

Correcting for Optical Aberrations using Multilayer Displays

Fu-Chung Huang¹

¹Computer Science Division, UC Berkeley

Douglas Lanman^{2,4}

²MIT Media Lab

Brian A. Barsky^{1,3}

Ramesh Raskar²

³School of Optometry, UC Berkeley

⁴NVIDIA Research

watch face corrected for presbyopia



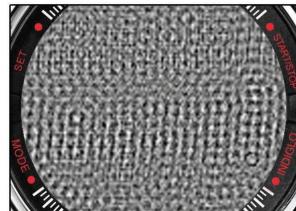
displayed images



perceived images



single-layer pre-filtering



two-layer pre-filtering

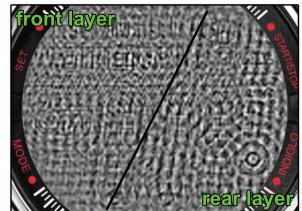


Figure 1: Correcting presbyopia using multilayer displays. A presbyopic individual observes a watch at a distance of 45 cm. The watch appears out of focus due to the limited range of accommodation. To read the watch, corrective eyewear (e.g., bifocals) must be worn with a +2.50 diopter spherical lens. (Left) As a substitute for eyewear, the watch can be modified to use a multilayer display containing two semi-transparent, light-emitting panels. The images displayed on these layers are pre-filtered such that the watch face appears in focus when viewed by the defocused eye. (Right) From left to right along the bottom row: the perceived image using a conventional display (e.g., an unmodified LCD), using prior single-layer pre-filtering methods, and using the proposed multilayer pre-filtering method. Corresponding images of the watch face are shown along the top row. Two-layer pre-filtering, while increasing the watch thickness by 6 mm, enhances contrast and eliminates ringing artifacts, as compared to prior single-layer pre-filtering methods. (Watch image © Timex Group USA, Inc.)

Abstract

Optical aberrations of the human eye are currently corrected using eyeglasses, contact lenses, or surgery. We describe a fourth option: modifying the composition of displayed content such that the perceived image appears in focus, after passing through an eye with known optical defects. Prior approaches synthesize pre-filtered images by deconvolving the content by the point spread function of the aberrated eye. Such methods have not led to practical applications, due to severely reduced contrast and ringing artifacts. We address these limitations by introducing multilayer pre-filtering, implemented using stacks of semi-transparent, light-emitting layers. By optimizing the layer positions and the partition of spatial frequencies between layers, contrast is improved and ringing artifacts are eliminated. We assess design constraints for multilayer displays; autostereoscopic light field displays are identified as a preferred, thin form factor architecture, allowing synthetic layers to be displaced in response to viewer movement and refractive errors. We assess the benefits of multilayer pre-filtering versus prior light field pre-distortion methods, showing pre-filtering works within the constraints of current display resolutions. We conclude by analyzing benefits and limitations using a prototype multilayer LCD.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms; I.4.4 [Image Processing and Computer Vision]: Restoration—Inverse filtering; K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities

Keywords: optical aberrations, deconvolution, pre-filtering, pre-compensation, multilayer displays, tailored displays, light fields, assistive devices

Links: DL PDF WEB VIDEO DATA

1 Introduction

Recent studies indicate an increasing prevalence of refractive errors: Vitale et al. [2009] found the incidence of myopia increased from 25.0% to 41.6% in the United States between 1971–1972 and 1999–2004. Today, individuals requiring correction have three options: eyeglasses, contact lenses, or refractive surgery. Ordinary eyeglasses can only correct lower-order aberrations (i.e., defocus and astigmatism). Higher-order aberrations, such as those induced by disorders including keratoconus or pellucid marginal degeneration, can be difficult to correct and are currently addressed using contact lenses or surgery. We describe a fourth option: modifying the composition of displayed imagery, as well as the underlying display hardware, to correct optical aberrations without eyewear or invasive surgery.

Our multilayer technique for the correction of optical aberrations using pre-filtering builds upon the work of Huang and Barsky [2011] and addresses visual artifacts occurring with earlier pre-filtering approaches proposed by Alonso and Barreto [2003].

ACM Reference Format

Huang, F., Lanman, D., Barsky, B., Raskar, R. 2012. Correcting for Optical Aberrations using Multilayer Displays. *ACM Trans. Graph.* 31, 6, Article 185 (November 2012), 12 pages. DOI = 10.1145/2366145.2366204 <http://doi.acm.org/10.1145/2366145.2366204>

Copyright Notice

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation.

Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or permissions@acm.org.

© 2012 ACM 0730-0301/12/11-ART185 \$15.00 DOI 10.1145/2366145.2366204 <http://doi.acm.org/10.1145/2366145.2366204>

and by Yellott and Yellott [2007]. These papers propose pre-filtering displayed imagery such that, when viewed by an aberrated eye, the perceived image appears in focus. The input image is first deconvolved by the point spread function, estimated from the viewer's refractive error (i.e., their optical prescription). As shown in Figure 1, such *single-layer pre-filtering* methods enhance the perceived image; however, two limitations have precluded practical applications: the perceived images exhibit ringing artifacts and severely reduced contrast.

1.1 Contributions

We address the limitations of single-layer pre-filtering by introducing the use of *multilayer displays* paired with a *multilayer pre-filtering* algorithm; such displays comprise stacks of semi-transparent, light-emitting panels (e.g., liquid crystal displays or organic light-emitting diodes). Our contributions include:

- By optimizing the separation between display layers, we demonstrate that multilayer pre-filtering preserves all spatial frequencies in the received image, eliminating the ringing artifacts appearing with prior single-layer pre-filtering methods.
- By optimizing the partition of spatial frequencies between layers, we show that multilayer pre-filtering increases image contrast, relative to single-layer pre-filtering.
- We describe design constraints, identifying light field displays as a preferred architecture; for such displays, we analyze the maximum resolution enhancement achievable with multilayer pre-filtering and light field pre-distortion methods.
- By performing simulations and experiments using a prototype multilayer LCD, we analyze the benefits and limitations of multilayer pre-filtering, including contrast enhancement and sensitivity to prescription and viewing parameters.

1.2 Overview of Benefits and Limitations

Multilayer pre-filtering not only corrects common lower-order aberrations, including defocus and astigmatism, but also has the potential to address higher-order aberrations, including coma. Multilayer pre-filtering provides two benefits over existing single-layer pre-filtering: enhanced image contrast and elimination of ringing artifacts. However, multilayer pre-filtering comes at the expense of added components and computational complexity, requiring two or more layers and additional operations to maximize contrast.

Multilayer pre-filtering also requires a display that is optically equivalent to a stack of semi-transparent, light-emitting layers. Implementation with physical layers (e.g., OLEDs or LCDs) increases the display thickness (typically no more than a few centimeters for moderate defocus or astigmatism). Ideal layer separations are dependent on both the refractive error and position of the viewer, with the latter requiring viewer tracking. To support binocular correction (i.e., different prescriptions for each eye), an autostereoscopic multilayer display is required, capable of delivering independent images to each eye. We identify existing autostereoscopic light field displays as a preferred architecture for meeting these design constraints. Such displays naturally support binocular viewing. Most significantly, virtual display layers can be synthesized beyond the display surface and in response to viewer movement, enabling thin form factors which are appropriate for mobile applications. However, such light field displays often reduce the spatial resolution of the received image, requiring an underlying high-resolution panel.

2 Related Work

Our approach builds on prior work in three areas: deconvolution methods for correcting camera and projector defocus, the construction of all-pass optical filters in computational photography, and emerging multilayer display architectures.

2.1 Deconvolution Methods

Image Restoration is applied to estimate an undistorted image given a received image degraded by camera shake, defocus, or object motion. The received 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, Wiener filtering, and the iterative Richardson-Lucy algorithm [Gonzalez and Woods 1992]. Recent developments in image deconvolution include exploiting natural image priors [Levin et al. 2007] and increasingly focus on blind deconvolution [Campisi and Egiazarian 2007], wherein the PSF is not known *a priori*.

