Arbitrarily-Controllable Programmable Aperture Light Field Cameras: Design Theory, and Applications to Image Deconvolution & 3-Dimensional Scanning

by

Thomas Bradford Milnes

B.S., Cornell University (2006)
S.M., Harvard University (2009)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2013

Author . . . .
Department of Mechanical Engineering
August 29th, 2013

Certified by. . . . .
Douglas P. Hart
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by . . . .
David E. Hardt
Professor of Mechanical Engineering
Chairman, Department Committee on Graduate Students

Arbitrarily-Controllable Programmable Aperture Light Field Cameras: Design Theory, and Applications to Image Deconvolution & 3-Dimensional Scanning

by

Thomas Bradford Milnes

Submitted to the Department of Mechanical Engineering
on August 29th, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

This thesis describes a new class of programmable-aperture light field cameras based on an all-digital, grayscale aperture. A number of prototypes utilizing this arbitrarily-controllable programmable aperture (ACPA) light field technology are presented. This new method of capturing light field data lends itself to an improved deconvolution technique dubbed "Programmable Deconvolution," as well as to 3D scanning and superresolution imaging. The use & performance of ACPA cameras in these applications is explored both in theory and with experimental results. Additionally, a framework for ACPA camera design for optimal 3D scanning is described.

Thesis Supervisor: Douglas P. Hart
Title: Professor of Mechanical Engineering

Acknowledgments

Like many significant undertakings, many people beyond the cited author of this work deserve credit for its contents. This section is an attempt to credit those to whom I am deeply grateful, and without whom the work in this thesis would have either suffered greatly or not have happened at all.

Firstly, I would like to thank my advisor, Doug Hart, whose signally flexible advising and mentoring style was exactly what I needed. My course through the Ph.D. program has been anything but usual–to the great benefit of both my academic and professional careers–and I don't believe it would have been possible with anyone else but Doug. It bears mentioning that Doug and I were colleagues (at Brontes Technologies), sailing partners, and friends before he was my advisor; that dynamic has colored positively my entire experience at MIT. Doug finally convinced me to apply to the Ph.D. program while we sipped wine and gutted a fish for dinner in the stern of his sailboat somewhere between Boston and the Cape Code Canal. "Come to MIT and do a Ph.D. with me," he said. "We'll do some cool stuff and have some fun doing it." He was very right.

I'd also like to thank the funders who made my time at MIT possible: to 3M, for seeing the value of our light field imaging work even when they had just purchased a competing technology in Brontes. To the United States Navy, who saw light field imaging as a new and exciting solution to underwater imaging problems. And most recently, to Mike Wardlaw and the Office of Naval Research, whose need for power generation underwater put an entirely different kind of research project in front of me, one that rivals the fascination I found in optics.

Thank you to Ramesh Raskar and Alex Techet for rounding out my thesis committee so well. Thank you for your support and guidance. Ramesh, thank you additionally for your help with Viztu through Imaging Ventures. Alex, thank you for showing me just how fun high-powered lasers, mutagenic dyes and jet fuel can be when put together in a research setting. Thank you also for staying on my committee even though we all would have understood if you wanted to focus all your time on

the triplets!

I'd also like to thank MIT for being the archetypal place where scientific theory meets real-world practice. MIT lives up to its charter's venerable charge, as borne by the dome in Lobby 7: "Established for [the] advancement and development of science [&] its application to industry, the arts, agriculture, and commerce." Though I hold only one degree from MIT (having earned my other two from competing universities) and spent only a relatively short time in the doctoral program, I nonetheless feel a great deal of "Nerd Pride!"

I'd like to thank Ash Martin, my business partner and friend, for the exciting journey we had together with Viztu Technologies. The story of much of my life has been one of divided attention, and I owe Ash a debt of gratitude for putting up with the constant demands made on me by the doctoral program, as well as the many demands I elected to make on myself with things like training at Fort Knox as a US Army Cadet, and training for, qualifying for and running the Boston Marathon. Despite my divided time, we built a fun company that did hard science and had a great culture (two words: "office archery!"). You pushed me in more positive ways than I could possibly enumerate, such that the company's eventual acquisition by 3D Systems was just the icing on the cake. I am very happy for and proud of what we did together, and grateful for having you to do it with.

I'd like to thank Eric Paley and Micah Rosenbloom, my two most formative mentors. Most of what I have achieved to date, both academically and professionally, can be traced back in one way or another to my early days at Brontes and to you. When I showed up for work that first day I thought I was getting a short-term, unpaid internship. Instead I got the opportunity of a lifetime. That opportunity continues to pay dividends today; thank you both.

In the same Brontes vein, I'd like to thank Janos Rohaly, Doug Johnston, Ed Tekeian and the rest of the Brontes team for the immensely positive roles they've played in my personal, professional and academic development. A special thanks to Meher Nerkizian for teaching me how to wire and program LCD drivers for my thesis work when I was in way over my head!

6

I'd like to thank my Beverly High School math & science teachers, Mrs. Brewster, Mrs. Dube, Mrs. Soghigian and Mr. Clement, for inspiring me toward and preparing me well for a life of science. The public education you provided to me and my classmates was unparalleled. Although Cornell, Harvard and MIT aren't known for slouching on scientific rigor, I have never felt anything but fully prepared for the academic journey I've had in the decade since I graduated from your halls. That is all to your credit. Thank you.

Penultimately, I'd like to thank my friends for their support and companionship. You've made the journey rich with new experiences, adventures, and much-needed distractions. Jim Brady, my oldest and closest friend, let me just say "thank you for everything," since compiling an exhaustive list of the things for which I am grateful to you is simply not possible. To Don Aubrecht & Ali Forsyth, thank you for us supporting each other in and out of the academy as we each found our own way through it. To Ralph & Taylor Gillis, Mary Nodine & Dave Breeding, for sharing many good times in the great outdoors as well as the urban jungle.

Finally, I'd like to thank my parents, who gave me every opportunity in life and then let me find my own way through it. You cultivated in me a love of knowledge, equipped me with a strong moral compass, and instilled in me a penchant for dispassion that has been–and ever will be–that which defines me most. I love you both. Thank you for being there for me when I needed it most, when I needed nothing at all, and for everything in between. If my mind is strong and if I am balanced and happy it is because you have done everything in your great capabilities to make me that way. This thesis is for you.

With Gratitude,

Tom

Lincoln Massachusetts, August 27th 2013

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction & Motivation

Cameras are ubiquitous in today's society, but for decades their core functionality remained unchanged: a camera recorded a single 2D image of a scene. Light field capture technology enables cameras to do much more [8], [9], [18]. For example, light field capture allows an image to be refocussed post-capture [35]. Calibration of a light field camera allows for the mathematical correction of aberrations post-capture [7]. Light fields also allow for the computation of the depth of visible objects in a scene [7]. None of this is possible by a regular camera, but with programmable-aperture light field (PALF) technologies all of this can be done in high resolution. Imagine imaging a scene or object without concern for focus, and getting a refocusable, aberration-free image in return. Envision a microscope that can computationally image beyond the limits placed on it by defocus, aberrations and diffraction. Further, imagine exporting a photographed scene or object in 3D to CAD software for digital manipulation, or directly to a 3D printer for digital replication. With PALF cameras, all of this is now possible.

1.1   Light Field Imaging

It is often said that a picture is worth "a thousand words," because it provides a clearer representation of a scene than a verbal description can achieve. An image, however, doesn't tell the whole story either: a camera's lens distorts and compresses the scene

it beholds. Imperfections in the lens's design and construction cause aberrations that distort shapes, and the 2-dimensional nature of an image means that the lens has collapsed away three of the scene's five native dimensions. Information about three of these native dimensions–surface irradiance and depth–are lost almost completely in a standard photograph; the only hints that remains are subtle cues like shading on which humans can pick up only after years of practice.

In order to capture this lost information with a camera, special hardware is needed; hardware whose output is worth "a thousand pictures," so-to-speak, capturing the depth and irradiance of the original scene. Enter the light field camera: an imaging device capable of sampling its own aperture plane, acquiring the images that arise from the light passing through individual regions of that plane. When the aperture of an imaging system is divided into a grid (which can be done either physically [7], [22], [21], [24], or synthetically through the use of a microlens array [44], [27],[3], pinhole array [15], or coded masks [35], [36], [17] [29]), the result is a 2-dimensional grid of (2D) images, yielding a 4-dimensional output. If this data is then used to calculate depth everywhere in these images and then the original image data is overlaid on this 3D model, the original 5-dimensional plenoptic function [39] can be recreated. This results in a true representation of the original scene; more information than words–or individual photographs–can express.

1.2   Depth from Light Field Data

To understand the data contained in a light field, one should first understand what happens when a scene is imaged with a regular lens. In Figure 1-1 the ray trace of a standard camera lens is depicted. The gray rays emanate from a point at the focal plane, while the blue and red rays emanate from points in front of and behind the plane, respectively. It is clear to see that while in-focus rays converge to point, out-of-focus rays do not. This defocus is the cause of the blurriness in images of objects that are out of focus. The diameter of the "blur circle" around a point varies directly with its distance from the focal plane.
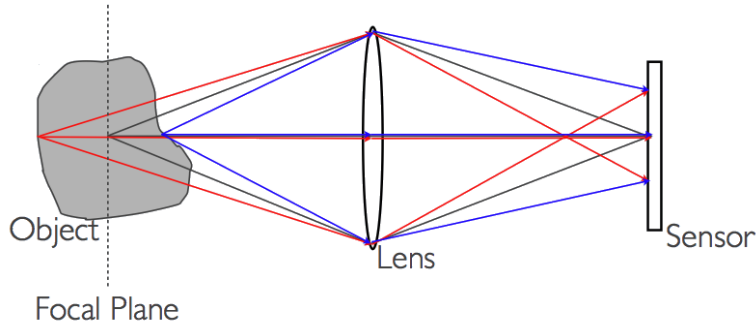
Figure 1-1: Ray trace for a standard photographic lens

An ideal lens obeys the thin lens formula,

$$\frac{1}{f} = \frac{1}{s_o} + \frac{1}{s_i} \qquad (1.1)$$

where f is the lens' focal length, s_o is the distance from an in-focus object to the lens, and s_i is the distance from the lens to the image plane. A geometric analysis of this system reveals that the diameter d of the blur circle around an out-of-focus point is

$$d = D\frac{s_i - \left(\frac{1}{f} - \frac{1}{s_{o'}}\right)}{\frac{1}{f} - \frac{1}{s_{o'}}} \qquad (1.2)$$

where D is the diameter of the lens and $s_{o'}$ is the distance from the out-of-focus point to the lens. So, a point's blur circle contains depth information about its position relative to the focal plane. By measuring that diameter and knowing the camera's parameters, that depth can be calculated. However, there are two practical problems with this. First, blur circles are just that–they're blurry. For scenes with complicated textures it's hard to determine where the blur from one feature starts and the blur from another begins. The deconvolution quickly becomes intractable as scene detail becomes elaborate (see Chapter 7 for more details). Secondly, the diameter of the blur circle alone isn't enough to determine whether the point lies in front of or behind the focal plane; because the circle diameter isn't signed, an ambiguity arises with relation to the focal plane.

To understand how to get around this ambiguity by sampling the aperture plane, it

is instructive to look at how rays passing through different regions of the lens behave. Figure 1-2, which is borrowed from [7], illustrates what off-center apertures look like. Figure 1-3 illustrates the image obtained by opening a single off-axis aperture. The blue ray arrives at the sensor above the red ray. By opening an aperture at the opposite end of the lens, as illustrated in Figure 1-4, the blue and red rays swap positions. If the position of each ray in one image is compared to its position in the opposite image, the diameter of the blur circle *and* whether the originating point lies in front of or behind the focal plane can be determined. Specifically, if the ray's image position moves in the same directions as the aperture position, then the object lies in front of the plane; if its image position moves opposite, the object lies behind the focal plane.

Figure 1-2: Two off-center apertures on a single aperture plate
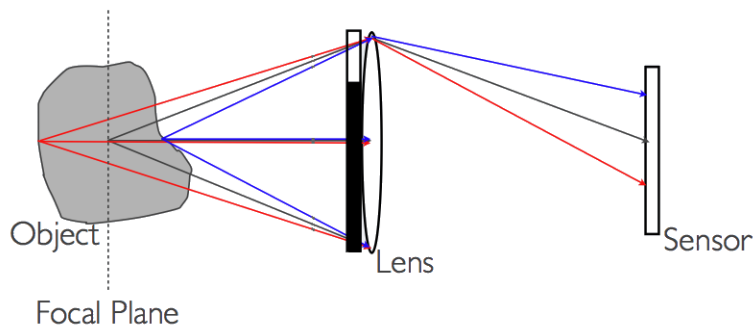
Figure 1-3: Ray trace of the light passing through the top section of a lens
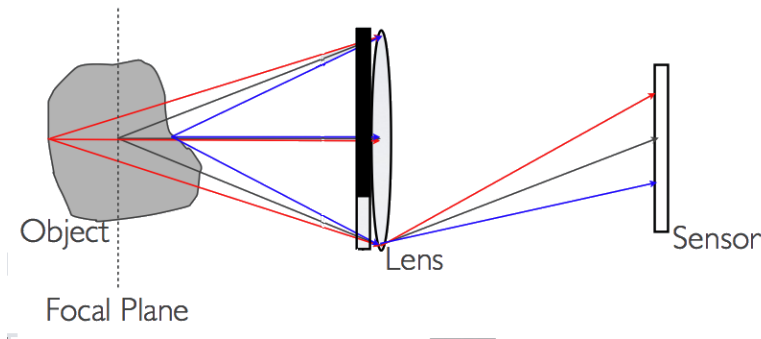
Figure 1-4: Ray trace of the light passing through the bottom section of a lens

This notion of sampling the aperture is the heart of light field imaging. In theory, any number of aperture configurations are possible, but pragmatically the ability to sample arbitrary positions is hard to implement. As previously cited in Section 1.1, there have been several implementations of light field technologies to date that select a sub-set of aperture positions from which to capture data, exploiting various geometrical properties and computational techniques. The main contribution of this thesis is to advance a technology that enables complete, arbitrary control over the aperture plane–which has not been achieved by other technologies–such that full advantage can be taken of the information coming through the lens across the widest range of possible imaging conditions. This new, "arbitrarily-controllable programmable aperture" (ACPA) technology has been dubbed the "digital-aperture."

1.3   Thesis Roadmap

The rest of this thesis is broken down into the following chapters:

Chapter 2 reviews many of the light field imaging techniques that have been developed to date. This review includes more detail on the technologies mentioned in Section 1.1

Chapter 3 describes the many ACPA prototypes that have been built and tested as part of this thesis work.

Chapter 4 demonstrates how ACPA cameras can be calibrated, and how calibrated ACPA cameras can be used to compute 3D data for a scene.

Chapter 5 advances a theoretical framework for ACPA camera design that optimizes design parameters for specific 3D scanning applications.

Chapter 6 describes a method for using ACPA technology with low-resolution sensors to achieve high-resolution imagery by leveraging known superresolution techniques.

Chapter 7 advances a technique for using ACPA technology–particularly in microscopes and telescopes–for deconvolving blurry scenes quickly and with better results than single-image methods. This technique has been termed "Programmable Deconvolution."

Finally, this thesis closes with some conclusions based on its presented work, directions in which this work can be taken in the future, and how these advances may impact our daily lives in the future. These topics are covered in Chapter 8.

Chapter 2

Review of Light Field Imaging Techniques

A number of light field camera types have been proposed and built to date. They can be grouped into three broad categories: the first category is comprised of single-lensed cameras that capture all of their information in a single snapshot. These cameras are covered in Section 2.1. The second group is comprised of camera arrays that capture a number of full-resolution images simultaneously from different perspectives; previous work in this area is discussed in Section 2.2. The third class of light field cameras spans the first two groups: these cameras achieve high-resolution capture in a single device, by taking multiple images with different aperture configurations. More detail on this group is available in Section 2.3; improvements to and applications of this paradigm of light field imaging–which are the central contribution of this thesis–are presented in Section 2.4 and in the chapters that follow it.

In addition to the development of specific devices, work has been done on the theoretical design and limitations of light field cameras in general. For example, [8] compares the abilities of a number of different light field camera paradigms with respect to their usefulness in integral imaging. Further examination by the same authors of the limits of spatial and angular resolution–and how they can sometimes be overcome computationally–is found in [9]. A thorough analysis of the performance of virtually all of the techniques in the first class of cameras can be found in [18].

2.1 Single-Shot Light Field Cameras

The first category of light field cameras in the literature is comprised of cameras that capture all of their information in a single snapshot: this has been accomplished with microlens arrays placed near the sensor plane [44], [27] and [3], pinhole apertures placed near the sensor plane [15], and coded masks placed near the sensor plane [35]. An additional, related method is to place coded masks at the aperture plane [36], [17] and [29]. Relatedly, some of these coded mask approaches have been used to demonstrate "optical heterodyning." The use of prisms in lieu of microdisplays has also been suggested [8], and the successful application of the microlens paradigm to microscopy has been demonstrated [19] and [20]. Figures 2-1 and 2-2 depict the microlens and coded sensor-mask approaches, respectively.
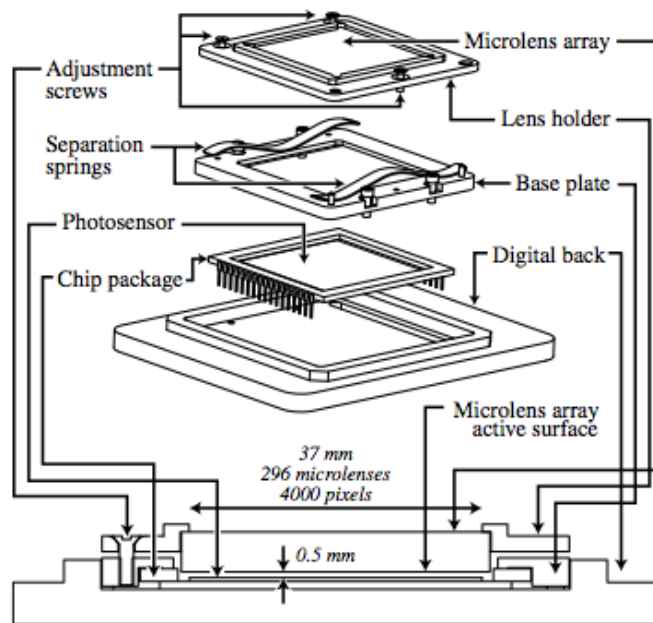
Figure 2-1: Schematic of the light field camera in [27], which utilizes a microlens array, and was one of the first camera technologies that demonstrated post-capture image refocusing.
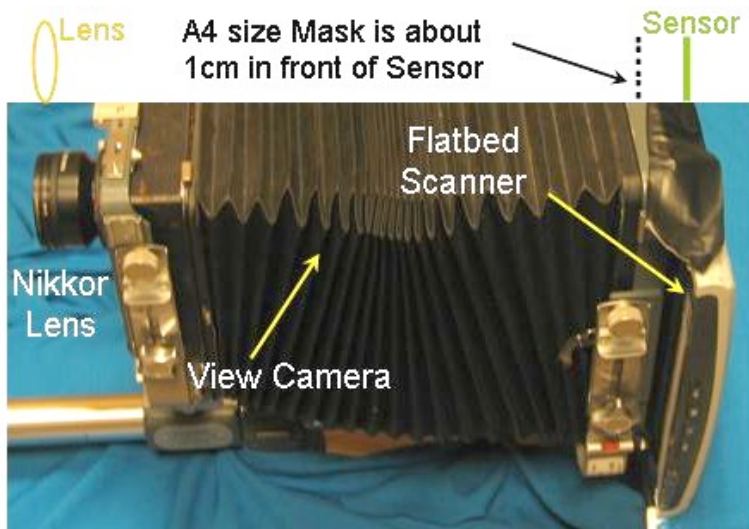
Figure 2-2: Image and schematic of coded sensor-mask camera described in [36]

2.2   Camera Arrays

The second major class of light field cameras is large camera arrays [43] and [45]. These embodiments utilize multiple cameras capturing in synchrony. Camera arrays are able to capture high resolution images for each perspective, at the expense of size, complexity and cost. These techniques have traditionally also taken advantage of the arrays' ability to record video at real-time frame rates, not just individual frames, in high-performance imaging applications [41], [42] and [33]. Two high-performance camera arrays are pictured in Figures 2-3 and 2-4.

2.3   Programmable-Aperture Light Field Cameras

The third class of light field cameras spans the first two groups: these cameras achieve high-resolution capture in a single device, by taking multiple images with different aperture configurations. Because of the use of multiple exposures, these cameras can generally only be used on stationary targets, but in return they capture higher resolution images than one-shot devices, without the added complexity of an array. Aperture manipulation has been demonstrated with rotating apertures [7], rotary aperture panels [22], punched paper rolls & static binary LCDs [21], and with com-
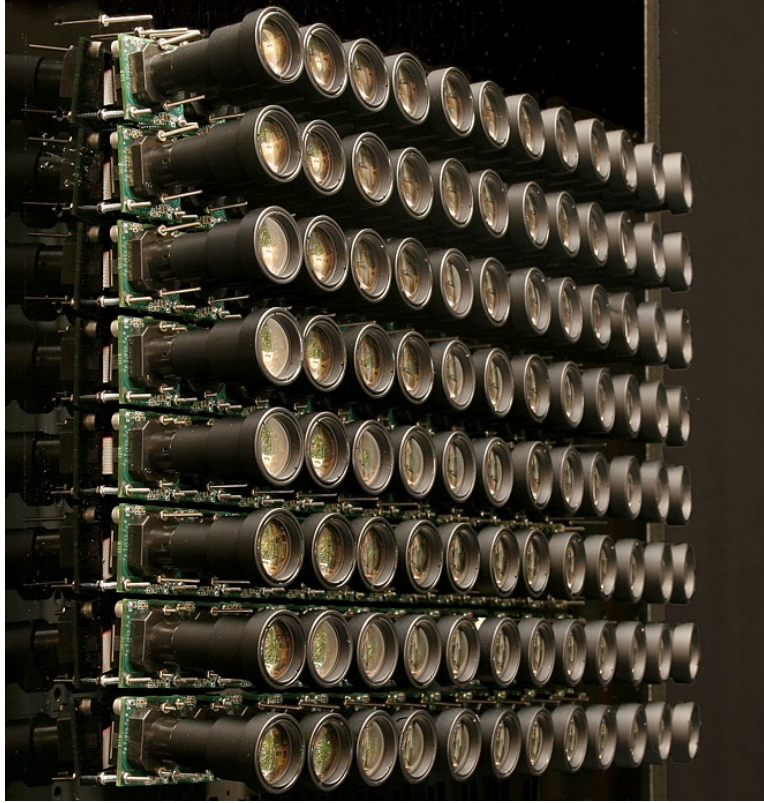
Figure 2-3: High-performance camera array from [42]

plex grayscale TFT LCD microdisplays [24]. Figure 2-5 shows a rotating-disk implementation of the programmable aperture, and Figure 2-6 shows implementations using paper scrolls and LCDs.

Contributions of this thesis to improving the programmable-aperture paradigm can be found in Section 2.4 and in the chapters that follow.

2.4 Improving the Programmable-Aperture Light Field Camera

Previous programmable aperture cameras relied on fixed subapertures whose shapes, sizes and positions were determined in advance and immutable by the operator. Programmable aperture techniques can be improved significantly by allowing control over aperture shapes, positions and opacities in real time and in response to the subject matter to be captured. Enabling these new modes of control is one of the main contri-
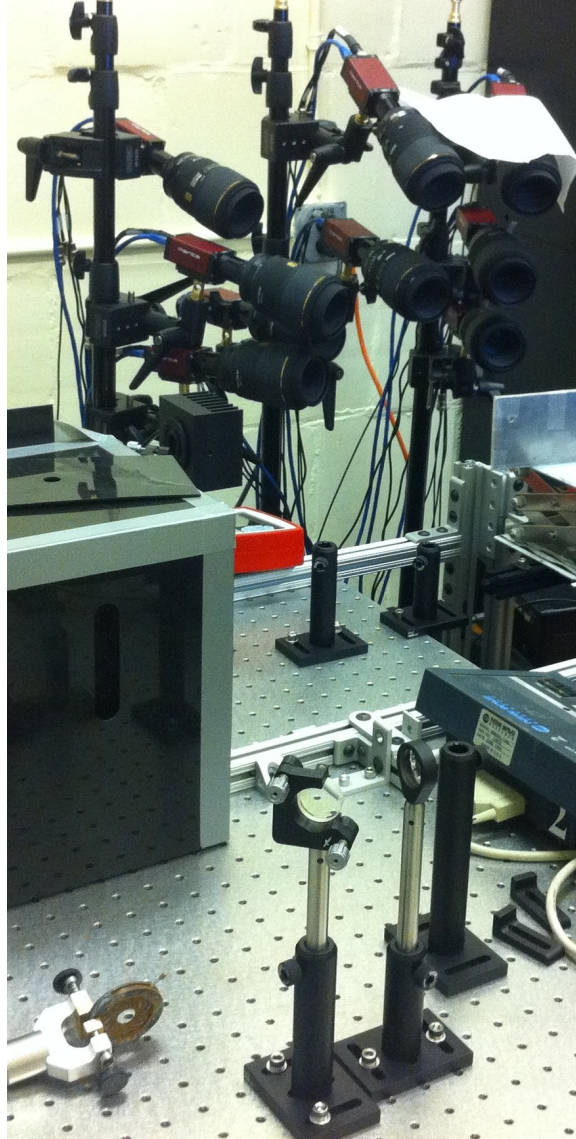
Figure 2-4: High-performance camera array (top) set up for laser-induced fluorescence imaging of two-phase fuel vaporization, from research in which author was involved.

butions of this thesis, and hardware implementations that offer this improved control are presented in Chapter 3.

Most of the previous work in light field imaging has focused on the qualitative aspects of light field photography–those that impact the visual quality of the captured data. Comparatively little has been done to use light field imaging for quantitative 3D scanning, deconvolution or superresolution imaging. This is understandable: the majority of work has been done on light field imaging technologies that capture low
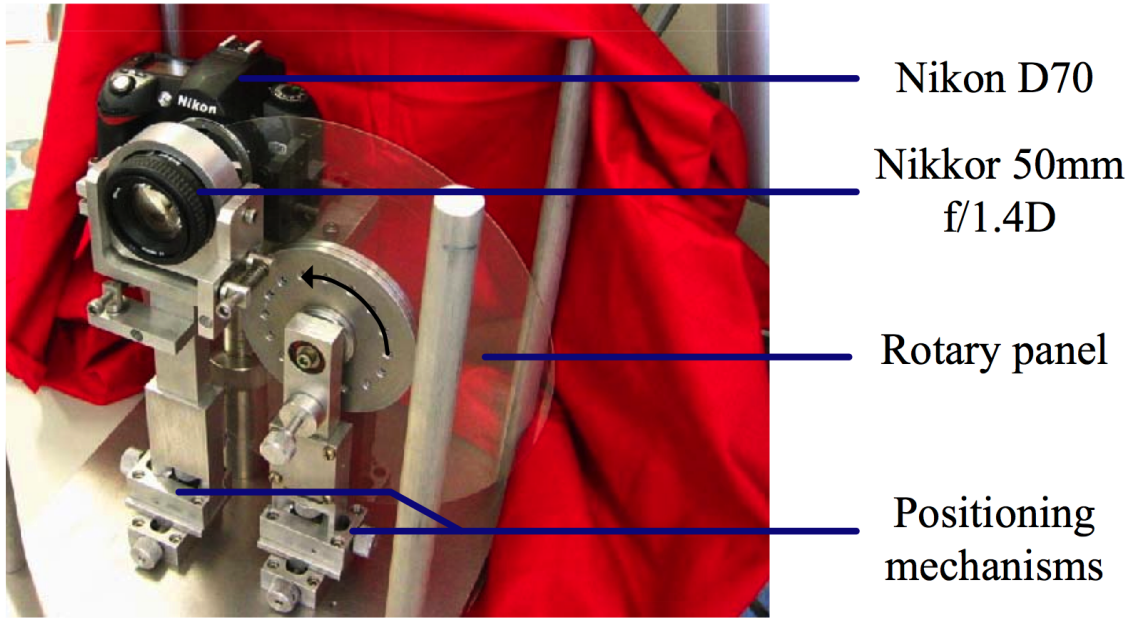
Figure 2-5: Image of rotating-disk prototype "programmable aperture" camera, excerpted from [22]
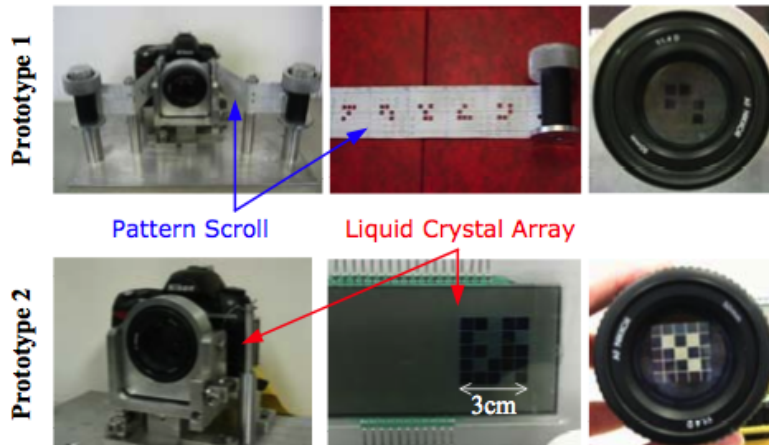
Figure 2-6: Images of prototype "programmable aperture" cameras, excerpted from [21]

spatial-resolution imagery and therefore aren't well suited to qualitative tasks. And the data that comes from large camera arrays can simply be treated by the host of multi-view stereo reconstruction techniques [31], without needing to exploit their light field properties *per se*.

But programmable aperture cameras offer the spatial resolution necessary for high

quality reconstruction in a true light field camera. For stationary objects of modest size, these new cameras are well suited to 3D reconstruction, as [7] demonstrates. With a few critical modifications that will be presented in the following chapters, they are also capable of new kinds of deconvolution and superresolution imaging. These hardware improvements and some of the new computational imaging methods they enable are the central contributions of this thesis, and will be presented in the proceeding chapters.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Digital-Aperture Programmable-Aperture Light Field Camera Prototypes

Figure 3-1: 4th-generation digital aperture light field camera.

A technology that seeks to capture high-quality light fields from a single, ordinary camera lens must be able to sample the aperture plane in an arbitrary way. To achieve this, an array of irises, all of which are individually manipulable, could be made It bears consideration, however, that this would come at the expense of extreme mechanical complexity. Alternatively, a single iris could be allowed to move

around in plane, with size and position adjustable by the user. With this solution, however, positional accuracy would be difficult to achieve and the camera would still be constrained to circular binary apertures.

