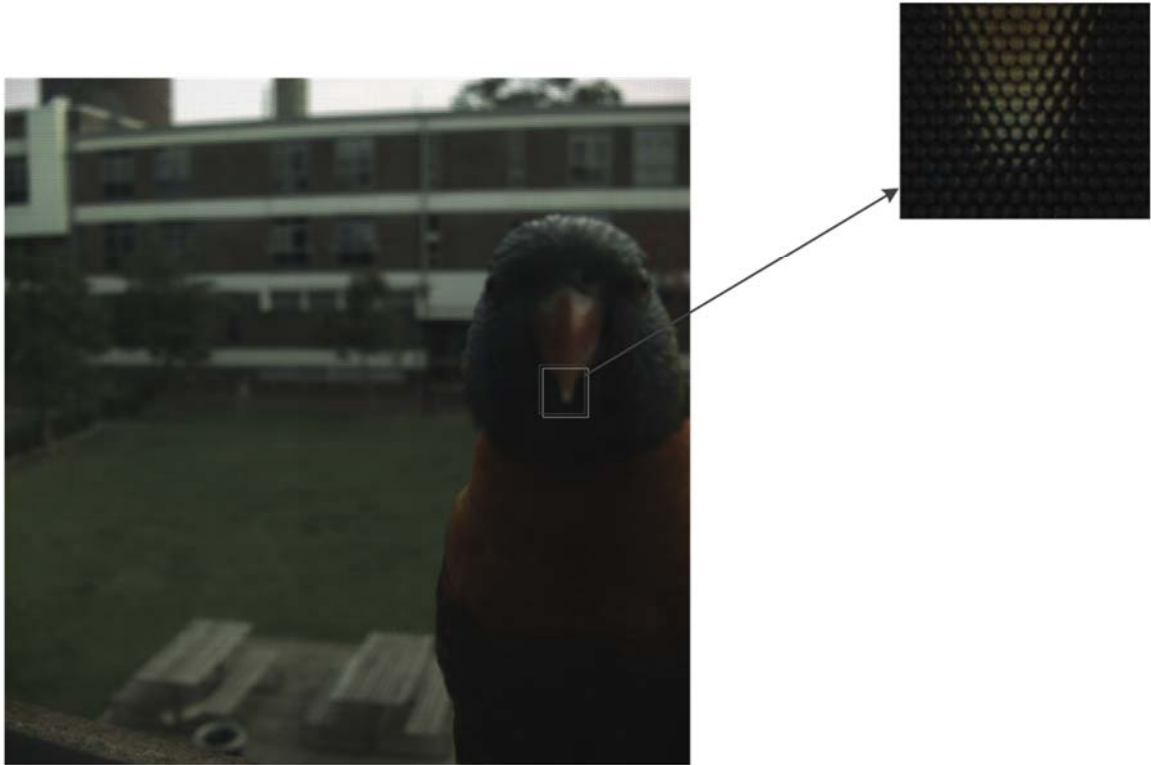


Figure 10 This figure is an example of a raw image captured with a light field camera that has a microlens array on top of its image sensor. The beak of the bird is enlarged to show the pattern that the microlens array produces at the level of the pixel. The raw image is from Light Field Toolbox of Matlab[35]

Given the properties explained above, each pixel in the imaging sensor can be mapped into the 4D light field equation. One way of determining this mapping is to calibrate it with an image of a plain white background. An example of an image calibrated with a plane white background is shown in Figure 11. At the micro level, one

can see that each microlens produces a circle on the image sensor. Each pixel within that circle is considered a sub-image. Circle detection can be used to calculate the center of each sub-image. As a result, the average pixel size of each sub-image can be determined. One way to determine these values is described in *Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras* [35]. As long as the relationship between the microlens array and the imaging sensor remains constant, the calibration process for a camera only needs to be completed once.

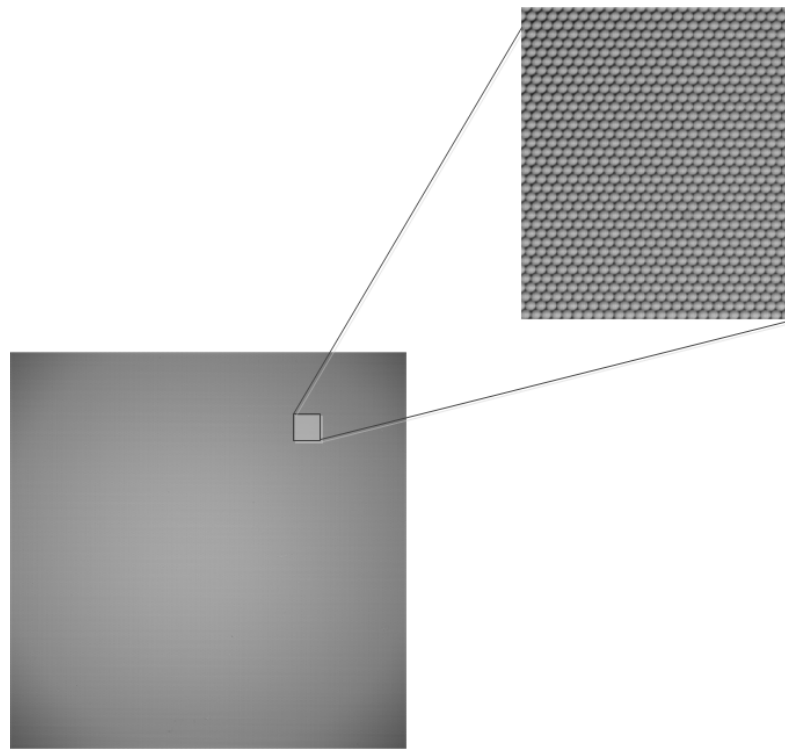


Figure 11 This figure is a raw light field image of a white background.

An additional method for the extraction of the 4D light field is the Fourier transform of raw images. The microlens array has a distinct optical effect in the frequency domain. This can be used to slice the image apart. The implementation of this method is described in the section titled “Capturing the 4D Light Field.” This method is



helpful when the size of the sub-image is unknown. When testing new camera prototypes, this proved to be a useful approach to raw images.

### Printed Mask Light Field Camera

The use of a microlens array is not the only means of capturing a light field image with a single-lens camera. An additional approach uses a printed film mask instead of a microlens array. A printed mask is much less expensive to produce than a microlens array. The printed mask offers more flexibility when it comes to choosing optical properties. An example of this type of hardware configuration is shown in Figure 12. A printed mask is placed directly in front of the imaging sensor. The rest of the camera is left unaltered. The printed mask can modulate and filter the light field before it hits the imaging sensor. Microlens arrays and printed masks serve similar functions. They both have known optical properties and those properties can be used to capture the 4D light field.

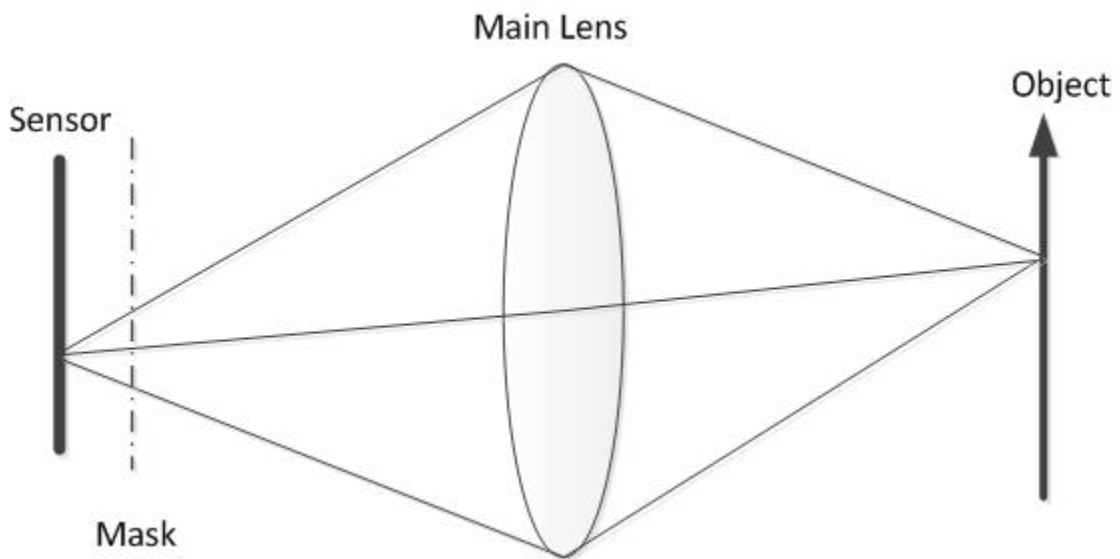


Figure 12 This figure depicts the modulation mask architecture.

Veeraraghavan showed that coded apertures on a printed mask can be used to reconstruct the 4D light field [36]. The coded mask modulates the light field coming into the camera with cosine functions that have different modulation frequencies. Figure 13 shows an enlarged view of a printed cosine mask with cosines at four different frequencies. As described by basic modulation theory, modulating with cosines will result in copies of the original signal. These copies can then be used to reconstruct the 4D light field.

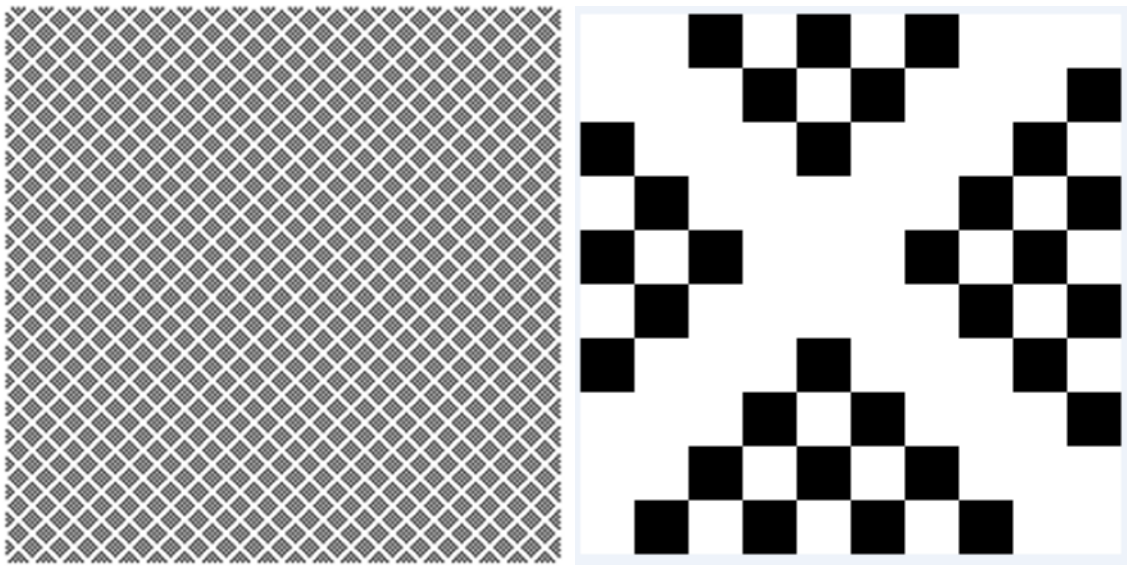


Figure 13 The left image illustrates an enlarged view of a printed cosine mask with four harmonics. The right image is an enlarged view of a single period of the mask.

According to the modulation theorem, a baseband signal multiplied by a cosine function in the time domain will produce copies in the frequency domain [37]. In Equation 9,  $s(x)$  is a one dimensional spatial domain function with the Fourier transform  $S(f_x)$ . The variable  $x$  can represent a location or a moment in time. By multiplying the original signal  $s(x)$  by a cosine function at frequency  $f_0$ , the resulting signal produces shifted copies that are centered at frequencies of positive and negative  $f_0$ . The Fourier

transform of a cosine wave is two delta functions that are offset at a negative and positive frequency of the cosine wave. By convoluting the two delta functions in the frequency domain, copies of the original frequency domain signal are made and centered on positive and negative  $f_0$ .

$$\mathcal{F}[\cos(2\pi f_0 x) s(x)](f_x) = \frac{1}{2}(S(f_x - f_0) + S(f_x + f_0))$$

9

The multiplication of the original signal with multiple cosines added together will produce copies at multiple frequencies at the same time in the frequency domain. Additionally, the multiple cosines can be in two different dimensions. Equation 10 depicts a desired frequency response for a 2D modulation mask that has multiple copies in both  $x$  and  $y$  directions. In this example, the mask is for one sub-image. This is equivalent to the sub-images produced by a single microlens. If the original image were bandlimited to half of  $f_0$ , this modulation would produce alias-free copies. The number of copies produced by this mask can be determined by  $4 * p_1 * p_2$ . The selection of  $f_0$ ,  $p_1$ , and  $p_2$  can be easily changed with printed masks. Microlens arrays do not possess this level of flexibility.

$$C(f_x, f_y) = \sum_{k_1=-p_1}^{k_1=p_1} \sum_{k_2=-p_2}^{k_2=p_2} \delta(f_x - k_1 f_0^x, f_y - k_2 f_0^y)$$

10

The printed mask shown in Figure 12 needs to be applied in the spatial domain and not in the frequency domain. The Fourier transform of Equation 10 produces a series of cosines. Because the image sensor cannot detect negative values, the entire transform needs to be offset. Equation 11 shows this transform when it is offset. When this mask is

applied directly over an array of pixels, it will produce a single sub-image. When copies of the mask are placed next to each other to cover the whole image sensor, the mask acts like a complete microlens array.

$$c(x, y) = 1 + \sum_{k_x=0}^{k_x=p_1} \sum_{k_y=0}^{k_y=p_2} \cos(2\pi f_0 k_x x) \cos(2\pi f_0 k_y y)$$

11

Reconstructing the 4D light field from an image captured with a printed mask is straightforward. However, it does require a multidimensional digital Fourier transform (DFT) and an inverse digital Fourier transform (IDFT). Because the imaging sensor takes discrete samples, a DFT and an IDFT must be used instead of their analog equivalents. First, a standard 2D DFT of the captured image is taken. The DFT will have copies of a standard 2D image modulated in the frequency domain. These copies will be centered on the delta functions described by Equation 10. Each delta function corresponds to a sub-aperture on the main lens. In this mapping, the variables  $k_1$  and  $k_2$  are the frequency transformation of the axes  $u$  and  $v$  in the light field equation. These modulated copies of the base image correspond to slices of the 4D light field. In the frequency domain, these slices can be mapped linearly into the DFT of the 4D light field equation. This mapping is shown in Equation 12.

$$\mathcal{LF}(f_x, f_y, k_1, k_2) = I(f_x - k_1 f_0^x, f_y - k_2 f_0^y)$$

12

In Equation 7,  $I$  is the 2D DFT of the captured image and  $LF$  is the DFT of the light-field equation. The valid range for  $k_1$  and  $k_2$  is the same as in Equation 11. Additionally,  $f_x$  and  $f_y$  are bandlimited by half of the value of  $f_0$ . This change in variables enables the 4D DFT of the 4D light field to be constructed from the 2D DFT of the

captured image. To return the 4D light field in the form expressed in Equation 7, apply a 4D IDFT to the results from Equation 12. As shown in later sections, the 4D light field is useful in both the spatial domain and the frequency domain. The intermediate results from calculating the DFT and IDFT of the light field can be saved for use in multiple applications.

## MANUFACTURING OF A PRINTED MASK

There are specific requirements for the manufacturing of printed masks. For these printed masks to work, they need to have a high number of printed pixels within a small printed area. Dots per inch (DPI) is the standard unit of measurement for the printed pixel density of printed images. A standard laser printer can print at  $600 \times 600$  DPI. However, this is not a high enough spatial resolution for this application. The Canon Rebel that was used for the prototype had roughly  $3450 \times 3450$  pixels per inch (PPI) [38]. For the modulation mask to be effective, the printed mask needs to have a DPI that is greater than  $3450 \times 3450$ .

A low-cost solution to printing a high-DPI film mask takes advantage of digital to film technology. There are certain applications for which film images are preferred over digital copies. Film and slide transparencies are often used for archival storage because they have a longer shelf life than digital media. Slide projectors that create very large projections, such as those used in planetariums, are cheaper than their digital equivalents. Large image formats, such as the 16k image, require multiple projectors to attain a full resolution projection. However, when a 16k image is printed onto slide film, it can be projected by a single projector and shown at full resolution. This type of film printing was available at the University of Iowa's Photo Services Department. Though this service is no longer available at the University of Iowa, there are a few online providers that print slides. One such provider is [gammatech.com](http://gammatech.com). This type of mask printing allowed for the production of inexpensive and quickly-produced prototypes.

The type of mask described by Equation 12 needs to be expressed so that it can be rendered into a printable image format. A Matlab implementation of the equation is shown in Figure 14. The variables  $h$  and  $w$  are the height and width of the mask image. The variable  $s$  is the width in pixels of one period of the lowest harmonic cosine in the mask. The variable  $p$  is the total number of harmonics in the mask. Each harmonic has a

period that is half the length of the next lowest harmonic. First, a blank image is created and each pixel is set to zero. Two nested for loops cycle over each pixel and sum the values of each cosine harmonic. Equation 11 added a constant to ensure that the mask never produced a negative value. In addition to adding an offset, this approach also scales the mask between a range of 0–255.

```
1 function mask = makemask( h, w, s, p )
2
3
4     mask = zeros(h,w);
5     fs = (2*pi) / s;
6
7
8
9     for i = 1:h
10        for j = 1:w
11
12            mask(i,j) = 4 + (cos(fs*i)) * cos(fs*j);
13            for q = 2:p
14                mask(i,j) = mask(i,j) + cos(fs*q*i) * cos(fs*q*j);
15            end
16
17        end
18    end
19
20    low = min(min(mask));
21    mask = mask - low;
22
23    high = max(max(mask));
24    mask = mask* (256/ high);
25
26 end
```

Figure 14 This figure details the Matlab function for a printed mask.

An example output of this function is shown in Figure 15. This figure shows a very low-resolution image. This image is only 500×500 pixels with an  $s$  value of 10. This pattern can be repeated to increase the scale for a larger image format. However, this pattern is difficult to see with the human eye when it is displayed as a large, full-

resolution image such as a 16k image. Figure 15 is an enlarged image of the pattern that can be printed onto film.

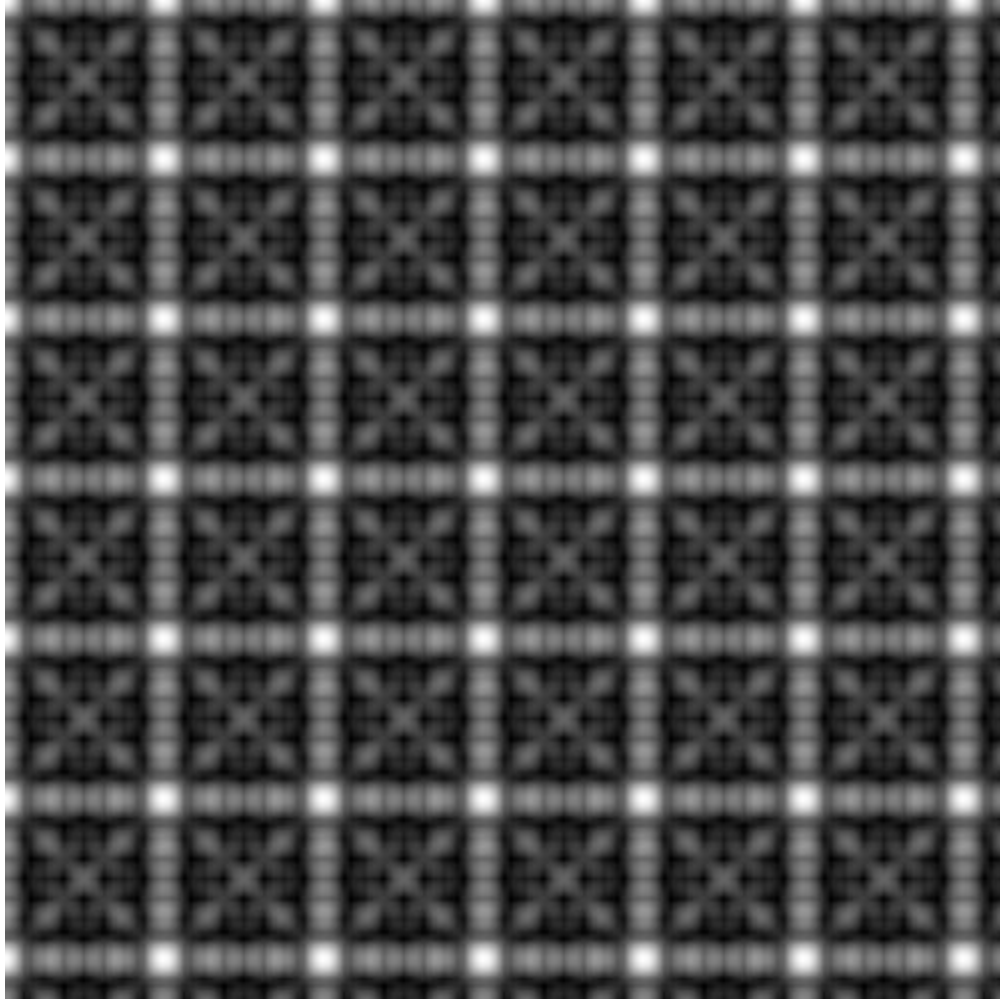


Figure 15 This figure is an example of a mask generated by the Matlab function shown in Figure 14. This mask is  $500 \times 500$  pixels with a minimum peak distance of 20 pixels.

### Printing Considerations

There are a couple of considerations that need to be taken into account when printing the masks onto film. The number of companies that develop 35mm film is decreasing because of a corresponding drop in the use of film cameras. During the



duration of this project, the University of Iowa stopped offering its digital to film service. As a result, the fabrication of the masks was delegated to online retailers. The following subsections describe important considerations for the ordering and production of printed masks.

### Pixel Density

Pixel density is the density of pixels in a given area. In this case, the area can be thought of as either the camera's image sensor or the printed mask. Table 2 shows pixel density for different formats. The camera that was used to produce this prototype does not have a full 35mm imaging sensor. The image sensor has  $3072 \times 2048$  pixels on a  $22.7\text{mm} \times 15.1\text{mm}$  chip. The first row of Table 2 shows the image sensor of the prototype camera. The sensor has a pixel density of 135.3 pixels/mm horizontally and 125.3 pixels/mm vertically. However, the digital to film process can handle much higher resolutions. Typically, in digital to film processing, images are printed in 4k, 8k, and 16k formats. These format sizes are also in shown in Table 2. These images are printed onto a full-frame 35mm film, which is  $36\text{mm} \times 25\text{mm}$ . The resulting pixel density for each of the formats is show in Table 2. When a 4k image is printed onto full-frame 35mm film, an image is produced with a pixel density that is similar to the prototype camera's image sensor. Images with resolutions of 8k and 16k have a pixel density that is approximately and respectively two and four times greater than that of the prototype camera's image sensor. Many of the theorems behind the extraction of the 4D light field rely on spatial measurements, and this implementation is accomplished at the pixel level. This table can be used to look up the conversion from pixels to millimeters and vice versa, depending on the application.

Table 2 This table shows the pixel density of important image formats.

<b>Image Format</b>	<b>Horizontal Rows</b>	<b>Vertical Rows</b>	<b>Horizontal Length (mm)</b>	<b>Vertical Length (mm)</b>	<b>Horizontal Pixels/mm</b>	<b>Vertical pixels/mm</b>
Camera Sensor	3072	2048	22.7	15.1	135.3	125.6
4k	4096	2732	36	24	113.8	113.8
8k	8192	6144	36	24	227.5	256
16k	16384	12288	36	24	455.1	512

### Printing Transparencies

There two different techniques for printing masks:

- 1) Print a 35mm color transparency
- 2) Print a 35mm black and white negative

Color transparencies are typically used in slide projectors. The first attempt at printing was to create a color transparency of Figure 15. The resulting slide appears to be the same as Figure 15 because it is a transparency. The image is directly mapped onto the film and there is no need for image inversion. The undeveloped transparency is initially clear. However, it reacts when exposed to a source of light. After development, the film becomes the color of the light to which it was exposed. There were several issues with the slides produced by this process, and the resulting slides were not functional. For instance, the locations that should have been completely black were neither dark nor opaque enough, and they allowed light to pass through. Additionally, there were issues with the locations that were printed gray. The darker shades of gray were not as dark as they should have been. Additionally, though the pixels that were the lighter shades of gray had proper coloring, they exhibited neither the desired transparency characteristics nor the ability to block intense levels of light. When testing the slide with

a back light, the slide was unable to successfully block light. The color black is needed to completely block light and this method was unable to produce a mask that could fully block a source of light. Overall, this method created issues that made it difficult to produce a mask with the desired characteristics.

### Printing Black and White Negatives

The next attempt at producing a printed mask involved black and white negatives. Black and white negatives are traditionally used to make picture prints. The negative holds an inverse of the final image that is to be developed into a print. The undeveloped negative is nearly completely opaque. After development, the area that was exposed to light is lighter, the value of which is based on the amount of exposure. When exposed completely, the negative becomes clear. This type of development is the opposite of the type used with color transparencies, for which exposure to light will darken the slide. With black and white negatives, the inverse of the desired image will develop on the slide. Therefore, to obtain the desired results, the negative of the image in Figure 15 was created and sent to be printed into a slide. Equation 13 shows how to take the negative of an image. In this equation,  $I$  is an original image with pixels at locations  $i$  and  $j$ . Additionally,  $N$  is the resulting negative with pixels at locations  $i$  and  $j$ . The negative of a pixel is the maximum value of that pixel subtracted from the original pixel value. In this case, the maximum value of a pixel is 255. The resulting image will have light values inverted so as to produce the corresponding dark values. Figure 16 is an example of the results of the negative image equation. It depicts the negative image of the sample mask shown in Figure 15.

$$N(i, j) = 255 - I(i, j)$$

13

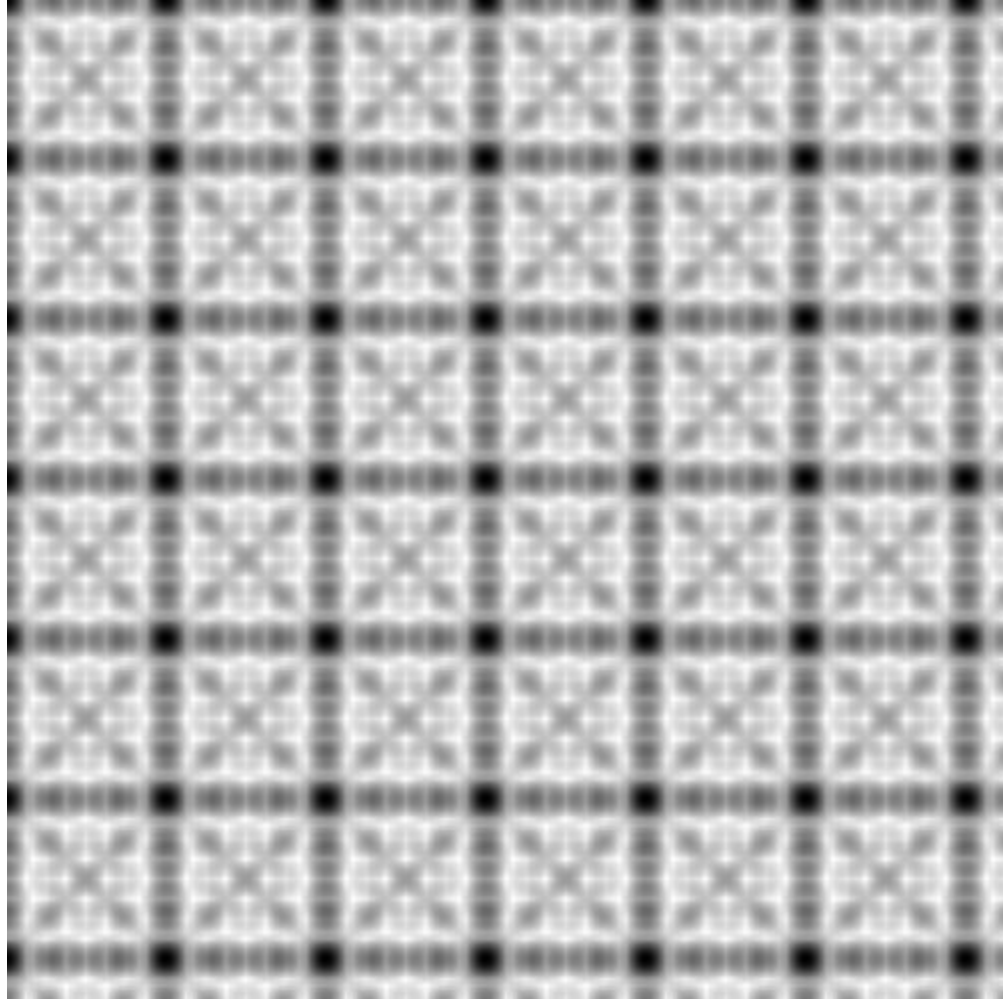


Figure 16 This figure shows a negative of a printed mask. This is the negative of the mask shown in Figure 15.

### Printed Mask Prototypes

Taking the considerations described in the previous sections, five prototype masks were printed onto black and white negatives. Using the Matlab function shown in Figure 14, the five masks were produced with the  $p$  and  $s$  values shown in Table 3, in which  $p$  is the number of cosine harmonics in the mask and  $s$  is the width in pixels of the lowest harmonic cosine wave. The masks were printed at a 4k resolution of  $4096 \times 2732$  pixels. The printer service was unable to print at a higher resolution because of the granularity of

the film. After the mask was generated, a negative was created using Equation 13. The resulting full scale image for the original mask and the negatives are shown in Appendix A: Printed Mask.

Table 3 This table displays prototype printed mask design values.

	Mask 1	Mask 2	Mask 3	Mask 4	Mask 5
$p$	2	2	4	10	5
$s$	5	10	10	20	20

Figure 17 is a photo of the first batch of printed masks produced with this process. This was the first mask that was able to diffract light. However, there were still a few printing issues. For example, the printing service focused the printer at half the size of a full 35mm film. The result was that the pattern was decreased by a factor of two. Even at this scaled, the mask was able to show light diffraction. An issue with these masks is that the modulation function is undeterminable because of the scaling. After this first batch of printed masks was completed, it was necessary to develop a method to test them to determine whether a mask was produced as originally intended. In the next chapter, “Testing Printed Masks,” a method of testing printed masked is discussed. Additionally, by comparing the first and second batches of masks, it was determined that the type of film used affects the color of the printed mask. Specifically, the film used for the second batch of masks possesses better grayscale coloring. An example of the better film is shown in Figure 18, which is the first figure depicted in the following chapter. The film in Figure 18 has better grayscale coloring then Figure 17.



Figure 17 This is an example of a printed mask on 35mm film.

## TESTING PRINTED MASKS

After the manufacturing of the printed masks, a process was necessary to test whether the resulting masks exhibited the characteristics of the original design. The primary concern with these masks relates to their ability to modulate incoming light with a desired modulation function. One way to test these masks is to apply a coherent visible light source. This allows for the determination of whether a mask produces the desired modulation pattern. Figure 18 shows an example of this test being performed.

