A Computational Light Field Display for Correcting Visual Aberrations



Fu-Chung Huang

Electrical Engineering and Computer Sciences University of California at Berkeley

Technical Report No. UCB/EECS-2013-206 http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-206.html

December 15, 2013

Copyright © 2013, by the author(s). All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

A Computational Light Field Display for Correcting Visual Aberrations

by

Fu-Chung Huang

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Brian Barsky, Chair Professor Carlo Sequin Professor Ravi Ramamoorthi Professor Austin Roorda

Fall 2013

A Computational Light Field Display for Correcting Visual Aberrations

Copyright 2013 by

Fu-Chung Huang

Abstract

A Computational Light Field Display for Correcting Visual Aberrations

by

Fu-Chung Huang

Doctor of Philosophy in Computer Science

University of California, Berkeley

Professor Brian Barsky, Chair

Vision problems such as near-sightedness, far-sightedness, as well as others, are due to optical aberrations in the human eye. These conditions are prevalent, and the population is growing rapidly. Correcting optical aberrations is traditionally done optically using eyeglasses, contact lenses, or refractive surgeries; these are sometime not convenient or not always available to everyone. Furthermore, higher order aberrations are not correctable with eyeglasses.

In this work, we introduce a new computation based aberration-correcting light field display: by incorporating the persons own optical aberration into the computation, we alter the content shown on the display, such that he or she will be able to see it in sharp focus without wearing eyewear. We analyze the image formation models; through the retinal light field projection, we find it is possible to compensate for the optical blurring on the target image by prefiltering with the inverse blur.

Using off-the-shelf components, we built a light field display prototype that supports our desired inverse light field prefiltering. The results show a significant contrast improvement and resolution enhancement over prior approaches. Finally, we also demonstrate the capability to correct for higher order aberrations. To my wife Ia-Ju Chiang and our parents.

Contents

Contents						
Li	List of Figures					
Li	List of Tables					
1	Intro 1.1 1.2 1.3	oductionBackgroundThe GoalOverview of Thesis Organization	1 2 5 7			
2	Rela 2.1 2.2 2.3	Ited WorkImage Restoration and Computational PhotographyDisplaysComputational Ophthalmology	8 8 9 11			
3	Mat 3.1 3.2 3.3	hematical Models of Optical BlurringImage Blurring as ConvolutionIntegrating Retinal Light FieldSummary	13 13 19 22			
4	Prel: 4.1 4.2 4.3 4.4	iminary Ideas with DeconvolutionFrequency Domain SolversSpatial Domain SolversTheoretical AnalysisSummary	 23 23 27 31 33 			
5	Mul 5.1 5.2 5.3 5.4 5.5	tilayer Displays Frequency Preservation via Multilayer Prefiltering Partition Function and Contrast Optimization Requirement and Construction Results Evaluations	34 34 37 41 45 50			

	5.6	Discussion	53			
6	Ligh 6.1 6.2 6.3	t Field Analysis Light Field Displays	54 54 60 63			
7	Ligh 7.1 7.2 7.3 7.4	t Field Based Inverse Prefiltering Analysis of Projection Matrix Prototype, Experiments and Results Evaluation Summary	64 65 69 74 79			
8	Ligh 8.1 8.2 8.3	t Field Wavefront Aberrometry Frequency Decomposition of Wavefront Geometry	80 80 83 86			
9	Futu 9.1 9.2 9.3	Ine Work Practicality.	93 93 95 96			
10	Con	clusion	97			
Bił	Bibliography					

iii

List of Figures

1.1	The Nimrud lens and early man-made reading stones.	3
1.2	The early eyeglasses (left) and bifocals (right)	3
1.3	Contact lenses.	4
1.4	Vision-correcting display.	5
3.1	Geometry of the optical setup for the eye.	13
3.2	Focal range of emmetropic, myopic, and hyperopic eye	15
3.3	Example wavefront surfaces.	16
3.4	Hartmann-Shack wavefront aberrometer.	17
3.5	Examples of non-planar wavefront generated point spread functions	18
3.6	Transformations of the light field.	20
3.7	Defocus light field.	21
4.1	Disk point spread function and its inverse.	24
4.2	Frequency domain prefiltering solutions using the inverse filter and the Wiener filter.	26
4.3	Spatial domain prefiltering solutions using Richardson-Lucy method	28
4.4	Prefiltering solution obtained using Levin et al. deconvolution.	29
4.5	Prefiltering using super-resolution technique.	30
4.6	Convolution weight matrices and their singular values.	31
4.7	Modulation transfer function for the disk PSF	32
5.1	Observation of multilayer PSFs and their MTFs	34
5.2	Winner-Take-All MTF and frequency assignment	37
5.3	Winner-Take-All binary frequency partitioning.	38
5.4	Contrast optimization via greedy partition function.	40
5.5	Comparing the perceived images from prefiltering algorithms using the eye chart ex-	
	ample	40
5.6	Prototype multilayer display, front view and side view	43
5.7	Comparing simulated single-layer prefiltering results with multilayer prefiltering results.	45
5.8	Comparing simulated results when brightness are equalized.	46
5.9	Camera photographs of prefiltering results.	47
5.10	Dynamic range variations of prefiltering	48

5.11 5.12 5.13 5.14	Dynamic range normalization in video example	49 50 51 52
6.1 6.2	Construction of light field displays	55 56
6.3	Depth of field analysis of light field displays.	58
6.4	Example of vision-correcting algorithms using light field displays	59
6.5	Light field projection in the frequency domain	61
7.1	Light field prefiltering with naive parameter setup.	65
7.2	Conditioning analysis.	66
7.3	Tradeoff between angular light field resolution and image contrast	67
7.4	Prototype parallax barrier light field display.	69
7.5	Optical experiment setup.	70
7.6	Photographs of results from prototype display.	71
7.7	Video examples comparing light field prefiltering.	73
/.8	Evaluation and comparison to previous work.	15
7.9	Comparing with preditering and light field based framework over a range.	/0 77
7.10	Accounting for a range of viewing distances.	// 70
/.11		/0
8.1	Wyant's ordering of Zerike polynomial.	81
8.2	ANSI standard Zernike ordering.	82
8.3	Ray-tracing wavefront	83
8.4	Derivatives	84
8.5	Derivative of Zernike polynomials	85
8.6	Optical arrangement to illustrate astigmatism.	86
8./	Examples showing astigmatism has depth-dependent blur	8/
8.8 8.0	Correcting for astigmatism.	88
0.9 8 10	Examples of single term Zernike polynomial point spread functions.	09
0.10 <u>8</u> 11	Physical setup to generate higher order aberrations	90
0.11	i hysical setup to generate inglier order aberrations	94

List of Tables

Acknowledgments

I want to thank my thesis advisor Prof. Brian Barsky for giving me freedom to explore many fun problems. Prof. Ramesh Raskar is like my second advisor who supported me in many ways, that I cannot complete my thesis without his help. Prof. Ravi Ramamoorthi not only served as my master thesis advisor and doctoral thesis committee, but he is also my role model as a researcher; I learned so much from him.

I also want thank my committee members. Prof. Carlo Sequin gave me so many suggestions in the final stage of thesis writing and how to building a career. Prof. Austin Roorda taught me my first class in vision, and has been an incredible source of knowledge to consult to.

People from MIT Media Lab helped me accomplish many experiments. Dr. Douglas Lanman and Dr. Gordon Witzstein are both my collaborators and mentors, and I owe them so much.

I also want to acknowledge people who encouraged me when research was not smooth. Dr. Manmohan Chandraker, Dr. Huamin Wang, Dr. Milos Hasan, and Prof. James O'Brien were my support. I'm also fortunate to meet with many other people: Prof. Laura Waller, Dr. Ramkumar Sabesan, Dr. Kar-Han Tan, and Dr. Sameera Poduri. I also want to thank my undergraduate advisors Bing-Yu Chen and Yung-Yu Chuang for leading me to the fun journey of research in computer graphics.

I had a happy time at UC Berkeley graphics group (Visual Computing Lab) with Jiamin, Michael, Florian, Armin, Soham, Jimmy, Sean, Sally, Woojong, Juan, Grace, Brandon, Eno, and all my fellow students during my 7 times teaching CS184 class!

Chapter 1

Introduction

According to an NIH report [MedlinePlus, 2008], "Most Americans report that, of all disabilities, loss of eyesight would have the greatest impact on their daily life, according to a recent survey by the NIH National Eye Institute (NEI). Vision loss ranks ahead of loss of memory, speech, arm or leg, and hearing. After all, 80 percent of the sensory information the brain receives comes from our eyes." Among the five sense organs, the eyes are probably the most important one; it is the most vital for our sensory inputs and memories. Learning, recognizing faces, perceiving colors, and appreciating the world, all depends on our vision. Among all vision related problems, the most common are refractive errors, such as myopia (nearsightedness), hyperopia (farsightedness), presbyopia, and astigmatism. These cause blurred vision, which is due to the shape of the eye preventing light from focusing precisely on the retina.

To see the world, light goes into the eye, refracted by a sequence of optical elements such as the cornea and crystalline lens. The light then lands on the retina and is converted to the electrical signal to the brain. When the refractive elements are not perfect, blurred images are fromed and decrease the quality of life. Although we already have eyeglasses, the technologies are improving and newer instruments are invented to rectify blurry vision, such as contact lenses or refractive surgeries. However, these are purely optical operation directly modifying the optics of the eye.

In this thesis, we present the first practical solution to "digitally" modify the content on a display device, that also corrects for blurred vision. We start by understanding the blurring process of the eye, and inversely preprocess the image content, such that the final perception is still sharp. This cannot be achieved without changing the display optics at a hardware level; we not only show the algorithmic operations but also build a hardware prototype using off-the-shelf components. We not only correct for myopia or hyperopia, more complicated blur induced by higher order aberrations are also considered.

This thesis not only presents solutions to correcting optical aberrations, but it is also the first to discuss different models of image blurring process under the same framework using light fields. The theoretical treatment on the image formation process provides insight to the problem of preprocessing based vision-correction, and directions for possible future improvements.

1.1 Background

1.1.1 Visual aberrations.

Optical aberrations are common in almost all optical systems, including the human eye. Global surveys estimate that 153 million people are visually impaired as a result of uncorrected refractive errors [Resnikoff *et al.*, 2008]. According to the World Health Organization (WHO), 246 million people have low vision (below 20/60), and 43% of them are due to simple uncorrected refractive errors such as myopia, hyperopia, or astigmatism.

Hyperopic vision prevents the observer from focusing at near range, and is common in the United States. About one quarter of population in the U.S. has hyperopia [Krachmer *et al.*, 2005]; most are mild until in their late 40s and early 50s where the accommodative ability decreases and presbyopia begins. Hyperopic vision is common for people over the age of 40, and its affected population, about 43%, is twice as large as that of myopia; in the Baltimore eye survey [Katz *et al.*, 1997], the population of hyperopia is over 68% for people older than 80. This is not surprising, since the accommodation ability of the eye decreases about one diopter every ten years, thus leading to presbyopia.

Recent studies also indicate an increasing prevalence of myopia; Vitale et al. [2009] found the incidence of myopia increased from 25% to 41.6% in the United States between 1971-1972 and 1999-2004, and some studies also show that the myopia increases with education level [Katz *et al.*, 1997]. The prevalence of myopia is increasing rapidly, especially in Asian populations [Rajan *et al.*, 1995]. It is especially relevant since the prevalence of myopia among young people, which varies by country and ethnic group, is approaching 60% to 90% in some Asian populations [Lin *et al.*, 2004][Wong *et al.*, 2000][Takashima *et al.*, 2001].

In addition to myopia, hyperopia and astigmatism due to lower order aberrations, there are also higher order terms due to spherical aberrations or irregularly shaped lenses. An early study shows that on average, most people have a non-zero amount of spherical aberration, with some non-zero variation in the higher order terms [Thibos *et al.*, 2002]; together with a recent large clinical study [Hartwig and Atchison, 2012] shows that, statistically, spherical aberrations and coma are dependent on the lower order terms, and higher order aberrations systematically happen in the human population [Bao *et al.*, 2009]. In the past decade, the advance in the measurement of wavefront geometry of the eye helps people understand visual acuity better; however, to date the correction for higher order terms is still experimental.



Figure 1.1: The Nimrud lens (left) and early man-made reading stones (center and right).

1.1.2 Correcting for Vision-Aberrations.

Lens. The history of correcting for visual aberrations could be traced back thousands of years¹. The first commercial man-made lens was the reading stone (Fig. 1.1 (center and right)), around 9 A.D., which was discovered during the process of making glass from sand by Abbas Ibn; the behavior of the convex shaped lens was first studied around 1021 A.D. in the "book of optics" by Alhazen (965 to 1038 A.D.), the father of modern optics.

Eyeglasses. An Italian named Salvino D'Armato started making wearable eyeglasses (Fig. 1.2 lower left) in 1284 A.D., and this led to the modern development of correcting for optical aberrations. During the middle of 15th century in Florence, the city that was the center for making eyeglasses, the convex lens for presbyopia and the concave lens for myopia were being mass-produced. At the end of 15th century, eyeglasses were a common commodity sold on the street. Later, around 1760, Benjamin Franklin invent the first bifocal lens (Fig. 1.2 lower right); and by 1825, George Airy was the first to design the concave astigmatism lens.

¹Debatable: the first lens ever found is the Nimrud lens (left of Fig. 1.1), around 934 B.C.; according Professor Giovanni Pettinato, Assyrian uses this to observer astronomical behavior. Another possible first lens is found on the Mountain Ida on Crete around 5th century B.C..



Figure 1.2: The early eyeglasses (left) and bifocals (right).



Figure 1.3: **Contact lenses.** Early contact lenses (left two frames) cover the entire ocular surface. Most modern contact lenses softly cover only the cornea and are sometimes also used for cosmetic purposes.

Contact Lens. Contact lenses also have a long history, dating back to the book "Codex of the eye" by Leonardo DaVinci around 1508, where he had the idea to modify the power of the cornea. DaVinci's idea was first implemented by René Descartes using a water-filled tube in 1636; how-ever, blinking is not possible due to the physical size of the prototype; the first working prototype was made by Thomas Young in 1801.

In early 1880's, the first glass contact lens that fit the anterior of the eye was made by Adolph Eugen Fick, Eugene Kalt, and August Mueller, independently. The lens was not convenient for use until 1949, when the first plastic contact lens was produced, and people could wear for many hours; since then, work has been directed toward better commercial products, e.g. improved oxygen-permeability rate, disposable contact lens, and even cosmetics contact lens (right of Fig. 1.3).

Refractive Surgery. Refractive surgery is a relatively new operation that directly modifies the optical power of the eye; the principle is to alter the shape of the cornea. The first idea was proposed by Lendeer Jans Lans in 1896 and was performed by Tsutomu Sato in 1930 by making cuts on the cornea, but the result was unsuccessful. Although there are also implant based surgeries, the most successful operation uses an excimer laser to reshape the cornea.

The early form of modern refractive surgery, called keratomileusis, was developed by Jose Barraquer in 1963, that removes the cornea layer and sculpts it into the desired shape. In 1983, Stephen Trokel and Rangaswamy Srinivasan performed the first PhotoRefractive Keratectomy (PRK) [Kugler and Wang, 2010] using excimer laser; later it became commonly known as Laser-Assisted in situ Keratomileusis (LASIK).



Figure 1.4: **Vision-correcting display.** We present the first practical-vision correcting display that does not require that observer wear eyeglasses. Our final light field based display (bottom right) enables both high resolution and high contrast, achieving a thin form factor and light-weight construction.

1.2 The Goal

Although the previous three options to correct optical aberrations are effective most of the time, they can sometimes be awkward. For example, going to a theater to watch 3D movies requires wearing double eyeglasses. Using a Head Mount Displays (HMD) to explore virtual reality requires wearing eyeglasses inside a tight device. For people with presbyopia, watching mobile devices requires another pair of glasses or bifocals. Although some of the problems can be solved via contact lenses, it is not always convenient, especially for a extended long period of time. Refractive surgery is not always an option for some people with a thinner cornea or for other reasons, and can also cause side effects. Finally, higher order aberrations are sometimes a relevant concern for some populations, and none of the above alternatives can correct the inherent blurriness.

Our goal in this work is to provide a fourth practical option that corrects optical aberration for the eye. In particular, as opposed to optically modifying the refractive power of the eye, we introduce a computational display that "presents the pre-distorted content according to the viewer's prescription, such that a sharp image is formed on the retina, without the need to modify the physical optics when the prescription changes." As opposed to the previous methods directly modifying the eye, the computational displays have no direct contact to the observer, and can be made from off-the-shelf hardware components. Once the display device is built, the refractive errors are corrected digitally; no further adjustment to the optical hardware component is required.

Figure 1.4 shows canonical cases for our computational displays; for the user having myopic eyes, he or she sees the blurred image when there is no correction. The multilayer method (Chapter 5) gives a sharp perceived image by digitally distorting the displayed images, but the contrast is low. Pamplona et al. [2012] use a light field display (Chapter 6) such that a sharp virtual image is shown within the focal range of the observer, but the spatial resolution of the display is low. Finally at the end of the thesis (Chapter 7), we introduce an integrated framework, called light field pre-filtering, that combines the advantages of the previous two methods. By designing the algorithm in concert with the optics of the device, light field prefiltering is able to provide both high resolution and high image contrast.

Contributions. The new framework involves understanding image formation by projecting light rays onto the retina. The frequency domain analysis of the light field gives insights into why the naive inversion of the point spread function on 2D images does not work, and higher dimensional methods are required, e.g. a multilayer display involves a stack of 2D images; a light field based methods require 4D light rays. We propose the conditioning rank analysis, which we will show is equivalent to the modulation transfer function (MTF) zeros analysis. The rank analysis reveals the design parameters of the hardware, and lightly modified hardware provides enough degrees of freedom to achieve the goal of correcting vision. Physical prototype hardwares are built to illustrate these benefits, and the light field prefiltering prototype using an iPhone 4 or iPod touch 4 corrects significant defocus blur. Finally, we also show that correcting for higher order aberrations is made possible using this technique.

Impacts, applications, and the future. The solution to the problems of our goal has a lot of potential, not just for correcting myopia or presbyopia. Correcting for higher order aberrations has an important impact to some populations, especially since currently there is no practical solution; we envision that, in the future, normal eyeglasses could correct the lower order aberrations, while the display will correct the higher order ones.

This is also the first work to show that focusing at more than one plane is possible; while the eye has its natural focus plane, the device plane is also capable of generating a sharp image on the retina or the camera sensor. This is useful for heads-up-display (HUD), like projecting useful information onto cars or objects when the windshield or Google glass is close to the observer.

The near-field display is also useful in some real world application, where wearing double glasses are cumbersome; the vision correcting display does two things in one device. There are some potential future directions that might give rise to interesting research, such as the vergences and the convergences problem of the eye when accommodating 3D content.

1.3 Overview of Thesis Organization

In Chapter 2, we describe prior work and fundamental building blocks of the theory. Chapter 3 reviews important background on image formation, especially the geometries for the imaging process of the eye; both 2D imaging and 4D light fields projection are considered. In Chapter 4, we first solve the problem with existing tools and methods, and then we show the fundamental limitations with them.

In Chapter 5, the first theoretical improvement over existing tools is presented using multilayer displays. In fact, the multilayer display is a subclass of the general light field displays; we discuss the generalization in Chapter 6. Chapter 7 presents another solution to the problem using 4D light field displays with significant contrast enhancement over that using multilayer displays.

Finally the measured prescription using a wavefront aberrometer replaces the thin lens model in Chapter 8; the extension to solve for higher order aberrations is shown in Chapter 8.3. We conclude the methodology to the problem and discuss potential future applications in Chapter 10.

Chapter 2

Related Work

2.1 Image Restoration and Computational Photography

Deblurring. Our vision-correcting idea is similar in principle to image deblurring with different constraints, and thus different difficulties. Image deblurring is applied to estimate an undistorted image given a received image degraded by camera shake, defocus, or object motion. The perceived image may be modeled as the convolution of the undistorted image by the optical point spread function (PSF), characterizing the degradation process. Deconvolution algorithms can be applied to approximate the undistorted image, including inverse filtering [Gonzalez and Woods, 1992], Wiener filtering [Wiener, 1964], and the iterative Richardson-Lucy algorithm [Richardson, 1972] [Lucy, 1974]. Recent developments in image deconvolution include exploiting natural image priors [Levin *et al.*, 2007], suppression of ringing artifacts [Shan *et al.*, 2008] [Yuan *et al.*, 2008], and increasingly focus on blind deconvolution [Campisi and Egiazarian, 2007], wherein the PSF is not known *a priori*.

Correcting Projector Defocus. Projecting onto an out-of-focus plane or a tilted wall introduces defocus blur, and it can be compensated by deconvolving the projected image with the projector's PSF. Brown et al. [2006] demonstrate extended depth of field projection using deconvolution. Inverse prefiltering sometimes introduces values outside the dynamic range of the projector, Oyamada et al. [2007] compare the performance of clipping values outside the dynamic range to normalization. Zhang and Nayar [2006] propose solving a constrained optimization problem to minimize artifacts while utilizing only the available dynamic range. Although this work considers unmodified projector optics, typically containing circular apertures, Grosse et al. [2010] introduce an adaptive coded aperture to ensure that the modulation transfer function (MTF), corresponding to the magnitude of the Fourier transform of the PSF, preserves all relevant spatial frequencies. In Chapter 5, we similarly seek to produce an *all-pass filter* by introducing a second display layer for correcting optical aberrations.

All-Pass Filters. Recent work in computational photography has also explored the notion of constructing all-pass optical filters, capable of preserving image information despite common distortions, including camera shake, defocus, or object motion. This work advocates modifying the optics or the capture process to synthesize an effective MTF that preserves all spatial frequencies. Raskar et al. [2006] rapidly modulate the aperture over the exposure to transform the PSF, due to motion blur, such that ringing artifacts are eliminated. Agrawal et al. [2009] capture two exposures, of slightly different durations, to accomplish a similar task. Veeraraghavan et al. [2007] introduce a coded aperture to create an all-pass MTF, allowing deconvolution algorithms to correct camera defocus without introducing ringing. Our development of multilayer prefiltering in Chapter 5 is inspired by these publications, with the goal of incorporating additional layers to ensure all spatial frequencies are preserved in the received image.

Light Field Analysis. Light fields have gained attention in computer graphics in the past two decades. Early pioneering work by Adelson and Bergen [1991], McMillian and Bishop [1995], Levoy and Hanrahan [1996], Gortler et al. [1996], Isaksen et al. [2000], and Chai et al. [2000] have focused on capture, sampling, and efficient rendering. Modern imaging uses a computational approach to capture 4D light fields. Ng [2005] demonstrates that 4D projection can be efficiently done by slicing a 2D plane in the frequency domain. Veeraraghavan et al. [2007] acquire the light field by inserting a coded plane inside the camera, and the idea is generalized by Liang et al. [2008] and Marwah et al. [2013], in a compressive sense, to capture the light field at full sensor resolution. Levin et al. [2009] show that the efficient use of the 3D manifold embedded in 4D by using a lattice focal lens can improve deconvolution quality, achieving extended depth of field. The rendering community also has begun to analyze the frequency domain property of light field for efficient sampling and rendering. Durand et al. [2005] analyze transport, occlusion, and shading effects. Kevin et al. [2009] reconstruct a sheared filter to efficiently sample the temporal domain for motion blur. Soft shadows from complex occluders are studied by Kevin et al. [2011], but the precomputation time outweighs the rendering time. Recently, Mehta et al. [2012] show axisaligned filters can be implemented for complex soft shadows to enable realtime rendering.