Ideally, a high-end light field camera would have a very fine grid of solid-state points, each individually programable to any value on the continuum from transparent to opaque. Fortunately, this functionality is ubiquitous in consumer electronics in the LCD displays that are used every day. While most LCDs are used to attenuate light coming from a backlight, many of the same principals apply for transmitting photographic light through the liquid crystal for use as a light switch. Indeed, many Spacial Light Modulators (SLM) exist in the optics world today for projection applications and laser optics. If an LCD with suitable qualities could be placed at the aperture plane, it would be an excellent way to "draw" apertures of arbitrary shape and size anywhere within the aperture plane of a lens. It would even be capable of drawing grayscale aperture profiles; Gaussian profiles, for example, are capable of beating the diffraction limit of conventional binary apertures. A simple optical train for such a system, including the polarizers necessary to "sandwich" the liquid crystal layer, is depicted in Figure 3-2.
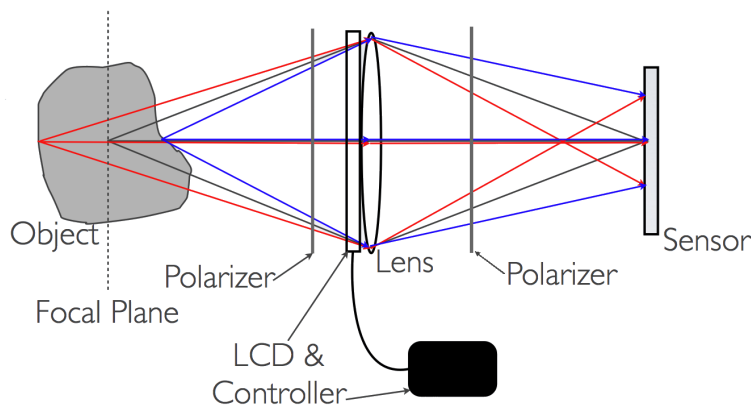
Figure 3-2: Schematic for an LCD-based digital aperture

Before going to great lengths re-engineering camera lenses and substituting aperture plates it was desirable to test whether such a system would work at all. Because sampling the front surface of a lens is similar to sampling at the aperture plane, a

simple first prototype was built by placing an off-the-shelf LCD in front of a standard camera sensor and lens; the system was tested by opening little clusters of pixels and observing qualitatively whether they resulted in favorable imaging conditions.

The first prototype relied on a 5-inch off-the-shelf color TFT-LCD for light switching. Many of the display's component layers–a normal LCD contains a backlight, diffusers, scratch-resistant coatings and other filters–needed to be removed for the experiment, as most were superfluous for this new application and degraded or totally destroyed the image. Even the two polarizing film layers–necessary for the LCD to adjust its transmittance–were not of optical quality and so contributed to a degraded image. They were replaced with separate photographic-quality glass polarizers. Once all of these modifications we made, the prototype system demonstrated exactly the kind of behavior that was predicted by theory.

Once these favorable results with the first LCD were observed, a finer, clearer, higher-resolution monochrome panel from TFT manufacturer Varitronix was purchased. This unit consisted only of the panel itself and a low-level display driver, and required that a custom microcontroller interface be built and that the driver interface code be written from scratch. On top of that effort, several similar layers from the new display need to be removed, as had been done with the earlier display. This again included the removal of the display's OEM polarizer films and their replacement with optical-grade photographic versions. A photo of this camera is shown in Figure 3-3, and salient excerpts from the display's datasheet can be found in Appendix A.

These prototype cameras were solid successes; they proved qualitatively that this new "digital aperture" approach to aperture sampling had great potential. An exemplar light field capture from the latter system can be found at

```
http://www.ThomasBradfordMilnes.com/DigitalAperture/aperture.swf?
lightfield=ColorMatchboxBajaCarGen2Point5.zip
```
While qualitative validation was intellectually satisfying, these early cameras also demonstrated all-to-well the unexpected phenomena of diffractive ghosting, resulting from the light passing through the small matrix of wire traces that serve as the boundaries between pixels in many LCDs, including the ones used in these exper-
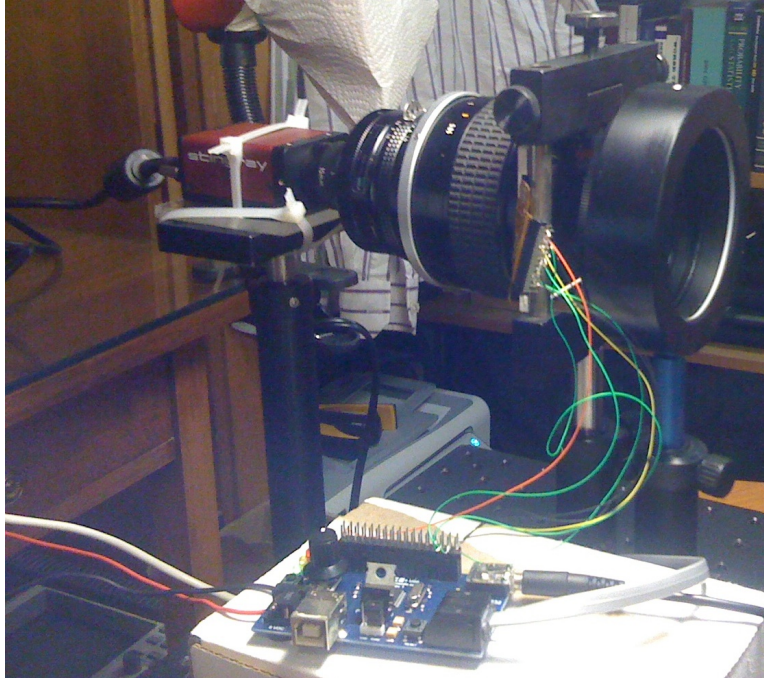
Figure 3-3: An image of the second prototype camera showing, from left to right, a stock AVT camera sensor, Nikkor lens, Varitronix display (in optical mount, pressed up against the lens), and photographic polarizer. A PIC micro-controller and the custom-made wiring harness that connects it to the Varitronix display are also shown.

iments. The author has come to understand this phenomena as one of the largest considerations in designing digital-aperture cameras.

3.1 Diffraction at the Aperture Plane

A standard LCD is composed of a grid of pixels which are transparent in the middle and opaque around the edges. The opaque grid is formed by conductive traces that allow for the electrical addressing of each pixel. The pixels and their traces form a regular 2D pattern of transparent and opaque regions–essentially a 2D diffraction grating. The Varitronix display from the previous section and its individual pixels are shown in Figure 3-4. These microstructures are hard to notice when looking at smartphones and computer screens because these displays are illuminated from behind by a diffuse light source. However, when a structured field of light passes through a diffraction grating and then through a lens, ghosting occurs in the resultant image.

38

Ghosting is the staggered superposition of an image with itself. Figure 3-5, taken from the light field linked at the URL mentioned above, demonstrates this effect clearly. The ghost images exhibit chromatic separation; the blue portion falls closest to the original image because of diffraction's wavelength dependence.
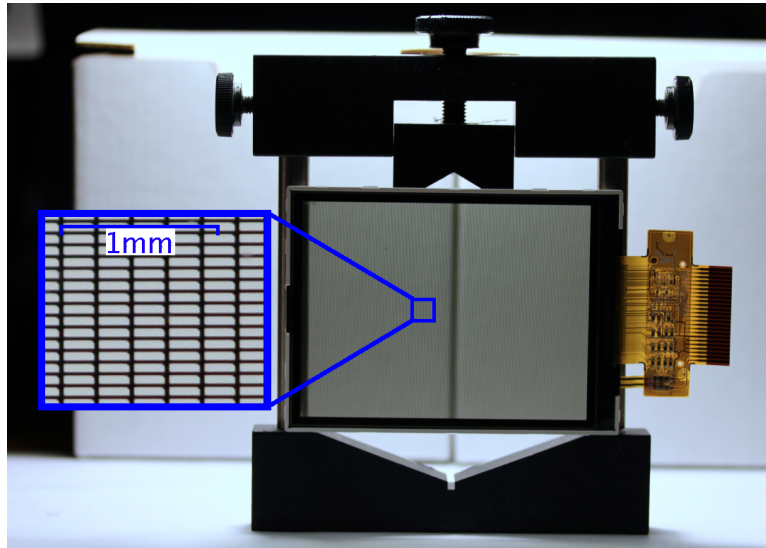
Figure 3-4: This image shows the Varitronix monochrome TFT LCD we used as an aperture; the inset shows individual pixels. Note that while monochrome, this display still has RGB subpixels (3 segments to a pixel), just not the color filter layer that would make it a color display.

The diffraction of collimated light passing through a grating is depicted in Figure 3-6. When a lens is placed immediately after the grating it creates an imaging condition that exhibits ghosting. This is depicted in Figure 3-7. The positions of the ghosted images can be derived from first principles, as follows: the diffraction equation for predicting angles of intensity maxima is

$$d(\sin\theta_m + sin\theta_i) = m\lambda \tag{3.1}$$

where d is the grating spacing, θ_m is the angle of the mth intensity maxima, θ_i is the angle of incidence of the light with the grating, m is any non-zero integer, and λ is the wavelength. Assuming that the field of view is small for the system in question (true in the case of the aforementioned prototype cameras), then the small angle approximation can be applied, simplifying this equation to

39

Figure 3-5: An image from one of our early light fields. The picture exhibits heavy ghosting, a result of the LCD behaving as a 2D diffraction grating. The ghosting effect is particularly noticeable around the Matchbox car's front fender.
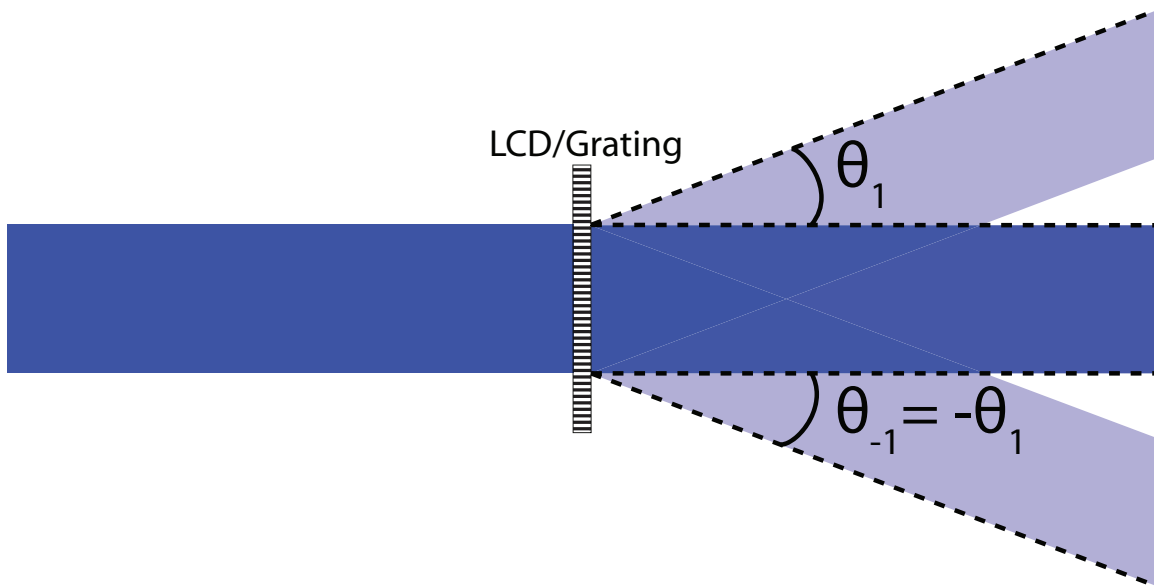
Figure 3-6: This figure depicts the diffraction of a collimated beam passing through a grating. Note that only the zeroth and first diffraction orders are shown.

$$d \times \sin \theta_m = m\lambda \tag{3.2}$$

Reorganizing this equation to solve for the first diffraction order angles yields
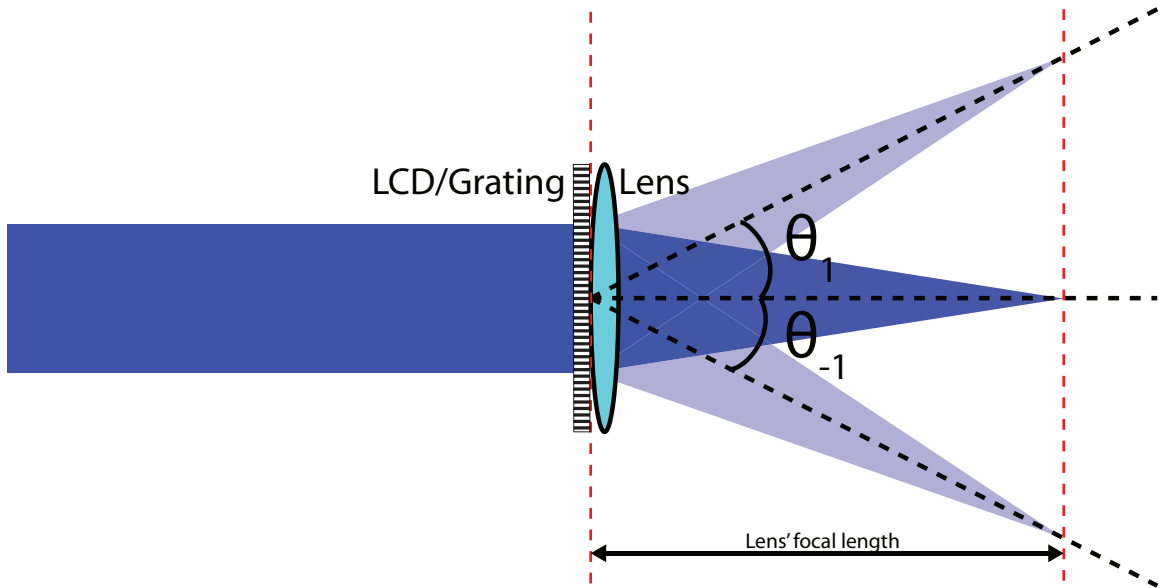
Figure 3-7: When collimated light passes through a diffraction grating and then immediately through a lens, an image is formed for each of the diffraction orders. Note that only the zeroth and first diffraction orders are shown.

$$\theta_{\pm 1} = \pm \sin^{-1} \frac{\lambda}{d} \tag{3.3}$$

This result means that ray bundles from the object being imaged pass through the diffraction grating and continue on as they would have, while also generating identical ray bundles emanating off to each side by angles of θ_m. Thus, a lens that immediately proceeds such a grating will form an image from the original bundle, and also form images from the diffracted bundles. Lenses behave as transducers–they convert angles to distances and vice versa–according the the equation

$$\tan \theta = \frac{d_i}{l} \tag{3.4}$$

where θ is the angle between two ray bundles, d_i is the distance at the focal plane between the images generated from the two ray bundles, and l is the distance between the lens and the focal plane. When the rays are coming from infinity they create parallel-ray bundles and the distance l simply becomes the focal length. To within a first-order approximation, the rays in each bundle do not need to be exactly parallel, so long as the two bundles are offset by an angle, like what happens at the LCD's

diffraction grating. Setting the transduction angle equal to the 1st diffraction order angle yields

$$d_i = l \times \tan \sin^{-1} \frac{\lambda}{d} \tag{3.5}$$

which can be simplified–by invoking the small-angle approximation–to

$$d_i = \frac{l\lambda}{d} \tag{3.6}$$

giving an elegant result: the offset of the first-order ghosting image is a simple function of the grating distance, wavelength and focal plane distance. This can be tested empirically with the values from the camera that captured the image in Figure 3-5. The relevant values for that camera are shown in Table 3.1. If an intermediate wavelength in the green region is chosen, and the relevant camera values are plugged into Equation 3.6, the result is

$$d_i = \frac{100mm \times 550 \times 10^{-6}mm}{0.0745mm} = 0.74mm = 90pixels \tag{3.7}$$

This compares favorably with Figure 3-5: measurement of the original image yields a distance of about 86 pixels from a point on the 0th-order image to that same point in the green portion of the 1st-order ghost. This is depicted in Figure 3-8.

Table 3.1: Selected properties of the Varitronix/Nikkor/AVT Stingray camera

Lens focal length:	85mm
Distance to focal plane:	~100mm
Vertical grating distance:	0.0745mm
Horizontal grating distance:	0.2235mm
Sensor size:	780x580 pixels
Sensor pixel pitch:	8.3μm

Figure 3-8: This is the same image as in Figure 3-5 with the vertical 1st-order ghosting distance measured in white.

3.2 Designing Away Diffractive Ghosting

The first two cameras were capable of taking light fields with as many 225 images (a 15x15 grid). These light fields demonstrated the perspective changes that theory predicted. However, the strong diffractive ghosting made it nearly impossible to use them for quantitative 3D reconstruction. In order to build a light field camera capable of reconstructing 3D, the ghosting should be eliminated entirely.

Equation 3.6 demonstrates that the ghosting distance is a straight-forward function of three camera specifications. It is therefore possible, by thoughtfully choosing those parameters, to make the ghosting distance equal to or larger than the size of the image sensor. In this case, an object that appears in the 0th-order image will not appear in a ghost anywhere else on the sensor. This is half the solution. It is also necessary to ensure that only the rays that can strike the sensor in the 0th-order image make it to the SLM. If both of these requirements are met, the 0th-order image will fill the sensor and the blue regions of the 1st-order ghosts will flank the sensor neatly above and below without spilling over into it. A schematic of how this can be achieved is shown in Figure 3-9.

To satisfy the first requirement, an SLM with a pixel pitch (the grating distance)
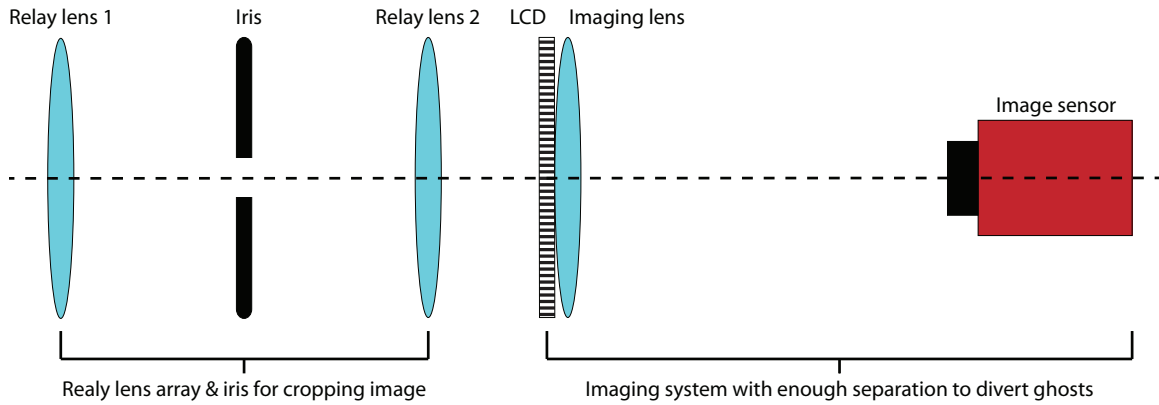
Figure 3-9: Schematic of a simple optical train capable of cropping an image and overcoming diffractive ghosting.

as small as possible was chosen, and the distance from the lens to the focal plane was made appropriately large. To satisfy the second requirement, relay lenses and field stops were arranged to create an intermediate image that could be cropped to exactly the size needed to fill the sensor.

Ultimately, a TFT microdisplay from Sony was selected. This microdisplay was designed for projector applications; its incredibly small pixel pitch was about as small as could be found anywhere. Some of its properties are listed in Table 3.2; additional details can be found in excerpts from the product's datasheet included in Appendix A. The microdisplay's pixel pitch allowed for a distance of about 350mm between the imaging lens and the image sensor. Given this distance and relying on Equation 1.1, a 300mm imaging lens was chosen, allowing for the imaging of objects from 0.6-2.1 meters without a ghost overlapping the sensor. Focus could be adjusted simply by sifting the camera and iris forward and backward appropriately. A prototype, which was referred to as the 3rd-generation camera, and which utilized the microdisplay, was tested and confirmed to work on an optical bread board; it is shown in Figure 3-10.

While this working distance was a good range for imaging the kinds of things that were of interest, there was one practical issue with the focal length: 300mm lenses zoom in extremely close to the imaged object, yielding a very narrow field of view. In order to get a normal field of view, the relay lens system was modified to provide

44
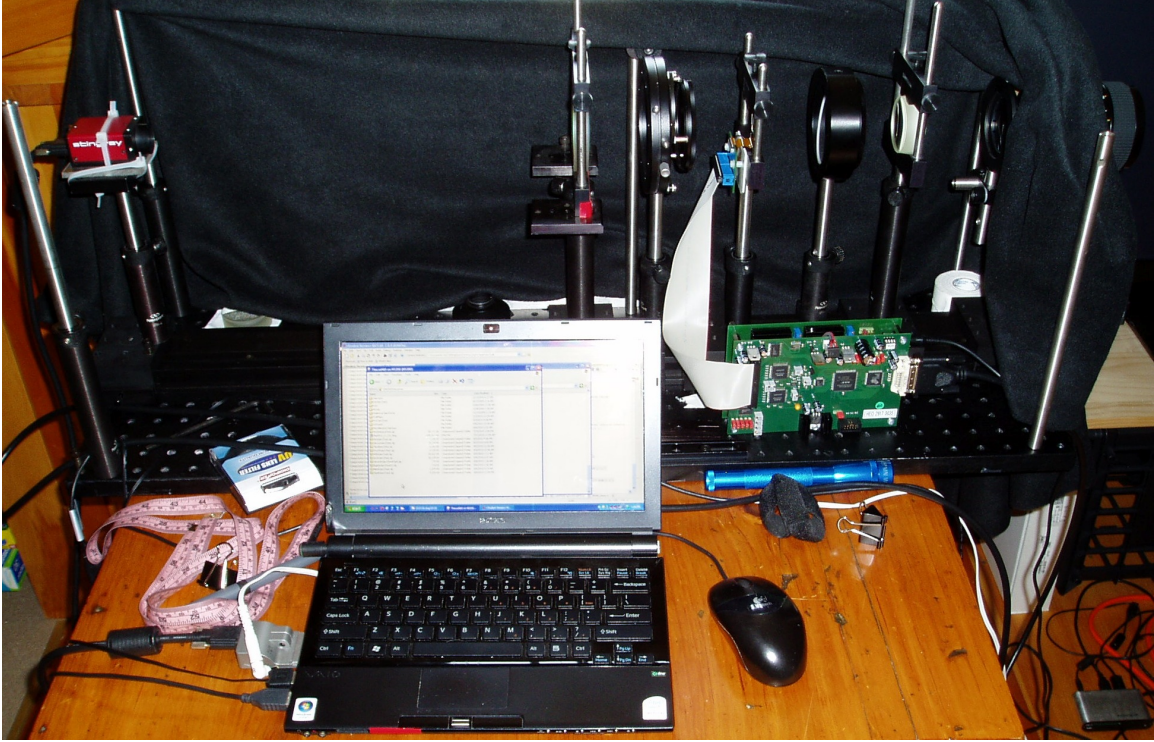
Figure 3-10: Photo of a preliminary prototype built to test the design paradigm. The microdisplay, driver board, image sensor and polarizers in this photo were used in the final camera.

demagnification as well as image cropping. By using two lenses with focal lengths f_1 and f_2, with $f_1 << f_2$ and a distance between them of $f_1 + f_2$, a de-magnifying telescope (or a beam condenser) was created that was also capable of cropping the image. Lenses of $f_1 = 75mm$ and $f_2 = 300mm$ were chosen, giving a magnification of $\frac{1}{4}$, and giving the overall system the effective field of view of a lens of focal length

$$f_{eff.FOV} = f_{imaging} \times \frac{f_2}{f_1} = 300mm \times \frac{75mm}{300mm} = 75mm \tag{3.8}$$

meaning that the system has the same field of view as a regular 75mm lens. This is a moderate value that falls right in the middle of the consumer grade camera lens range, and has proven to be reasonable for testing.

The precise specifications of each component in the optical train are listed in tables below. The properties of the microdisplay can be found in Table 3.2. The specifications for the lenses and polarizers used in the optical cavity can be found in

Table 3.3. Finally, the specifications for the AVT Stingray camera that serves as the image sensor can be found in Table 3.4. The final setup is depicted schematically in Figure 3-11.

Table 3.2: Sony LCX023 microdisplay specifications

Resolution:	1024x768 pixels
Diagonal size:	33mm
RGB subpixels:	No
Pixel pitch:	0.026mm x 0.026mm
Aperture ratio:	67%
Optical transmittance:	23%
Driver board:	Holoeye HED-0017 board
Driver board interface:	DVI

Table 3.3: 4th-Generation Camera Optical Elements

Lens 1:	Edmund Optics NT49-291 achromatic doublet
	f=75mm, $\varnothing$=50mm, MgF_2 coating
Lenses 2 & 3:	Edmund Optics NT45-418 achromatic doublet
	f=300mm, $\varnothing$=75mm, MgF_2 coating
Polarizers:	Hoya photographic linear polarizing filters
	77mm, 30% transmission, 95% efficiency
Cropping mechanism:	$\sim$ 15mm adjustable steel-leaved iris

Table 3.4: AVT Stingray Camera Properties

Model:	AVT Stingray F-046C
Resolution:	780x580 pixels
Sensing area:	6.4x4.8 mm
Pixel pitch:	0.0083x0.0083 mm
Pixel depth:	8-14 bits
Interface:	Firewire 800
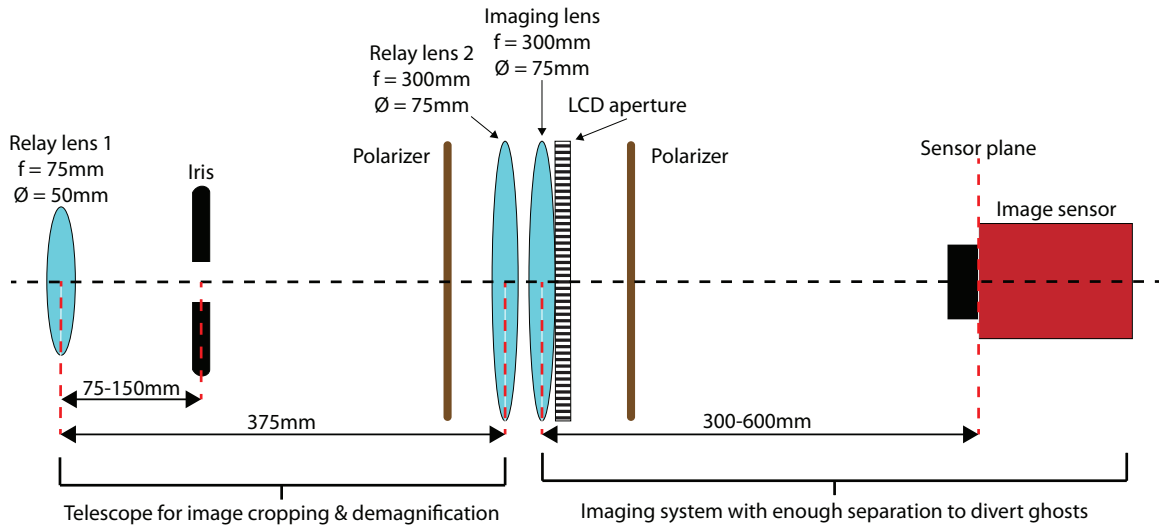
Figure 3-11: Schematic of the final optical train design, capable of cropping an image, overcoming diffractive ghosting, and increasing the imaging lens' field of view.

3.3 Making It Portable

While the successful design of the camera was the largest, most import acheivement in the project, constraining the prototype to a permanent setup on an optical bench severely limited the variety and composition of data that could be captured. To allow the widest range of potential objects that could be imaged for testing, the camera needed to be portable and compatible with a standard photographic tripod so that it could be used in much the same was as a normal consumer camera. But the prototype design was large–nearly a meter long–and required high precision for alignment of its many optical parts. How could a camera body be designed and built that could safely, securely, and precisely carry the many parts in its optical cavity while being portable, light-tight, relatively inexpensive and easy to manufacture?

A basic design was chosen based on the optical table prototype, which was built around a dovetail rail running parallel to the optical axis with each component mounted on standard rail carriers. The rail and carriers can be seen in Figure 3-10. For the portable design, it would be best if each component was mounted to a custom carrier that rode along a single rail running down the body. This would make adjustments easy and allowed for a modular design architecture that was flexible about which parts were used and in which order; it would also facilitate adjusting the

47

camera and iris positions to maintain sharp focus anywhere in the camera's working distance. The long body also needed to have a relatively small cross-section to make it as light and portable as possible while still accommodating the largest of its optical parts. Finally, the design would need to be light-tight so that even when out in bright sunlight the sensor would only be registering light from the image.

While a machined metal camera body would certainly be the most rigid, it would also be heavy, difficult to machine intricately, and expensive. Machining the body out of plastic would make it lighter and cheaper, but would still limit the geometries that could be produce. For this application, 3D printing seemed like the best choice. However, a printer that used inexpensive filament plastic was preferred over the very expensive plastic cartridges used by industrial printers. This would allow for the construction and testing of many parts and the rapid iteration of designs at very low cost. The highest-quality printer with these qualities was the Bits From Bytes BFB 3000, one of the then-new open-source printers originally designed at labs at MIT, Cornell and others. One was purchased, pre-made, for $4000. A stock photo of the printer is shown in Figure 3-12; the actual printer used to build our system, hard at work, is shown in Figure 3-13. Specifications for the printer can be found in Table 3.5.

Figure 3-12: Stock photo of the Bits For Bytes BFB 3000 3D printer, which was used to build most of the parts in the new camera.

As can be seen from Table 3.5, the printer isn't capable of printing the meter-long

Figure 3-13: A photo of the BFB 3000 used for this project hard at work.

Table 3.5: BFB 3000 3D Printer Specifications

Model:	Bits From Bytes BFB 3000 Panther Dual-Head
X & Y positioning accuracy:	0.05mm
Z layer minimum thickness:	0.1mm
Printable area:	285x300x200mm
Max print speed:	$15\text{mm}^3/\text{s}$
Interface:	SD Card

body in one shot. This necessitated a re-design of the body into discrete segments. However, discrete sections would greatly reduce the structural rigidity of the final body. To solve this issue the new body was designed to have a steel backbone: four integrated press-fit 1/4-20 threaded rods pulling all of the sections together firmly, with section lengths of roughly 200mm, each containing its own length of rail. An STL model of a section is shown in Figure 3-14.