Correcting Optical Aberrations of the Eye requires applying deconvolution before the image is displayed, rather than after it is received. This discrepancy profoundly impacts the quality of received images; as derived in Section 3, a pre-filtered image typically includes both negative and positive values of equal amplitude. Following Alonso and Barreto [2003] and Yellott and Yellott [2007], pre-filtered images must be normalized to the dynamic range of the display, resulting in a severe loss in contrast. In a closely related work, Huang and Barsky [2011] introduce multilayer pre-filtering and evaluate its performance via simulations. Recently, Archand et al. [2011] apply single-layer pre-filtering to commercial displays.

Correcting Projector Defocus can be achieved by applying pre-filtering to deconvolve the projected image by the projector's PSF. Brown et al. [2006] demonstrate extended depth of field projection using pre-filtering. Similar to correcting optical aberrations of the eye, pre-filtering 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. While 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 this work, we similarly seek to produce an *all-pass filter* by introducing a second display layer for correcting optical aberrations.

2.2 All-Pass Filtering in Computational Photography

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 pre-filtering is inspired by these publications, with the goal of incorporating additional layers to ensure all spatial frequencies are preserved in the received image.

2.3 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 beam splitters 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, respectively. 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. As described in Section 5, we employ a similar architecture. However, time multiplexing enables the multilayer LCD to operate in a mode that is optically equivalent to the required stack of light-emitting, rather than light-attenuating, layers. Multilayer LCDs have also found applications in computational photography [Zomet and Nayar 2006]. We demonstrate a new application for autostereoscopic displays: in addition to depicting 3D scenes, such displays are ideally suited for correcting optical aberrations.

3 Aberration-Correcting Multilayer Displays

This section describes how optical aberrations can be corrected by pre-filtering content for conventional single-layer displays or emerging multilayer displays. Section 3.1 assesses single-layer pre-filtering. In Section 3.2, we extend pre-filtering to multilayer displays comprising stacks of semi-transparent, light-emitting panels. Prior single-layer pre-filtering methods result in severely reduced contrast and image artifacts; in Section 3.3, we demonstrate how multilayer pre-filtering mitigates these limitations—providing a practical means for correcting for optical aberrations at the display device, rather than in front of the imaging apparatus.

3.1 Single-Layer Displays

3.1.1 Pre-filtering

Consider an imaging apparatus (e.g., a camera or an eye) located in front of a planar display (e.g., an LCD panel). In the following analysis we model the imaging apparatus as a linear shift-invariant (LSI) system [Gonzalez and Woods 1992]. The image $i(x, y)$, formed in the plane of the display, is approximated such that

$$i(x, y) = s(x, y) * p(x, y), \quad (1)$$

where $s(x, y)$ is the displayed radiant emittance profile, $p(x, y)$ is the point spread function (PSF), and $*$ is the convolution operator. The cumulative effect of optical aberrations is fully characterized, under this model, by the point spread function. As introduced by Alonso and Barreto [2003], an undistorted image $\tilde{i}(x, y)$ can be formed by displaying a pre-filtered image $\tilde{s}(x, y)$ such that

$$\tilde{s}(x, y) = s(x, y) * p^{-1}(x, y), \quad (2)$$

where $p^{-1}(x, y)$ is the inverse point spread function: defined such that $p^{-1}(\cdot) * p(\cdot) = \delta(\cdot)$, where $\delta(\cdot)$ is the Dirac delta function.

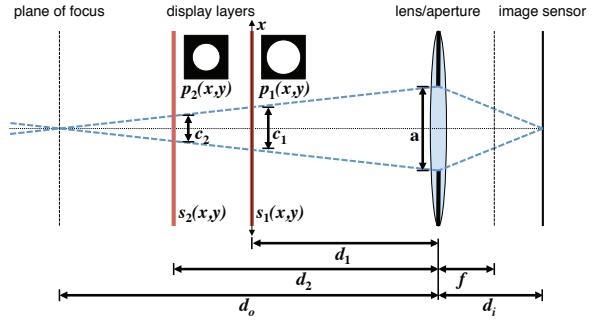


Figure 2: A defocused camera observing a multilayer display. We model a simplified camera containing a thin lens with focal length f and aperture diameter a . It observes a two-layer display, with layers located at distances d_1 and d_2 in front of the lens. When focused at a distance of d_o , the images of the display layers are defocused, resulting in point spread functions $p_1(x, y)$ and $p_2(x, y)$ with circles of confusion of diameter c_1 and c_2 , respectively.

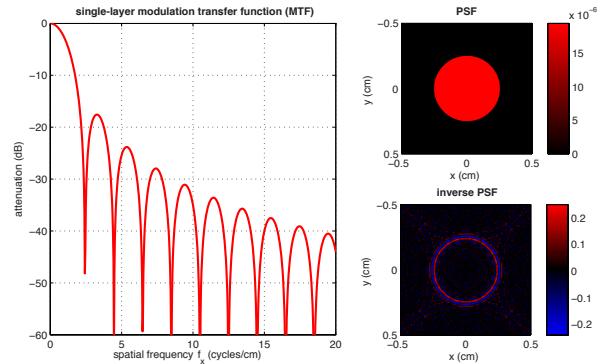


Figure 3: Modulation transfer function with a single-layer display. The human eye is modeled as a camera, following Figure 2, with $f = 17$ mm and $a = 4$ mm. A single-layer display is separated by a distance of $d = 35$ cm, with the eye focused at $d_o = 40$ cm. (Left) The MTF acts as a low-pass filter. Note the zeros (i.e., nulls) of the MTF correspond to frequencies that cannot be depicted in the received image $i(x, y)$ given by Equation 1. (Right) The resulting PSF and inverse PSF are shown at the top and bottom, respectively. Negative values in the inverse PSF result in negative values in single-layer pre-filtered images $\tilde{s}(x, y)$, requiring normalization via Equation 10—causing a significant loss of contrast in Figure 4.

Substituting Equation 2 into Equation 1 yields the following expression for the received image $\tilde{i}(x, y)$ using pre-filtering.

$$\tilde{i}(x, y) = [s(x, y) * p^{-1}(x, y)] * p(x, y) = s(x, y) \quad (3)$$

In summary, single-layer pre-filtering allows an undistorted image $\tilde{i}(x, y)$ to be formed by displaying the pre-filtering image $\tilde{s}(x, y)$, found by deconvolving the target image $s(x, y)$ by the PSF $p(x, y)$.

3.1.2 Frequency-Domain Analysis of Pre-filtering

Correcting for optical aberrations in this manner requires that the pre-filtered image $\tilde{s}(x, y)$ be non-negative, since the display only emits light with positive radiance. However, in practice, the inverse PSF $p^{-1}(x, y)$ often has the form of a high-pass filter, yielding both negative and positive values in the pre-filtered image [Yellott and Yellott 2007]. As a deconvolution method, the limitations of pre-filtering can be characterized through a frequency-domain analysis.

Taking the two-dimensional Fourier transform of Equation 1 yields the following relationship:

$$I(f_x, f_y) = S(f_x, f_y)P(f_x, f_y), \quad (4)$$

where $I(f_x, f_y)$ and $S(f_x, f_y)$ are the spectra of the received and displayed images, respectively, $P(f_x, f_y)$ denotes the optical transfer function (OTF), and f_x and f_y are the spatial frequencies along the x and y axes, respectively. Similarly, the spectrum of the single-layer pre-filtered image $\tilde{S}(f_x, f_y)$ is given by

$$\tilde{S}(f_x, f_y) = S(f_x, f_y)P^{-1}(f_x, f_y). \quad (5)$$

As described in Section 2.1, deconvolution algorithms can be applied to estimate the inverse optical transfer function $P^{-1}(f_x, f_y)$. For correcting optical aberrations, the target image $s(x, y)$ is free of noise; as a result, direct inverse filtering can be applied. In practice, this approach significantly expands the dynamic range of the pre-filtered image, leading to reduced contrast. As an alternative, we follow a similar approach to Brown et al. [2006] and Oyamada et al. [2007] and apply Wiener deconvolution, such that

$$P^{-1}(f_x, f_y) \approx \frac{1}{P(f_x, f_y)} \left(\frac{|P(f_x, f_y)|^2}{|P(f_x, f_y)|^2 + K} \right), \quad (6)$$

where K denotes the inverse of the signal-to-noise ratio (SNR), effectively serving as a regularization parameter in this application. By adjusting K , the dynamic range of the pre-filtered image can be reduced in comparison to direct inverse filtering.