2.2 Displays

Glasses-Free 3D Displays. Glasses-free 3D was invented in the beginning of the 20th century. The two dominating technologies are *lenslet arrays* [Lippmann, 1908] and *parallax barriers* [Ives, 1903]. Today, a much wider range of different 3D display technologies are available, including volumetric displays by Cossairt et al. [2007] and by Jones et al. [2007]. Volumetric displays create the illusion of a virtual 3D object floating inside the physical device enclosure; an observer's eye can accommodate within this volume. Although these devices support binocular disparity and motion parallax, no evidence has been provided that a sufficiently high spatial and angular resolution is provided to support accommodation or, similarly, move a virtual 2D image outside the device and into the focal range of the observer.

CHAPTER 2. RELATED WORK

Multilayer Displays. Multilayer display is an emerging technology targeted towards autostereoscopic glasses-free 3D display. Such panels represent content on superimposed, semi-transparent layers, providing a faithful reproduction of perceptual depth cues. However, to achieve an extended range of depths, multiple panels must be distributed within a thick enclosure. To preserve thin form factors, research focuses on achieving the illusion of an extended volume with a compact device [Urey *et al.*, 2011]. Multilayer displays are one such family of autostereoscopic displays, divided into those that employ stacks of *light-emitting* versus *light-attenuating layers*.

For example, Akeley et al. [2004] place a series of beamsplitters at 45 degree angles with respect to a single LCD; viewed from in front of the stack, the eye perceives superimposed light-emitting layers. In contrast, Wetzstein et al. [2011] and Holroyd et al. [2011] consider thin stacks of light-attenuating films for synthesizing high dynamic range light fields and 3D scenes. Lanman et al. [2011] and Gotoda [2012] evaluate stacks of LCD panels; these publications describe a mode where the virtual scene extends beyond the display enclosure. With temporal multiplexing, the multilayer LCD can operate in a mode optically equivalent to a stack of layers that are light-emitting rather than light-attenuating. Multilayer LCDs have also found applications in computational photography [Zomet and Nayar, 2006].

Compressive Light Field Displays. Zwicker et al. [2006] studied the depth of field capability of the 3D display, and showed that angular sampling resolution is the limiting factor; some later research in 3D displays strives to increase the resolution limits. Most recently, the compressive nature is exploited not only in photography via compressed sensing, but also in the displayed content to enhance the angular resolution of the display. Lanman et al. [2010] decompose a 3D scene into a series of images shown on a multilayer display with high refresh rate. The decomposition is further studied using low rank approximation by Wetzstein et al. [2012], and better image quality is achieved by adding additional directional backlighting. By exploring the redundancy in the 3D scene, these recent efforts show significant improvements on spatial and angular resolution.

Adaptive Optics. The idea to compensate for the turbulence of the atmosphere to achieve better imaging system was first discussed by Babcock [1953], but was not made practical until the modern development of computer system [Liang *et al.*, 1997]. Optical aberrations are compensated by reflecting off a deformable mirror in astronomical telescopes or retinal imaging [Roorda *et al.*, 2002]. Artal et al. [2004] and Sabesan and Yoon [2010] showed patients adaptive optics compensated images to study neural compensations and adaptations to optical blur. When combined with high frequency eye tracking, adaptive optics corrected stimuli can be painted onto the desired retinal location with accuracy of a single cone photoreceptor [Yang *et al.*, 2010]. Adaptive optics provides an accurate control of the retinal stimulus that allows people to track the chromatic cone responses [Hofer *et al.*, 2005] and the retinal limits on the acuity of spatial vision [Rossi and Roorda, 2009]. A extensive historical survey of adaptive optics can be found in [Roorda, 2011].

2.3 Computational Ophthalmology

Diagnosis of refractive errors. The exact aberrations of the eye can be measured using a Shack-Hartmann wavefront sensing aberrometer [Shack and Platt, 1971] [Platt and Shack, 2001]; the essence is to detect the normal directions of the wavefront surface using a sensor mounted with a lenslet array, which is equivalent to a light field camera. Pamplona et al. [2010] proposed a novel cellphone add-on device, called NETRA, to diagnose refractive errors by inversely tracing rays from the screen to the eye; later, another device, called CATRA, was proposed to measure the cataract map [Pamplona *et al.*, 2011] in the human eye. These latter applications are examples of computational ophthalmology, where interactive techniques are combined with computational photography and display for medical applications.

Vision-Correcting Displays. Devices tailored to correct the visual aberrations of an observer have recently been introduced. Correcting the optical aberrations of the eye requires applying a preprocess, e.g. deconvolution, before the image is displayed, rather than after it is received by the visual system. This discrepancy significantly impacts the quality of received images. Early approaches attempt to prefilter a 2D image presented on a conventional screen by using the inverse point spread function (PSF) of the observer's eye; the preprocessed image typically includes both negative and positive values of equal amplitude. Following Alonso and Barreto [2003], Yellott and Yellott [2007], and Archand et al. [2011], preprocessed images must be normalized to the dynamic range of the display. Although these methods slightly improve image sharpness, contrast is reduced. Fundamentally, the PSF of an eye with refractive errors is a low-pass filter—high image frequencies are irreversibly canceled out in the optical path from display to the retina.

To overcome this limitation, Pamplona et al. [2012] proposed the use of conventional light field displays with lenslet arrays or parallax barriers to correct visual aberrations. For this application, these devices must provide a sufficiently high angular resolution so that multiple light rays emitted by a single lenslet enter the pupil. This resolution requirement is similar for light field displays supporting accommodation cues [Takaki, 2006]. Whereas these displays support accommodative depth cues in addition to binocular disparity and motion parallax, for the application of 3D image display, vision-correcting displays present a 2D image at a distance from the physical device on which the viewer can focus. Unfortunately, conventional light field displays as used by Pamplona et al. [2012] are subject to a spatio-angular resolution tradeoff: an increased angular resolution decreases the spatial resolution. Hence, the viewer sees a sharp image but at a significantly lower resolution than that of the screen, as shown in Figure 1.4.

Computational Vision-Correcting Displays. To mitigate the effects of low contrast or low spatial resolution, we will introduce two approaches using multilayer displays [Huang and Barsky, 2011][Huang *et al.*, 2012] and light field displays [Huang *et al.*, 2013] in Chapter 5 and 7. The approach developed in this thesis combines previous methods by employing 4D light field prefiltering

with hardware designs that have previously only been used in a "direct" way; that is, where each screen pixel corresponds to one emitted light ray.

We demonstrate that the new light field design enables significantly higher resolution as compared to the "direct" method because angular resolution demands are decreased. Ultimately, we will show that light field based inverse prefiltering is a promising, high-resolution approach that significantly increases the resulting image contrast and the perceived resolution.

Chapter 3

Mathematical Models of Optical Blurring

3.1 Image Blurring as Convolution

Image blurring is traditionally modeled as spatial filtering, or convolution [Gonzalez and Woods, 1992], of the image with a point spread function(PSF), which is the response of an imaging system to a point light source in the scene. There are some limitations with this approach: Lambertian surface assumption, linear spatially invariant (LSI) model, and single depth objects; at the end of this chapter, we will derive image formation as convolution using a light field.

With a simple formulation, various forms of the convolution model are commonly used, and the reader can refer to the book by Gonzalez [Gonzalez and Woods, 1992]. In this section, we will discuss two of models: using the circle of confusion and the wavefront map to represent the point spread functions.



Figure 3.1: When a hyperopic or presbyopic eye is focused farther, the closer object is blurred by the circle of confusion. The size of the circle is determined by the pupil size and the geometry. The traditional model for image blurring convolves the object with the point spread function, which is the circle of confusion here.

3.1.1 Circle of Confusion

For simplicity, we first introduce the traditional 2D geometrical imaging model using a ray tracing diagram, as shown in Figure 3.1. The distance d_f to the plane of focus can be determined by the thin lens equation:

$$\frac{1}{d_f} + \frac{1}{d_e} = \frac{1}{f},$$
(3.1)

where the eye has an internal diameter d_e , and the focal length f can be derived from the optical power of the eye. When the display is close to the hyperopic or presbyopic eye, at the distance d_o , the content is perceived as an image blurred by a circle of confusion, which depends on the pupil diameter a, and the diameter r of the circle is given by $r = a|d_f - d_o|/d_f$. The retinal converging point receives the superposition of all objects within the circle, and thus the perceived image is blurred.

To model the blurring process, the displayed image i is convolved with the circle of confusion, and the object space blurred image i^b is given by

$$i^{b}(x,y) = i(x,y) \ast k(x,y) \quad \iff \quad i^{b}(\omega_{x},\omega_{y}) = \widehat{i}(\omega_{x},\omega_{y}) \cdot \widehat{k}(\omega_{x},\omega_{y})$$
(3.2)

where k is the kernel point spread function defined by the circle of confusion, * is a convolution operator, $\hat{\cdot}$ denotes the Fourier transformed variables, and Equation 3.2 is the result of the convolution theorem. Note that this models the blurring process in the object space; the retinal space formulation can also be derived similarly by scaling.

Diopter and Focal Length. In Equation 3.1, the focal length f determines the diameter of the circle, and f is given by the optical power of the eye. Consider the diagram of the eye in Figure 3.2. A relaxed normal eye (a) has a power of 60 diopters to focus parallel rays, which yields a focal length f = 1000/60 = 16.67, and the internal diameter d_e of the eye is about 17mm¹. The crystalline lens provides up to 8 additional diopters on demand for the eye to focus at near objects (b); in this case the focal length f = 1000/68 = 14.7, and by Equation 3.1 the nearest plane of focus is 125mm. The diagram of the focal range is depicted in Figure 3.2(e).

Myopia. For a myopic eye, as shown in Figure 3.2(c), the lens has optical power more than 60 diopters in the relaxed state, so the observer cannot focus on an object at infinity; the far plane at the relaxed state is drawn closer to the eye. The focal range of the myopic eye is shown in Figure 3.2(f); the overpowered lens pulls the entire range closer to the observer, and negative powered corrections, such as with eyeglasses, are required to compensate for this effect.

¹Note that actual diameter is about 22mm to 24mm, but the internal liquid refracts the light more than that in air. http://www.blackwelleyesight.com/eye-math/reduced-schematic/



Figure 3.2: Focal range of emmetropic, myopic, and hyperopic eye. The lens has optical power between +60 to +68 diopters to focus from infinity to 125mm objects. The myopic eye is over-powered and the focal range is pulled closer to the observer; the hyperopic eye is underpowered and cannot focus at near range.

Hyperopia. On the other hand, a hyperopic eye (Fig. 3.2(d)) has a lens of power less than +60 diopters when relaxed, even with the addition of +8 diopter the observer still cannot focus at the normal near range. Figure 3.2(g) illustrates the focal range of the underpowered eye, and positive powered corrections, such as eyeglass, are need. Similar to hyperopia, the crystalline lens of a presbyopic eye loses its ability to provide additional focusing power.

3.1.2 Measured Wavefront Map

The previous section uses the circle of confusion to describe the point spread function; however, this is sometimes insufficient to model the aberrations that are caused by imperfections in the optical lens, such as astigmatism, spherical aberrations, or aberrations from even higher order terms. In this section, more general point spread functions that are based on measured data are derived.

Light is an electromagnetic wave and can be expressed in the form of $E(x) = A(x) \cdot e^{i\phi(x)}$, where the function $A(\cdot)$ describes the amplitude, and the function $\phi(\cdot)$ describes the phase. For the 2D extension to capture the light propagating into the eye, we first express the light as:

$$E(x^a, y^a) = A(x^a, y^a) \cdot e^{-i\frac{2\pi}{\lambda}\omega(x^a, y^a)}$$
(3.3)

where $A(x^a, y^a)$ is the 2-D amplitude function defined on the pupil aperture plane, and is shaped by the pupil; $\omega(x^a, y^a)$ is the phase wavefront map, which can be analytically decomposed into a series of commonly known terms, such as defocus, astigmatism, spherical aberrations, and so on.

Wavefront. The wavefront is the surface perpendicular to all rays passing through it at the same phase; Figure 3.3 shows 3 scenarios. The wavefront characterizes the directions of rays leaving an optical system, and hence the optical aberrations of the lens can be described. Since the wavefront gives a compact representation of the ray directions, it is useful to measure the wavefront using the Hartmann-Shack wavefront sensor for the eye.

Hartmann-Shack Aberrometer. By starting from a point light source at the focal point (e.g. by shooting a laser beam into the relaxed eye, it will be scattered by the retina, as illustrated in Figure 3.4(a)), we measure the wavefront leaving the eye, and the configuration is called the Hartmann-Shack wavefront aberrometer. At the pupil plane, the wavefront encodes the directions of the rays leaving the eye, which is also the phase function. Since the directions of light rays is given by the normals of the wavefront surface, a pinhole array or a lenslet array can be used to measure the wavefront surface, as shown in Figure 3.4(b) and (c).



Figure 3.3: Wavefront coming out of the eye: (a) the planar wavefront where all rays are parallel and in-phase, (b) the spherical wavefront where rays are converging to a point, and (c) the wavefront where rays are pointing irregularly, resulting a irregular perpendicular surface.



Figure 3.4: (a): Measuring wavefront using Hartmann-Shack wavefront sensor. After registration with a planar wavefront (b), the irregular wavefront generates offsets from the registration points (c), and can be used to derive the original function for the wavefront.

We first measure the planar wavefront (b), where the converging points after the lenslet array give the registration. We then capture deviation from the registration of the converging points for an irregular wavefront (c), and then the wavefront surface is fit with a discrete set of measurement. The sparse fitting process involves the projection onto a finite series of frequency basis functions, called the Zernike polynomials. With sufficient observations, we can fit a smooth surface function $z = \omega(x^a, y^a) = \sum_j Z^j(x^a, y^a)$, and this is the principle of the Hartmann-Shack wavefront aberrometer. We will formally describe the basis functions, $Z^j(x^a, y^a)$, in Chapter 8. With reciprocity, the measured outgoing wavefront can also be used to describe how parallel coherent light rays entering the eye change their directions, and we can find the point spread functions from that information.

Generating Point Spread Function. To infer the convergence of the light after the lens, it is useful to consider the lens acting as a complex optical Fourier transform [Goodman, 2004]. Using Equation 3.3 to express the light before the optical transform, the retinal cone cells perceive the converged rays as:

$$k(x,y) = \left| \int \int A(x^a, y^a) \cdot e^{-i\frac{2\pi}{\lambda}\omega(x^a, y^a)} \cdot e^{-i(xx^a + yy^a)} dx dy \right|^2.$$
(3.4)

CHAPTER 3. MATHEMATICAL MODELS OF OPTICAL BLURRING

For a perfect optical system with a planar wavefront where $\omega(x^a, y^a) = 0$, the converged retinal light rays produce a PSF that is an Airy disk function at the origin. A non-planar wavefront forms more general point spread functions. Some examples are given in Figure 3.5; each is generated from a single-term Zernike polynomial fitted wavefront map, and in Chapter 8 we will further discuss the decomposition and how to utilize a wavefront map to solve our problem.



Figure 3.5: Examples of non-planar wavefront generated point spread functions

It is worth noting that Equation 3.4 generates point spread functions with coherent light sources and Airy disk; these are in the regime of wave optics. In Chapter 8, we will generate point spread functions with geometric optics, i.e. with classical ray tracing geometry. In practice, the point spread functions will be slightly different, especially the ripples and alternating bands due to wave interferences are not apparent.

3.2 Integrating Retinal Light Field

The previous sections deals with image blurring by convolving the target image with a point spread function. However, the light field is a more general tool to model the overall process. Similar to the wave-based formulation (Eq. 3.3), the light field also models the intensities and directions of the light rays, but under geometric optics². The most general light field, also known as the plenoptic function, is a high dimensional field that characterizes the rays of light. For simplicity, most techniques ignore wavelength and temporal variations, and assume constancy in free space; this results in a 4 dimensional function involving positions (2D) and directions (2D).

Following the "relative" parameterization by Chai et al. [2000] using two planes, we represent the starting light field with the plane for spatial coordinates (x^d, y^d) on the display, and the 1-unit separated plane (u^d, v^d) for angular parameterization. For simplicity, we derive the formulation using a 1D flatland light field. The extension to full 4D is straightforward, and can be found in [Durand *et al.*, 2005] and [Liang *et al.*, 2011]; the optical setup is shown in the top of Fig. 3.6.

Light Field Imaging and Transforms. The image irradiance i on the retina can be expressed as an integral of the retinal light field radiances l_e inside the eye as:

$$i(x) = \int_{-\infty}^{\infty} l_e(x, u) A(u) du.$$
(3.5)

We model a finite pupil of diameter r as a multiplication with the aperture function A(u), which is modeled as a *rect* function: $A(u) = rect(\frac{u}{r})$. A straightforward extension to model the circular pupil in a 4D light field is to have the aperture as a disk function A(u, v) = 1 where $\sqrt{u^2 + v^2} \le \frac{r}{2}$.

To obtain the retinal light field inside the eye, the display light field $l(x^d, u^d)$ first propagates a distance d_o to the pupil and is then refracted by the lens of focal length f; we illustrate the ray tracing diagram and light field transforms in the bottom of Figure 3.6. Inside the eye, the light field propagates another distance d_e to the retina, and the radiances are finally integrated on the retinal point, as indicated by the yellow arrow and the box.

The propagation and the refraction can be represented with linear transforms³ T and R (Fig. 3.6 bottom) on the light field, and it is sometimes useful to re-parameterize the light field inside the eye with another operator Q such that the angular plane lies on the pupil.

Note the pupil aperture is located right after the first propagation, as the top of Figure 3.6 shows. It initially blocks the light field in the spatial coordinates; for derivation simplicity (Eq. 3.5), we include it at the "final stage" of the retina light field; it can be shown that after the second propagation and re-parameterization, the aperture blocks the angular coordinates.

²extension to wave optics can be found in [Oh et al., 2010]

³Gaussian 1st order approximation, or paraxial approximation



Figure 3.6: **Top:** Transformations of the light field. **Bottom:** Ray-tracing diagrams, corresponding transformations on the light field, and the linear operators.

The composite transform M that relates the display light field $[x^d, u^d]^T$ to the retinal light field $[x, u]^T$ is:

$$\begin{pmatrix} x \\ u \end{pmatrix} = \mathbf{Q}(d_e) \quad \mathbf{T}(d_e) \quad \mathbf{R}(f) \quad \mathbf{T}(d_o) \begin{pmatrix} x^d \\ u^d \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 1 & -d_e \end{pmatrix} \begin{pmatrix} 1 & d_e \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d_o \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^d \\ u^d \end{pmatrix}$$

$$= \begin{pmatrix} 1 - \frac{d_e}{f} & d_o d_e \Delta \\ 1 & d_o \end{pmatrix} \begin{pmatrix} x^d \\ u^d \end{pmatrix} = \mathbf{M} \begin{pmatrix} x^d \\ u^d \end{pmatrix}$$
(3.6)

where $\Delta = \frac{1}{d_o} + \frac{1}{d_e} - \frac{1}{f}$, and similarly the inverse transform:

$$\begin{pmatrix} x^{d} \\ u^{d} \end{pmatrix} = \begin{pmatrix} -\frac{d_{o}}{d_{i}} & d_{o}\Delta \\ \frac{1}{d_{e}} & \frac{1}{f} - \frac{1}{d_{e}} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} x \\ u \end{pmatrix}$$
(3.7)

When the object is in focus, $\Delta = 0$ and the transform on the spatial coordinates is simply an upside down scaling by a factor of $1 - \frac{d_e}{f} = -\frac{d_e}{d_o}$, a ratio typically found in the pinhole camera model.

Defocus Imaging. When the plane of focus is not within the focal range of the eye, the spatial coordinates add the directional components; this causes the entire light field to be slanted, as shown in Figure 3.7. The final angular integration in Equation 3.5 involves integrating over neighboring pixels around the target location. The size of the integration window corresponds to the convolution kernel width, and it depends on the pupil aperture and the slope of the sheared light field, given by $1/(d_e\Delta)$.



Figure 3.7: Defocus light field. When the object lies exactly on the focus plane, its retinal light field is straight lines. When object is out of focus, the light field is sheared. Integrating the angular coordinates requires summing up the neighboring pixels in the object, resulting a blur.

Combining the aperture size and the light field slope, we obtain the diameter of the convolution kernel as the following:

$$r = \frac{a}{\frac{1}{d_c \Delta}} \frac{-d_o}{d_e} = -ad_o \Delta, \tag{3.8}$$

where the ratio $-\frac{d_o}{d_e}$ accounts for the backward projection from retina to the image plane. By substituting $\frac{d_e f}{d_e - f} = d_f$, the diameter r is:

$$r = ad_o\left(\frac{d_e - f}{fd_e} - \frac{1}{d_o}\right) = a\left(\frac{d_o}{d_f} - 1\right) = a\frac{|d_o - d_f|}{d_f},\tag{3.9}$$

which is identical to the circle of confusion diameter as we have derived in Section 3.1.1. Here we conclude this chapter by showing the connection between different mathematical models of the image blurring.

3.3 Summary

In Chapter 4 and 5, we will develop the idea of inverse prefiltering based on 2D image operator. The light field inverse projection will be explored as a generalization in Chapter 6 and 7. Finally the integration with wavefront data will be presented in Chapter 8 and 8.3.

Chapter 4

Preliminary Ideas with Deconvolution

To generate an image that appears to be sharp to the observer, inverse blurring is applied to preprocess the image content. Following the mathematical models for image blurring in Equation 3.2 of Chapter 3, the convolution process is inversely applied, called deconvolution [Gonzalez and Woods, 1992]. Deconvolution is well known to be ill-posed, and the inverse problem has been discussed for decades. Using the idea from the prior literature, we will first derive some mathematical tools that preprocess images i^p to be shown on a conventional display. Given the preprocessed image, the blurring of the observer reveals a sharp retinal image:

$$i^{p}(x,y) * k(x,y) = i(x,y).$$
(4.1)

At the end of this chapter, we will discuss theoretical limitations of the overall problem.

4.1 Frequency Domain Solvers

Since the convolution theorem states that the convolving operator in the spatial domain is equivalent to the element-wise multiplication in the frequency domain [Gonzalez and Woods, 1992], the inverse blurring, or preprocessing, can be easily defined.

Inverse Filter. The frequency domain solvers have evolved significantly in signal processing. The first widely known tool for deconvolution is the inverse filter [Gonzalez and Woods, 1992]. Following Equation 3.2 where the blurred image is obtained by convolving the image with the point spread function, the inverse filter $h(\cdot)$ is obtained by inverting the kernel in the frequency domain:

$$\widehat{h}(\omega_x, \omega_y) = \begin{cases} \frac{1}{\widehat{k}(\omega_x, \omega_y)} & \text{for } |\widehat{k}(\omega_x, \omega_y)| \ge \epsilon, \\ 0 & \text{otherwise} \end{cases}$$
(4.2)



Figure 4.1: Disk point spread function and its inverse. (a): disk point spread function, (b): the frequency domain spectrum, and (c): the inverse point spread function of (a). Due to the zero-valued frequency response, the inverse kernel has interspersed positive and negative rings.

where $|\cdot|$ represents the magnitude of a complex number, and ϵ is a small threshold to avoid divideby-zero. With the inverse kernel, we "**inversely prefilter**¹" the image such that the prefiltered image i^p is seen as a sharp retinal copy after the blurring process of the eye:

$$i^{p}(x,y) * k(x,y) \equiv i(x,y) * [h(x,y) * k(x,y)] \approx i(x,y).$$
(4.3)

Although the inverse filter is fundamentally simple, it does not account for the fact that higher frequencies are usually weaker, and a single small threshold is not appropriate for the entire frequency spectrum for every content source.