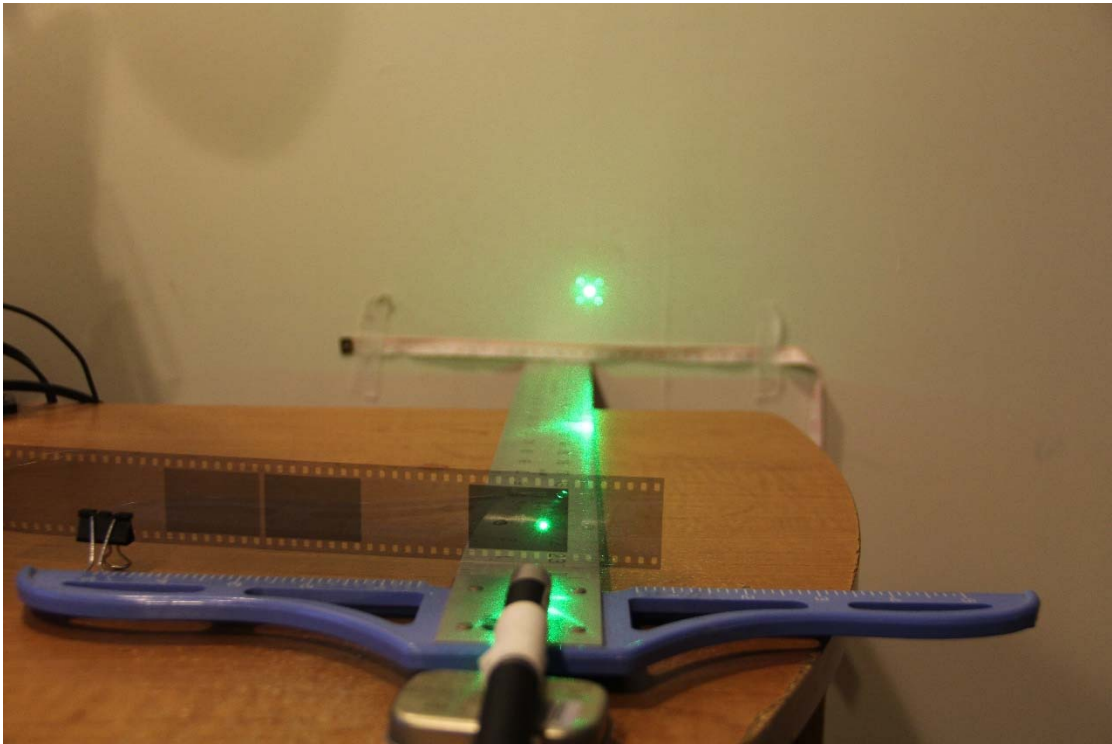


Figure 18 The following photograph demonstrates how to test a printed mask with a laser pointer.

In the test, a 532nm laser with a maximum output of 5 mW was pointed at one of the printed masks. The coherent light from the laser is modulated with the mask. The resulting modulated light is projected onto a flat surface. In this case, the flat surface is a

wall. An example of the projected mask is shown in Figure 19. This pattern relates to the Fourier transform of the mask. Each major dot in this projected pattern corresponds to an impulse function in the Fourier transform of the original mask. The impulses located closest to the center of the pattern correspond to the carrier frequencies of the lowest harmonic cosine waves. The next most proximal impulses correspond to the second harmonic cosine waves. This relationship continues to the most distal impulses.

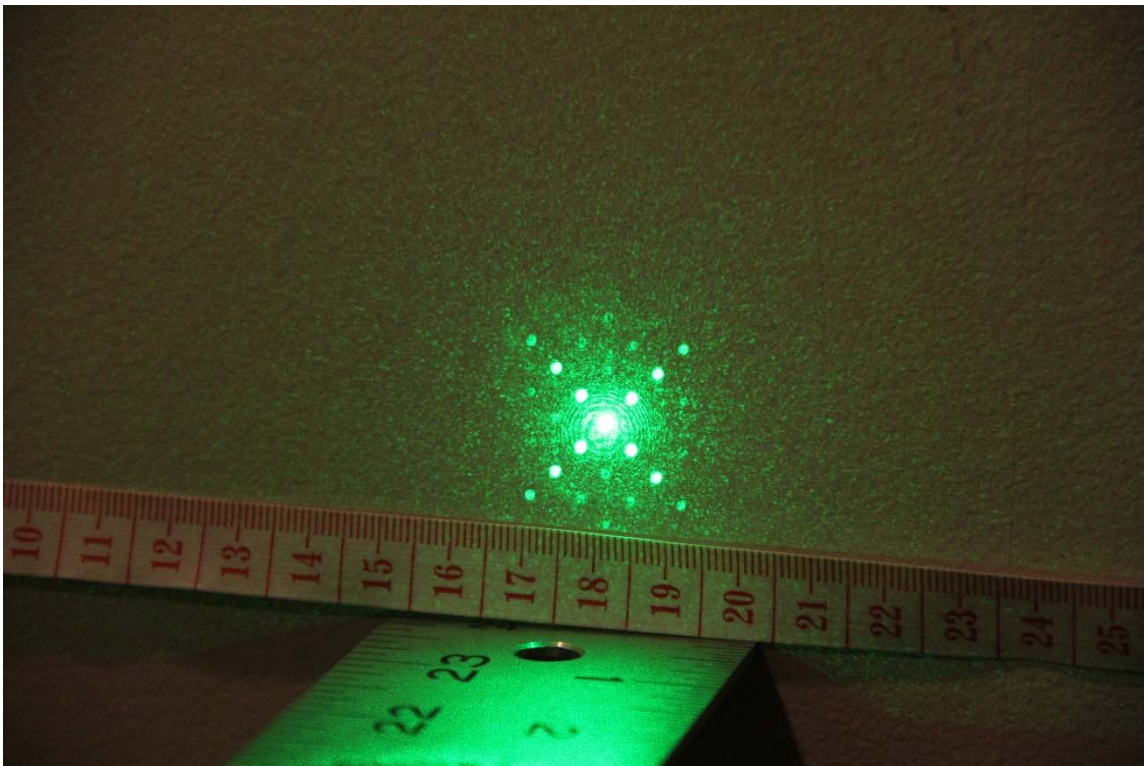


Figure 19 This is an enlarged image of the diffraction pattern. It is produced by coherent visible light passing through the printed mask.

Depending on the distance between the mask and the wall, the pattern formed by the laser passing through the mask varies in scale. The scaling of the pattern is an example of diffraction. Generally, instances of diffraction can be grouped into two categories. These categories are near-field and far-field diffraction. Fresnel diffraction describes near-field diffraction, and Fraunhofer diffraction describes far-field diffraction.



The diffraction pattern depicted in Figure 19 is a pattern produced by near-field diffraction. To be considered far-field diffraction, the distance between the mask and the wall would have to be much greater than the wavelength of the applied visible light. Therefore, the projection of the pattern on the wall is primarily governed by the Fresnel diffraction. The definition of Fresnel diffraction, identifies, and several relevant equations and forms are described in *Introduction to Fourier Optic* [26].

Using the Fresnel approximation for near-field diffraction, the diffraction equation reduces to Equation 14. In this equation,  $E(x', y', z')$  is the electric field at the printed mask, and  $E(x, y, z)$  is the electric field at location  $z$  down range. The  $z$  axis is the optical axis that travels between the mask and the wall. The  $x$  and  $y$  axes are orthogonal to the  $z$  axis and to each other. They both lie in plane with the printed mask and the projection wall. The frequency of the coherent test light is  $\lambda$ . For this test,  $\lambda$  is a constant. The wavelength of the coherent test light is  $k$  and is equal to  $2\pi/\lambda$ .

$$E(x, y, z) = \frac{e^{jkz}}{j\lambda z} \iint_{-\infty}^{+\infty} E(x', y', z') e^{\frac{jk}{2z}[(x-x')^2 + (y-y')^2]} dx' dy'$$

14

In this form, the equation for Fresnel diffraction is too complex for the application. It is possible to express the integral in Equation 14 in terms of a Fourier transform. By changing variables and through a substitution, it is possible to simplify this equation. Equation 15 defines  $G(p, q)$  as the 2D Fourier transform of  $g(x, y)$ . In a few steps, this equation will be used as a substitution to reduce Equation 14. The  $F\{ \}$  notation is used to indicate a 2D Fourier transform. This notation is also used in other equations.

$$G(p, q) = F\{g(x, y)\} = \iint_{-\infty}^{+\infty} g(x, y) e^{-j2\pi(px+qy)} dx dy$$

15

The impulse response of light in free space propagation is useful in reducing the Fresnel diffraction equation. It is defined in Equation 16. When this value is convoluted with  $E(x, y, 0)$ , it results in any possible value of  $E(x, y, z)$ . The Fourier transform of  $h(x, y, z)$  results in the frequency response of light in free space propagation.

$$h(x, y, z) = \frac{e^{jkz}}{j\lambda z} e^{i\frac{k}{2z}(x^2 + y^2)}$$

16

Using a change of variables, it is possible to express Equation 14 in a reduced form with a Fourier transform. The result is shown in Equation 17. The variable  $z'$  only has a value of zero at the mask, and it is set to that value.

$$E(x, y, z) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{\pi}{\lambda z}(x^2 + y^2)} F\{E(x', y', 0) e^{j\frac{\pi}{\lambda z}(x'^2 + y'^2)}\}$$

17

With the change of variables  $p = x \setminus \lambda z$  and  $q = y \setminus \lambda z$ , it possible to express this equation in terms of the frequency response of light in free space and Equation 15. The result is shown in Equation 18. In equations 18 and 19,  $H(p, q)$  is the frequency response of light in free space propagation, and  $G(p, q)$  is the frequency response of the printed mask. The conversion from frequency response of free space light propagation to the impulse response of free space light propagation is similar to a Fourier transform and is shown in Equation 19.

$$E(x, y, z) = H(p, q) \cdot G(p, q) \Big|_{p=\frac{x}{\lambda z}; q=\frac{y}{\lambda z}}$$

18

$$H(p, q) = F\{h(x, y)\}$$

19

With Equation 18, it is possible to explain the diffraction pattern that is produced when a coherent light source is directed through the printed mask. The term  $g(x,y)$  in Equation 15 is the printed mask. The printed mask consists of different cosine harmonics. Taking the Fourier transform of these cosines produces a sequence of impulses that are defined at  $G(p,q)$ . Those impulses only have nonzero values at specific locations  $(p,q)$ . Changing the variables that are defined in Equation 18, in which  $\lambda$  is the wavelength of the laser and  $z$  is the distance between the mask and the wall, determines the locations of the impulses. All of these terms are defined in meters.

The expected locations of these impulses can be used to test whether the mask was printed correctly and whether it will produce the desired results. Additionally, when a mask shows the proper diffraction pattern at 1 m down range, it can be inferred that it will also behave correctly when placed on top of the imaging sensor, which is a distance of approximately 1 mm.

Example Analysis of Printed Mask 1  $p = 2$   $s = 5$  Using  
Fresnel Diffraction

The theory behind the diffraction patterns produced by a coherent light source passing through a printed mask was discussed in the previous section. As such, this section uses the example of Mask 1 from Table 3 to provide detailed steps regarding analysis. The steps provided below are applicable to the other masks, all of which follow a similar process. For this mask,  $p = 2$  and  $s = 5$  indicate that there are 2 harmonics of cosines and the frequency of the lowest harmonic is 5 pixels in length. The printed mask generation function shown in Figure 14 generates a printed mask that is described by Equation 20. The variables  $i$  and  $j$  are pixels within the mask image. For the value of  $s = 5$ ,  $f_s$  has a value of  $(2\pi)/5$ . The value  $c$  is a constant. For the purposes of this analysis,

this value is unimportant. The function described in Figure 14 uses a constant value of 4. However, for this analysis it is left as the constant variable  $c$ .

$$M(i, j) = \cos(f_s i) * \cos(f_s j) + \cos(2f_s i) * \cos(2f_s j) + c \quad 20$$

Equation 20 is an equation for a digital mask function. All of the Fresnel diffraction approaches described in the previous section assume that the mask is defined by analog functions. Because the mask is printed with a relatively high-pixel density, the granularity of the film smooths out any drastic color changes between pixels. As a result, it will appear as though the mask was generated by means of analog functions. This approximation has proven to be reasonable in testing. With this approximation, Equation 20 needs to be expressed in terms of meters with the center of the filter being at location (0,0). This filter was generated using a 4k image printed onto a full-frame 35mm back and white negative. Using the values from Table 2, the mask has a pixel density of 113.8 pixel/mm both horizontally and vertically. Using this value for pixel density, Equation 21 expresses the mask in meters. The  $x$  and  $y$  are the horizontal and vertical axes in plane with the mask. They share an origin at the center of the mask. This is the same expression of  $g(x,y)$  that is used in Equation 15.

$$g(x, y) = \cos(45520x\pi) * \cos(44520y\pi) + \cos(91040x\pi) * \cos(91040y\pi) + c \quad 21$$

Taking the 2D Fourier transform of  $g(x,y)$  as it is defined in Equation 15, results in a sequence of delta functions. The result is shown in Equation 22. The result has nonzero values at 9 different locations. The equation is expressed in terms of 1-D delta functions  $\delta(a)$ . In this equation, the function is equal to zero for all values of  $a$ , except when  $a = 0$  and the function has a value of one. This equation becomes more intuitive

and easier to use when the function is expressed in terms of 2D delta functions  $\delta(a,b)$ .

When this is the case, the function equals zero for all values of  $(a,b)$ , except when  $a$  and  $b$  are both equal to zero. Equation 23 expressed  $G(p,q)$  in terms of 2D delta functions.

$$\begin{aligned}
 G(p, q) = & c\delta(p)\delta(q) \\
 & + \frac{1}{4}(\delta(p - 22760) + \delta(p + 22760))(\delta(q - 22760) + \delta(q + 22760)) \\
 & + \frac{1}{4}(\delta(p - 45520) + \delta(p + 45520))(\delta(q - 45520) + \delta(q + 45520))
 \end{aligned}$$

22

$$\begin{aligned}
 G(p, q) = & c\delta(p, q) \\
 & + \frac{1}{4}(\delta(p - 22760, q - 22760) + \delta(p - 22760, q + 22760) \\
 & + \delta(p + 22760, q - 22760) + \delta(p + 22760, q + 22760)) \\
 & + \frac{1}{4}(\delta(p - 45520, q - 45520) + \delta(p - 45520, q + 45520) \\
 & + \delta(p + 45520, q - 45520) + \delta(p + 45520, q + 45520))
 \end{aligned}$$

23

There are 9 locations where  $G(p,q)$  has a nonzero value, and they are shown in Table 4. This mask should be able to produce 9 light field slices as the result of these points. These points form the corners of two concentric boxes with a single point in the center. Regardless of the distance between the wall and the mask, this pattern should be visible.

Table 4 This table depicts  $p$  and  $q$  locations, in which  $G(p,q)$  is a nonzero for Mask 1.

$p$	$Q$
22760	22760
22760	-22760
-22760	22760
-22760	-22760
45520	45520
45520	-45520
-45520	45520
-45520	-45520
0	0

Applying a change of variables to  $G(p,q)$  will produce the expected pattern that should be projected onto the wall. The relations  $p = x / (\lambda z)$  and  $q = y / (\lambda z)$  express the pattern projected on the wall for a given wavelength of light and a given distance between the light source and the wall. For this test, the distance  $z$  is equal to 0.5842 m and the wavelength of the test light source  $\lambda$  is equal to 532 nm. Table 5 shows the expected  $x$  and  $y$  locations for the delta functions projected against the wall. These values are found by using the relations  $x = p * (\lambda z)$  and  $y = q * (\lambda z)$  and applying the values from Table 4.

Table 5 This table details the expected and actual  $x$  and  $y$  values of the delta functions for a diffraction pattern produced by Mask 1 at  $z$  distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters.

Expected $x$	Expected $y$	Actual $x$	Actual $y$
7.07	7.07	3	3
7.07	-7.07	3	-3
-7.07	7.07	-3	3
-7.07	-7.07	-3	-3
14.15	14.15	7	7
14.15	-14.15	7	-7
-14.15	14.15	-7	7
-14.15	-14.15	7	-7
0	0	0	0

The two columns on the left contain the actual values from a test. The setup was shown earlier in Figure 18. The resulting diffraction pattern is shown in Figure 20. On the project's surface, there was a tape measure attached. All measurements in the table were measured in reference to the tape measure. The measurements were only accurate within 1 mm of precision. Because of the test setup, it was difficult to identify the center of an impulse with a high-level of precision. The locations of the impulses were at half of their expected value. This shows that the printing process was unable to produce Mask 1 in a manner that would allow the mask to diffract light as expected.

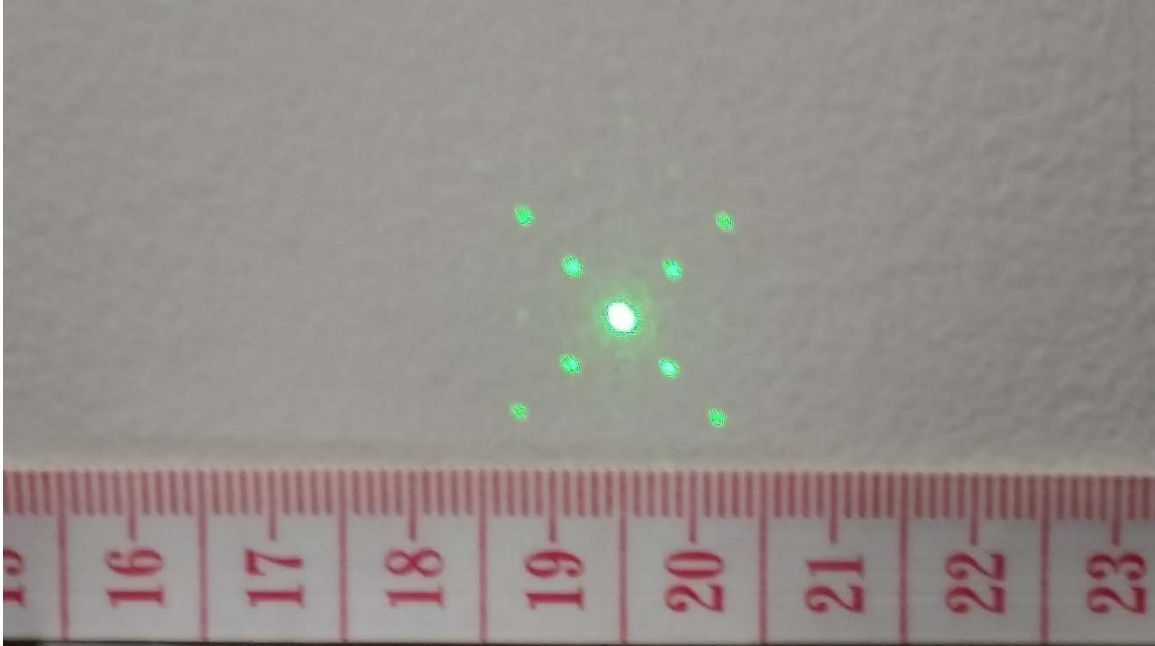


Figure 20 This figure shows the Mask 1( $p = 2$   $s = 5$ ) test pattern at a distance of 0.5842 m down range.

Analysis of Printed Mask 2:  $p = 2$   $s = 10$

The same approach of analysis used for Mask 1 can be applied to Mask 2. Mask 1 and Mask 2 have the same number of cosine harmonics. The values for the frequencies of the carrier cosines for Mask 2 are half of the values for the frequencies of the carrier cosines for Mask 1. Mask 1's lowest harmonic cosine has a period of 5 pixels. Mask 2's lowest harmonic cosine has a period of 10 pixels. With a pixel density of 113.8 pixel/mm in both the vertical direction and the horizontal direction, the printed mask has the transfer function shown in Equation 24. The Fourier transform of the mask is shown in Equation 25.

$$g(x, y) = \cos(22260x\pi) * \cos(22260y\pi) + \cos(44520x\pi) * \cos(44520y\pi) + c \quad 24$$



$$\begin{aligned}
G(p, q) = & c\delta(p, q) \\
& + \frac{1}{4}(\delta(p - 11130, q - 11130) + \delta(p - 11130, q + 11130) \\
& + \delta(p + 11130, q - 11130) + \delta(p + 11130, q + 11130)) \\
& + \frac{1}{4}(\delta(p - 22260, q - 22260) + \delta(p - 22260, q + 22260) \\
& + \delta(p + 22260, q - 22260) + \delta(p + 22260, q + 22260))
\end{aligned}$$

25

This equation has the same number of impulses as Mask 1, but the locations are different. The locations at which  $G(p,q)$  has a nonzero value are shown in Table 6. These nonzero locations are half the value of the expected locations for Mask 1. In the spatial domain, Mask 2's pattern is twice as large as Mask 1's pattern. In the frequency domain, this corresponds to Mask 2 producing a frequency response that is half that of the frequency response produced by Mask 1 [37].

Table 6 This table details the  $p$  and  $q$  locations where  $G(p,q)$  is nonzero for Mask 2.

$p$	$q$
11130	11130
11130	-11130
-11130	11130
-11130	-11130
22260	22260
22260	-22260
-22260	22260
-22260	-22260
0	0

With the change of variables described by Fresnel diffraction, it is possible to know where the impulses should be located when a test light is applied. Using the relations  $x = p * (\lambda z)$  and  $y = q * (\lambda z)$ , with  $\lambda$  equal to 532 nm and  $z$  equal to 0.5842 m, the expected impulse locations are shown in the two columns on the right in Table 7, and the

measured values are located in the two columns on the left. The actual test pattern is shown in Figure 21. The actual values for the impulse response are within 0.5 mm of the expected values. This falls within a reasonable margin of error and shows that Mask 2 diffracts light as expected. This makes the mask a suitable candidate to be tested in a prototype camera.

Table 7 This table shows the expected and actual  $x$  and  $y$  values of delta functions for a diffraction pattern produced by Mask 2 at  $z$  distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters.

Expected $x$	Expected $y$	Actual $x$	Actual $y$
3.5	3.5	3	3
3.5	-3.5	3	-3
-3.5	3.5	-3	3
-3.5	-3.5	-3	-3
7.07	7.07	7	7
7.07	-7.07	7	-7
-7.07	7.07	-7	7
-7.07	-7.07	7	-7
0	0	0	0

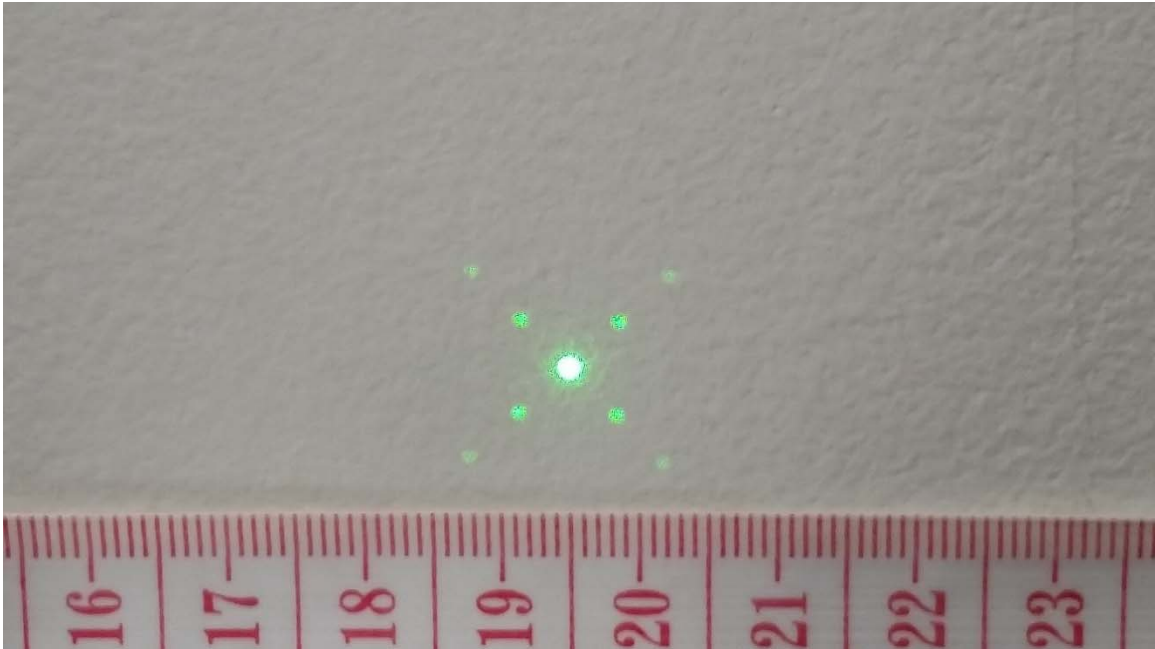


Figure 21 This figure shows the test pattern produced by Mask 2( $p = 2$   $s = 10$ ) at a distance of 0.5842 m down range.

### Analysis of Printed Mask 3: $p = 4$ $s = 5$

Printed Mask 3 had 3 harmonics of cosines. The lowest harmonic cosine had a period of 5 pixels. Originally, this was designed to be a test mask with multiple cosine waves, all of which were intended to possess very short periods. This result of the test is shown in Figure 22. Figure 22 shows that the mask was unable to produce a clean diffraction pattern. Therefore, it was determined that Mask 3 was not a good test candidate. Because this mask did not produce a diffraction pattern with the desired properties, the detailed analysis for this mask is not included. The analysis does follow the same method as Mask 1 and Mask 2, but it has twice as many delta functions that are not located at location (0,0). Mask 3 appears to have only one impulse, which is located at location (0,0). The printing setup was unable to render this pattern. A different

printing technology may be able to produce this mask. However, this method was not able produce this mask.

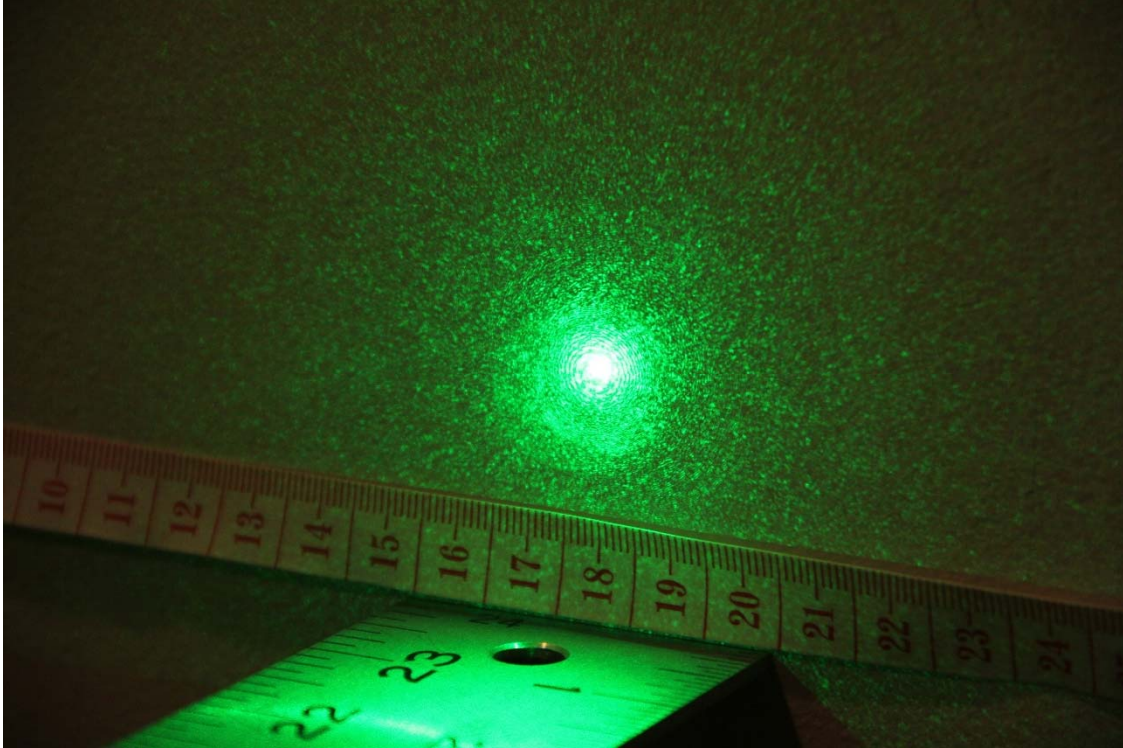


Figure 22 This image shows the Mask 3( $p = 4$   $s = 5$ ) test pattern at a distance of 0.5842 m down range.

#### Analysis of Printed Mask 4: $p = 10$ $s = 20$

The goal with Mask 4 was to test how many harmonics could reasonably be printed on the film. Mask 4 had 10 harmonics of cosine waves with the lowest harmonic having a period of 20 pixels. Some of the higher frequency cosine waves have periods that are smaller than 1 pixel. These cosines should not produce any frequency responses because the printer was not able to print at a resolution-level capable of producing a subpixel. With 10 harmonics of cosine waves, the  $g(x,y)$  and  $G(p,q)$  functions for this mask are quite large. In Mask 4,  $G(p,q)$  consists of only impulse

functions. Table 8 shows the locations of some of these delta functions. Only the 25 of the most proximal impulses in the pattern are shown in Table 8. Additionally,  $G(p,q)$  has a total of 41 delta functions. However, considering the resolution capabilities of the printer, the mask should only be able to produce the 25 most proximal delta.

Table 8 This table details the  $p$  and  $q$  locations where  $G(p,q)$  is nonzero for Mask 4.

$p$	$q$
5690	5690
5690	-5690
-5690	5690
-5690	-5690
11390	11380
11390	-11380
-11390	11380
-11390	-11380
22760	22760
22760	-22760
-22760	22760
-22760	-22760
45520	45520
45520	-45520
-45520	45520
-45520	-45520
91040	91040
91040	-91040
-91040	91040
-91040	-91040
182080	182080
182080	-182080
-182080	182080
-182080	-182080
0	0

Similar to the other masks, the change of variables to show Fresnel diffraction is  $x = p * (\lambda z)$  and  $y = q * (\lambda z)$ , with  $\lambda$  is equal to 532 nm and  $z$  equal to 0.5842 m. The result of applying this change of variables to Table 8 is shown in the two columns on the right of Table 9. While performing the actual test, only some of these impulses were visible. The diffraction pattern for Mask 4 is shown in Figure 23. Only the 13 delta functions in the center are visible. Their measured location is shown in the two columns on the left of Table 9. One limitation of this type of testing setup is that impulses must be visible to the human eye. Some of the delta functions for the more distal locations may have been created. However, they were not visible to the human eye.

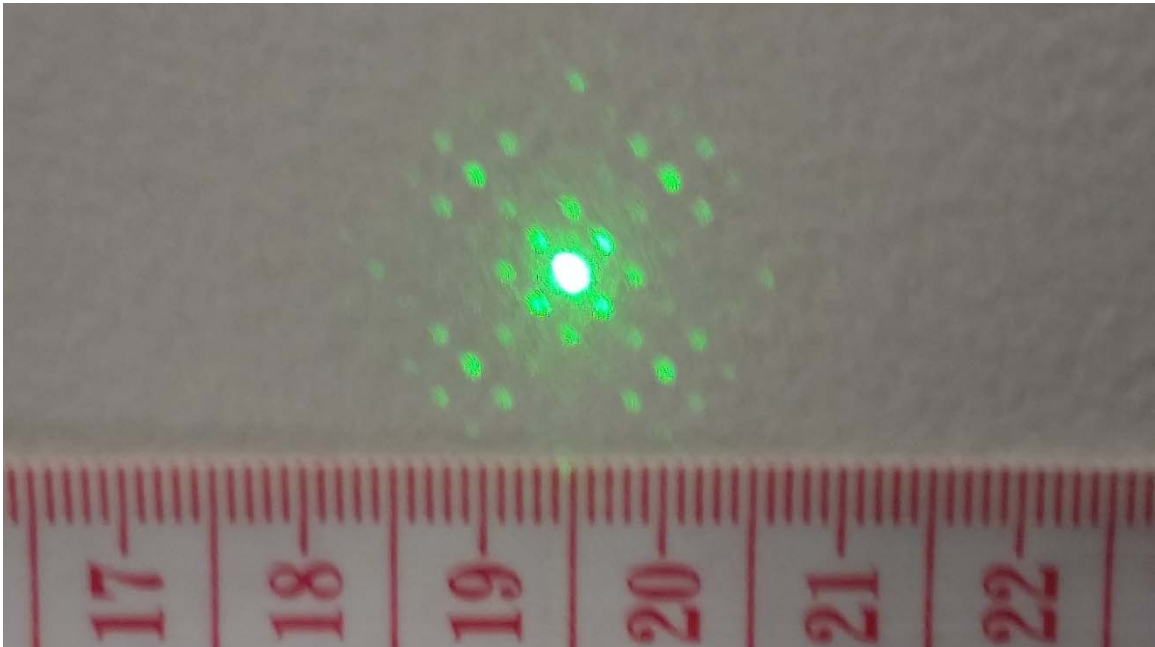


Figure 23 This image is of the Mask 4( $p = 10$   $s = 20$ ) test pattern at a distance of 0.5842 m down range.