Equations 5 and 6 reveal the first limitation of single-layer pre-filtering: **the modulation transfer function of the aberrated imaging apparatus must not have zeros**; spatial frequencies at these nulls cannot be preserved in the received image $\tilde{s}(x, y)$.

3.1.3 Analysis of Pre-filtering for a Defocused Camera

Consider the camera in Figure 2, separated a distance d from a single-layer display and composed of a thin lens with focal length f and aperture diameter a . The sensor and display are centered on the optical axis. By the *Gaussian thin lens equation*, the sensor is located a distance d_i behind the lens, such that a focused image is formed of the object plane, located a distance d_o in front of the lens.

A defocused camera (i.e., one for which $d_o \neq d$) records a blurred image of the display, as modeled by Equation 1. Under the geometrical optics approximation [Goodman 2004], the point spread function is a uniform disk with unit area, given by

$$p(x, y) = \begin{cases} 4/\pi c^2 & \text{for } \sqrt{x^2 + y^2} < c/2, \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

where c is the diameter of the circle of confusion such that

$$c = \left(\frac{|d_o - d|}{d_o} \right) a. \quad (8)$$

Taking the two-dimensional Fourier transform yields an approximation of the optical transfer function for a defocused camera:

$$P(f_x, f_y) = 2 \operatorname{jinc}(\pi c \sqrt{f_x^2 + f_y^2}) \equiv \frac{2 J_1(\pi c \sqrt{f_x^2 + f_y^2})}{\pi c \sqrt{f_x^2 + f_y^2}}, \quad (9)$$

where $J_1(\xi)$ denotes the first-order Bessel function of the first kind. As shown in Figure 3, the OTF acts as a low-pass filter, interspersed with null frequencies.

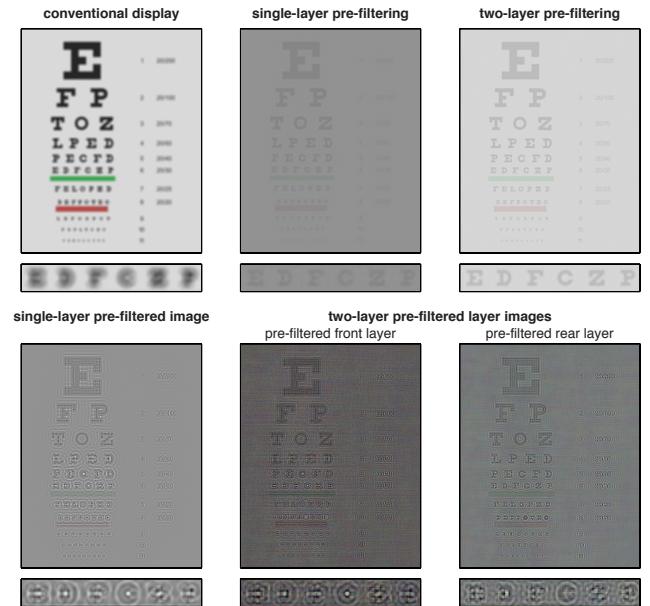


Figure 4: Correcting defocus with pre-filtering. The human eye is modeled as a defocused camera following Figures 2 and 3. A $3.0 \text{ cm} \times 2.4 \text{ cm}$ Snellen chart is presented at 35 cm . We simulate a presbyopic or hyperopic individual requiring a $+1.50$ diopter spherical corrective lens. Single-layer and two-layer displays are considered, with layers separated by $d_1 = 35 \text{ cm}$ and $d_2 = 35.22 \text{ cm}$ (optimized via Equation 20). (Top) From left to right: the received image without correction, using single-layer pre-filtering, and using two-layer pre-filtering. (Bottom Left) Single-layer pre-filtered image $\tilde{s}(x, y)$ given by Equation 5. (Bottom Right) Two-layer pre-filtered images $\tilde{s}_1(x, y)$ and $\tilde{s}_2(x, y)$ given by Equation 17 with the greedy partition given by Equation 21. Note that two-layer pre-filtering improves the legibility and contrast, eliminating artifacts observed with single-layer pre-filtering. Inset regions demonstrate correction to 20/30 vision. (Snellen chart courtesy of Jeff Dahl.)

Application of Equations 5 and 6 yields the pre-filtered image $\tilde{s}(x, y)$; yet, without subsequent processing, this image includes both negative and positive values (roughly of equal magnitude). This is understood by evaluating the structure of the inverse PSF, given by substitution of Equation 9 into Equation 6. Following Yellott and Yellott [2007], the inverse PSF comprises nested rings of positive and negative values with radii of $c/2$. Note that similar structures appear in the pre-filtered image in Figure 4.

In summary, analysis of a defocused camera reveals a second limitation of single-layer pre-filtering: **the pre-filtered image $\tilde{s}(x, y)$ has an expanded dynamic range, with negative and positive values of similar magnitude**. To ensure this image is non-negative, Alonso and Barreto [2003] normalize the pre-filtered image:

$$\tilde{s}_{\text{normalized}}(x, y) = \frac{\tilde{s}(x, y) - \min(\tilde{s}(x, y))}{\max(\tilde{s}(x, y)) - \min(\tilde{s}(x, y))}. \quad (10)$$

As shown in Figures 1 and 4, normalization results in severely reduced contrast. Following Oyamada et al. [2007], clipping outlying values improves contrast, but also introduces additional ringing artifacts. Due to dynamic range expansion, high dynamic range (HDR) displays best support pre-filtering, with standard dynamic range (SDR) displays exhibiting decreased brightness and increased quantization noise. These limitations, in addition to the attenuation of null frequencies of the OTF, have prevented practical applications of single-layer pre-filtering for correcting optical aberrations.

3.2 Multilayer Displays

In this section, we develop pre-filtering for emerging multilayer displays. Such displays comprise stacks of semi-transparent, light-emitting panels separated by small gaps (e.g., layered LCDs). Alternative display architectures, including light field displays and layered organic light-emitting diodes (OLED), are discussed in detail in Section 4. We demonstrate that such displays mitigate the primary limitations of single-layer displays for correcting optical aberrations, improving contrast and eliminating image artifacts.