In the end, the camera design required about 30 unique printed parts and nearly 100 fasteners, tape lengths, and other off-the-shelf parts. The total construction time for the camera was about 3 weeks, start to finish, including printing and assembly. Rather than go into the details of the design and construction here, it will suffice to
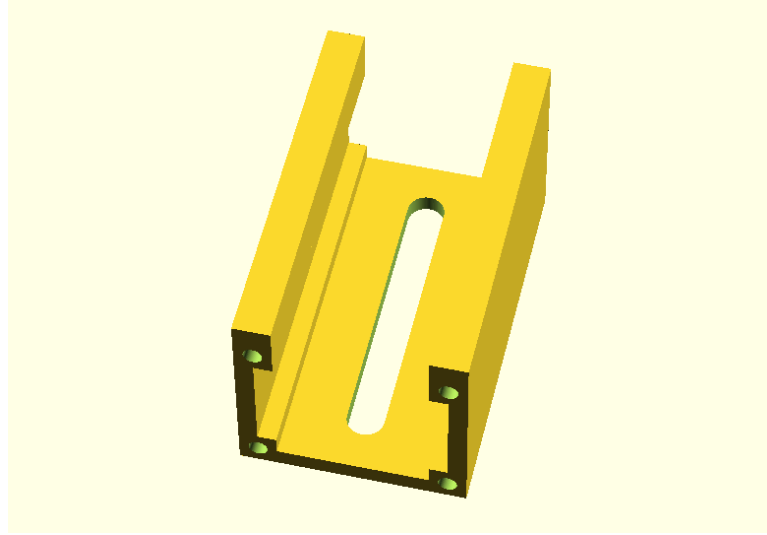
Figure 3-14: A CAD model of a 200mm section of the camera body. Note the rail slot on the bottom for carrying optical components and the through-holes running along each corner for the threaded rod backbone.

show several images depicting the camera's salient features. After all, "a picture is worth a thousand words:"

Figure 3-1 at the beginning of this section shows the camera with its lids attached. Removing the lids reveals the contents of the optical cavity inside, shown in Figure 7-4a. At the front of the cavity are the 75mm-focal length lens and the image-cropping iris, shown up close in Figure 7-4b. The polarizers, 2nd and 3rd lenses, the Sony microdisplay and their carriers are shown in Figure 7-4i, and the microdisplay itself can be seen through the 2nd and 3rd lenses in Figure 7-4j. The AVT Stingray and its sensor can just barely be seen in the back of the cavity behind the ribbon cable in Figure 3-16a. The Holoeye HED-0017 driver board is shown in Figure 3-16c; it is connected to the microdisplay by the gray ribbon cable. Figure 3-16d shows the camera's electrical connections as well as some textual finishing touches enabled by the 3D printer. Finally, Figure 3-16b, taken of the rear underside of the camera, shows the mounting bolts for the rail carriers and the black duct tape used to cover the empty portions of the rail slots in order to keep the optical cavity light tight.

(a) 4th-generation camera with lids removed to reveal the optical cavity

(b) First lens and the image-cropping iris

(c) Over-head view of polarizers, 2nd & 3rd lenses and the Sony microdisplay

(d) Close-up of the Sony microdisplay as seen through the 2nd & 3rd lenses

Figure 3-15: 4th generation camera images

(a) AVT Stingray sensor at rear of optical cavity, barely visible under the ribbon cable that connects the microdisplay to the driver board
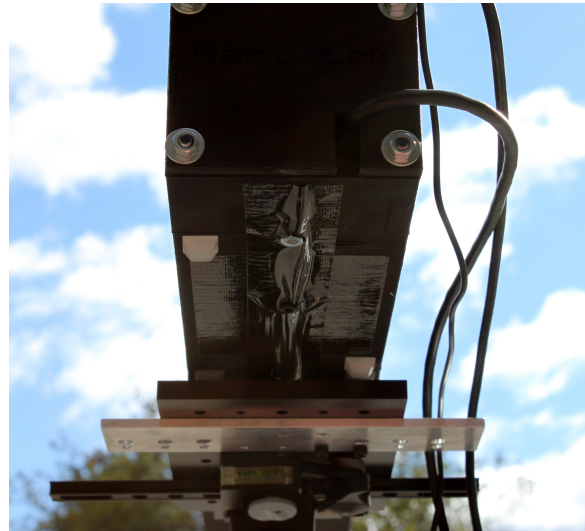
(b) Rear underside of the camera, tape covering optical track for light-tightness

(c) Holoeye HED-0017 driver board. The apparent warping of the parts is actually radial distortion from the wide-angle lens that capture the photo.
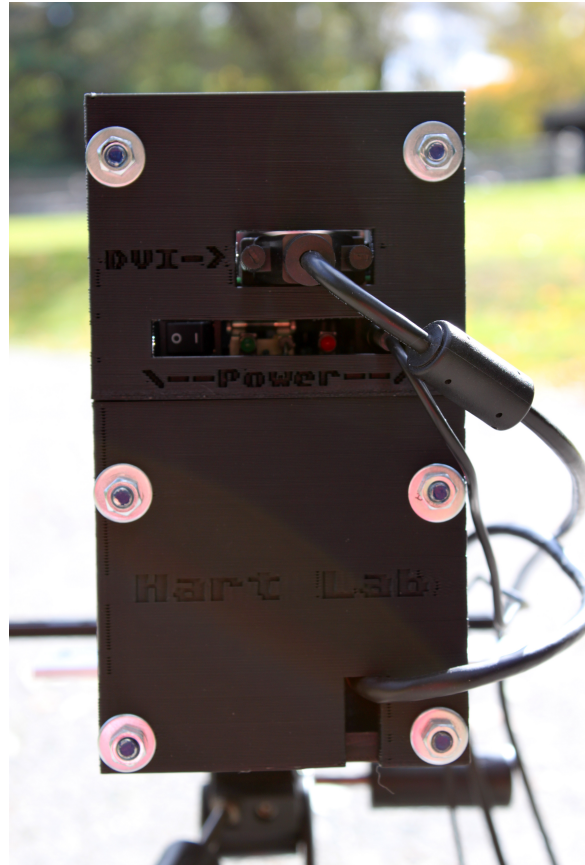
(d) Camera's rear panel & electrical connections

Figure 3-16: 4th generation camera images

3.4 Results From Portable Prototype Camera

The data quality achieved by this new camera is far beyond the quality achieved by its predecessors. There is no diffractive ghosting, and the ability to adjust the iris and sensor positions for sharp focus makes for very clear images. The use of achromats throughout keeps chromatic aberration to a minimum. Two tests were sufficient to prove that the camera's image quality met desired standards. The first was to capture a light field and note that 1) the images are ghost free 2) they have minimal chromatic aberrations and 3) the light field exhibits pronounced, aperture-dependent shifting in its images. Readers are invited to confirm these results for themselves by viewing the light field at the link below.

```
http://www.ThomasBradfordMilnes.com/DigitalAperture/aperture.swf?
lightfield=ThirdGen4LightField.zip
```

The second test was to take several center-aperture images from different angles of the same wooden puzzle and see whether these images were of high-enough quality to build a 3D model using multi-view stereo reconstruction techniques. The answer was resoundingly yes–all the images taken of the wooden puzzle were usable for the reconstruction. The 3D reconstruction can be viewed at the link below, and a screenshot of the model can be found in Figure 3-17.

```
http://www.thomasbradfordmilnes.com/Research/WoodenPuzzle/MITViewer.
html
```

3.5 Notes on Hardware Shortcomings

The results from Section 3.4 demonstrate the viability of the ACPA camera paradigm. Furthermore, they demonstrate that such a system can be made cost-effectively with consumer-grade parts. However, it is worth noting that the restriction to off-the-shelf parts did create some compromises in quality.

Firstly, the optical train design required to crop the ghost images in our system had an extremely small field of view and numerical aperture. As a result, though
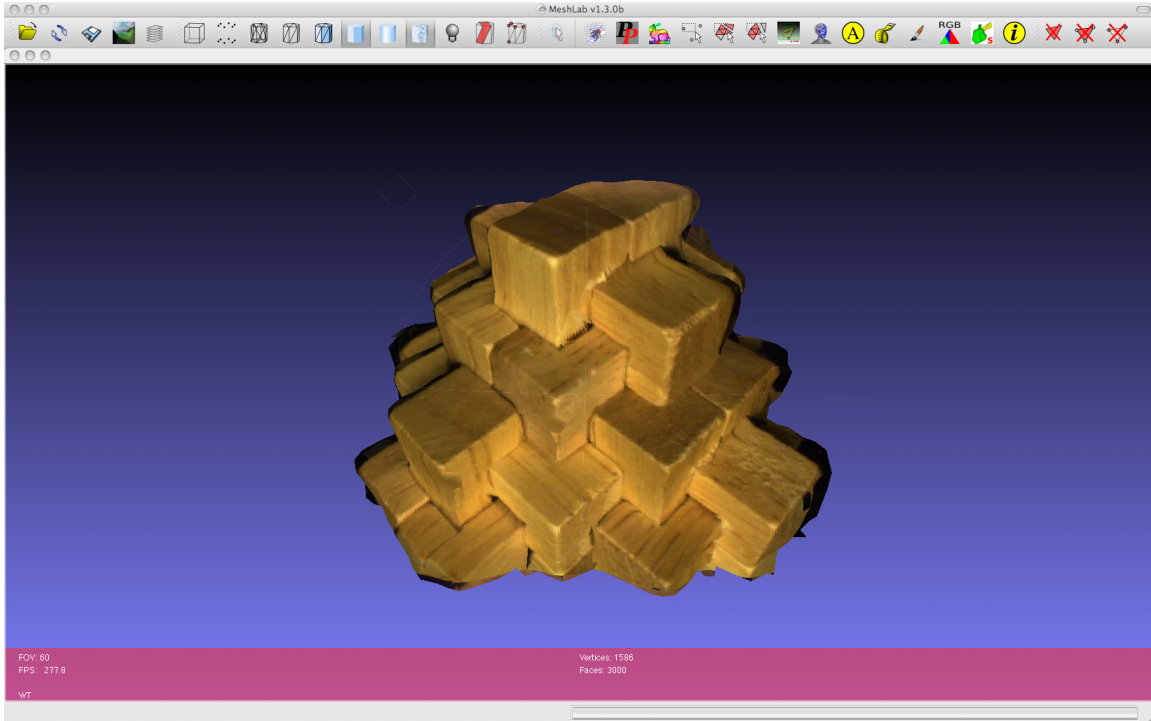
Figure 3-17: Screenshot of a 3D model made from images taken by the digital-aperture light field camera.

diffractive ghosting was eliminated, the system exhibited an unusually low-resolution diffraction limit because of the small numerical aperture. The Airy disk size for the system was larger (on the order of 10 pixels) than the pixel pitch in the AVT sensor, so even when properly focused the images exhibited some noticeable blurriness. In order to make small tweaks to these settings, an additional few prototype camera setups were eventually built using standard optical table mounts and a long dovetail rail system for linear adjustments.

The second compromise on quality came from the opaque traces in the LCD. Even though the diffractive ghost images had been removed, the tiny matrix of traces still had an attenuating effect in the frequency domain, creating an optical transfer function that suppressed a lot high spatial-frequencies.

Both of these compromises could be solved with custom-made LCD panels that have high contrast ratios *and* transparent traces. The author consulted LCD manufactures, including Varitronix, about this issue, and the problem is not insurmountable. All that is needed is an LCD with transparent traces, which exist in many

low contrast-ratio displays like those used in digital watches and small dot-matrix displays. Such an LCD isn't necessary for (or particularly well suited to) common back-lit applications such as phone screens and computer monitors, so they are vanishingly rare. But they could be made–at a unit price close to those of current displays–if and when productization of a ACPA camera is desired. There is a large up-front cost on the order of $1MM for setting up a new display fabrication line (which is why it was cost prohibitive to make only a handful of displays for testing purposes), but the production costs at high volume would be similar to LCDs of similar sizes and specifications.

3.6   Future Work

Once a high-quality digital-aperture camera had been built and demonstrated, research could be turned to developing computational techniques for utilizing its captured data. There has been much work to date on reconstructing 3D from a collection of regular images. This is commonly know and multi-view stereo reconstruction. One goal was to modify the best-in-class multi-view stereo techniques to work with light field data. Because light field data are different from regular images, there are some salient modifications that need to be made. For example, while epipolar lines constrain the location of a point in one image to a line in another, in a light field that constraint is instead along the vector of aperture movement. Further, while multi-view stereo relies on triangulation to determine depth, depth in a light field is encoded as a function of image movement and aperture movement, derivable from Equation 1.2.

The porting of high-performance multi-view stereo techniques for use with light field data will be a powerful step forward for digital-aperture technology. One of multi-view stereo's biggest weaknesses–and largest sources of error–is determining the original positions of the camera that took the images. With a digital aperture, the original position of the aperture is known *a priori*, and known to a very high degree of accuracy because of the regularity of the LCD's fabricated pixel structure.

A successful port would leverage all of the strengths of multi-view stereo while eliminating a great weakness. In the long run, this would mean that regular consumer cameras could be equipped with digital apertures and be used for high-accuracy 3D scanning on top of their regular photographic functionality.

Additional projects of great interest were in using the light field data from a digital aperture camera for superresolution imaging and for deconvolution. The image manipulation available with a digital aperture allows for the encoding of important information into captured images that isn't available with traditional imaging systems. These manipulations include spatial-frequency response manipulation and sub-pixel image shifting. Two of the following chapters will go into detail about how such manipulations can be utilized for superresolution imaging and deconvolution.

Chapter 4

Calibration and Point Tracking with Digital-Aperture, Programmable-Aperture Light Field Cameras

4.1 Introduction to Calibrations Techniques

The work done on Active Wavefront Sampling (AWS) in [7] provides a solid foundation for the calibration of rotating-aperture cameras. These techniques can be extended directly to "digital-aperture" (i.e. arbitrarily-controllable programmable-aperture (ACPA)) light field cameras. The author of [7] offers two techniques for calibrating rotating-aperture cameras: one based on look-up tables (LUTs), and one based on creating a direct lens aberration model. For the purposes of this thesis, and because of the diffraction-based limitations of prototype ACPA cameras (as discussed in Chapter 3), only the look-up table approach will be covered in this chapter. The lens aberration model, however, could be applied as soon as better-quality ACPA cameras can be produced.

A LUT, in the context of AWS and ACPA imaging, is a "volume" of points

in space whose 3D coordinates are know and whose light field image positions and disparities are also known. With this "volume," an arbitrary point's position in 3D space can be interpolated by comparing its image position and disparity with those points in the LUT whose image positions and disparities most closely match it. The author of [7] created look-up tables by imaging at various distances in front of the camera a machined aluminum plate with a grid of points on its face. By knowing the grid spacing and the various positions of the plate in Z, the 3D coordinates of each grid point at each plate position could be calculated. Then, the image disparity for each of these grid points at each plate position was calculated; the 3D coordinate coupled with that point's image position and disparity served as a reference point in the calibration volume, and together all of these points formed the calibration LUT.

4.2   Calibrating an ACPA Camera

To generate LUTs for an ACPA camera, a system similar to the one used in [7] was constructed. Rather than use a plate with points on a grid though, a standard checkerboard target (Figure 4-1) was used instead. The corners of the checkerboard pattern were used as the reference points, and their image locations where determined using standard corner-detection algorithms available in the OpenCV [5] library. An image of the checkerboard target captured by an ACPA prototype using the Sony-microdisplay described in Section 3.2, with corners detected and marked, is shown in Figure 4-2.

For an ideal lens with a rotating aperture, points in the image sweep out circular paths as the aperture rotates. For real lenses, however, that ideal path is corrupted by lens aberrations. Figure 4-3, excerpted from [7], shows these non-ideal paths for a near-axis and off-axis point in a rotating aperture system. Not surprisingly, the aberrations of points far from the optical axis are aberrated more than those that lie near the image center.

This point-tracking approach can be extended to ACPA cameras by moving the programmable aperture in a grid pattern, and tracking the image points from one
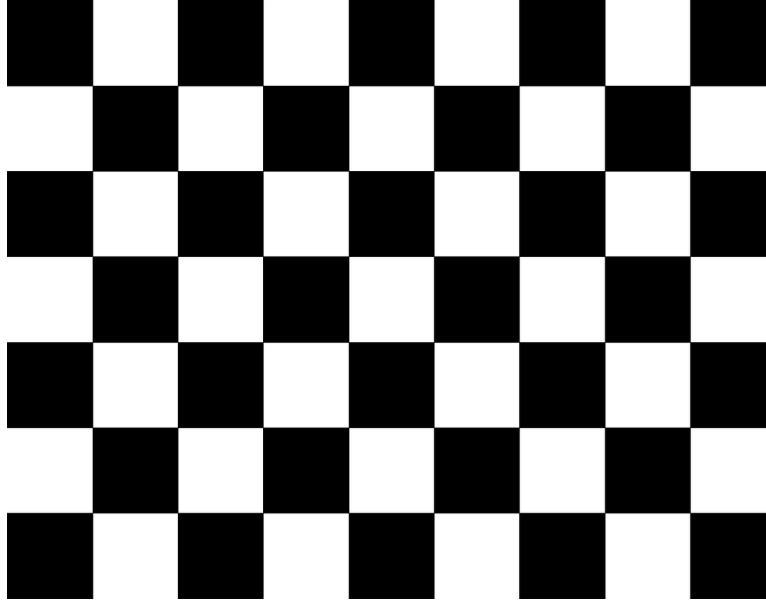
Figure 4-1: Common checkerboard calibration target.

position to the next. Ideally this would result in a square grid of disparities, and in reality results in an aberrated grid. Results from a Sony-microdisplay based system (with sampling specifications shown in Table 4.1) are shown in Figure 4-4, with a close up of four of these regions shown in Figure 4-5. These graphs demonstrate that ACPA cameras can be calibrated using standard machine-vision techniques and a simple extension of previous work.

Table 4.1: Specifications for ACPA camera with Sony-microdisplay aperture, used to capture point-tracking data

Microdisplay aperture:	Sony LCX023
Microdisplay resolution:	1024x768 pixels
Microdisplay driver board:	Holoeye HED-0017 board
Imaging sensor:	AVT Stingray F046C
Sensor resolution:	780x582 pixels
Light field sampling:	5x5 grid of rectangular adjacent apertures

4.3 Tracking Points on an Object for Scanning

Once a camera is calibrated, it should also be able to scan objects other than checkerboards. One of the potential values of ACPA cameras is that they can be used for
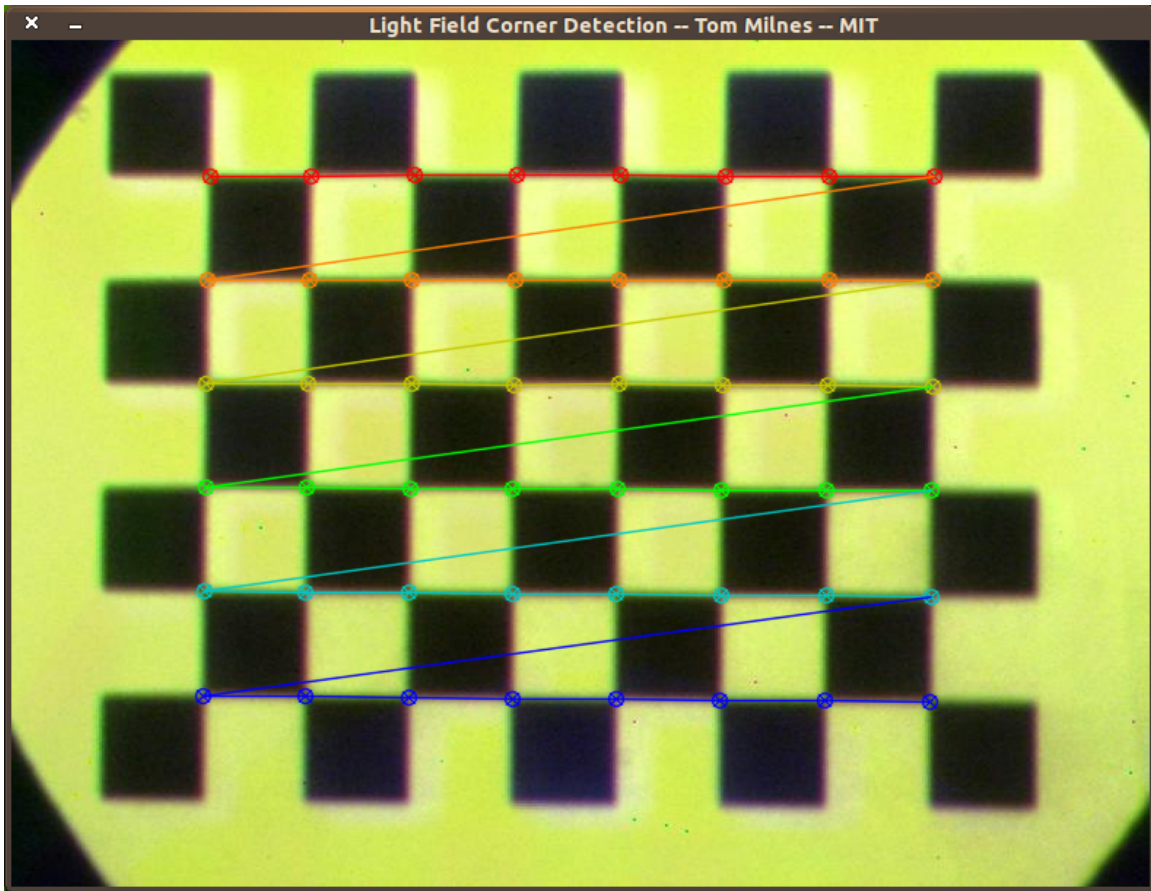
Figure 4-2: Corner detection in OpenCV [5] performed on a calibration image from a Sony-microdisplay based ACPA prototype camera.
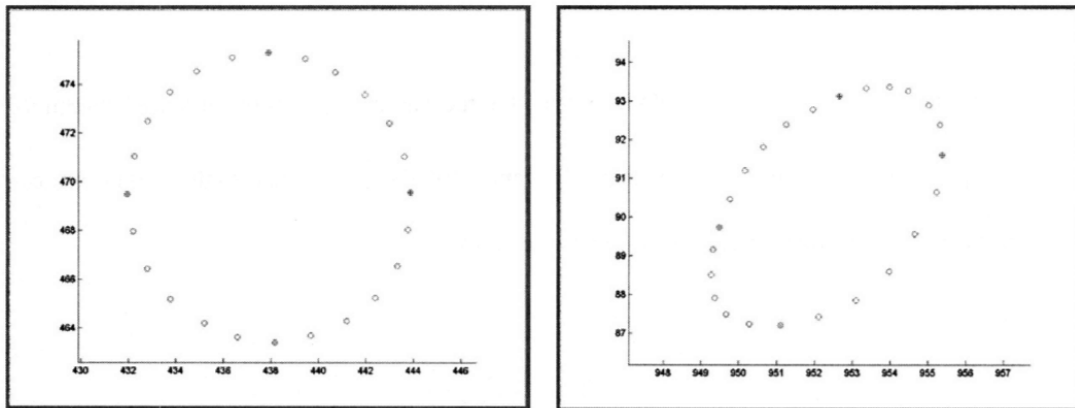
Figure 4.2. Interrogation area image paths, for an experiment where the aperture sampled 24 evenly spaced positions on the sampling circle. On the left is the path of an interrogation area located near the center of the image. On the right is the path of an interrogation area located far from the axis. All dimensions are in pixels.

Figure 4-3: Point tracking with an Active Wavefront Sampling camera, from [7].
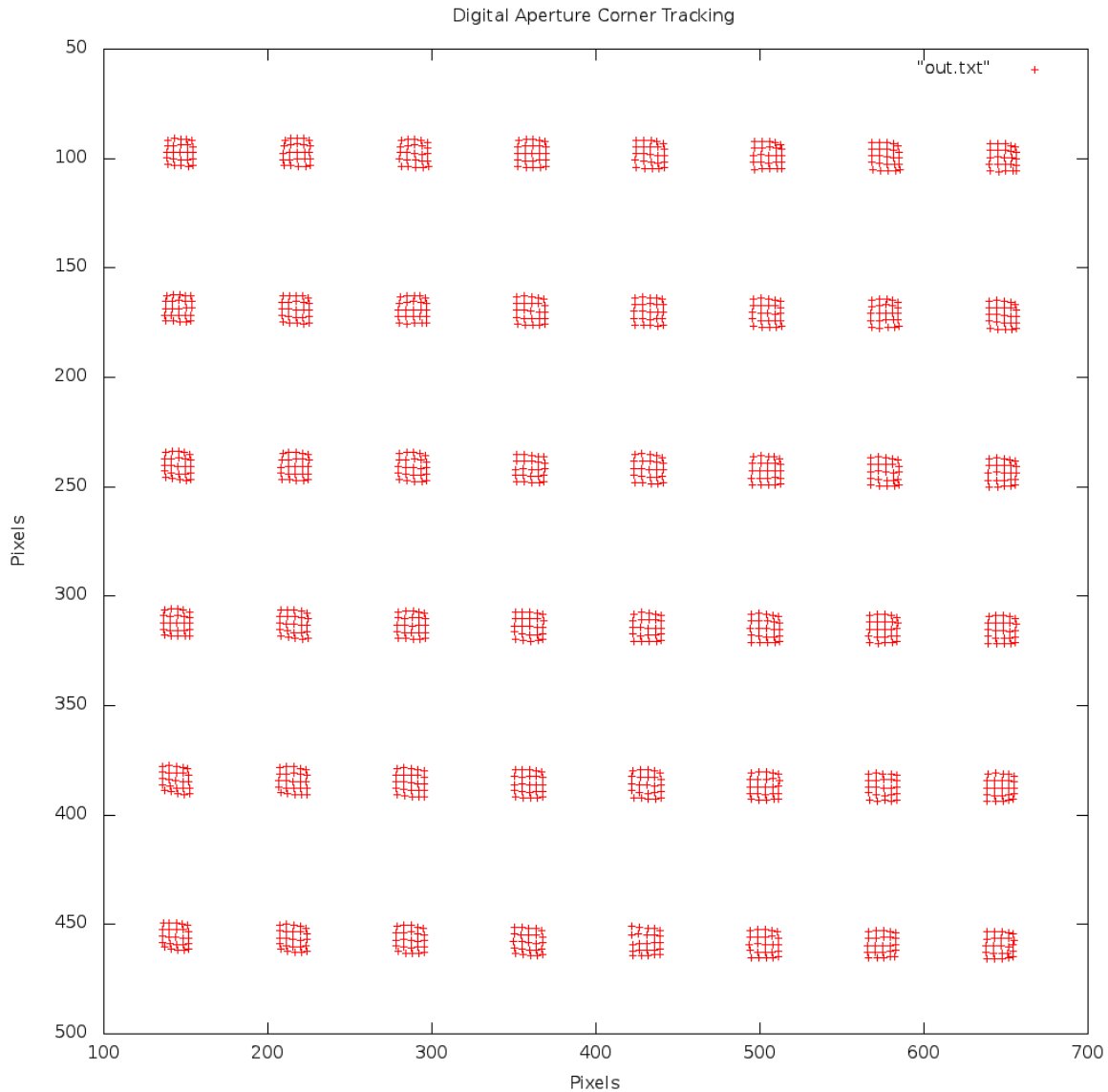
Figure 4-4: Tracking the corners of a checkerboard target across the whole sensor frame of an ACPA camera using a 5x5 grid of apertures

scanning without the need for structured illumination. In order to track points on the surfaces of objects, the object must have some amount of texture. Additionally, since most objects aren't covered in "checkerboard" wrapping, a different set of algorithms must be employed to identify unique points on the textured surfaces. The OpenCV library [5] contains a number of effective optical-flow algorithms that can be used for this purpose, and they were utilized for the test results that follow.

Because texture is important for strong tracking, a feature-rich target with in-

61
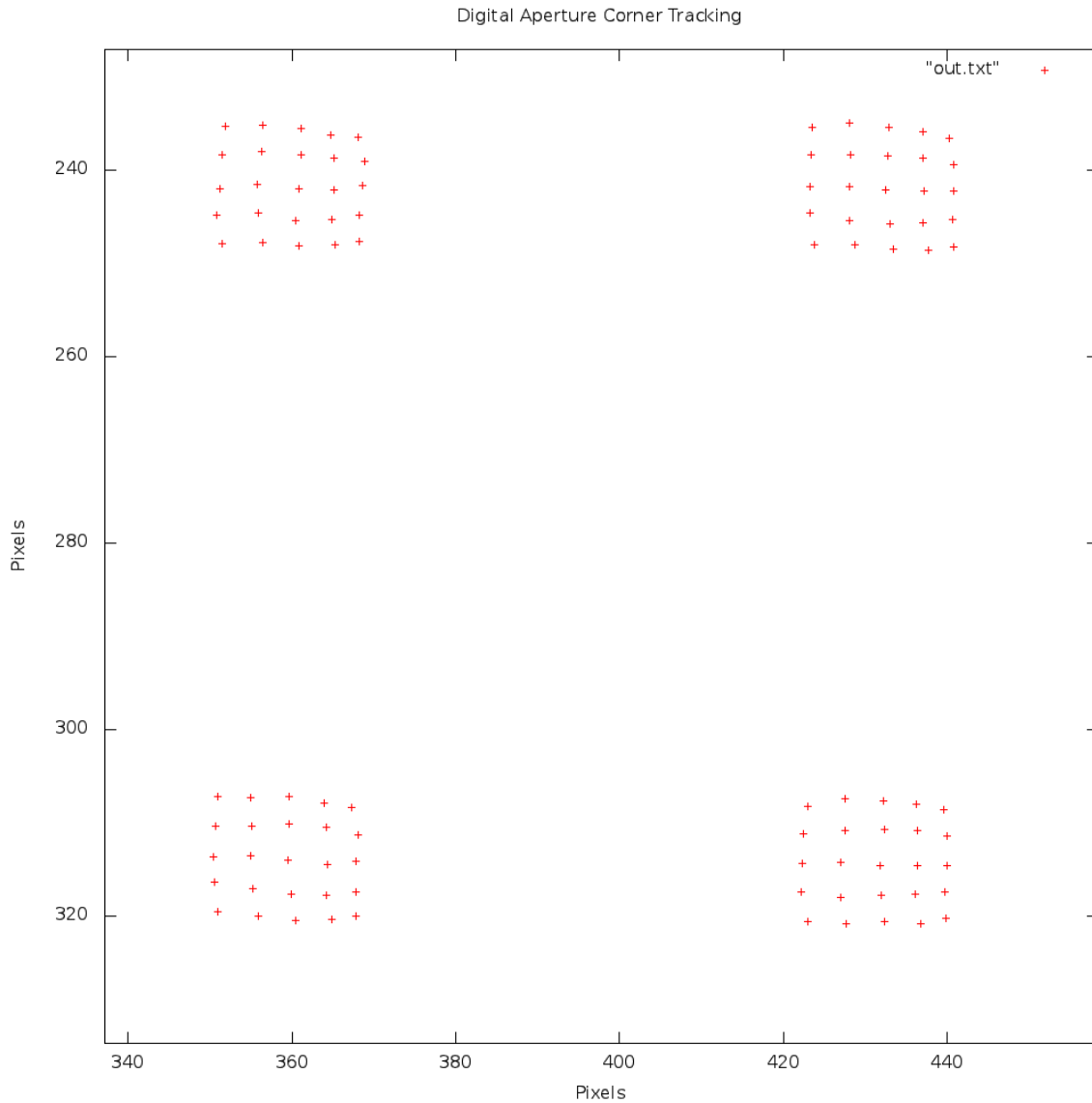
Figure 4-5: Close-up of tracked corner points for four corners near center of frame.

teresting surface geometry was desired. A wooden 3D puzzle with strong woodgrain texture was selected and imaged by the ACPA camera. One of the central images from the light field scan is shown in Figure 4-6. The major grains can be seen, as well as the many right-angle intersections of the orthogonal surfaces.