Table 9 This chart depicts the expected and actual  $x$  and  $y$  values of the delta functions for a diffraction pattern produced by Mask 4 at  $z$  distance of 0.5842 m. All distances are measured from the center of the pattern in millimeters.

Expected $x$	Expected $y$	Actual $x$	Actual $y$
1.77	1.77	1.8	1.8
1.77	-1.77	1.8	-1.8
-1.77	1.77	-1.8	1.8
-1.77	-1.77	-1.8	-1.8
3.54	3.54	3.6	3.6
3.54	-3.54	3.6	-3.6
-3.54	3.54	-3.6	3.6
-3.54	-3.54	-3.6	-3.6
7.07	7.07	7.1	7.1
7.07	-7.07	7.1	-7.7
-7.07	7.07	-7.1	7.1
-7.07	-7.07	-7.1	-7.1
14.15	14.15	-	-
14.15	-14.15	-	-
-14.15	14.15	-	-
-14.15	-14.15	-	-
28.29	28.29	-	-
28.29	-28.29	-	-
-28.29	28.29	-	-
-28.29	-28.29	-	-
56.59	56.59	-	-
56.59	-56.59	-	-
-56.59	56.59	-	-
-56.59	-56.59	-	-
0	0	0	0

The test for Mask 4 had some unintended artifacts. There appears to be impulses at locations that were not expected. The intensity of the light at these locations is much weaker than anticipated. It is likely that these impulses are the result of imperfect



diffraction. This is difficult to describe using the diffraction model posited in the previous section.

Though Mask 4 did not perform as expected, it did exhibit some interesting characteristics. This made it worth testing in an actual camera. Only a portion of the designed impulses showed up in the test. However, the mask did produce a dense array of impulses that were evenly distributed throughout the test area. Therefore, though there were unexpected impulses, the locations were beneficial because they occupied sections of the 2D frequency response that were not being used. The locations of these points should not cause aliasing. Knowing the diffractive properties of the mask helps explain the mask's behavior when it is installed in a camera.

#### Analysis of Printed Mask 5: $p = 5$ $s = 20$

The final test mask was Mask 5. Mask 5 had 5 harmonics of cosine waves, the lowest of which had a period of 20 pixels. With this design, the period of the 5 cosine waves were 20, 10, 5, 2.5, and 1.75 pixels. The results of the application of the test light are shown in Figure 24. The diffraction pattern shown in the figure is not usable as a modulation mask. There are only 5 impulses visible and only 4 of them appear to be clean. The impulse at (0, 0) is too large. Therefore, it is unusable. The 4 impulses that appear clean are located at (-7,-7), (-7, 7), (7, 7), and (7,-7). These points have units in millimeters. These impulses correspond to the cosine waves that have a period of 10 pixels. Though these 4 impulses are very clean, the rest of the mask will not be a viable light-field modulation mask because of the diffraction pattern located at (0,0). The pulse size is too large to be useful.

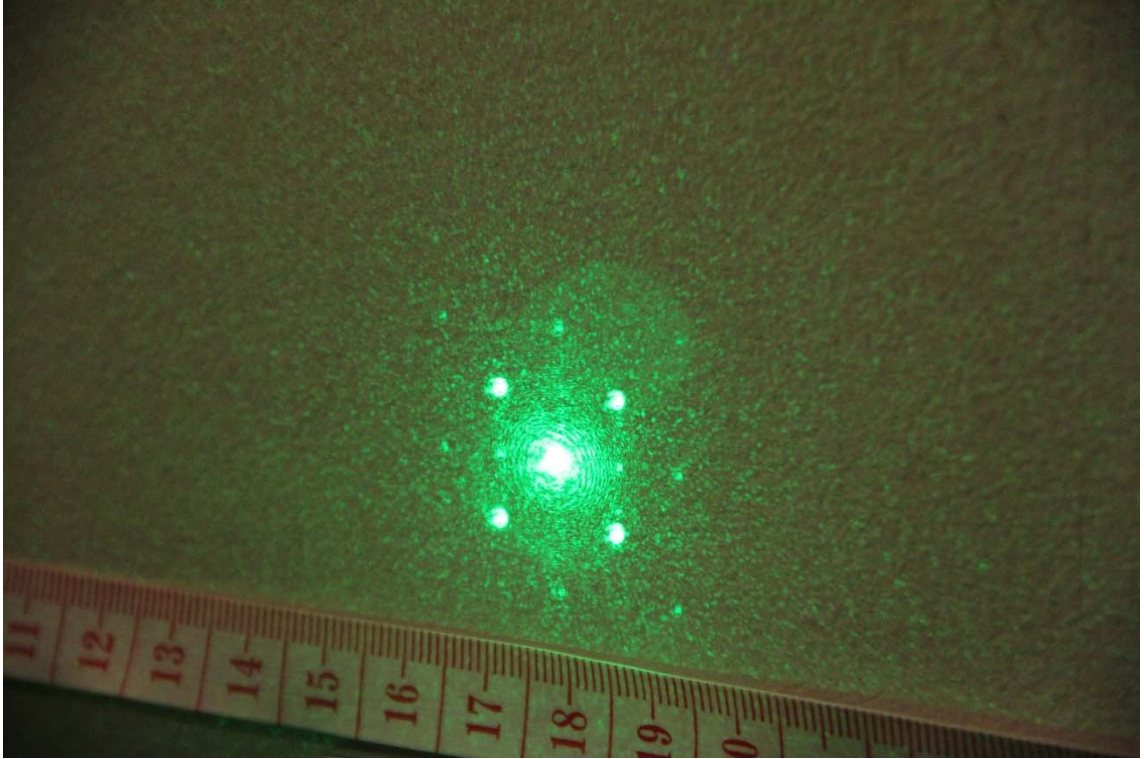


Figure 24 This image is of Mask 5's test pattern ( $p = 5$   $s = 20$ ) at a distance of 0.5842 m down range.

### Potential of Using Fresnel Diffraction to Design New Mask

#### Patterns

Prior to this study, several approaches have been taken in the construction of a light field camera. All previous approaches share one commonality. They all filter incoming light in a manner that allows for individual rays to hit individual pixels. This allows for the origin of each ray to be determined. In several designs, ray tracking is used to describe the behavior of the light field within the camera [1, 2]. There have been attempts to place filters on each individual pixel of a CMOS chip. This design would allow for light to be filtered before it hit the pixels according to specified angles [39]. Each pixel could be then tuned to a specific ray. Printed mask light field cameras use concepts from communication theory to explain the origin of each ray [36, 40, 41]. The

effects that these printed masks have on the light field can also be explained by Fresnel diffraction. Traditionally, these mask were not thought of as a diffractive optics system. This chapter has shown that these printed masks operate as a diffractive optics system in a manner that is described by Fresnel diffraction.

Because the masks are devices that are capable of producing diffraction, different modulation functions are possible. However, modulation theory does not take into consideration the distance between the mask and the sensor. This distance could be modified in the designs of other printed masks. Working with an existing camera, it will not be easy to modify the distance. If a camera is specifically designed with this consideration, it might be possible to produce a light field camera that is more suitable for airborne platforms. Using a number of design considerations, it might be possible to create a single-lens light field camera that is optimized for flying at a specific altitude above an imagine area. Those design considerations are the focal length of a main lens, the distance between the main lens and the printed mask, the distance from the printed mask to the imaging sensor, and the PPI on the imaging sensor. Testing a camera with these considerations was not possible for a project of this size. However, it would be a feasible endeavor for a camera manufacturer.

## INSTALLATION OF PRINTED MASK

The mask that successfully demonstrated Fresnel diffraction was determined to be a potentially viable modulation mask for a light field camera. This mask was installed into a camera to determine whether it was functional. The goal was to mimic the camera setup in *Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing* [36] and *Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses* [41]. The cameras that were used in these papers were professional-grade cameras. These cameras have a common property. They all have an easily accessible image sensor. Two camera models were tested to determine whether they were suitable surrogate cameras, the Canon EOS Rebel Xsi and the Fujifilm Finepix S5200.

Canon EOS Rebel Xsi was a very popular camera when it came out in 2010. At the time of the study, the resale price of the camera was low because the model was a few generations old. This camera has a large do-it-yourself (DIY) community [42]. This camera is popular among astronomers because it is possible to remove the IR filter. The imaging sensor is 22.3mm×14.9mm and has 12.2 megapixels [13]. This image sensor is accessible with some effort. The goal is to place the printed mask onto the image sensor in a manner that is as close as possible. To get as close as possible to the image sensor, the printed mask was placed between the image sensor and the IR filter. The steps that were taken to access the camera's image sensor are shown in Figure 25. After removing the black panel of the camera, the image sensor's mounting becomes visible. It is the large silver object in the center of the Subfigure ii. To access the image sensor, the top-layer circuit board needed to be removed. This was done by disconnecting a number of ribbon cables and unscrewing the mounting points. The result is shown in Subfigure iii. At this point, there are only a few screws that need to be removed to disconnect the image sensor's mounting. The image sensor's mounting is shown in Subfigure iv. The active IR

filter was press-fitted onto the image sensor's mounting. It was carefully removed with a pair of tweezers. Once removed, the printed mask was placed on the image sensor and the IR filter was placed back onto its original position. The placement of the printed mask is shown in Subfigure vi.



Subfigure i



Subfigure ii



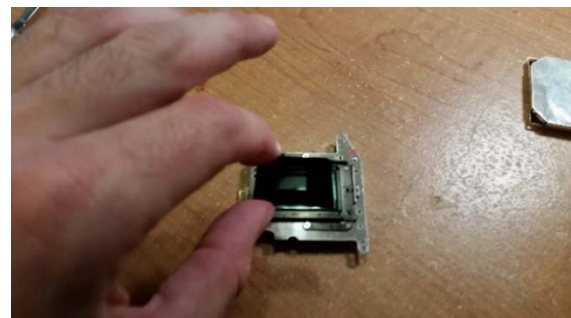
Subfigure iii



Subfigure iv



Subfigure v



Subfigure vi

Figure 25 This figure is comprised of images that show the different steps of the installation of the printed mask for a Canon EOS Rebel Xsi. Subfigure i) shows a fully assembled camera. Subfigure ii) is a photo of a camera with the back panel removed. Subfigure iii) depicts the image sensor's mounting after the removal of the top-layer circuit board. Subfigure iv) is a photo of the removed image sensor's mounting. Subfigure v) is a photo of the image sensor's mounting with the IR filter removed. Subfigure vi) illustrates the placement of a printed mask onto the image sensor.

After the printed mask was installed, the camera was reassembled. The reassembly was straightforward. The printed mask did not require the camera to be

modified in any major way. The ribbon cables in this camera were very difficult to work with because they were fragile and easily broken. After a few printed mask installations, this model of camera exhibited irregular behavior. Producing a reliable camera with a printed mask installed proved to be very difficult with this model of camera.

The other camera was the Fugifilm Finepix S5200. It is easy to repair a Fugifilm Finepix S5200. Additionally, there are many DIY websites that demonstrate techniques for the removal of the IR filter on this camera. This camera is not a DSLR camera. This camera has some properties that are different from a typical 35mm camera. The image sensor in this camera contains  $3688 \times 2776$  pixels. The image sensor in this camera has a physical size of  $5.76\text{mm} \times 4.29\text{mm}$ . The camera has a magnetically controlled lens. The focal length of this lens is between 6.3 mm and 63 mm [43]. This camera has a much higher PPI than the Canon EOS Rebel Xsi. The higher PPI was beneficial and is described later in the section on the processing of these images.

Some of the steps for installing the printed mask are shown in Figure 26. Subfigure i shows the fully assembled camera. Once the back panel is removed, the internal components of this camera are much easier to work with. Subfigure ii shows the inside of the camera when the back cover is removed. All of the cables are quick-release cables and there are only a couple of screws securing each component. Subfigure iii shows the inside of the camera when the top circuit board is removed. At this point, the metal mounting bracket that holds the image sensor in place is visible. Subfigure iv shows the camera with the second circuit board removed. Additionally, Subfigure iv shows that the mounting bracket is only secured to the device by a couple screws. Once this bracket was removed, the image sensor was pulled back as shown in Subfigure v. The printed mask was then placed onto the lens, and the image sensor was returned to its original position.





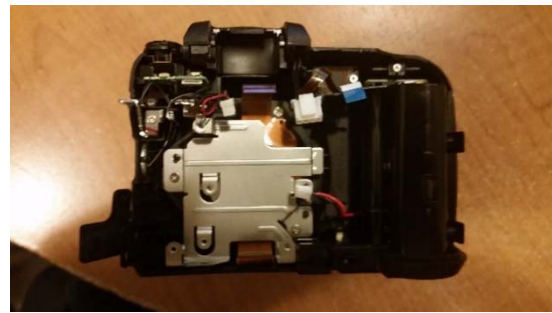
Subfigure i



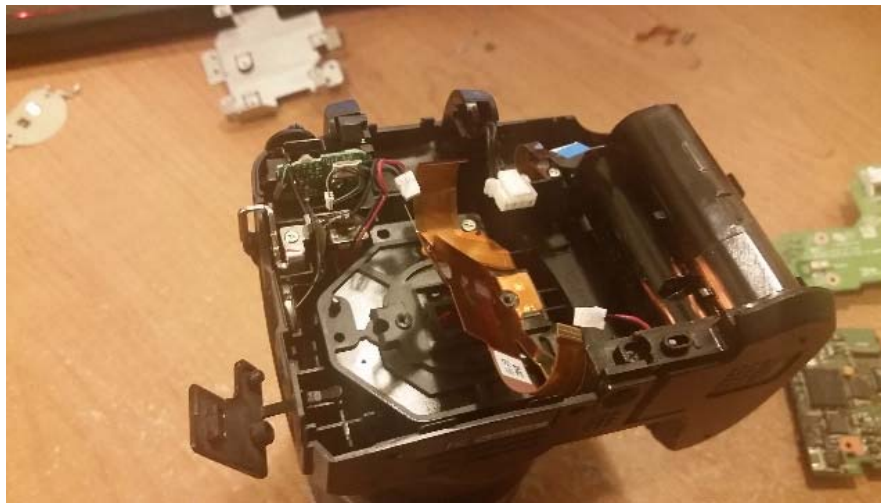
Subfigure ii



Subfigure iii



Subfigure iv



Subfigure v

Figure 26 These images depict different steps of the installation of the printed mask into a Fujifilm Finepix S5200. Subfigure i) is a view of the fully assembled camera, ii) is a view of the camera with the back panel removed, iii) is a view with the top-layer circuit board removed, and iv) is a view of the camera with the bottom-layer circuit board removed and the image sensor back plate exposed. Subfigure v) is an image of the camera with the back plate removed and with an elevated image sensor.



The camera was then reassembled. This camera is easy to reassemble because it is held together by a small number of screws and cables. Overall, the layout of the inside of this camera is easy to work with. After it was reassembled, the camera was determined to be functional. Additionally, it did not exhibit any unexpected behavior. Because this camera is a few years old at the time of the study, it was relatively inexpensive. Printed Mask 2 and Mask 4 were installed into two separate cameras. The analysis of the images produced by these two cameras is described in “Testing Printed Masks.”

## PROCESSING CAPTURED LIGHT FIELD IMAGES

This chapter details the methods by which raw light-field images are processed. In doing so, this chapter introduces a new method to extract the 4D light field from the Fourier domain. Images captured from a Lytro camera were used to validate this method. This chapter explains the method that was applied to the images captured by the prototype cameras. Additionally, this chapter details the mapping of the Fourier slices into the light-field equation. The technique for extracting the 4D light field from Lytro images is compared to the method of extraction used by the Matlab Light Field Toolbox. It was determined that the Matlab Light-Field Toolbox cannot process the raw images produced by the prototype cameras. The new method that is described in this chapter can be used for both camera types.

### Fourier Domain Light Field Extraction

The method described by Equation 12 for extracting the 4D light field from an image captured by a printed-mask lightfield camera has some inherent assumptions. It is assumed that the intrinsic properties of the camera are completely known. Additionally, traditional intrinsic properties of cameras are also assumed. These properties include the focal length, the format of the image sensor, and the principle point. Determining whether the printed mask is aligned with the imaging sensor is necessary. In practice, the manufacturing and installation of the filter is important. However, it was not practical for this application to attempt pixel-level accuracy in the alignment of the imaging sensor and the printed mask. To test whether a printed mask was manufactured and installed correctly, it was necessary to implement a 4D light field extraction algorithm that worked with a microlens array light field camera and a camera with a printed mask. When an algorithm is implemented that works with a microlens light field camera, it can be used to

detect imperfections in the process of printing the masks. The sample images in Veeraraghavan's paper on doppler photography look similar to raw images produced by a microlens camera [36]. The following section describes an attempt to modify Veeraraghavan's approach so that it translates to microlens light field images.

The image that was used to develop this algorithm is shown in Figure 27. This image was taken from a MI-2 helicopter that was flying at a low altitude above the Iowa City Municipal Airport (K10W). The image was transformed into grayscale to reduce the complexity of the algorithm. There are three distinct depth regions that are present in the image. In the background of the image is the horizon of Iowa City. For this part of the image to be in focus, the camera would have to be focused at infinity. Capturing the light field is not practical at infinity, and this is a condition in which all slices of the light field should be uniform. In the middle range are two calibration targets that have a known pattern printed onto them. These targets were used to calibrate imaging sensors for the study, and their size and pattern is known. The bottom half of one of the targets is enlarged on the right side of Figure 27. In each lens, slices of the light field are visible. The straight lines of the calibration pattern can be used for test cases. There was a mounting bracket placed on the camera to hold it during the collection process. Part of the mounting bracket is visible in the upper-right corner of the image. This is an example of an extremely close object. The bracketing has straight edges that can be used as registration points in the algorithm. Overall, this image has a diverse set of features that can be used to help determine the robustness of this calibration process. Figure 28 shows the captured scene. This figure was produced by taking the raw light field image and producing a focused and color balanced image with a Lytro camera's software.

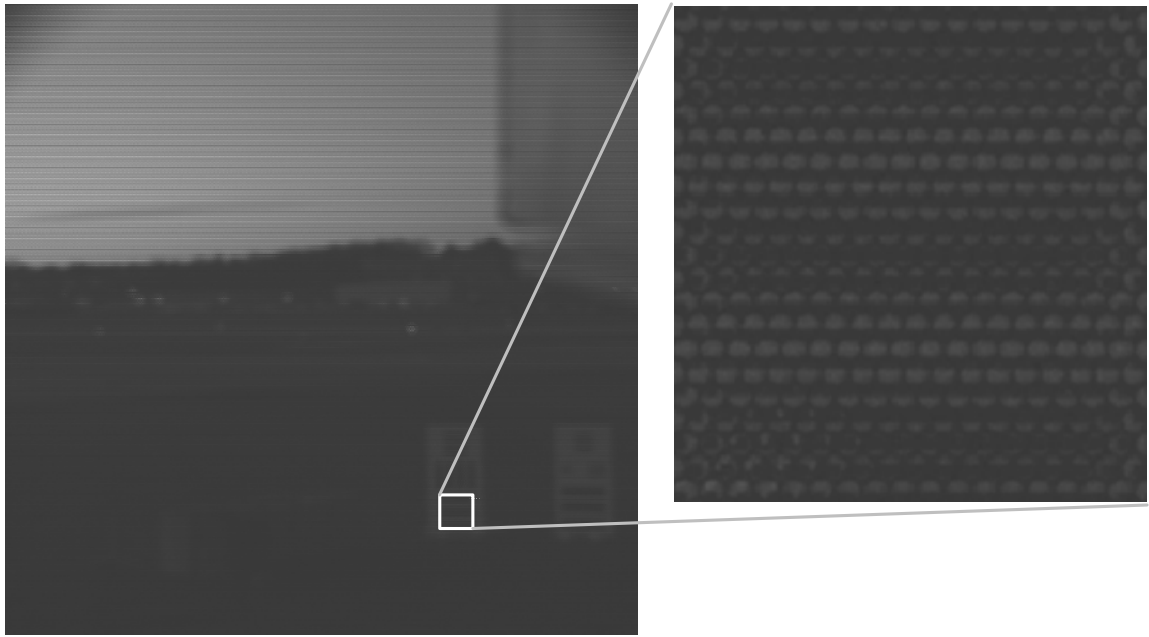


Figure 27 This figure is an image captured from a Lytro camera that has been converted to grayscale. The image on the left is the full image. The image on the right is an enlarged section of the calibration target.



Figure 28 This figure is a color version of the test image collection with the Lytro camera. The image was color balanced and focused with the Lytro software.

The first step in constructing the 4D light field is to perform a 2D discrete Fourier transform (DFT) on the original image. A DFT converts a discrete spatial signal into a frequency domain representation. A 2D DFT is able to produce a discrete frequency domain representation of a 2D image. Equation 26 is a 2D DFT function for a monochromatic image [44]. The original image is represented by the function  $f(m,n)$ , in which  $m$  and  $n$  are the locations of pixels within the image. The variables  $m$  and  $n$  index across the two axes of the image. Either variable can represent either axes. In the following examples,  $m$  is described as indexing from the top to the bottom across the image, and  $n$  is indexing from the left to the right across the image. In this equation,  $F(u,v)$  is the resulting image of a discrete frequency domain. The  $u$  dimension in this function represents the frequency components along the  $m$  dimension. Likewise, the  $v$  dimension of the function represents the frequency components along the  $n$  dimension. In this equation,  $M$  and  $N$  are the total number of pixels along the  $m$  and  $n$  dimensions of the image, respectively. For this application, the input image contains only real pixel

values. Each output pixel in  $F(u,v)$  is the result of the summing of complex exponentials that are scaled by the pixel values in the original image. These resulting values contain complex values and are stored as a complex value.

$$F(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \exp \left[ -2\pi j \left( \frac{mu}{M} + \frac{nv}{N} \right) \right]$$

26

The frequency image that is produced with 2D DFT is periodic, with period  $2\pi$ . The equation above returns only values for one period. The variables  $u$  and  $v$  are discrete frequency values. The width of each discrete sample is  $2\pi/M$  for  $u$  and  $2\pi/M$  for  $v$ . The highest possible frequency that can be represented by a DFT function is determined by the sampling rate of the original analog system. Figure 29 shows the mapping between analog frequency, angular frequency and samples for a DFT. The value  $\pi$  is the highest possible frequency component. Both 0 and  $2\pi$  are DC components of the signal. In the  $m$  direction, the  $M/2$  sample represents the highest frequency component. Both 0 and  $M$  are DC components of the signal. The same is true in the  $n$  direction. All of 2D DFT images in this dissertation will use angular measurements to show frequency.

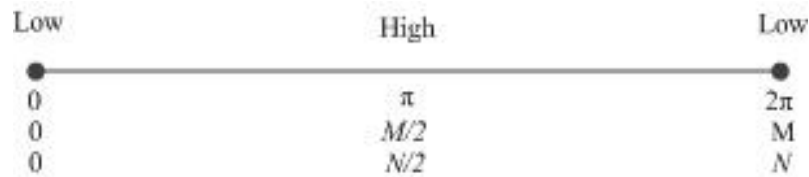


Figure 29 Analog frequency, angular frequency and sample frequency mapping

With this visualization of how low and high frequency components are encoded, it is possible to look at the DFT of an image. Equation 26 was applied to the raw image shown in Figure 27. Figure 30 depicts the results. The original image was a grayscale

image that contained no complex values. To visualize complex numbers in the resulting image, color was used to represent the phase of a pixel value. Yellow represents a purely real value. Red represents a purely imaginary value. These two colors represent both ends of the phase spectrum. The gradient between red and yellow represents all possible phase values. The intensity of a pixel represents the magnitude of the complex pixel value. In this visualization, the low-frequency components are located at the outer edges of the image. The high-frequency components are located near the center of the image. There are large values at the locations of the low-frequency pixels and small values at the locations of the high-frequency pixels. This is typical for most images because the majority of neighboring pixels in a spatial image exhibit relatively minute differences. Changes within the image happen slowly. Often, high-frequency changes appear as noise.

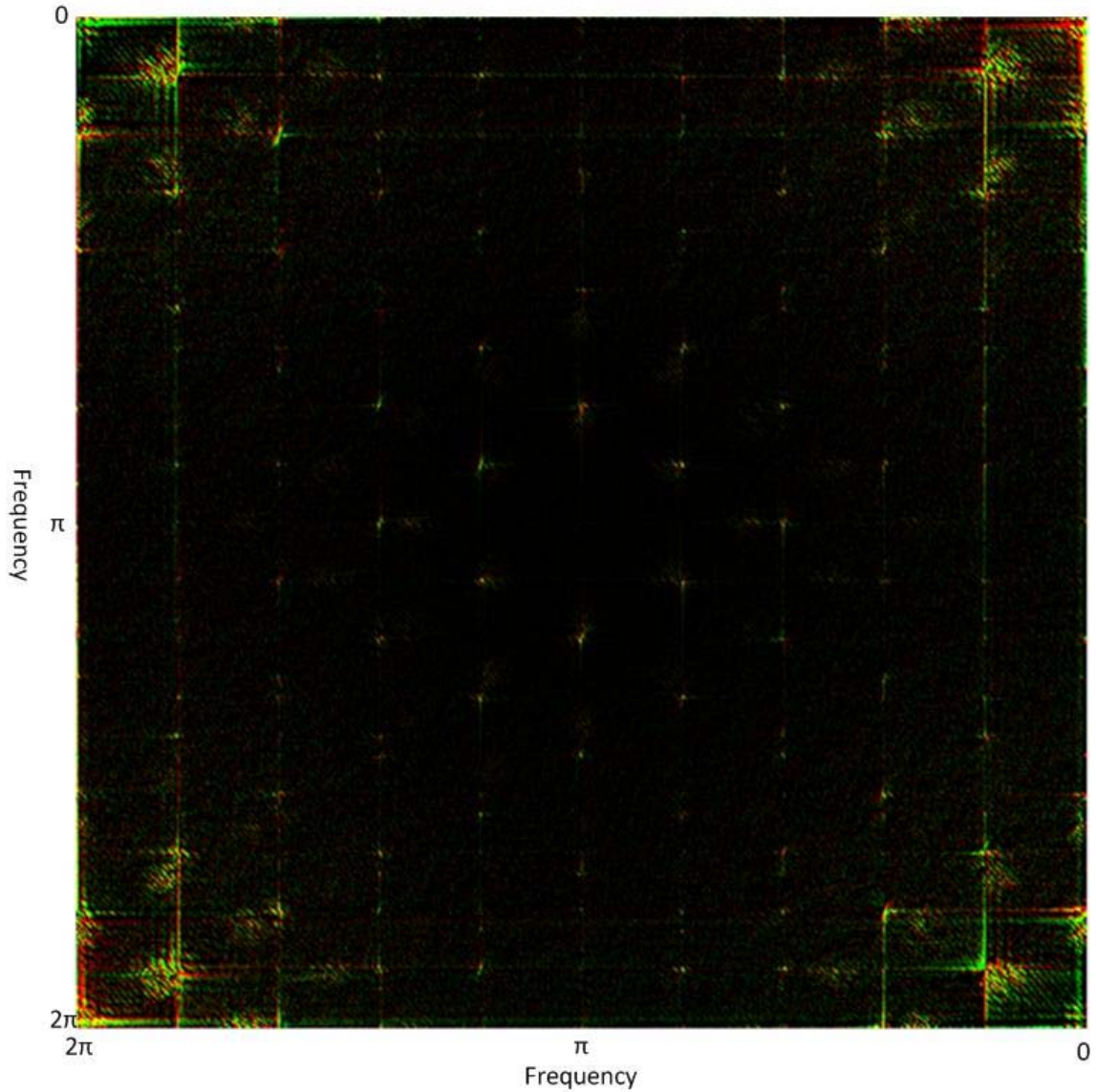


Figure 30 This figure depicts an unshifted DFT of the test image.

This type of frequency representation contains low-frequency components located in the distal areas of the image. However, this is not convenient for some of the final steps of the algorithm. It is much more useful to represent low-frequency components at the center of the image and to represent the high-frequency components in the distal areas. This allows the representation to remain consistent with the analog Fourier



transform, in which the low-frequency components are at the center of the image. This is a fairly common practice and is referred to as a DFT shift [44].

As discussed before the result of a DFT function is periodic. The result is shifted copies of the analog frequency response. Because the transform is periodic, it is possible to shift the period. Figure 31 shows a 1D frequency signal produced by a DFT function. The frequency response of the system is a triangle wave. The peak of the triangle is at zero frequency. The lower corners of the triangle are high frequency components. There are copies of this transfer function every  $2\pi$ . The DFT function that was described before samples the frequency signal between 0 and  $2\pi$ . This is shown by the sampling window in the figure. The middle of the sampling window is the highest frequency component. The outer edges contain the low frequency components.

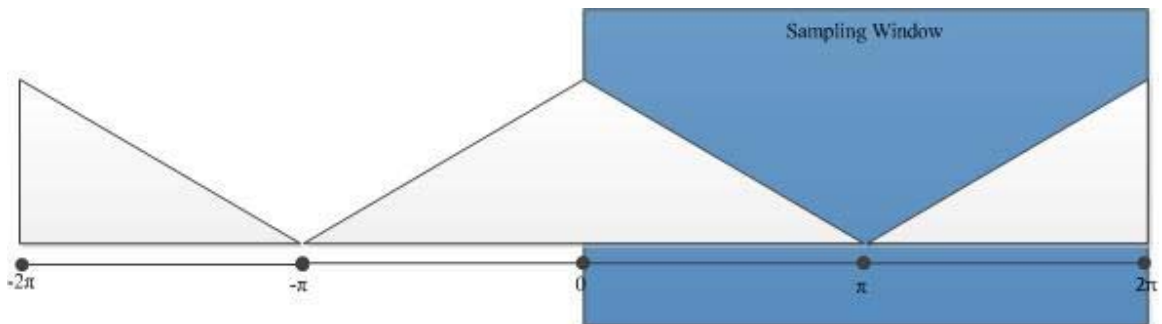


Figure 31 Sampling window for an unshifted 1D DFT signal

The sampling window can be shifted so that the middle of the window contains the low frequency components. This is shown in Figure 32 where the periodic function is sampled from  $-\pi$  to  $\pi$ . The analog Fourier Transform has the low frequency components at the center of the image, and with this shift, the digital frequency signal has the same layout.