3.2.1 Multilayer Pre-filtering

Consider an N -layer display with planar screens separated by increasing distances $\{d_1, d_2, \dots, d_N\}$ from an imaging apparatus. Modeled as an LSI system, the received image $i(x, y)$ is given by

$$i(x, y) = \sum_{n=1}^N \bar{s}_n \left(\frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right) * \bar{p}_n \left(\frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right), \quad (11)$$

where $\bar{s}_n(x, y)$ is the image displayed on the n^{th} layer and $\bar{p}_n(x, y)$ is the point spread function for the n^{th} layer (see Figure 2). Assuming a perspective projection of the layers onto the image sensor, each layer is magnified by a factor of d_1/d_n , relative to the front layer. Let $s_n(x, y)$ and $p_n(x, y)$ denote the projections of the n^{th} layer image and PSF onto the first layer, respectively, such that

$$s_n(x, y) = \bar{s}_n \left(\frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right) \text{ and } p_n(x, y) = \bar{p}_n \left(\frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right). \quad (12)$$

Thus, expressed in the plane of the first layer, the received image is

$$i(x, y) = \sum_{n=1}^N s_n(x, y) * p_n(x, y). \quad (13)$$

Equation 13 reveals the first benefit of multilayer displays for correcting optical aberrations; we observe that this expression is equivalent to N collocated, independent single-layer displays, separated by a distance $d = d_1$ from the imaging apparatus. Unlike conventional single-layer displays, the effective point spread function $p_n(x, y)$ applied to each image $s_n(x, y)$ differs. For the defocused camera analyzed in Section 3.1.3, the effective PSFs are given by

$$p_n(x, y) = 2 \text{jinc}(\pi c_n \sqrt{f_x^2 + f_y^2}), \text{ for } c_n = \left(\frac{d_1 |d_o - d_n|}{d_n d_o} \right) a. \quad (14)$$

As shown in Figure 5, due to the varying diameters c_n of the effective circles of confusion, the zeros of the corresponding effective OTFs $P_n(f_x, f_y)$ do not overlap—opening the door to constructing a multilayer pre-filter capable of preserving all spatial frequencies.

Consider the case for which the layer images are identical, such that $s_n(x, y) = s(x, y)$. Equation 13 reduces to the following form.

$$i(x, y) = s(x, y) * p'(x, y), \text{ for } p'(x, y) = \sum_{n=1}^N p_n(x, y). \quad (15)$$

Thus, a multilayer display can be operated in a mode akin to a single-layer display, but where the effective PSF $p'(x, y)$ is given by the linear superposition of the PSFs for each layer. As shown in Figure 5, with an appropriate choice of the layer separations (e.g., one maximizing the minimum value of $p'(x, y)$), the effective PSF $p'(x, y)$ becomes an *all-pass filter*. **Since the nulls of the effective OTFs differ, all spatial frequencies can be preserved in the multilayer pre-filtered image**, given by $\tilde{s}_n(x, y) = s(x, y) * p'^{-1}(x, y)$. An example of this operation mode is shown in Figure 4, eliminating artifacts seen with single-layer pre-filtering.

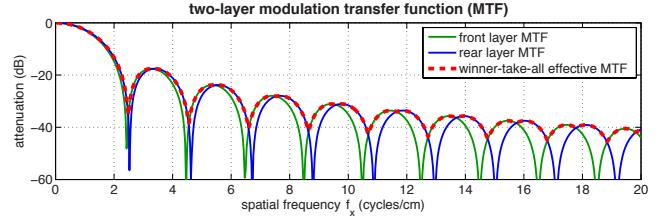


Figure 5: Modulation transfer function with a two-layer display. We again consider a human eye, modeled as in Figures 3 and 4, with focal length $f = 17 \text{ mm}$, aperture $a = 4 \text{ mm}$, and focused at $d_o = 40 \text{ cm}$. This example depicts the MTF for a two-layer display, with layers separated by $d_1 = 35 \text{ cm}$ and $d_2 = 35.22 \text{ cm}$. The green and blue lines depict the effective MTFs, given by the magnitude of the Fourier transform of Equation 14, for the front and rear layer, respectively. Layer positions are optimized via Equation 20, maximizing the minimum value of the effective MTF. Note that the addition of a second layer allows all spatial frequencies to be preserved in the received image, as shown in Figure 4.

3.2.2 Frequency-Domain Analysis of Multilayer Pre-filtering

Multilayer displays also support modes with dissimilar layer images $\tilde{s}_n(x, y)$, while ensuring the received image $\tilde{i}(x, y)$ equals the target image $s(x, y)$. In this section we apply a frequency-domain analysis to show that this added degree of freedom enables a second benefit: **the received image contrast can exceed that achievable with single-layer pre-filtering**.

Taking the two-dimensional Fourier transform of Equation 13 yields the following expression for the received image spectrum:

$$I(f_x, f_y) = \sum_{n=1}^N S_n(f_x, f_y) P_n(f_x, f_y). \quad (16)$$

Extending Equation 5 to multilayer displays indicates the pre-filtered layer image spectrum $\tilde{S}_n(f_x, f_y) = S(f_x, f_y) P_n^{-1}(f_x, f_y)$, such that $\tilde{I}(f_x, f_y) = N \tilde{S}(f_x, f_y)$. This operation mode assumes, as in the preceding section, that each layer contributes equally to the received magnitude of each spatial frequency. However, since the structures of the null frequencies differ for the effective OTFs $\{P_n(f_x, f_y)\}$, a more flexible allocation is possible. For full generality, we allow the pre-filtered layer image spectrum to be:

$$\tilde{S}_n(f_x, f_y) = S(f_x, f_y) [W_n(f_x, f_y) P_n^{-1}(f_x, f_y)], \quad (17)$$

where $W_n(f_x, f_y)$ is the partition function determining the relative contribution of each layer to each spatial frequency component. Note that the partition function must satisfy the constraint:

$$\sum_{n=1}^N W_n(f_x, f_y) = 1, \text{ for } 0 \leq W_n(f_x, f_y) \leq 1. \quad (18)$$

To ensure that the pre-filtered layer images $s_n(x, y)$ are real-valued, the partition function must also be odd-symmetric such that $W_n(f_x, f_y) = W_n(-f_x, -f_y)$.

3.3 Optimizing Image Contrast

The partition function $W_n(f_x, f_y)$ should be defined to maximize the contrast of the received image $\tilde{i}(x, y)$, while preserving all spatial frequencies. Section 3.2 analyzed the partition function $W_n(f_x, f_y) = 1/N$, assigning equal weight to each layer. However, in this section we assess alternative partition functions that, while preserving all spatial frequencies, achieve enhanced contrast.

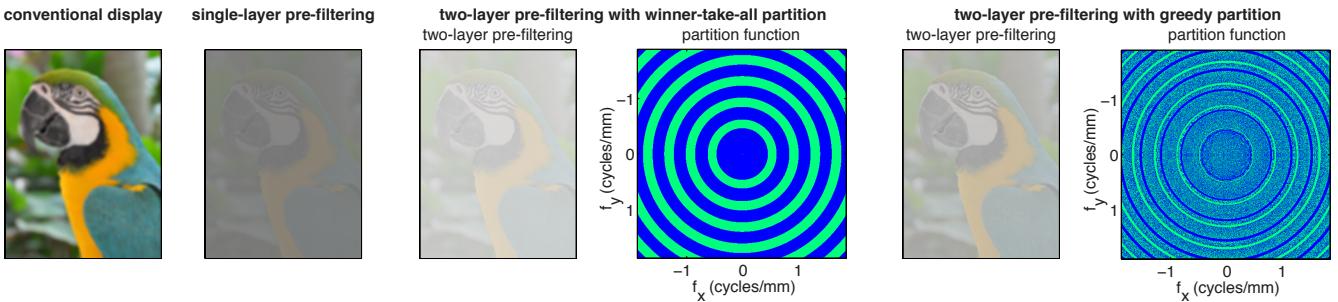


Figure 6: Enhancing image contrast using optimized multilayer partition functions. A postcard-sized (*i.e.*, 13.6 cm × 10.2 cm) image of a bird is presented at 35 cm. We simulate a presbyopic or hyperopic individual requiring a +3.00 diopter spherical corrective lens, such that the closest focus distance is $d_o = 1.0$ m. (Left) The received images without correction and using single-layer pre-filtering are shown on the left and right, respectively. (Middle) Two-layer pre-filtering results are shown using the winner-take-all partition function given by Equation 19. The partition function $W_1(f_x, f_y)$ is shaded from blue to green, corresponding to a value of 0 and 1, respectively. (Right) Two-layer pre-filtering results are shown using the greedy partition function described in Section 3.3. Pre-filtering restores fine structures including the pupil, feathers, and stripes. Note that ringing artifacts observed with single-layer pre-filtering (*e.g.*, along the bird’s silhouette) are eliminated with two-layer pre-filtering. However, contrast is most enhanced with the greedy partition function, which more fully exploits the added degrees of freedom afforded by multiple display layers. (Macaw image courtesy of Michael Gwyther-Jones.)