Applying a leading optical flow algorithm [6] from [5] to a set of light field images like the one in Figure 4-6 yielded a strong flow field, even though the images weren't perfectly sharp. The flow field is shown in Figure 4-7, in which red points indicate the raw locations of the flow, and the green points represent the grid-of-best-fit such

that the spacing and location of the grid minimizes the error between itself and the raw points. This fitting was done because the depth of a point in space is encoded as the spacing between grid points, and grid fitting was a good way to average out the noise and aberations present in the raw data. This is a simple extension into two dimensions of the radius-of-best fit used for the same purposes for the rotating aperture systems in [7]. The mathematical relationship between the grid spacing and the depth of a point in 3D space relative to the camera is described in simplest terms by Equation 5.1, and covered in detail generally in Chapter 5.
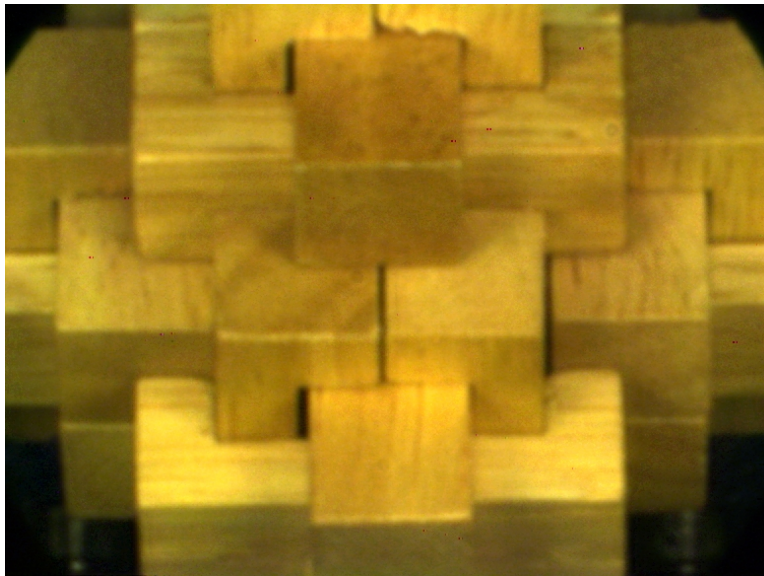
Figure 4-6: Example image from light field capture of object for purposes of 3D scanning

4.4   Conclusions and Future Work

The results in the preceding sections demonstrate the viability of using ACPA cameras as high-resolution 3D scanners. Because of the short-comings in quality resulting from off-the-shelf hardware discussed in Section 3.5, the images used for testing and presented here lacked sufficient sharpness to create strong 3D reconstructions. As a result, the experimental results of that work have been omitted in this thesis. It is left to future research projects to construct ACPA cameras of superior quality,
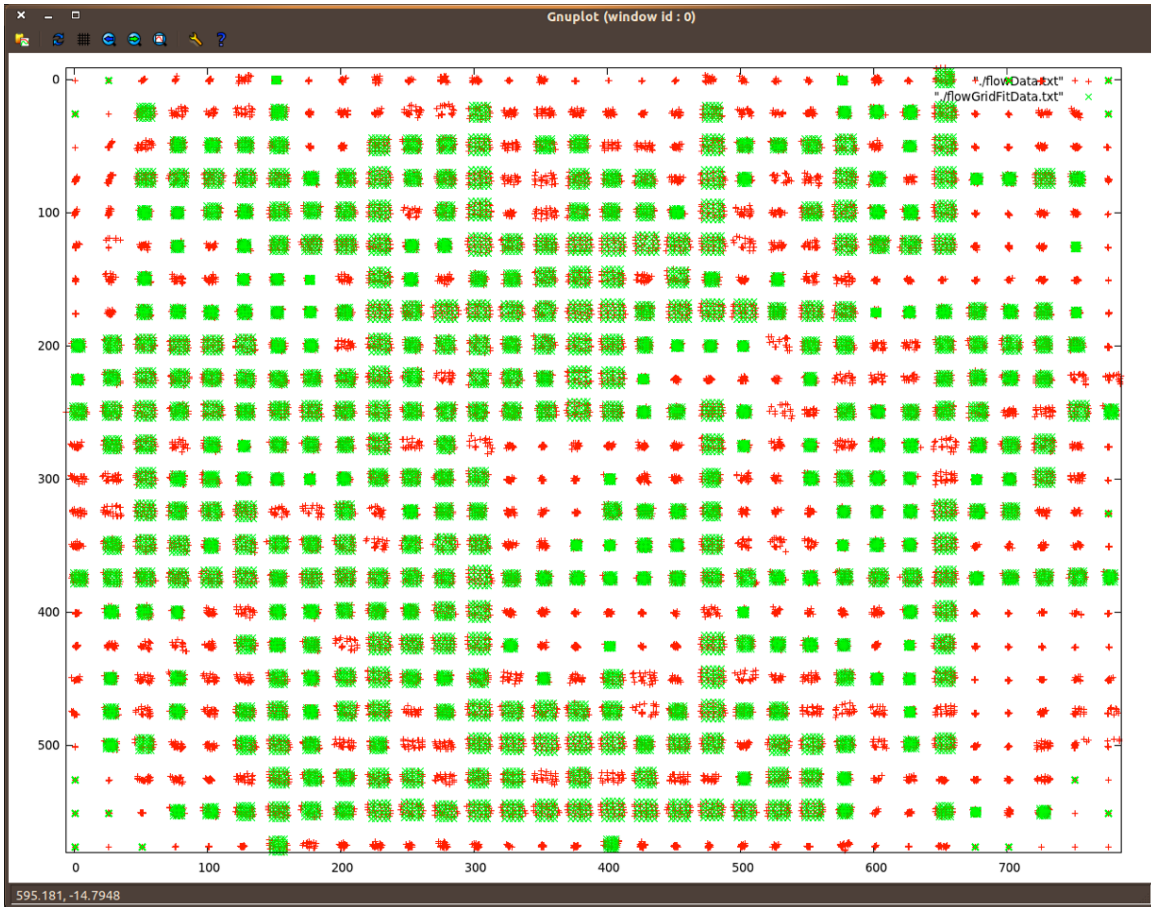
Figure 4-7: Texture point tracking of object in Figure 4-6 using optical flow for purposes of 3D scanning

acquire data with them, and then process that data in the same ways as have been demonstrated in the preceding sections. The theoretical and practical groundwork has been laid; the projects that remain concern new and improved implementations of the principals already demonstrated here.
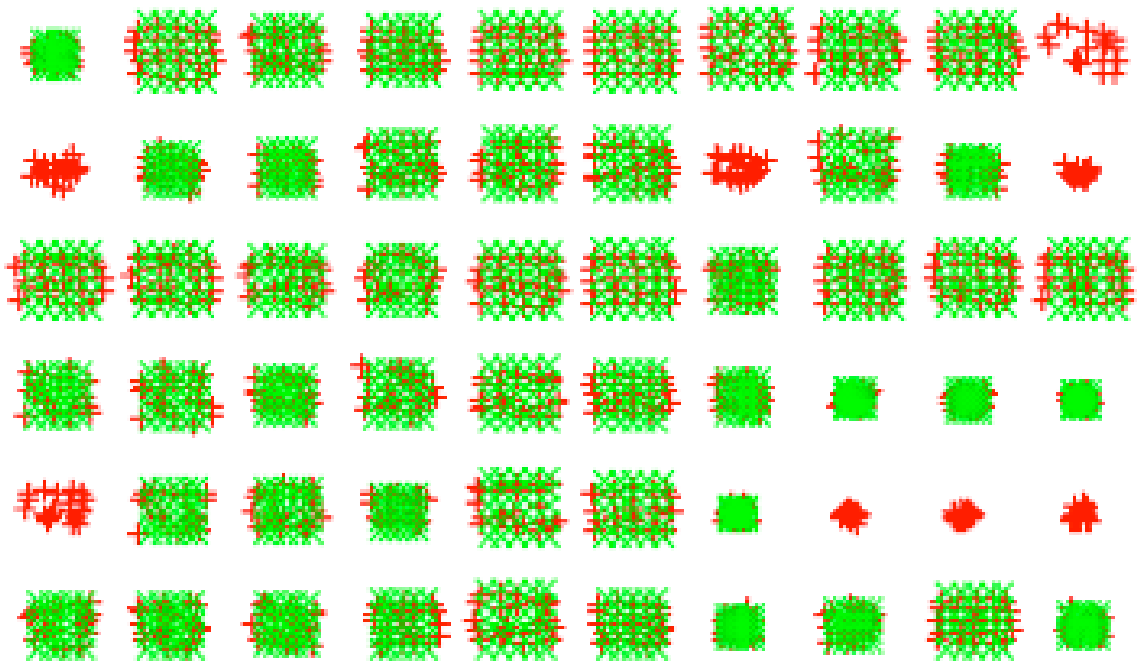
Figure 4-8: Close-up of region of texture point tracking from Figure 4-7

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Design Theory of Programmable-Aperture Light Field Cameras for 3D Scanning Applications

5.1 Introduction

This chapter advances an analytical framework for choosing optimal programmable-aperture light field (PALF) camera parameters for a given 3D scanning application. Programmable-aperture cameras–particularly arbitrarily-controllable programmable aperture (ACPA) cameras–have great potential for use as 3D-scanning tools. They are not without their limitations (see Section 5.3), however, and so a framework for optimizing their performance is of great value. The method put forward in this chapter deals with diffraction-limited systems only, and presents only a theoretical framework for how to think about the tradeoffs between design criteria. Section 5.8 deals with how this work can be verified and improved upon in the future.

5.2 Principles of Light Field Imaging

A programable aperture light field camera acquires depth information by sampling the aperture plane of the lens. Figure 5-1 illustrates this. The principle behind this technique is that the depth of a point in space is directly related to the size of its blur circle on the sensor plane. By subdividing the aperture plane and sampling the image blur at each "sub-aperture," the position of the corresponding target point in 3D space can be computed. Assuming that the camera is focused at infinity, and ignoring depth-of-field and diffraction-limit considerations (these considerations will be addressed later), then the relationship between the depth of a point in space and the disparity between two apertures' images of that point is

$$Z = \frac{Bf}{x_{tb}} \tag{5.1}$$

where B is the light field "baseline," or distance between the centers of the two apertures, f is the focal length of the lens (i.e. the distance between the lens and sensor planes), and x_{tb} is the disparity at the sensor plane between the images of the point from the top and bottom apertures, respectively.

5.3 Light Field Depth Performance–Overcoming Small Baselines

Given the relatively limited baselines available to a single-lensed light field camera, the performance of the two-image paradigm must be augmented by exploiting the arbitrary sampling abilities of the camera. This can be accomplished by taking multiple photos (many more than two), and using the aggregated data to get better depth approximations.

The AWS approach presented in [7] tackles this problem by rotating an aperture at constant radius around the optical axis. By sampling many points in a ring and fitting a circle in the least squares sense, the authors achieve accuracies that exceed
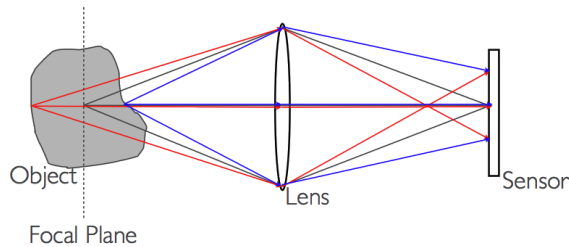
(a) Rays passing through camera without programmable aperture.
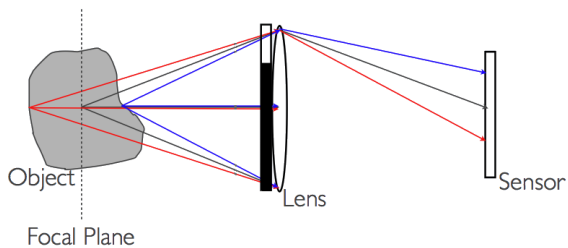
(b) PALF camera sampling the top of the aperture.
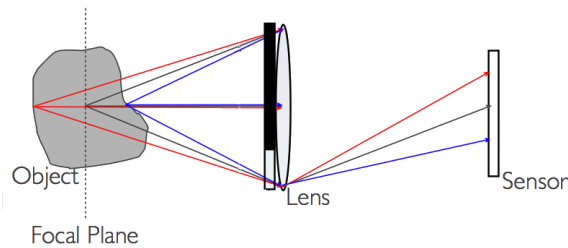
(c) PALF camera sampling the bottom of the aperture.

Figure 5-1: How rays pass through a normal camera, and how a programmable aperture can sample them

what was possible with a single pair of images from opposed apertures. While very effective, this approach has one major drawback: it does all of its imaging through the periphery of the lens, where baseline is maximized but also where aberrations are largest. We can advance this approach significantly by utilizing a PALF camera to sample across the entire aperture plane. To that end, a method is now proposed for determining the theoretically optimal parameters of a PALF camera given a particular 3D scanning application.

5.4 Light Field Camera Design Theory

This section presents the theory necessary for a quantified approach to light field camera design. Before beginning, it is important to call attention to one key premise that lies at the heart of the coming framework: that both the Airy disk size, and the blur circle or "the circle of confusion" at the edge of the depth of field, should be equal to the pixel pitch.

The importance of these equalities is intuitive if not immediately obvious: if the Airy disk were smaller than the pixel pitch (i.e. a large sub-aperture), then the system would have a limited depth of field without any gain in image sharpness. Conversely, if the Airy disk were larger than the pixel pitch, then the image would be blurred according to the diffraction PSF and the effectiveness of good feature extractors [23] and [1] and optical flow algorithms [4], [6], [32] used for 3D reconstruction would be reduced.

Similarly, it is desirable to set the circle of confusion size equal to the pixel pitch so that the image disparity from any two adjacent sub-apertures falls within the range [0, 1] pixels. If the sub-aperture were smaller, then the disparity range would be truncated and the system SNR (disparity to error, in this case) would drop. If the sub-aperture were larger, then points on the rear-most and front-most portions of the scanned object would fall out of focus, and disparity measurements would incur a higher level of uncertainty due to blurring.

For these reasons equality between the pixel pitch, Airy disk diameter and circle of confusion diameter will be enforced for the remainder of this analysis. For similar reasons we'll also constrain the apertures to be tightly packed yet non-overlapping, as the disparity from overlapping pitch-matched sub-apertures would suffer from low SNR, and loosely packed apertures would unnecessarily waste aperture space. With these constraints in hand, derivation of the framework from first principles can begin.

5.4.1 Airy Disk Size

Diffraction of light through the aperture of a perfectly focused lens still results in a finite spot size on the focal plane. This spot is called the Airy disk, and according to the Rayleigh Criterion it has a diameter of

$$d_{Airy} = 2.44\lambda\frac{v}{d} \qquad (5.2)$$

where λ is the wavelength, v is the distance from the lens to the sensor plane, and d is the diameter of the pupil. For distant objects, this can be approximated as

$2.44\lambda N$, where N is the f-number (defined as $N = \frac{f}{d}$), and f is the focal length. For polychormatic imaging, λ may be set to the middle of the system's design wavelengths; for visible light imaging this is green at $\sim565nm$.

5.4.2 F-Number

The f-number, N, of a lens is defined as

$$N = \frac{f}{d} \tag{5.3}$$

where f is the focal length and d is the "clear aperture" diameter of the lens. As previously shown in Equation 5.1, the larger the baseline distance, B, between two light field apertures, the further off in Z we can measure for a given minimum measurable disparity. For a given lens, B is maximized when two small, opposed apertures are located at the very edges of the lens (i.e. at a distance of d apart). Additionally, for a multitude of apertures, a smaller N means that more non-overlapping sub-apertures of a given f-number can be packed into the lens' clear aperture. Thus for a given focal length and imaging scenario, the smaller the lens' f-number, the better the depth resolution that can be achieved.

5.4.3 Depth of Field

A critical design criterion is that the object of interest be entirely in focus, i.e. that the depth of field given by the sub-aperture size, focal length and pixel pitch be adequate to capture the entire depth of our object sharply. The figure below depicts the geometry of depth of field,

where D_N and D_F are the near and far limits of focus when s is the distance to the object-side focal plane, d is the pupil diameter, c is the acceptable circle of confusion diameter (the pixel pitch), and v is the distance from the lens to the sensor. The "depth of field" is then calculated simply as $DoF = D_F - D_N$. The formulae for D_N and D_F are
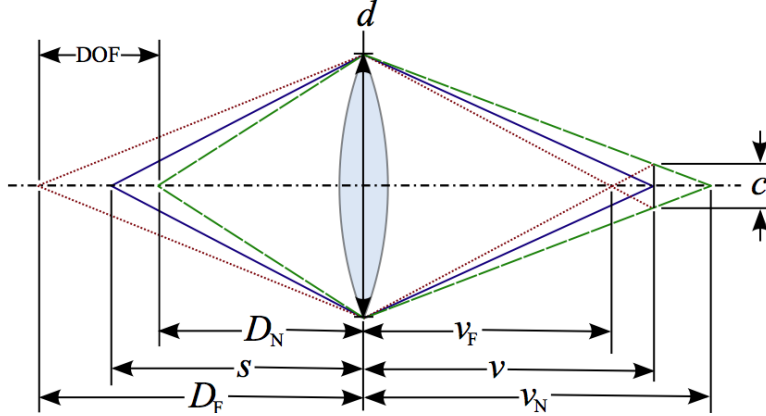
Figure 5-2: Depth of field diagram, CCA 3.0 License, from [37].

$$D_N = \frac{sf^2}{f^2 + Nc(s - f)}, \qquad D_F = \frac{sf^2}{f^2 - Nc(s - f)} \tag{5.4}$$

If c, the acceptable circle of confusion diameter, is set to be the same as the Airy disk diameter for a given f-number, then the focus of the lens at the boundary of the depth of field is matched to its diffraction-limited performance. The depth of field can then be plotted in 3D against f-numbers and object-side focal distances. A common 50mm Nikon lens (f/1.4, 0.45m near-focus) yields the plot in Figure 5-3.

5.4.4 Hyperfocal Distance

The hyperfocal distance of an imaging system is the focal distance at which objects at infinity are just barely in focus (and so increasing the focal distance sacrifices near-field focus without a corresponding gain in the far field). It is calculated as

$$H = \frac{f^2}{Nc} + f \tag{5.5}$$

The plateau in Figure 5-3 indicates that for the given f-number, the corresponding focal distance lies beyond the hyperfocal distance.

There are clearly many points on the plot in Figure 5-3 that would satisfy any arbitrary required depth of field. How should a focal distance/f-number pair be chosen? A similar plot of error vs. focal distance and f-number makes the decision
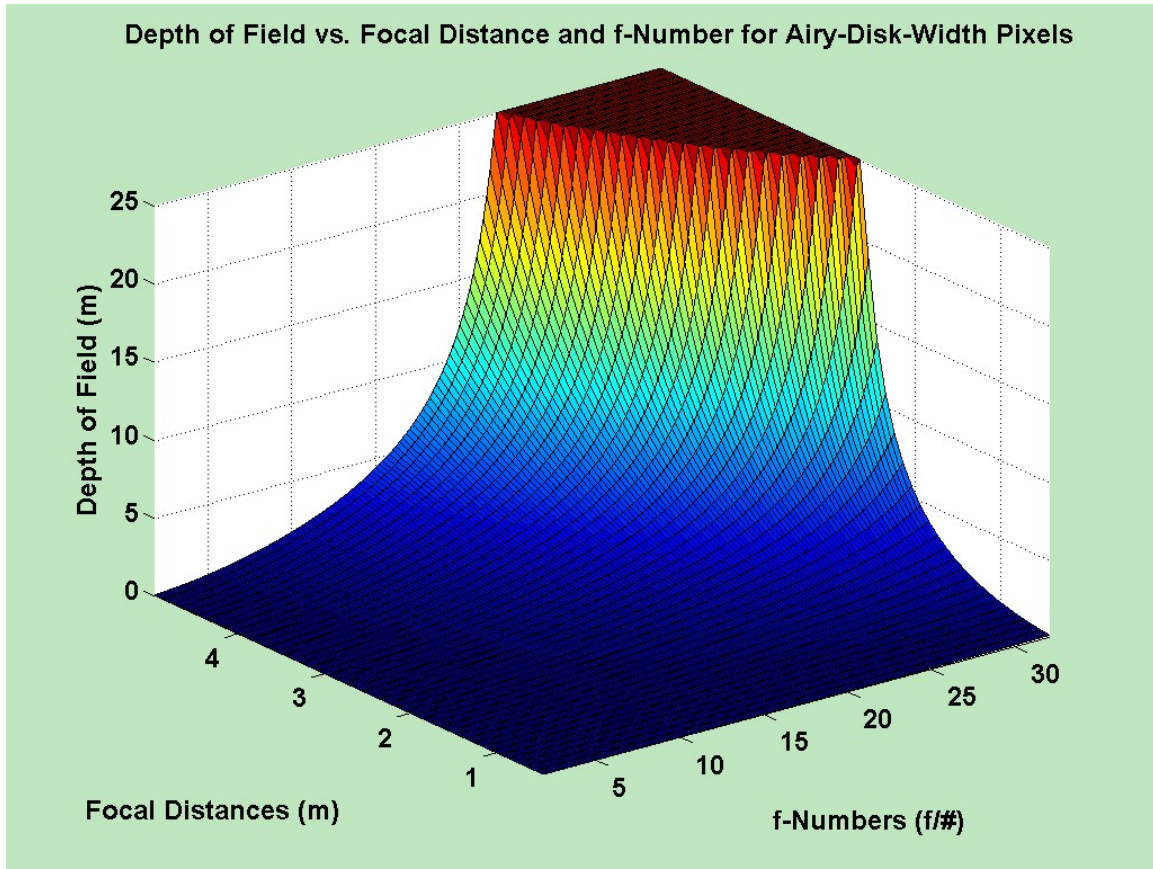
Figure 5-3: Plot of depth of field vs. focal distances and f-numbers for a Nikon 50mm lens.

clearer.

5.4.5 Disparity Error Analysis

As shown in Section 5.2, depth measurement in a PALF camera is done by estimating the disparity between corresponding image points. These estimates contain error. According to the central limit theorem (CLT) this error can be reduced by taking more samples. This "standard error of the mean" (SEM) goes as $SEM = \frac{\sigma}{\sqrt{n}}$, where σ represents the standard deviation of error for a single sample. However, this is not the whole story.

As the aperture size is made smaller and smaller to pack in more samples, the spacial-frequency bandwidth of the system is correspondingly reduced due to the diffraction limit. Quartering the area of an aperture quadruples the number of aper-

tures that could fit in the aperture plane, but it also doubles the diameter of the Airy disk, according to Equation 5.2. The plot in Figure 5-4 shows the surprising theoretical result that depth error is invariant on f-number for packed, non-overlapping sub-apertures.
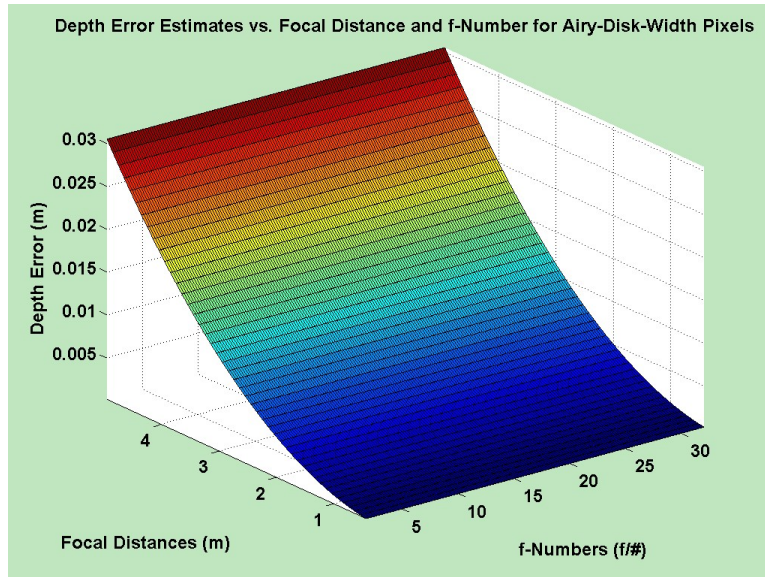
Figure 5-4: Plot of depth error estimation vs. focal distances and f-numbers for packed, non-overlapping sub-apertures in a Nikon 50mm f/1.4 lens.

Figure 5-4 also shows that the depth estimate error increases as the focal distance increases, because of the decreasing parallax resulting from the fixed "baselines" of the aperture positions. For the depth-of-field decision, this means that, all other things being equal, it is best to minimize focal distance. For the Nikon lens used in Figures 5-3 and 5-4, this value is $450mm$. With that information in hand, an appropriate sub-aperture f-number must be chosen.

5.4.6   Resolution, Sensor Size and Field of View

Based on the results from Section 5.4.5, an arbitrarily small aperture size could be chosen to accommodate the lowest focal distance offered by the lens. But there are two limitations: the first is that taking a large number of small-aperture (long-exposure) images would be a lengthy process. The second and more concrete limitation is that such an aperture size would yield correspondingly low-resolution image data because

of the diffraction limit. Therefore, by prescribing a minimum resolution (1280 pixels), a maximum f-number at a given field of view (FOV) is also prescribed.

The necessary sensor height for a given FOV can be calculated from the thin lens equation and similar triangles:

$$s_i = \left(\frac{1}{f} - \frac{1}{s_o}\right)^{-1} \quad \text{and} \quad \frac{s_i}{s_o} = \frac{h_i}{h_o} \quad \text{yield the sensor height:} \quad h_{sensor} = \frac{\left(\frac{1}{f} - \frac{1}{s_o}\right)^{-1}}{s_o} h_o$$

$$(5.6)$$

This value and the prescribed image resolution (r_{image}) can be used to calculate the pixel pitch:

$$p_{pixel} = \frac{h_{sensor}}{r_{image}} \tag{5.7}$$

Finally, by rearranging the Rayleigh Criterion (Equation 5.2) and matching the pixel pitch to the diameter of the Airy disk ($p_{pitch} = d_{Airy}$) the corresponding sub-aperture f-number can be calculated:

$$N = \frac{f}{d}, \quad d = 2.44\lambda\frac{s_i}{d_{Airy}} \quad \Rightarrow \quad N_{Airy} = \frac{f \times p_{pitch}}{2.44\lambda s_i} \tag{5.8}$$

5.5   System of Equations for Design

The analyses in Section 5.4 produce a four-equation, four-unknown system of nonlinear equations. Solving the system of non-linear equations yields the optimal camera design specifications for a given set of design parameters. The equations are:

$$N = \frac{f \times p_{pitch}}{2.44\lambda}\left(\frac{1}{f} - \frac{1}{f_d}\right) \tag{5.9}$$

$$p_{pixel} = \frac{h_{sensor}}{r_{image}} \tag{5.10}$$

75

$$h_{sensor} = \frac{\left(\frac{1}{f} - \frac{1}{f_d}\right)^{-1}}{f_d} \times h_{object} = \frac{h_{object}}{\frac{f_d}{f} - 1} \tag{5.11}$$

$$DoF = D_F - D_N = \frac{f_d f^2}{f^2 - N \times p_{pixel} \times (f_d - f)} - \frac{f_d f^2}{f^2 + N \times p_{pixel} \times (f_d - f)} \tag{5.12}$$

The variables in this system can be found in Table 5.1.

Variables in system of equations		
Sub-aperture f-Number (N)	Pixel pitch / Airy Disk ⌀ (p_{pitch})	Design wavelength (λ)
Focal length (f)	Focal distance (f_d)	Sensor height (h_{sensor})
Image resolution (r_{image})	Object height (h_{object})	Depth of field (DoF)

Table 5.1: Variables found in the system of non-linear equations

The system contains nine variables and four independent equations. This means that values for the five most important (i.e. independent) variables should be pre-scribed by the user, and the system will solve for the optimal values of the remaining four, dependent, variables. For example, the design wavelength, λ, will likely be pre-scribed by the application (visible light, IR, UV, etc.). Sub-aperture f-Number, on the other hand, may commonly be sought as a dependent variable.

5.6 Application Example

Now that the framework is laid out, it can be demonstrated by applying it to a common real-world application: face scanning. To start, the common Nikon 50mm, f/1.4 lens that was used in Figures 5-3 and 5-4, which has a minimum focal distance of 450mm, is selected as the imaging lens. A typical human face is approximately 300mm high, with about 100mm of depth (tip-of-nose to middle-of-ears). Visible-spectrum imaging is desired. The dependent and independent variables, along with the independent variable values for this application, are listed in Table 5.2.

The system of equations converges to a solution with these values, and these results are listed in Table 5.3.

Dependent and Independent Variables		
Dependent Variables (unknowns)	Independent Variables	
Sub-aperture f-Number (N)	Focal length (f):	50mm
Pixel pitch / Airy Disk $\varnothing$ (p_{pitch})	Focal distance (f_d):	450mm
Sensor height (h_{sensor})	Object height (h_{object}):	300mm
Image resolution (r_{image})	Depth of field (DoF):	100mm
	Design wavelength (λ):	565nm

Table 5.2: Dependent and independent variables for the system of non-linear equations in Equations 5.9-5.12

Optimal Primary Design Parameters	
Sub-aperture f-Number (N)	f/21.0
Pixel pitch / Airy Disk $\varnothing$ (p_{pitch})	32.6μm
Sensor height (h_{sensor})	37.5mm
Image resolution (r_{image})	$1150 pixels$

Table 5.3: Optimal camera specifications from the solution to the system of non-linear equations

If, for example, it was felt that specifying the depth of field was less important than specifying an output image resolution, then the dependent and independent variables could be rearranged as shown in Table 5.4. This system also converges to a solution (similar to the first but not the same), and those values are listed in Table 5.5.

Dependent and Independent Variables		
Dependent Variables (unknowns)	Independent Variables	
Sub-aperture f-Number (N)	Focal length (f):	50mm
Pixel pitch / Airy Disk $\varnothing$ (p_{pitch})	Focal distance (f_d):	450mm
Sensor height (h_{sensor})	Object height (h_{object}):	300mm
Depth of field (DoF)	Image resolution (r_{image}):	1280pixels
	Design wavelength (λ):	565nm

Table 5.4: Alternative dependent and independent variables

5.7 Conclusion

This chapter has demonstrated from first principles that the optimal design of a diffraction-limited light field camera for use in a specific 3D scanning application can

Optimal Primary Design Parameters	
Sub-aperture f-Number (N)	f/18.9
Pixel pitch / Airy Disk $\varnothing$ (p_{pitch})	29.3μm
Sensor height (h_{sensor})	37.5mm
Depth of field	$80.3mm$

Table 5.5: Optimal camera specifications for alternative independent variables

be reduced to a system of non-linear equations and solved directly. The theory has been applied to a common scanning application using a popular commercial lens, and the returned design values for sub-aperture f-Number, pixel pitch, sensor size and field of view are well within the range of today's common imaging hardware.