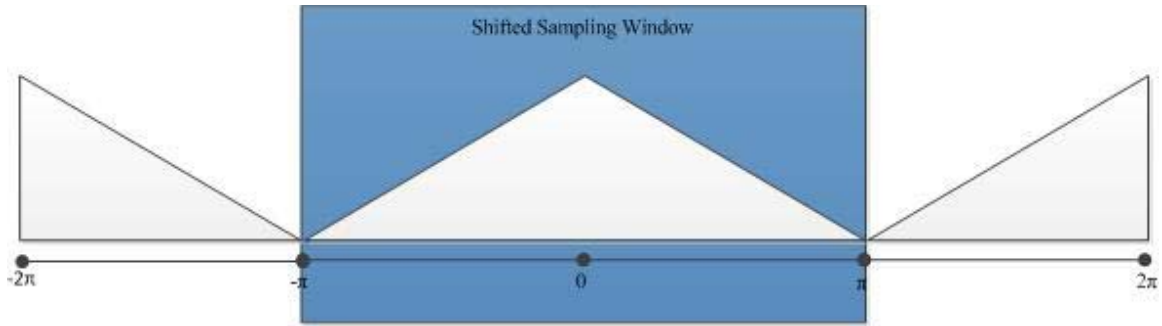


Figure 32 Sampling window for a shifted 1D DFT function

This type of shift can also be applied to the 2D DFT. One way to perform this shift is to split the image into four sub-images and shift each sub-image's location. Figure 33 shows one means of shifting the sub-images to center the low-frequency components. Each sub-image is labeled 1 through 4. Sub-image 1 is copied to the location of sub-image 4. Likewise, sub-image 4 is copied to the location of sub-image 1. The same type of sub-image shift is deployed with sub-image 2 and sub-image 3. Swapping the four sub images has the same effect as shifting the sampling window that was shown in Figure 32.

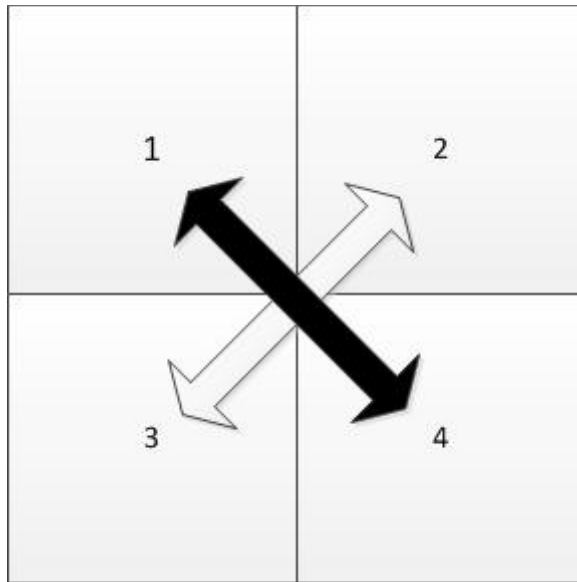


Figure 33 This figure illustrates a DFT shift function.

Figure 34 depicts a DFT shift that has been applied to the image in Figure 30. The magnitude of this image was uniformly scaled to make the spikes in the image more apparent. The large low-frequency values are located at the center of the image. This visualization shows that the low-frequency components in each of the four sub-images are similar. Though the phase is different in each of the sub-images, their magnitudes are very similar. When low-frequency components were represented in the corners of the image, this was less apparent. In the unshifted DFT image, the axes went from 0 to  $2\pi$ . After applying the shift, the axes now go from  $-\pi$  to  $\pi$ .

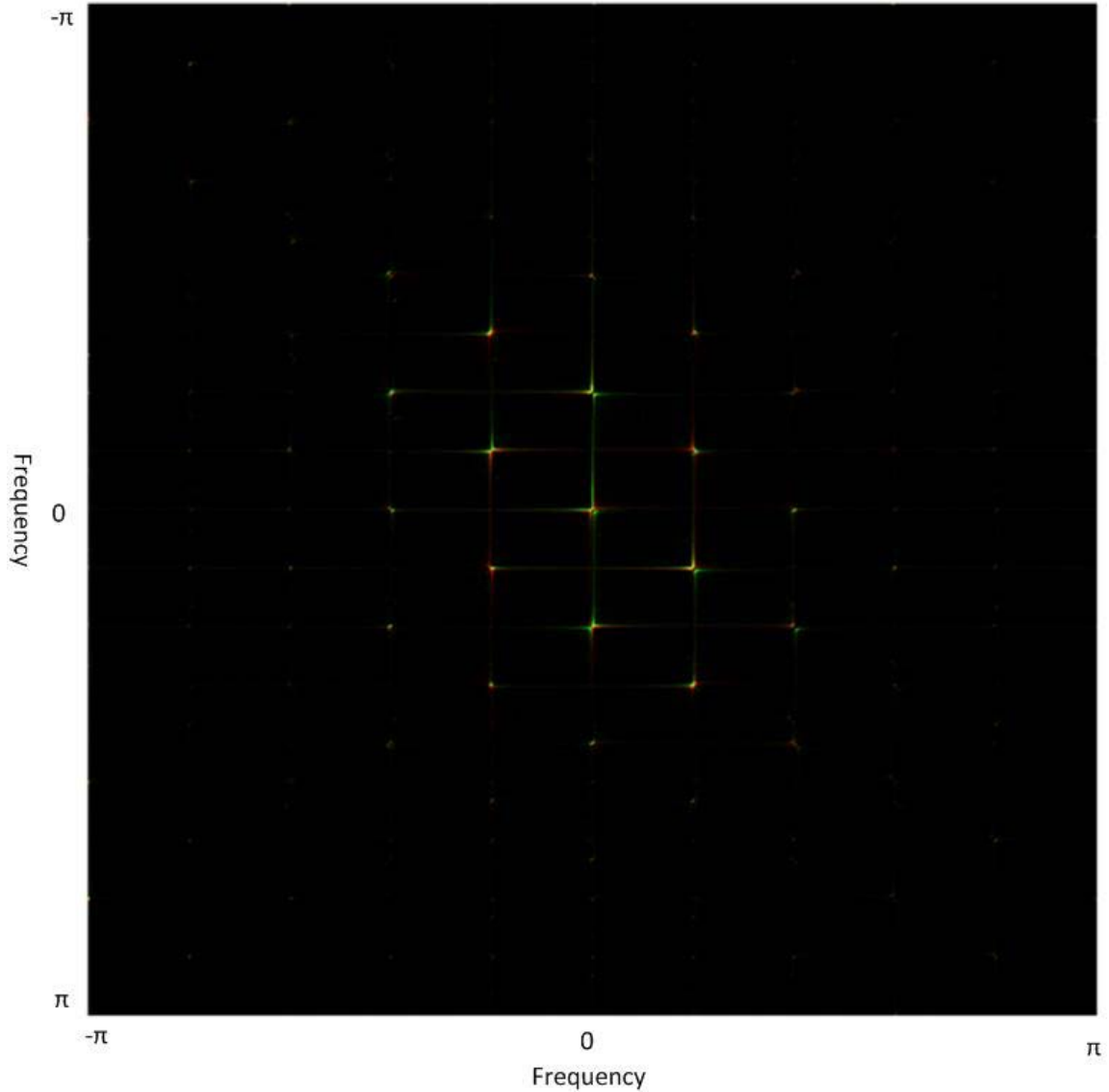


Figure 34 This image shows the DFT of the target image.

The pixels at the center of Figure 30 represent the components with the lowest frequency in the image. The value stored in these pixels represents a constant value component that exists in every pixel of the original spatial image. The constant value component can be thought of as a DC offset for the whole image. The color coding of these pixels near the center of the image shows a phase response that is expected with the DC values of an image. The DC component pixels in sub-image 2 and sub-image 3 have

the same phase. Likewise, the DC component pixels in sub-image 1 and sub-image 4 have the same phase. This expected behavior can be explained by the 2D DFT function that is described in Equation 26. The exponential portion of the transfer function dictates the phase. To make this apparent, think of an input image where the upper left corner of the image is labeled  $(m = 0, n = 0)$  and the lower right pixel is labeled  $(m = M, n = N)$ . The exponential term in the transfer function is equal at these two corners and is equal to  $\exp(0)$ . This is a purely real phase. This means that their phases are equal, since the exponential term is the only thing that contributes to phase. Likewise, the exponential term will have the same value for the upper right corner and the lower left corner. The exponential term has a value of  $\exp(j\pi)$  for these two corners. This is a purely imaginary phase. After the shift, these phase values are shown at the center of the image. Figure 35 shows an enlarged view of the center of Figure 34. The phase pattern is more apparent in this enlarged view. Sub-image 1 and sub-image 4 are primarily composed of pixels with real phases, which are indicated by their yellow coloring. Sub-image 2 and sub-image 3 are primarily composed of pixels with imaginary phases, which are indicated by their red coloring.

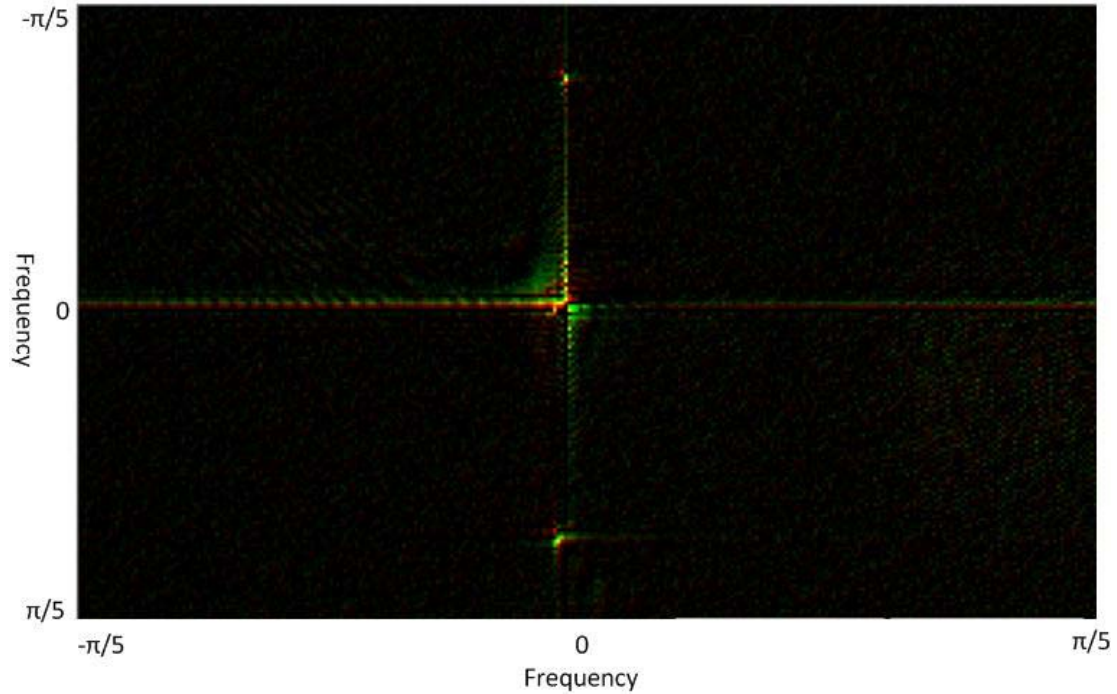


Figure 35 This figure shows an enlarged view of the center of a DFT image.

In this type of phase pattern, there are 4 pixels that form a square and each pixel's phase type matches the phase type of its diagonal neighbor. This pattern is visible in locations other than the center of Figure 34. This phase pattern can identify the center of the sub-aperture images described in the section titled "A Microlens Array Light Field Camera." This property was discovered by inspecting sample images. This property was present in all images that were inspected. Equation 12 describes how to theoretically demodulate the light field from a raw image captured by a printed mask light field camera. A microlens array light field camera performs a similar type of light field modulation. It is difficult to identify the carrier frequencies that allow for each slice of light field to be modulated. Finding exact carrier frequency locations is specific to each camera. Because the lenses in a microlens array are very small in relation to the tolerances of the mounting bracket, the alignment of the microlens array and the imaging sensor differs from camera to camera.

The DC component of each light field slice is centered on the carrier frequency for the slice. It is relatively easy to identify the locations of each light field slice's DC component. This can be used to determine the corresponding carrier frequencies for each light field slice. The carrier frequency for each slice has the following two properties:

- 1) The magnitude of the 4 pixels that form the carrier frequency is much larger than the magnitude of the pixels in the surrounding area
- 2) Each of the 4 pixels at the center of the carrier frequency has a phase sign that matches the phase sign of its diagonal neighbor

This means that diagonal pixels are either primarily real values or primarily imaginary values. The quickest way to identify the carrier frequency locations for the test image was to visually inspect the pixels and record the locations where these identities occurred. The locations for the carrier frequency in Figure 34 are at (985, 505), (1640, 505), (2295, 505), (657, 694), (1312, 694), (1968, 694), (2623, 695), (985, 883), (1640, 883), (2295, 884), (657, 1072), (1312, 1072), (1967, 1072), (2623, 1073), (984, 1262), (1640, 1261), (2295, 1262), (657, 1449), (1312, 1450), (1967, 1451), (2622, 1451), (329, 1639), (987, 1639), (1639, 1640), (2295, 1641), (2951, 1641), (657, 1828), (1311, 1828), (1967, 1829), (2622, 1829), (984, 2017), (1639, 2018), (2295, 2018), (657, 2207), (1312, 2207), (1967, 2207), (2622, 2209), (984, 2396), (1639, 2396), (2294, 2397), (656, 2585), (1311, 2585), (1967, 2585), (2622, 2586), (984, 2774), (1639, 2275), and (2294, 2775). It is possible to automatically identify these locations through automatic processes such as a convolution filter [44, 45]. Because this has to be performed only once for the camera, it was quicker to manually determine the locations.

Figure 36 depicts each sub-aperture image identified by the red boxes within the 2D DFT of the original image. The sub-aperture images were discovered to be a fixed size. This was determined by examining the minimum vertical and horizontal distances between carrier frequencies. The distance was fairly constant. However, it occasionally

shifted by a couple of pixels. The minimum distances were used for the slice size for the following two reasons:

- 1) It is much easier to work with light field slices when each slice is the same size

- 2) Using the smallest image size will reduce the risk of aliasing between slices

High-frequency components are not important in this application. It is better to exclude them than to introduce aliasing. Most of the image's content is contained in low frequency components. Therefore, they do not affect the final output image. For this image, the slices were fixed at a size of  $320 \times 320$  pixels.



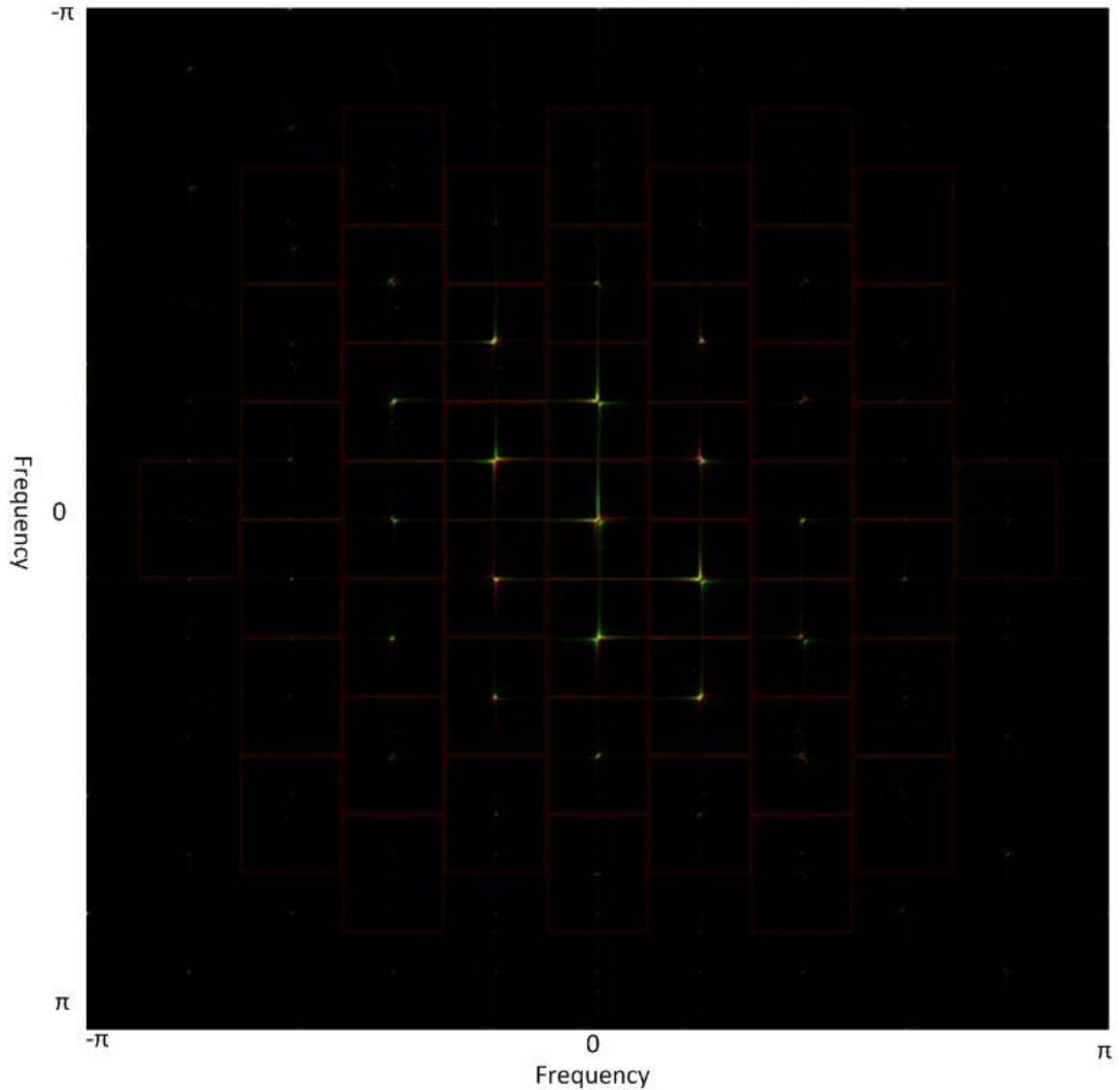


Figure 36 This figure shows each sub-image that is identified in the DFT image.

A 2D inverse digital Fourier transform (IDFT) is the inverse to the 2D DFT. A 2D IDTF transforms a frequency domain image back to a visual image. Equation 27 is a 2D IDFT function, and it is the inverse of the 2D DFT defined in Equation 26 [44]. In the 2D IDFT equation,  $F(u, v)$  is the frequency domain image produced by Equation 26. The exponential term in the 2D IDFT is the complex conjugate of the exponential in the forward 2D DFT. The return value  $f(m, n)$  is a spatial image. The 2D IDFT is able to

return the original image  $f(m, n)$  that was passed into the 2D DFT. Any differences in the images are the result of round off and truncation errors. When the DC components of an image are located at the center of the image, a DFT shift needs to be performed before applying the inverse transform. The DFT shift described in a previous section is a self-inverting function. By applying the same shift operation, the 4 DC pixels will be moved to the appropriate locations at the outer corners of the image.

$$f(m, n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[ 2\pi j \left( \frac{mu}{M} + \frac{nv}{N} \right) \right]$$

27

Each light field slice shown in Figure 36 is a frequency domain image. Applying the 2D IDFT function to each of the slices produces a spatial representation of the modulated light field. Figure 37 shows multiple slices of the light field after it was transformed in a manner that returned it to the spatial domain. Only the center  $5 \times 5$  squares from Figure 36 are being displayed. The 2D IDFT function was performed on each  $320 \times 320$  pixel sub-aperture image. A DFT shift was performed on each sub-aperture image to place the 4 DC pixels at the outer corners of the sub-image. The sub-images were then placed in the same orientation as their slices were located in the frequency domain image. This was done to illustrate how each slice appears spatially at each modulation frequency location. The 4D light field is not easily visualized, and this is one way to present multiple slices.

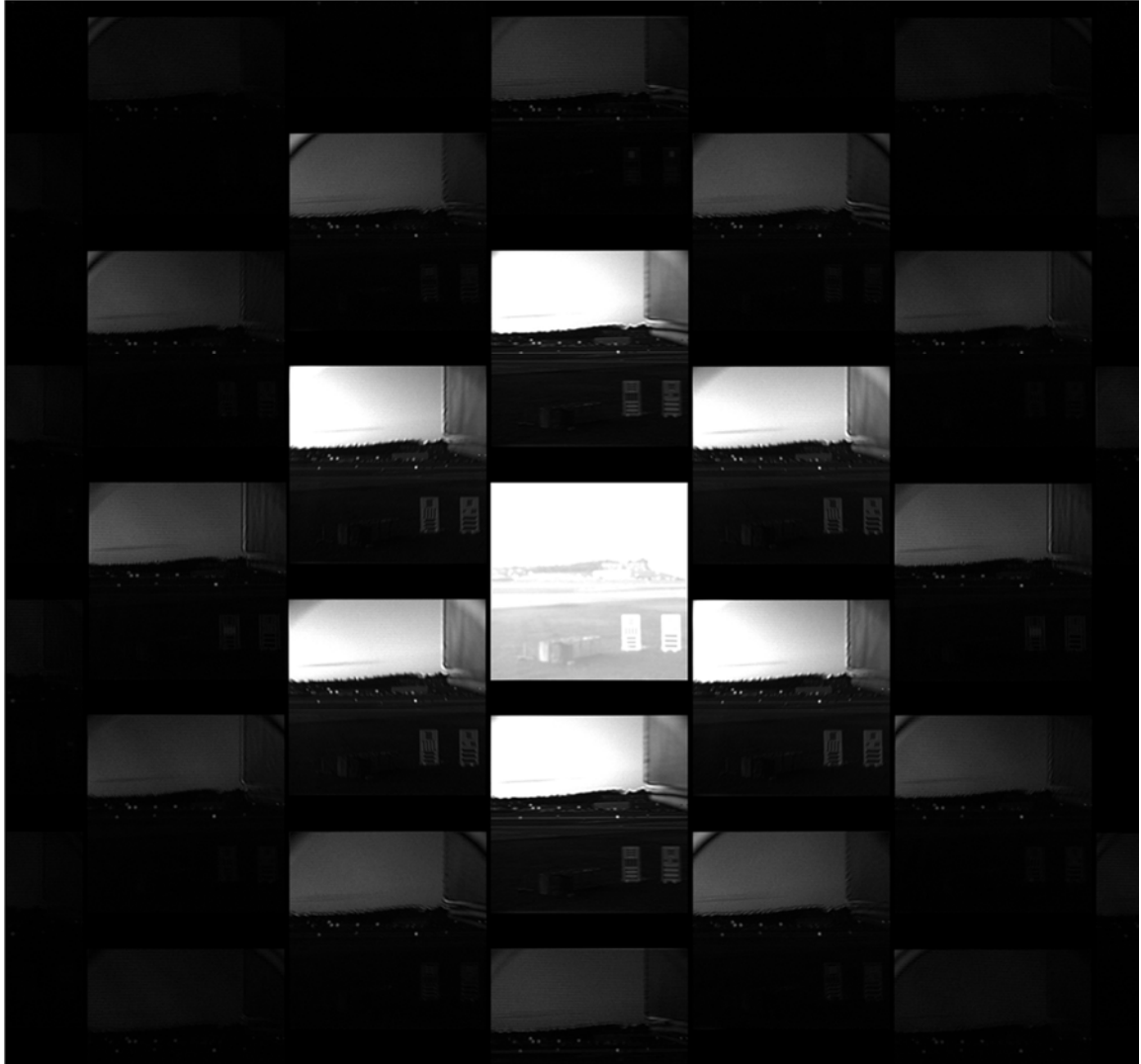


Figure 37 This figure depicts spatial representation of light-field slices shown at their modulation frequency locations.

These light field slices show a number of anticipated and expected characteristics. On average, the center slice of the light field is brighter than the surrounding slices. The center image represents light that is coming through the center sub-aperture of the main lens. In most cases, this is the primary component of a captured light field. Light that passed through the center of the main lens has been less affected by spherical aberrations on the microlens. In the original image, which is shown in Figure 27, small groups of

pixels formed circular images. Generally, the pixels at the center of these circles were brighter than those at the outer edges.

### Constructing the Light Field Equation from Fourier Slices

The modulated slices shown in Figure 37 can be mapped into the 4D light field equation. The general form of the 4D light field equation is shown in Equation 8. Equation 8 is indexed over the variables  $x$ ,  $y$ ,  $u$ , and  $v$ . The variables  $u$  and  $v$  express the location of the modulation impulses within the DFT image. The variables  $x$  and  $y$  show the individual pixel values at each of these locations.

This mapping of the light field can be thought of as:

- 1) Taking the 2D DFT of an image and identifying the locations of all of the modulation pulses, which are stored in the variables  $u$  and  $v$
- 2) Taking the 2D IDFT for the sub-image around the pulse for every modulation pulse  $u$  and  $v$ , a value which is then stored in the variables  $x$  and  $y$

This method of constructing the 4D light field is different than the method used in the Matlab Light Field Toolbox. The Matlab Light Field Toolbox only examines at the raw spatial images and the circles generated by each of the microlenses. The Matlab Light-Field Toolbox populates the light field in the following fashion. The variables  $x$  and  $y$  are used to identify the locations of each circle's center. The variables  $u$  and  $v$  are used to identify the location of a pixel within a circle, which requires a well-working and robust circle detection to be performed on every camera. The circle detection method does not work with images produced by printed mask light field cameras because they do not produce these circles. One benefit of the Fourier approach is that it works well for both types of cameras. Additionally, the calibration process can be applied to actual images, and those images are not required to have been calibrated beforehand.

## Digital Refocusing

Ng [2] took the imaging equation and modified it to allow the focal plane of an image to be moved. In a traditional single-lens camera, this is not possible post-image acquisition. However, with a 4D light field camera, it is possible to refocus an image post-acquisition because a 4D light field contains information about the paths light rays take within the camera. A traditional 2D image does not contain this information. Post-acquisition digital refocusing is made possible by manipulating the imaging equation shown in Equation 8. The original focal length  $F$  of a captured image can be altered by applying a scaling factor  $\alpha$ . When  $\alpha$  equals 1, the camera's original geometry is returned. When  $\alpha$  is less than 1, the new imaging plane will have a focal length that is shorter than the original geometry. Likewise, when  $\alpha$  is greater than 1, the new imaging plane will have a focal length that is longer than the original geometry.

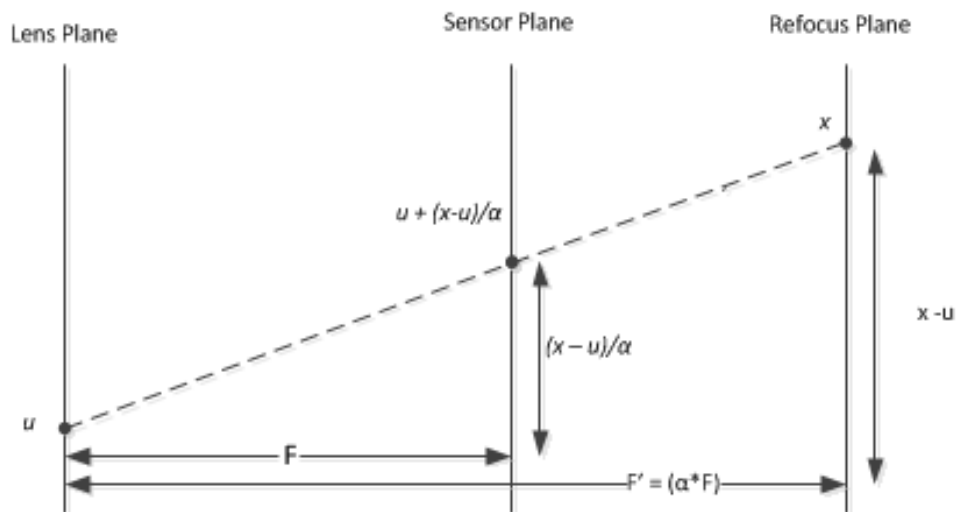


Figure 38 This image shows the geometry required for refocusing light field images.

The imaging equation is defined in terms of a light ray's intersection with the main lens and the imaging sensor. In the original image, the focal plane is located at the imaging sensor. By a change of variables, Equation 8 can be rewritten for any focal plane determined by  $\alpha$ . The  $x$ -coordinate in the refocused imaging plane maps to a coordinate on the sensor plane as shown in Figure 38. Likewise, a similar mapping can be performed for the  $y$ -coordinate. The application of this parameterization produces Equation 28.

$$E(x, y) = \iint L\left(u\left(1 - \frac{1}{\alpha}\right) + \frac{x}{\alpha}, v\left(1 - \frac{1}{\alpha}\right) + \frac{y}{\alpha}, u, v\right) dudv$$

28

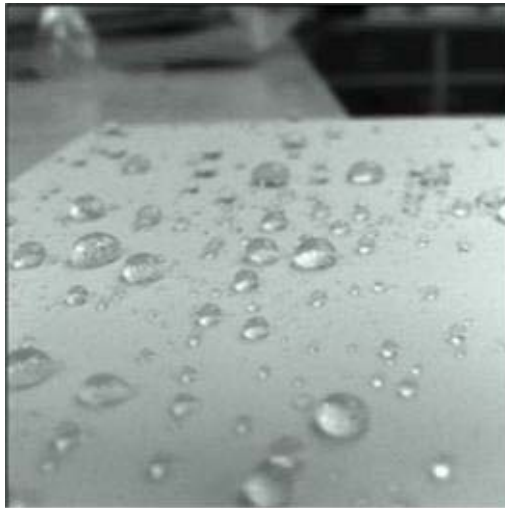
The captured 4D light field is only valid for discrete values of  $x$ ,  $y$ ,  $u$ , and  $v$ . In previous sections, the methods of 4D light field acquisition resulted in a discrete sampling of the light field. For discrete samples, Equation 28 was generalized and rewritten as a summation. This is shown in Equation 29. With the use of this equation, a 2D image is returned and can have an altered focal plane that is defined by  $\alpha$ .

$$E(x, y) = \sum_u \sum_v L\left(u\left(1 - \frac{1}{\alpha}\right) + \frac{x}{\alpha}, v\left(1 - \frac{1}{\alpha}\right) + \frac{y}{\alpha}, u, v\right)$$

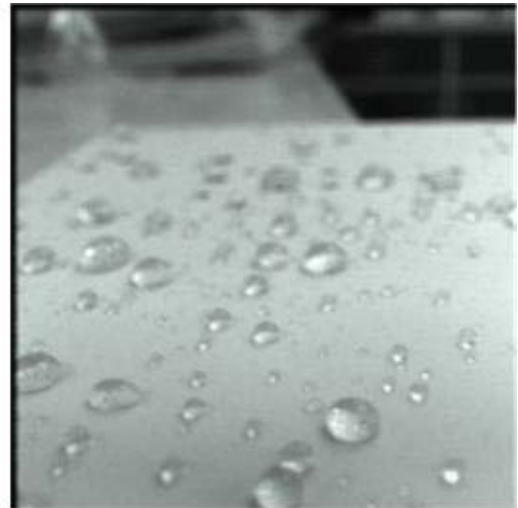
29

Equation 29 was applied to a 4D light field image. This image was an example image in the test dataset of the Matlab Light Field Toolbox [35]. The application of Equation 29 to the example image produced the images in Figure 35. When the  $\alpha$  value was 1, the image returned was a 2D image with the original focal length. When the  $\alpha$  value was 0.8, the focal plane was 80% of the original image, and the focus of the image was moved into the foreground. For images with an  $\alpha$  value greater than 1, the focal

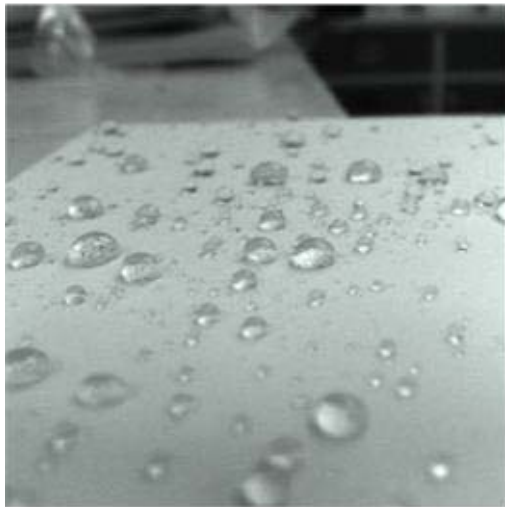
plane of the image was extended, and the focus of the image was pushed into the background. These images were produced from a single captured image and the focus did not need to be determined until post-acquisition.