Wiener filter. By assuming orthogonality between signal and noise, zero mean noise, and linearity, it can be shown [Gonzalez and Woods, 1992] that the Wiener filter [Wiener, 1964] is a minimum mean square error estimator and has a similar form to Equation 4.2:

$$\widehat{h}(\omega_x, \omega_y) = \frac{1}{\widehat{k}(\omega_x, \omega_y)} \cdot \left[\frac{|\widehat{k}(\omega_x, \omega_y)|^2}{|\widehat{k}(\omega_x, \omega_y)|^2 + \frac{1}{\operatorname{snr}(\omega_x, \omega_y)}} \right]$$
(4.4)

where $\frac{1}{snr(\omega_x,\omega_y)}$ is the inverse signal-to-noise-ratio (SNR) acting as the new regularization term. For spatial frequencies where the SNR is high, Equation 4.4 quickly reduces to the inverse filter; when the SNR is small, the denominator becomes large and suppresses the overall value to zero. Although Wiener filter is fundamentally more powerful than an inverse filter, the signal-to-noise ratio for the entire spectrum is usually unknown or hard to obtain in advance, and thus a user specified constant is applied.

Figure 4.1 illustrates the point spread function kernel k(x, y) and its inverse h(x, y) for a defocus point spread function based on the circle of confusion. Notice that the magnitude plot (MTF)

¹in the later text, for brevity we will omit the "inverse" and simply use "prefilter"
of the frequency domain spectrum has several zero-valued frequency responses; this is because the response oscillates between the real and the imaginary part. The threshold ϵ in Equation 4.2 or the SNR in Equation 4.4 avoids inverting responses at these locations. While the kernel has the relative simple structure of a *disk* function, its inverse h(x, y) has infinitely vanishing concentric rings, as shown on the right of Figure 4.1.

One particularly interesting structure in the inverse kernel is the alternating positive and negative rings: prefiltering images by convolving with the inverse kernel, as used in the papers by Yellott and Yellott [2007] and Adrien [2008], produces images with negative intensities. The approaches using wavefront based point spread function (Eq 3.4) adopted by Alonso and Barreto [2003], Legras et al.[2004], and Mohammadpour et al.[2012] share similar problems. Since physical display pixels can only be set to positive states, the overall prefiltered image is shifted by a constant bias to compensate for the negative pixels and is normalized by the maximum pixel value:

$$i_{normalized}^{p} = \frac{i*h - \min(i*h)}{\max(i*h) - \min(i*h)}.$$
(4.5)

The darkest pixel in the final perceived image is also shifted. For human observers, image contrast is largely determined by the darkest pixel dividing the brightest pixel, and it is significantly lost in this case.

Figure 4.2 shows examples using the inverse filter and the Wiener filter at various defocus levels: the first row shows the normalized prefiltered images (Eq. 4.5), and the second row shows the perceived images, with and without the renormalization step. As the defocus level increases, the prefiltering distorts the structure more significantly. We show both versions "without" and "with" the normalization for the perceived image. For the image without the normalization (assuming negative light in the simulation²), we can clearly see the ringing artifacts, and this is because the number of zero-valued responses in Figure 4.1(b) grows as the diameter of the PSF increases. For an image with normalization that corresponds to real human perception, the contrast is significantly reduced due to the negative rings in the inverse point spread function (Fig. 4.1(c)) or inverting the "weak" frequency responses³, and the contrast loss worsens as the defocus level increases.

The small threshold ϵ in Equation 4.2 for the inverse filter corresponds to an infinite signalto-noise ratio. For the Wiener filter, the SNR can be treated as a user specified regularization control knob: a higher SNR respects the original signal but the contrast loss is severe; a lower value regularizes the process so contrast is higher, but the original content is distorted and is not as sharp. In both cases, the artifacts prohibit the readability of the original content.

²Note that negative light can be achieved using coherent light, and the wave effect provides the negative radiances. This type of display is, however, impractical for a real life scenario; but we still provide the simulation as a reference.

³Since each spatial frequency is a sinusoidal function oscillating between positive and negative, inverting a weak signal creates a strong one that also magnifies the negative part; we will discuss the behavior in the end of this chapter



Figure 4.2: Frequency domain prefiltering solutions using the inverse filter and the Wiener filter. This type of solvers generally have contrast loss problem. Depending on the regularization, the ringing artifacts can be attenuated and the contrast can be enhanced, but the image can also become more blurry.

4.2 Spatial Domain Solvers

Instead of operating in the frequency domain, spatial domain solvers provide the option to incorporate desired properties to the system or solution, such as non-negativity, gradient distribution, or even subpixel rendering. Adding constraints to the convolution is non-trivial in the spatial domain, and it usually requires an iterative solver.

Richardson-Lucy Solver. Applying Bayes's theorem, Richardson [1972] and Lucy [1974] methods iteratively solve for the most probable hidden prefiltered image. By defining the pixel probability $P(i^p(x)) = \sum_s P(i^p(x) \cap i(s)) = \sum_s P(i^p(x) | i(s)) P(i(s))$, we can use the Bayes posterior probability to evaluate the right hand side:

$$P(i^{p}(x)) = \sum_{s} \frac{P(i(s) \mid i^{p}(x)) P(i^{p}(x))}{\sum_{t} P(i(s) \mid i^{p}(t)) P(i^{p}(t))} \cdot P(i(s)).$$

$$(4.6)$$

Since the probability distribution $P(i^p(x))$ appears on both side of Equation 4.6, we solve it with an iterative strategy so that the right hand side evaluates to be the next iteration probability. By substituting the conditional probability distribution $P(i(s) | i^p(x))$ with the point spread function kernel density P(k(s-x)), we obtain the following iterative expression:

$$P_{t+1}(i^{p}(x)) = P_{t}(i^{p}(x)) \sum_{s} \frac{P(k(s-x))P(i(s))}{\sum_{t} P(k(s-t))P_{t}(i^{p}(t))} \quad \Rightarrow \quad i^{p}_{(t+1)} = i^{p}_{t} \cdot \frac{i}{i^{p}_{t} * k} * k', \quad (4.7)$$

where i_t^p is the *t*-th iteration solution to the inverse problem, k' is the flipped point spread function, and the multiplication and division are element-wise operations. One disadvantage for the iterative method is the number of the iterations it takes. Furthermore, it is generally hard to prove the convergence condition; the solution can sometimes oscillate between good and poor solutions. Because the convolution with the kernel *k* always produces positive values, the Richardson-Lucy method is a non-negative solver, and there is no contrast loss⁴; due to the non-negative property, the solution space is not exploited, and additional artifacts are presented, as shown in Figure 4.3.

As the figure shows, the Richardson-Lucy method generally produces a fairly good perceived image contrast, but artifacts make the image barely recognizable as the diameter of the PSF increases. We also show different iteration numbers, and sometimes the solution improves the image with sharper features as we increase the computation time; however, there is no clear stopping condition giving the best result, and the user has to visually judge between the high frequency details and the ringing artifacts.

⁴It can becomes darker due to the expanded dynamic range.



Figure 4.3: **Spatial domain prefiltering solutions using Richardson-Lucy method.** Although the Richardson-Lucy solver gives non-negative solution, perceived images are generally more blurry and have more ringing. More iterations can potentially reduce the artifacts, but there is no obvious stopping criteria, and the user has to determine it empirically.

Sparse Gradients Natural Image Statistics. Recent tools for image deblurring employ additional regularizations such as natural image statistics [Fergus *et al.*, 2006] or sparse gradients [Olshausen and Field, 1996] constraints; specifically, the gradient distribution of a natural image is quite different from white noise, and is either heavy-tailed or sparse in the image.

Following the derivation by Levin et al. [2007], we assume Gaussian noise distribution and have

$$\underset{i^{p}}{\operatorname{argmin}} \|i - k * i^{p}\|^{2} + \sum \rho(\nabla i^{p}), \tag{4.8}$$

where the function ρ is either a smooth prior or a sparse estimate on the gradient of the hidden image, and the smooth prior where $\rho(x) = ||x||^2$ gives us better results⁵.

Figure 4.4 shows the results obtained using the solver provided by Levein et al.[2007], and the results are slightly better than that from Richardson-Lucy solver. Note that even artifacts are suppressed, the perceived images are also more blurry as a consequence. We also compare with the frequency domain Wiener filter solution by showing both the simulated negative light and real perception.



Figure 4.4: **Prefiltering solution obtained using Levin et al. deconvolution.** By employing the smooth gradient distribution prior, the perceived image is slightly sharper, but without nonnegative constraints, the image contrast is lost similar to that of the frequency domain solver.

⁵Levin et al.[2007] use 0.8 for sparse prior. In their paper, the second order derivative is also constrained; alternatively we can constrain the gradient of the perceived image, i.e. $\|\nabla(i^p * k)\|^{0.8}$.

Super Resolution Techniques. Higher resolution images can be obtained by interleaving lower resolution ones. The technique is called "wobulation" [Allen and Ulichney, 2005], and as shown in Figure 4.5 left. Although the technique receives lots of attention such as the super resolution projector proposed by Damera-venkata and Chang [2007] [2009] and optical pixel sharing by Sajadi et al. [2012], the results appear as if there is a high resolution display.

Didyk et al. [Didyk *et al.*, 2010] and Berthouzoz and Fattal [Berthouzoz and Fattal, 2012] extend these methods by manipulating the way eye moves, and their methods directly form a high resolution retinal image using just a single low resolution display, as shown on the right of Figure 4.5. Based on "alias-cancellation", these techniques encode the higher frequency information in the low frequencies; however, the human eye has finite integration time and the high display refresh-rate decodes the hidden higher frequency.

Although the technique is made relatively simple by inverting the sub-frame integration (Fig. 4.5 center), the fundamental limitation is still how the frequency information passes through the lens [Damera-Venkata and Chang, 2009]; as shown in Figure 4.1, some of the spatial frequencies will not be transfered to the retina and thus irrecoverable.



Figure 4.5: **Prefiltering using super-resolution technique.** Super-resolution uses the "wobulation" concept (left) where the sub-frames are integrated and are perceived as having higher resolution. The resolution enhancement exploits human eye perception and encodes high frequencies into low frequencies through "alias-cancellation". The decoding is done by using a display with high refresh-rate.

4.3 Theoretical Analysis

Generalized Linear Solver Without loss of generality, we can unroll the 2D image into a 1D vector, and express the convolved pixels as linearly weighted neighboring pixels from the original images; this leads to the following expression:

$$\mathbf{i} = \mathbf{W} \cdot \mathbf{i}^p. \tag{4.9}$$

Then the solution to the inverse problem simply requires inverting the weight matrix \mathbf{W} . It is also possible to include additional constraints and make it an overall optimization problem:

$$\underset{\mathbf{i}^{p}}{\operatorname{argmin}} \|\mathbf{W} \cdot \mathbf{i}^{p} - \mathbf{i}\|^{2}, \text{ s.t. constraints } \boldsymbol{C} \text{ are satisfied.}$$
(4.10)

Typical least square solutions to the inverse problem require finding the pseudo-inverse of the weight matrix W via singular value decomposition (SVD). Additional constraints such as non-negativity, sparse gradients, and desired distributions can also be incorporated into a unified constrained linear solver system. However, inspecting the singular values of W reveals some important structure of the inverse problem: when the display object is in focus, the point spread function is a delta function and the weight matrix is a diagonal identity matrix, as shown in Figure 4.6(a); when the display is out of focus, the circle of confusion PSF expands and the weight matrix becomes banded(b), thus making an ill-posed problem.

Here we simulate a 1D case scenario where the eye has 4mm pupil, the distance to the display is 30cm, and the observer requires +7.67Diopter hyperopic eyeglasses, resulting in a near plane at 3m distance. The singular values of banded matrix is shown in Figure 4.6(c): since the observer has large blurring kernel, the singular value quickly drops significantly and becomes nearly zero around the 500-th singular value, and *the rank of the system is thus reduced*.



(a) weight matrix in-focus (b) weight matrix out-of-focus

(c) singular values of the matrix in (b)

Figure 4.6: **Convolution weight matrices and their singular values.** A focused optical system has pure diagonal identity matrix. A defocused system is banded, causing its singular values to degenerate, and it is thus harder to invert.

Frequency Domain Analysis. Another interesting perspective from which to consider the problem is in the frequency domain; following the definition of a *disk* function:

$$k(x,y) = \begin{cases} 4/\pi r^2 & \text{for } \sqrt{x^2 + y^2} < r/2, \\ 0 & \text{otherwise;} \end{cases}$$
(4.11)

the frequency domain response is given by the *jinc* function [Goodman, 2004]:

$$\widehat{k}(\omega_x, \omega_y) = 2jinc(\pi r\sqrt{\omega_x^2 + \omega_y^2}) \equiv \frac{2J_1\left(\pi r\sqrt{\omega_x^2 + \omega_y^2}\right)}{\pi r\sqrt{\omega_x^2 + \omega_y^2}},$$
(4.12)

where J_1 is the Bessel function of the first kind, and there are zeros in the function, as shown in Figure 4.7.

Since the Fourier transform is linear, *the modulation transfer function is the unsorted singular values, and the zero-valued spatial frequencies match the zero-valued singular values,* and the inverse filter and the Wiener filter are like the pseudo-inverse approximation: inverting at the lost frequencies is undefined and different regularization strategies are employed.

These zero-valued spatial frequencies lead to the first theoretical limitation of ringing artifacts and slight blurriness. *The lost frequencies or singular-values cannot be recovered since the blurring kernel is applied at the last step of the process*; generally, this information loss causes blurriness. The lack of certain frequencies at sharp edges generates some ringing artifacts, as shown in Figure 4.2.

The second problem with inverting the blurring kernel is the extremely expanded dynamic range. Due to the heavily attenuated frequency responses that are close to zero, their inversion is close to infinity, causing these frequencies to exhibit dominating sinusoidal structures in the spatial domain images, and generating negative and overwhelmingly positive pixel states; intensity renormalization is required to show the prefiltered image on the display, but it also introduces the contrast loss.



Figure 4.7: **Modulation transfer function for the disk PSF.** The frequency domain responses of a disk PSF is a *jinc* function and has zeros. These zeros transfer no frequency content to the retina, and the weak responses introduce the image contrast loss.

4.4 Summary

In the beginning of this chapter, we formally introduced the formulation of the problem (Eq. 4.1) in the context of image convolution (Sec. 3.1). Using the circle of confusion assumptions(similarly the wavefront based PSF) and a traditional display devices, we show that prior solutions obtained under different assumptions and spatial/frequency domains can roughly achieve the goal with some artifacts.

At the end of the previous section, we revealed two fundamental limitations of the prefiltering method using a traditional display:

- Frequency loss: cause slight blurriness and ringing artifacts.
- Contrast loss: the expanded dynamic range and the negative pixel values require intensity re-normalization.

While these drawbacks are lightly documented in the paper by Alonso and Barreto [2003], no theoretical improvement has been proposed in the following years. With the theoretical analysis for the prefiltering method on a traditional display, we conclude that modification to the assumption has to be introduced, and new hardware has to be built. In the next chapter, we will introduce the prefiltering method with "multilayer" type of displays that addresses the fundamental limitations.

Chapter 5

Multilayer Displays

In this chapter, we develop inverse prefiltering for emerging multilayer displays. Following the limitations on image prefiltering using the conventional display, as discussed in the previous chapter, the multilayer display enables an "all-pass kernel": there will be no zero-valued frequency responses [Huang and Barsky, 2011]. We will first discuss the observations and intuitions behind the idea, and then develop the theory of multilayer prefiltering [Huang *et al.*, 2012]; a contrast optimization will be introduced, and finally we will discuss the hardware design alternatives and our prototypes.

5.1 Frequency Preservation via Multilayer Prefiltering

For the case of image prefiltering using a conventional display, we now call it "single-layer prefiltering" when only one display panel is used. Following Section 3.1.1, the point spread function is a *disk* function of diameter r, and it is a function of the distance from the eye to the focus plane.

On the left of Figure 5.1 are two cases of defocus blur: for the display (blue) closer to the plane of focus, its point spread function is smaller. On the right of Figure 5.1, we plot the modulation transfer functions (MTFs) of the two PSFs. Since their MTFs are both *jinc* functions with periods determined by the diameter of their PSFs, they both inevitably have the zero-valued problem.



Figure 5.1: Observation of multilayer PSFs and their MTFs.

However, one interesting observation is that these zero-valued frequencies generally do not align. With a carefully chosen separation between the two layers, we can say that the coincident frequency will not happen within human perceivable frequencies (around 60 cycles per degree). So the engineering question is how to make both layers visible to the eye, and the two layers still maintain the property of different point spread functions. Fortunately, nowadays there are transparent display panels, and they are capable of refreshing at high speed (120Hz panel are commercially available). By utilizing the critical flickering rate (around 40Hz) limiting the temporal integration of the eye, we quickly interchange contents on the two displays, and the eye will fuse the two images on the retina; this is the fundamental idea of our multilayer inverse prefiltering.

Multilayer image convolution. Let's now consider the eye integrates N layers of display panels, each separated from the eye by the distance $d_1, d_2, ..., d_n$; we can formulate a fused image blurring operation:

$$i(x,y) = \sum_{n=1}^{N} \overline{i}_n^p \left(\frac{d_n}{d_1}x, \frac{d_n}{d_1}y\right) * \overline{k}_n \left(\frac{d_n}{d_1}x, \frac{d_n}{d_1}y\right).$$
(5.1)

Here, x and y are defined on the first layer display plane. We first assume that the *n*-th layer "original" prefiltered image \bar{i}_n^p is given; the actual derivations for prefiltering will be presented in the next section. To account for perspective viewing, we scale the images by d_n/d_1 according to the separation; in this way, the *n*-th layer image and point spread function are respectively projected onto the first layer. To simplify the notation, the scaled prefiltered images and point spread functions are redefined as:

$$i_n^p(x,y) = \bar{i}_n^p\left(\frac{d_n}{d_1}x, \frac{d_n}{d_1}y\right) \quad \text{and} \quad k_n(x,y) = \bar{k}_n\left(\frac{d_n}{d_1}x, \frac{d_n}{d_1}y\right).$$
(5.2)

We substitute the image formation at the first layer, and the perceived image can be expressed as the following:

$$i(x,y) = \sum_{n=1}^{N} i_n^p(x,y) * k_n(x,y).$$
(5.3)

Equation 5.3 reveals the first important benefit of a multilayer display for correcting optical aberrations: this expression is equivalent to N collocated independent single-layer displays, and they are all separated by a distance d_1 to the eye. What makes the idea different from a single-layer prefiltering problem is that, even the conceptual display panels are all located at the same distance, the effective point spread function $k_n(x, y)$ applied to each layer differs, opening the door to an "all-pass kernel" prefiltering algorithm.

Multilayer All-pass kernel. Consider the case where layer images are all identical, such that $i_n^p(x, y) = i^p(x, y)$, Equation 5.3 reduced to the following:

$$i(x,y) = i^p(x,y) * k'(x,y), \text{ where } k'(x,y) = \sum_{n=1}^N k_n(x,y).$$
 (5.4)

This demonstrates that the multilayer prefiltering can operate in a mode similar to the single-layer prefiltering, but the effective point spread function is now a linear superposition of all circle of confusions for each layer. By Equation 4.12, the frequency spectrum of the circle of confusion is given by¹

$$\widehat{k_n}(x,y) = 2jinc(\pi r_n \sqrt{\omega_x^2 + \omega_y^2}) \quad \text{where} \quad r_n = \left(\frac{d_1}{d_n} \frac{|d_f - d_n|}{d_f}\right) a.$$
(5.5)

As shown on the right of Figure 5.1, with carefully chosen layer separations, the zero-valued frequencies of the MTFs differ; all spatial frequencies are preserved in the multilayer fashion, and the effective point spread function is thus an "all-pass" filter. Applying this operation can eliminate artifacts in single-layer prefiltering.

Frequency domain analysis of multilayer kernel. The multilayer display can also operate in a mode with dissimilar layer images, while maintaining the same perceived image. By understanding the frequency domain property of the multilayer display, we show the second benefit: the perceived image contrast can exceed that achievable with single-layer prefiltering.

Taking the two-dimensional Fourier transform of Equation 5.3 yields the following expression:

$$\widehat{i}(\omega_x, \omega_y) = \sum_{n=1}^{N} \widehat{i}_n^p(\omega_x, \omega_y) \cdot \widehat{k}_n(\omega_x, \omega_y)$$
(5.6)

In Equation 5.4, we obtain for each layer the prefiltered image $\hat{i}_n^p(\omega_x, \omega_y) = i(\omega_x, \omega_y) \cdot \hat{k}_n^{-1}(\omega_x, \omega_y)$, such that the perceived image is simply the summation of the identical filtered layer images. However, as described in the previous paragraph, this operation mode assumes equally distributed image content; since the zero-valued frequencies differ in each layer, a more flexible formulation is required.

In the general form, we allow the prefiltered layer image spectrum to be expressed as:

$$\widehat{i}_{n}^{p}(\omega_{x},\omega_{y}) = \widehat{i}(\omega_{x},\omega_{y}) \cdot \left[\Psi_{n}(\omega_{x},\omega_{y}) \cdot \widehat{k}_{n}^{-1}(\omega_{x},\omega_{y})\right],$$
(5.7)

where $\Psi_n(\omega_x, \omega_y)$ is the partition function determining the relative contribution of each layer to each spatial frequency component. Note that the partition function must satisfy the unity constraint:

$$\sum_{n=1}^{N} \Psi_n(\omega_x, \omega_y) = 1, \quad \text{where} \quad 0 \le \Psi_n(\omega_x, \omega_y) \le 1.$$
(5.8)

Finally to ensure the prefiltered images are real-valued, we further impose the partition function must be odd-symmetry such that $\Psi(\omega_x, \omega_y) = \Psi(-\omega_x, -\omega_y)$.

¹Equation 14 in the multilayer paper is P_n instead of $p_n!!!$

. .

5.2 Partition Function and Contrast Optimization

In this section, we will describe different partition function strategies. Consider the prefiltered multilayer image perception model:

$$\widehat{i}(\omega_x, \omega_y) = \sum_{n=1}^N \left(\widehat{i}(\omega_x, \omega_y) \cdot \left[\Psi_n(\omega_x, \omega_y) \cdot \widehat{k}_n^{-1}(\omega_x, \omega_y) \right] \right) \cdot \widehat{k}_n(\omega_x, \omega_y),$$
(5.9)

Here we model the whole effect as a linear superposition in the frequency domain. As discussed in the previous chapter on the property of the point spread function in the frequency domain, zerovalued frequencies introduce undefined inversion; the partition function $\Psi_n(\omega_x, \omega_y)$ helps us avoid such cases, and the simplest inverse filter can be used. In practice we require the hard constraint that

$$\Psi_n(\omega_x, \omega_y) = 0$$
, when the corresponding $|\hat{k}_n(\omega_x, \omega_y)| = 0$, (5.10)

and this way the spatial frequency will be zero for the prefiltered image at that layer: $\hat{i}_n^p(\omega_x, \omega_y) = 0$; information presented in that frequency must be preserved in other layers.