5.8 Future Work

This chapter has laid out a framework for determining optimal PALF camera design parameters given a specific scanning application. Future work would provide a generalized framework for a larger class of light field design problems. This larger framework would handle a wider variety of constraints and design parameters, and allow for constrains ranges (e.g. design for scanning any object within a range of sizes). Additionally, experiments should be run to verify the accuracy and optimality of the current technique with real-world hardware. Finally, an extension to the framework that accounts for non-negligible lens aberrations should be developed.

Chapter 6

Superresolution Imaging with Digital-Aperture Programmable-Aperture Light Field Cameras

6.1   Introduction to Superresolution Imaging

Superresolution (SR) imaging is the process of combining multiple low-resolution (LR) images of the same object from slightly different perspectives into one single, higher resolution (HR) image [46]. SR techniques can be used when the resolution of an imaging source is much lower than the imaging capacity of the optical train (e.g. lower resolution than the diffraction limit and lens aberrations). In this case, the sensor is under-sampling the information coming through the lens, and some of the 'lost' information can be recovered by taking many images of the same object with sub-pixel shifts.

This short chapter discusses the theory of how "digital apertures" (i.e. arbitrarily-controllable programmable-aperture (ACPA)) light field cameras can be used for su-perresolution imaging. No demonstrable results will be presented; the general prin-

cipals are compelling on their own and can be elucidated schematically with relative ease. The demonstration of these techniques with experimental data is reserved for future work.

6.2   Principles of Geometric Sub-Pixel Superresolution

Superresolution is an umbrella term for a myriad techniques related to increasing the resolution of an image source. For the purposes of this chapter, "superresolution" refers to the process of generating a high resolution digital image from a stack of low resolution images of the same object. The requirements for these LR images are that they be shifted by sub-pixel amounts from each other, and that they represent an undersampling of the scene (e.g. their resolution is not limited by the diffraction limit).

There exist many good introductions and technical reviews of the superresolution field [28], [34], [46], so an overview is omitted here. For the purposes of this brief chapter, it is necessary only to know that SR techniques perform best with as many images with different sub-pixels shifts as possible. Furthermore, they perform best when the movement causing the subpixel shifting occurs normal to the optical axis (such that the shifting is uniform across the sensor). With this information in hand, it can be demonstrated schematically that ACPA cameras are well-suited to the task of creating superresolution-friendly imaging data.

6.3   Creating Sub-Pixel Shifts with an ACPA Camera

Figure 6-1 demonstrates the functioning of a programmable aperture light field camera when imaging a target that is well-focused. The ray bundle in (A) focuses to a point on pixel P_i. Because of this, selecting only a portion of the lens, as in diagrams (B)

and (C), has no effect on the distribution of light falling on the sensor. Regardless of which parts of the lens are used for imaging, all of the light falls on pixel P_i. This invariance on aperture configuration means that no change can be made to the captured images (aside from trivial changes in brightness) and so the light field data won't contain the shifting needed for SR imaging techniques.
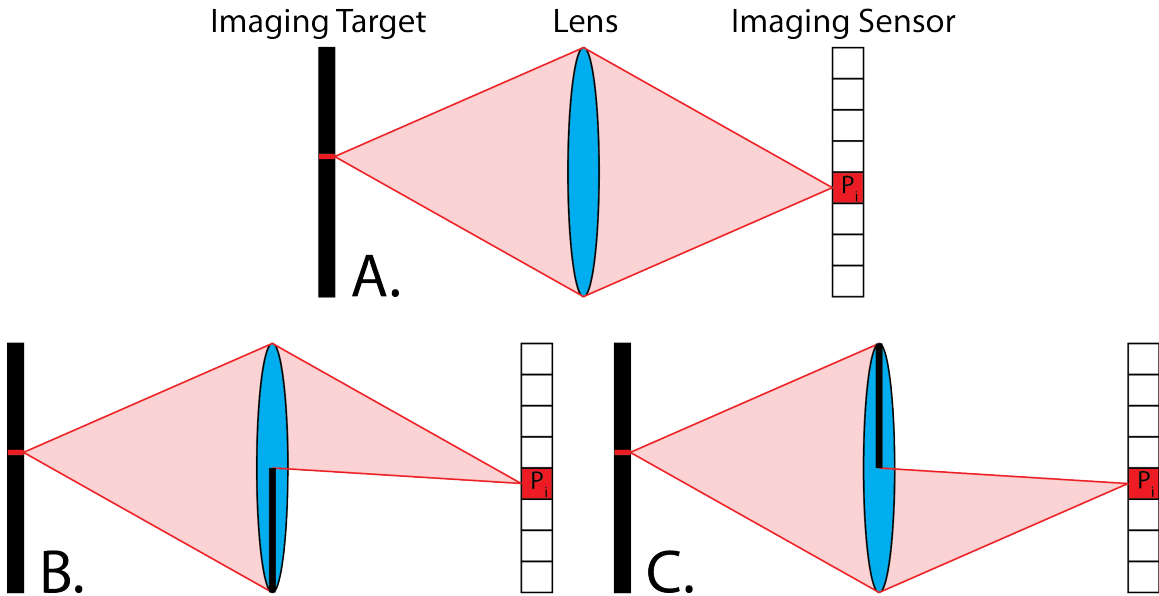

Figure 6-1: Lens sampling in a well-focused programmable-aperture camera.

If, however, the camera is made to be slightly out of focus (defocus on the order of a few pixels), then the image is no longer invariant upon aperture configuration. Figure 6-2 illustrates how subpixel shifts can be induced in this case. In diagram (A) the lens is fully open and light is falling on two adjacent pixels, P_i and P_{i+1}, as the result of slight defocus. As the lens is selectively sampled by a programmable aperture, the light coming from the imaged point can be made to fall completely onto pixel P_{i+1} (B), be distributed between P_i and P_{i+1} (C), or fall completely onto P_i (D). In this way, an ACPA camera can create arbitrarily many subpixel-shifted images. Furthermore, because the shifting is the result of blocking light at the aperture plane (which is mechanically parallel to the sensor), the shifting occurs exactly perpendicular to the optical axis, which is ideal for SR algorithms.
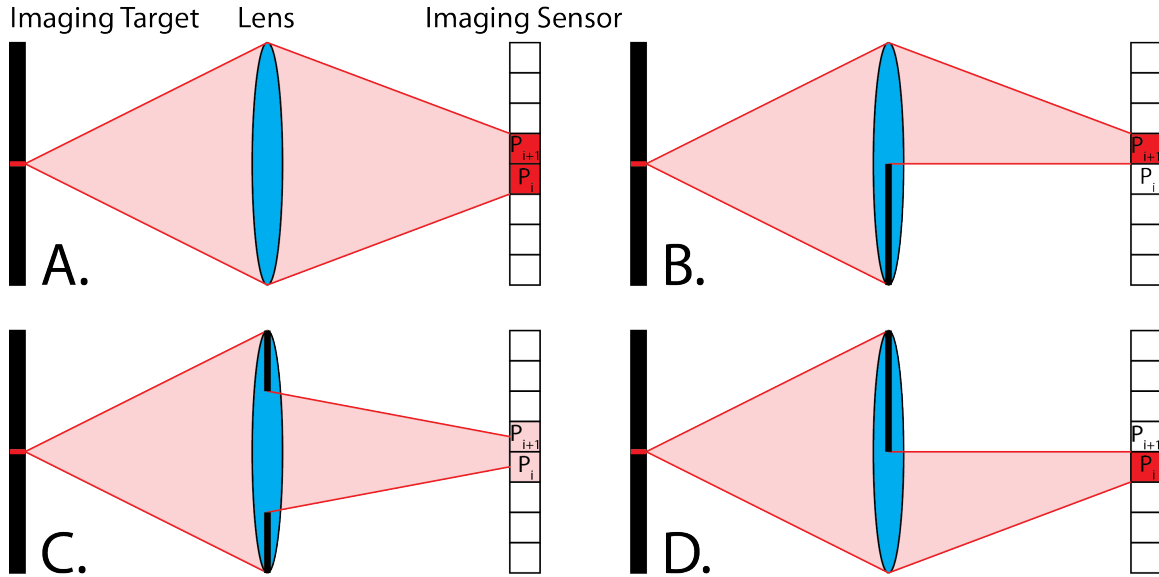
Figure 6-2: Lens sampling in a slightly out-of-focus programmable-aperture camera.

6.4 Conclusions and Future Work

Though this is a brief chapter, its contents illustrate that ACPA cameras are well-suited to superresolution imaging tasks. One possible application of these principles is in low-cost document scanners that use low-resolution sensors and low-resolution digital apertures, yet achieve very high resolution scans. Another possible application is for surveillance camera equipment. SR is only applicable to surveillance video when the video has movement, like a car driving away whose license plate needs to be superresolved. In these cases, the results are at the mercy of just how many frames of video caught the action (and just how much motion blur corrupted those precious few images). With ACPA light field surveillance, on the other hand, a large stack of LR images could be compiled while the car of interest was parked motionless in its parking spot hours before it was driven away.

It is left to future work to build a prototype ACPA camera for testing this technique. Of particular interest is just how beneficial the guaranteed-parallel aperture-based shifting is compared to other more stochastic sources of shifting. Also of great interest is what the incremental improvement in high-resolution recovery is as the image stack gets larger. Answering questions like what subpixel shifting intervals are

optimal would be of considerable value as well.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 7

"Programmable Deconvolution" with Digital-Aperture Programmable-Aperture Light Field Cameras

7.1 Introduction to Deconvolution

Deconvolution is the mathematical process of deblurring an image and is well studied in the literature [16]. The practical need for deconvolution arrises because physical imaging systems are imperfect: an infinitesimal point-source of light imaged by a real system will focus the point to a spot with finite (or, as in the case of diffraction, quasi-infinite) extent. The pattern of this spot is referred to as the "point spread function" (PSF), and can result from poor focus, various lens aberrations, and/or diffraction. Figure 7-1, taken from [40], shows some example PSFs that result from spherical aberration and defocus.

Additionally, in astronomical observations, small index changes in the atmosphere between earth's surface and space cause rapidly-changing observational distortions. This effect is called "atmospheric seeing" and frequently is the dominant distortion
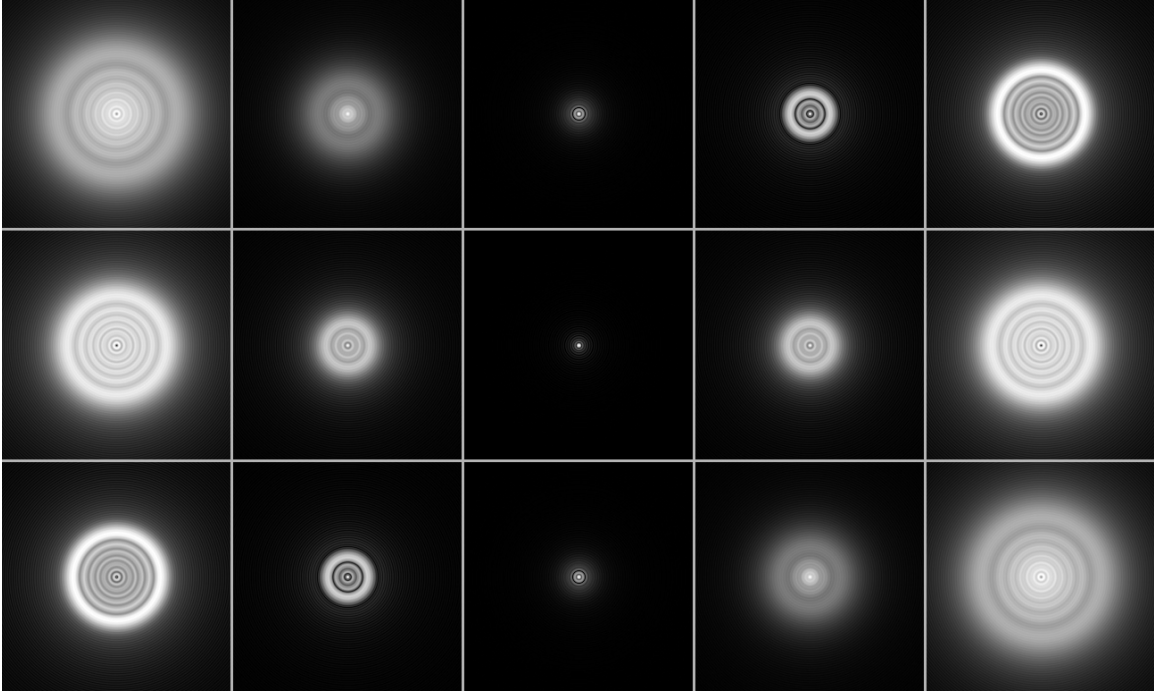
Figure 7-1: A point source as imaged by a system with negative (top), zero (center), and positive (bottom) spherical aberration. Images to the left are defocused toward the inside, images on the right toward the outside.

in land-based high-end telescopes.

In its simplest form, the mathematical expression for a convoluted image is

$$i(x, y) = o(x, y) \otimes h(x, y) \tag{7.1}$$

where $i(x, y)$ is the image observed at the sensor, $o(x, y)$ is the perfect image of the object, and $h(x, y)$ is the function for the PSF. The goal of deconvolution is to recover $o(x, y)$. Additionally, for real sensors which exhibit pixel noise, a noise term can be added (variables dropped for simplicity):

$$i = o \otimes h + n \tag{7.2}$$

We can take the Fourier transform of the values, yielding this commonly-known equation in the frequency domain:

$$I = HO + N \tag{7.3}$$

In this equation, H is the Fourier transform of the PSF and is called the "optical transfer function" (OTF). This equation suggests a straightforward inversion solution to the deconvolution problem:

$$O = (I - N)H_{inv} \qquad (7.4)$$

where, in the simplest case, H_{inv} is the element-wise inversion of H:

$$H_{inv} = 1/H \qquad (7.5)$$

Unfortunately, this simplistic approach has two main drawbacks: first, it assumes that we know much about N, which in practice is quite challenging. Secondly, and at least as importantly, most common PSFs have OTFs which contain one or more regions with zero or near-zero values. As a result, taking the straight inverse of the OTF isn't effective because these near-zero values dominate all other values when inverted. To make matters worse, it is precisely at these zeros where the signal is weakest: the straight inverse over-emphasizes the very frequencies with the lowest signal-to-noise ratios (SNR). Much work has been done on filter design to mitigate these issues [13] [48], but single-image deconvolution in the face of OTF zeros remains a challenge today.

7.2 Multi-Image Deconvolution With Zero-Canceling

To address the issue of near-zero information for certain OTF frequencies, the authors of [47] developed a technique for deconvolving multiple images of the same object simultaneously. Taking a stack of images:

$$\forall j \in [1, 2, ..., n] : i_j(x, y) = o(x, y) \otimes h_j(x, y) \qquad (7.6)$$

such that:

$$\forall j, k \in [1, 2, ..., n], j \neq k : h_j \neq h_k \implies \{zeros(H_j)\} \neq \{zeros(H_k)\} \qquad (7.7)$$

the authors could guarantee that a good signal for each frequency was present in at least one of the images in the stack. Note that the observed object, $o(x, y)$, is constant across the stack. By weighting the reconstruction of each deconvolved frequency toward the image(s) in the stack whose OTF(s) had the strongest signal at that frequency, the authors attained deconvolutions of superior quality across the frequency spectrum. This was a seminal work, and many authors have built upon and extended it [30] [12] [11] [2] [10] & [26].

7.3   Traditional Methods for Acquiring Varying PSFs

A common source of images with varying $h_j(x, y)$s is land-based rapid, sequential captures of an astronomical body, which suffer from atmospheric seeing. The seeing creates varying PSFs whose differences vary based on the turbulence of the atmosphere along the optical path between the body and the telescope. The presence and degree of non-overlapping OTF zeros therefore also depends on this turbulence. Frequently, to ensure good OTF coverage across the frequency spectrum, hundreds or thousands of images of a body will be taken. This Monte Carlo-style sampling can take a long time and often requires complicated telescope tracking to follow the body through the sky as the earth rotates on its axis. The long capture window also means that this technique can not be used to observe short-timescale events like solar flares.

The authors of [25] developed a deconvolution framework called MISTRAL that implicitly took advantage of frequency information across these multiple-image astronomical data sets, while also incorporating a regularization scheme to balance edge preservation against smoothness for frequencies that were not known confidently. Though computation took a long time, this technique returned superior results for these datasets. To help alleviate the long computation times, the authors of [14] im-

proved MISTRAL with an automatic parameter tuning routine, and released it under the name AIDA as an open source project.

Most recently, an author of AIDA and the author of this thesis have collaborated to extend AIDA to light field imaging data, and the results that follow in this chapter were computed using this improved program. It is these two authors' intention that this new version of AIDA, which supports light field data, will eventually be released as an open-source project.

7.4 Modulating the PSF with a Programmable Aperture

Relying on atmospheric seeing or other pseudo-random processes to generate sufficient variability in PSFs is not ideal. In addition to the time required for capture, these methods can be intractable by requiring additional layers of complexity (like star tracking) to support long capture windows, and are unable to capture instantaneous events and pseudo-stationary objects whose appearance changes materially over the course of the window.

In order to remove the limitations of long-duration captures, a new technology is required that allows PSFs to be manipulated in highly deterministic ways so as to achieve complete zero cancelation with only a small number of captures. Light field imaging with an arbitrarily-controllable programmable-aperture (ACPA) camera offers this ability.

In an optical system that is limited by various aberrations and defocus, the PSF shape is directly related to the aperture shape. By blocking parts of the lens pupil one can selectively block that region's contribution to the PSF. To demonstrate this, an ACPA camera prototype was implemented mechanically (to avoid the LCD limitations discussed in Section 3.5) from a standard iris and a vertical slit. This compound aperture was placed in front of an achromatic doublet and coupled with an AVT image sensor, and PSFs were taken across a range of slit widths. The parts list for the

prototype is found in Table 7.1, and an image of the setup is shown in Figure 7-2. Figure 7-3 demonstrates the shape of this composite aperture as the slit is varied in width. Note that the aperture's shape is only varied along the x-axis as the vertical slit opens and closes.

Table 7.1: Parts list for mechanically-implemented ACPA camera with single-axis control.

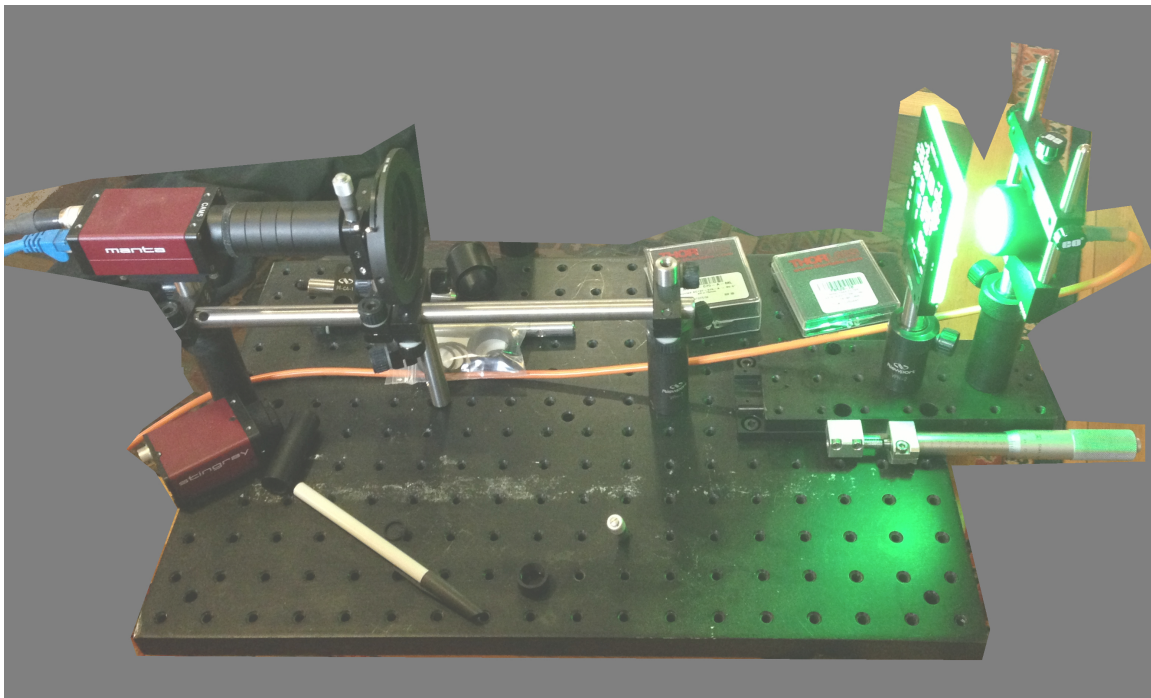|	
Imaging sensor:	AVT Manta G-125B
Lens:	Thorlabs f=75mm Achromatic Doublet, #AC127-075-A-ML
Iris:	Traditional multi-blade @ 6mm open diameter
Vertical slit:	Thorlabs 0mm-6mm adjustable mechanical slit, #VA100
Light source:	Thorlabs green LED light source, #LIU002
Imaging target:	Thorlabs USAF 1951 Test Target, #R3L3S1N

Figure 7-2: Mechanical implementation of one-dimensional ACPA camera.

To test just how much of an effect this changing aperture shape has on the PSF of a real optical system that exhibits aberrations and sever defocus, the PSF at various slit widths was measured. As Figure 7-4 shows, the PSF is modulated quite dramatically over the course of the slit movement. To determine whether these visually clear variations translate into an effective image stack, the OTF was taken for each PSF,
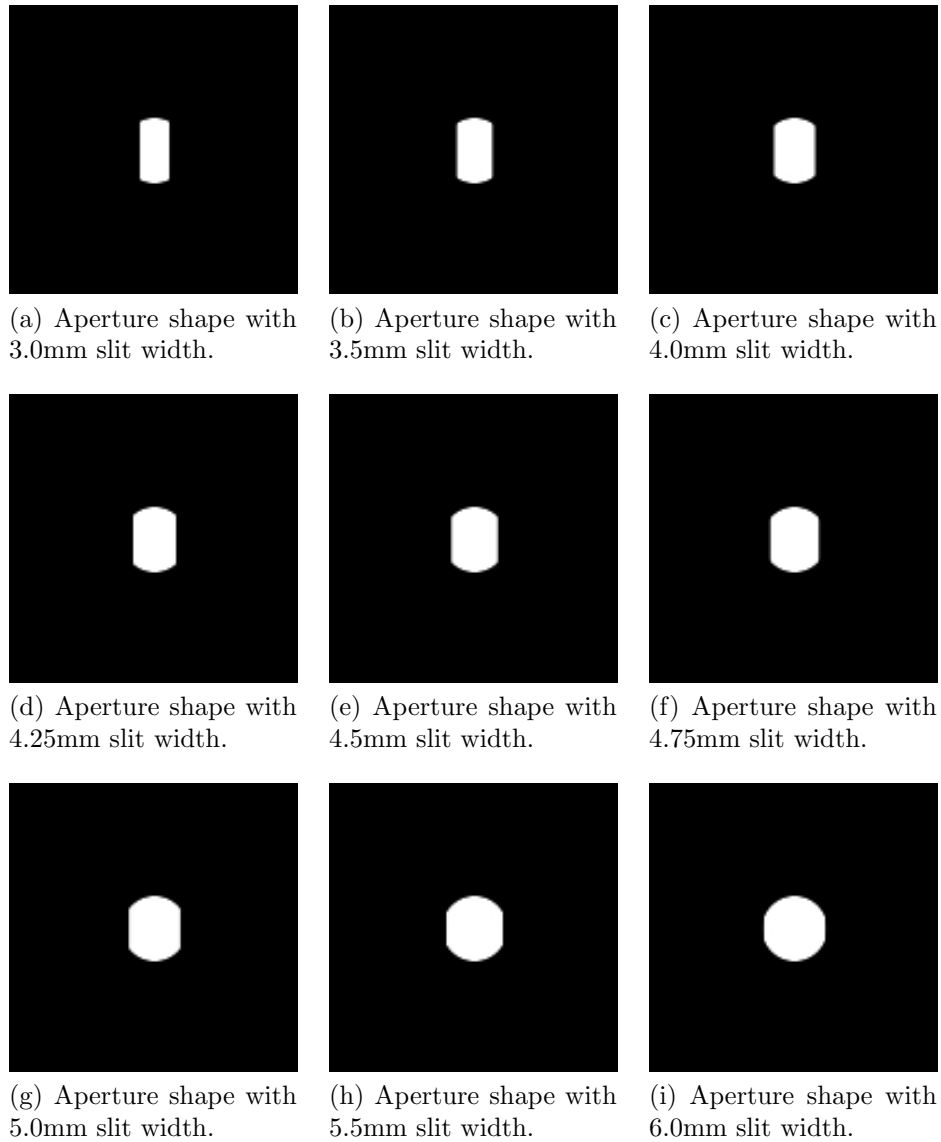
(a) Aperture shape with 3.0mm slit width.

(b) Aperture shape with 3.5mm slit width.

(c) Aperture shape with 4.0mm slit width.

(d) Aperture shape with 4.25mm slit width.

(e) Aperture shape with 4.5mm slit width.

(f) Aperture shape with 4.75mm slit width.

(g) Aperture shape with 5.0mm slit width.

(h) Aperture shape with 5.5mm slit width.

(i) Aperture shape with 6.0mm slit width.

Figure 7-3: Example aperture shapes created by a 6mm circular iris superimposed with a vertical metal slit varying in width from 3.0mm to 6.0mm.

and a plot of the central horizontal profile of each OTF was made. That plot is shown in the top plot of Figure 7-5, and exhibits excellent zero-cancellation. Perhaps most saliently, the first zero for the 4.0mm slit occurs at almost exactly the same spatial frequency (33 cycles/image) as the first local maximum of the 5.0mm slit width, demonstrating nearly-perfect cancelation. Practically speaking, this means that it is possible to get *sharper* deconvolutions by incorporating *blurrier* images into the dataset. Furthermore, the plot demonstrates that the frequency responses from the various PSFs can be tuned simply and directly by utilizing the aperture modulation available from an ACPA camera. Finally, the bottom plot in Figure 7-5 plots the

vertical profiles of the OTFs, demonstrating that an ACPA system can modulate the frequency response of one axis with minimal disruption to the other.
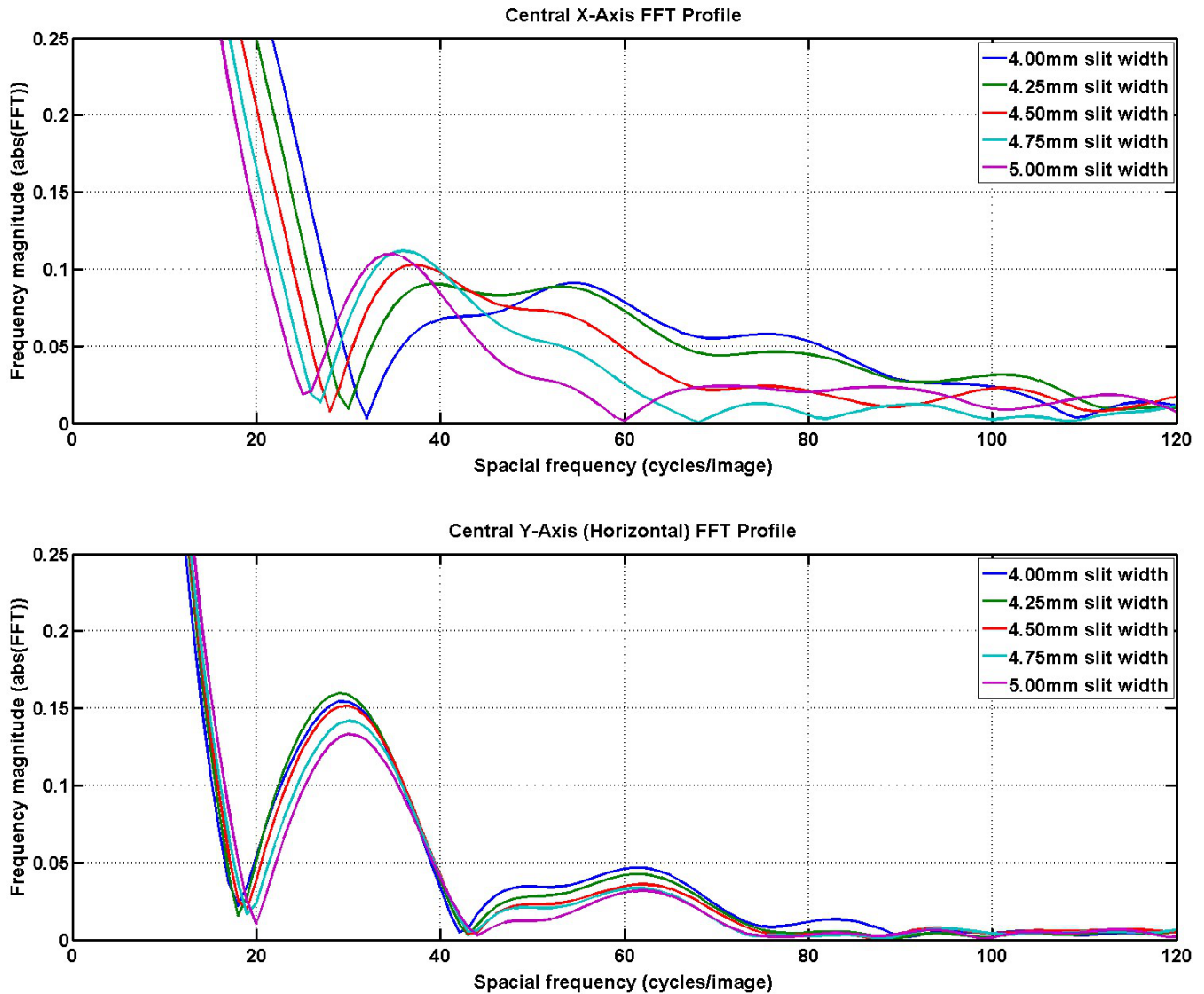
Figure 7-5: Central horizontal & vertical OTF profiles for all observed PSFs in the image stack with slit widths ranging from 4.0mm to 5.0mm in 0.25mm steps

7.5   Demonstration of Principle with Real Data

Having demonstrated that the modulation of an ACPA aperture can yield favorable zero-cancelation, a dataset of test images was taken for deconvolution testing. The

well-known USAF 1951 test target [38] (shown in Figure 7-6) was used, with a focus on the #2 & #3 groups. This target contains strong signals at the spatial frequencies that correspond to the spacings within the individual elements. Additionally, because the target is rich in sharp-edge transitions from white-to-black, all possible spatial frequencies are represented in the target to at least some degree (i.e. because ideal step functions contain all spatial frequencies).
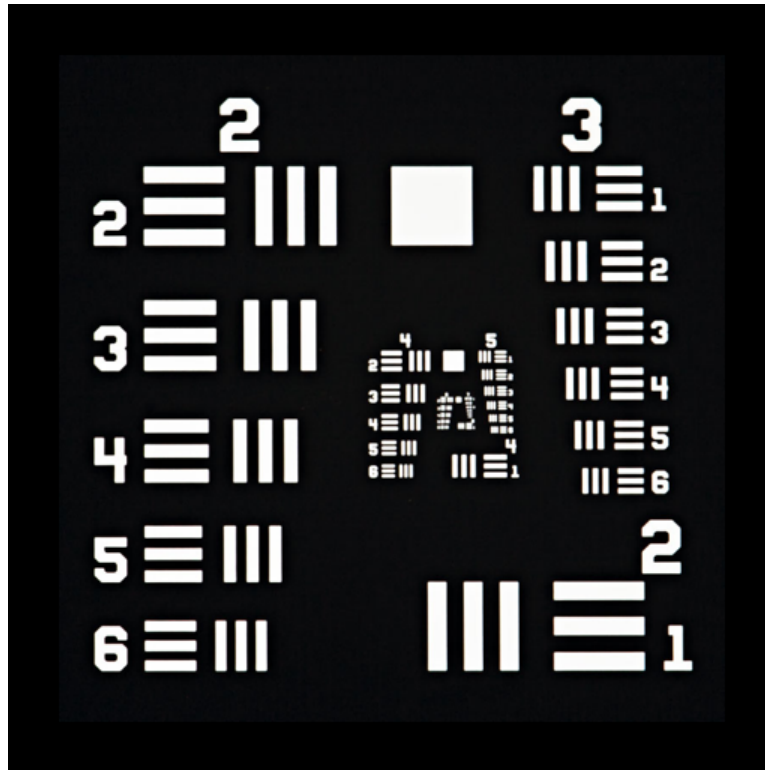

Figure 7-6: USAF 1951 test target used for deconvolution testing.