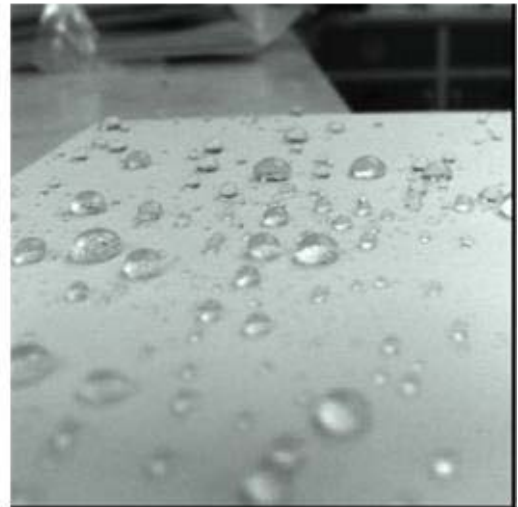
$\alpha = 1$



$\alpha = .8$



$\alpha = 1.2$



$\alpha = 2$

Figure 39 This figure depicts four images captured from a light-field camera and then refocused at different values of  $\alpha$ . The raw light field image used to produce the four images was part of the Light Field Toolbox for Matlab[35].

## PROCESSING LIGHT FIELD IMAGES FROM PRINTED MASK LIGHT FIELD CAMERA

This chapter explains the processing of images captured by the two printed mask light field camera prototypes. The general concepts that are required to understand how the images were processed are explained. First, this chapter will provide an overview of sampling theory and how it applies to imaging sensors. This theory has implications for transferring analog system concepts to the digitally captured images that the cameras produce. Additionally, the theory behind the diffraction patterns produced by the printed mask will be discussed. With these concepts understood, analysis is performed on the images captured by the two prototype cameras. Two images from each camera are examined to show their light field modulations. After these methods of processing have been performed, images captured by printed mask light field cameras can be treated in a similar fashion to images captured with a microlens camera.

### Sampling Theory for Imaging Sensors

An imaging sensor takes a digital sample of the light it is exposed to inside the camera. The sampling rate is defined by the distance between individual pixels. When it comes to 1D signals, the Nyquist-Shannon sampling theorem states that the DFT can only represent frequencies that are less than half of sampling frequency [37]. This can be extended to multidimensional signals using the Petersen-Middleton theorem [46]. In the case of the image sensor this dissertation chooses to look at the sampling rate as separable. The vertical sampling rate is produced when the number of vertical pixels is divided by the vertical size of the chip in millimeters. A horizontal sampling rate is calculated in the same way. Two prototype masks were installed into two different FujiFilm PinePix S5200s. The sampling rates for these camera's imaging sensors are



show in Table 10. These sampling rates will be used to examine the bandwidth of the modulated images that the camera captures.

Table 10 This table shows the sensor sampling rate for a prototype light field camera.

Direction	Pixels Along Axes	Size of Axes (mm)	Sampling Rate (1/mm)
Horizontal	3688	5.76	640
Vertical	2776	4.29	647

### Lens Effects on Printed Mask Modulation

The chapter “Testing Printed Masks” described the use of a spatially coherent light to test whether the printed masks met the design considerations. This test is not necessary once a mask has been placed into a camera because a converging lens is able to perform a 2D analog Fourier transform. With a thin lens, an image that is present on the projected plane of the lens surface will result in a Fourier transform of the image, which will be produced on the focal point of the lens [26]. This property is also invertible. When a mask is placed at the focal point of the lens, a Fourier transform of the image will be produced on the project plane.

Placing a printed mask directly at the focal point of a lens is feasible. However, it is not desirable. Figure 40 shows the geometry of a printed mask that has been placed inside a camera. It was assumed that the lens of the camera acted like a thin lens. The approximation of a thin lens was sufficient for the aims of this test. The imaging sensor was a distance  $f$  away from the focal point of the main lens. The printed mask was placed

a distance  $d$  away from the imaging sensor. This setup will be used to explain the diffraction caused by the printed mask inside the camera.

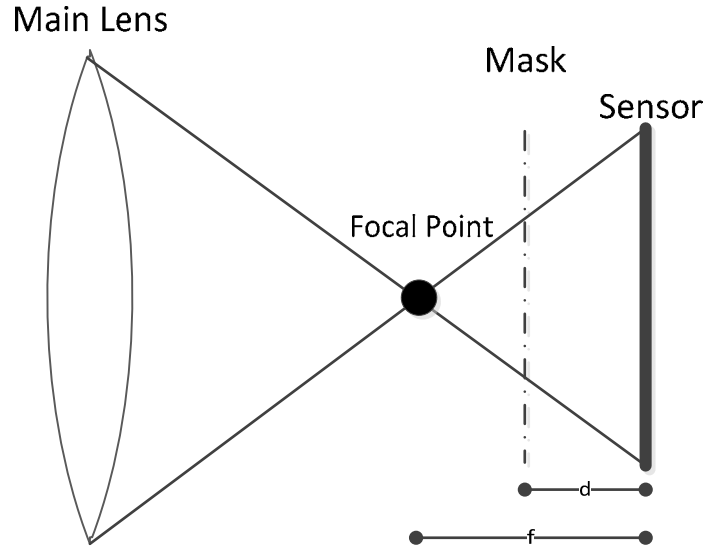


Figure 40 This figure shows the mask and main lens geometry to explain the diffraction pattern spreading.

There was an issue regarding the placement of the printed mask on the exact focal point of the lens. This is where  $f$  equals  $d$ . The carrier frequencies of the cosines used to produce the printed mask are relatively large. Mask 2 has 2 harmonics of cosine waves. The 2 carrier frequencies for the mask are 22.76 (1/mm) and 11.38 (1/mm). The Fourier transform of the mask will result in delta functions in which zeros are located distances of 22.76 mm and 11.38 mm from the center of the imaging sensor. The prototype camera's imaging sensor is only 5.76 mm wide and 4.29 mm tall. There is a delta function located at the center of the sensor that represents the DC components of the mask. However, the other modulated slices do not land on the image sensor.

Fortunately, changing the geometry so that  $d$  does not equal  $f$  allows the delta functions to land on the imaging sensor. The Fresnel diffraction integral can be modified to take into account a setup with a thin lens and a printed mask. Equation 30 describes the Fresnel diffraction for the geometry shown in Figure 40 [26]. This equation is for a normal incident, monochromatic plane wave that is back lighting the lens with an amplitude of  $A$ . The mask is placed a distance  $d$  away from the imaging sensor. The distance between the focal point of the lens and the image plane is  $f$ . The term in front of the double integral is a quadratic phase factor for the light. This term can be ignored at this time. The term  $t(\xi, \eta)$  is the Fourier transform of the printed mask. The term  $P$  is the Fourier transform of an aperture function of the lens. This corresponds to the  $f$  stop on the lens. The exponential term inside of the integral matches the Fresnel diffraction integral that was previously described. However, its rate of change has been scaled by a factor of  $1/d$ .

$$U_f(u, v) = \frac{A * \exp\left(j \frac{k}{2d} (u^2 + v^2)\right)}{j\lambda d} * \frac{f}{d} * \iint_{-\infty}^{\infty} t(\xi, \eta) P\left(\xi \frac{f}{d}, \eta \frac{f}{d}\right) \exp\left(-j \frac{2\pi}{\lambda d} (u\xi + v\eta)\right) d\xi d\eta$$

30

A couple of items are produced by this diffraction integral. These items are useful because they allow for a better understanding of how the light field is modulated with the printed mask. First, the function  $P$  can be replaced with the light field that is entering the camera. The light field is modulated with the printed mask  $t$ . The scale of the modulated image is governed by the ratio  $f/d$ . As  $d$  becomes smaller, the resulting modulated image decreases in size. When  $d$ 's value gets closer to  $f$ , the modulated image gets larger and cannot fit onto the imaging sensor. As previously described, the sampling rate of the imaging sensor limits the recovery of modulated frequencies above a certain

threshold. For example, it is necessary for the ratio  $f/d$  to be in a range where the modulated slice falls onto the camera sensor. The next two sections will apply these properties to the cameras that have Mask 2 and Mask 4 installed. The results are then compared with images captured by the cameras.

### Mask 2 Processing

The function that generated Mask 2 is a separable function. Approaching the horizontal and vertical components separately is necessary because the sampling rate of the imaging sensor is different in the two directions. In the horizontal direction, there are 2 cosine waves that make up the mask. One of these waves has a period of 10 pixels and the other has a period of 5 pixels. Because the printed mask was printed with a pixel density of 113.8 pixels per millimeter, the frequency of these cosine waves is 11.38 and 22.76 respectively. When the Fourier transform of the mask is taken, the resulting impulses have horizontal distances of 0 mm, 11.38 mm, or 22.76 mm away from the center. These frequencies will be referred to as  $f_0$ ,  $f_1$ , and  $f_2$ , respectively. These distances will not correspond to an imaging sensor that is only 5.76 mm wide. A similar approach can be applied to cosine waves in the vertical direction.

Manipulating the  $f/d$  ratio can cause these points to land onto the imaging sensor. The distance  $d$  is fixed because this is the distance between the mask and the imaging sensor. This distance is approximately 1.01 mm. The  $f$  distance can be changed because the focal length of the camera is adjustable. The lens of the camera is magnetically controlled and can be adjusted to have a focal distance between 6 mm and 63 mm. Increasing the value of  $f$  will cause the impulses of the transform to land closer to the center of the imaging sensor.

Table 11 shows the effect of focal length on the location of the impulses that are caused by the printed mask. The first column shows the settings of the camera's focal length when an image is captured. The second column and the third column show where

frequencies  $f1$  and  $f2$  will land as a result of the  $f/d$  ratio. The  $f0$  location remains the same because it is at the zero frequency. There are different pixel densities in the vertical and horizontal directions of the imaging sensor. The fourth, fifth, sixth, and seventh columns show the locations of  $f1$  and  $f2$  in terms of their distances from the sensor's center in both the horizontal and vertical direction.

Table 11 This table illustrates the modulation frequency locations for Mask 2.

Focal length	$f'1$ (mm)	$f'2$ (mm)	Horizontal Pixels $f1$	Horizontal Pixels $f2$	Vertical Pixels $f1$	Vertical Pixels $f2$
6	1.91	3.83	1226	2453	1240	2479
8	1.44	2.87	920	1840	930	1859
10	1.15	2.30	736	1472	744	1487
11	1.04	2.09	669	1338	676	1352
12	0.96	1.91	613	1227	620	1240
13	0.88	1.77	566	1132	572	1144
15	0.77	1.53	490	981	496	992
19	0.60	1.21	387	775	391	783
25	0.46	0.92	294	589	297	595
29	0.40	0.79	253	507	256	513
33	0.35	0.70	223	446	225	451
39	0.29	0.59	188	377	191	381
52	0.22	0.44	141	283	143	286

To test whether adjusting the focal length of the main lens would shift the locations of the modulation frequencies, two images were examined. The first image was captured with a focal length of 12 mm. The second image was captured with a focal length of 33 mm. The expected locations of the modulation impulses were determined by the values in Table 11. The locations were defined in terms of the distances between the

impulses and the center of the DFT image. After the expected locations were determined, they were used to identify the locations of the actual impulses.

On occasion, the impulses were difficult to see because the center of the DFT image has a much larger value than the other modulation impulses. One way to identify the locations of these impulses was to perform a Laplacian operator on the image. A Laplacian operator takes the second derivative of the image and is typically used in edge detection [44]. When the Laplacian operator of the DFT image was taken, the impulses became more distinct compared to pixels in a neighboring area. The reason for this is because the impulses have a very large gradient in relation to the pixels that surround them. Figure 41 shows one of these Laplacian images. The image was captured when the focal length of the camera was set to 12 mm. When looking closely at this image, the center of each square has a much brighter pixel value than the surrounding pixels. This is the location of the modulation impulse. The red boxes show the location of each of the slices of the light field.

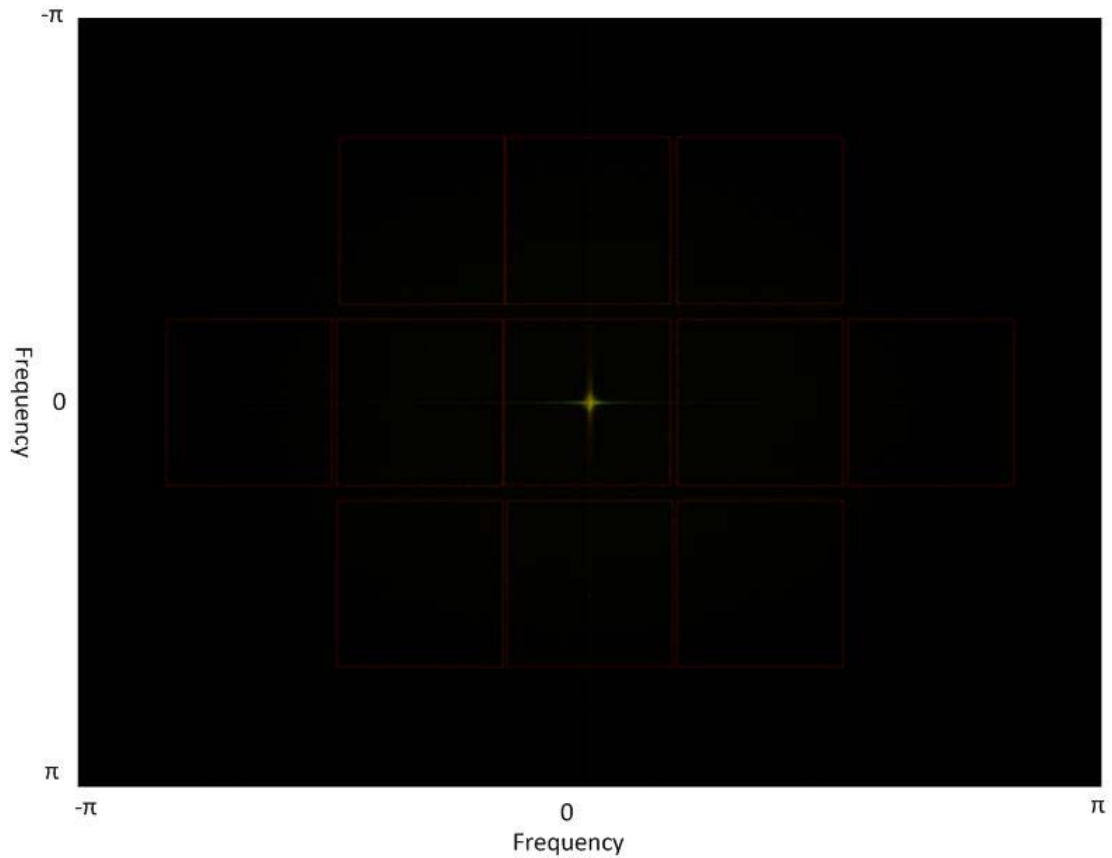


Figure 41 This figure shows a Laplacian operator of a DFT image. The original image had a focal length of 12 mm. The modulation mask was Mask 2. The slices of the light field are segmented by the red boxes.

Taking the locations identified in Figure 41, the slices of the light field are shown in the DFT image in Figure 42. The red boxes identify each slice of the light field and the center of each box is the location of the modulation impulse. The impulses have the same phase pattern that the microlens array light field camera produced.

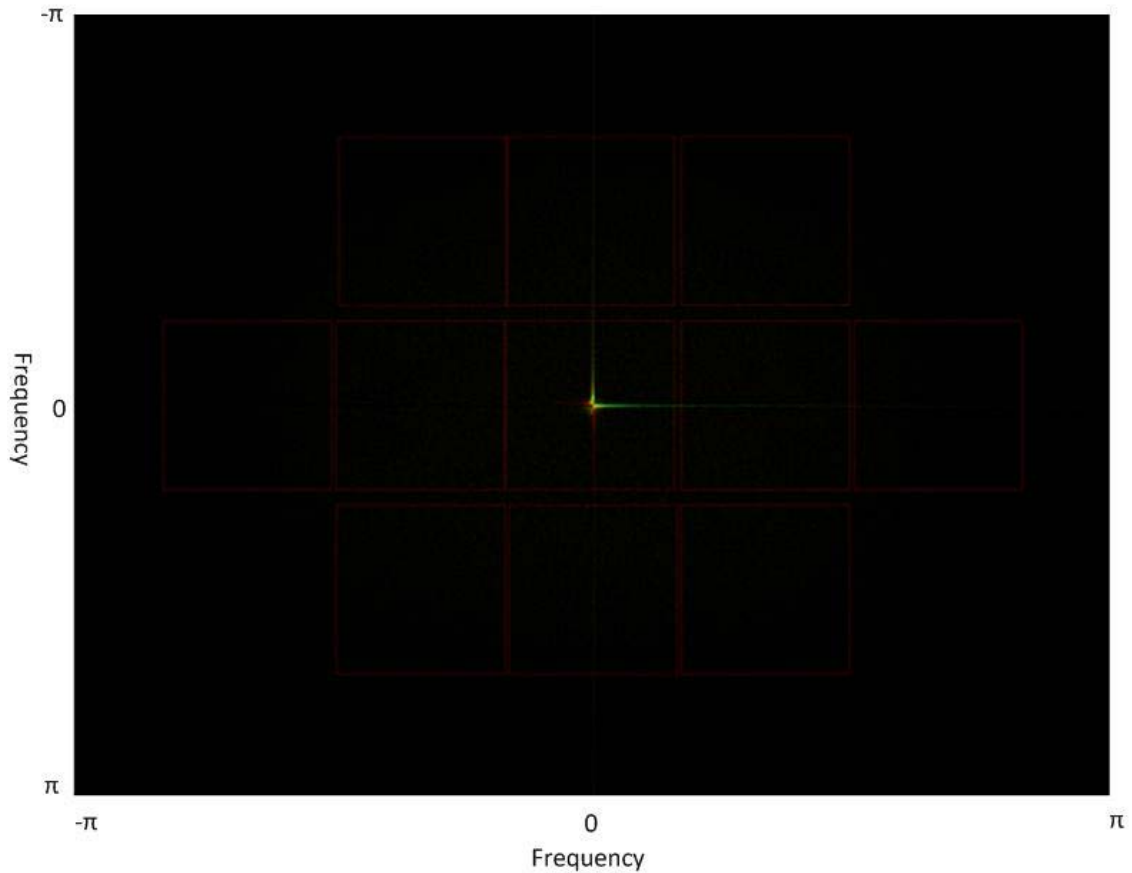


Figure 42 This figure is a DFT of an image that was captured with a focal length of 12 mm. This image was modulated with Mask 2. The slices of the light field are segmented by the red boxes.

Taking the 2D IDFT of each slice of the light field will result in a spatial image of the captured scene. Figure 43 shows each of the light field slices placed at the locations of their modulation frequencies. As made evident by the images visual properties, the center slices contain a majority of the light field. Additionally, this is the same type of image that the Lytro camera produces.



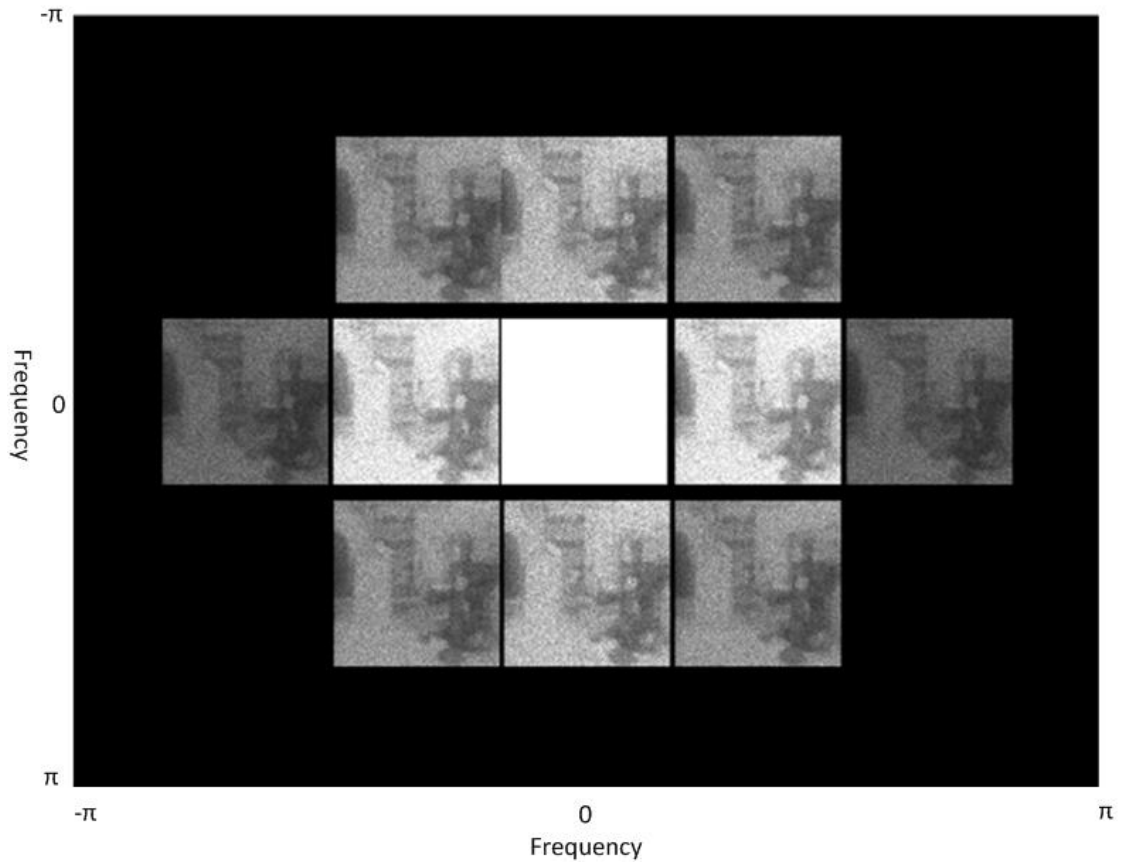


Figure 43 This figure shows the spatial representation of light field slices arranged according to their modulation frequency locations. This image was captured by the camera with Mask 2 installed. The camera had a focal length of 12 mm.

Table 12 shows the expected and actual location of all of the modulation frequency impulses. The value is expressed as pixels. It is apparent in the test images that vertical modulation frequencies occasionally crossed with an unintended horizontal modulation frequency. When testing with a laser pointer, this unintended cross-modulation was not apparent. When Mask 2 was installed, some of these cross-modulations became apparent. When the column's main pulse has a value of one, the horizontal and vertical frequencies modulate as intended. When the value is zero, this represents a case in which a horizontal frequency modulated with an unintended vertical

frequency. Though these slices were not originally intended, they are still useful. In this case, they do not introduce aliasing because the slices do not overlap. The raw image from the camera is shown in Figure 44. Objects in the raw image are apparent in the light field slices. The scene in this image is the same scene as all test images captured with printed mask light field camera in this dissertation.



Figure 44 This is a raw image from a test camera with Mask 2 installed. The image was captured with a focal length of 13mm.

Table 12 This table contains the modulation frequency locations for Mask 2 when the images are captured with a focal length of 12 mm.

Main Pulse	Expected $i$	Expected $j$	Actual $i$	Actual $j$	Offset $i$	Offset $j$	Offset Distance
1	1388	1844	1388	1844	0	0	0
1	768	1231	734	1240	-34	9	35
1	768	2464	732	2457	-36	-7	37
1	2008	1231	2041	1232	33	1	33
1	2008	2464	2041	2457	33	-7	34
0	768	1844	733	1843	-35	-1	35
0	2008	1844	2042	1844	34	0	34
0	1388	1231	1388	1229	0	-2	2
0	1388	604	1388	615	0	11	11
0	1388	2464	1388	2458	0	-6	6
0	1388	3084	1388	3073	0	-11	11

The second and third columns show the expected modulation frequency locations. Table 11 shows that the horizontal modulation frequencies are 0, 613, and 1227. The vertical modulation frequencies are 0, 620, and 1240. The expected values can be determined by adding or subtracting these distances from the center of the image. The actual values are stored in columns four and five. These values were determined by examining the Figure 38 and Figure 39. The sixth and seventh column display the differences between the expected and actual values. The last column displays the Euclidean distance between the modulation frequency and the intended location.

A majority of the offset error occurred in the vertical direction. The error is constant and is roughly comprised of 33 pixels. There are a number of possible reasons as to why this occurred. For example, the vertical sampling rate of the imaging sensor may not be accurate. Additionally, the  $d$  and  $f$  ratio may be not accurate in the vertical

direction. Regardless of the actual cause of the offset, the center of the light field slices was located. This type of calibration can compensate for the aforementioned errors.

A similar approach can be applied to an image that was captured when the focal length was set to 33 mm. Table 11 depicts the modulation frequencies of an image that was taken with the focal length of 33 mm. These vertical modulation frequencies are 0, 191, and 881 pixels. The horizontal modulation frequencies are 0, 188, and 377. Table 13 shows the expected modulation frequency locations. These are arranged in the same manner that was used in Table 12. Expected and measured locations are provided along with the differences between the values. In this case, there appears to be a constant scaling factor that is being applied to all modulation frequencies. This is most likely the result of the imprecise measurements of the  $d$  and  $f$  ratio. If a different  $d/f$  ratio is used, the modulation frequencies might correspond to one another more accurately.

Table 13 This tables shows the modulation frequency locations for Mask 2 when images are captured with a focal length of 33 mm.

Main Pulse	Expected $i$	Expected $j$	Actual $i$	Actual $j$	Offset $i$	Offset $j$	Offset Distance
1	1388	1844	1388	1844	0	0	0
1	1163	1621	1201	1660	38	39	55
1	1163	2069	1201	2026	38	-43	58
1	1613	1621	1569	1667	-44	46	64
1	1613	2069	1570	2029	-43	-40	59
1	937	1393	1019	1476	82	83	116
1	937	2295	1022	2211	85	-84	119
1	1839	1393	1757	1472	-82	79	114
1	1839	2295	1754	2212	-85	-83	118
0	1163	1844	1200	1843	37	-1	37
0	937	1844	1016	1843	79	-1	79
0	1613	1844	1573	1844	-40	0	40
0	1839	1844	1756	1843	-83	-1	83
0	1388	1621	1388	1659	0	38	38
0	1388	1393	1388	1475	0	82	82
0	1388	2069	1388	2028	0	-41	41
0	1388	2295	1388	2212	0	-83	83

Overall, the results are very positive, in that the  $d/f$  ratio shifts the location of the modulation frequencies in an expected manor. This is more apparent when looking at Figure 45 and Figure 46. The first figure gives the Laplacian of the DFT image. The second figure depicts the DFT of the image. The slices of the light field are segmented by the red boxes. The slices on this image are smaller than 12 mm. This is to be expected. With a larger focal length, the modulation pattern decreases in size. This should hold for all values of  $f$  that the camera possesses.

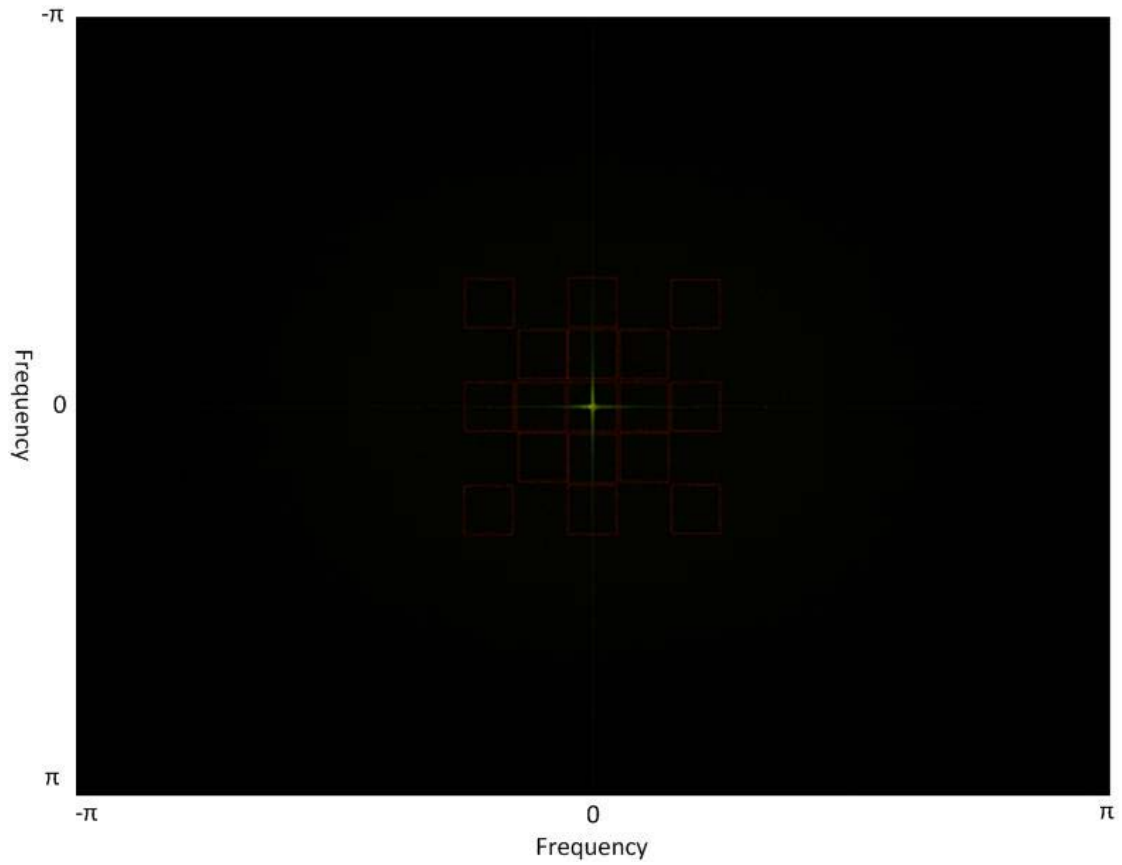


Figure 45 This figure illustrates the Laplacian of a DFT image. The original image had a focal length of 33 mm. The modulation mask that was used was Mask 2. The slices of the light field are segmented by the red boxes.