Winner-Take-All Strategy To avoid the zero-valued frequency assignment, we define a simple rule that the contribution is always given to the layer with the maximum response at the specific spatial frequency:

$$\Psi_n^{wta}(\omega_x, \omega_y) = \begin{cases} 1 & \text{for } n = \operatorname{argmax}_m |\widehat{k}_m(\omega_x, \omega_y)|, \\ 0 & \text{otherwise} \end{cases}$$
(5.11)

and we call it the "Winner-Take-All" strategy; graphical illustration is shown in Figure 5.2. In simple words, the combined effective kernel is the envelope of kernels from all layers, as shown in the red dashed line, and the assignment of spatial frequency to layers is shown on the bottom.



Figure 5.2: Winner-Take-All MTF and frequency assignment.



Figure 5.3: **Winner-Take-All binary frequency partitioning.** First attempt to solve the inverse prefiltering problem using Winner-Take-All frequency assignment strategy. Although we are able to preserve all spatial frequencies and avoid the extreme attenuations, the contrast is, however, not better than that of the single-layer prefiltering, and this is due to the isolated sinusoidal wave structure induced by the binary frequency assignment, as shown on the right.

The essence of the Winner-Take-All strategy is that the layer with the maximum effective MTF $|\hat{k}_n(\omega_x, \omega_y)|$, for a given spatial frequency (ω_x, ω_y) , is always assigned with the full dominance, and the remaining layers makes no contribution to this component. Under this choice of the partition function, one can optimize the layer distances $\{d1; d2; \ldots d_N\}$ such that the minimum value of the overall MTF (i.e., the envelope of the effective MTFs) is maximized, and this corresponds to the solution to the following optimization problem:

$$\underset{\{d_2,\dots,d_N\}}{\operatorname{argmax}} \left[\min\left(\sum_{n}^{N} |\widehat{k}_n(\omega_x,\omega_y;d_n)| \cdot \Psi_n^{wta}(\omega_x,\omega_y;d_n)\right) \right].$$
(5.12)

This is equivalent to finding the max-min value in a high dimensional tensor, and can be implemented efficiently with a brute force search when the number of layers (dimensions) is small.

Figure 5.3 shows the first comparison of the Winner-Take-All strategy with the single-layer prefiltering (second column). The third column is an example of the perceived two-layer prefiltering result using the Winner-Take-All partition function, which is the fourth column. Although the entire spatial frequency spectrum is preserved, the contrast is, however, only slightly improved from the result obtained using the single-layer prefiltering. The binary frequency assignment strategy might not be always the optimal choice of partition function.

Optimal contrast optimization. It is interesting to see why the contrast in Figure 5.3 is not much improved given that the zero-valued and weak frequencies are avoided. In fact, as we separate spatial frequencies into different layers, the sinusoidal wave structure also becomes apparent. As

shown in Figure 5.2, the DC term (the lowest spatial frequency) is always assigned to the the rear layer, so the front layer image inevitably has negative values due to the negative part of the wave. Ideally, by carefully assigning fractional DC term contribution, the negative values can be maximized for all layers.

Without lose of generality, the partition function $\Psi_n^{opt}(\omega_x, \omega_y)$ allowing fractional values for all spatial frequencies can potentially improve the image contrast and can allow for a more general formulation. Below we show a generalized optimization that maximizes the final image contrast by maximizing the minimum value of the prefiltered images:

$$\underset{\{d_2,\dots,d_N,\Psi_1,\dots,\Psi_n\}}{\operatorname{argmax}} \quad \left[\min\left(iFT\{\widehat{i}(\omega_x,\omega_y)\cdot\Psi_n^{opt}(\omega_x,\omega_y;d_n)\cdot\widehat{k}_n^{-1}(\omega_x,\omega_y;d_n)\}, \quad \forall n\right)\right], \quad (5.13)$$

where the $iFT\{\cdot\}$ is the inverse Fourier transform and gives the prefiltered image. While we want to maximize the minimum negative intensities, the objective function has a nested dependency inside a Fourier transform, which makes the problem intricate; some relaxation is needed to make the optimization tractable.

Greedy Optimization Optimizing both partition functions and the separations between layers, with them depending on each other, is hard. There are two important observations that lead to our greedy optimization: the natural image spectrum usually has stronger low frequency energy, and the optimization depends more on the separations of the layers.

These observations lead to a two-step greedy strategy. First, we find the separations between layers following the Winner-Take-All optimization, that maximizes the minimum response of the envelope; this strategy gives the best potential optimization space for fractional assignment. Second, we assign the spatial frequencies following a monotonically decreasing order that the larger normalized energy of a spatial frequency will be processed first. The normalized energy is defined as the following:

$$|\widehat{i}^{norm}(\omega_x, \omega_y)| = \sum_{n=1}^N \frac{|\widehat{i}(\omega_x, \omega_y)|}{|\widehat{k}_n(\omega_x, \omega_y)|}.$$
(5.14)

In Equation 4.2, we find the prefiltered image using the inverse of the kernel, and we also use the reciprocal here; the kernel response in the denominator encourages zero-valued spatial frequencies to be processed first, and the magnitude of image spectrum in the numerator prioritizes the stronger lower frequencies.

We define the following iterative assignment optimization for the spatial frequency (ω_x, ω_y) with each step where $|\hat{i}^{norm}(\omega_x, \omega_y)|$ is the strongest among all other unassigned spatial frequencies:

$$\Psi_{\{1\sim N\}}^{GR}(\omega_x,\omega_y) \leftarrow \operatorname*{argmax}_{\Psi_{\{1\sim N\}}^*(\omega_x,\omega_y)} \left[min\left(i_1^p(x,y),\ldots,i_n^p(x,y)\right) \right].$$
(5.15)

In the spatial domain, the frequency (ω_x, ω_y) is a sinusoid wave, and the *min* is a nonlinear function, we implement the optimization using a line search; at each step on the line, we add the



Figure 5.4: **Contrast optimization via greedy partition function.** Compared with the singlelayer prefiltering and the Winner-Take-All strategy, the greedy optimization strategy not only preserve all spatial frequencies, but it also provides better image contrast. The greedy strategy allows fractional frequency assignment to different layers.

"fractional" sinusoidal waves, from the original image $i(\omega_x, \omega_y)$, to the N determined prefiltered images among which we then find the minimum intensity value. Note that under this scheme, the DC term is optimized in the end, since there is no sinusoidal wave structure.

In Figure 5.4 we compare the Winner-Take-All strategy with the greedy strategy using two layers. The image contrast is improved as we allow for fractional partition weights, as shown on the right. Notice the solid concentric lines in the partition function; this is because the zero-valued spatial frequency occurs in the opposite layer, and thus the corresponding layer is assigned with full dominance. In Figure 5.5, we compare the prefiltering results using the eye chart example, and their corresponding prefiltered image is shown on the right; notice that the greedy strategy multilayer prefiltering can provide much better image contrast than that of the single-layer prefiltering.



Figure 5.5: **Comparing the perceived images from prefiltering algorithms using the eye chart example**. We compare the single-layer prefiltering with our multilayer prefiltering, and their corresponding prefiltered images are shown on the right.

5.3 **Requirement and Construction**

After describing the principle of multilayer prefiltering, the requirement for hardware construction will be discussed in this section. The only requirement for the desired display is the optical equivalence to a Lambertian multilayer display.

5.3.1 Design alternatives

Any practical multilayer display must meet four design criteria. It should: (1) be optically equivalent to a stack of semi-transparent, lighting-emitting layers, (2) be thin, (3) support binocular correction, since refractive errors may differ between eyes, and (4) support a wide field of view. In addition, the display should ideally support HDR modes, due to the expansion in dynamic range. In this section, we assess the ability of various display technologies to meet these constraints. We observe that most of these constraints are shared by autostereoscopic displays. We propose adapting these emerging architectures to the task of optical aberration correction.

Multilayer OLEDs OLEDs contain an organic film enclosed between electrode arrays that emit light proportional to the applied voltage. Transparent OLEDs incorporate semi-transparent contacts [Görrn *et al.*, 2006] providing an ideal architecture for multilayer prefiltering. However, such displays do not support binocular correction. To address this limitation, we propose placing a parallax barrier or a lenslet array in front of an OLED stack; such elements ensure each eye views different pixels on each layer, enabling binocular correction at the cost of reduced resolution.

Beam-splitter Trees LCDs dominate consumer applications, with OLEDs restricted to smaller form factors. Large-format OLEDs are poised for introduction, yet a multilayer display incorporating LCDs currently possesses greater commercial potential. An LCD contains two primary components: a backlight and a spatial light modulator (SLM). The SLM comprises a liquid crystal layer enclosed between electrode arrays and surrounded by a pair of crossed linear polarizers. The SLM acts as a light-attenuating layer, with opacity varying depending on the applied voltage. Layering multiple SLMs implements a stack of semi-transparent, light-attenuating layers [Bell *et al.*, 2008].

Viewing multiple LCDs through a set of half-silvered mirrors (i.e., beamsplitters) is optically equivalent to a stack of semi-transparent, light-emitting layers [Akeley *et al.*, 2004]. Although providing a practical embodiment for multilayer prefiltering, such a design falls short of our design criteria by requiring a large enclosure, prohibiting binocular correction, and restricting viewer movement.

Multilayer LCDs We observe that multilayer LCDs can be operated in another manner that is optically equivalent to a stack of light-emitting layers, while achieving a thin form factor. High-speed LCDs allow stereoscopic viewing with shutter glasses [Urey *et al.*, 2011]. For this application, the panels are refreshed at 120 Hz, with the left-eye and right-eye images sequentially displayed while a shutter is opened over the corresponding eye. We propose a similar time-multiplexed display mode, wherein the prefiltered images are sequentially displayed on each layer, while the other layers are rendered transparent. Assuming a flicker fusion threshold of 60 Hz [Kaufman and Alm, 2002], a viewer will perceive an N-layer display, composed of semi-transparent, emissive layers, if the refresh rate of the panels is 60 * N Hz. In Section 5.3.2, we analyze a two-layer LCD prototype. Similar to multilayer OLEDs, additional optical elements are required to support binocular correction. We propose incorporating directional backlighting to ensure that each eye perceives a different image. As described by Urey et al. [2011], directional backlighting consists of a rear-illuminating light guide capable of directing illumination independently to each eye in a time sequential manner. As a result, we conclude that viewer tracking will be required to ensure that the layer images are compensated for changes in perspective.

Light Field display Practical multilayer displays require increasing the display thickness, limiting mobile applications. By Equation 5.13, the optimal layer separation depends on the viewer's refractive error and position. Although a fixed separation can be employed, dynamic adjustment is preferred. Rather than constructing multiple physical display layers, we observe that emerging light field displays can synthesize virtual layers at arbitrary distances from the display surface. Furthermore, since such displays are optimized for autostereoscopic viewing, binocular correction is naturally supported; we will formally analyze the operation mode in Chapter 6

A light field display can control the radiance of emitted light rays as a function of both position and direction [Urey *et al.*, 2011]. For autostereoscopic modes, the light field replicates that produced by a 3D scene. To date, commercial light field displays primarily rely on two technologies: parallax barriers [Ives, 1903] and integral imaging [Lippmann, 1908]. Each pixel of the light field display is capable of emitting light rays within a field of view of degrees, creating a multi-view display.

We propose a new operation mode for light field displays; rather than replicating a 3D scene, we propose emitting a light field that replicates a virtual stack of semi-transparent, light-emitting layers. Such virtual layers can be displaced dynamically to account for viewer movement. Yet, light field displays suffer from two limitations. First, increasing angular resolution requires decreasing the spatial resolution; the underlying display requires a greater resolution than an equivalent multilayer display constructed with physical panels. Second, light field displays exhibit a finite depth of field, limiting the range over which virtual layers can be synthesized. In the next chapter, we will discuss the theoretical light field analysis, the implementations of the light field displays, and the comparison with contemporary light field distortion based "Tailored Display" [Pamplona *et al.*, 2012].

5.3.2 Prototype construction

Hardware. We built an early prototype using a beam-splitter tree similar to the construction by Akeley et al. [2004], but its form factor and field of view was not satisfying. Here we will show a more sophisticated construction using multilayer LCDs. As described in Chapter 2, PureDepth, Inc. markets two-layer LCDs [Bell *et al.*, 2008]. However, the separation between panels cannot be altered and additional layers are not available. As a result, we employ a multilayer LCD following the design of Lanman et al. [2011].

As shown in Figure 5.6, the prototype comprises four modified 40.8cm-by-30.6cm Barco E-2320 PA LCD panels, supporting 8-bit grayscale display with a resolution of 1600-by-1200 pixels and a refresh rate of 60 Hz. Each panel was disassembled and mounted on an aluminum frame. The panels are arranged on a stand and suspended from a set of four rails, allowing their separation to be continuously adjusted. The front and rear polarizing films are removed from each panel and replaced with American Polarizers AP38-006T linear polarizers; a pair of crossed polarizers enclose the rear layer, with successively-crossed polarizers affixed to the front of the remaining layers. The stack is illuminated using a single backlight. With this configuration, each LCD behaves as an unmodified panel when the other panels are rendered transparent. As described in Section 5.3.1, the stack is operated in a time-multiplexed manner such that only one panel displays content at any given time. With a sufficiently long exposure (i.e., $\geq N/60$ seconds when N layers are used), the prototype appears as a semitransparent stack of light-emitting layers. A 2.8 GHz Intel Core i7 workstation with 8 GB of RAM controls the prototype, and a four-head NVIDIA Quadro NVS 450 graphics card synchronizes the panels.



Figure 5.6: Prototype multilayer display, front view and side view.

We briefly outline the limitations of the proof-of-concept prototype, relative to a preferred commercial embodiment. First, the panels only support a 60 Hz refresh rate; for two-layer prefiltering, the effective refresh rate is reduced to 30 Hz, falling below the 60 Hz human flicker fusion threshold. As a result, our ability to conduct user studies is hindered, due to flicker being perceived when using multiple layers. Yet a long camera exposure allows multilayer prefiltering experiments. Second, the panels only support grayscale modes. This has the benefit of mitigating moire's patterns resulting from layering LCDs [Bell *et al.*, 2008] and increasing the brightness by eliminating attenuation across multiple color filter arrays. We record color images by simulating a field sequential color (FSC) backlight (i.e., a strobed backlight that illuminates the stack with time-varying color sources); for the results in Figure 5.9, we combine three separate photographs, each recorded while displaying a different color channel of the prefiltered images.

Software We implemented the single-layer and multilayer prefiltering algorithms described in Section 5.1 using a combination of Matlab scripts and compiled C/C++ routines. The FFTW discrete Fourier transform library is used to accelerate prefiltering. For color images, each channel is processed independently in a separate thread. For a 256-by-256 color image, single-layer prefiltering requires an average of 1 second for processing; two-layer prefiltering takes 5 seconds, when using the Winner-Take-All partition function, and 150 seconds when using the greedy partition function. All run times are reported using the same workstation used to control the prototype.

5.4 Results

Simulated results. We first show some simulated corrections using both single-layer prefiltering and two-layer prefiltering. The results shown in Figure 5.7 are simulated with a 50mm f/1.8 lens with object at a distance 100cm, and the camera focuses at 84cm; the separation between layers is 3.4cm. By comparing with the blurred images, clearly the perceived images preprocessed with the prefiltering algorithms are sharper. For the first two examples, we show both the negative light simulation and the real perceived images. The negative light simulation enables us to see more clearly how the ringing artifacts are successfully removed with multilayer prefiltering. In the meantime, the image contrast are greatly improved, as shown in the third row. On the right of Figure 5.7, the corresponding prefiltered layer images are shown.



Figure 5.7: **Comparing simulated single-layer prefiltering results with multilayer prefiltering results.** For the first two rows, we also show the comparison with negative light assumed to delineate the ringing artifacts due to the lost frequencies.

CHAPTER 5. MULTILAYER DISPLAYS

Comparisons with equal brightness. As one might argue that the brightness setting (dynamic range) of the multilayer display may be different from a conventional display used for single-layer prefiltering, in Figure 5.8 we also show direct comparisons when the brightnesses are equalized to match each other. With the Michaelson contrast, which is defined as $(I_{max} - I_{min})/(I_{max} + I_{min})$, as our evaluation metric, the perceived image contrast is thus "independent" of display brightness. In both examples, the multilayer prefiltering results are still better than the single-layer prefiltering results in the image contrast, and as discussed in Section 5.1 and Section 5.2, this is achieved by avoiding the zero-valued/weak spatial frequencies and greedy contrast optimization. On the bottom left of Figure 5.8, we also show the ringing artifacts in the single-layer prefiltering due to the zero-valued spatial frequencies and the weak frequencies affected by the regularization. These problems are eliminated by inverting the "all-pass-kernel" of the multilayer prefiltering algorithm.



Figure 5.8: **Comparing simulated results when brightness are equalized.** For a fair comparison, the brightness of the results from single-layer and multilayer prefiltering are equalized to match with each other. On the bottom left of each example, we manifest the ringing artifacts from single-layer prefiltering with negative light assumed.

CHAPTER 5. MULTILAYER DISPLAYS

Camera photographs Using the same parameters as the previous simulation, a Canon EOS Rebel T3 digital camera, with a Canon EF 50mm f/1.8 II lens, was separated by 100 cm from the front layer of the prototype. The camera is focused at 16cm in front of the display, with the minimum f-number setting of f/1.8, resulting in an aperture diameter of 2.8cm.

Figure 5.9 summarizes experimental results achieved with the multilayer LCD prototype. The same three sample images are evaluated. As described in Section 5.3.2, three exposures are combined to synthesize color images using the grayscale panels. Comparing Figure 5.9 with Figure 5.7 confirms the predicted contrast enhancement and elimination of ringing artifacts. For example, the inset region of the bird appears brighter and with higher contrast using multilayer prefiltering, rather than the prior single-layer prefiltering algorithm. Also note that the outline of the eye and the black stripes appear with less distortion using multilayer prefiltering. Ringing artifacts, visible on the left-hand side of the face of the blue toy, are eliminated with multilayer prefiltering.



Figure 5.9: **Camera photographs of prefiltering results.** With the same parameters as the simulated experiments, the multilayer (two-layer) prefiltering has better image contrast and has no obvious ringings. The additional ringing artifacts are due to spatially varying point spread functions, spherical aberrations, non-circular camera aperture, and residuals due to the non-linear gamma correction and diffraction.

Experimental results also reveal limitations of the linear spatially invariant (LSI) model introduced in Section 3.1. First, the medical display panels used in the prototype do not produce a linear radiometric response; gamma compression was applied to the displayed images, with a calibrated gamma value 2.2, to approximate a radiometrically linear display. Remaining radiometric non-linearities contribute to ringing artifacts in the experimental imagery. Second, the lens produces a spatially-varying PSF, as analyzed by Kee et al. [2011]; as seen in the bottom left of the currency image, differences between the modeled and experimental PSFs result in ringing artifacts in the periphery. However, the central region is well approximated by the defocused camera model introduced in Section 3.1.1. The camera lens aperture, consisting of several blades, used in the experiment does not produce a circular symmetric point spread function, and thus the optical transfer functions are different. Finally, the Canon EF 50mm f/1.8 lens has some spherical aberrations, which is not modeled in the current experiments; we will have more discussions about the modeling of higher order aberrations in Chapter 8.3.

We quantitatively assess the received image using the Michelson contrast metric, given by the ratio of the difference of the maximum and minimum values, divided by their sum. Michelson contrast is increased by an average of 44% using multilayer prefiltering rather than single-layer prefiltering. Following Section 4.1, prefiltering expands the dynamic range both above and below the range of radiance values that are physically supported by the display. We quantify this effect by evaluating the dynamic range compression (DRC) of the prefiltered images, given by the difference of the maximum and minimum values before normalization using Equation 4.5. By convention, the displayed normalized images always have a dynamic range of unity. For these examples, the dynamic range is reduced by an average of 42%, enabling contrast to be enhanced with multilayer prefiltering, despite normalization.

Video Data Prefiltering can also apply to video sequences. Without modifications, processing each frame independently produces videos with rapid intensity variations, as shown in Figure 5.10. We attribute this to the fact that normalization changes the mean received image value, due to variations in the minimum and maximum values of the prefiltering images. For a pre-recorded sequence, perceived flashing can be removed by normalizing each frame by the global minimum and maximum values of the prefiltered sequence, as shown in Figure 5.11. For interactive or streaming content, we propose applying an adaptive filter to recursively estimate a temporally smoothed estimate of the necessary normalization range.



Figure 5.10: Dynamic range variations of prefiltering in a video sequence.

CHAPTER 5. MULTILAYER DISPLAYS



Figure 5.11: Dynamic range normalization in video example.

5.5 Evaluations

5.5.1 Error in refractive error and calibration

Slight prescription errors are common, and the precise distance to the display is hard to obtain without external monitoring devices; knowing how sensitive the prefiltering methods are is important. Both existing single-layer and the proposed multilayer prefiltering algorithms are sensitive to perturbations in the viewer's refractive error. As shown on the left of Figure 5.12, if the corrective power differs from the viewer's eye refractive error, then the received image will be degraded; this is because the wrong optical transfer function is used in the kernel inversion. In general, there are only small changes in the lower frequencies due to the differences, but the higher frequencies will, as shown in Figure 5.1.

Both single-layer and multilayer prefiltering require tracking the viewer. With single-layer prefiltering, the distance to the viewer must be estimated to model the PSF in the plane of the display; however, unlike single-layer prefiltering, multilayer prefiltering also requires tracking lateral motion to ensure that the multiple layers are rendered with the correct perspective. The sensitivity to lateral tracking error is depicted on the right of Figure 5.12. Although the features are sharper when compared with the prescription error, the ringing artifacts coming from the misalignment are not canceled out.



Figure 5.12: Sensitivity analysis due to refractive errors and imprecise calibrations. Both single-layer prefiltering and multilayer prefiltering require precise prescriptions and calibrations. The multilayer also requires an additional dimension on lateral motion to account for perspective projection.

5.5.2 Performance over different levels of blurring

Increasing contrast in the received image lies at the heart of enabling practical applications of single-layer and multilayer prefiltering. The prototype results demonstrate moderate improvements over single-layer prefiltering, while achieving the goal of eliminating ringing artifacts. Similar to the strong dependence on depth of field for light field pre-distortion [Pamplona *et al.*, 2012], Figure 5.13 assesses the dependence of contrast enhancement on the required corrective power.

From this analysis, we identify a key limitation of the proposed multilayer prefiltering algorithm: the received image contrast is significantly reduced for large amounts of defocus, as shown on the left of Figure 5.13. Note that this assesses the required diopters "in addition" to the supplied modulation of the crystalline lens, and in practice, we find out that for small amount of blur, single-layer prefiltering may just be good enough; however, the performance of single-layer prefiltering decreases rapidly, and the multilayer prefiltering has about 44% better contrast on average. On the right of Figure 5.13, we show that the required dynamic range of the multilayer prefiltering is about 42% less than single-layer prefiltering, and this is manifested in the video sequence example.



Figure 5.13: Contrast over different levels of blurring.