The target was imaged at each of the aperture positions in the last section: 4.0mm to 5.0mm in 0.25mm steps. The resulting images are shown in Figure 7-7 alongside the reference target image. Notice that the blurring is so severe that many of the 3-line elements appear to contain 2 or 4 lines.

To set a baseline for the quality of deconvolutions achievable with this data, each image was deconvolved with its corresponding observed PSF using AIDA's [14] single-image myopic deconvolution routine. The results of these deconvolutions are shown in Figure 7-8. More detail about these deconvolutions can be found in Appendix B

in Figures B-2 and B-3, which show the horizontal and vertical (respectively) central FFT profiles of the ground truth image, observed images & PSFs, and myopically deconvolved images (those shown in Figure 7-8) & corrected PSFs for each aperture position.

AIDA was run with a value of $lambda\_object\_scaling$ set to 0.01; this setting was chosen for its visually pleasing balance between edge preservation and relatively low high-spatial-frequency noise. See Figure B-1 in Appendix B for an example of how changes in the value of $lambda\_object\_scaling$ effect the deconvolution result.

Despite these settings creating a "good balance," there is still a noticeable amount of high-frequency noise present in the deconvolved image, as Figures B-2 and B-3 show. This is due in large part to the weak or missing frequency components in the OTF. If each image in the stack contains different combinations of frequency components with good zero-cancelation, then it should be possible to mitigate this high frequency noise by deconvolving all the images together into a single result.

To test whether the additional frequency data available in the image stack was capable of suppressing the spurious high frequencies found in the single-image deconvolutions while maintaining the sharp edges that exist in the test target, a single, simultaneous deconvolution of all five images (using each of their corresponding observed PSFs) was done. AIDA again was used, this time in 'NPSFS' mode which is meant for this kind of deconvolution. All other AIDA settings were maintained from the myopic runs, including the setting of $lambda\_object\_scaling = 0.01$. The resulting image is shown in Figure 7-9, alongside the sharpest single-image deconvolution. A frequency-domain comparison of the NPSFS result and the sharpest single-image decon are available in Figure B-6.

It can be seen in the side-by-side comparison that the high-frequency artifacts found in the single-image deconvolutions have been noticeably suppressed in the multi-image result. The reduction is particularly noticeable for spurious high frequencies along the x-axis. This is an expected effect, since the modulation of the PSF/OTF was done in the x-direction. Reduction in noise in the y-axis is the result of the small changes in the OTF that resulted from the slight changes of aperture

shape in the y-axis (i.e. loosing area along the neutral axis), and from any noise suppression that came from averaging several images together. That this reduction isn't as strong as the reduction along the x-axis demonstrates that the increase in performance along the x-axis must be due to the deliberate zero-cancelation that came from modulating the PSF, and in turn demonstrates that the "programmable deconvolution" technique available from ACPA cameras is effective. Figure B-4, a replication of Figure 7-5, demonstrates the selective modulation of the OTF along the x-axis compared to the y-axis for the images used in the image stack. Figure B-5 confirms that this selective modulation effect exists well beyond the stack's 4.0mm-5.0mm aperture range.

A summary of the progression from original target image to NPSFS result is shown in Figure 7-10. The purpose of this figure is simply to demonstrate in one place the transformation that occurs at each step of the image corruption and reconstruction process. To aid in an apples-to-apples comparison, the brightness of each image has been normalized to the full dynamic range.

7.6   Conclusions and Future Work

Consistent with the findings in [47] and the work done by many researches that builds upon them, the results presented in this chapter demonstrate that attaining good frequency-domain coverage and OTF zero-cancelations across an image stack yields deconvolutions that are superior to single-image results. In addition, this work makes a new contribution to the field of deconvolution by developing a new technique, "Programmable Deconvolution," that can be achieved with ACPA cameras. This new technique allows for the rapid capture of data for use in high-resolution deconvolutions. By requiring only a handful of images which can be taken in rapid succession with known, deterministic aperture shapes (as opposed to Monte Carlo-style stochastic methods), this new technique allows for unprecedentedly fast capture times. This speed-up can reduce or eliminate the need for layers of complexity that compensate for scene changes over long time scales. Perhaps more importantly, it opens up the

possibility of imaging and deconvolving quasi-stationary scenes and objects which change too quickly for the timescales required by previous techniques.

In the future, this work can be extended in several important ways. For example, a thorough analysis of the PSF modulations that can be achieved with a fully-capable digital aperture (as opposed to our simplified vertical-slit-and-iris aperture) imaging system would be a valuable contribution. Additionally, a framework and codebase for calibrating ACPA cameras for programmable deconvolution work would be valuable. Finally, a framework for optimizing the shapes and minimizing the number of images that need to be captured to yield a strong deconvolution would be an important advancement as well.
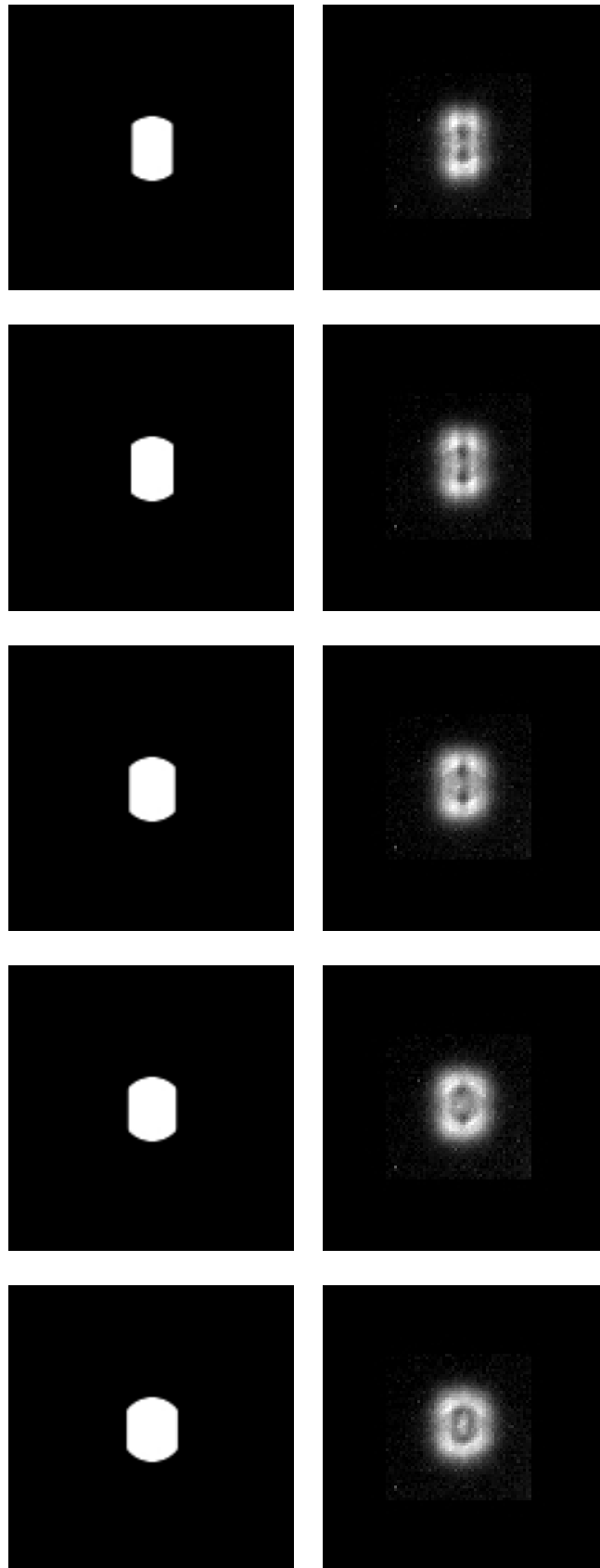
Figure 7-4: Aperture shapes created by slit widths varying from 4.0mm to 5.0mm in 0.25mm steps, alongside their resulting observed PSFs
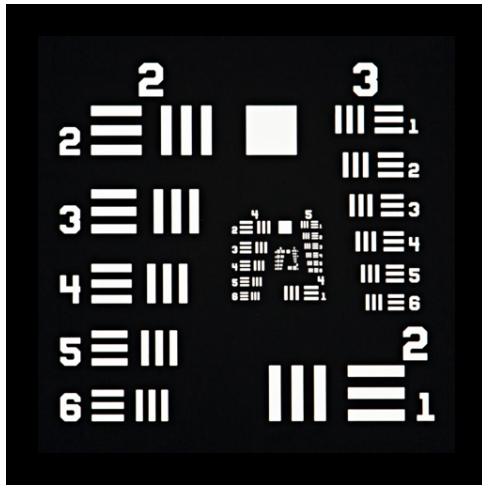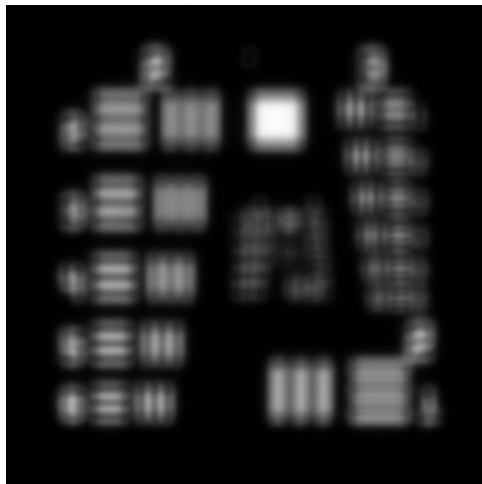
(a) Reference image
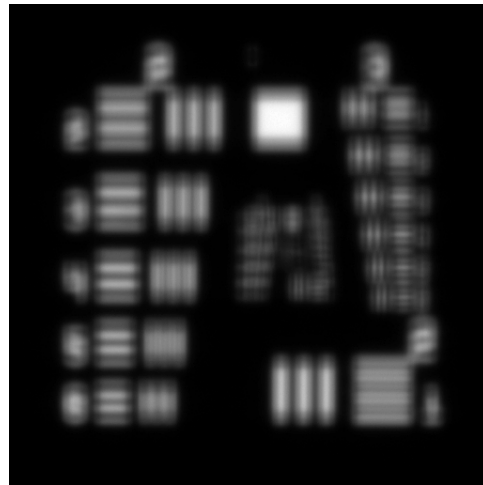
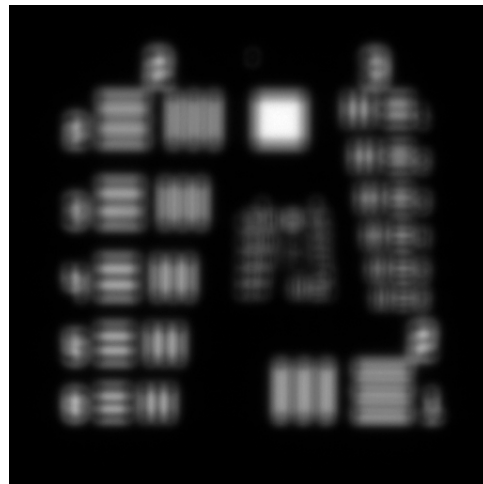(b) Observation @ 4.0mm slit width
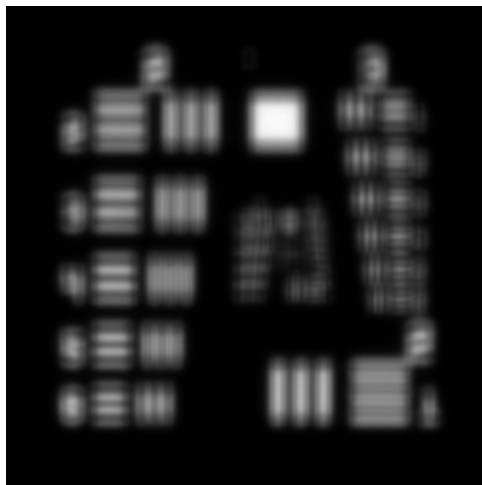
(c) Observation @ 4.25mm slit width
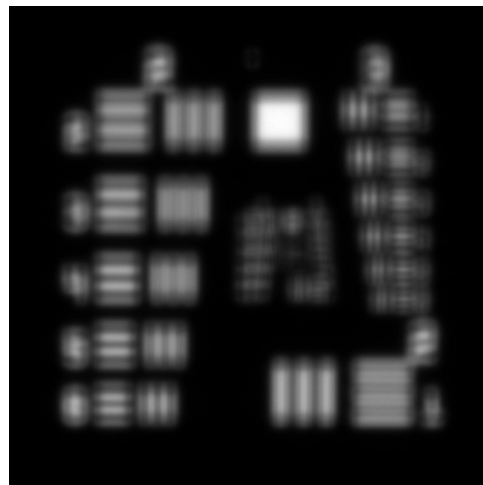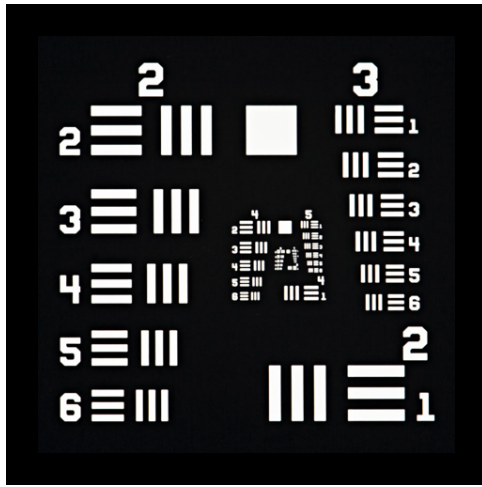
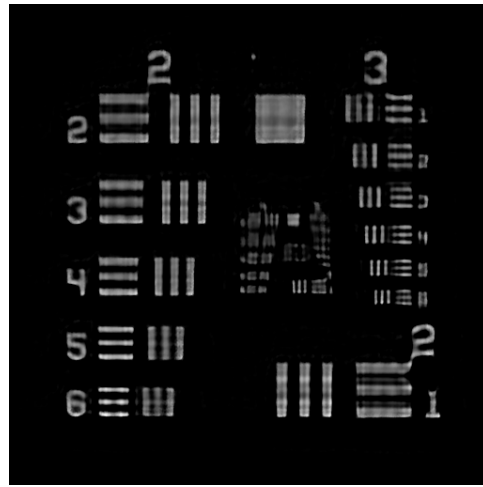(d) Observation @ 4.5mm slit width

(e) Observation @ 4.75mm slit width
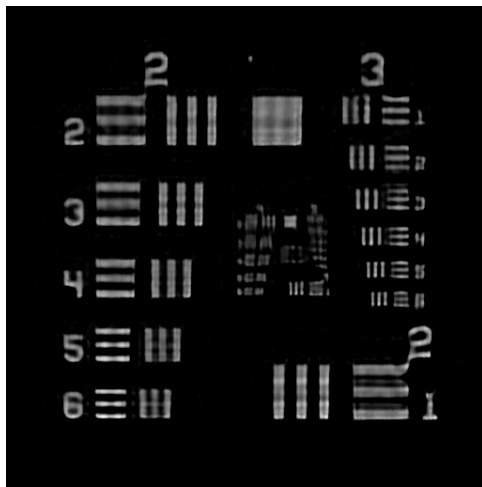
(f) Observation @ 5.0mm slit width

Figure 7-7: Observed blurry test target images captured at aperture shapes ranging from 4.0mm to 5.0mm in 0.25mm steps.
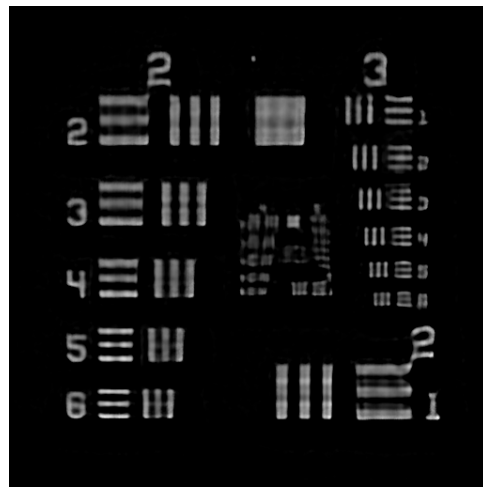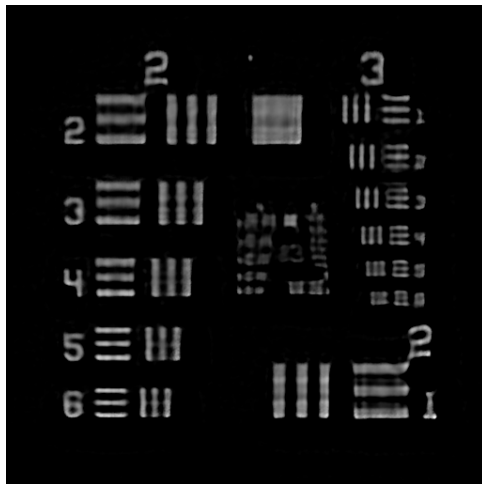
(a) Reference image

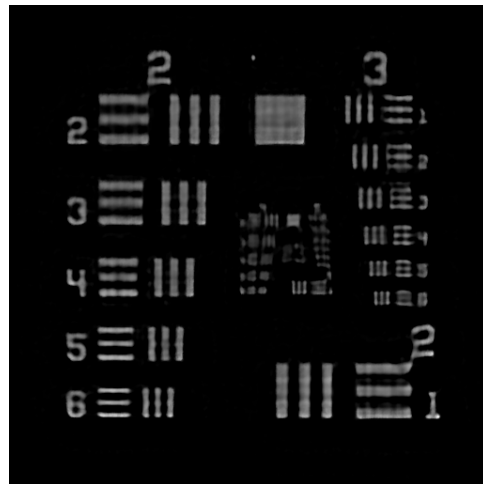(b) Myopic decon @ 4.0mm slit width

(c) Myopic decon @ 4.25mm slit width
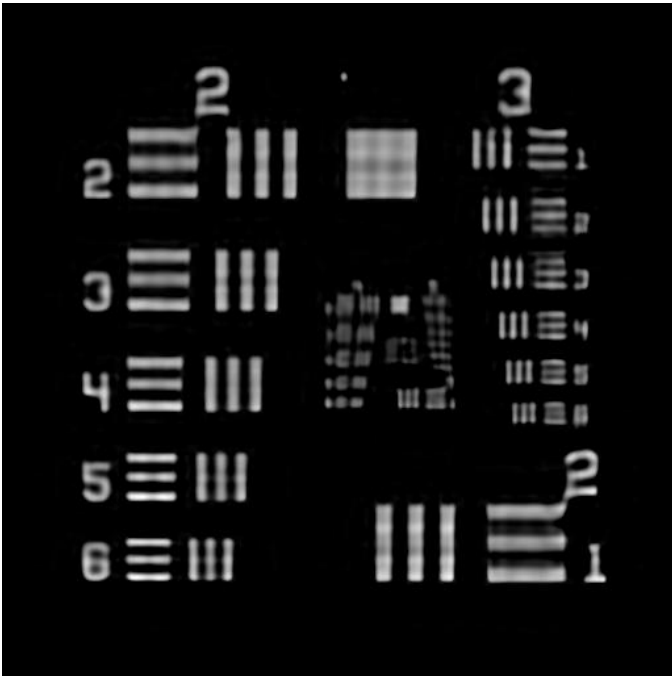
(d) Myopic decon @ 4.5mm slit width

(e) Myopic decon @ 4.75mm slit width

(f) Myopic decon @ 5.0mm slit width
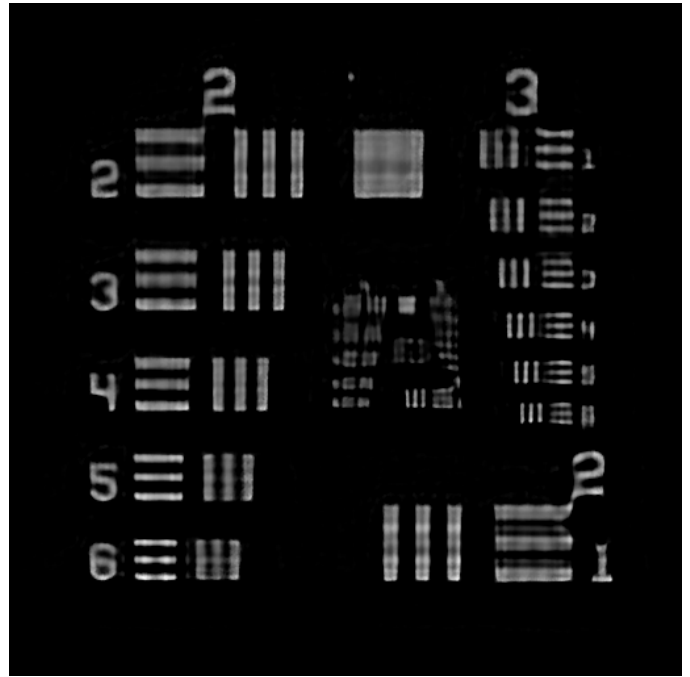
Figure 7-8: Single-image myopic deconvolution of blurry test target images captured at aperture shapes ranging from 4.0mm to 5.0mm in 0.25mm steps.
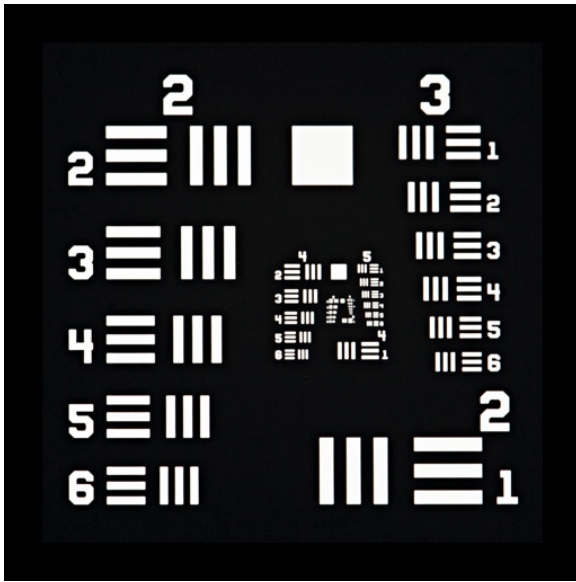
(a) Simultaneous 'NPSFS' multi-image deconvolution result

(b) Sharpest single-image deconvolution result (4.0mm slit width)

Figure 7-9: Comparison of multi-image deconvolution result to the best single-image deconvolution in the stack.

(a) Reference Test Target Image
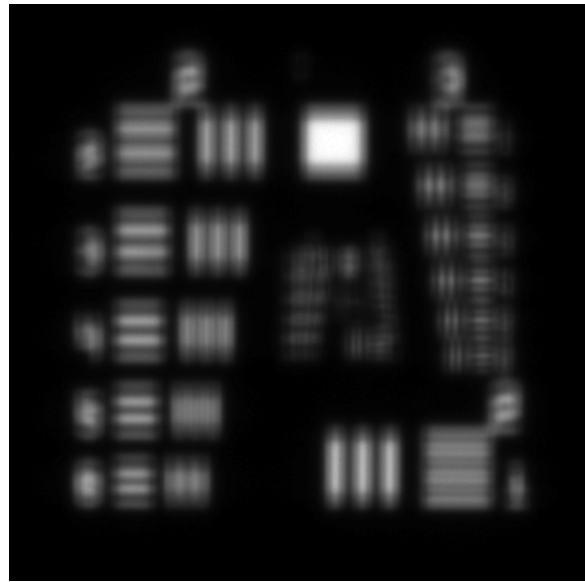
(b) Sharpest blurred image, taken @ 4.0mm slit width
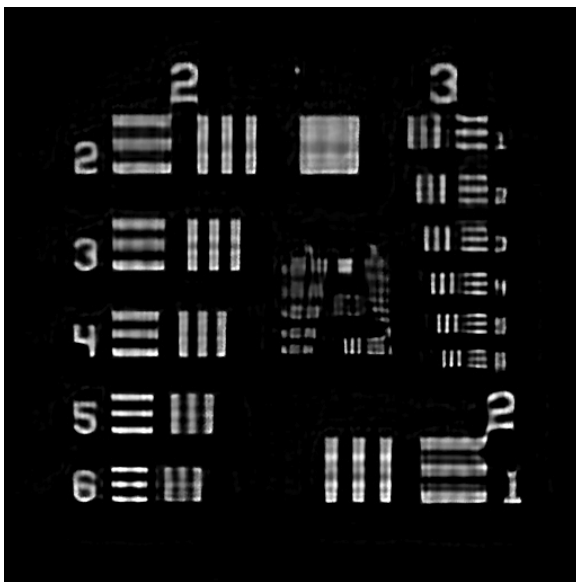
(c) Single-image deconvolution of image taken @ 4.0mm slit width
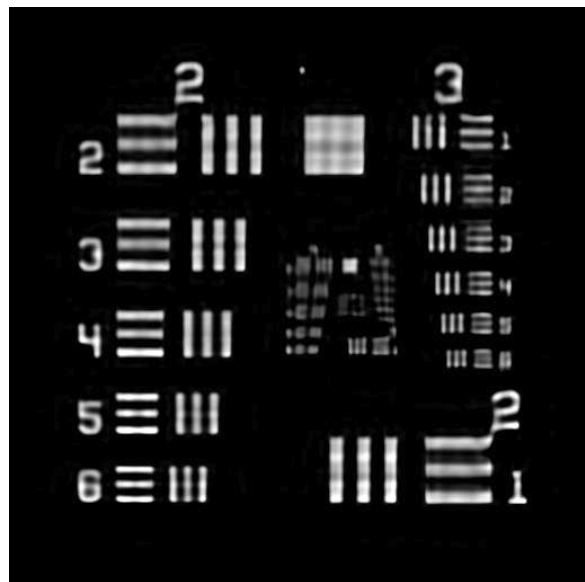
(d) Multi-image deconvolution incorporating image data from 4.0mm to 5.0mm slit widths in 0.25mm steps, and exhibiting reduced high-frequency artifacts.

Figure 7-10: Demonstration of the effects of the various steps of the corruption and reconstruction process. The white levels of each image have been normalized to the full dynamic range.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 8

Conclusions, Future Work & Discussion

8.1 Conclusions

From the work presented in the previous chapters, it is clear that light field imaging offers some exciting advantages over traditional imaging techniques. These advantages include the ability to re-focus images and adjust the depth of field after capture, and the ability to tweak perspective after capture. Furthermore, the arbitrarily-controllable programmable-aperture (ACPA) technology described in this thesis enables the acquisition of light field data for these purposes with unparalleled flexibility. This important hardware advancement also enables new imaging techniques that were off-limits to previous light field camera designs, for example by modulating the frequency response of the optical train for purposes of "programmable Deconvolution." Together, these hardware and software techniques open new doors to imaging our world.

8.2 Future Work

The work presented in this thesis can be extended in many exciting ways. What follows in this section is a collection of the research paths that the author believes

have particular merit.

8.2.1 Aberration Correction

It was demonstrated in [7] that lens aberrations could be quantified and corrected for by sampling the aperture plane of a lens. This work was applied to a single aperture rotating at fixed radius around the optical axis. This approach could be generalized to ACPA cameras for the purpose of correcting aberrations across the entire lens. This would be an exciting advancement because it would enable high quality imaging with cheap lenses. Using cheap glass or even plastic lenses in place of high-cost, precision-ground, many-element glass lenses could drastically lower the price of high-end imaging systems when equipped with an ACPA.

8.2.2 3D Scanning

As has been mentioned in numerous places in this thesis already, ACPA cameras have great potential for 3D scanning, especially on small size scales. While this thesis advances some work on designing ACPA cameras for scanning and possible algorithmic approaches to computing depth, there is much room for improvement. Perhaps most saliently, this work can be improved by utilizing the latest advancements in machine vision algorithms (like cutting edge feature detection and optical flow), and quantifying the practical limitations on resolutions placed on systems with real (i.e. non-ideal) lenses.

8.2.3 Extensions of "Programmable Deconvolution"

Chapter 7 introduces the concept of ACPA-based frequency modulation for the purposes of null-filling and deconvolution. The results presented there are based on single-axis modulation. Extending this work to demonstrate its effectiveness with two dimensions of modulation, and a theoretical framework for calculating optimal modulations for a given imaging condition, would be an important step forward. Finally, extending this work to color images, which was not done in this thesis, would

be additional validation of the value of this new technique.

8.2.4 Superresolution Imaging

Though short, Chapter 6 opens the dialogue on the use of ACPAs for certain kinds of superresolution imaging. Specifically, an ACPA camera can be used to augment the resolution of a low (i.e. subsampling) resolution sensor. Like the correction of lens aberrations mentioned above, this ability of ACPAs could enable high-quality imaging from increasingly inexpensive hardware.

8.2.5 Extensions to Non-Visible Wavelengths

Finally, it is worth mentioning that none of the new techniques presented in this thesis is necessarily constrained to the visible spectrum of light. There is potentially very interesting work to be done by extending these techniques to non-visible regimes. One could imagine this being particularly interesting for astronomical observations and biological studies, where features of interest are frequently found outside of the spectrum visible by the human eye.

8.3 Discussion

Since one very good metric for the value of research lies in how this research can improve our lives, it seems appropriate to conclude this thesis with a brief discussion of how the technologies advanced herein could effect us in positive ways. To that end, the author would like to conclude by drawing attention to three interesting and important avenues.