Consider the *winner-take-all partition function*, defined such that

$$W_n(f_x, f_y) = \begin{cases} 1 & \text{for } n = \arg \max_m |P_m(f_x, f_y)|, \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

As shown in Figure 5, this partition function ensures that each spatial frequency is only reproduced on one layer of the display; the layer with the maximum effective MTF $|P_n(f_x, f_y)|$, for a given spatial frequency (f_x, f_y) , is assigned a unit weight, with the remaining layers making no contribution to this component. Under this choice of the partition function, one can optimize the layer distances $\{d_1, d_2, \dots, d_N\}$ such that the minimum value of the overall MTF (*i.e.*, the envelope of the effective MTFs) is maximized. This corresponds to the solution of the following optimization problem.

$$\arg \max_{\{d_2, \dots, d_N\}} [\min(|P_1(f_x, f_y; d_1)|, \dots, |P_N(f_x, f_y; d_N)|)] \quad (20)$$

In practice, one desires a partition function that minimizes the loss of contrast that occurs when applying Equation 10 to normalize the pre-filtering layer images $\tilde{s}_n(x, y)$. Generally, the minimum value of $\{\tilde{s}_n(x, y)\}$ should be maximized, such that a small bias can be added to the pre-filtered image to restore non-negativity. Solving for the optimal choice of the partition function to achieve this goal requires a combinatorial search, assuming a discrete image with a finite set of spatial frequencies. To accelerate this search, we propose the following iterative *greedy partition function* algorithm. We initialize the partition function using Equation 19. Afterward, a spatial frequency (f_x, f_y) is selected, in decreasing order, based on the magnitude of the target image spectrum normalized by the average of the effective MTFs. The following update rule is applied:

$$\{W_n(f_x, f_y)\} \rightarrow \arg \max_{\{W_n(f_x, f_y)\}} [\min(\tilde{s}_1(x, y), \dots, \tilde{s}_N(x, y))], \quad (21)$$

ensuring that the smallest value on any layer is maximized. Updates continue until all spatial frequencies have been considered.

Figure 6 summarizes the performance of multilayer pre-filtering for various choices of the partition function. Note that multilayer pre-filtering does not exhibit the ringing artifacts previously observed with single-layer pre-filtering, due to the preservation of all spatial frequencies in the received image. For the case of black text on a white background (shown in Figure 4), the greedy partition function, implemented on a two-layer display, significantly enhances contrast and legibility relative to prior methods.

4 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.

4.1 Multilayer Displays

Multilayer displays contain stacks of semi-transparent panels, such as liquid crystal displays (LCDs) or organic light-emitting diodes (OLEDs). We assess how current displays can meet our criteria.

4.1.1 Multilayer OLEDs

OLEDs contain an organic film enclosed between electrode arrays that emits light proportional to the applied voltage. Transparent OLEDs incorporate semi-transparent contacts [Görn et al. 2006], providing an ideal architecture for multilayer pre-filtering. 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; as described in Section 4.2, such elements ensure each eye views different pixels on each layer, enabling binocular correction at the cost of reduced resolution.

4.1.2 Beamsplitter 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].

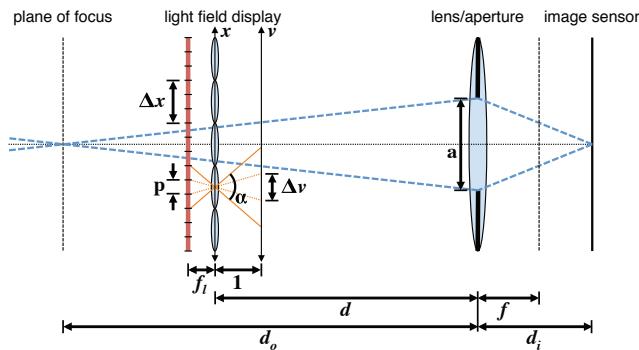


Figure 7: A defocused camera observing a light field display. A light field display, here depicted as an integral imaging display, is separated by a distance d from the lens. A lenslet array, of focal length f_l , is affixed such that A display pixels of width p are covered, allowing control of A light rays within the field of view α .

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

4.1.3 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 pre-filtered 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 $60N$ Hz. In Section 6, 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.

4.2 Light Field Displays

Practical multilayer displays require increasing the display thickness, limiting mobile applications. By Equation 20, 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.

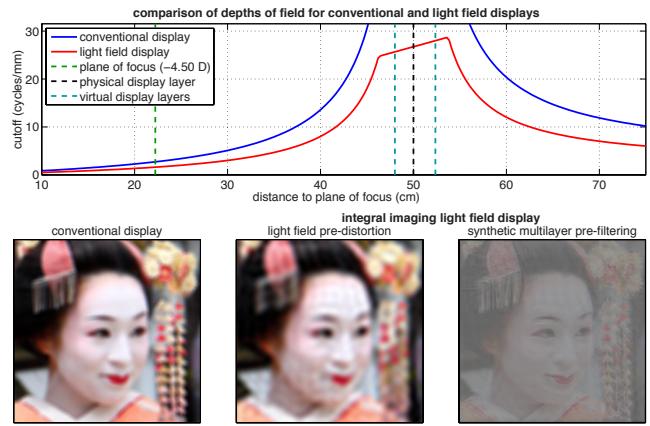


Figure 8: Correcting defocus using light field displays. An integral imaging display is constructed by modifying a 14.8 cm × 19.7 cm display with 104 pixels per cm (equivalent to a 2012 Apple iPad). A lenslet array is affixed with 19.7 lenses per cm and focal length $f_l = 7.5$ mm. The display is separated from the eye by $d = 50$ cm. We simulate a myopic individual requiring a -4.50 diopter spherical corrective lens, such that the far point is at $d_o = 22.2$ cm. (Top) The depths of field for light field and conventional displays, given by Equations 22 and 23. (Bottom) From left to right: the received image using a conventional display, light field pre-distortion, and synthetic multilayer pre-filtering. As analyzed in Section 4.2.2, pre-filtering only uses the available depth of field: placing virtual layers at $d_1 = 48.0$ cm and $d_2 = 52.4$ cm. In contrast, light field pre-distortion exceeds the depth of field, placing a virtual layer 27.8 cm in front of the display. (Maiko image courtesy of Zhang Wenjie.)

4.2.1 Parallax Barrier and Integral Imaging Displays

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]. As shown in Figure 7, affixing a lenslet array to a conventional 2D display creates an integral imaging display. Each lenslet is separated by its focal length f_l and covers A pixels, each of width p . Thus, each lenslet is capable of emitting A light rays within a field of view of α degrees, creating a multiview display supporting A views. A parallax barrier display functions similarly, with a grid of pinholes substituting for the lenslet array.

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.

4.2.2 Correcting Defocus with Light Field Displays

In this section, we assess the ability of light field displays to correct for defocus. 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 *pre-distorted* light field such that, 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. In contrast, synthetic multi-layer pre-filtering requires synthesizing two or more virtual layers, generally in close proximity to the light field display.