5.5.3 Comparison with prior methods

We also assess the performance of single-layer prefiltering, relative to multilayer prefiltering, with different choices for the underlying deconvolution method. Figure 5.14 tabulates the perceived images, again assuming a "negative light enabled" display emitting rays with both positive and negative radiance to manifest the differences in terms of frequency preservation. The first column tabulates the results obtained using Richardson-Lucy deconvolution. Note the severe ringing artifacts . The second column tabulates the results obtained using the deconvolution method introduced by Levin et al. [2007]. While ringing is reduced, sharp features appear blurred (e.g., the coin and the figures on the bills). Given knowledge of the precise point spread function, the

CHAPTER 5. MULTILAYER DISPLAYS

Wiener filter is capable of preserving such features, as shown in the third column. However, with single-layer prefiltering, certain frequencies cannot be preserved, resulting in ringing artifacts. As shown in the fourth column, multilayer prefiltering, employing Wiener deconvolution, eliminates ringing and preserves sharp features.



Figure 5.14: **Comparison of single-layer deconvolution.** We assess the relative benefits and limitations of the Wiener filter, the iterative Richardson-Lucy algorithm, and the method introduced by Levin et al. [2007].

5.6 Discussion

As established by theory and experiment, multilayer prefiltering achieves our primary goal: mitigating contrast loss and eliminating ringing artifacts observed with single-layer prefiltering. Yet, multilayer prefiltering comes at a cost of added components, increased computational complexity, and expanded display thickness. However, to our knowledge, our introduction of the multilayer partition function is the first avenue to allow demonstrable increases in the contrast of images presented with prefiltered displays. A promising direction for future work is to explore the potential for three or more layers to achieve further increases in contrast; in addition, our greedy partition function is but one choice for enhancing contrast. We anticipate that further research may reveal computationally efficient alternatives that achieve greater contrast, through refined optimization algorithms, than our iterative approach.

In this work, we have optimized the Michaelson contrast and the dynamic range of the received image, as measured in a linear radiometric domain. A promising direction for future work is to explore alternative, possibly non-linear, perceptual optimization metrics. Following Grosse et al. [2010], incorporating the human contrast sensitivity function (CSF) [Kaufman and Alm, 2002] may allow further perceived gains in contrast.

As described in Section 5.3.1, emerging light field displays are a compelling platform for achieving practical applications of multilayer prefiltering. By utilizing synthetic, rather than physical, layers, display thicknesses can be reduced and layers can be virtually displaced to account for viewer movement. We will start immediately in the next Chapter to discuss this viable option.

Chapter 6

Light Field Analysis

In this chapter, we will extend the analysis of the light field introduced in Chapter 3. First, we will present the construction of the light field display and the analysis of the depth of field, which is crucial to the multilayer prefiltering and the light field predistortion [Pamplona *et al.*, 2012] methods when using the virtual layers emitted from light field displays. We will show analytically and qualitatively how the multilayer prefiltering outperforms the light field predistortion by Pamplona *et al.*, 2012].

Inspired by the prefiltered virtual layer methods, we hypothesize that image quality can be further improved with better use of the light field display. We will formally analyze image inverse blurring in the frequency domain of the light field; the insight will lead us to a better understanding of the problem. In the next chapter, we will describe a generalized algorithm, called the "light field prefiltering", that significantly improves the results.

6.1 Light Field Displays

Light field displays are traditionally considered as autostereoscopic multi-view 3D displays. We will briefly describe the principles of the integral based and parallax barrier based constructions. The frequency domain analysis of the depth of field helps us understand the virtual layer in comparison with the conventional display, and the limitation of it.

6.1.1 Principles and Constructions

Light field display can be considered as a virtual window to a 3D scene, and the display is capable of emitting the corresponding light field as a function of both position and direction [Urey *et al.*, 2011]. This type of display does not require tracking and glasses to emulate a 3D scene to the viewer; at different viewing directions, the eye receives different images. Each pixel of the display is capable of emitting different light rays toward different directions; we consider two common constructions, namely parallax barrier [Ives, 1903] and integral imaging [Lippmann, 1908]. We

also follow the relative two-plane parameterization [Chai *et al.*, 2000], as discussed in Section 3.2, where the x-plane denotes the spatial coordinates, and the u-plane, 1-unit separated from the x-plane, denotes the angular directions.

On the left of Figure 6.1, a parallax barrier based light field display is shown. This type of display requires a thin pinhole array laying on top of the high resolution display panel. The pinholes are separated by Δx that determines the macro pixel pitch. Underlying the pinhole, the high resolution panel has pixel pitch p; this gives a spatial-angular trade-off ratio $\Delta x/p$, and this light field display has $\Delta x/p$ views. One important design parameter is the field of view θ , and is determined by the separation distance f_l between the pinhole mask and the high resolution panel:

$$\theta = 2atan\left(\frac{\Delta x}{2f_l}\right). \tag{6.1}$$

Thus, the construction transforms the high spatial resolution conventional display panel with small pixel pitch p into a low spatial resolution light field display with pixel pitch Δx , but the device is capable of emitting different $\Delta x/p$ rays/views across θ degree viewing zone. Note the angular sampling rate Δu on the angular u-plane also determines the capability of the field of view of the display, as we will discuss in the next section. The construction using integral imaging is shown on the right of Figure 6.1, where the pinhole array is replaced by the lenslet array enabling brighter images. Instead of specifying the separation from the display, the field of view is directly encoded into the specification of the lenslet array.

6.1.2 Depth of Field Analysis

Similar to a camera and human eye, objects lying outside the range of depth of field are also blurred. The supported depth of field of the light field display is of great interest to us: it tells how the image quality degrades as a function of separation from the display [Wetzstein *et al.*, 2012]; Perwass and Wietzke [2012] also derived similar results for the light field camera, Raytrix.



Figure 6.1: Construction of light field displays. The light field display can be constructed with a parallax barrier (left) or an integral imaging (right). Both use a high resolution display panel, and a pinhole mask or lenslet array converts the spatial pixels into angular rays by trading spatial resolution for angular views at the ratio of $\Delta x/p$. The light field display has a field of view θ degrees and depth of field $\Delta x/\Delta u$.

We consider again the light field display as shown in Figure 6.1, where the spatial sampling interval is Δx and the angular sampling interval is Δu . Following the definition of Nyquist rate [Gonzalez and Woods, 1992], the maximum bandwidth of the display in the spatial frequency is $|\omega_x^d| \leq 1/(2\Delta x)$, and the maximum bandwidth of the display in the angular frequency is $|\omega_u^d| \leq 1/(2\Delta u)$. The frequencies that can be shown on this light field display are bounded by a rectangular region, as shown on Figure 6.2(a); frequencies outside the region cause aliasing artifacts [Zwicker *et al.*, 2006]. The maximum spatial frequency of a virtual plane floating outside the display enclosure can be derived as a function of separation *d* from the display.

First, for a virtual object in the center depth of the display (d = 0), its corresponding light field, as we have discussed in Section 3.2, is a set of vertical lines as shown on Figure 6.2(b). When the object is separated by a non-zero distance d from the display, its light field is sheared by the transport matrix $\mathbf{T}(d)$ defined in Figure 3.6; now these lines have a slope 1/d, as shown in Figure 6.2(c). In the frequency domain, the light field is a straight line, of slope -d, orthogonal to the spatial domain representation.

Since the display cannot reproduce frequencies outside the bounding rectangular region, the perceived virtual object has the maximum spatial frequency determined by the intersections of the rectangle and the line of the light field in the frequency domain. Consider two cases where transition is happening when the line passes the corners of the bounding rectangle: (1) When the separation d is smaller than the ratio $\Delta x/\Delta u$, the maximum spatial frequency of the virtual object is always bounded by the left and right edges of the rectangular box, and the bandwidth is given by $1/(2\Delta x)$, as shown in Figure 6.2(c). (2) When the separation is larger than the threshold ratio, its maximum spatial frequency is clipped by the top and bottom edges, as shown in Figure 6.2(d), and is smaller than that in (c). The bandwidth is now given by the clipped spatial frequency $1/(2d\Delta u)$, and it degrades as the reciprocal of the separation distance.



Figure 6.2: **Deriving image quality as a function of the depth of field.** The light field displays have fixed spatial and angular bandwidth (a). For a virtual object shown exactly on the display, its light field is straight lines (b). When it is set apart from the display surface, its light field is sheared, and so is its frequency domain representation (c). Intersection of the frequency domain representation with the bandwidth gives the perceived maximum spatial frequency of the virtual image (c) and (d).

Now we can formally describe the maximum spatial frequency of the virtual object: when it is placed at distance d_n other than the light field display location d_o , the bandwidth is a function of separation $|d_n - d_o|$ from the display as the following:

$$\omega_{x-max}^{d} = \begin{cases} \frac{d_{n}}{d_{e}} \frac{1}{2\Delta x} & \text{for } |d_{n} - d_{o}| \leq \frac{\Delta x}{\Delta u} & \text{(within the field of view),} \\ \frac{d_{n}}{d_{e}} \frac{1}{2|d_{n} - d_{o}|\Delta u} & \text{otherwise} \end{cases}$$
(6.2)

The ratio d_n/d_e accounts for the scaled projection onto the retina, and the function is shown by the red curve in Figure 6.3. This is a crucial analysis of the depth of field of the light field display; it depicts the principal capability on how far the object can be virtually shown outside the display enclosure without sacrificing image quality. Equation 6.2 suggests that the maximum spatial frequency is a nearly constant line for virtual objects within a threshold $\Delta x/\Delta u$ separation from the center depth of the display. As the virtual object moves far from the display enclosure, it becomes blurred with a rapidly decreased spatial bandwidth, similar to the out-of-focus objects found in traditional photographs.

6.1.3 Vision-Correcting Light Field Displays

In Chapter 5, we described the multilayer prefiltering algorithm using two physical LCD panels as the prototype construction; we now observe the physical LCDs could be substituted with virtual layers simulated by the light field display. Rather than constructing multiple physical display layers, we observe that light field displays can synthesize virtual layers at "arbitrary distances" from the display surface. The optimal layer separation, as given by Equation 5.13, depends on the viewer's refractive error and position, and using the light field display allows for the dynamic separation in the optimization. We propose a new operation mode for light field displays: rather than replicating a 3D scene, the display emits a light field that replicates a virtual stack of semi-transparent, light-emitting layers.

In this section, we assess the capability of a light field display to correct for defocus blur. We compare two operation modes: light field pre-distortion and synthetic multilayer pre-filtering. As recently introduced by Pamplona et al. [2012], given a light field display of sufficient resolution, the former operation mode involves emitting a predistorted light field such that, when viewed by the optics of the eye, an undistorted image is formed on the retina. This mode of operation is similar to existing wavefront correction methods [Kaufman and Alm, 2002]. For example, defocus is corrected by displaying a virtual layer at the closest plane of focus to the light field display surface. Depending on the magnitude of defocus, this virtual layer may be located far from the surface and the image quality degrades, as we have shown in Equation 6.2. In contrast, synthetic multilayer prefiltering requires synthesizing two or more virtual layers, generally in close proximity to the display; when the virtual layers are within the field of view, they can fully utilize the maximum bandwidth without degradation.

We formally assess these modes by comparing the depth of field expressions describing conventional displays and light field displays. As a baseline point of comparison, we consider the depth of field for a conventional display (e.g., an LCD) located a distance d_o from the viewer. Similar to Equation 5.5, the diameter of the circle of confusion for a defocused camera, projected onto the retina, is given by $r = (d_e/d_o)(|d_f - d_o|/d_f)a$. Thus, the maximum spatial frequency in a defocused image of a conventional display is:

$$\omega_{x-max}^{d} = \min\left(\frac{d_o}{d_e}\frac{1}{2p}, \frac{d_o}{d_e}\frac{d_f}{2|d_f - d_o|a}\right),\tag{6.3}$$

where the two arguments denote the sampling rate given by half the reciprocal of the projected display pixel width and the circle of confusion diameter, respectively.

Substituting d_f for d_n , the ratio of Equation 6.2 to Equation 6.3 provides an analytic expression for the maximum resolution enhancement r_{max} that can be achieved by depicting a virtual layer using a light field display, rather than a conventional display; this expression characterizes the benefit of affixing a lenslet array or parallax barrier to the underlying display. When the virtual layer is significantly separated from the display surface (i.e., $|d_o - d_f| \ge \Delta x / \Delta u$), this ratio is given by:

$$r_{max} = \frac{a}{d_o \Delta u} \tag{6.4}$$

We observe that r_{max} is equal to the number of light rays entering the pupil aperture from a single lenslet or pinhole. This provides formal intuition into the primary limitation of light field pre-distortion: a high angular resolution light field display is required when virtual planes are significantly separated from the surface.



Figure 6.3: **Depth of field analysis of light field displays.** We plot the bandwidth degradation of different displays, and the bandwidth is relatively flat within the depth of field of the light field display. The predistortion based algorithm simulated a virtual layer at the plane of focus far from the display surface, so the degradation is severe. The virtual multilayer method prefilters two virtual planes within the depth of field, thus the spatial bandwidth is higher.

In Figure 6.4, we consider a specific example using current-generation LCDs and lenslet arrays. Note that even with a state-of-the-art LCD with 264 pixels per inch (PPI), affixing a lenslet array slightly decreases the received image resolution, relative to an unmodified display. This is because, using light field pre-distortion, the virtual layer must be displaced well beyond the high-resolution region of the depth of field. In contrast, multilayer pre-filtering only requires virtual layers within the high-resolution region, enabling a high resolution image to be received, albeit with decreased contrast.

We conclude that light field displays present a compelling platform that meets our design constraints. As observed by Pamplona et al. [2012], light field pre-distortion will be feasible only once resolutions significantly exceed current commercial panels (approaching 1900 PPI). While reducing contrast, multilayer pre-filtering can be implemented using current-generation displays with 264 PPI.

6.1.4 Summary

The light field display has proven to be able to correct for visual aberrations using either virtual multilayer prefiltering or light field predistortion algorithms. However, the optical blur can be corrected at the cost of contrast loss or at the expense of a high density panel. One question left unanswered is: "Can can the prefiltering method be combined with the full 4D light field, instead of just layers, such that higher contrast will be achieved, and a lower resolution panel will be required ?" In the next section, we will formally describe all the prior work in the light field domain; the analysis provides insights that redundant information can be avoided and the unexploited area in the frequency domain can be utilized.



Figure 6.4: **Example of vision-correcting algorithms using light field displays.** The predistortion based methods produce lower spatial resolution than without correction; a higher resolution panel is required to correct for aberrations. On the other hand, the prefiltering based method simulates virtual layers within the depth of field of the light field display; thus the bandwidth is higher.

6.2 Image Formation in the Frequency Domain Light Field

Following Section 3.2 on the spatial domain light field image formation, in this section, we will derive the frequency domain interpretation. Under the frequency domain representation, the prior work on inverse blurring preprocessing can be explained using the same framework, and this leads to the next chapter on light field prefiltering.

6.2.1 Light field image formation

Frequency analyses have become standard tools to generate an intuitive understanding of performance bounds of computational cameras and displays (e.g., [Durand *et al.*, 2005; Levin *et al.*, 2009; Wetzstein *et al.*, 2011]); we follow this approach. First, we note that the coordinate transform in Equation 3.6 between the display and the retina can be used to model the coordinate transtransform in the frequency domain as

$$\begin{pmatrix} \omega_x \\ \omega_u \end{pmatrix} = \mathbf{Q}^{-T}(d_e) \quad \mathbf{T}^{-T}(d_e) \quad \mathbf{R}^{-T}(f) \quad \mathbf{T}^{-T}(d_o) \begin{pmatrix} \omega_x^d \\ \omega_u^d \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \frac{1}{d_e} \\ 0 & -\frac{1}{d_e} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -d_e & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & \frac{1}{f} \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ -d_o & 1 \end{pmatrix} \begin{pmatrix} \omega_x^d \\ \omega_u^d \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{d_o}{d_e} & \frac{1}{d_e} \\ d_o \Delta & \frac{1}{f} - \frac{1}{d_e} \end{pmatrix} \begin{pmatrix} \omega_u^d \\ \omega_u^d \end{pmatrix}$$
(6.5)

where ω_x and ω_u are the spatial and angular frequencies of the light field inside the eye; the corresponding frequencies of the display are ω_x^d and ω_u^d .

One of the most interesting results of the frequency analysis is the effect of the pupil outlined in Equation 3.5. The multiplication with the pupil function in the spatial domain becomes a convolution in the frequency domain:

$$\widehat{i}(\omega_x) = \left(\widehat{l}_e * \widehat{A}\right)(\omega_x, 0).$$
(6.6)

Based on the Fourier photography slice theorem [Ng, 2005], the angular projection in the spatial domain becomes a slicing along $\omega_u = 0$ in the frequency domain.

Lambertian surfaces. Embedded within a 4D light field, a Lambertian plane has no angular variation, and its frequency domain representation is a 2D plane embedded in 4D; in a flatland light field, it is a line, as shown in Figure 6.5(a). For an object in focus, its spatial light field is a set of vertical lines; thus, its frequency representation is a horizontal line. The convolution with the aperture response \hat{A} is reduced to a vertical spreading. In this case, all spatial frequencies are preserved at $\omega_u = 0$, and the final axis slicing captures all information.
CHAPTER 6. LIGHT FIELD ANALYSIS



Figure 6.5: Light field projection in the frequency domain. The image formation in the frequency domain is a slicing on the axis to which light field spectrum are duplicated by convolving with the aperture response.

6.2.2 Single-layer inverse prefiltering

For the purpose of vision-correcting displays, the observer is interested in seeing a 2D image within his or her focal range. For a defocused object, its light field is sheared and so is its frequency representation, as illustrated by Figure 6.5(b); 1D images at different depths correspond to differently sheared lines.

By combining Equations 6.5 and 6.6 with the Lambertian surface assumption, the Fourier transform of the retinal light field projection can be modeled as

$$\widehat{i}(\omega_x) = \widehat{l}\left(-\frac{d_e}{d_o}\omega_x\right)\operatorname{sinc}\left(rs\omega_x\right),\tag{6.7}$$

where $s = -d_e \Delta$ is the slope of the emitted light field in the frequency domain as defined by the defocus of the eye. Conceptually, this is a spreading centered at the sheared light field, as shown on the bottom of Figure 6.5(b); the perceived image spectrum is now the replica of the scaled light field spectrum at $\omega_u = 0$. Inverse prefiltering requires solving for the corresponding light field contents with respect to the scaled aperture spectrum. Note the replica obtained on the slicing axis by the aperture spreading is exactly the scaled modulation transfer function as we have described in Figure 4.7. Unfortunately, *sinc* functions contain a lot of zero-valued positions, which makes the solution of an inverse problem ill-posed. This is the case for correcting aberrations with conventional (Lambertian based) 2D screens.

6.2.3 Multilayer inverse prefiltering

The earlier multilayer method discussed in Chapter 5 proposed to remedy this ill-posedness by adding additional layers to the display, or by simulating virtual layers using a light field display. As illustrated in Figure 6.5(c), the image formation is changed to summing two differently-sheared lines in the frequency domain.

Generalizing Eq. 6.7 to multiple display layers results in the following frequency representation of the retinal projection:

$$\widehat{i}(\omega_x) = \sum_{n=1}^{N} \widehat{l}_n \left(-\frac{d_e}{d_n} \omega_x \right) \operatorname{sinc} \left(r s_n \omega_x \right),$$
(6.8)

where s_n is the shear of display layer n and \hat{l}_n is the light field emitted by each of the layers. The offsets between display layers are chosen so that the envelope of the differently sheared *sinc* functions has no zeros.

Conceptually, the position with zero-valued spreading is now covered by the other layers, as shown on the bottom of Figure 6.5(c); the spectrum of the target image is obtained without information loss, and light field spectrum at each layer is inversely redistributed by a contrast optimization algorithm. While this is conceptually effective, physical constraints of the display, including nonnegative pixel states and the limited dynamic range, still result in a severe loss of contrast in practice.

6.2.4 Light field predistortion

Pamplona et al. [2012] present a ray tracing solution to solve the inverse solution to Equation 3.5; each ray emitted from a retinal pixel maps to exactly one angular view of one spatial pixel on the light field display. This ray tracing method virtually presents a 2D image outside the physical device enclosure but within the focal range of the observer.

In the frequency analysis (Section 6.1.2), the method simulates a conventional in-focus display (see Fig. 6.3). Due to the practical limitations of currently available light field displays, showing a 2D image at a distance to the device enclosure requires an increased angular resolution of the display, as we have derived in Equation 6.4. Construction of light field displays, however, are subject to the spatio-angular resolution tradeoff: angular resolution can only be achieved by trading off spatial resolution. Hence, an in-focus object can be shown but at a significantly lower resolution than that of the display panels.

Frequency analysis in the light field domain reveals a more versatile formulation than what can be achieved with a naive ray tracing approach. As outlined by Equation 6.6, the underlying image formation from the emitted light field to the retinal projection can now be extended to prefiltering

CHAPTER 6. LIGHT FIELD ANALYSIS

using an infinite number of layers. Effectively, a light field display has the capability to display many different 2D planes at varying depths simultaneously, and interpreting Equation 6.6 under this aspect shows that a light field display generalizes Equation 6.8 as

$$\hat{i}(\omega_x) = \int_{\Omega_s} \hat{l}(\omega_x, s\omega_x) \operatorname{sinc}(rs\omega_x) \, ds.$$
(6.9)

While the above formulation generalizes multilayer prefiltering, the slopes occupy a double-wedge shaped space in the frequency domain; there are still some lower frequencies that are left unused, and a better utilization can improve both the perceived spatial resolution and the image contrast.

6.3 Summary

In summary, the presented frequency analysis reveals that a conventional display cannot present sharp imagery to an observer who cannot focus on the screen—the inversion problem is ill-posed due to zeros in the frequency domain.

While a multilayer approach is theoretically capable of converting this into a well-posed inversion problem, physical limitations in dynamic range and nonnegativity constraints of pixel states result in a significant loss of image contrast in practice. Pamplona et al.'s approach employs light field displays but uses a direct ray-traced solution, which does not exploit all available degrees of freedom in computation and, therefore, must make tradeoffs in the image resolution. Light field displays offer significantly more degrees of freedom compared to simple multilayer displays, but these are unlocked by solving the inverse light field projection problem, which we call "light field prefiltering".

With full generality, we apply prefiltering directly to the 4D light field, rather than a subset of possible light fields (i.e., those produced by synthetic multilayer displays). With added degrees of freedom, inverse prefiltering (or deconvolution) yields further benefits in image contrast and reduces the demand in resolution, as we will show in the next chapter.

Chapter 7

Light Field Based Inverse Prefiltering

Combining Equation 3.5 and 3.7, we obtain the projection of the retinal image in the spatial domain as the following:

$$i(x) = \int_{-\infty}^{\infty} l_e(x, u) A(u) du$$

=
$$\int_{-\infty}^{\infty} l\left(\mathbf{M}^{-1}(x, u)\right) A(u) du,$$
 (7.1)

where the retinal light field l_e is inversely transformed to the display side light field l. Discretizing the integral in Equation 7.1 results in a linear forward model:

$$\mathbf{i} = \mathbf{Pl} \tag{7.2}$$

where P is the projection from a 4D display light field l onto a 2D retinal image i.

The objective of an aberration-correcting display is to present a 4D light field to the observer that results in a desired 2D retinal projection. Assuming that viewing distance, pupil size, and other parameters are known, the emitted light field can be found by optimizing the following objective function:

Here, i is the target image and the constraints of the objective account for physically feasible pixel states of the screen. Equation 7.3 can be solved using standard non-negative linear solvers, we employ LBFGSB [Byrd *et al.*, 1995].

While Equation 7.3 allows for optimal display pixels states to be determined, natural questions that remain are "Which display type is best suited for aberration-correction? And what trade-off do we have to make?" We attempt to answer these questions with an analysis of the conditioning of the projection matrix **P** and contrast optimization.