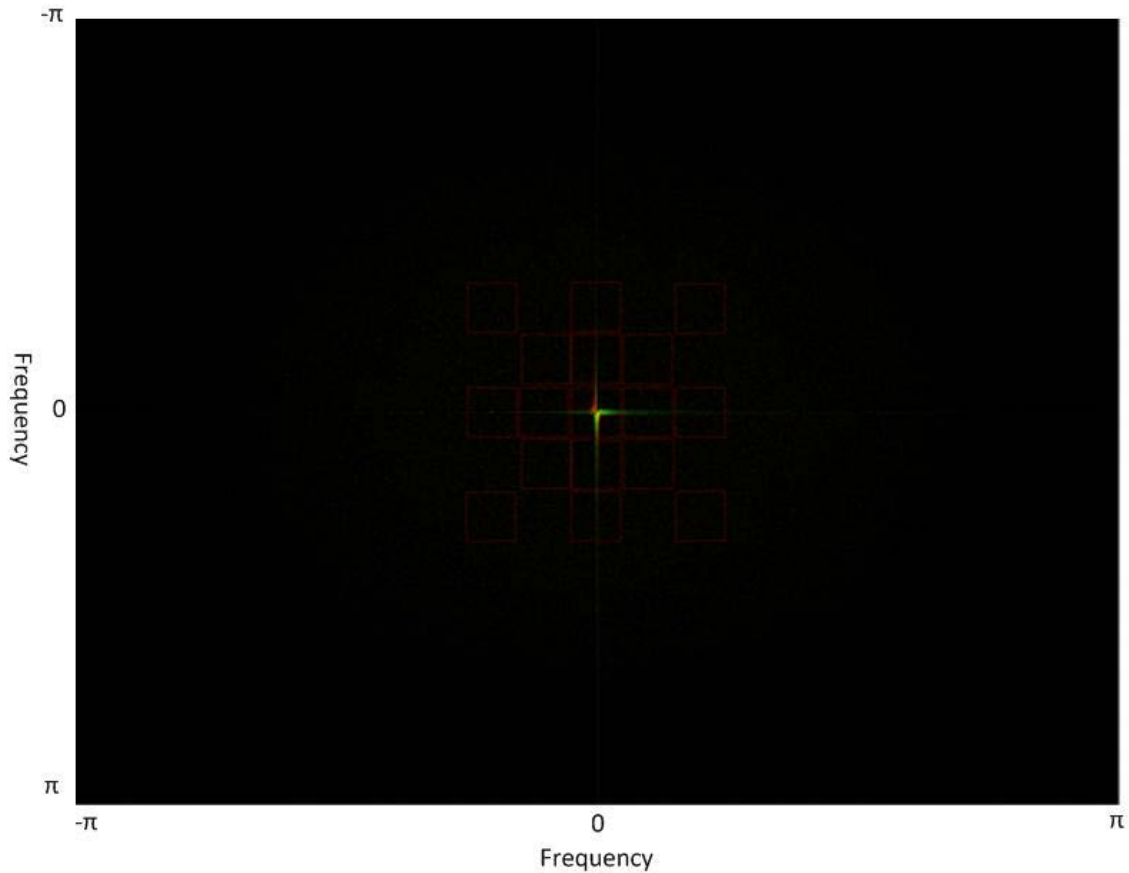


Figure 46 This figure depicts a DFT of an image that was captured with a focal length of 33 mm. The modulation mask that was used was Mask 2. The slices of the light field are segmented by the red boxes.

The amount of the mask that is utilized can be defined by the amount of modulation frequencies that fall onto the imaging sensor. When all the modulation frequencies fall onto the imaging sensor and are spread out as far as possible, the mask is being fully utilized. These two test images show that the ratio of  $d/f$  governs the location of the impulses. Additionally, they show that the measurements for the camera are reasonably accurate. As depicted by the figures in Table 11, a focal length of 13 mm is the shortest possible focal length in which all of the slices will land on the imaging sensor. When a focal length is over 13 mm, all slices land on the imaging sensor. However, there are two negative effects that are associated with focal lengths larger than

13 mm. The first negative effect is that the full spectral range of the image is not being utilized. The second negative effect is that the size of each slice is not as large as it could be, because of aliasing. When  $f$  increases, the size of each slice decreases. These effects are readily apparent in Figure 47. Figure 42 depicts an image captured with a focal length of 33 mm. Each slice is smaller than the image captured at 12 mm. Additionally, the high-frequency components of the image are unused. Though this focal length is not optimal, the image is still usable.

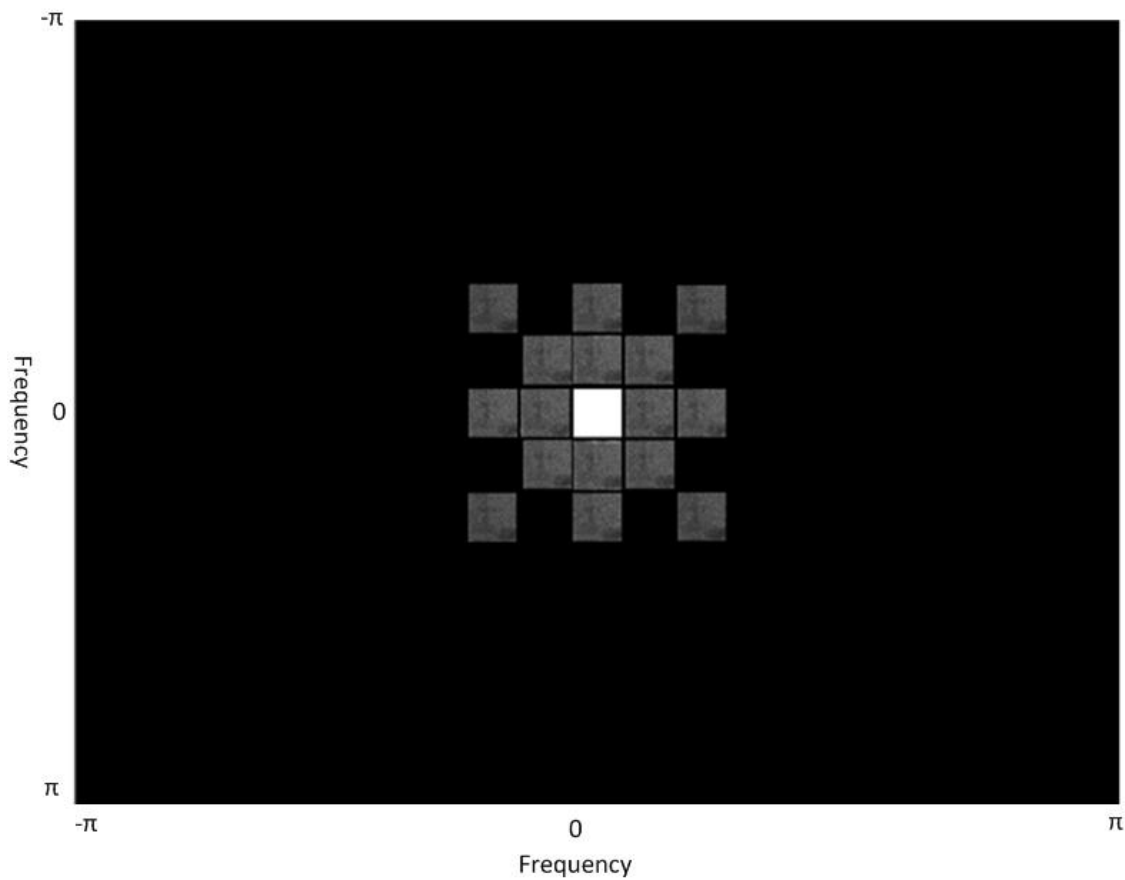


Figure 47 This figure shows the spatial representation of light-field slices arranged according to their modulation frequency locations. This was captured with a focal length of 33 mm. The modulation mask that was installed was Mask 2.



## Mask 4 Processing

The same type of process that was applied to Mask 2 can also be applied to Mask 4. Mask 4 had multiple harmonics of cosine waves in both horizontal and vertical directions. This section will state the expected locations of the modulation impulses produced by the cosine waves. Two images captured at different focal lengths are examined to see whether they produce the expected light-field modulation effects. This method of light field modulation works with two different mask designs, which indicates that this method could be generalized to other masks.

Table 14 contains the expected modulation frequencies for Mask 4. Only the first 6 harmonics are presented for the mask. The cosine waves that produced these harmonics have periods of 0.625, 1.25, 2.5, 5, 10, and 20 pixels. This corresponds to  $f_1$ ,  $f_2$ ,  $f_3$ ,  $f_4$ ,  $f_5$ , and  $f_6$  respectively. The top row of the table shows the focal length of the main lens when an image is captured. The length of the lens is measured in millimeters. The second, third, fourth, fifth, sixth, and seventh rows show the locations of each of the modulation frequencies. These locations are defined by the distance between the modulation frequency and the center of the image in millimeters. When the length of the lens is increased, these impulses become located closer to the center of the image.

The capturing camera possesses different pixel densities in the horizontal and vertical directions. The remaining rows of the table express the locations of the modulation pulses in pixels. In this table, H denotes the horizontal direction and V denotes the vertical direction. Using the values laid out in this table, it is possible to determine both the horizontal and vertical locations of the modulation pulses. For many lens lengths, these impulses will not fall onto the imaging sensor. Because of this property, the mask does not have an ideal design. However, it is able to show the effects of the modulation. Additionally, some of these modulation frequencies can cause

aliasing. Regardless of this aliasing, the modulation properties of this type of mask are apparent.

Table 14 This table contains the modulation frequency locations for Mask 4.

<b>Lens</b>	<b>6</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>15</b>	<b>19</b>	<b>25</b>	<b>33</b>	<b>39</b>	<b>52</b>
<b>f1</b>	30.65	18.39	16.72	15.33	14.15	12.26	9.68	7.36	5.57	4.72	3.54
<b>f2</b>	15.33	9.20	8.36	7.66	7.07	6.13	4.84	3.68	2.79	2.36	1.77
<b>f3</b>	7.66	4.60	4.18	3.83	3.54	3.07	2.42	1.84	1.39	1.18	0.88
<b>f4</b>	3.83	2.30	2.09	1.92	1.77	1.53	1.21	0.92	0.70	0.59	0.44
<b>f5</b>	1.92	1.15	1.04	0.96	0.88	0.77	0.60	0.46	0.35	0.29	0.22
<b>f6</b>	0.96	0.57	0.52	0.48	0.44	0.38	0.30	0.23	0.17	0.15	0.11
<b>H f1</b>	19625	11775	10704	9812	9058	7850	6197	4710	3568	3019	2264
<b>V f1</b>	19833	11900	10818	9917	9154	7933	6263	4760	3606	3051	2288
<b>H f2</b>	9812	5887	5352	4906	4529	3925	3099	2355	1784	1510	1132
<b>V f2</b>	9917	5950	5409	4958	4577	3967	3132	2380	1803	1526	1144
<b>H f3</b>	4906	2944	2676	2453	2264	1962	1549	1177	892	755	566
<b>V f3</b>	4958	2975	2705	2479	2288	1983	1566	1190	902	763	572
<b>H f4</b>	2453	1472	1338	1227	1132	981	775	589	446	377	283
<b>V f4</b>	2479	1487	1352	1240	1144	992	783	595	451	381	286
<b>H f5</b>	1227	736	669	613	566	491	387	294	223	189	142
<b>V f5</b>	1240	744	676	620	572	496	391	297	225	191	143
<b>H f6</b>	613	368	335	307	283	245	194	147	112	94	71
<b>V f6</b>	620	372	338	310	286	248	196	149	113	95	72

The ideal focal length that was used to capture images with Mask 2 was approximately 13 mm. This focal length was chosen for the first test image for Mask 4. With this length, the slices of the light field are a reasonable size for comparison. With this focal length, 4 harmonics should be visible in both the horizontal and vertical directions. The harmonics that land on the imaging sensor are  $f0$ ,  $f4$ ,  $f5$ , and  $f6$ . Figure

48 shows the Laplacian operator of the 2D DFT image. However, in the image produced by the Laplacian operator, the impulses were not as distinct as they were in the previous mask. Because these impulses were fairly difficult to identify, only the slices that were cross-modulated with  $f_0$  are identified with red squares.

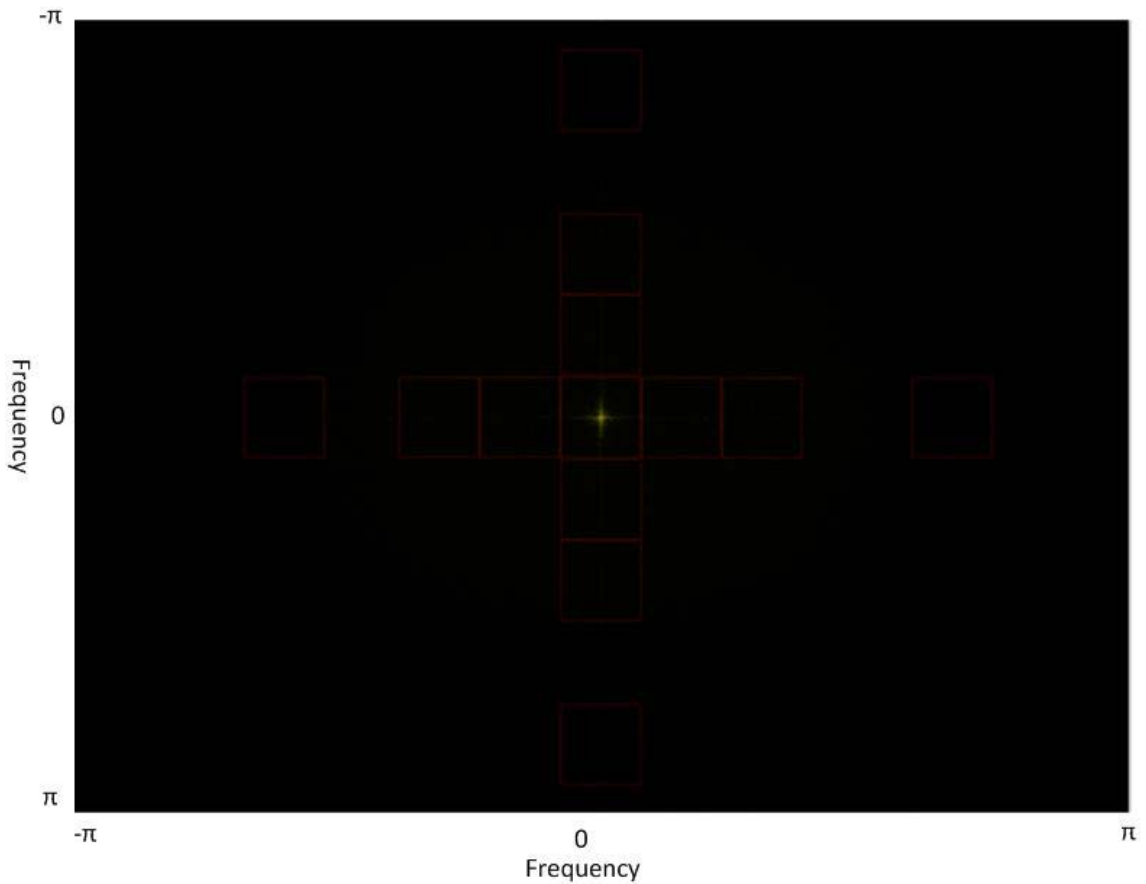


Figure 48 This figure shows a Laplacian operator of a DFT image. The original image had a focal length of 13 mm. The modulation mask used was Mask 4. The slices of the light field are segmented by the red boxes.

Figure 49 depicts the DFT of the image with the slices identified. These slices possess the same phase pattern as the slices captured by the Lytro camera and the camera with Mask 2 installed. This image shows that the small impulses have the correct phase

pattern at all of the cross-frequency locations. However, there is a fair amount of noise around those areas. For this reason, the slices are not identified in Figure 44.

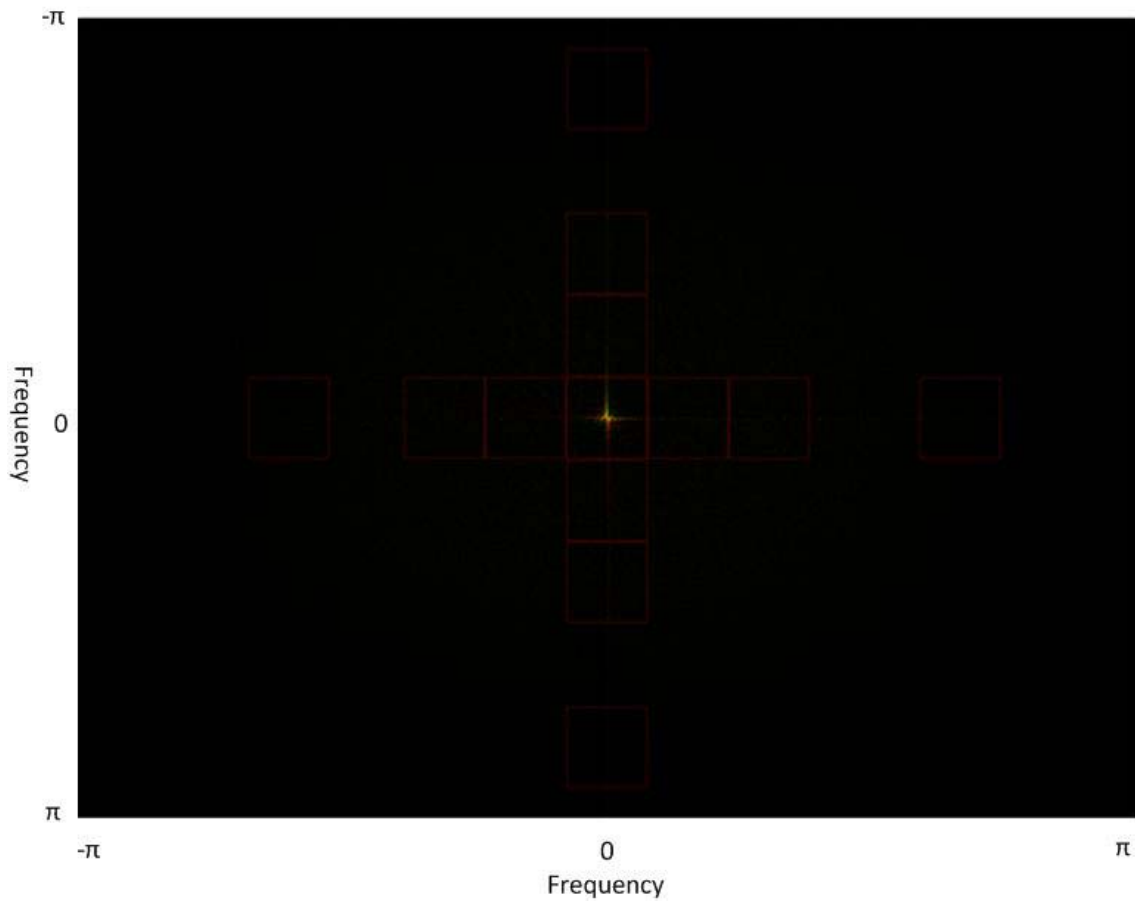


Figure 49 This is a DFT of an image that was captured with a focal length of 13 mm. This image was modulated with Mask 4. The slices of the light field are segmented by the red boxes

Figure 50 shows the 2D IDFT of the slices segmented in Figure 44. This image was taken indoors. The objects in the image were approximately 10 m down range. The objects are apparent in each of the slices. This was expected. The center slice has the strongest signal. In this rendering, the image has a gain value that caused the center slice to be oversaturated. This gain value allows for the other slices to be visible. The outer

slices have lower-values of intensity than the inner slices. This was expected because a similar response was observed in the Fresnel diffraction test of the printed mask.

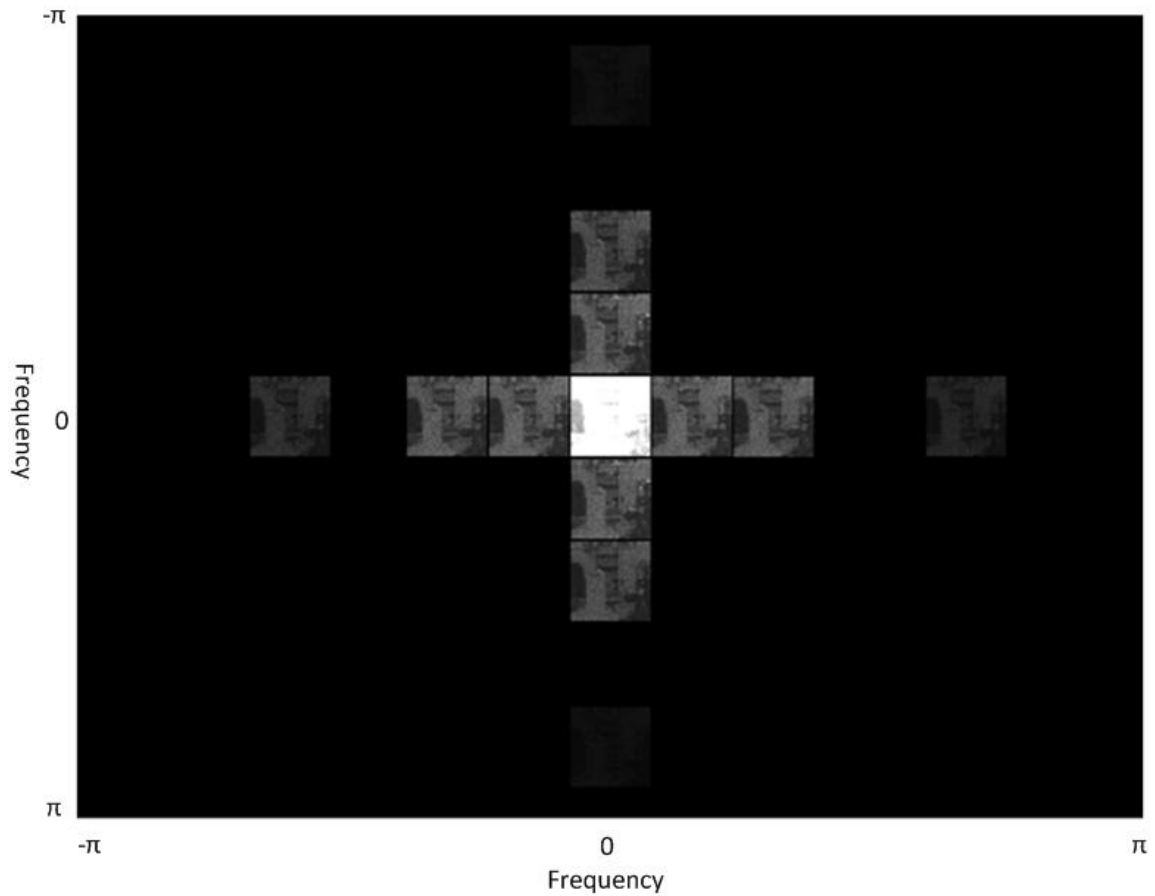


Figure 50 This image shows the spatial representation of light-field slices arranged according to their modulation frequency locations. This is an image that was captured with Mask 4. The focal length was 13 mm.

Because each of these slices can produce a 2D IDFT image, it can be inferred that the center of each slice is located correctly. Table 15 depicts both the expected and actual pixel locations in the center of each slice. This was only applied to the slices identified in Figure 43, Figure 44, and Figure 45. The expected locations were determined by using the values in Table 14. Table 14 contains the modulation frequency locations for an image captured with a focal length of 13 mm. In this table,  $i$  is the pixel

index in the vertical direction and  $j$  is the pixel index in the horizontal direction. The fifth and sixth columns contain the offsets in each direction. In the horizontal direction, the modulation frequencies were only a few pixels off. In the vertical direction, the frequencies were slightly off. However, these are reasonable values given the clarity of the slices. The values in these channels are reasonably low. The last column shows the Euclidian distance from the expected value. All of these values are reasonably low.

Table 15 This table shows the printed Mask 4 modulation frequency locations for images captured with a focal length of 13 mm.

Expected $i$	Expected $j$	Actual $i$	Actual $j$	Offset $i$	Offset $j$	Offset Distance
1844	1388	1844	1388	0	0	0
2130	1388	2126	1388	-4	0	4
2416	1388	2408	1389	-8	1	8
2988	1388	3073	1388	85	0	85
1558	1388	1561	1387	3	-1	3
1272	1388	1278	1387	6	-1	6
700	1388	736	1388	36	0	36
1844	1105	1843	1101	-1	-4	4
1844	822	1843	816	-1	-6	6
1844	1671	1845	1674	1	3	3
1844	1954	1844	1958	0	4	4
1844	2520	1843	2531	-1	11	11
1844	256	1844	244	0	-12	12

The second image that was examined for Mask 4 was captured with a focal length of 6 mm. This caused large slices of the light field to be captured. However, fewer slices landed on the imaging sensor. Horizontally, only  $f_0$ ,  $f_5$ , and  $f_6$  were able to be rendered. Vertically, only  $f_0$  and  $f_6$  landed on the imaging sensor. Figure 51 is the Laplacian of the

2D DFT image, and Figure 52 is the 2D DFT of the captured image. At this focal length, the modulation impulses are apparent. Each of the slices are approximately  $600 \times 600$  pixels without aliasing. In both of the figures, the slices are denoted by the red rectangles.

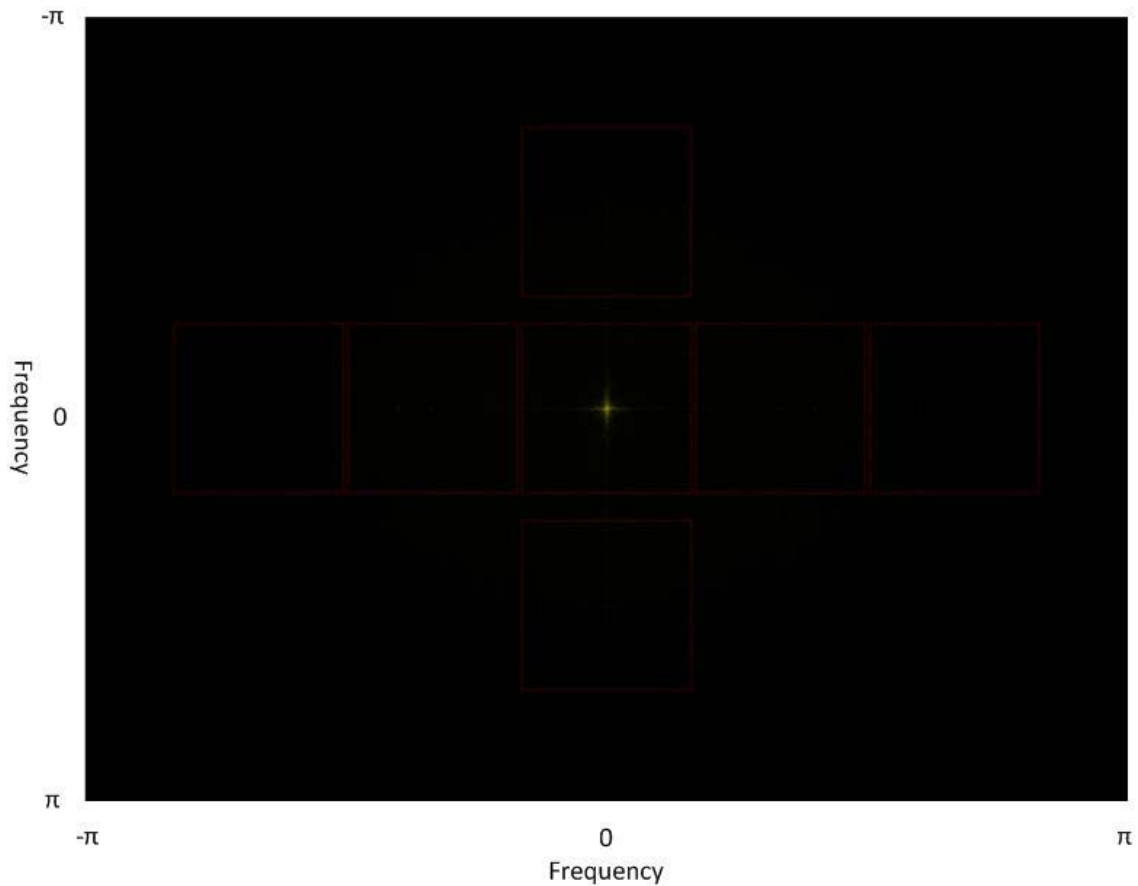


Figure 51 This image shows a DFT image after a Laplacian operator has been applied. The original image had a focal length of 6 mm and the modulation mask was Mask 4. The slices of the light field are segmented by the red boxes

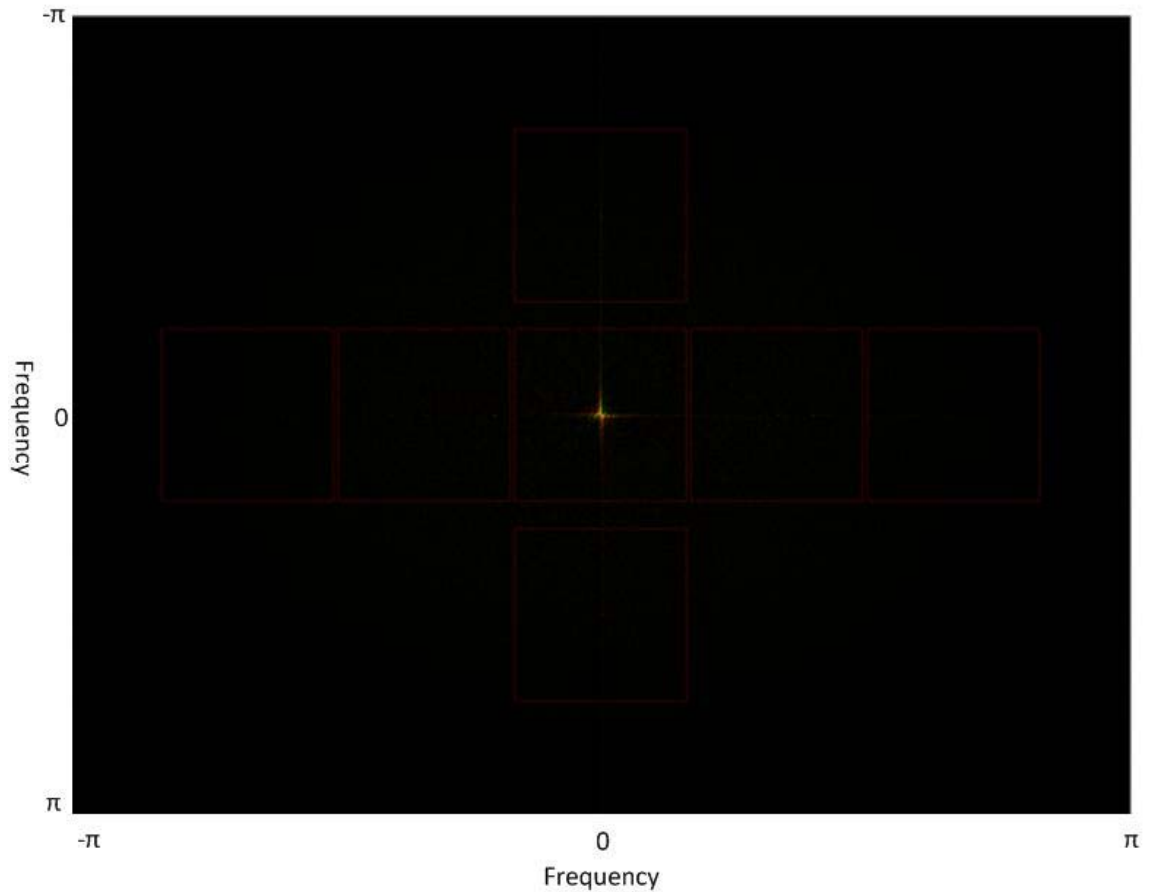


Figure 52 This is a DFT of an image captured with a focal length of 6 mm. The image was modulated with Mask 4. The slices of the light field are segmented by the red boxes.

Figure 53 depicts a 2D IDFT of each slice. These are arranged according to their modulation locations within the whole image. In the figure, it is very apparent that the slices are correctly demodulated. The difference between each slice is marginal. The center slice is brighter than its neighbors. Because there are no large-banked surfaces in the image, this was expected. Steep angles would be more apparent in the slices. There is slightly more noise present in these images when compared to images produced by the Lytro camera. However, images produced by a Lytro camera and a printed-mask camera with Mask 4 installed exhibit the same modulation properties.