We formally assess these modes by comparing the depth of field expressions describing conventional displays and light field displays. As characterized by Zwicker et al. [2006], the depth of field defines the maximum spatial frequency $f_{\max}(d_o)$ that can be depicted in a virtual plane separated by a distance d_o from a light field display. As shown in Figure 7, we adopt a two-plane parameterization of the light field [Chai et al. 2000], where ray (x, v) is defined by its intersection with the x -axis, coincident with the display surface, and the v -axis, located a unit distance in front. As derived by Wetzstein et al. [2012], the depth of field for a parallax barrier or integral imaging display, evaluated in the plane of the image sensor, is given by

$$f_{\max}(d_o) = \begin{cases} \frac{d_o}{d_i} \frac{1}{2\Delta x} & \text{for } |d_o - d| \leq \left(\frac{\Delta x}{\Delta v}\right), \\ \frac{d_o}{d_i} \frac{1}{2|d_o - d|\Delta v} & \text{otherwise,} \end{cases} \quad (22)$$

where Δx is the lenslet width, $\Delta v = (2/A)\tan(\alpha/2)$ is the width of the projection of a display pixel onto the v -axis, and the factors of d_o/d_i account for the projection onto the image sensor. As shown in Figure 8, the image resolution is nearly constant near the light field display surface, but rapidly decreases as the distance d_o to the plane of focus (i.e., the virtual layer) moves away from the surface.

As a baseline point of comparison, we consider the depth of field for a conventional display (e.g., an LCD) located a distance d from the viewer. Similar to Equation 8, the diameter of the circle of confusion for a defocused camera, projected onto the image sensor, is given by $c = (d_i/d)(|d_o - d|/d_o)a$. Thus, the maximum spatial frequency in a defocused image of a conventional display is:

$$f_{\max}(d_o) = \min\left(\frac{d}{d_i} \frac{1}{2p}, \frac{d}{d_i} \frac{d_o}{2|d_o - d|}a\right), \quad (23)$$

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

The ratio of Equations 22 or 23 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., $|d_o - d| \geq \Delta x/\Delta v$), this ratio is given by

$$r_{\max} = \frac{a}{d\Delta v}. \quad (24)$$

We observe that r_{\max} is equal to the number of light rays entering the aperture of the camera from a single lenslet. 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. In Figure 8, we consider a specific example using current-generation LCDs and lenslet arrays. Note that, even with a state-of-the-art LCD with 104 pixels per cm (ppcm), affixing a lenslet array slightly *decreases* the received image resolution, relative to an unmodified display.

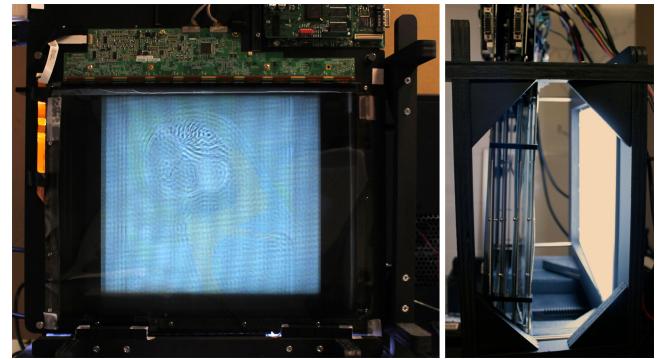


Figure 9: Prototype multilayer LCD. (Left) The two-layer pre-filtered images, for the target image shown in Figure 6, are displayed on the second and fourth layers. The images are presented such that a viewer directly in front of the display perceives a focused image (see the top row of Figure 10). (Right) Four LCD panels are mounted on rails, supporting arbitrary layer separations.

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 meeting our design constraints. As observed by Pamplona et al. [2012], light field pre-distortion is feasible only once resolutions significantly exceed current commercial panels (approaching 1,000 ppcm). While reducing contrast, multilayer pre-filtering can be implemented using current-generation displays with 100 ppcm.

5 Implementation

This section describes the multilayer LCD prototype. Section 5.1 details the construction and Section 5.2 reviews the software.

5.1 Hardware

As described in Section 2.3, 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 Lamm et al. [2011]. As shown in Figure 9, the prototype comprises four modified 40.8 cm \times 30.6 cm Barco E-2320 PA LCD panels, supporting 8-bit grayscale display with a resolution of 1600 \times 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 were 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 4.1, 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 semi-transparent stack of light-emitting layers. A 2.8 GHz Intel Core i7 workstation with 8 GB of RAM controls the prototype. A four-head NVIDIA Quadro NVS 450 graphics card synchronizes the panels.

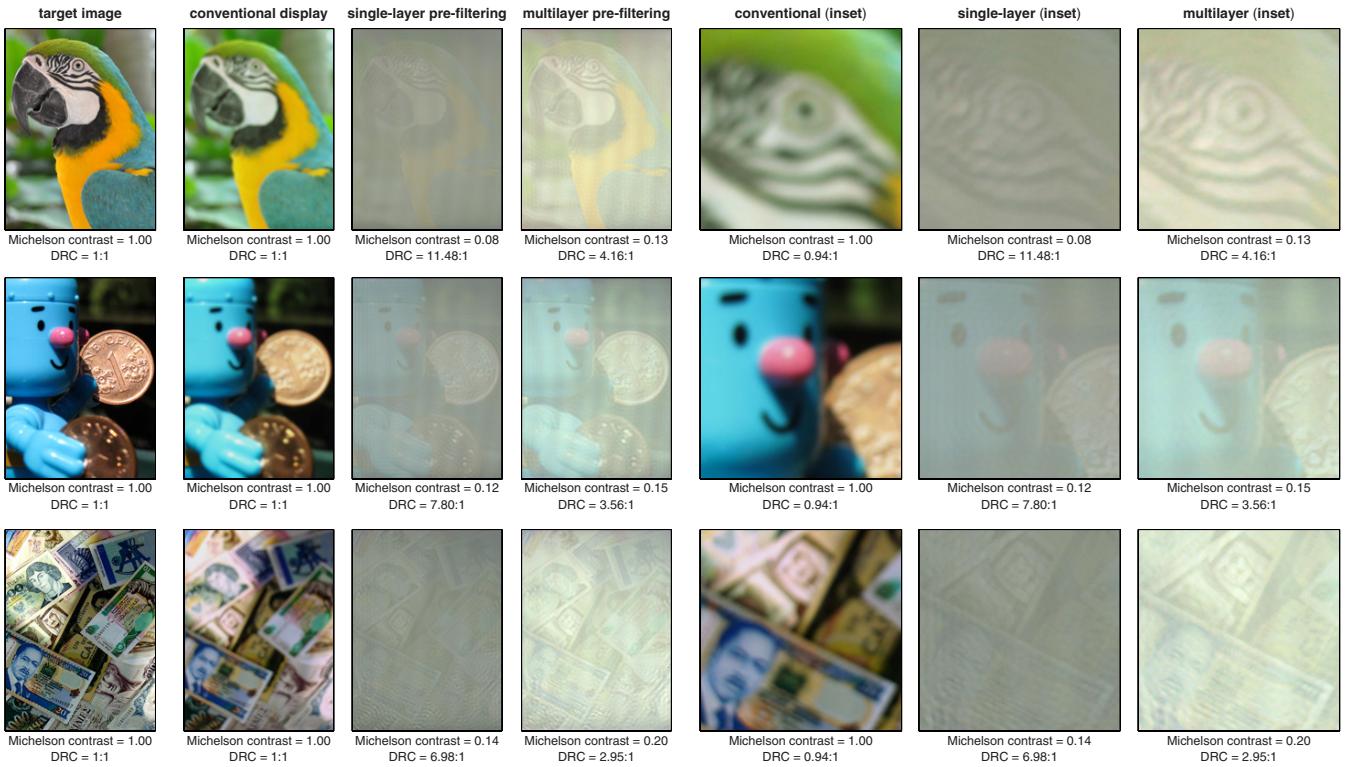


Figure 10: Correcting defocus using the multilayer LCD prototype. The prototype was photographed using the defocused camera and display parameters described in Section 6.1. The first four columns depict, from left to right: the target image and the received images without pre-filtering, using single-layer pre-filtering, and using two-layer pre-filtering. The remaining three columns show inset regions of the second through fourth columns. Michelson contrast is reported for each received image. Dynamic range compression (DRC) refers to the ratio of the maximum dynamic range of the pre-filtered layer images (before normalization) to the displayed layer images. Note that Michelson contrast is enhanced using multilayer pre-filtering. As shown in the inset regions, ringing artifacts are mitigated with multilayer pre-filtering. As described in Section 6.1, ringing artifacts remain visible in the periphery, due to the off-axis variation of the PSF. (Source images courtesy of Michael Gwyther-Jones (top row), Flickr user “palo” (Dragon Papillon Photography) (middle row), and Flickr user “earlg” (bottom row).)