7.1 Analysis of Projection Matrix

Figure 7.1 shows an example of a prefiltered light field with 3-by-3 views for an example scene. In this example, the different views contain overlapping parts of the target image, allowing for increased degrees of freedom for aberration compensation. Within each view, the image frequencies are amplified, which is the typical structure of single-layer prefiltering. When the views are optically projected onto the retina of an observer, all views are integrated, which results in a perceived image that has significantly improved sharpness, shown in Figure 7.1(c), as compared to an image observed on a conventional 2D display, shown in Figure 7.1(b).

We illustrate the concept by the flatland light field in Figure 7.1(d) and (e). In (d), the angular sampling resolution is too low and the overlapping regions cannot cover the entire image, and the uncovered regions suffer from single-layer prefiltering: singularities due to the zero-valued frequencies cause low contrast and frequency loss. In (e), with higher angular sampling, the overlapping regions occupy the entire image such that every single pixel receives more than one view; the singularities are remedied with added degree of freedom. The plot reveals one important fact to achieve better conditioning: "At least more than one view entering the pupil aperture is required."



Figure 7.1: Light field prefiltering with naive parameter setup. The proposed prefiltering approach computes a light field (here with 3-by-3 views) that results in a desired 2D projection on the retina of an observer. The yellow box overlapping region is the primary benefit of the method. In the flatland light field example, the yellow boxes illustrate why 4D light field prefiltering is more powerful than single-layer image prefiltering: a single region of the target image receives contributions from different views (e), which removes singularities present in conventional single-layer prefiltering (d) approaches and improves perceived image quality.



Figure 7.2: **Conditioning analysis.** The light field projection matrix corresponding to a defocused eye is ill-conditioned. With more angular resolution available in the emitted light field, more degrees of freedom are added to the system, resulting in lower condition numbers (lower is better). The condition number of the projection matrix is plotted for a varying defocus distance (kernel size) and angular resolution (number of light field views). We observe that even as few as 1.5 angular light field samples entering the pupil of an observer significantly decrease the condition number.

Conditioning Analysis. Given the minimum requirement on the light field angular sampling rate, it is worth noting the rank behavior of the projection matrix P in Equation 7.2. Consider the degenerate case of Lambertian surfaces where angular resolution is just one. For images shown on a 2D display outside the focal range of the observer, Equation 7.2 quickly reduced to an image filtering operation: the target pixel is given by weighted neighboring pixels. Solving for the pre-filtered light field simply requires the banded projection matrix P be inverted, which is similar to inverting the frequency domain spectrum of a point spread function; this is a rank deficient system, as we have discussed in Section 4.3. From Figure 4.6, most singular values are quite small, and the system is ill-conditioned. Inverting the corresponding linear system will not help the observer to see a sharper image.

Employing a 4D light field display instead adds degrees of freedom that become visible by plotting the condition number of the projection matrix **P**. Figure 7.2 shows the matrix conditioning for varying amounts of defocus and angular light field resolution (lower condition number is better). Increasing the angular resolution of the light field passing through the observer's pupil significantly decreases the condition number of the projection matrix for all amounts of defocus. This results in an interesting observation: "increasing the blur kernel size increases the condition number¹, but increasing the angular sampling rate does the opposite."

¹There are some slight fluctuations in the plot due to aliasing caused by the DFT discretization

The condition number drops significantly after it passes the 1.1 and 1.3 marks where the angular sampling enables more than one view to enter the pupil, effectively using available angular light field variation; this is expected from the previous observation that more than one view has to enter the pupil. As more than 1.7 and 1.9 light field views pass through the pupil, the condition number keeps decreasing but at a much slower rate. With an extreme number of views around 7 to 9, the system becomes the setup of light field predistortion: each ray hits exactly one retinal pixel, and the system becomes a diagonal identity matrix. However, as we have discussed in the previous chapter, the spatial-angular trade-off reduces the image resolution, and the trade-off might not be desirable. Our light field prefiltering method is located in between these two extremes. Usually, fewer than 2 views are required to maintain a sufficiently low condition number.

The conditioning analysis is a generalized insight into the effective modulation transfer function (Fig. 5.2) as in the winner-take-all multilayer display, where the minimum of the envelope response is maximized. As we mentioned in Section 4.3 where the frequency response of the MTF and the singular values are equivalent, the condition number is simply the inverse of the minimum response, and "*lowering the condition number is equivalent to maximizing the minimum envelope response*." In the convolution based image modeling, maximizing envelope response is restricted to a few layers by brute-force searching the minimum in a high dimensional tensor; the condition number analysis of the light field reveals only marginal benefits by increasing angular sampling rate passing the 1.7 mark.

Contrast optimization trade-off. In multilayer prefiltering, we developed a greedy contrast optimization (Sec. 5.2), which is non-intuitive and computationally expensive. A critical insight into the contrast optimization is how we handle the negative pixel values. In Equation 7.3 we can avoid such problems via non-negative least square solvers. At the defocus level shown in Figure 7.3 (a), naively applying the nonnegative constraint results in additional artifacts as shown in (b, top), and this is similar to the artifacts found in the Richardson-Lucy solver in Section 4.2.



Figure 7.3: Tradeoff between angular light field resolution and image contrast.

Solution to the inverse problem of Equation 7.2 can be done by using standard SVD solver (in Sec. 4.2, we show that this is equivalent to inverting the optical transfer function), which produces negative pixel values: we shift and renormalize the intensity afterward. Alternatively, we can shift and scale the target image before solving the system, effectively scaling the target image into the target range space of the projection matrix. Although this is a user-defined process, observed image quality can be enhanced. In particular, Equation 7.3 can be modified as:

argmin l
$$\|(\mathbf{i}+b)/(1+b) - \mathbf{Pl}\|^2$$

subject to $0 \le l \le 1$, (7.4)

where b is a user specified bias term that reduces the image contrast to 1/(b+1).

Similar to the conditioning analysis, we plot achieved image quality measured in PSNR for all contrast levels at various angular sampling rates on the left of Figure 7.3. We reconstruct a test image with different combinations of angular resolution and image contrast. With a conventional display that emits a light field with no angular variation, shown in Figure 7.3(b), we either obtain a low quality but high contrast image (b, top) or a high quality but low contrast image (Figure 7.3(b), bottom). At the same angular sampling rate, higher image contrast can be obtained with a lower quality or vice versa. With 1.5 or more views entering the pupil, shown in Figure 7.3(c), similar trade-offs are observed but overall reconstruction quality is significantly increased.

7.2 Prototype, Experiments and Results

The proposed aberration-correcting display can be implemented using most light field display technologies, including lenslet arrays [Lippmann, 1908] and multilayer designs [Wetzstein *et al.*, 2012]. For the purpose of this paper, we demonstrate the feasibility of our techniques with a parallax barrier display [Ives, 1903] because the required hardware is readily available. Nevertheless, the proposed displays are not limited to this particular architecture.

Prototype hardware. The prototype device is shown in Figure 7.4. A pinhole based parallax barrier mask is printed, with high contrast at a resolution of 5080 DPI, on a transparency with a Heidelberg Herkules imagesetter (www.pageworks.com). To optimize light throughput and avoid diffraction, the pinholes have a size of 75 microns each. This mask is mounted at an offset of 5.4 mm in front of a conventional 2D screen using a clear acrylic spacer. The screen is an Apple iPod touch 4th generation display with a pixel pitch of 78 microns (326 PPI). The dimensions of our prototype allow 1.66 light field views to enter a human pupil with a diameter of 6 mm at a distance of 25 cm. Higher-resolution panels are commercially available and could improve spatial and angular resolution of our prototype and facilitate larger viewing distances.



Figure 7.4: **Prototype parallax barrier light field display.** The barrier mask contains a pinhole array (left) that is mounted at a slight offset in front of an Apple iPod touch 4 screen (lower right). The display emits a light field with a high-enough angular resolution so that at least two viewing zones enter the pupil of a human observer. This effect is illustrated on the top right: multiple Arabic numerals are emitted in different viewing directions; the finite pupil size then creates an average of multiple different views on the retina (here simulated with a camera).

CHAPTER 7. LIGHT FIELD BASED INVERSE PREFILTERING

Software. The light field prefiltering algorithm is implemented in Matlab on a PC with a 2.7GHz 2-core CPU and 8GB of RAM. The light field projection matrix is precomputed in about 3 minutes with radiances sampling the pupil at 25 rays/mm, resulting in approximately 7800 effective rays per retinal pixel. We use the non-negative least squares solver package LBFGSB [Byrd et al., 1995] to solve Equation 7.4 in about 20 seconds for each image shown on the prototype. The projection matrix only needs to be computed once for each viewing distance and we believe that an optimized GPU implementation of the solver could achieve real-time frame rates in the future.

Photographs of Prototype. We show a variety of results captured from our prototype display using the optical setup in Figure 7.5. Photographs are captured with a DSLR camera equipped with a 50mm lens at f/8, simulating a 6mm human pupil. The display is placed at a cellphone reading distance of 25cm to the camera. The camera is focused at 45 cm, placing the screen 25 cm away from the focal plane. This setup closely resembles a 6.75D hyperopic eye requiring a +1.78D glasses to focus 20cm closer.



(d)prefiltered light field for examples shown in (b) and (c)

Figure 7.5: Optical experiment setup. In (a), we perform an experiment closely simulating a 6.75D hyperopic eye viewing a cellphone at 25cm away, where he need an +1.78D glasses focus. In (b), we include a battery in the scene to show how blurred it is in comparison with the sharp image in the display. The color banding is due to aliasing and can be removed using a transparent diffuser (e.g. diffuse screen protecting film). (c) is a comparison of the bunny without any correction and with our light field prefiltering correction. The corresponding prefiltered light field sources are shown in (d).



Figure 7.6: **Photographs of results from prototype display.** The hyperopic camera simulates a human pupil with a diameter of 6 mm at a distance of 25cm to the screen, while the camera is 45cm to the focus plane.

In Figure 7.6, we illustrate the effectiveness of the light field prefiltering method using two groups of examples: natural images and art work. Within each group, we first show the simulated correction shown on the first row, which is significantly sharper than that without correction shown on the second row. On the third row, we also compare with the "ray-traced" method proposed Pamplona et al. [2012] for the same display resolution and spatio-angular tradeoff. Our light field prefiltering on the fourth row allows for significantly increased resolutions without requiring the observer to wear glasses, and closely resembles simulations; the minor artifacts are due to imperfect display screen calibration.

Video results. Similar to the multilayer display results in Section 5.4, we also show a sequence of animation to compare the prototype results. Unlike what we did in the multilayer prefiltering, the contrast value b in Equation 7.4 is chosen to be a constant zero so that the contrast is 100%, and there is no need for temporal normalization.

Figure 7.7 shows a comparison of results without post-processing to remove the color banding artifacts². The first column shows the direct capture without any correction. Results obtained from Pamplona et al. [2012] are shown on the second column; as we have discussed earlier, there is not much improvement without enough angular resolution sampling. The third column shows the results obtained using light field prefiltering, and the images are much sharper than the previous methods. Note that since the prefiltering algorithm also reduces the dynamic range, the image appears dimmer; however, the contrast is not lost, and this dimmed image can be compensated by a brighter display panel.

Finally the corresponding source images are shown on the last two columns. The light field predistortion by Pamplona et al. [2012] is simply a resampling of the sheared light field; without enough angular resolution, the predistorted light field image (fourth column) looks very similar to a blurred image. On the other hand, the prefiltered light field image (fifth column) contains the amplified frequency information. Note that due to printing resolution constraints, we are only able to make 5-by-5 spatial angular trade-off; our method only requires 3-by-3, and the rest are left unused and the prefiltered image looks much darker.

²The color banding is due to the aliasing, and can be removed by placing a diffuser on top of the panel.



Figure 7.7: Video examples comparing light field prefiltering.

7.3 Evaluation

7.3.1 Visual Performance

We evaluate the quality achieved with the proposed approach using the four simulated examples shown in Figure 7.8. Similar to the prototype, we simulate a 10 inch tablet with a 300 PPI panel and a pinhole parallax barrier with 6.5 mm offset. The tablet is held at a distance of 30 cm and viewed with the eye requiring +6.75D hyperopic glasses; images are shown on the center of the display in a 10.8cm-by-10.8cm area. For each example, we compare our approach with multilayer prefiltering and the light field predistortion approach [Pamplona *et al.*, 2012]. The first two columns show the target images and the perceived images on a conventional display, respectively, when the physical location is outside the focal range of the eye.

Michaelson contrast metric. Prefiltering involves modulating the image content by enhancing weaker frequencies. Without utilizing the full degrees of freedom in the light field sense, the results obtained using multilayer prefiltering suffer from extreme contrast loss, here measured in Michelson contrast. Light field predistortion does not depend on content modifications but on resampling of the light field, so the contrast is not sacrificed.

By efficiently using all views, the proposed light field prefiltering approach restores contrast by a factor of 3 to 5 higher than that of the multilayer prefiltering. We note that the contrast from light field prefiltering is not quite as good as the predistortion algorithm, which always gives full contrast. However, when closely inspecting the image content, the light field predistortion always gives blurred images, which is due to insufficient resolution. Using the same hardware setup, prefiltering the light field enables sharper perceived images at the cost of slight contrast loss.

Perceptual Metric. To assess both contrast and sharpness, we resort to HDR-VDP2 [Mantiuk *et al.*, 2011], a perceptually-based image metric. The quality mean opinion score (QMOS) gives an evaluation of overall perceived image quality, and in most examples we score 2 to 3 times higher than light field predistortion.

The images in the third row are a particular difficult example for prefiltering based algorithms, because performance depends on the frequency content of the image which in this case does not allow prefiltering to achieve a higher quality. Lots of high frequencies in the example tend to reduce image contrast so that even our light field prefiltering scores slightly lower. Visually our result still looks sharp. In the last row of Figure 7.8, we show a probabilistic map on whether a human can detect per pixel differences for the fourth example. Clearly our result has a much lower detection rate.



Figure 7.8: **Evaluation and comparison to previous work.** We compare simulations of conventional and vision-correcting image displays qualitatively and quantitatively using contrast and quality-mean-opinion-square (QMOS) error metrics. A conventional out-of-focus display always appears blurred (second column). Multilayer displays with prefiltering improve image sharpness but at a much lower contrast (third column). Light field displays without prefiltering require high angular resolutions, hence have a low spatial resolution (fourth column). The proposed method combines prefiltering and light field display to optimize image contrast and sharpness (right column). The QMOS error metric is a perceptually linear metric, predicting perceived quality for a human observer. We also plot maps that illustrate the probability of an observer detecting the difference of a displayed image to the target image (bottom row). Our method performs best in most cases.

7.3.2 Kernel Size Performance

We also assess performance under different sizes of the blur kernel by moving the display further away from the focal range of the eye. Similar to the previous subsection, image contrast and the perceptual score are treated separately.

With the perceived image contrast, we only compare the prefiltering based framework, since light field predistortion always produces an image with full contrast. On the left of Figure 7.9, our light field prefiltering generates the same sharpness as multilayer prefiltering, but the contrast is always higher; we also include the single-layer prefiltering as a base reference. As the kernel size increases, the prefiltered image contrast degrades for both methods; at a distance where the PSF diameter is 11 pixels over a 256 pixels wide image, the remaining contrast is less than 20% for both image based prefiltering methods. Our light field based framework generates on average two to three times better contrast across all different sized kernels.

When making comparisons using the perceptual metric HDR-VDP2, we omit the image based methods since the metric significantly penalizes the contrast loss in the multilayer prefiltering. On the right of Figure 7.9, we compare with light field predistortion. Even with slight contrast loss, our method has higher score than that of the light field predistortion across all levels of blurring. To improve light field predistortion, even when we use a higher resolution panel with higher angular sampling (7-to-1 spatial-angular trade-off on a 456 PPI panel), the perceptual score only increases slightly.



Figure 7.9: **Comparing with prefiltering and light field based framework over a range.** Left: Image based prefiltering methods have poor contrast. At the same sharpness as multilayer pre-filtering, the light field prefiltering generates significantly higher image contrast. Right: the light field prefiltering also produces image with higher perceptual scores. The performance of light field predistortion only increases slowly over increased resolution on a higher density display panel.

7.3.3 Error in Calibrations.

Extending depth of field. One advantage of our approach is its ability to incorporate additional constraints, such as to provide extended depth of field. One of the concerns of precorrection for optical aberrations is that the algorithms are traditionally considered only for a fixed blur kernel or, equivalently, viewer distance. Although this can be dynamically tracked and incorporated into the inverse problem, we can alternatively account for a range of viewing distances simultaneously.

We account for variable viewing distances by stacking multiple light field projection matrices into Equation 7.4 with incremental defocus distances. The resulting system becomes overconstrained, so the solution tries to satisfy all viewing distances equally well, which results in slight image degradations as compared to a single viewing distance. Nevertheless, image quality for all other viewing distances within the predefined range is significantly increased as shown in Figure 7.10.



Figure 7.10: Accounting for a range of viewing distances. Top row: when considering a fixed viewing distance, defocus errors are compensated at that distance (top center) but image quality degrades when the observer moves (top left and right). The proposed method can account for a range of viewing distances (second row).

Off-axis movement. Another dimension is lateral movement off the axis; in this case, the head can move in a plane parallel to the display. The second row of Figure 7.11 shows that, without off-axis optimization, the perceived image is distorted due to the uncalibrated viewing condition. Similar to variable viewing distance optimization, we also stack several off-axis movement projection matrices, and obtain stable perceived images, as show on the bottom row of Figure 7.11.

We formally assess the performance of the off-axis viewing condition using the PSNR plot (top row of Figure 7.11). The performance of the perceived image without constraining the off-axis behavior is shown in blue, which peaks at zero movement; the PSNR degrades rapidly and significantly as the head moves just a bit, and reaches the minimum at 9.5mm. Since light field displays have repeated viewing zones, the PSNR curve is able to reach another peak at the 19mm mark; the difference between the maximum and minimum PSNR is 30dB, introducing an unacceptable viewing condition. When the display is optimized for off-axis viewing, the peak performance drops about 4dB, but the overall variation is small and smooth. Although the minimum also occurs at 9.5mm, it still achieves 27dB, as opposed to 5dB compared to that without optimization.



Figure 7.11: Similar to the extended depth of field, we can optimize for off-axis viewing.

7.4 Summary

In this chapter, we present a computational light field display approach to correcting for visual aberrations of a human observer. Our display architecture employs off-the shelf hardware components, such as printed masks or lenslet arrays, combined with computational light field prefiltering techniques.

Benefits and Limitations The proposed techniques offer significantly increased resolution and contrast compared to previously proposed aberration-correcting displays. Intuitively, light field pre-filtering minimizes demands on angular light field resolution, which directly results in higher spatial resolutions. For device implementations with lenslet arrays, the reduced angular resolution, compared to Pamplona et al.[2012], allows for shorter focal lengths of the employed lenslets, resulting in thinner form factors and easier fabrication. For implementations with parallax barriers, pinhole spacings are reduced, allowing for increased image brightness.

However, the proposed system requires modifications to conventional display hardware and increased computational resources. Our current implementation does not allow real-time frame rates. Although the proposed approach provides increased resolution and contrast as compared to previous approaches, achieving the full target image resolution and contrast is not currently possible. We evaluate all system parameters and demonstrate prototype results under conditions that realistically simulate a human pupil; however, we do not perform a user study. Slight artifacts are visible on the prototype, these are mainly due to imprecise device calibration. Finally, we only demonstrated successful compensation for defocus, and we will discuss correcting for higher-order aberrations in the next chapter.

Chapter 8

Light Field Wavefront Aberrometry

In the previous chapter, we have shown that a prefiltered light field can successfully compensate for the defocus aberrations in the observer's eye. It is interesting to see how the vision-correcting display behaves when higher-order aberrations are presented. We will formally introduce the standard decomposition of wavefront geometry into Zernike polynomials, and the ray tracing mechanism to support all optical aberrations.

8.1 Frequency Decomposition of Wavefront Geometry

Following Section 3.1.2, the wavefront map describes the phase of rays leaving the eye's pupil, and gives the new directions of propagation.

The wavefront map is simply a height field defined on a circular plane where $z = \omega(x^a, y^a) = \omega(rcos(\theta), rsin(\theta))$. Similar to the Fourier decomposition of a spherical function using spherical harmonics, the wavefront phase map $\omega()$ can be decomposed into a series of mutually orthogonal circular basis functions. Among many indexing variants, one that is similar to the spherical harmonic is called Fringe labeling and is used by [Wyant and Creath, 1992] [Ruoff and Totzeck, 2009] with the following:

$$\begin{aligned}
\omega(x^{a}, y^{a}) &= c_{0}^{0} Z_{0}^{0}(x^{a}, y^{a}) \\
&+ c_{1}^{0} Z_{1}^{0}(x^{a}, y^{a}) + c_{1}^{-1} Z_{1}^{-1}(x^{a}, y^{a}) + c_{1}^{+1} Z_{1}^{+1}(x^{a}, y^{a}) \\
&+ c_{2}^{0} Z_{2}^{0}(x^{a}, y^{a}) + c_{2}^{-1} Z_{2}^{-1}(x^{a}, y^{a}) + c_{2}^{+1} Z_{2}^{+1}(x^{a}, y^{a}) + c_{2}^{-2} Z_{2}^{-2}(x^{a}, y^{a}) + c_{2}^{+2} Z_{2}^{+2}(x^{a}, y^{a}) \\
&+ \cdots \\
&= \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} c_{n}^{m} Z_{n}^{m}(x^{a}, y^{a}) \end{aligned}$$
(8.1)

where n is the radial index, m is the azimuth index, c_n^m are the scaling weights, and $Z_n^m(x^a, y^a)$ are the Zernike polynomials, named after Nobel prize laureate Frits Zernike for his contribution in phase contrast microscope imaging. These Zernike polynomial basis functions are mutually orthogonal, and the frequencies increases with the radial index, as shown in Figure 8.1.

n	m	polynomial	name
	0	$2r^2 - 1$	defocus
1	+1	$rcos(\theta)$	tilt
	-1	$rsin(\theta)$	tilt
	0	$6r^4 - 6r^2 + 1$	spherical
	+1	$(3r^2-2)rcos(\theta)$	coma
2	-1	$(3r^2-2)rsin(\theta)$	coma
	+2	$r^2 cos(2\theta)$	astigmatism
	-2	$r^2 sin(2\theta)$	astigmatism

Figure 8.1: Wyant's ordering of Zerike polynomial.

Since the normalized series of the basis functions forms an orthonormal basis family, the corresponding scaling coefficients can be obtained by projecting the original wavefront map onto the Zernike polynomials and vice versa.

Standardized Ordering and Naming. Early Zernike polynomials have many indexing and ordering schemes, and from 1999 to 2004, the VSIA task force [Thibos *et al.*, 2000] from the Optical Society of America (OSA) and ANSI made a standard Z80.28 to describe the ordering and meaningful naming for each term, such as defocus and astigmatism, as shown in Figure 8.2.

