Correcting for Optical Aberrations using Multilayer Displays

Fu-Chung Huang¹ Douglas Lanman²,³ Brian A. Barsky¹ Ramesh Raskar²
University of California, Berkeley¹ MIT Media Lab² NVIDIA Research³

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 formally assess the benefits of multilayer pre-filtering vs. 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.

Links: _DL  _PDF  _WEB  _VIDEO

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. Eyeglasses only correct common lower-order aberrations (i.e., defocus and astigmatism) that occur with myopia, hyperopia, or presbyopia. Higher-order aberrations, such as the artifacts induced by disorders including keratoconus or pellucid marginal degeneration, can be difficult to correct and is currently attempted 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.

We are not the first to propose correction of optical aberrations using novel display devices. Our approach builds upon that introduced by Alonso and Barreto [2003] and 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 displayed image is first deconvolved by the known 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:

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 top 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 bottom 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.
We demonstrate that, by optimizing the separation between display layers, multilayer pre-filtering preserves all spatial frequencies in the received image, eliminating the ringing artifacts appearing with prior single-layer pre-filtering methods.

We show that, by optimizing the partition of spatial frequencies between layers, 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 formally analyze resolution enhancement for multilayer pre-filtering vs. light field pre-distortion methods for aberration correction.

Through 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 from 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. Recently, Archand et al. [2011] apply single-layer pre-filtering to commercial displays. In a closely related work, Huang and Barsky [2011] introduce multilayer pre-filtering and evaluate its performance via simulations.

Correcting Projector Defocus can be achieved by applying pre-filtering to deconvolve the project 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] evaluate the performance of clipping values outside the dynamic range vs. normalization. Zhang and Nayar [2006] propose solving a constrained optimization problem to minimize artifacts while utilizing only the available dynamic range. While these works consider 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. These works advocate 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 introduce ringing. Our development of multilayer pre-filtering is inspired by these works, with the goal of incorporating additional layers to ensure all spatial frequencies are preserved in the received image.

2.3 Multilayer Displays

Multilayer displays are 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 vs. 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. Lamman et al. [2011] and Gotoda [2012] evaluate stacks of LCD panels; these
works 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. Multi-layer 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. While 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

\[
i(x,y) = s(x,y) * p(x,y),
\]

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

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

where \( p^{-1}(x,y) \) is the inverse point spread function: defined such that \( p^{-1}(x,y) * p(x,y) = \delta(x,y) \), where \( \delta(x,y) \) is the Dirac delta function. Substituting Equation 2 into Equation 1 yields the following expression for the received image \( \tilde{i}(x,y) \) using pre-filtering.

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

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; 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 in a frequency-domain analysis.

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

\[
I(f_x, f_y) = S(f_x, f_y)P(f_x, f_y),
\]

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

\[
\tilde{S}(f_x, f_y) = S(f_x, f_y)P^{-1}(f_x, f_y).
\]

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

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

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 $\tilde{i}(x, y)$.

### 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 $d_s$ 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, the point spread function is a uniform disk with unit area, given by

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

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

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

(8)

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

$$P(f_x, f_y) = 2\text{jinc}(\pi c\sqrt{f_x^2 + f_y^2}) = \frac{2J_1(\pi c\sqrt{f_x^2 + f_y^2})}{\pi c\sqrt{f_x^2 + f_y^2}}.$$  

(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. 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:

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

(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 $\{d_1, d_2, \ldots, d_N\}$ from an imaging apparatus. Modeled as an LSI system, the received image $i(x, y)$ is given by

$$i(x, y) = \sum_{n=1}^{N} \tilde{s}_n \left( \frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right) \ast \tilde{p}_n \left( \frac{d_n}{d_1} x, \frac{d_n}{d_1} y \right).$$  

(11)

where $\tilde{s}_n(x, y)$ is the image displayed on the $n^{th}$ layer and $\tilde{p}_n(x, y)$ 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

\[ s_n(x, y) = \bar{s}_n \left( \frac{d_n}{dx} - \frac{d_n}{dy} \right) \quad \text{and} \quad p_n(x, y) = \bar{p}_n \left( \frac{d_n}{dx} - \frac{d_n}{dy} \right). \]  

(12)

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

\[ i(x, y) = \sum_{n=1}^{N} s_n(x, y) * p_n(x, y). \]

(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

\[ p_n(x, y) = 2 \text{jinc}(c_n \sqrt{x^2 + y^2}), \]

for \( c_n = \left( \frac{d_1 d_n - d_1 d_n}{d_1 d_n} \right) \). \( \text{(14)} \)

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.

\[ i(x, y) = s(x, y) * p(x, y), \]

for \( p(x, y) = \sum_{n=1}^{N} p_n(x, y). \) \( \text{(15)} \)

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 \( \bar{s} \bar{x}(x, y) = s(x, y) * p^{-1}(x, y) \). An example of this operation mode is shown in Figure 4, eliminating artifacts seen with single-layer pre-filtering.