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 pre-filtering, 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, as shown in Figure 10, a long camera exposure allows multilayer pre-filtering experiments. Second, the panels only support grayscale modes. This has the benefit of mitigating moiré 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 10, we combine three separate photographs, each recorded while displaying a different color channel of the pre-filtered images.

5.2 Software

We implemented the single-layer and multilayer pre-filtering algorithms described in Section 3 using a combination of Matlab scripts and compiled C/C++ routines. The FFTW discrete Fourier transform library was used to accelerate pre-filtering. For color images, each channel is processed independently in a separate thread. For a 256×256 color image, single-layer pre-filtering requires an average of 1 second for processing; two-layer pre-filtering takes 5 seconds, when using the winner-take-all partition function, and 150 seconds

when using the greedy partition function. Section 7 describes procedures to accelerate greedy partitions. All run times are reported using the same workstation used to control the prototype.

6 Performance Assessment

6.1 Experimental Results

Figure 10 summarizes experimental results achieved with the multilayer LCD prototype. A Canon EOS Rebel T3 digital camera, with a Canon EF 50 mm f/1.8 II lens, was separated by 100 cm from the front layer of the prototype. The camera was focused 16 cm in front of the display, with the minimum f-number setting of 1.8, resulting in an aperture diameter $a \approx 2.8$ cm. We compare two modes of operation: single-layer pre-filtering and two-layer pre-filtering, with the two panels separated by a gap of 3.4 cm. Three sample images were evaluated. As described in Section 5.1, three exposures were combined to synthesize color images using the grayscale panels.

Comparing the top row of Figure 10 to Figure 6 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 pre-filtering, rather than the prior single-layer pre-filtering algorithm. Also note that the outline of the eye and the black stripes appear with less distortion using multilayer pre-filtering. Ringing artifacts, visible on the left-hand side of the face of the blue toy, are eliminated with multilayer pre-filtering.

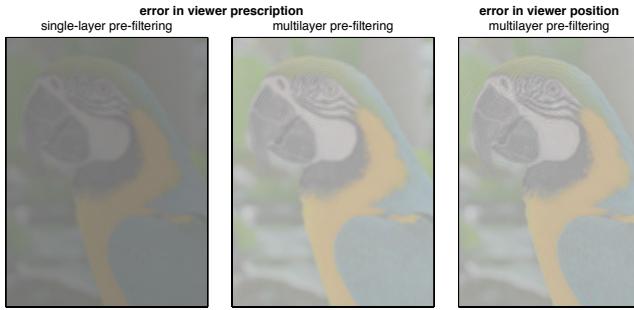


Figure 11: Sensitivity to perturbations in viewer prescription and position. We consider calibration errors for the example in Figure 6. (Left) Pre-filtering requires accurate estimates of the viewer’s prescription. Pre-filtering is performed assuming a +2.50 diopter correction, rather than the true value of +3.00 diopters. (Right) Multilayer pre-filtering also requires tracking the lateral viewer position. In this case, the viewer is displaced 10.0 cm to the right of the estimated position. Note the appearance of high-frequency artifacts for both prescription and position errors. (Macaw image courtesy of Michael Gwyther-Jones.)

Experimental results also reveal limitations of the linear spatially-invariant (LSI) model introduced in Section 3.1. First, the 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 $\gamma = 2.2$, 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 3.1.3.

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 pre-filtering rather than single-layer pre-filtering. Following Section 3.1.3, pre-filtering expands the dynamic range both above and below the range of radiance values that is physically supported by the display. We quantify this effect by evaluating the dynamic range compression (DRC) of the pre-filtered images, given by the difference of the maximum and minimum values before normalization using Equation 10. 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 pre-filtering, despite normalization.

6.2 Limitations of Multilayer Pre-filtering

Both existing single-layer and the proposed multilayer pre-filtering algorithms are sensitive to perturbations in the viewer’s refractive error. As shown in Figure 11, if the corrective power differs from the viewer’s true refractive error, then the received image will be degraded. Both single-layer and multilayer pre-filtering require tracking the viewer. With single-layer pre-filtering, the distance to the viewer must be estimated to model the PSF in the plane of the display; however, unlike single-layer pre-filtering, multilayer pre-filtering also requires tracking lateral motion—ensuring that the multiple layers are rendered with the correct perspective. Sensitivity to lateral tracking errors are depicted in Figure 11.

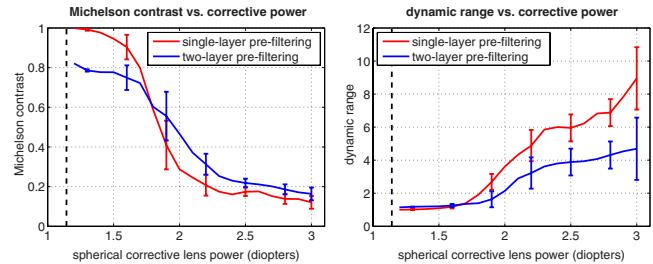


Figure 12: The Michelson contrast of the received image and the dynamic range of the pre-filtered images are shown on the left and right, respectively. The three test images in Figure 10 were processed as a function of varying corrective power; error bars have lengths equal to twice the standard deviation measured over the three images. The viewing and display parameters correspond to the postcard example in Figure 6. Note that, for moderate to severe presbyopia or hyperopia (i.e., requiring a correction of greater than +1.75 diopters), two-layer pre-filtering enhances contrast by 40% and decreases the dynamic range by 60%, compared to single-layer pre-filtering. The dashed black line denotes a correction of +1.14 diopters, less than which focusing is possible without correction.

Increasing contrast in the received image lies at the heart of enabling practical applications of single-layer and multilayer pre-filtering. The prototype results demonstrate moderate improvements over single-layer pre-filtering, while achieving the goal of eliminating ringing artifacts. Similar to the strong dependence on depth of field for light field pre-distortion (see Section 4.2.2), Figure 12 assesses the dependence of contrast enhancement on the required corrective power. From this analysis, we identify a key limitation of the proposed multilayer pre-filtering algorithm: the received image contrast is significantly reduced for large amounts of defocus. In Section 7, we discuss potential refinements for further improving contrast using multilayer pre-filtering.

6.3 Multilayer Pre-filtering for Videos

Results obtained by applying pre-filtering to videos are included in the supplementary video. Without modifications, processing each frame independently produces videos with rapid intensity variations. 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 pre-filtering images. For a pre-recorded sequence, perceived flashing can be removed by normalizing each frame by the global minimum and maximum values of the pre-filtered sequence. For interactive or streaming content, we propose applying an adaptive filter to recursively estimate a temporally-smoothed estimate of the necessary normalization range.

7 Discussion and Future Work

As established by theory and experiment, multilayer pre-filtering achieves our primary goal: mitigating contrast loss and eliminating ringing artifacts observed with single-layer pre-filtering. Yet, multilayer pre-filtering 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 pre-filtered 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 Michelson 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 4.2.2, emerging light field displays are a compelling platform for achieving practical applications of multi-layer pre-filtering. By utilizing synthetic, rather than physical, layers, display thicknesses can be reduced and layers can be virtually displaced to account for viewer movement. Future work includes constructing a working prototype using off-the-shelf parts. A particularly promising direction is to combine the benefits of multi-layer pre-filtering with those of light field pre-distortion. With full generality, we propose applying pre-filtering 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, deconvolution may yield further benefits in contrast.