Under the standard scheme, the Zernike polynomials are defined as:

$$Z_n^m(r,\theta) = \begin{cases} +N_n^m R_n^{|m|}(r) cos(m\theta) & \text{for } m \ge 0, \\ -N_n^m R_n^{|m|}(r) sin(m\theta) & \text{for } m < 0 \end{cases}$$
(8.2)

In this definition, N_n^m is a normalization scaling constant, and is given by the following:

$$N_n^m = \begin{cases} \sqrt{(n+1)} & \text{for } m = 0, \\ \sqrt{2(n+1)} & \text{for } m \neq 0. \end{cases}$$
(8.3)

The radial polynomial function $R_n^{|m|}(r)$ is given by

$$R_n^{|m|} = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! (\frac{(n+m)}{2} - s)! (\frac{(n-m)}{2} - s)!} r^{n-2s} = \sum_{s=0}^{(n-|m|)/2} (-1)^s \binom{n-s}{s} \binom{n-2s}{\frac{(n-m)}{2} - s} r^{n-2s}$$
(8.4)



Figure 8.2: ANSI standard Zernike ordering.

Single Indexing. It is sometimes useful to refer to a specific Zernike polynomial using just a single index, or describe the expansion coefficients as a vector. Without a specific standard, ordering can be arbitrary and introduce confusion. The standard [Thibos *et al.*, 2000] also defines the single indexing scheme by using the double indices n and m to be:

$$j = \frac{n(n+2) + m}{2}.$$
(8.5)

The single indexing scheme follows the top to down and left to right ordering of the pyramid in Figure 8.2.

Function Representation. Given the Zernike basis family, it is often desired to succinctly represent the original circular function by a few coefficients; this can be written as

$$\omega(x^a, y^a) \approx \sum_{j=1}^N c_j Z_j(x^a, y^a).$$
(8.6)

With sufficiently a large number of terms, N, the original wavefront phase map can be represented faithfully. In fact, since the measurement of the wavefront using a lenslet based sensor has finite resolution, the wavefront map is usually fitted using Zernike polynomials directly; in addition, the projection onto the basis functions also allows us to write the analytic expressions of the measured wavefront, which is useful, as we will discuss in the next section.

8.2 Ray-tracing Wavefront Geometry.

So far in the thesis, we have only considered the thin lens model of the eye, which limits the ability to model more complicated blurs, such as astigmatism or spherical aberration; those are traditionally treated analytically or through convolution using wavefront generated point spread function, as discussed in Section 3.1.2. Since we want to correct for higher order terms, it is essential to integrate the wavefront geometry into the light field prefiltering architecture. In Section 3.2, we describe the transformations of the light field as a propagation, followed by a refraction due to the eye. To model the wavefront geometry, the only part that is modified is the refraction; now replaced with the wavefront geometry.

The wavefront surface describes the directions of rays leaving the pupil at the same phase, and we only need to propagate these rays back to the display and determine the intersections. Figure 8.3 illustrates the simplified two-step processes: (a) generate rays whose directions are obtained using the wavefront surface, and then (b) propagate to the display. When the directions are normal to the pupil plane, they focus at infinity; when the directions are normal to a spherical wavefront, they focus at the center. However, generating directions due to an irregular wavefront surface is not straightforward, and a numerical solution is time consuming and subject to limited resolution. Fortunately, each term of the Zernike polynomials is a differentiable smooth function, and we can leverage the decomposable nature to obtain the directions at sampled locations, and combine normals for the entire expansion series.

For each Zernike polynomial function, we calculate the normal directions by using tangents at the sampled locations. We first uniformly sample locations (x^a, y^a) at the pupil plane; this directly generates the positions part of the light field coming out of the pupil (Fig. 8.3 (c)). The normal is obtained by forming the cross product of the tangent \mathbf{T}_j^x along the *x*-direction with \mathbf{T}_j^y along the *y*-direction:

$$\mathbf{N}_{j}(x^{a}, y^{a}) = \mathbf{T}_{j}^{x}(x^{a}, y^{a}) \times \mathbf{T}_{j}^{y}(x^{a}, y^{a}) = \begin{bmatrix} 1\\0\\dx \end{bmatrix} \times \begin{bmatrix} 0\\1\\dy \end{bmatrix} = \begin{bmatrix} -dx\\-dy\\1 \end{bmatrix}.$$
(8.7)



Figure 8.3: Ray-tracing wavefront

The analytic forms of the tangent functions are calculated by directly differentiating the Zernike polynomials in Cartesian coordinates. Some terms are shown in Figure 8.4. Figure 8.5 shows the 3D surfaces of tilt, astigmatism, trefoil, and their corresponding derivative functions.

j	n	m	Z_n^m in Cartesian coordinate	dZ_n^m/dx	dZ_n^m/dy
0	0	0	1	0	0
1	1	-1	2 <i>y</i>	0	2
2		+1	2x	2	0
3	2	-2	$2\sqrt{6}xy$	$2\sqrt{6}y$	$2\sqrt{6}x$
4		0	$2\sqrt{3}(x^2+y^2)-\sqrt{3}$	$4\sqrt{3}x$	$4\sqrt{3}y$
5		+2	$\sqrt{6}(x^2-y^2)$	$2\sqrt{6}x$	$-2\sqrt{6}y$
6	3	-3	$\sqrt{8}(3x^2y - y^3)$	$6\sqrt{8}xy$	$\sqrt{8}(3x^2 - 3y^2)$
7		-1	$\sqrt{8}(3x^2y + 3y^3 - 2y)$	$6\sqrt{8}xy$	$\sqrt{8}(3x^2 + 9y^2 - 2)$
8		+1	$\sqrt{8}(3xy^2 + 3x^3 - 2x)$	$\sqrt{8}(9x^2 + 3y^2 - 2)$	$6\sqrt{8}xy$
9		+3	$\sqrt{8}(x^3 - 3xy^2)$	$\sqrt{8}(3x^2 - 3y^2)$	$-6\sqrt{8}xy$
10	4	-4	$\sqrt{10}(4x^3y - 4xy^3)$	$\sqrt{10}(12x^2y - 4y^3)$	$\sqrt{10}(4x^3 - 12xy^2)$
11		-2	$\sqrt{10}(8x^3y + 8xy^3 - 6xy)$	$\sqrt{10}(24x^2y + 8y^3 - 6y)$	$\sqrt{10}(8x^3 + 24xy^2 - 6x)$
12		0	$\sqrt{5}(x^2 + y^2)^2 - 6\sqrt{5}(x^2 + y^2) + \sqrt{5}$	$\sqrt{5}(24x^3 + 24xy^2 - 12x)$	$\sqrt{5}(24y^3 + 24x^2y - 12y)$
13		+2	$\sqrt{10}(x^2 - y^2)(x^2 + 7y^2)$	$\sqrt{10}(16x^3 - 6x)$	$\sqrt{10}(-16y^3+6y)$
14		+4	$\sqrt{10}(x^4 - 6x^2y^2 + y^4)$	$\sqrt{10}(4x^3 - 12xy^2)$	$-\sqrt{10}(12x^2y - 4y^3)$

Figure 8.4: Derivatives

The normal direction on the original wavefront surface is the sum of the normals from all Zernike terms, and the expression is given by:

$$\mathbf{N}(x^a, y^a) = \sum_{j=1}^N c_j \mathbf{N}_j(x^a, y^a),$$
(8.8)

Finally, we can combine the positions with directions to find the 4D light field coming out of the pupil: $(x^a, y^a, \mathbf{N}_1, \mathbf{N}_2)^T$.

Limitations Although we have introduced wavefront based light field prefiltering, it is just another tool when compared to the traditional Fourier transformed point spread function (Sec. 3.1.2). Light field modeling is usually done in the regime of geometric optics, i.e. the particle nature of the light; wave effects like diffraction and interference are not modeled. The point spread functions generated from ray-tracing based and Fourier optics based are visually different; the user has to select a usage scenario such as coherent/incoherent light source and diffraction due to small pupil. Extensions to include wave optics have been briefly studied in [Oh *et al.*, 2008] [Raskar *et al.*, 2009] [Oh *et al.*, 2010], and we anticipate future research to unify the architecture.



Figure 8.5: Derivative of Zernike polynomials

The required computation is another factor to consider; sampling a 4D function is much more expensive than transforming a 2D complex image. The quality of the ray-traced solution depends on the sampling rate; while there are many variance reduction techniques in the rendering literature, the application to wavefront aberrations is still unclear, and a large number of angular samples are still required.

8.3 Correcting Zernike Polynomials Aberrations

In this section, we formally use light field prefiltering to correct for aberrations other than purely the defocus term. As we have shown in Section 3.1.2, higher-order terms cause the point spread function to have a non-circular shape; in fact, the shape of the point spread function is also depth-dependent, and traditional single-layer image prefiltering is not very suitable for this application.

We will first show cases by using astigmatism to describe what causes the problem. The depthdependency is solved using the concept from the extended depth of field concept in Section 7.3.3, and then we will further show results for other aberrations.

8.3.1 Astigmatism

Astigmatism in an optical system refers to the case where the two axes have different focusing power. In the Zernike polynomial, astigmatism is denoted by Z_2^{-2} and Z_2^{+2} , as shown in Figure 8.2, where the two axes are 90 degrees to each other. A graphical illustrations of the optical arrangement is shown in Figure 8.6 for Z_2^{+2} ; in this example, the vertical focusing plane is farther away than that of the horizontal axis, and objects are always blurred as there is no zero distortion spot. As we further investigate this case, objects at the near plane have a vertical blur, and it becomes a horizontal blur at the far plane in Figure 8.6. The perceived images for Z_2^{-2} , a 45 degrees rotation from Z_2^{+2} , are illustrated in Figure 8.7, where the focus plane sweeps from one end to another. When an object is closer to the observer, the point spread function is squashed and oriented 45 degree, as shown in the top left corner on the first row. As the object moves further away, the point spread function is restored to a circular shape, and then squashed into an ellipsoid oriented in the other direction.



Figure 8.6: **Optical arrangement to illustrate astigmatism.** Astigmatism has two different focus planes that cause a depth-dependent point spread function, which is an ellipsoid when objects sit away from the center between focus planes.



Figure 8.7: **Examples showing astigmatism has depth-dependent blur.** The PSF is an ellipsoid at one end on the optical axis; it gradually changes to a circle, and then deforms into another ellipsoid on the other end. The image is thus blurred non-uniformly depending on the location where it is shown.

Correcting for Depth-dependent Blur. The depth-dependent nature, like astigmatism, introduces non-uniform point spread functions, and is generally hard to correct with traditional singlelayer prefiltering methods. Similar to what we have introduced in Section 7.3.3 to support extended depth of field, we can optimize for an extended range to support the depth-dependent distortion.

In the center of Figure 8.8, we show a matrix of images at the locations where we optimized for versus at locations from which images are viewed. It is not surprising that the performance of the prefiltering excels when the locations we optimized for match where the images are viewed; the perceived images for this case are shown in the diagonal red boxes. However, the off-diagonal cases perform badly; even when we choose to optimize for the compromised center location, the circular point spread function cannot account for ellipsoid kernel (as shown Fig. 8.7) occurring at either end of the focus plane.

Without losing image quality and using the same hardware component, we can optimize for extended depth of field with depth-dependent non-uniform point spread functions. Similar to what we have introduced in Section 7.3.3, multiple projection matrices for different focusing planes are stacked into a giant set of linear equations, and we solve the over-constrained system in a consistent way as we solve for a single location. We generate projection matrices starting from the center location with a circular kernel, and expand to both ends with differently oriented ellipsoids.



Figure 8.8: **Correcting for astigmatism.** Direct optimization at a single location does not scale well for all other locations, and the depth-dependent blur badly distorts the artifacts. Applying the same principle for the extended depth of field, the perceived images are much more stable across different viewing locations.

In the right column of Figure 8.8, we show the perceived images viewed from different locations, so that the display sits at different locations on the the optical axis. The direct comparison to the naive optimization without extended depth of field shows slight degradation and blur around the edges (take the last row for example), but the overall quality is maintained. This concludes how we deal with general depth-dependent point spread functions with extended depth of field.

8.3.2 Higher Order Terms

Correcting higher order aberrations is like what we have introduced before, but the sampling can be more expensive. Due to the non-uniform shape of the Zernike polynomials wavefront, higher sampling of rays and time are required. However, once the precomputed sampling is done, prefiltering the light field can be computed in a few seconds. We will show two categories of blur using Zernike polynomials.

The first category uses only a single term of the Zernikes, such as spherical, coma, trefoil, and tetrafoil, as shown in Figure 8.9. In the first column, the defocus blur based on circle of confusion is shown as a reference. The spherical aberration (2nd column) shows a different blurring structure, with a stronger contour around the edges. Coma (3rd column) displays a significant "smear" downward, and image details are lost due to strong spreading. The trefoil term produces rounded triangles, such as the eyes and the trumpet; tetrafoil, on the other hand, gives a very different blur around the eye even though they differ by just one additional leg.

Correcting these kinds of blur with traditional single-layer prefiltering methods is hard, as shown in the second row of Figure 8.9. We found it particularly difficult to invert point spread functions with strong spreading, like spherical aberration and coma. However, not all higher order terms are difficult; correcting for trefoil and tetrafoil aberrations is much more successful. The results is not surprising, since the latter preserves more high frequency information in the direction orthogonal to the legs. Finally, the light field based architecture generates results with much better quality, even for the very difficult coma case.

Finally in Figure 8.10, we show results for randomly generated higher order aberrations. These are in general quite impossible for single-layer prefiltering (red box), but the light field prefiltering (blue box) still excels at these cases.



Figure 8.9: Examples of single term Zernike polynomial point spread functions. Here we show comparisons of spherical, coma, trefoil, and tetrafoil with the defocus term.



Figure 8.10: Correcting for randomly generated higher order aberrations.

8.3.3 Future Work for Higher Order Aberrations

Importance sampling and closed form solution. For the current software implementation, we did a brute force ray tracing, which is computationally intensive; for each retinal pixel, we sample around 22000 rays across the pupil aperture on a regular grid, and the computation requires many hours on our Matlab implementation. This is not a smart strategy; even the simplest stratified sampling can improve the variances. However, the deeper insight we look for is how rays actually converge or diverge through the wavefront geometry. Fortunately, the Zernike frequency decomposition of the wavefront geometry is a set of well-known polynomial functions, and we believe a smarter importance sampling strategy can be developed.

Another way to accelerate the computation is to derive the wave propagation. Since the Zernike wavefront basis functions are mathematically well defined, their propagations to the display could be derived in a closed form solution as opposed to the expensive ray tracing; the angular variation on the display is simply the derivative of the propagated wavefront function. Applying the Huygens-Fresnel principle, the propagation could be derived using Fresnel diffraction at far field. Since the Fresnel diffraction only considers up to quadratic terms and we need to deal with the high frequency waves, the validity in these approximations should be re-examinated.

Experiments with higher order terms. In simulation, we have shown that higher order aberrations are possible to correct with light field prefiltering. However, physical experiments still need to be implemented in a way that the lens generates higher order terms, and this is a challenging optical setup. The main challenge comes from the optical lens with the desired higher order aberrations. The current technology generates higher order aberrations through deformable mirrors or phase plates, but they are not easy to integrate into our optical setup. Also, since an aspherical lens is required to generate a pure defocus term to focus at the display, a commercial camera lens usually consists of several lenses, which obscure the optical center to augment the desired higher order terms. To the author's knowledge, there is no lens that comes with known higher order aberrations, thus custom lenses have to be built.

Ideally, we could still construct a lens with higher order aberrations with a phase plate. One possible solution is placing the phase plate in front of the lens, and we back propagate the wavefront from the phase plate to the hypothetical optical center, which involves some trials and errors. After finding the propagated wavefront Zernike polynomial coefficients C'_j , the defocus term can be added to the new coefficients as one "hypothetical" lens with higher order aberrations. It is still unclear how to find the back propagation and reconstruct the coefficients; the Fresnel diffraction gives only approximated solutions, and the ray tracing is subject to the sampling and resolution problems.

The other possible optical setup is shown in Figure 8.11, where multiple lenses are used to replicate the virtual phase plate at the main lens. In this particular setup, the rays from the display follow the green paths to the image sensor or the eye. Rays starting from the the phase plate follow



Figure 8.11: **Physical setup to generate higher order aberrations.** The phase plate (in blue) used to create higher order aberrations is placed in the center that generates a virtual phase plate on the main lens in front of the image sensor.

the blue paths that converge on the main lens in front of the sensor. Since every point forms a virtual correspondence on the main lens, we can say it creates the exact wavefront we require to verify the correction for higher order aberrations; this optical setup provides future directions to implement physical experiments.

Chapter 9 Future Work

In this thesis, we demonstrate the possibility of correcting optical aberrations through computational light field displays. This is an emerging field of research, not just for computational light field displays, but also for correcting optical aberrations through computation. Although we show that the introduction of computation for optical correction potentially offers many promises, there are also challenges. We will discuss the future work with respect to different criteria: practicality, application, and computational optics. The discussions will give the reader ideas on what challenges are faced by the vision-correcting displays, what are the potential applications that use similar principles, and what is the future of correcting aberrations using algorithmic solutions.

9.1 Practicality.

Resolution enhancement. The major challenge for vision correcting displays is to exceed the performance of optical solutions. Since there is a sacrifice of the spatial resolution in favor of the angular resolution for the light field display, we also want to ask: "Is the spatial-angular resolution trade-off reasonable and acceptable?" We have reduced the requirement of high angular resolution by Pamplona et al. [2012], but a reduction by half or a third is still considered a high loss in the resolution, where the resolution is reduced from a "4K resolution" display to a "Full-HD" or just "HD" display. An angular sampling rate of 1.5 and above is now still considered very high for commercial 3D TVs or light field displays, where each view spans a few degrees; in our prototype, each view spans less than 1 degree.

The high angular sampling rate, in fact, favors the correction for hyperopia and presbyopia, where people want to be able to read the devices at a close distance. Myopic patients would like to be able to see objects at a further distance, and high pupil angular sampling rate is harder to achieve when the light field display is further away. Wetzstein et al. [2012] demonstrate the possibility of constructing a high resolution light field display using the compressive nature of the light field content, but it is unknown how well that can be applied to vision correcting displays.

However, perhaps the most important part is the spatial acuity of the eye; we are limited by the spatial resolution of the display, which is much lower than the capability of the human eye. The retinal sampling is able to discriminate about 60 cycles per degree, and to achieve such high spatial resolution on the light field display would be an incredible engineering challenge. Pamplona et al. [2012] use a high density panel, around 2000 pixel per inch, to generate high angular resolution, but optical challenges such as diffraction and other wave effects also arise.

Sensitivity to the parameter space. Many of our results are based on the assumption that perfect calibrations are achieved; however in practice, the viewer's head can move, and the distance to the display can change. There is also the high frequency scanning and movement of the eye, called saccades, that will cause the retinal light field and image to be unstable. The human eye is a wonderful optical device for imaging, and it responds to the environment, such as illumination changes and accommodation changes. Overall, this parameter space has many dimensions, and even though we introduce the idea of using an over-constrained solver to incorporate an extended depth of field, developing a lab prototype to commercial products would still require a significant effort.

Correcting for Binocular Vision Although the primary context discussed in this thesis is mainly for correcting aberrations with the same prescription for both eye, the light field based solution is a perfect candidate for this binocular purpose, as it was used for stereoscopic display. The binocular architecture requires sending different images to different eyes, thus the constrained optimization shown in Section. 7.3.3 now solves the light fields for two projection locations: the left eye and the right eye. The added constraints require doubling the underlying panel resolution of the light field display; the resolution requirement is now from two pixels to four pixels. However, it could be further reduced to three by compacting the two views with shared pixels if two eyes have similar prescription. Unfortunately, correcting for binocular vision complicates the parameter spaces as it makes the solution space highly nonlinear, and further investigation to make it a practical product should consider the intersection of both problems.

Mobile computing platform. Ultimately, our ideal target platform would be the mobile device with close reading distance. The computational challenges are in two parts: the sampling of projection matrices and the computation of the non-negative solutions. Even though in most cases, the sampling of projection matrices can be precomputed, the parameter space is huge, and real-time matrix sampling is preferred. The computation for non-negative solutions involves non-convex optimization, but in practice, our prototype results are generated in a few seconds. The required time and space increase with higher spatial resolution, but recent advances in parallel computing (e.g. using GPU) show promising directions.

9.2 Applications

HMD, HUD, Augmented Reality. For many cases, a head-mount-display requires users to wear their own eyeglasses, which can be cumbersome and uncomfortable. Pre-correcting for such a display can improve the willingness to use such devices. We also found applications where the user needs to focus at more than one place can benefit from this research, and applications like heads-up-displays are a potentially rich field in which many appliances can use the method to increase user experiences, such as projecting information on windshields or helmets. Another concrete example is augmented-reality eye-glasses, where information is always projected onto a fixed virtual plane; the prefiltering can take the plane to whichever depth the user is focusing, and the world is always "augmented" with the information.

Light field displays. There are some human psychophysical factors that are also important, especially with the common 3D TVs/theaters that show two separate images to the eyes; prolonged viewing can cause dizziness due to the differences in accommodation and vergence cues. We feel that combining the true light field with real depth of field to support accommodation and image prefiltering techniques should be able to enhance the viewing experiences. In the long run, we believe that flexible display architectures will allow for multiple different modes, such as glasses-free 3D image display, vision corrected 2D image display, and combinations of vision-corrected and 3D image display.

Optical lithography. Lithography is the process of printing, and optical lithography is now widely used to print circuits. With advances in semiconductors, the features become extremely tiny and the diffraction effect comes into play. Research in photo-lithography often finds printing patterns that will become the desired image after the diffraction, and this is exactly the inverse prefiltering problem we have discussed earlier.

Health care. Quality of life is an important indicator of civilization, but unfortunately the loss of vision is an inevitable process as we age. In fact, elder people with presbyopic eye are the biggest potential market for handheld vision-correcting mobile display, since the eye cannot focus at near range, and the displays tend to be small. We hope that, in the future, the vision-correcting display can help people with uncorrected eyes to see sharp images; they either use this as an alternative besides eyeglasses, or for people with higher order aberrations can start to see display in sharp focus. In the study by Katz et al. [1997], it is shown that myopia has a positive correlation with education level, and this may be due to prolonged near range viewing; perhaps by showing virtual content at further distances thus may relax the eye for the "long range reading" experience, but this has yet to be proven to work even with psychophysical experiments; We feel that "vision rectification" might be another interesting direction.

Cryptography Hiding information by convolving an image with a complex point spread function is another interesting direction to explore, where the viewer is required to wear the inverse decrypting lens to decipher the information. As opposed to initial goal of the thesis, we want to design a complex point spread function composed of arbitrary higher order aberrations, and decrypting requires a specifically designed phase plate to inversely cancel the prefiltering of the image; this forms a secret key paradigm for data encryption. The encryption can be made stronger by enforcing the point spread function to be the composition of the person's aberrations and the phase plate.

9.3 Computational Optical Correction.

Correcting optical aberrations through computation is a new research area that has not been explored, and we have shown that it is possible to compensate for the aberrations by sacrificing some contrast. This capability is also content dependent: for a constant image nothing will be sacrificed. But the prefiltering might struggle with a white noise image; comprehensive theoretical analysis is an interesting direction. Although there is a plethora of research in computational optics stemming from electrical engineering or optical engineering, little of this research is from computer science and uses computation to solve similar problems. One immediate direction is to explore wave effects, and the pioneering work by Ou et al. [2010] shows promising results, although applying their principles on prototypes or products will still require significant effort.
Chapter 10 Conclusion

A large number of people have eye problems, and most of them are due to uncorrected refractive error. Studies show some refractive errors like myopia are correlated with higher levels of education, and presbyopia increases with age. Higher order aberrations are traditionally very difficult to correct since aspherical lenses are required. The need to correct for optical aberrations has a long history, and has been primarily addressed purely by optical means. Although refined optical elements give better image quality, this is fixed hardware directly changing the optics of the eye, and can sometimes be intrusive/invasive. In this thesis, we introduce a "computation-based" display that corrects the optical aberrations using computer programs; this is done by reversing the projection process in our eye, and showing the preprocessed contents on a computer display, such as a multilayer display or a light field display.