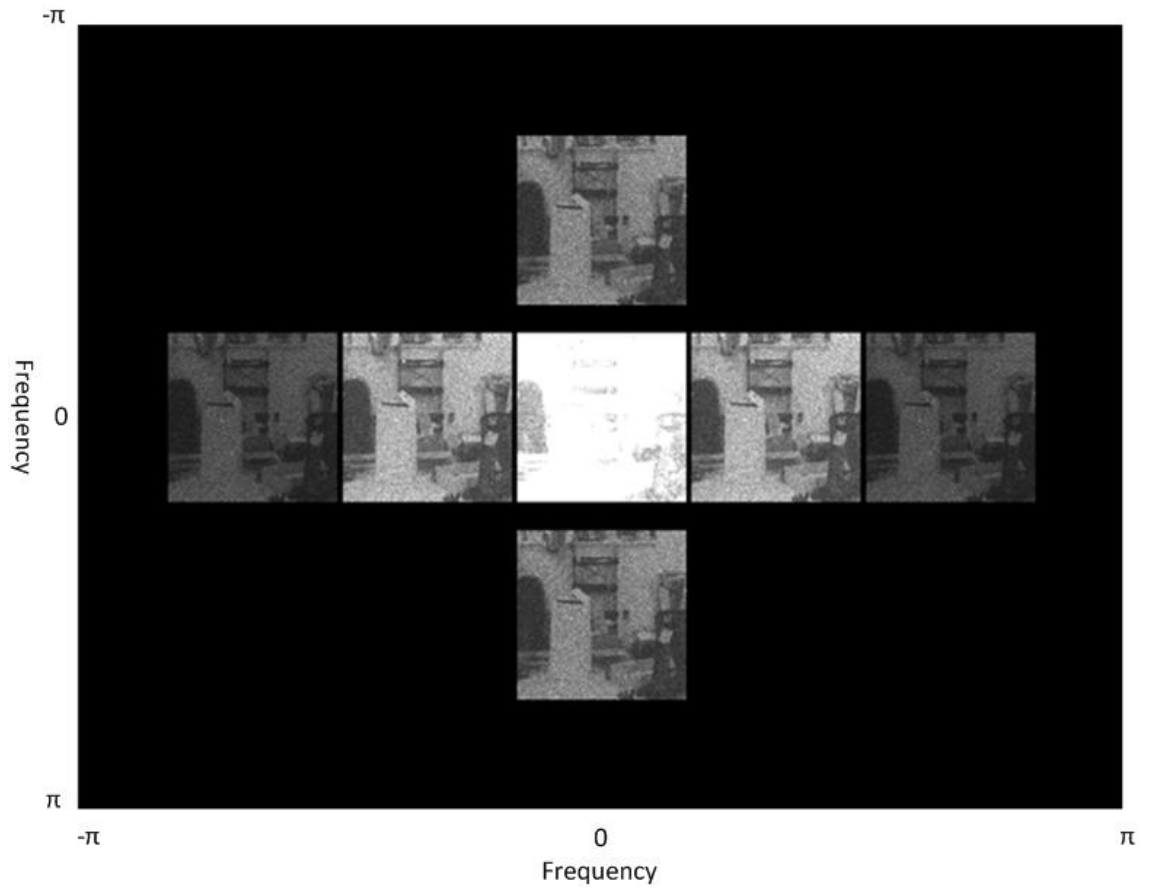


Figure 53 This image shows spatial representation of light-field slices arranged according to their at their modulation frequency locations. This image was captured at focal length of 6 mm by a camera with Mask 4 installed.

Table 16 shows the expected and actual locations of the modulation frequencies. The expected locations of the modulation frequencies can be derived from the values in the first and second columns of Table 14. The third and fourth columns show the actual values. The fifth and six columns contain the differences between the expected and actual values. Each offset is small and symmetric. The last column contains the total offset distances in terms of their Euclidian distance. In all cases, the offset distances are relatively small.

Table 16 This table contains the modulation frequency locations for images modulated by Mask 4. These images were captured with a focal length of 6 mm.

Expected $i$	Expected $j$	Actual $i$	Actual $j$	Offset $i$	Offset $j$	Offset Distance
1844	1388	1844	1388	0	0	0
2457	1388	2458	1388	1	0	1
3071	1388	3073	1388	2	0	2
1231	1388	1229	1388	-2	0	2
617	1388	614	1388	-3	0	3
1844	768	1843	693	-1	-75	75
1844	2008	1844	2083	0	75	75

### Summary of Results

Overall, the results from the analysis of images captured by the printed mask cameras were positive. For example, it was discovered that the approximate location of modulation frequencies can be determined by comparing the expected modulation pattern, the placement of the mask within the camera, and the focal length of the camera's main lens. Additionally, it was discovered that it is possible to find the actual locations of the modulation pulses using the expected locations of modulation pulses, the Laplacian operator of the DFT image, and the DFT image. Significantly, the modulation pulses exhibited the same phase pattern as those produced by a Lytro camera.

Though a Lytro camera produces slices that are cleaner, the test did show that a printed mask camera is able to do the same type of modulation. Additionally, the printed mask offers more flexibility than a microlens array. This trade off in terms of the quality might be mitigated by using a camera that was designed specifically for printed masks. For example, if the design of a camera's lens took this application into consideration, a printed mask camera may be able to produce cleaner slices.

### 3D DEPTH ESTIMATION

When multiple images are captured from different locations and are pointed at the same object can produce a 3D image [44]. These methods require global knowledge of where the cameras are located in reference to each other. When this spatial relationship has been established, corresponding points between the images can be used to map a 3D world. This setup is very useful in computer vision applications [22]. However, the requirement of global knowledge and the need for multiple images of the same area is an impediment to the creation of multiple-view configurations. A light field camera simplifies this process by allowing a user to detect depth from a single capture frame [1]. Global positioning of the camera is necessary only when a user desires to stitch multiple frames together.

An additional way to produce a 3D image is to use stereoscopic imaging. Stereoscopic images mimic the human vision system, in that they consist of two cameras pointed at the same location [47]. When one camera's image is shown to a person's left eye and the other camera's image is shown to a person's right eye, the person will perceive a 3D image. For this visualization to work, the cameras need to possess the same optical properties. For example, if the two cameras have adjustable focal lengths, both must be adjusted equally each time a change is made. For this method to work well, the focus of the images needs to be set at the time of capture. A light field camera enhances the functionality of this method, in that it allows for focusing to occur post-processing.

Similar to human eyes, cameras set close together are unable to produce 3D images of objects at considerable distances. However, cameras can be placed far apart,

which allows them to capture a 3D image of distant objects. This technique is referred to as hyperstereo photography [47]. In some past applications of hyper stereo photography, the distance between the cameras was nearly 15 m. In these instances, the captured image of the distant object is similar in appearance to a regular stereoscopic image. The images produced by a hyperstereo setup are useful when they are employed in wearable display systems, specifically a head-mounted display. Head-mounted displays that use hyperstereo images allow for users to perceive depth of 3D objects better than with the unaided human visual system [48]. One issue with hyper stereo photography, however, is that it sacrifices 3D imaging and clarity in the foreground [47]. Single lens light field cameras will not be able to do the same kind of 3D imaging and depth estimation for long distance objects as hyper stereo systems can, but the light field camera is better suited for shorter-range distances, and might be an option for low-flying vehicles or indoor navigation.

A 4-D light field image contains much of the same information as a stereoscopic image. Additionally, light field cameras can produce depth estimations. This is possible by using the displacement and object scaling of objects between sub-images [1]. By determining the optical properties of the capturing camera, the displacement between sub-images, and the object scaling between sub-images, it is possible to use elementary optics to infer the depth of objects.

An additional means of estimating the depth of objects in light field images involves using the gradient of the 4D light field [14]. The gradient of the light field can be used to describe the directions the light rays took when passing through the camera. With this information, it is possible to map light rays out of the camera to possible

sources. Figure 54 depicts a 2D slice of a light field image. The left side of the image is the main lens. The right side of the image is the imaging sensor. By determining the gradient of this image, it is possible to show the direction that the light rays took when it passed through the capturing camera. From the direction vectors produced by a gradient operator of the image, it is possible to infer the origin of the light source.

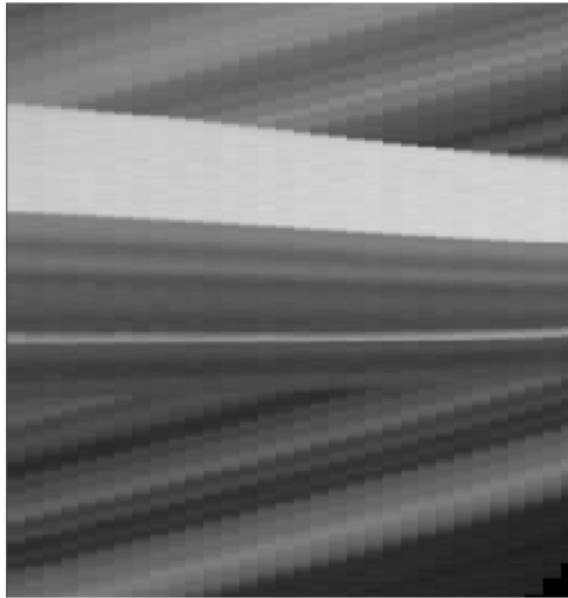


Figure 54 This figure illustrates a single slice of a light-field image. The left side of the image is the main lens. The right side of the image is the imaging sensor. This image was taken from *Gradient-Based Depth Estimation from 4D Light Fields* © 2004 IEEE [14].

### Stereo Correspondence

The gradient based method described in the previous section has some properties that are not desirable. This method does not perform well when images contain a high-level of noise. The images processed by the two prototype cameras contain a fair amount of noise. It is difficult to visualize these slopes in a 4D space. One attempt to attain a better understanding of the depth of field in regards to light field images was to use

stereo-correspondence. Typically, stereo correspondence is used to depict differences in stereo images. Though the 4D light field slices are not stereo images, they do possess similar properties. Stereo correspondence examines the optical flow between two images. The disparity between two images can be used for estimating the depth of objects in a scene. Because of this property, stereo correspondence of light field slices might contain beneficial properties like the disparity between light rays.

To determine whether stereo correspondence delivers any interesting insight, an OpenCV implementation of disparity estimation was applied to slices of the light field [49]. Figure 55 is the raw light field image used for the test. Figure 56 shows the light field of the test image. The image was captured with a Lytro camera. The lower-right corner of the image contains objects that were located close to the camera. The upper-right corner of the image contains objects that were located at a distance of approximately 10 m from the camera. The color gain on the image is relatively high. This allowed for each slice to be visible. The center slice is oversaturated because of the high gain. This image contains multiple sharp edges at different depths.



Figure 55 This is a raw light field image captured with a Lytro camera. This was used for stereo correspondence test.

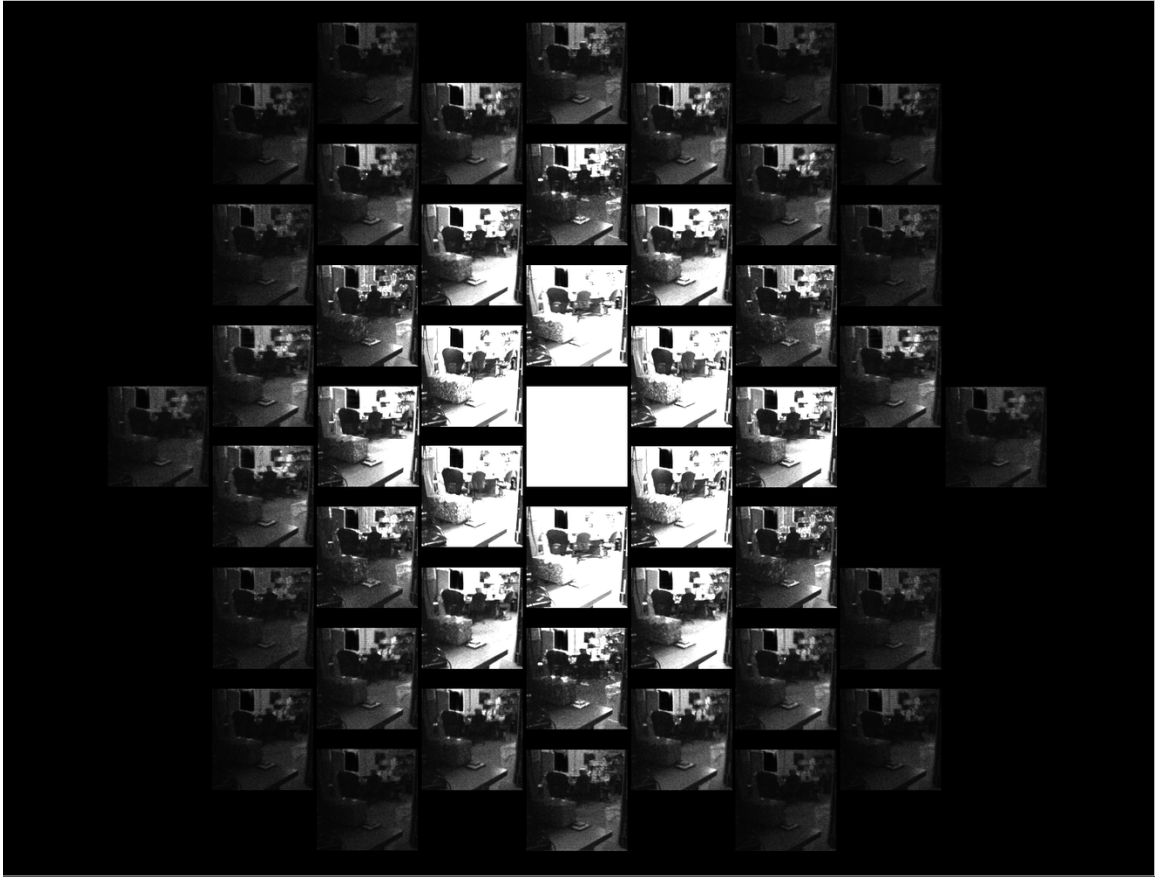


Figure 56 This figure shows light-field slices of a test image for a stereo correspondence test.

Figure 57 depicts a rendering of the stereo correspondence test. All slices were compared with the center slice. The picture visualizes the disparities between each slice and the center slice. Each disparity corresponds to a slice at the same location in the original image. Disparity is measured in terms of the pixel disparity between images. In this figure, brighter values correspond to larger disparities. The largest disparity values were measured along a wall that was located at a distance of approximately 10 m from the camera. The objects that were closer to the camera tended to have smaller disparities.



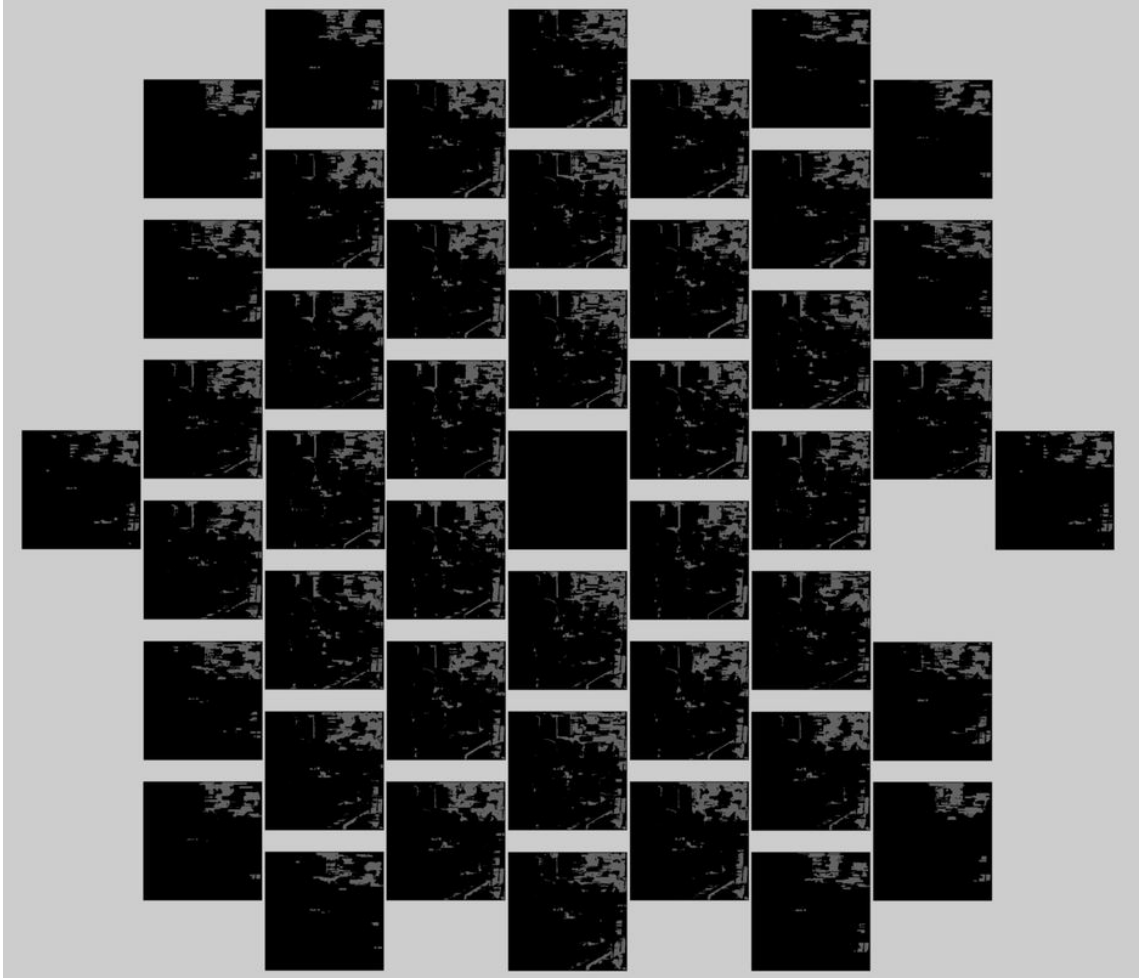


Figure 57 This figure depicts a disparity image for each slice of the light field. The disparity is measured between the center slices and the slices located at the center of each disparity image.

In this simple case, the stereo correspondence identified objects that were located at distances farther away. Though the measurements produced were very coarse measurement, this process shows potential. It might be possible to use disparity to perform ray-tracing for objects in a field. There are multiple ways to perform a gradient operation on a 4D light field image and the best solution has yet to be determined. These disparity images deliver similar information. There is still a lot of work that can be done to fully understand the potential of these disparity images.

## CONCLUSIONS AND FUTURE WORK

A range of useful information was generated regarding light field cameras and the considerations that need to be taken to make them viable for airborne platforms. This dissertation used commercially available cameras and lenses that had adjustable focal lengths. Capturing the light field was demonstrated with two different modulation masks inserted into prototype cameras. The masks were able to function with the camera's main lens set to different focal lengths. The effects of mask placement and focal length of the main lens was quantified. Although the prototype camera was never mounted to an airborne platform during this dissertation, demonstrating its ability to capture light fields with varying focal lengths makes the prototype viable for a low-flying platform. The principles demonstrated with these masks and cameras could be applied to future camera designs. Specifically, the designs of three components should be considered. These components are the imaging sensor, the modulation mask, and the main lens. These components are depicted in Figure 58.

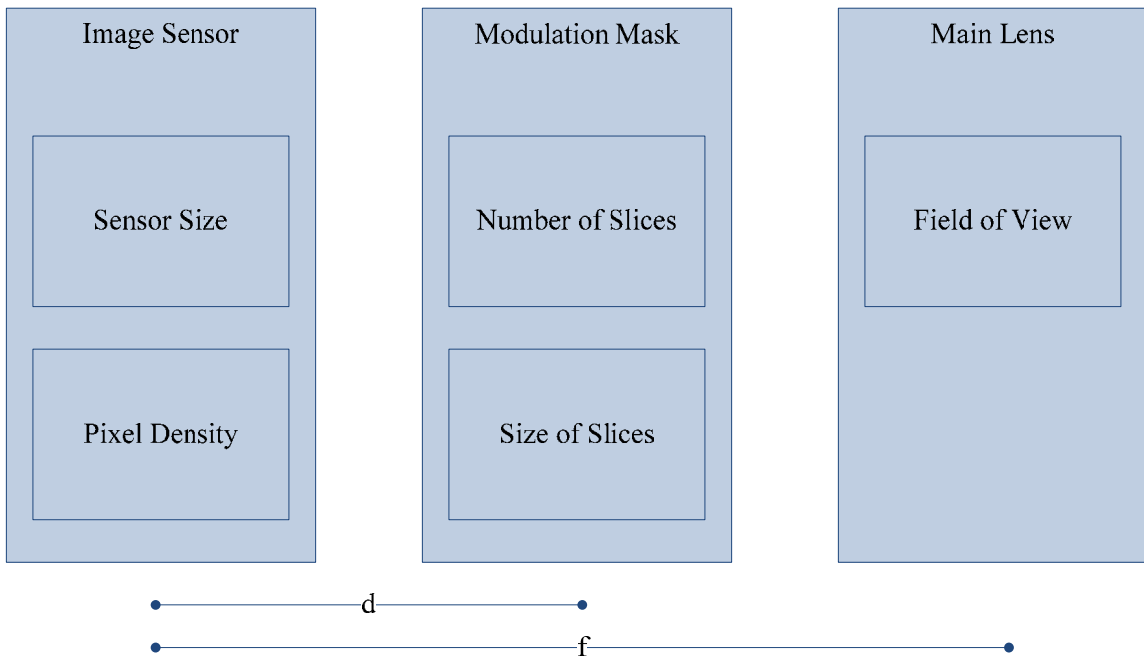


Figure 58 This figure shows considerations that should be taken when designing light-field cameras.

Regarding imaging sensors, there are two critical aspects. These aspects are sensor size and pixel density. Large sensor sizes allow for more modulation pulses to land on the imaging sensor. However, larger imaging sensors are only more effective than smaller imaging sensors when the modulation masks and main lenses fully utilize the surface of the imaging sensor. Yet, the physical size of the sensor can limit a camera's design more than the pixel density of the sensor. The modulated light field slices need to fall onto the image sensor in order to be captured. Pixel density dictates the possible resolution of the light field slices. If the slices do not land on the image sensor, the resolution of the slices does not matter. A small sensor with a high pixel density requires a very small  $d/f$  ratio. In general, this is not very practical unless the mask is integrated into the fabrication of the image sensor so that  $d$  value is very small. The other way is to have a lens with a large focal length but that will increase the cost and size of the camera.

Two types of modulation masks were discussed. These types were the microlens array and the printed mask. It was discovered that there were limitations relating to the printing method and to the use of 35mm film. Other printing methods may produce better results. The tests performed with a laser pointer can be performed on any printed mask to test its viability. It is easier to control the placement of the mask in the camera than it is to control manufacturing of the mask itself. The  $d/f$  ratio can be manipulated to allow the modulated slices to land on the imaging sensor. The printed mask or microlens array governs the number of slices produced. When a mask is able to diffract light as expected, the  $d/f$  ratio can be manipulated in a manner that causes the slices to land on the imaging sensor.

The main lens governs the field of view of the camera. Thus, the field of view of the camera should be the primary concern when designing a lens. The focal length of the lens and the size of the imaging sensor control the field of view. With a given focal length of the main lens, the modulation mask can be placed in such a way that allows the whole imaging sensor to be used.

A new method of extracting the 4D light field from raw images was also presented. The method works for raw images captured with both a microlens array light field camera and a printed mask light field camera. This method involves manipulating DFT images based on the properties of the camera's lens and modulation device (microlens or mask). Even without knowing the exact properties of a camera, it was possible to successfully use this method.

There is a lot of additional research that can be performed regarding the means by which an object's depth can be determined from light field images. This dissertation primarily focused on the method of capturing the 4D light field with a printed-mask 4D light field camera. There is still a lot of additional research that can be done in terms of the post-capture processing of the light field, as well. Stereo correspondence between light field slices shows promise. However, a larger dataset is required before any claim

of efficacy can be made. The preprocessing method described in this dissertation may help advance this effort.

Overall, the principles demonstrated in this dissertation could be used to develop small, light-weight and low-cost light field cameras that would be used on a low flying airborne platform. Printed masks offer more flexible modulation and filtering capabilities than tradition microlens arrays and they are considerably easier and of lower cost to manufacture. The research presented in this dissertation opens the door for other types of light field modulations with cameras.

## WORKS CITED

1. Adelson, E.H. and J.Y.A. Wang, *Single lens stereo with a plenoptic camera*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1992. **14**(2): p. 99-106.
2. Ng, R., *Fourier slice photography*. ACM Trans. Graph., 2005. **24**(3): p. 735-744.
3. Bell, D.G., et al. *NASA World Wind: Opensource GIS for Mission Operations*. in *Aerospace Conference, 2007 IEEE*. 2007.
4. Butler, D., *Virtual globes: The web-wide world*. Nature, 2006. **439**(7078): p. 776-778.
5. *Maps, Imagery, and Publications*. 2013 June 21, 2013 [cited 2013 September 13, 2013]; Available from: <http://www.usgs.gov/pubprod/index.html>.
6. Kihwan, K., et al. *Augmenting Aerial Earth Maps with dynamic information*. in *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*. 2009.
7. Bhanu, B., et al., *Distributed Video Sensor Networks*. 2011, Springer-Verlag London Limited,: London.
8. *Aerial Imaging Market - 2014 Global Industry Analysis*. 2014, M2PressWIRE.
9. *History of Aerial Photography*. [cited 2013 September 25, 2013]; Available from: <http://www.papainternational.org/history.asp>.
10. Ritter, N. and M. Ruth, *The GeoTiff data interchange standard for raster geographic images*. International Journal of Remote Sensing, 1997. **18**(7): p. 1637-1647.
11. Ng, R., et al., *Light field photography with a hand-held plenoptic camera*. Computer Science Technical Report CSTR, 2005. **2**(11).
12. Levoy, M. and P. Hanrahan, *Light field rendering*, in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. 1996, ACM. p. 31-42.
13. *EOS REBEL T1i/EOS 500D Instruction Manual*. [cited 2015 1/23/2015]; Available from: <http://gdip01.c-wss.com/gds/0/0300004280/01/eosrti-eos500d-im2-en.pdf>.
14. Dansereau, D. and L. Bruton. *Gradient-based depth estimation from 4D light fields*. in *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*. 2004.
15. Atashgah, M.A. and S. Malaek, *Prediction of aerial-image motion blurs due to the flying vehicle dynamics and camera characteristics in a virtual environment*. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of aerospace engineering, 2013. **227**(7): p. 1055-1067.
16. Hoffmann, G., et al. *The Stanford testbed of autonomous rotorcraft for multi agent control (STARMAC)*. in *Digital Avionics Systems Conference, 2004. DASC 04. The 23rd*. 2004.
17. Huerta, M.P., *Interpretation of the Special Rule for Model Aircraft*, F.A. Administration, Editor. 2014.
18. Lear, A.C., *Digital orthophotography: mapping with pictures*. Computer Graphics and Applications, IEEE, 1997. **17**(5): p. 12-14.
19. Nale, D., *Digital orthophotography: The foundation of GIS*. The American City & County, 1995. **110**(8): p. 26.
20. Zhu, J., *Conversion of Earth-centered Earth-fixed coordinates to geodetic coordinates*. Aerospace and Electronic Systems, IEEE Transactions on, 1994. **30**(3): p. 957-961.
21. Lijun, Z., et al. *Evaluation of GPS/ IMU Supported Aerial Photogrammetry*. in *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*. 2006.

22. Hartley, R. and A. Zisserman, *Multiple view geometry in computer vision*. 2000, Cambridge, UK ; New York: Cambridge University Press. xvi, 607 p.
23. Fogel, D.N. *Image rectification with radial basis functions: Application to RS/GIS data integration*. in *Proceedings of the Third International Conference on Integrating GIS and Environmental Modelling, CDROM*. 1996.
24. Brown, M. and D.G. Lowe, *Automatic panoramic image stitching using invariant features*. *International Journal of Computer Vision*, 2007. **74**(1): p. 59-73.
25. Fathima, A.A., R. Karthik, and V. Vaidehi, *Image Stitching with Combined Moment Invariants and Sift Features*. *Procedia Computer Science*, 2013. **19**(0): p. 420-427.
26. Goodman, J.W., *Introduction to Fourier optics*. 3rd ed. 2005, Englewood, Colo.: Roberts & Co. xviii, 491 p.
27. Yariv, A., P. Yeh, and A. Yariv, *Photonics : optical electronics in modern communications*. 6th ed. The Oxford series in electrical and computer engineering. 2007, New York: Oxford University Press. xii, 836 p.
28. Devernay, F. and O. Faugeras, *Straight lines have to be straight*. *Machine Vision and Applications*, 2001. **13**(1): p. 14-24.
29. Ricolfe-Viala, C. and A.-J. Sánchez-Salmerón, *Using the camera pin-hole model restrictions to calibrate the lens distortion model*. *Optics & Laser Technology*, 2011. **43**(6): p. 996-1005.
30. Adelson, E.H. and J.R. Bergen, *The plenoptic function and the elements of early vision*. *Computational models of visual processing*, 1991. **91**(1): p. 3-20.
31. Levin, R., *Photometric Characteristics of Light Controlling Apparatus*. *Illuminating Engineering*, 1971. **66**(4): p. 205-215.
32. Gortler, S.J., et al. *The lumigraph*. in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. 1996. ACM.
33. Wilburn, B., et al. *High-speed videography using a dense camera array*. in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. 2004.
34. *PRESS RELEASE: LYTRO REDEFINES PHOTOGRAPHY WITH LIGHT FIELD CAMERAS*. 2011 [cited 2014; Available from: [https://www.lytro.com/press/releases/2011-06-22\\_lytro-redefines-photography-with-light-field-cameras/](https://www.lytro.com/press/releases/2011-06-22_lytro-redefines-photography-with-light-field-cameras/)].
35. Dansereau, D.G., O. Pizarro, and S.B. Williams, *Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras*.
36. Veeraraghavan, A., et al., *Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing*. *ACM Trans. Graph.*, 2007. **26**(3): p. 69.
37. Oppenheim, A.V., R.W. Schafer, and J.R. Buck, *Discrete-time signal processing*. 2nd ed. Prentice Hall signal processing series. 1999, Upper Saddle River, N.J.: Prentice Hall. xxvi, 870 p.
38. *Canon EOS Rebel Ti*. 2014 [cited 2014; Available from: [http://www.usa.canon.com/app/pdf/slr/Rebel\\_Ti\\_manual.pdf](http://www.usa.canon.com/app/pdf/slr/Rebel_Ti_manual.pdf)].
39. Wang, A., P.R. Gill, and A. Molnar. *An angle-sensitive CMOS imager for single-sensor 3D photography*. in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*. 2011.
40. Chih-Chieh, C., et al. *Depth estimation of light field data from pinhole-masked DSLR cameras*. in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. 2010.
41. Raskar, R., et al., *Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses*. *ACM Trans. Graph.*, 2008. **27**(3): p. 1-10.
42. *Life Pixel – Canon DRebel XSi (450D) DIY Digital Infrared Conversion Tutorial*. [cited 2015 1/23/2015]; Available from: <http://www.lifepixel.com/tutorials/infrared-diy-tutorials/canon-drebel-xsi-ir>.