First, light field imaging enables an unprecedented amount of information to be captured in a photo. This ability to capture "more information" means better images of our friends & family and of life's most meaningful events. In this vein, light field imaging can improve the way that we capture, share and remember our lives.

Secondly, and on a related note, new light field imaging technologies may usher in

a new age of audio-visual entertainment. Current methods of 3D cinematography are unsatisfying, relying on stereography techniques that are usually divorced from the way a human would actually have observed a scene in real life. Light field imaging offers potential solutions to this problem by enabling both the capture and playback of a scene in a richer and more natural manner.

Finally, this thesis' light field imaging technologies allow new inroads into observing our microscopic and astronomic world. As exploration is a very human pursuit, the author takes pride in having made a contribution–however small–to our ability to better explore the world around us.

Appendix A

Selected Data for LCDs Used As Digital Apertures

Two different LCDs were used as the primary digital apertures for the prototype cameras presented in this thesis. The first of the two, used in earlier camera versions, is made by Varitronix Limited, model COG-T350M6080. This display was driven by a PIC18 microcontroller, running custom code written by the author of this thesis according to the display's controller/driver chip's specifications.

The second display, used in later prototypes, is a Sony Corporation LCX023 microdisplay (without microlenses), which was driven by a DVI-based driver board from HOLOEYE Photonics AG.

The following excerpted pages from the two displays' datasheets contain physical specifications for the displays. They are included here to provide insight into how the prototype cameras presented in this thesis were built and controlled, and also to document the optical properties of these displays that had critical effects (most notably diffractive ghosting) on the imaging systems in which they were placed.

The excerpted pages that follow were published by Varitronix Limited and Sony Corporation, respectively, as part of complete datasheets. The complete datasheet for the Sony microdisplay can be downloaded from `http://www.datasheetcatalog.com/`. The Varitronix display's datasheet is available from the manufacturer.

VARITRONIX LIMITED

Preliminary Specification
of
LCD Module Type
Model No.: COG-T350M6080-01

1. **General Description**

   - 3.5"(diagonal), 240 x 3 x 320 dots, portrait 3:4, transmissive, mono TFT LCD module.
   - Viewing angle: 6 o'clock.
   - LCD controller/driver: "NOVATEK" NT39122B or equivalent.
   - 8080 MCU 8-bit parallel interface, 6800 MCU 8-bit parallel interface or 3-pin serial interface.
   - Logic voltage: 2.8V (Typ.).
   - With anti-glare front polarizer.
   - White LED backlight.
   - FPC connection.

2. **Mechanical Specifications**

   The mechanical detail is shown in Fig. 1 and summarized in Table 1 below.

Table 1

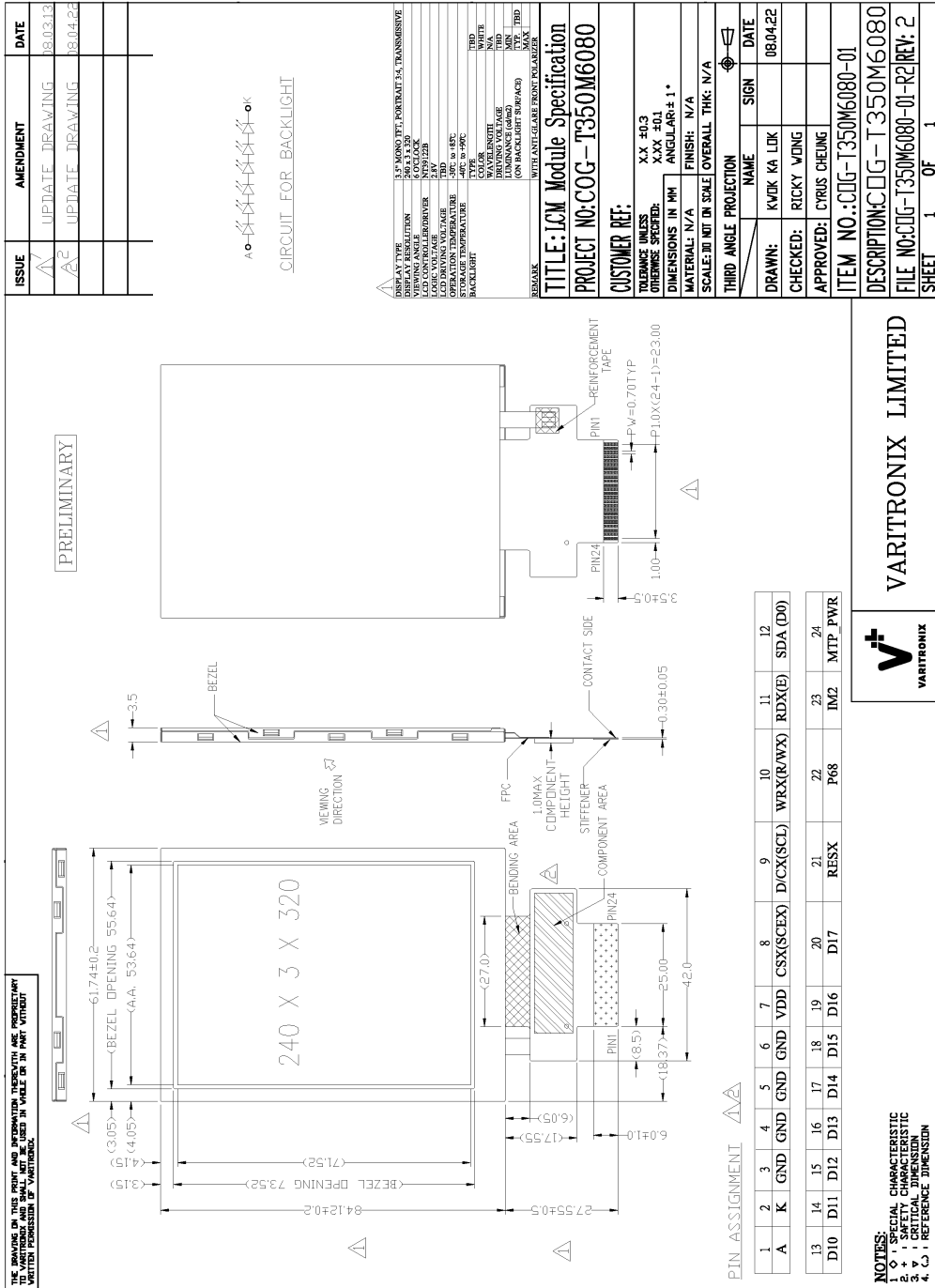| Parameter		Specifications	Unit
Outline dimensions		61.74(W) x 84.12(H) x 3.5(D) (excluded FPC and component area)	mm
Mono TFT 240x3x320	Bezel opening	55.64(W) x 73.52(H)	mm
	Active area	53.64(W) x 71.52(H)	mm
	Display format	240 x 3 x 320	dots
	Dot pitch	0.2235(W)(3 successive horizontal dots) x 0.2235(H)	mm
Weight		TBD	grams

Figure 1: Outline Drawing

6. Optical Characteristics

6.1 Specification

Table 10 (Ta=25 $^{\circ}$C)

Item		Symbol	Condition	Values			Units	Remarks
				Min	Typ	Max		
Viewing Angle	Horizontal	ϕ1(3 o'clock), ϕ2(9 o'clock)	CR>10	55	60	-	deg	Note 1
	Vertical	θ2(12 o'clock)		35	40	-		
		θ1(6 o'clock)		50	55	-		
Contrast Ratio		CR	At optimized viewing angle	-	TBD	-	-	Note 2
Response time	Rise	Tr	θ=0°,ϕ=0°	-	15	30	ms	Note 3
	Fall	Tf		-	25	50		
Brightness			θ=0°,ϕ=0°	-	600	-	cd/m^2	

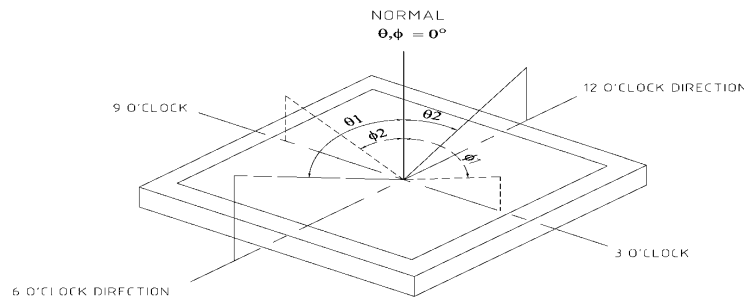Note 1: The definitions of viewing angles

Figure 18

Note 2:

$$CR = \frac{\text{Luminance when Testing point is White}}{\text{Luminance when Testing point is Black}}$$

Contrast ratio is measured in optimum common electrode voltage.

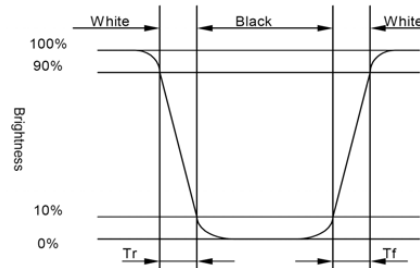Note 3: The definition of response time:

Figure 19

110

SONY

LCX023CMT

## 3.3cm (1.3-inch) Black-and-White LCD Panel

Description

The LCX023CMT is a 3.3cm diagonal active matrix TFT-LCD panel addressed by polycrystalline silicon super thin film transistors with a built-in peripheral driving circuit. Use of three LCX023ALB panels provides a full-color representation. The striped arrangement suitable for data projectors is capable of displaying fine text and vertical lines.

The adoption of DMS (Dual Metal Shield) structure realizes a high luminance screen. And new cross talk free and ghost free structures contribute to high picture quality.

This panel has a polysilicon TFT high-speed scanner and built-in function to display images up/down and/or right/left inverse. The built-in 5V interface circuit leads to lower voltage of timing and control signals.

The panel contains an active area variable circuit which supports S-XGA 5:4 and PC-98 8:5 data signals by changing the active area according to the type of input signal. Also incorporating microlens can increase efficiency of incident light.
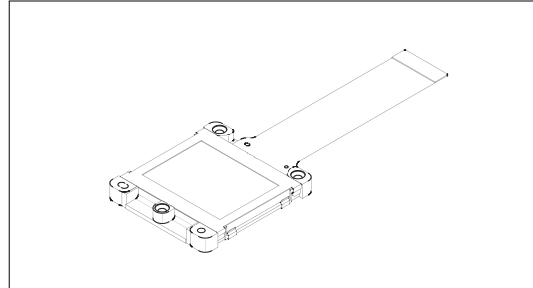
Features

- Number of active dots: 786,432 (1.3-inch, 3.3cm in diagonal)
- XGA display
- SXGA viewable
- High optical transmittance: 30% (typ.)
- New high light resistance DMS (Dual Metal Shield) structure adopted
- Built-in-new cross talk free circuit and ghost free circuit
- High contrast ratio with normally white mode: 250 (typ.)
- Built-in H and V drivers (built-in input level conversion circuit, 5V driving possible)
- Up/down and/or right/left inverse display function
- Antidust glass package

Element Structure

- Dots: 1024 (H) × 768 (V) = 786,432
- Built-in peripheral driver using polycrystalline silicon super thin film transistors

Applications

- Liquid crystal data projectors
- Liquid crystal multimedia projectors
- Liquid crystal rear-projector TVs, etc.

* The company's name and product's name in this data sheet is a trademark or a registered trademark of each company.

– 1 –

E98831A93-PS

Electro-optical Characteristics (XGA mode)

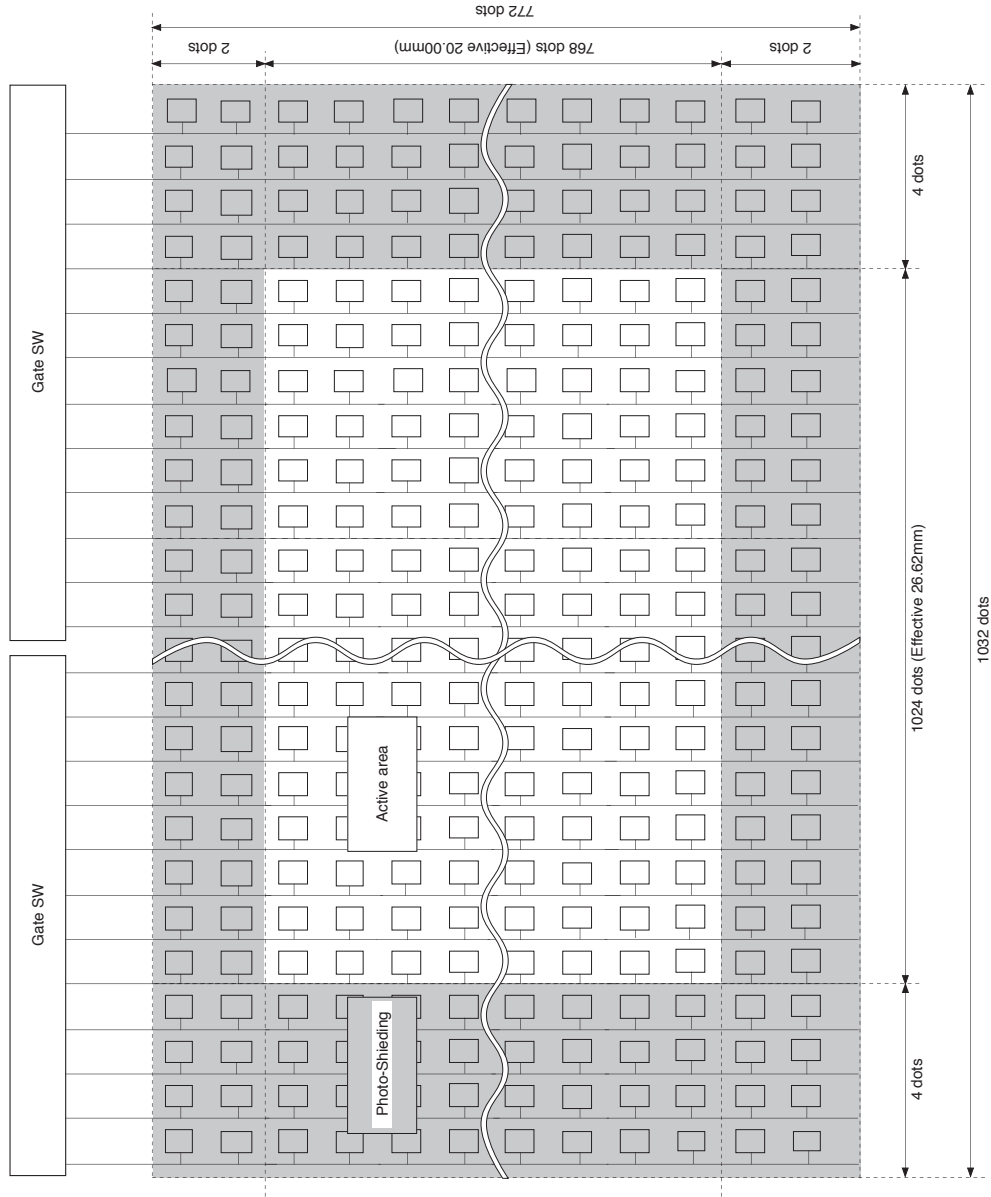| Item			Symbol	Measurement method	Min.	Typ.	Max.	Unit
Contrast ratio		25°C	CR	1	150	250	—	—
Optical transmittance		25°C	T	2	26	30	—	%
V-T characteristics	V_{90}	25°C	RV_{90-25}	3	0.8	1.2	1.5	V
			GV_{90-25}		0.9	1.3	1.6	
			BV_{90-25}		1.0	1.4	1.7	
		60°C	RV_{90-60}		0.7	1.1	1.4	
			GV_{90-60}		0.9	1.3	1.6	
			BV_{90-60}		0.9	1.3	1.6	
	V_{50}	25°C	RV_{50-25}		1.1	1.5	1.8	
			GV_{50-25}		1.2	1.6	1.9	
			BV_{50-25}		1.3	1.7	2.0	
		60°C	RV_{50-60}		1.1	1.5	1.8	
			GV_{50-60}		1.2	1.6	1.9	
			BV_{50-60}		1.2	1.6	1.9	
	V_{10}	25°C	RV_{10-25}		1.6	2.0	2.3	
			GV_{10-25}		1.7	2.1	2.4	
			BV_{10-25}		1.8	2.2	2.5	
		60°C	RV_{10-60}		1.6	2.0	2.3	
			GV_{10-60}		1.7	2.1	2.4	
			BV_{10-60}		1.7	2.1	2.4	
Response time	ON time	0°C	ton0	4	—	28.0	80.0	ms
		25°C	ton25		—	14.0	40.0	
	OFF time	0°C	toff0		—	72.0	200.0	
		25°C	toff25		—	34.0	70.0	
Flicker		60°C	F	5	—	−67.0	−40.0	dB
Image retention time		25°C	YT60	6	—	0	—	s
Cross talk		25°C	CTK	7	—	—	5	%

Reflection Preventive Processing

When a phase substrate which rotates the polarization axis is used to adjust to the polarization direction of a polarization screen or prism, use a phase substrate with reflection preventive processing on the surface. This prevents characteristic deterioration caused by luminous reflection.

– 12 –

Viewing angle characteristics (Reference value without microlens)

Measurement method

– 16 –

113
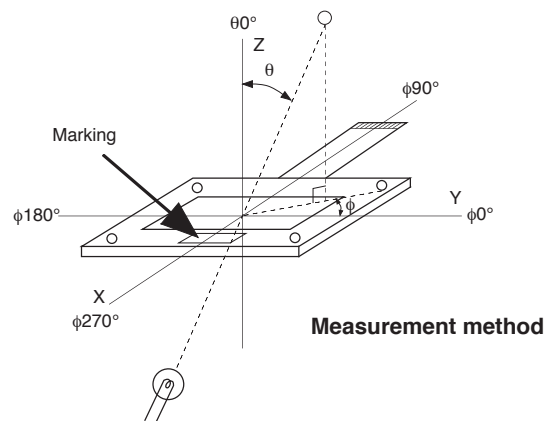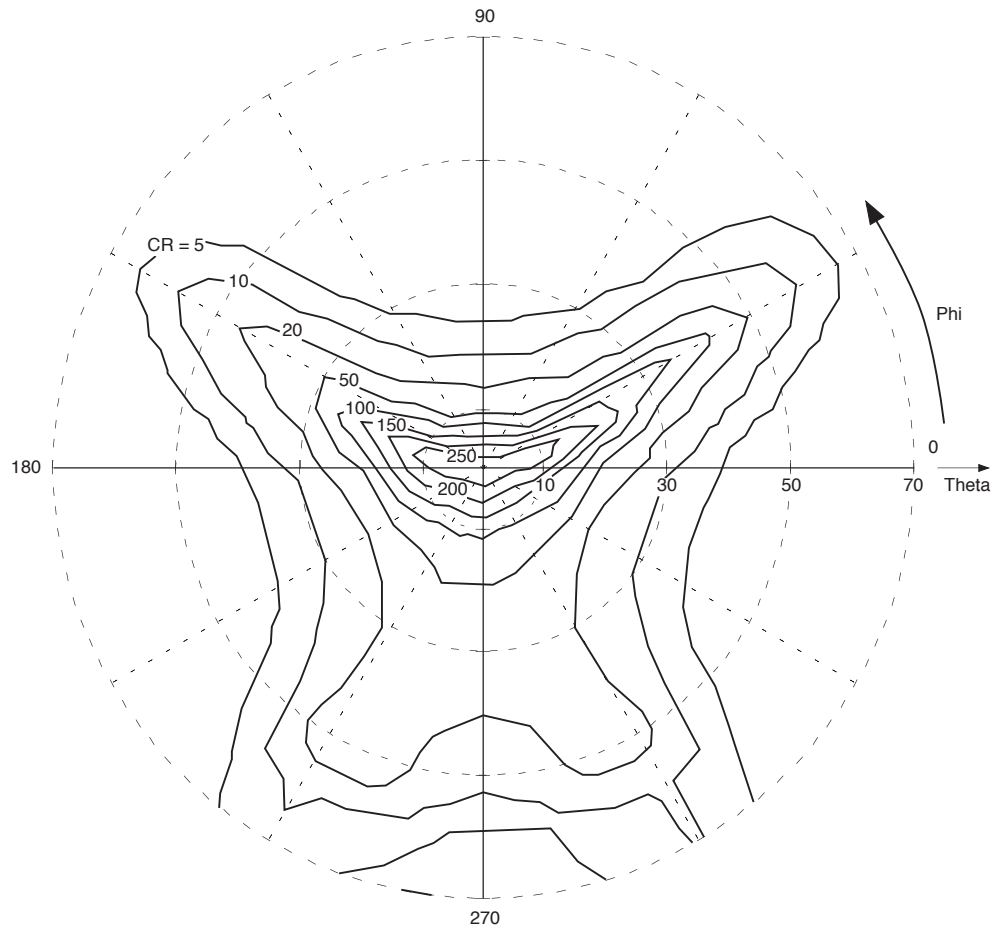
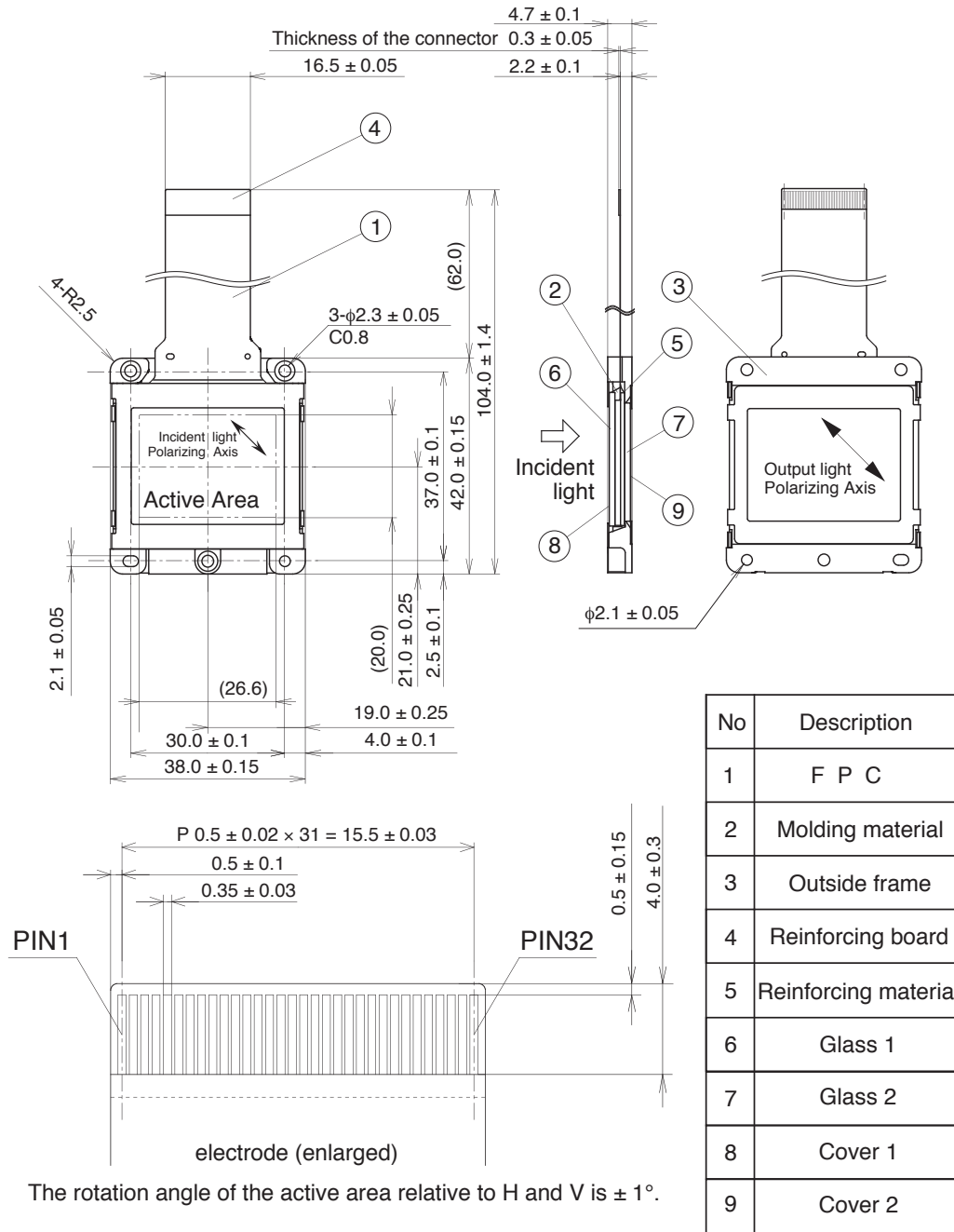**1. Dot Arrangement**

The dots are arranged in a stripe. The shaded area is used for the dark border around the display.

Package Outline   Unit: mm

| No	Description
1	F P C
2	Molding material
3	Outside frame
4	Reinforcing board
5	Reinforcing material
6	Glass 1
7	Glass 2
8	Cover 1
9	Cover 2

The rotation angle of the active area relative to H and V is ± 1°.

weight 12.4g

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix B

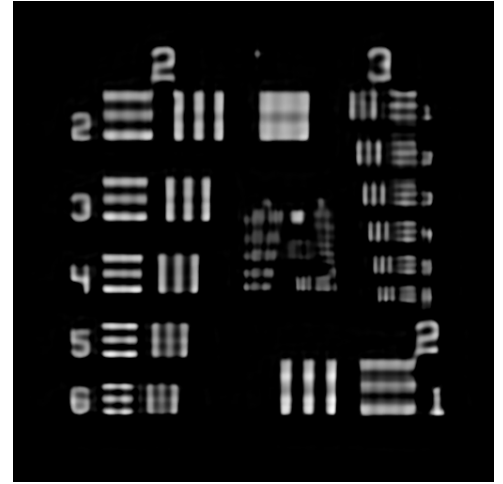# Supplemental Material for Programmable Deconvolution

This appendix contains additional material relating to programmable deconvolution; it is intended to be a companion to Chapter 7. It contains many relevant images and plots that provide important details about the results presented in Chapter 7, but these images and plots were deemed too voluminous to fit cleanly into the chapter prose without detracting from the higher-level results presented there. All of the deconvolution image results were computed with AIDA [14], and the plots were generated in MATLAB. All of the frequency-domain results presented in Chapter 7 and this Appendix are based on images with a size of 512x512 pixels. The X-axis in most graphs is cropped to 120 or 150 (as relatively little information existed beyond there), but the full range is 0-256.
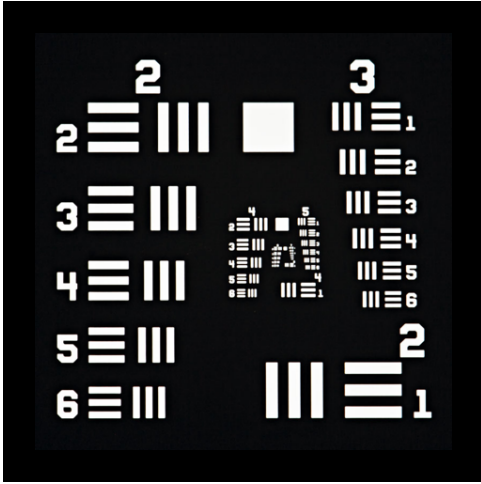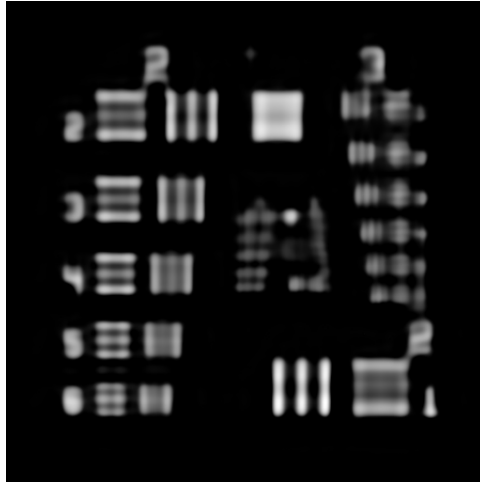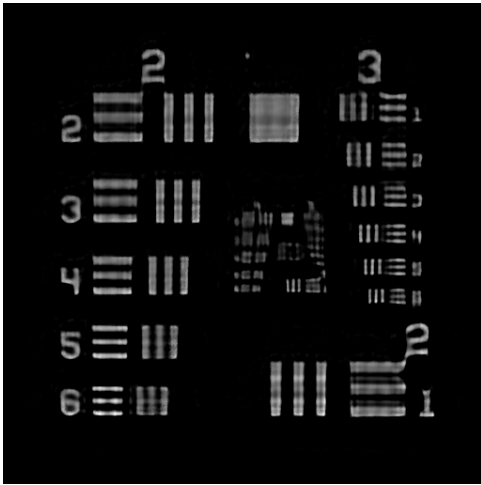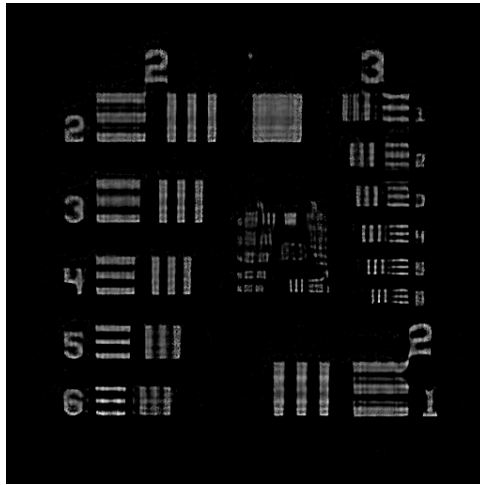
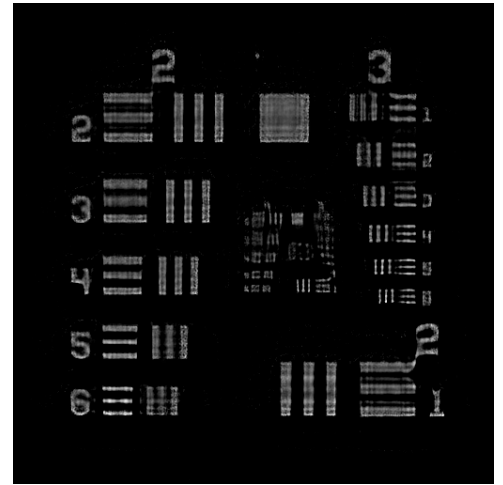(a) Reference image (b) lambda_object_scaling = 1.0 (c) lambda_object_scaling = 0.1
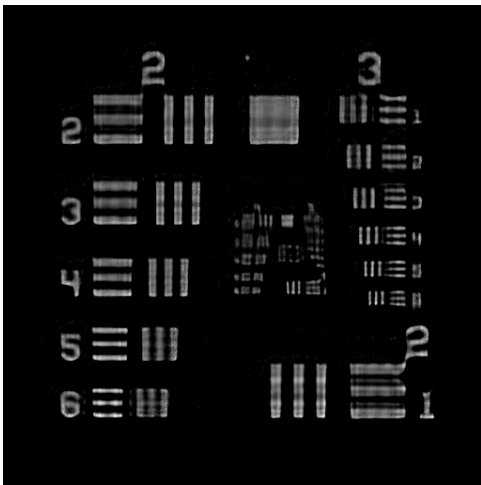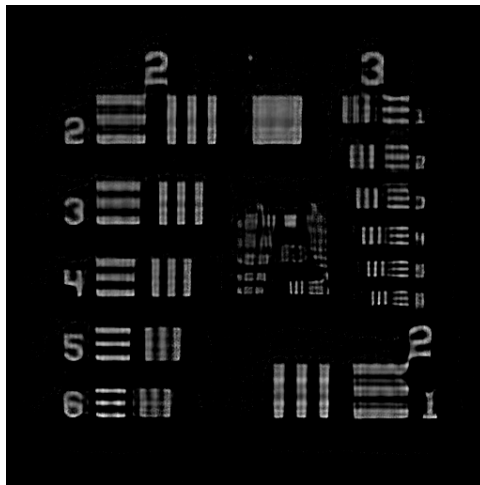
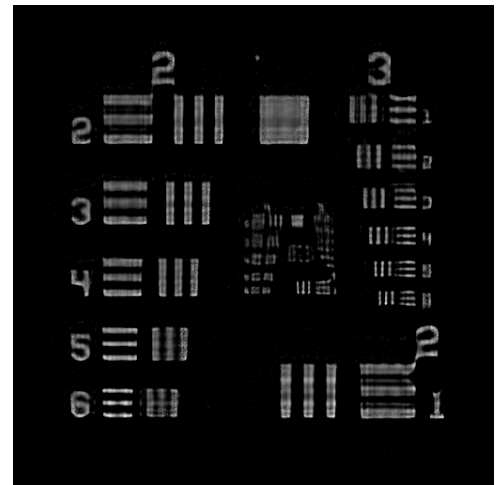(d) lambda_object_scaling = 0.01 (e) lambda_object_scaling = 0.001 (f) lambda_object_scaling = 0.0001