In a multilayer display, we analyze the frequency domain behavior of the convolution kernel, and show that the zero and extremely attenuated frequencies are covered by another layer; image contrast can also be further optimized via fractional frequency assignment. We have built a multilayer LCDs prototype to show the effectiveness of the multilayer prefiltering idea, and we have demonstrated the removal of ringing artifacts and the enhancement in image contrast. We also have shown that a traditional light field display can be used to emulate a virtual multilayer display, and with frequency analysis, our virtual multilayer prefiltering outperforms prior work using the same light field display.

To account for better light field display utilization, we found that directly prefiltering the full 4D light field gives better results. We analyze the image formation in the light field domain, and demonstrate a light field prefiltering framework. When compared with prior work, the hardware to achieve the required angular sampling rate is shown to be much simpler, and the image contrast is also improved significantly by using a non-negative linear solver with the light field architecture. We have demonstrated the construction of the light field display prototype using off-the-shelf components such as an iPod Touch and a printed pinhole mask, and with a thin form factor, we can successfully correct the defocus blur. Evaluations are done using Michaelson contrast and a perceptual metric, called HDR-VDP2, and light field prefiltering finds a good balance between image contrast and perceptual quality.

Traditional image based methods are unable to compensate for higher order aberrations; we have shown that higher order terms can be corrected in the same way as defocus. The sampling of the projection matrix is done by inverse ray tracing, with the directions of the rays given by forming the cross product of the tangents of the Zernike polynomials. The result shows that not all higher order terms are difficult to correct; e.g. trefoil and tetrafoil preserve more high frequency information, making the overall prefiltering easier. Randomly generated point spread functions with higher order aberrations are also corrected with our method in simulation, and the results are promising.

Finally, we discussed several future directions to enhance the vision-correcting display, several potential applications that share similar ideas as the "deconvolution" idea, and also possible topics to further extend the capability of the generalized prefiltering architecture.

Bibliography

- [Adelson and Bergen, 1991] Edward H. Adelson and James R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991.
- [Agrawal *et al.*, 2009] Amit Agrawal, Yi Xu, and Ramesh Raskar. Invertible motion blur in video. *ACM Trans. Graph.*, 28(3):95:1–95:8, 2009.
- [Akeley *et al.*, 2004] Kurt Akeley, Simon J. Watt, Ahna Reza Girshick, and Martin S. Banks. A stereo display prototype with multiple focal distances. *ACM Trans. Graph.*, 23(3):804–813, 2004.
- [Allen and Ulichney, 2005] Will Allen and Robert Ulichney. 47.4: Invited paper: Wobulation: Doubling the addressed resolution of projection displays. *SID Symposium Digest of Technical Papers*, 36(1):1514–1517, 2005.
- [Alonso Jr. and Barreto, 2003] M. Alonso Jr. and A. B. Barreto. Pre-compensation for high-order aberrations of the human eye using on-screen image deconvolution. In *IEEE Engineering in Medicine and Biology Society*, volume 1, pages 556–559, 2003.
- [Archand *et al.*, 2011] Pierre Archand, Eric Pite, Herve Guillemet, and Loic Trocme. Systems and methods for rendering a display to compensate for a viewer's visual impairment. International Patent Application PCT/US2011/039993, 2011.
- [Artal *et al.*, 2004] Pablo Artal, Li Chen, Enrique J. Fernandez, Ben Singer, Silvestre Manzanera, and David R. Williams. Neural compensation for the eyes optical aberrations. *Journal of Vision*, 4(4), 2004.
- [Babcock, 1953] H. W. Babcock. The Possibility of Compensating Astronomical Seeing. *Pub. Astron. Soc. Pacific*, 65, October 1953.
- [Bao et al., 2009] Jinhua Bao, Rongrong Le, Jiangxiu Wu, Yeyu Shen, Fan Lu, and Ji C. He. Higher-order wavefront aberrations for populations of young emmetropes and myopes. *Journal* of Optometry, 2(1):51 – 58, 2009.

- [Bell et al., 2008] Gareth P. Bell, Robert Craig, Robert Paxton, Gordon Wong, and Duncan Galbraith. Beyond flat panels: Multi-layered displays with real depth. SID Digest, 39(1):352–355, 2008.
- [Berthouzoz and Fattal, 2012] Floraine Berthouzoz and Raanan Fattal. Resolution enhancement by vibrating displays. *ACM Trans. Graph.*, 31(2):15:1–15:14, April 2012.
- [Bousseau, 2008] Adrien Bousseau. Virtual glasses: The myopic revenge. 2008.
- [Brown et al., 2006] Michael S. Brown, Peng Song, and Tat-Jen Cham. Image pre-conditioning for out-of-focus projector blur. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1956–1963, 2006.
- [Byrd et al., 1995] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput., 16(5):1190– 1208, September 1995.
- [Campisi and Egiazarian, 2007] Patrizio Campisi and Karen Egiazarian. *Blind Image Deconvolution: Theory and Applications.* CRC Press, 2007.
- [Chai *et al.*, 2000] Jin-Xiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum. Plenoptic sampling. In *ACM SIGGRAPH*, pages 307–318, 2000.
- [Cossairt et al., 2007] Oliver S. Cossairt, Joshua Napoli, Samuel L. Hill, Rick K. Dorval, and Gregg E. Favalora. Occlusion-capable multiview volumetric three-dimensional display. Applied Optics, 46(8):1244–1250, 2007.
- [Damera-venkata and Chang, 2007] Niranjan Damera-venkata and Nelson L. Chang. On the resolution limits of superimposed projection. In *In Proc. IEEE International Conference on Image Processing (ICIP, 2007.*)
- [Damera-Venkata and Chang, 2009] Niranjan Damera-Venkata and Nelson L. Chang. Display supersampling. *ACM Trans. Graph.*, 28:9:1–9:19, February 2009.
- [Didyk et al., 2010] Piotr Didyk, Elmar Eisemann, Tobias Ritschel, Karol Myszkowski, and Hans-Peter Seidel. Apparent display resolution enhancement for moving images. In ACM SIGGRAPH 2010 papers, SIGGRAPH '10, pages 113:1–113:8, New York, NY, USA, 2010. ACM.
- [Durand *et al.*, 2005] Frédo Durand, Nicolas Holzschuch, Cyril Soler, Eric Chan, and François X. Sillion. A frequency analysis of light transport. In *ACM SIGGRAPH*, pages 1115–1126, 2005.
- [Egan et al., 2009] Kevin Egan, Yu-Ting Tseng, Nicolas Holzschuch, Frédo Durand, and Ravi Ramamoorthi. Frequency analysis and sheared reconstruction for rendering motion blur. In ACM SIGGRAPH 2009 papers, SIGGRAPH '09, pages 93:1–93:13, New York, NY, USA, 2009. ACM.

- [Egan et al., 2011] Kevin Egan, Florian Hecht, Frédo Durand, and Ravi Ramamoorthi. Frequency analysis and sheared filtering for shadow light fields of complex occluders. ACM Transactions on Graphics, 30(2):1–13, April 2011.
- [Fergus et al., 2006] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. In ACM SIGGRAPH 2006 Papers, SIGGRAPH '06, pages 787–794, New York, NY, USA, 2006. ACM.
- [Gonzalez and Woods, 1992] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (Third Edition)*. Addison-Wesley, 1992.
- [Goodman, 2004] Joseph W. Goodman. *Introduction to Fourier Optics (Third Edition)*. Roberts & Company Publishers, 2004.
- [Görrn *et al.*, 2006] P. Görrn, M. Sander, J. Meyer, M. Kröger, E. Becker, H.-H. Johannes, W. Kowalsky, and T. Riedl. Towards see-through displays: Fully transparent thin-film transistors driving transparent organic light-emitting diodes. *Advanced Materials*, 18(6):738–741, 2006.
- [Gortler *et al.*, 1996] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pages 43–54, New York, NY, USA, 1996. ACM.
- [Gotoda, 2012] Hironobu Gotoda. Implementation and analysis of an autostereoscopic display using multiple liquid crystal layers. In *SPIE Stereoscopic Displays and Applications XXIII*, volume 8288, 2012.
- [Grosse *et al.*, 2010] Max Grosse, Gordon Wetzstein, Anselm Grundhöfer, and Oliver Bimber. Coded aperture projection. *ACM Trans. Graph.*, 2010.
- [Hartwig and Atchison, 2012] A. Hartwig and D.A. Atchison. Analysis of higher order aberrations in a large clinical population. *Invest Ophthalmol Vis Sci*, 2012.
- [Hofer *et al.*, 2005] Heidi Hofer, Ben Singer, and David R. Williams. Different sensations from cones with the same photopigment. *Journal of Vision*, 5(5), 2005.
- [Holroyd *et al.*, 2011] Michael Holroyd, Ilya Baran, Jason Lawrence, and Wojciech Matusik. Computing and fabricating multilayer models. *ACM Trans. Graph.*, 30(6):187:1–187:8, 2011.
- [Huang and Barsky, 2011] Fu-Chung Huang and Brian Barsky. A framework for aberration compensated displays. Tech. Rep. UCB/EECS-2011-162, University of California, Berkeley, December, 2011.
- [Huang et al., 2012] Fu-Chung Huang, Douglas Lanman, Brian A. Barsky, and Ramesh Raskar. Correcting for optical aberrations using multilayer displays. ACM Trans. Graph., 31(6):185:1– 185:12, November 2012.

- [Huang et al., 2013] Fu-Chung Huang, Gordon Wetzstein, Brian A. Barsky, and Ramesh Raskar. Computational light field display for correcting visual aberrations. In ACM SIGGRAPH 2013 Posters, SIGGRAPH '13, pages 40:1–40:1, New York, NY, USA, 2013. ACM.
- [Isaksen et al., 2000] Aaron Isaksen, Leonard McMillan, and Steven J. Gortler. Dynamically reparameterized light fields. In Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00, pages 297–306, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.
- [Ives, 1903] Frederic E. Ives. Parallax stereogram and process of making same. U.S. Patent 725,567, 1903.
- [Jones et al., 2007] Andrew Jones, Ian McDowall, Hideshi Yamada, Mark Bolas, and Paul Debevec. Rendering for an interactive 360° light field display. ACM Trans. Graph., 26:40:1–40:10, 2007.
- [Katz et al., 1997] J. Katz, J.M. Tielsch, and A. Sommer. Prevalence and risk factors for refractive errors in an adult inner city population. *Investigative Ophthalmology and Visual Science*, 38(2):334–40, 1997.
- [Kaufman and Alm, 2002] P.L. Kaufman and A. Alm. Adler's Physiology of the Eye (Tenth Edition). Mosby, 2002.
- [Kee *et al.*, 2011] Eric Kee, Sylvain Paris, Simon Chen, and Jue Wang. Modeling and removing spatially-varying optical blur. In *IEEE International Conference on Computational Photogra-phy*, 2011.
- [Krachmer *et al.*, 2005] J.H. Krachmer, M.J. Mannis, and E.J. Holland. *Cornea*. Elsevier Mosby, 2005.
- [Kugler and Wang, 2010] Lance J. Kugler and Ming X. Wang. Lasers in refractive surgery: history, present, and future. *Appl. Opt.*, 49(25):F1–F9, September 2010.
- [Lanman *et al.*, 2010] Douglas Lanman, Matthew Hirsch, Yunhee Kim, and Ramesh Raskar. Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization. *ACM Trans. Graph.*, 29:163:1–163:10, 2010.
- [Lanman *et al.*, 2011] Douglas Lanman, Gordon Wetzstein, Matthew Hirsch, Wolfgang Heidrich, and Ramesh Raskar. Polarization fields: dynamic light field display using multi-layer LCDs. *ACM Trans. Graph.*, 30(6), 2011.
- [Legras *et al.*, 2004] Richard Legras, Nicolas Chateau, and W Neil Charman. A method for simulation of foveal vision during wear of corrective lenses. *Optom Vis Sci*, 81(9):729–38, 2004.
- [Levin *et al.*, 2007] Anat Levin, Rob Fergus, Frédo Durand, and William T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3), 2007.

- [Levin et al., 2009] Anat Levin, Samuel W. Hasinoff, Paul Green, Frédo Durand, and William T. Freeman. 4D frequency analysis of computational cameras for depth of field extension. ACM Trans. Graph., 28(3), 2009.
- [Levoy and Hanrahan, 1996] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proceedings* of the 23rd annual conference on Computer graphics and interactive techniques, SIGGRAPH '96, pages 31–42, New York, NY, USA, 1996. ACM.
- [Liang et al., 1997] J. Liang, D.R. Williams, and D.T. Miller. Supernormal vision and high-resolution retinal imaging through adaptive optics. J Opt Soc Am A Opt Image Sci Vis, 14(11):2884–92, 1997.
- [Liang et al., 2008] Chia-Kai Liang, Tai-Hsu Lin, Bing-Yi Wong, Chi Liu, and Homer H. Chen. Programmable aperture photography: multiplexed light field acquisition. ACM Trans. Graph., 27(3):55:1–55:10, August 2008.
- [Liang *et al.*, 2011] Chia-Kai Liang, Yi-Chang Shih, and H.H. Chen. Light field analysis for modeling image formation. *Image Processing, IEEE Transactions on*, 20(2):446–460, feb. 2011.
- [Lin *et al.*, 2004] L.L.K. Lin, Y.F. Shih, C.K. Hsiao, and C.J. Chen. Prevalence of myopia in taiwanese schoolchildren: 1983 to 2000. *Ann Acad Med Singapore*, 33(1):27–33, 2004.
- [Lippmann, 1908] Gabriel Lippmann. Épreuves réversibles donnant la sensation du relief. *Journal* of *Physics*, 7(4):821–825, 1908.
- [Lucy, 1974] L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745+, June 1974.
- [Mantiuk *et al.*, 2011] Rafat Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. Hdr-vdp-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions. In *Proc. ACM SIGGRAPH*, pages 40:1–40:14, 2011.
- [Marwah *et al.*, 2013] Kshitij Marwah, Gordon Wetzstein, Yosuke Bando, and Ramesh Raskar. Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Trans. Graph.*, 32(4):46:1–46:12, July 2013.
- [McMillan and Bishop, 1995] Leonard McMillan and Gary Bishop. Plenoptic modeling: an image-based rendering system. In *Proceedings of the 22nd annual conference on Computer* graphics and interactive techniques, SIGGRAPH '95, pages 39–46, New York, NY, USA, 1995. ACM.
- [MedlinePlus, 2008] MedlinePlus. How to keep your sight for life. *MedlinePlus*, 3(3):12, 2008.
- [Mehta *et al.*, 2012] Soham Uday Mehta, Brandon Wang, and Ravi Ramamoorthi. Axis-aligned filtering for interactive sampled soft shadows. *ACM Trans. Graph.*, 31(6):163:1–163:10, November 2012.

- [Mohammadpour *et al.*, 2012] Soheil Mohammadpour, Alireza Mehri Dehnavi, Hossein Rabbani, and Vasudevan Lakshminarayanan. A pre-compensation algorithm for different optical aberrations using an enhanced wiener filter and edge tapering. In *ISSPA'12*, pages 935–939, 2012.
- [Ng, 2005] Ren Ng. Fourier slice photography. In *ACM SIGGRAPH 2005 Papers*, SIGGRAPH '05, pages 735–744, New York, NY, USA, 2005. ACM.
- [Oh *et al.*, 2008] Se Baek Oh, George Barbastathis, and Ramesh Raskar. Augmenting light field to model wave optics effects. *Tech. Report*, abs/0907.1545, 2008.
- [Oh *et al.*, 2010] Se Baek Oh, Sriram Kashyap, Rohit Garg, Sharat Chandran, and Ramesh Raskar. Rendering wave effects with augmented light field. *Comput. Graph. Forum*, 29(2):507–516, 2010.
- [Olshausen and Field, 1996] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, June 1996.
- [Oyamada and Saito, 2007] Yuji Oyamada and Hideo Saito. Focal pre-correction of projected image for deblurring screen image. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [Pamplona et al., 2010] Vitor F. Pamplona, Ankit Mohan, Manuel M. Oliveira, and Ramesh Raskar. Netra: interactive display for estimating refractive errors and focal range. ACM Trans. Graph. (SIGGRAPH), 29:77:1–77:8, 2010.
- [Pamplona et al., 2011] Vitor F. Pamplona, Erick B. Passos, Jan Zizka, Manuel M. Oliveira, Everett Lawson, Esteban Clua, and Ramesh Raskar. Catra: interactive measuring and modeling of cataracts. ACM Trans. Graph. (SIGGRAPH), 30(4):47:1–47:8, 2011.
- [Pamplona et al., 2012] Vitor F. Pamplona, Manuel M. Oliveira, Daniel G. Aliaga, and Ramesh Raskar. Tailored displays to compensate for visual aberrations. ACM Trans. Graph., 31(4):81:1– 81:12, July 2012.
- [Perwass and Wietzke, 2012] Christian Perwass and Lennart Wietzke. Single lens 3d-camera with extended depth-of-field. In *Proc. SPIE*, volume 8291, pages 829108–829108–15, 2012.
- [Platt and Shack, 2001] B. C. Platt and R. Shack. History and principles of Shack-Hartmann wavefront sensing. *Journal of refractive surgery*, 17(5), October 2001.
- [Rajan et al., 1995] U Rajan, FT Tan, and TK Chan. Increasing prevalence of myopia in singapore school children. Proceedings of the Fifth International Conference on Myopia, pages 41–46, 1995.
- [Raskar *et al.*, 2006] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. Graph.*, 25(3):795–804, 2006.

- [Raskar et al., 2009] Ramesh Raskar, Se Baek Oh, Anthony Accardi, and Zhengyun Zhang. Light field: Present and future. CVPR2009 Course, 2009.
- [Resnikoff *et al.*, 2008] Serge Resnikoff, Donatella Pascolini, Silvio P. Mariotti, and Gopal P. Pokharel. Global magnitude of visual impairment caused by uncorrected refractive errors in 2004. *Bulletin of the World Health Organization*, 86(1):63–70, 2008.
- [Richardson, 1972] Hadley W. Richardson. Bayesian-Based Iterative Method of Image Restoration. *Journal of the Optical Society of America*, 62(1):55–59, January 1972.
- [Roorda *et al.*, 2002] Austin Roorda, Fernando Romero-Borja, William Donnelly, Hope Queener, Thomas Hebert, and Melanie Campbell. Adaptive optics scanning laser ophthalmoscopy. *Optics Express*, 10(9):405–412, May 2002.
- [Roorda, 2011] Austin Roorda. Adaptive optics for studying visual function: A comprehensive review. *Journal of Vision*, 11(5), 2011.
- [Rossi and Roorda, 2009] Ethan A Rossi and Austin Roorda. The relationship between visual resolution and cone spacing in the human fovea, 2009.
- [Ruoff and Totzeck, 2009] Johannes Ruoff and Michael Totzeck. Orientation zernike polynomials: a useful way to describe the polarization effects of optical imaging systems. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 8(3):031404–031404–22, 2009.
- [Sabesan and Yoon, 2010] Ramkumar Sabesan and Geunyoung Yoon. Neural compensation for long-term asymmetric optical blur to improve visual performance in keratoconic eyes. *Investigative Ophthalmology and Visual Science*, 2010.
- [Sajadi et al., 2012] Behzad Sajadi, M. Gopi, and Aditi Majumder. Edge-guided resolution enhancement in projectors via optical pixel sharing. ACM Trans. Graph., 31(4):79:1–79:122, July 2012.
- [Shack and Platt, 1971] R.V. Shack and B.C Platt. Production and use of a lenticular hartmann screen. *The Journal of the Optical Society of America*, pages 656–660, 1971.
- [Shan et al., 2008] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. In ACM SIGGRAPH 2008 papers, SIGGRAPH '08, pages 73:1–73:10, New York, NY, USA, 2008. ACM.
- [Takaki, 2006] Y. Takaki. High-Density Directional Display for Generating Natural Three-Dimensional Images. *Proc. IEEE*, 94(3), 2006.
- [Takashima *et al.*, 2001] T. Takashima, T. Yokoyama, S. Futagami, K. Ohno-Matsui, H. Tanaka, T. Tokoro, and M. Mochizuki. The quality of life in patients with pathologic myopia. *Jpn J Ophthalmol*, 45(1):84–92, 2001.

- [Thibos et al., 2000] Larry Thibos, Raymond A. Applegate, James T. Schwiegerling, and Robert Webb. Standards for reporting the optical aberrations of eyes. In Vision Science and its Applications, page SuC1. Optical Society of America, 2000.
- [Thibos *et al.*, 2002] Larry N. Thibos, Arthur Bradley, and Xin Hong. A statistical model of the aberration structure of normal, well-corrected eyes. *Ophthalmic and Physiological Optics*, 22(5):427–433, 2002.
- [Urey *et al.*, 2011] H. Urey, K.V. Chellappan, E. Erden, and P. Surman. State of the art in stereoscopic and autostereoscopic displays. *Proceedings of the IEEE*, 99(4):540–555, 2011.
- [Veeraraghavan *et al.*, 2007] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.*, 26(3), 2007.
- [Vitale et al., 2009] Susan Vitale, Robert D. Sperduto, and Frederick L. Ferris, III. Increased prevalence of myopia in the United States between 1971-1972 and 1999-2004. Arch. Ophthalmology, 127(12):1632–1639, 2009.
- [Wetzstein *et al.*, 2011] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4), 2011.
- [Wetzstein *et al.*, 2012] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph*, 2012.
- [Wiener, 1964] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series.* The MIT Press, 1964.
- [Wong et al., 2000] T Y Wong, P J Foster, J Hee, T P Ng, J M Tielsch, S J Chew, G J Johnson, and S K Seah. Prevalence and risk factors for refractive errors in adult chinese in singapore. *Invest Ophthalmol Vis Sci*, 41(9):2486–94, 2000.
- [Wyant and Creath, 1992] James C. Wyant and Katherine Creath. Basic wavefront aberration theory for optical metrology. *Applied Optics and Optical Engineering*, 11:1–53, 1992.
- [Yang et al., 2010] Q. Yang, D.W. Arathorn, P. Tiruveedhula, C.R. Vogel, and A. Roorda. Design of an integrated hardware interface for aoslo image capture and cone-targeted stimulus delivery. *Opt Express*, 18(17):17841–58, 2010.
- [Yellott and Yellott, 2007] John I. Yellott and John W. Yellott. Correcting spurious resolution in defocused images. *Proc. SPIE*, 6492, 2007.
- [Yuan et al., 2008] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Progressive interscale and intra-scale non-blind image deconvolution. In ACM SIGGRAPH 2008 papers, SIG-GRAPH '08, pages 74:1–74:10, New York, NY, USA, 2008. ACM.

- [Zhang and Nayar, 2006] Li Zhang and Shree Nayar. Projection defocus analysis for scene capture and image display. *ACM Trans. Graph.*, 25(3), 2006.
- [Zomet and Nayar, 2006] A. Zomet and S.K. Nayar. Lensless imaging with a controllable aperture. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2006.
- [Zwicker *et al.*, 2006] Matthias Zwicker, Wojciech Matusik, Frédo Durand, and Hanspeter Pfister. Antialiasing for automultiscopic 3D displays. In *Eurographics Symposium on Rendering*, 2006.