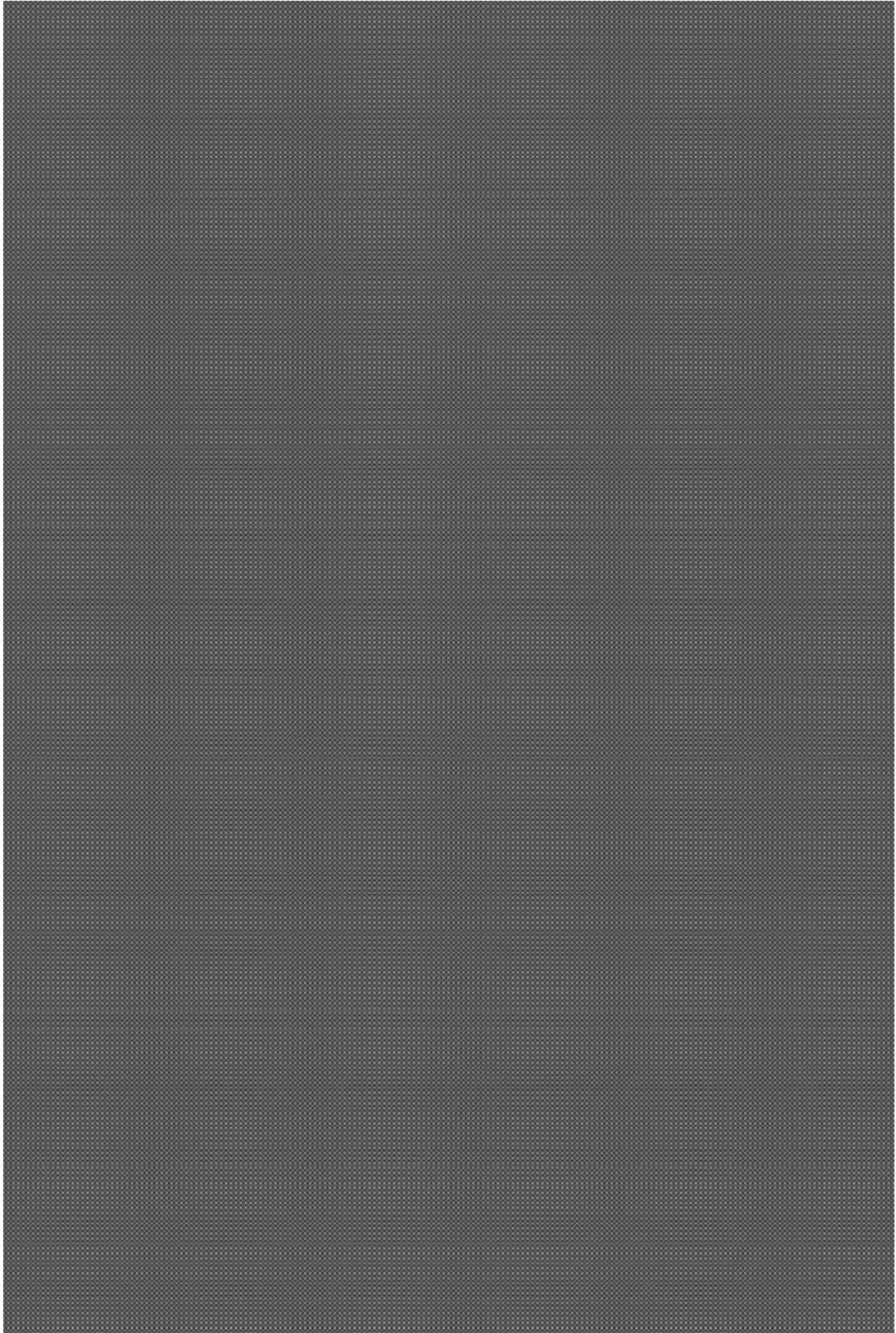
43. *FinePix S5200 Owners Manual*. 2005 [cited 2015 2/18/2015]; Available from: <http://www.fujifilmusa.com/shared/bin/S5200ownermanual.pdf>.
44. Sonka, M., V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*. 3rd ed. 2008, Toronto: Thompson Learning. xxv, 829 p.
45. Shapiro, L.G. and G.C. Stockman, *Computer vision*. 2001, Upper Saddle River, NJ: Prentice Hall. xx, 580 p.
46. Petersen, D.P. and D. Middleton, *Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean spaces*. *Information and Control*, 1962. **5**(4): p. 279-323.
47. Morgan, W.D.L.H.M., *Stereo realist manual*. 1954, New York: Morgan & Lester.
48. McMillan, L. and G. Bishop. *Head-tracked stereoscopic display using image warping*. 1995.
49. Kosov, S., T. Thormählen, and H.-P. Seidel, *Accurate Real-Time Disparity Estimation with Variational Methods*, in *Advances in Visual Computing*, G. Bebis, et al., Editors. 2009, Springer Berlin Heidelberg. p. 796-807.



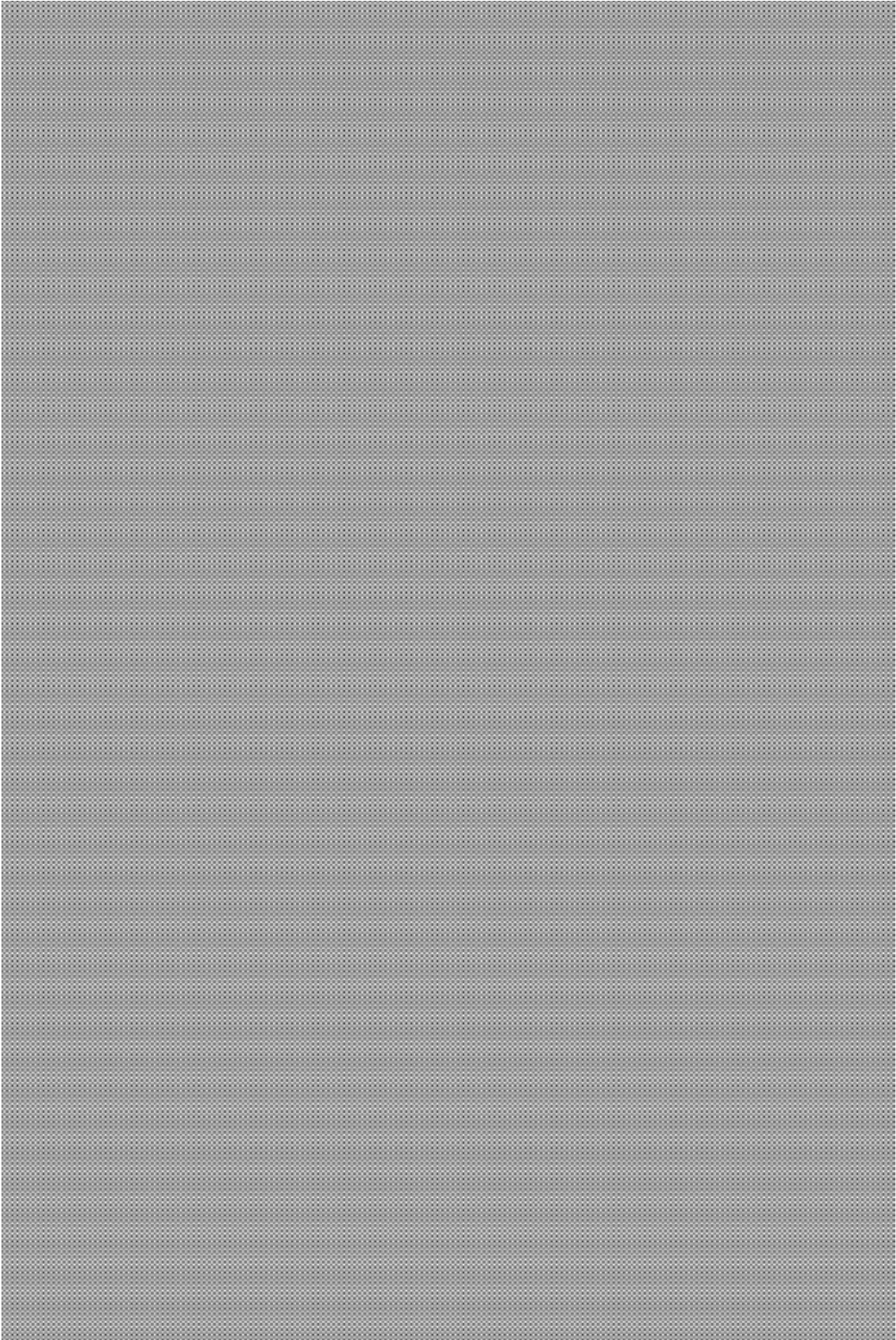
## APPENDIX A: PRINTED MASKS

This appendix contains full-resolution printed images. Each mask has the following variables defined:  $p$  is the number of cosine harmonics in the mask,  $s$  is the frequency in pixels for the lowest harmonic cosine. It is also indicated whether the image is a mask or the negative of a mask. All masks have an original resolution of  $2732 \times 4096$  pixels. The reproduction of this document does not support the full resolution of the mask images. The masks should only be used as approximations due to this limitation.

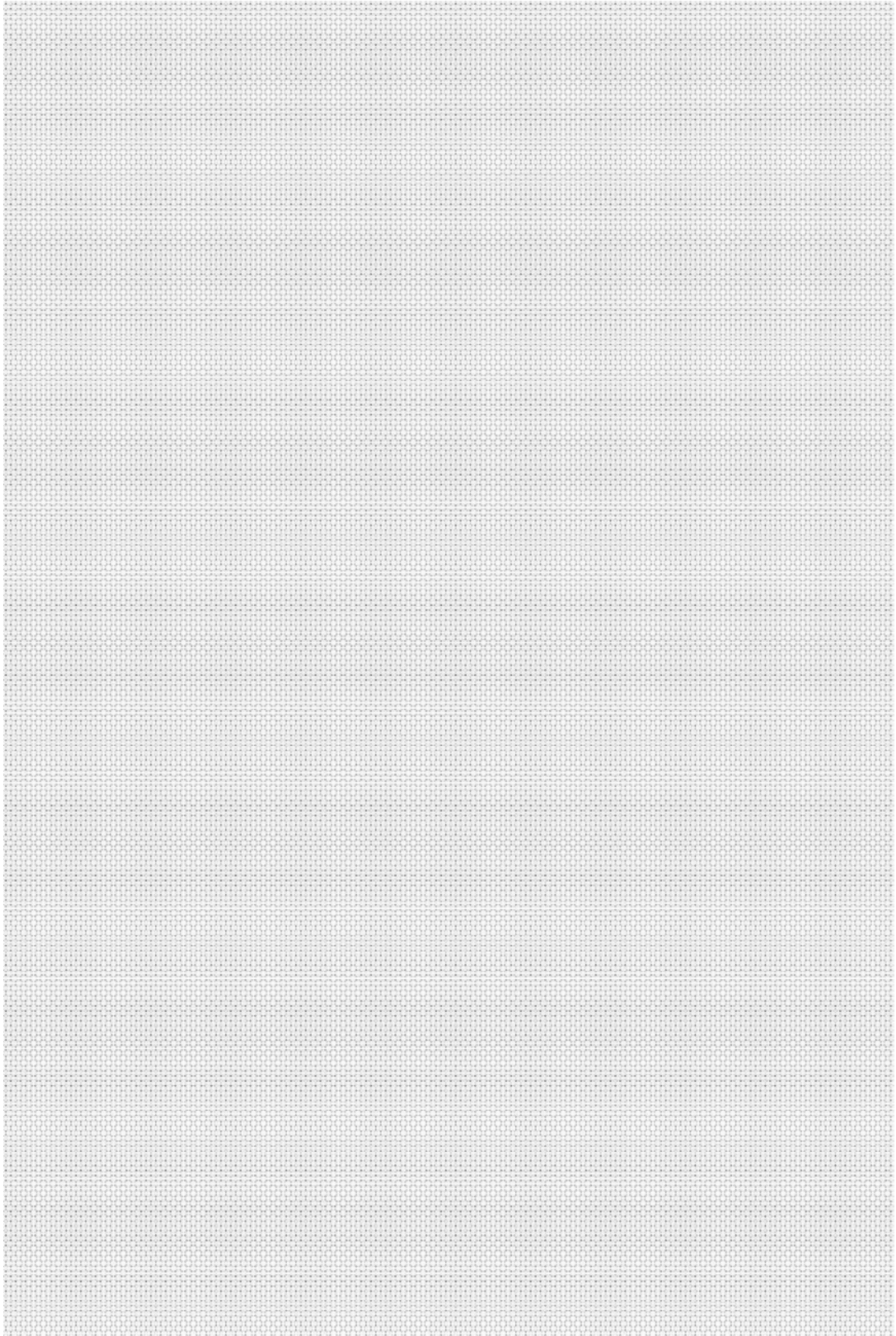
Printed Mask  $p = 2, s = 5$



Negative Printed Mask  $p = 2, s = 5$

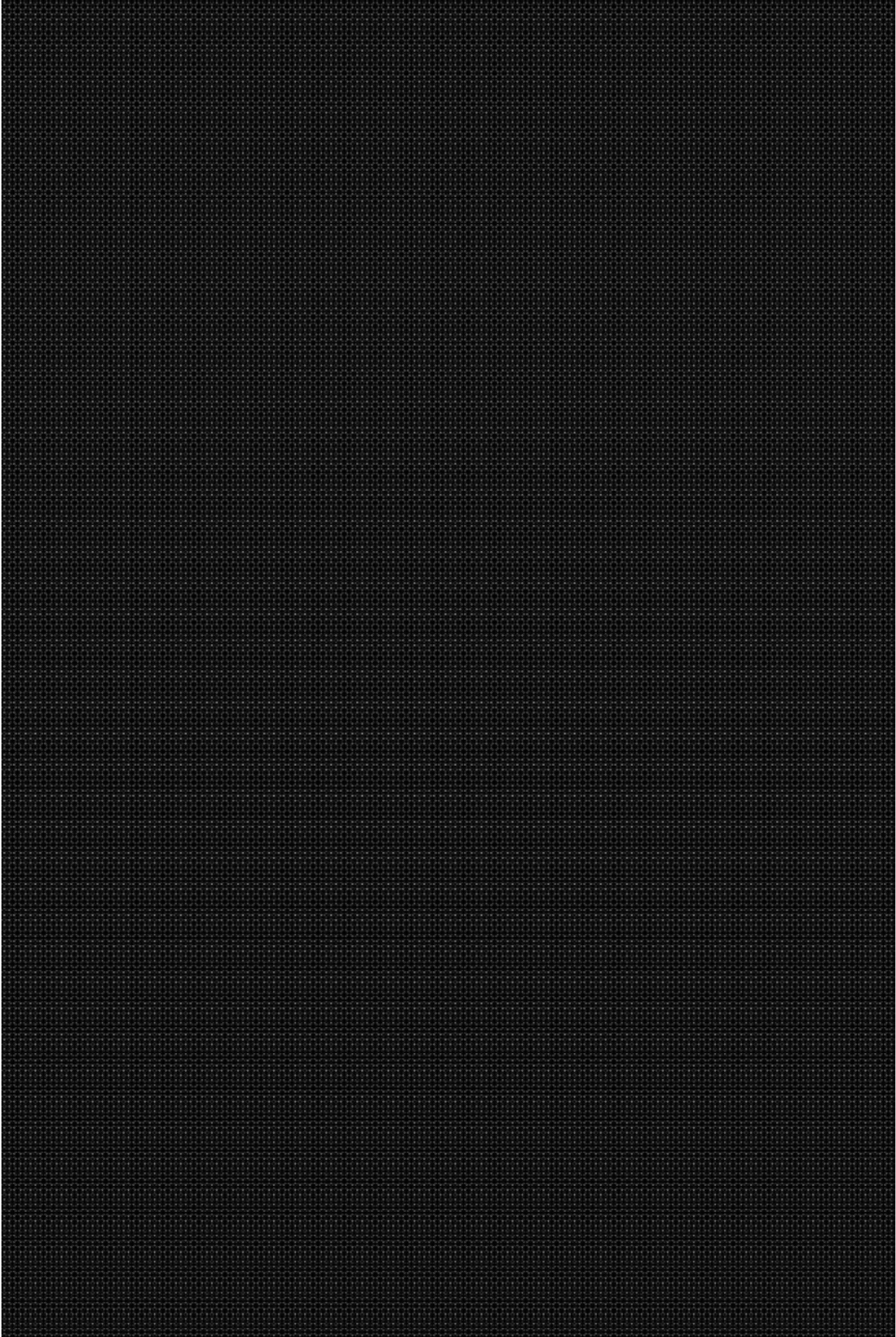


Printed Mask  $p = 2, s = 10$

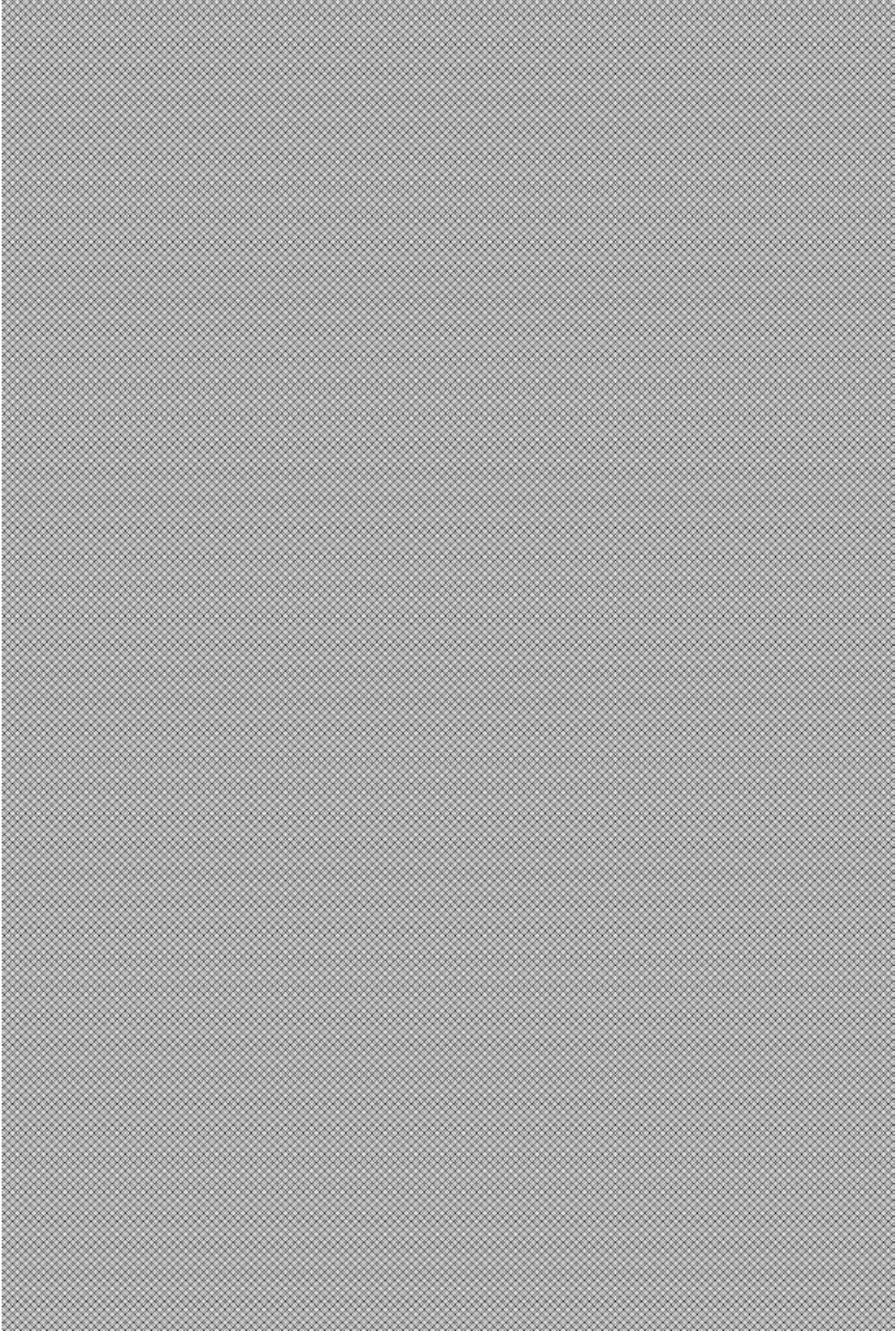




Negative Printed Mask  $p = 2, s = 10$

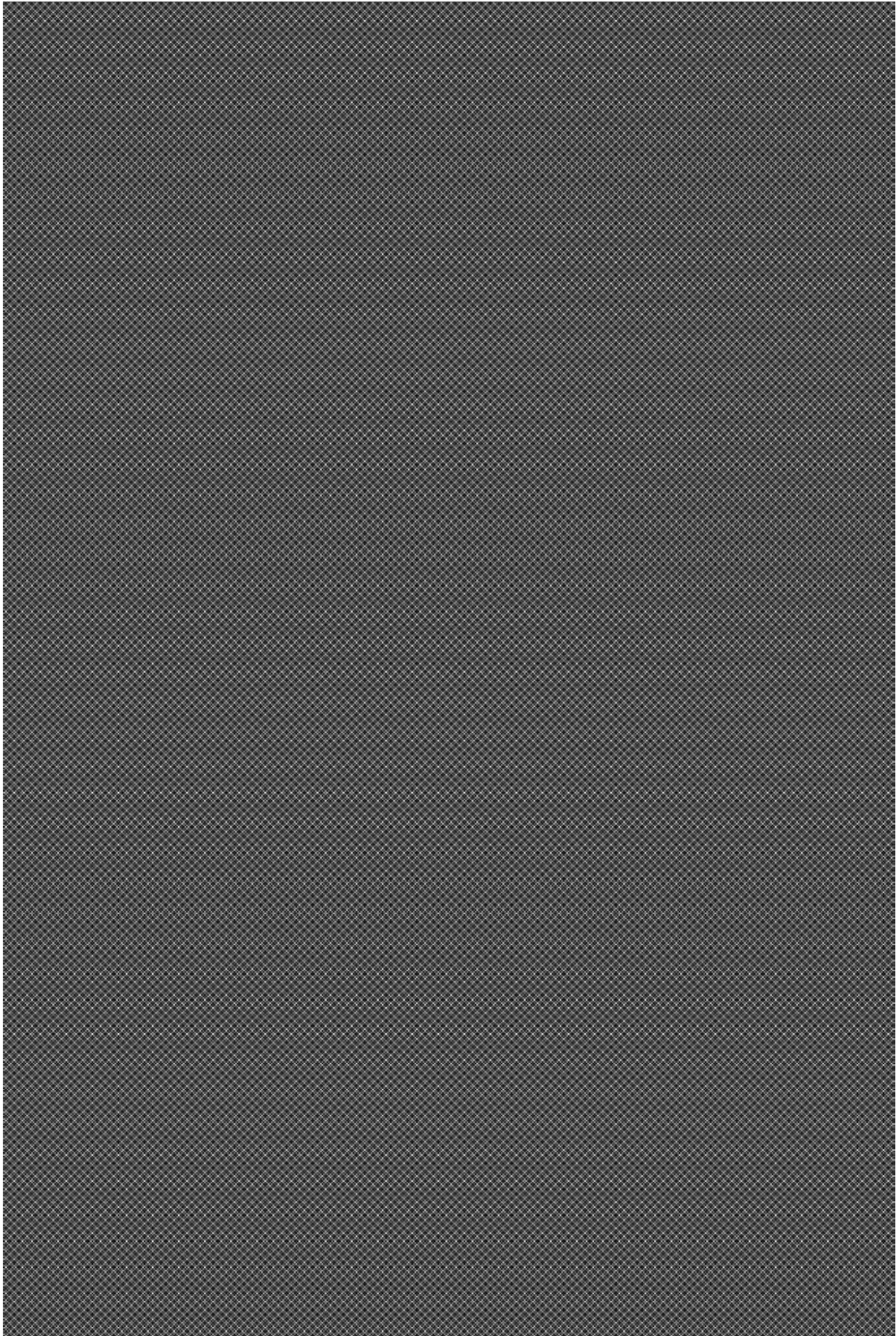


Printed Mask  $p = 4, s = 10$

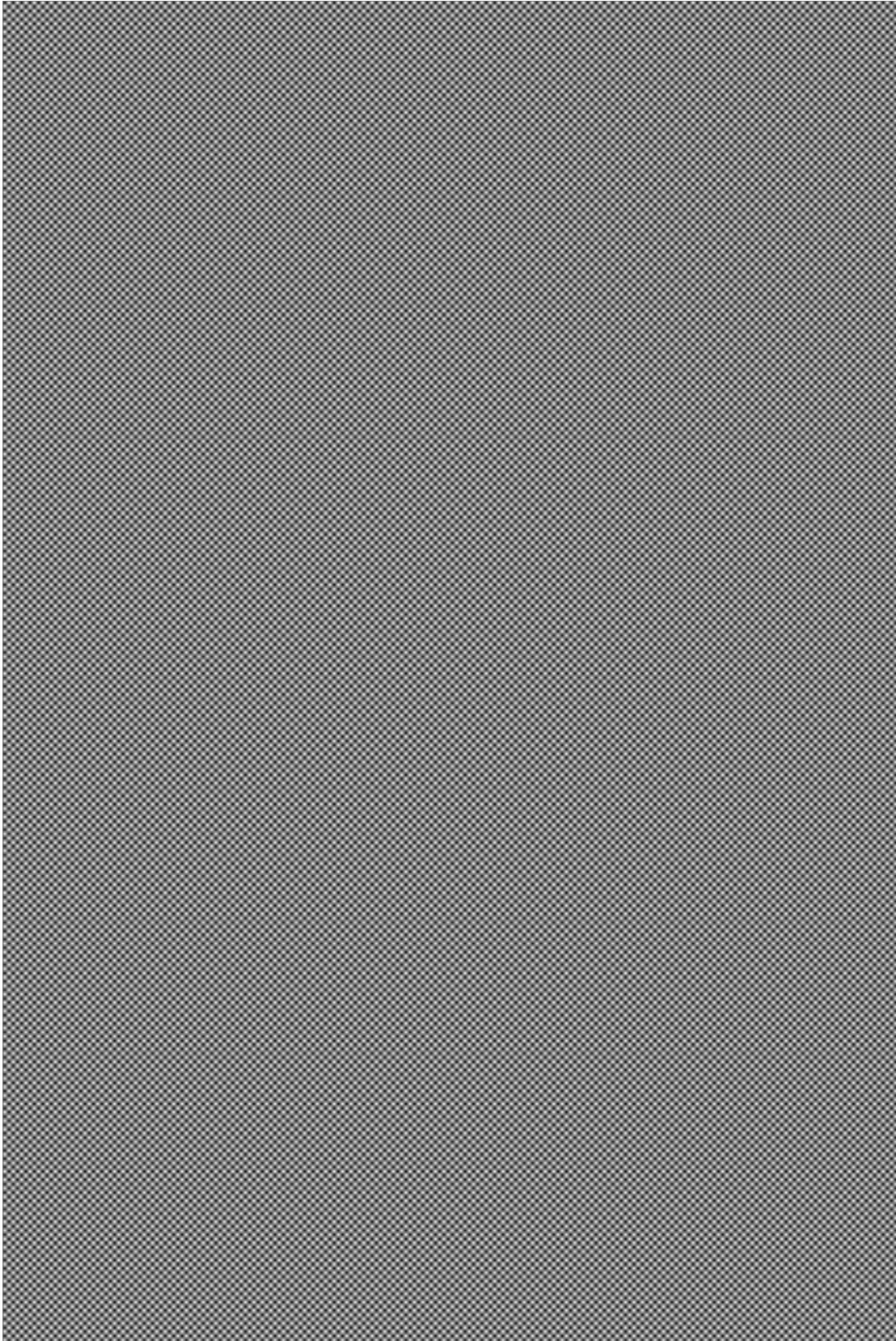




Negative Printed Mask  $p = 4, s = 10$



Printed Mask  $p = 10, s = 20$

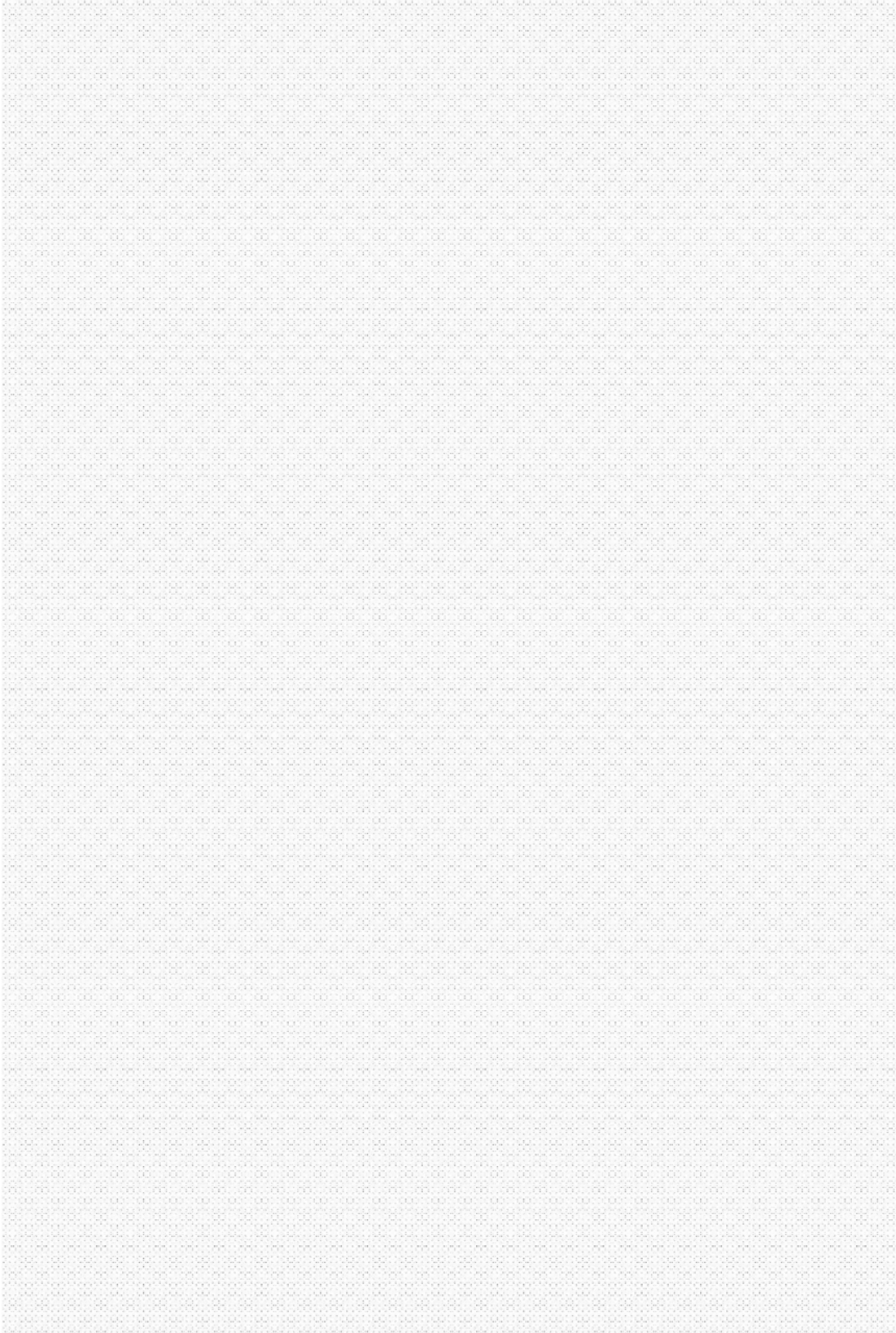




Negative Printed Mask  $p = 10, s = 20$



Printed Mask  $p = 5, s = 20$



Negative Printed Mask  $p = 5, s = 20$