(g) lambda_object_scaling = 0.00001 (h) lambda_object_scaling = 0.000001 (i) lambda_object_scaling = 0.0000001

Figure B-1: Effects of different lambda_object_scaling values on AIDA myopic decon-
volution results

118

Figure B-2: Central horizontal OTF profiles for ground truth, observed image & PSF, and myopically deconvolved image & corrected PSF
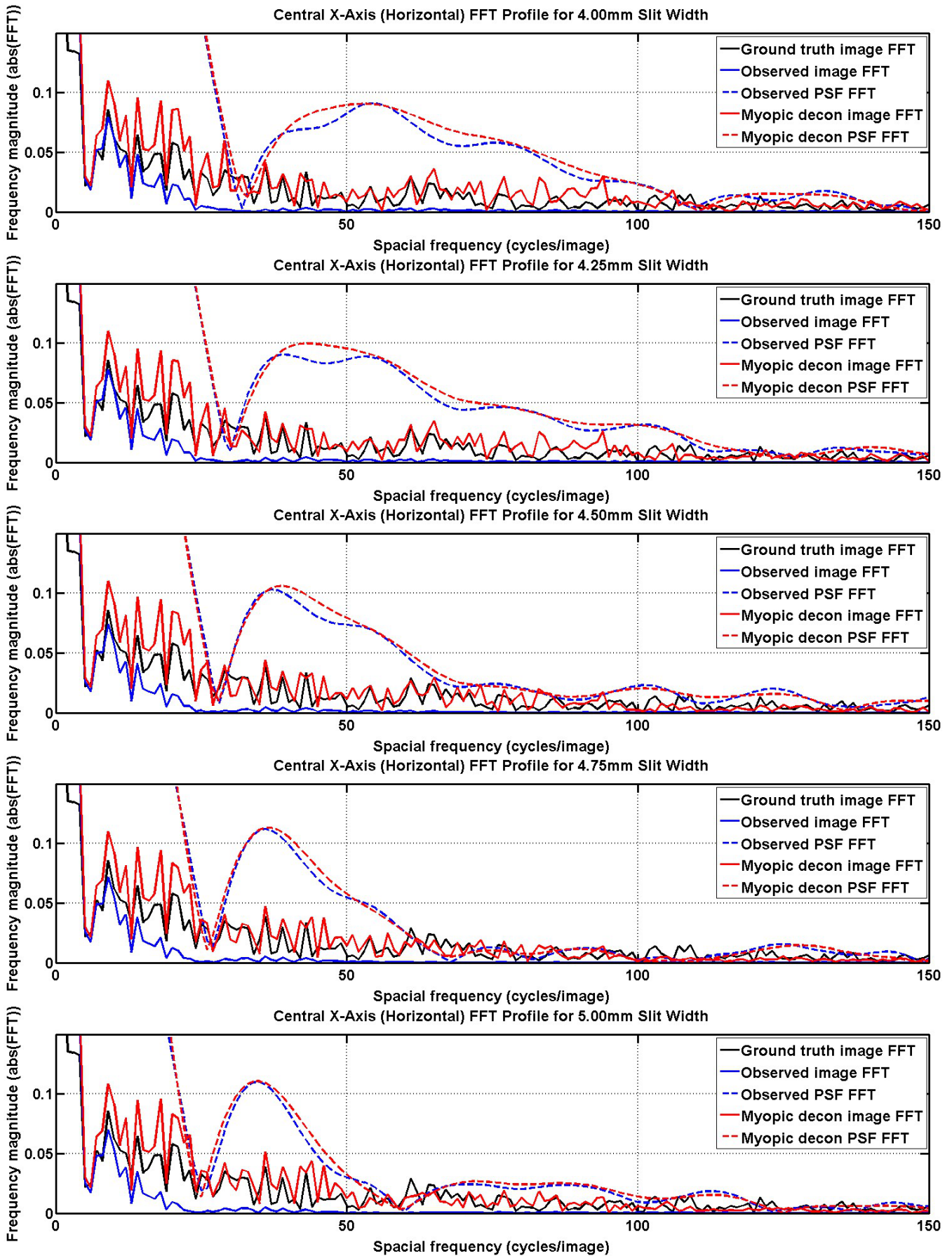
Figure B-3: Central vertical OTF profiles for ground truth, observed image & PSF, and myopically deconvolved image & corrected PSF
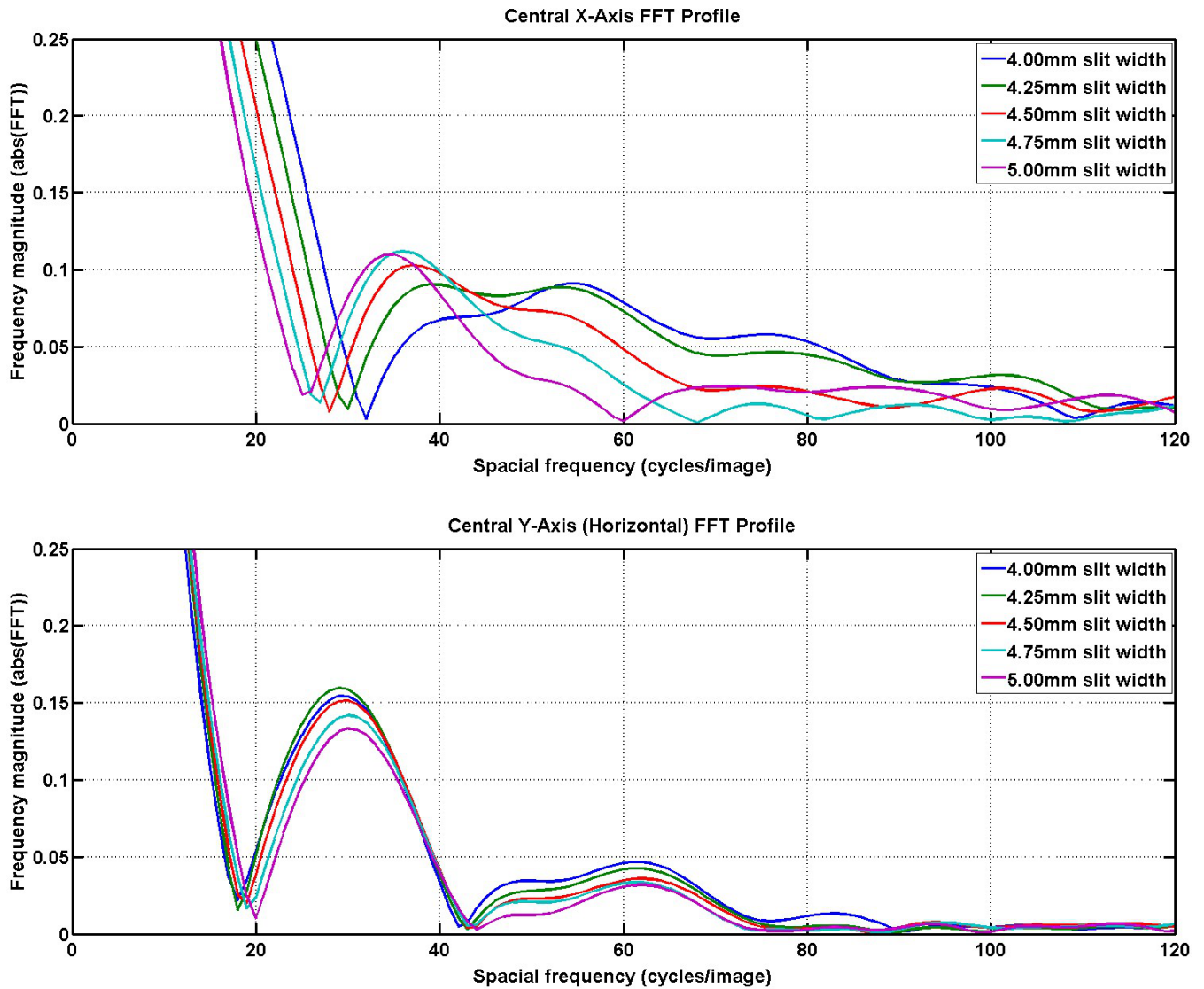
Figure B-4: Central horizontal & vertical OTF profiles for all observed PSFs in the image stack with slit widths ranging from 4.0mm to 5.0mm in 0.25mm steps
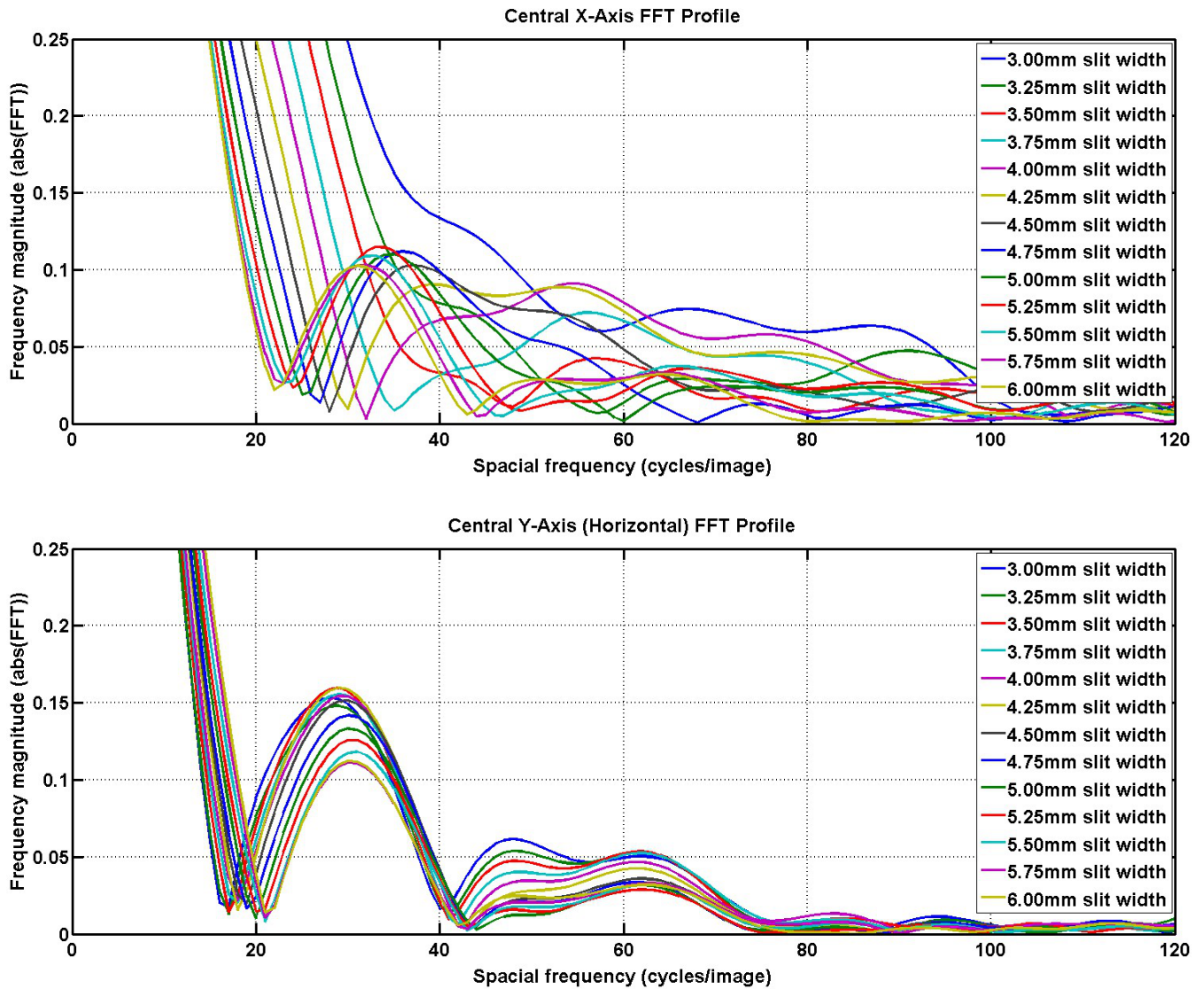
Figure B-5: Central horizontal & vertical OTF profiles for all observed PSFs in the image stack with slit widths ranging from 3.0mm to 6.0mm in 0.25mm steps
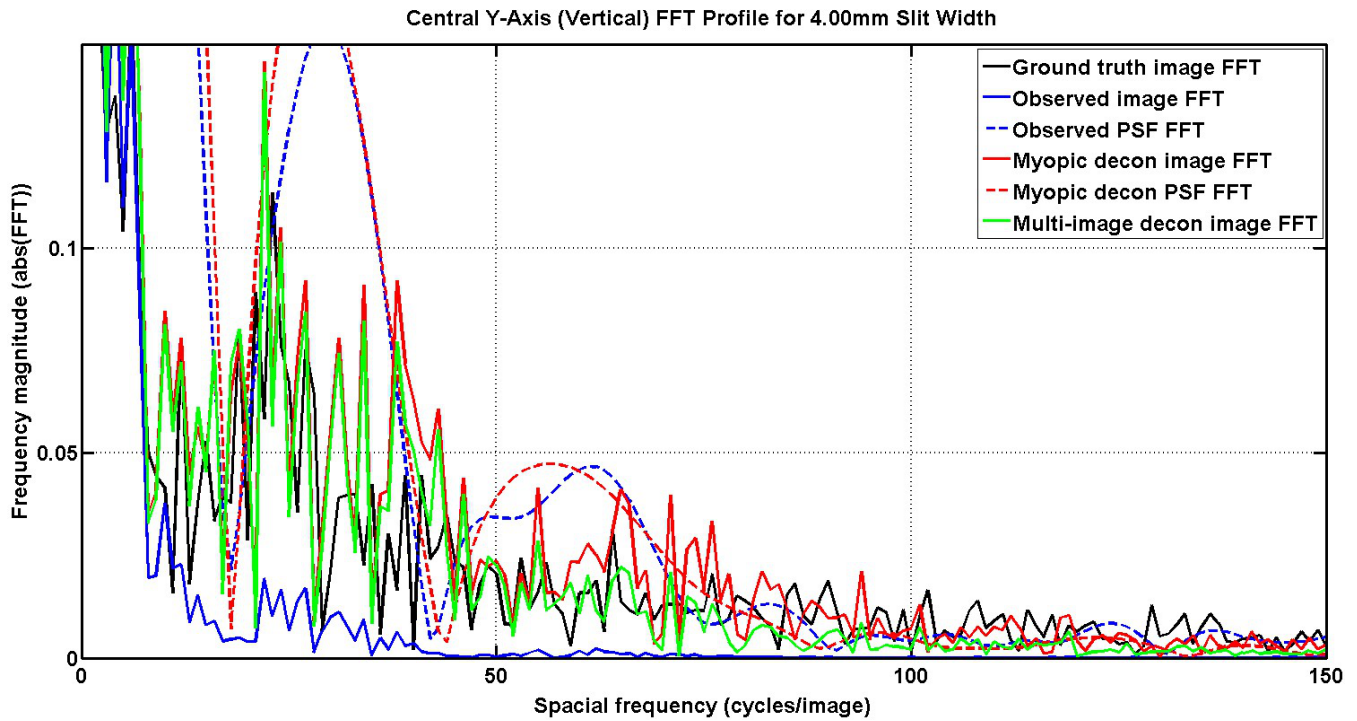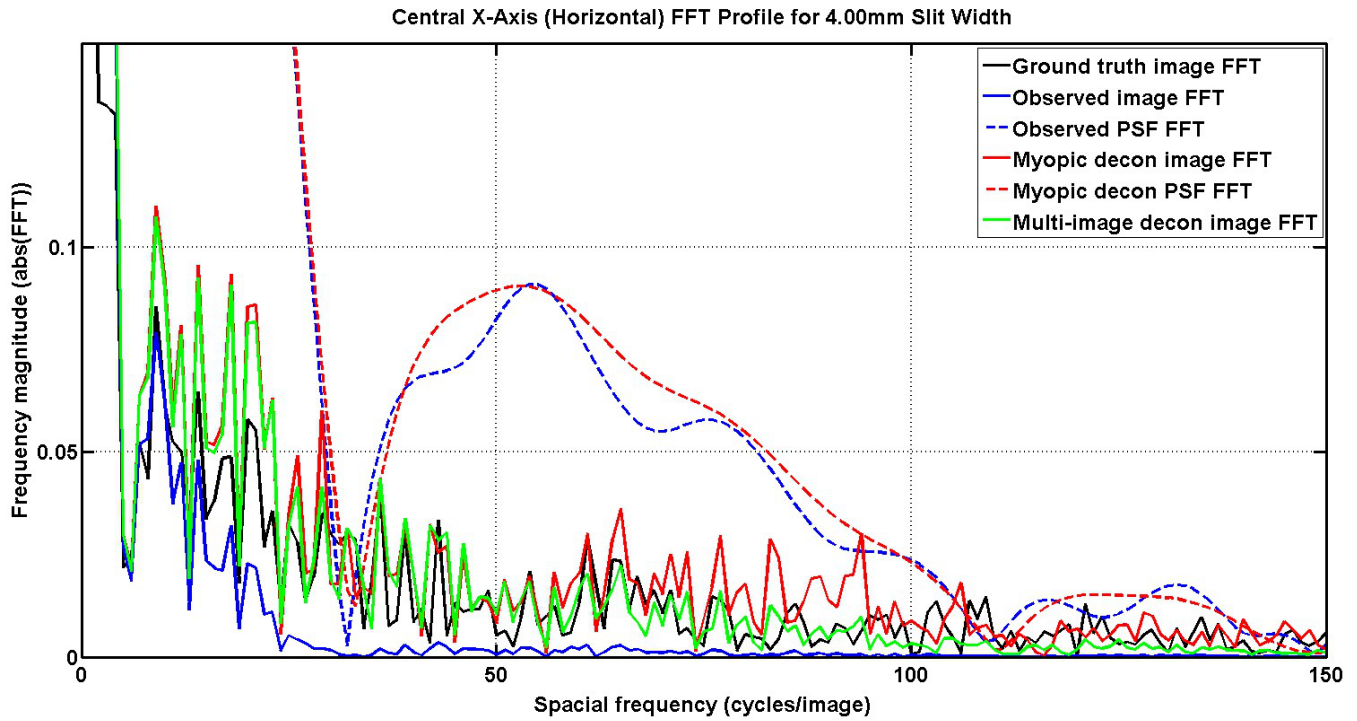
Figure B-6: Central horizontal & vertical OTF profiles for ground truth, observed image & PSF, myopically deconvolved image & PSF, and the Multi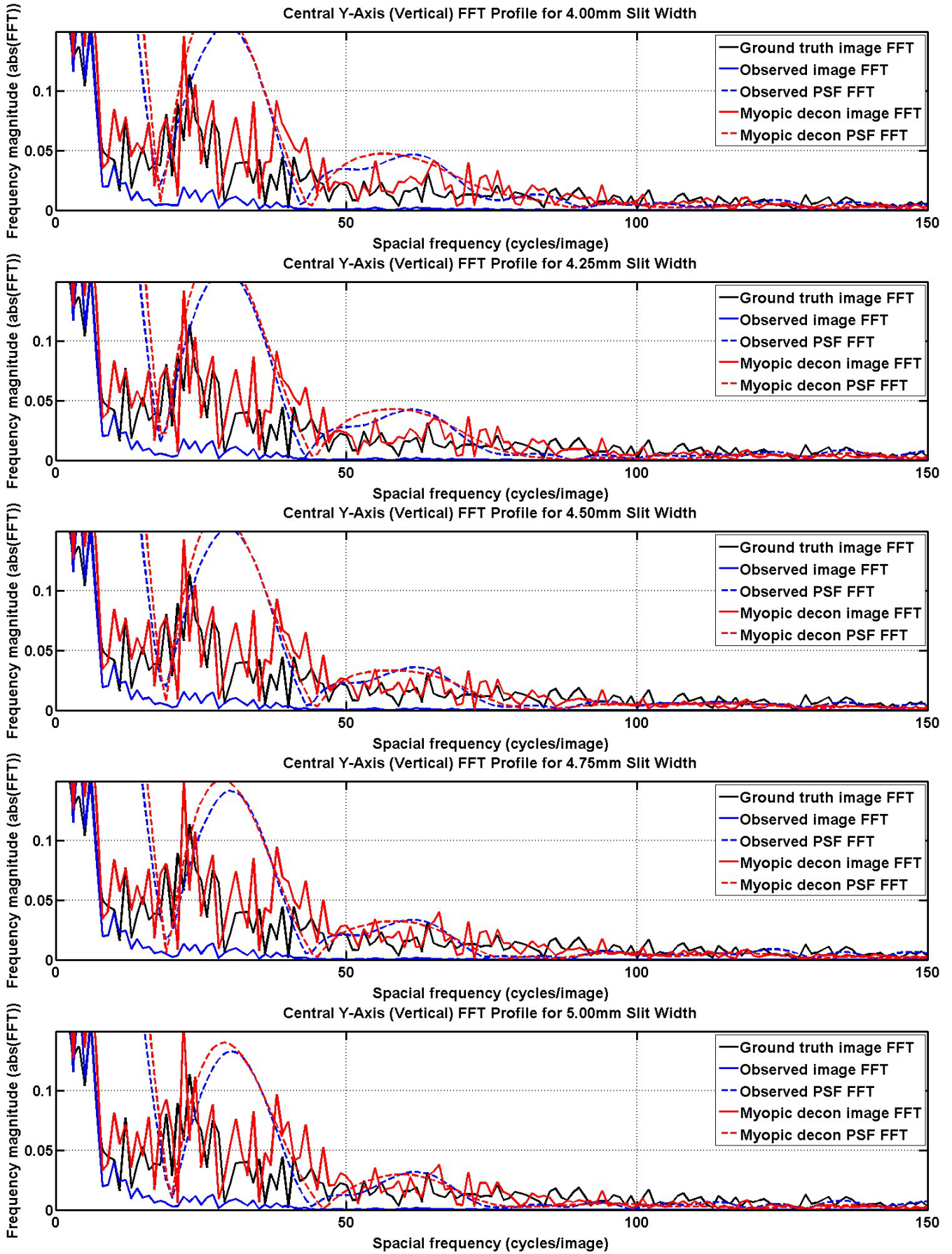-image deconvolution image result. The single-image myopic decon was done with the sharpest image (4.0mm), while the multi-image decon included four more images with *increased* blurriness.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

[1] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359, 2008.

[2] M Bertero and P Boccacci. Image restoration methods for the large binocular telescope (lbt). *Astronomy and Astrophysics Supplement Series*, 147:323–333, 2000.

[3] T.E. Bishop, S. Zanetti, and P. Favaro. Light field superresolution. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–9. IEEE, 2009.

[4] JY Bouguet. Pyramidal implementation of the lucas kanade feature tracker, intel corporation, microprocessor research labs, 2000.

[5] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[6] G. Farnebäck. Two-frame motion estimation based on polynomial expansion. *Image Analysis*, pages 363–370, 2003.

[7] Federico Frigerio. *3-Dimensional Surface Imaging Using Active Wavefront Sampling*. PhD thesis, Massachusetts Institute of Technology, 2006.

[8] T. Georgiev and C. Intwala. Light field camera design for integral view photography. Technical report, Citeseer, 2006.

[9] T. Georgiev, C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala. Spatio-angular resolution tradeoffs in integral photography. In *Eurographics Symposium on Rendering*, pages 263–272, 2006.

[10] Jiann-Ching Guey and Mark R Bell. Diversity waveform sets for delay-doppler imaging. *Information Theory, IEEE Transactions on*, 44(4):1504–1522, 1998.

[11] Gopal Harikumar and Yoram Bresler. Exact image deconvolution from multiple fir blurs. *Image Processing, IEEE Transactions on*, 8(6):846–862, 1999.

[12] Gopal Harikumar and Yoram Bresler. Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms. *Image Processing, IEEE Transactions on*, 8(2):202–219, 1999.

[13] Carl W Helstrom. Image restoration by the method of least squares. *JOSA*, 57(3):297–303, 1967.

[14] Erik FY Hom, Franck Marchis, Timothy K Lee, Sebastian Haase, David A Agard, and John W Sedat. Aida: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data. *JOSA A*, 24(6):1580–1600, 2007.

[15] R. Horstmeyer, G. Euliss, R. Athale, and M. Levoy. Flexible multimodal camera using a light field architecture. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–8. IEEE, 2009.

[16] Peter A Jansson. *Deconvolution of images and spectra*. DoverPublications. com, 2009.

[17] A. Levin, R. Fergus, F. Durand, and W.T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics (TOG)*, 26(3):70, 2007.

[18] A. Levin, W. Freeman, and F. Durand. Understanding camera trade-offs through a bayesian analysis of light field projections. *Computer Vision–ECCV 2008*, pages 88–101, 2008.

[19] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz. Light field microscopy. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 924–934. ACM, 2006.

[20] M. Levoy, Z. Zhang, and I. McDowall. Recording and controlling the 4d light field in a microscope using microlens arrays. *Journal of Microscopy*, 235(2):144–162, 2009.

[21] C.K. Liang, T.H. Lin, B.Y. Wong, C. Liu, and H.H. Chen. Programmable aperture photography: multiplexed light field acquisition. In *ACM Transactions on Graphics (TOG)*, volume 27, page 55. ACM, 2008.

[22] C.K. Liang, G. Liu, and H.H. Chen. Light field acquisition using programmable aperture camera. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 5, pages V–233. IEEE, 2007.

[23] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[24] T.B. Milnes and D.P. Hart. Actively addressable aperture light field camera, 2011. US Patent Application 12/952,670.

[25] Laurent M Mugnier, Thierry Fusco, and Jean-Marc Conan. Mistral: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *JOSA A*, 21(10):1841–1854, 2004.

[26] Laurent M Mugnier, Clélia Robert, Jean-Marc Conan, Vincent Michau, and Sélim Salem. Myopic deconvolution from wave-front sensing. *JOSA A*, 18(4):862–872, 2001.

[27] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2, 2005.

[28] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, 2003.

[29] R. Raskar and A.K. Agrawal. 4d light field cameras, September 7 2010. US Patent 7,792,423.

[30] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 795–804. ACM, 2006.

[31] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519–528. IEEE, 2006.

[32] M. Tao, J. Bai, P. Kohli, and S. Paris. Simpleflow: A non-iterative, sublinear optical flow algorithm. In *Computer Graphics Forum*, volume 31, pages 345–353. Wiley Online Library, 2012.

[33] AH Techet, BE Scharfman, TB Milnes, and DP Hart. Light field imaging of fuel droplets and sprays. In *16th Int. Symp on Appl. Laser Techniques to Fluid Mechanics*, 2012.

[34] JD Van Ouwerkerk. Image super-resolution survey. *Image and Vision Computing*, 24(10):1039–1052, 2006.

[35] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Coded aperture and optical heterodyning: A mask-based approach for digital refocusing and light-field acquisition by conventional cameras.

[36] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Transactions on Graphics*, 26(3):69, 2007.

[37] Wikipedia. Depth of field — Wikipedia, the free encyclopedia, 2012. [Online; accessed 28-September-2012].

[38] Wikipedia. 1951 usaf resolution test chart — wikipedia, the free encyclopedia, 2013. [Online; accessed 21-July-2013].

[39] Wikipedia. Light field — wikipedia, the free encyclopedia, 2013. [Online; accessed 24-July-2013].

[40] Wikipedia. Point spread function — wikipedia, the free encyclopedia, 2013. [Online; accessed 20-July-2013].

[41] B. Wilburn, N. Joshi, V. Vaish, M. Levoy, and M. Horowitz. 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*, volume 2, pages II–294. IEEE, 2004.

[42] B. Wilburn, N. Joshi, V. Vaish, E.V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*, 24(3):765–776, 2005.

[43] B.S. Wilburn, M. Smulski, H.H.K. Lee, and M.A. Horowitz. Light field video camera. In *Electronic Imaging 2002*, pages 29–36. International Society for Optics and Photonics, 2001.

[44] J. Yang. *A light field camera for image based rendering*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2000.

[45] J.C. Yang, M. Everett, C. Buehler, and L. McMillan. A real-time distributed light field camera. In *Proceedings of the 13th Eurographics workshop on Rendering*, pages 77–86. Eurographics Association, 2002.

[46] Jianchao Yang and Thomas Huang. Image super-resolution: Historical overview and future challenges. *Super-resolution imaging*, 2010.

[47] Leonid P Yaroslavsky and H John Caulfield. Deconvolution of multiple images of the same object. *Applied optics*, 33(11):2157–2162, 1994.

[48] Leonid P Yaroslavsky and LP Yaroslavskij. Digital picture processing. an introduction. *Digital picture processing. An introduction.. LP Yaroslavsky (LP Yaroslavskij). Springer Seriesin Information Sciences, Vol. 9. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo. 12+ 276 pp. Price DM 112.00 (1985). ISBN 3-540-11934-5 (FR Germany), ISBN 0-387-11934-5 (USA).*, 1, 1985.