Our ultimate goal is to augment, or substitute for, corrective eyewear and invasive surgery by pre-processing images to correct for optical aberrations of the human eye. In this paper, we have restricted our demonstrations to correcting lower-order defocus aberrations using two-layer displays. A promising direction for future work is to extend our approach to address higher-order aberrations. As described by Barsky [2004], wavefront aberrometry (e.g., using a Shack-Hartmann aberrometer) can be applied to characterize higher-order aberrations. Such systems typically quantify the wavefront deviation due to refractive errors by reporting a series of Zernike polynomial coefficients. We propose using the mapping introduced by Barsky to transform wavefront aberrometer measurements to effective PSFs, as required for multilayer pre-filtering. We anticipate that correction of higher-order aberrations may require more than two layers to eliminate ringing artifacts (i.e., to obtain all-pass optical filters) and to maximize received image contrast.

8 Conclusion

The development of multilayer pre-filtered displays has the potential to offer an assistive device that may reduce the need for eyewear or surgery. For individuals with common lower-order aberrations, our proposed method increases visual acuity without the need for eyewear when viewing a programmable display such as a phone, tablet, laptop screen, workstation monitor, or other information display. In this work, we establish the benefits and limitations of multilayer pre-filtering for such applications, mitigating contrast loss and eliminating ringing artifacts occurring with single-layer approaches. The concepts we present also have the potential to address unmet needs for individuals with higher-order optical aberrations, which cannot be corrected with ordinary eyeglasses and remain difficult to correct without advanced contact lenses or surgery. We hope to inspire others to pursue similar efforts to combine the benefits of jointly optimizing displays and image-encoding algorithms to correct the full range of aberrations of the human eye.

Acknowledgements: We thank the reviewers for their insightful feedback. Fu-Chung Huang and Brian A. Barsky were supported by NSF Grant IIS-1219241. Douglas Lanman was supported by NSF Grant IIS-1116452. Ramesh Raskar was supported by a Sloan Research Fellowship, a DARPA Young Faculty Award, and a Vodafone Wireless Innovation Award.

References

- AGRAWAL, A., XU, Y., AND RASKAR, R. 2009. Invertible motion blur in video. *ACM Trans. Graph.* 28, 3, 95:1–95:8.
- AKELEY, K., WATT, S. J., GIRSHICK, A. R., AND BANKS, M. S. 2004. A stereo display prototype with multiple focal distances. *ACM Trans. Graph.* 23, 3, 804–813.
- ALONSO JR., M., AND BARRETO, A. B. 2003. Pre-compensation for high-order aberrations of the human eye using on-screen image deconvolution. In *IEEE Engineering in Medicine and Biology Society*, vol. 1, 556–559.
- ARCHAND, P., PITE, E., GUILMET, H., AND TROCME, L., 2011. Systems and methods for rendering a display to compensate for a viewer's visual impairment. International Patent Application PCT/US2011/039993.
- BARSKY, B. A. 2004. Vision-realistic rendering: simulation of the scanned foveal image from wavefront data of human subjects. In *Applied Perception in Graphics and Visualization*, 73–81.
- BELL, G. P., CRAIG, R., PAXTON, R., WONG, G., AND GALBRAITH, D. 2008. Beyond flat panels: Multi-layered displays with real depth. *SID Digest* 39, 1, 352–355.
- BROWN, M. S., SONG, P., AND CHAM, T.-J. 2006. Image preconditioning for out-of-focus projector blur. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1956–1963.
- CAMPISI, P., AND EGIAZARIAN, K. 2007. *Blind Image Deconvolution: Theory and Applications*. CRC Press.
- CHAI, J.-X., TONG, X., CHAN, S.-C., AND SHUM, H.-Y. 2000. Plenoptic sampling. In *ACM SIGGRAPH*, 307–318.
- GONZALEZ, R. C., AND WOODS, R. E. 1992. *Digital Image Processing (Third Edition)*. Addison-Wesley.
- GOODMAN, J. W. 2004. *Introduction to Fourier Optics (Third Edition)*. Roberts & Company Publishers.
- GÖRRN, P., SANDER, M., MEYER, J., KRÖGER, M., BECKER, E., JOHANNES, H.-H., KOWALSKY, W., AND RIEDL, T. 2006. Towards see-through displays: Fully transparent thin-film transistors driving transparent organic light-emitting diodes. *Advanced Materials* 18, 6, 738–741.
- GOTODA, H. 2012. Implementation and analysis of an autostereoscopic display using multiple liquid crystal layers. In *SPIE Stereoscopic Displays and Applications XXIII*, vol. 8288.
- GROSSE, M., WETZSTEIN, G., GRUNDHÖFER, A., AND BIMBER, O. 2010. Coded aperture projection. *ACM Trans. Graph.*
- HOLROYD, M., BARAN, I., LAWRENCE, J., AND MATSIK, W. 2011. Computing and fabricating multilayer models. *ACM Trans. Graph.* 30, 6, 187:1–187:8.
- HUANG, F.-C., AND BARSKY, B. A. 2011. A framework for aberration compensated displays. Tech. Rep. UCB/EECS-2011-162, University of California, Berkeley, December.
- IVES, F. E., 1903. Parallax stereogram and process of making same. U.S. Patent 725,567.
- KAUFMAN, P., AND ALM, A. 2002. *Adler's Physiology of the Eye (Tenth Edition)*. Mosby.
- KEE, E., PARIS, S., CHEN, S., AND WANG, J. 2011. Modeling and removing spatially-varying optical blur. In *IEEE International Conference on Computational Photography*.

- LANMAN, D., WETZSTEIN, G., HIRSCH, M., HEIDRICH, W., AND RASKAR, R. 2011. Polarization fields: dynamic light field display using multi-layer LCDs. *ACM Trans. Graph.* 30, 6.
- LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.* 26, 3.
- LIPPmann, G. 1908. Épreuves réversibles donnant la sensation du relief. *Journal of Physics* 7, 4, 821–825.
- OYAMADA, Y., AND SAITO, H. 2007. Focal pre-correction of projected image for deblurring screen image. In *IEEE Conference on Computer Vision and Pattern Recognition*.
- PAMPLONA, V. F., OLIVEIRA, M. M., ALIAGA, D. G., AND RASKAR, R. 2012. Tailored displays to compensate for visual aberrations. *ACM Trans. Graph.* 31, 4, 81:1–81:12.
- RASKAR, R., AGRAWAL, A., AND TUMBLIN, J. 2006. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. Graph.* 25, 3, 795–804.
- UREY, H., CHELLAPPAN, K., ERDEN, E., AND SURMAN, P. 2011. State of the art in stereoscopic and autostereoscopic displays. *Proceedings of the IEEE* 99, 4, 540–555.
- VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.* 26, 3.
- VITALE, S., SPERDUTO, R. D., AND FERRIS, III, F. L. 2009. Increased prevalence of myopia in the United States between 1971–1972 and 1999–2004. *Arch. Ophthalmology* 127, 12, 1632–1639.
- WETZSTEIN, G., LANMAN, D., HEIDRICH, W., AND RASKAR, R. 2011. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.* 30, 4.
- WETZSTEIN, G., LANMAN, D., HIRSCH, M., AND RASKAR, R. 2012. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.* 31, 4, 80:1–80:11.
- YELLOTT, J. I., AND YELLOTT, J. W. 2007. Correcting spurious resolution in defocused images. *Proc. SPIE* 6492.
- ZHANG, L., AND NAYAR, S. 2006. Projection defocus analysis for scene capture and image display. *ACM Trans. Graph.* 25, 3.
- ZOMET, A., AND NAYAR, S. 2006. Lensless imaging with a controllable aperture. In *IEEE Conference on Computer Vision and Pattern Recognition*.
- ZWICKER, M., MATUSIK, W., DURAND, F., AND PFISTER, H. 2006. Antialiasing for automultiscopic 3D displays. In *Eurographics Symposium on Rendering*.