3.2.2 Frequency-Domain Analysis of Multilayer Pre-filtering

Multilayer displays also support modes with dissimilar layer images \( \bar{s}_n(x, y) \), while ensuring the received image \( \bar{i}(x, y) = s(x, y) \). 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:

\[ I(f_x, f_y) = \sum_{n=1}^{N} S_n(f_x, f_y)P_n(f_x, f_y). \]  

(16)

Extending Equation 5 to multilayer displays indicates the pre-filtered layer image spectrum \( \bar{S}_n(f_x, f_y) = S(f_x, f_y)P_n^{-1}(f_x, f_y) \), such that \( I(f_x, f_y) = NS(f_x, f_y) \). 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:

\[ \bar{S}_n(f_x, f_y) = S(f_x, f_y) \left[ W_n(f_x, f_y)P_n^{-1}(f_x, f_y) \right]. \]  

(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:

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

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 \( \bar{i}(x, y) \), 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.

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

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

(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 \( \{d_1, d_2, \ldots, d_N\} \) 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.

\[ \arg \max_{\{d_1, \ldots, d_N\}} \left[ \min \{|P_1(f_x, f_y; d_1)|, \ldots, |P_N(f_x, f_y; d_N)|\} \right]. \]

(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:

\[
\{ W_n(f_x, f_y) \} \rightarrow \arg \max_{\{ W_n(f_x, f_y) \}} \left[ \min \left( \tilde{s}_1(x, y), \ldots, \tilde{s}_N(x, y) \right) \right]. \tag{21}
\]

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. First, it should be optically equivalent to a stack of semi-transparent, lighting-emitting layers. Second, it should be thin. Third, it should support binocular correction, since refractive errors may differ between eyes. Fourth, it should 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].

Viewing multiple LCDs through a set of half-silvered mirrors (i.e., beamsplitters) is optically equivalent to the required 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: 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 60 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 separation between layers depends on the viewer’s refractive error and position. While 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.

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 $f_1$ and covers $A$ pixels, each of width $p$. Thus, each lenslet is capable of emitting a light ray within a field of view of $\alpha$ 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 multilayer 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 $f_{\text{max}}(d_o)$ 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, 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

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

where $\Delta x$ 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 
\[
c = (d_o/d)(|d_o - d|/d) o. \text{ Thus, the maximum spatial frequency in a defocus image of a conventional display is:}
\]
\[
f_{\text{max}}(d_o) = \min \left( \frac{d}{d_1}, \frac{1}{2p} \frac{d}{d_2} |d_o - d| o, \right), (23)
\]
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 of 23 provides an analytic expression for the maximum resolution enhancement \(r_{\text{max}}\) 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
\[
r_{\text{max}} = \frac{a}{d \Delta v}. \quad (24)
\]
We observe that \(r_{\text{max}}\) 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. This is due to the fact that, 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 Lanman 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.

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 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\(\times\)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 on the target image, and then 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 Michael Gwyther-Jones (top row), Flickr user “palo” (Dragon Papillon Photography) (middle row), and Flickr user “earlg” (bottom row).

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.

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, as analyzed by Kee et al. [2011], the lens produces a spatially-varying PSF; 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 41% using multilayer pre-filtering vs. single-layer pre-filtering with these examples. 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 divided by their sum. Michelson contrast is increased by an average of 41% using multilayer pre-filtering vs. single-layer pre-filtering with these examples. 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.

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 on contrast enhancement vs. 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 further research may reveal computationally-efficient alternatives that achieve a greater contrast enhancement through refined optimization algorithms vs. our iterative approach.

We optimize contrast, as measured in a linear radiometric domain and quantified by the Michelson contrast of the received image or the dynamic range of the pre-filtered layers. 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 multilayer 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 multilayer 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 eliminate the need 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 augment or eliminate the need for eyewear or surgery. For individuals with common lower-order aberrations, such displays provide a convenient consumer technology: allowing a casual glance at a watch, phone, or other information display to be accomplished without eyewear. 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. Yet, the concepts presented in this work also have the potential to address unmet needs for individuals with higher-order aberrations, which 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.

References


