

Beyond Parallax Barriers: Applying Formal Optimization Methods to Multi-Layer Automultiscopic Displays

Douglas Lanman^{a†}, Gordon Wetzstein^a, Matthew Hirsch^a,
Wolfgang Heidrich^b, and Ramesh Raskar^a

^aMIT Media Lab, 75 Amherst Street, Cambridge MA, USA;

^bUniversity of British Columbia, 2366 Main Mall, Vancouver BC, Canada

ABSTRACT

This paper focuses on resolving long-standing limitations of parallax barriers by applying formal optimization methods. We consider two generalizations of conventional parallax barriers. First, we consider general two-layer architectures, supporting high-speed temporal variation with arbitrary opacities on each layer. Second, we consider general multi-layer architectures containing three or more light-attenuating layers. This line of research has led to two new attenuation-based displays. The High-Rank 3D (HR3D) display contains a stacked pair of LCD panels; rather than using heuristically-defined parallax barriers, both layers are jointly-optimized using low-rank light field factorization, resulting in increased brightness, refresh rate, and battery life for mobile applications. The Layered 3D display extends this approach to multi-layered displays composed of compact volumes of light-attenuating material. Such volumetric attenuators recreate a 4D light field when illuminated by a uniform backlight. We further introduce Polarization Fields as an optically-efficient and computationally-efficient extension of Layered 3D to multi-layer LCDs. Together, these projects reveal new generalizations to parallax barrier concepts, enabled by the application of formal optimization methods to multi-layer attenuation-based designs in a manner that uniquely leverages the compressive nature of 3D scenes for display applications.

Keywords: computational displays, automultiscopic displays, light fields, parallax barriers, multi-layer LCDs

1. INTRODUCTION

Thin displays that present the illusion of depth have become a driving force in the consumer electronics and entertainment industries, offering a differentiating feature in a market where the utility of increasing 2D resolution has brought diminishing returns. In such systems, binocular depth cues are achieved by presenting different images to each of the viewer’s eyes. An attractive class of 3D displays, known as automultiscopic or multi-view autostereoscopic displays, present view-dependent imagery without requiring special eyewear. These displays provide binocular and motion parallax cues. Conventional designs use either optically-attenuating parallax barriers, introduced by Ives in 1903, or refracting lens arrays (also known as integral imaging), introduced by Lippmann in 1908.^{1,2} In both cases an underlying 2D display (e.g., an LCD) is covered with a second optical element. For barrier designs, a second light-attenuating layer is placed slightly in front of the first, whereas lens arrays are directly affixed to the underlying display. In most commercial systems either parallax barriers or lenticular sheets are used for horizontal-only parallax. However, pinhole arrays or integral lens sheets can achieve simultaneous vertical and horizontal parallax. While automultiscopic displays realized with lenses are brighter than barrier designs, lens-based designs reduce spatial resolution; as a result, we seek to address the limitations of multi-layer attenuation-based displays, particularly diminished brightness and reduced spatial resolution.

All automultiscopic displays share three components: display elements, viewers, and multi-view content. Prior works consider the benefits of prefiltering multi-view content for a particular display device.³ Others consider adapting display elements (e.g., the spacing within parallax barriers) depending on viewer position.^{4,5} We propose two generalizations of conventional parallax barriers, which we dub *High-Rank 3D (HR3D)*⁶ and *Layered 3D*⁷, for which display elements are optimized for the multi-view content. The resulting generalized parallax barriers significantly differ from existing heuristics (e.g., grids of slits or pinholes) and achieve increased optical

[†]E-mail: dlanman@media.mit.edu, Telephone: 401-477-2233

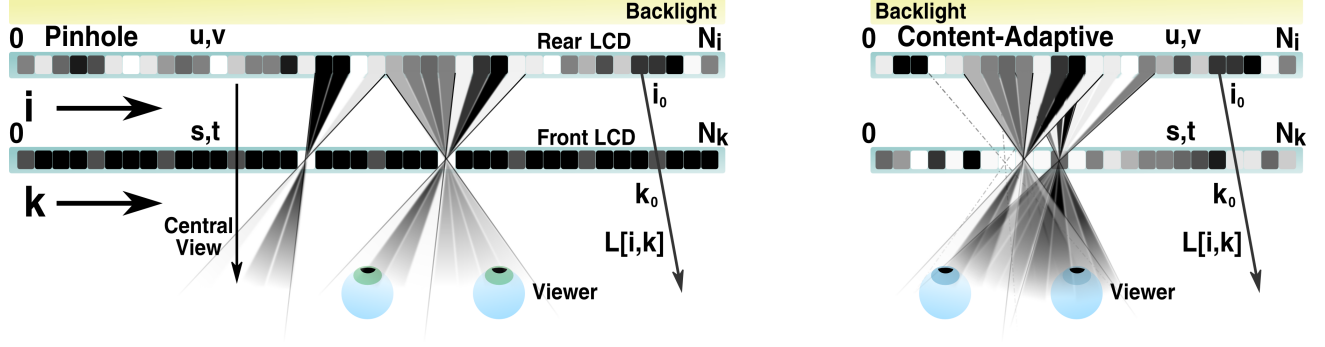


Figure 1. Conventional versus content-adaptive parallax barriers. (Left) In a conventional parallax barrier display the front panel contains a uniform grid of slits or pinholes. The viewer sees each pixel on the rear panel through this grid, selecting a subset of visible pixels depending on viewer location. A uniform backlight, located behind the rear layer, enables the rear layer to act as a conventional 2D display. (Right) Rather than heuristic barriers, we consider two-layer LCDs as general spatial light modulators that act in concert to recreate a target light field by attenuating rays emitted by the backlight. Unlike conventional barriers, both masks can exhibit non-binary opacities.

transmission, enhanced spatial resolution, greater depth of field, and higher refresh rates, while preserving the fidelity of displayed images. We further introduce *Polarization Fields*⁸ as an optically-efficient implementation of our generalized parallax barriers using multi-layer LCDs. Most significantly, displays employing these generalized barriers can similarly benefit from prefiltered multi-view content and can adapt to the viewer position, allowing simultaneous optimization of all three significant system components.

In this paper we summarize our recent work on generalizing parallax barriers. In Sec. 2 we review HR3D, demonstrating how to optimize two-layer display architectures, wherein each layer acts as a general spatial light modulator. Unlike conventional parallax barriers, we allow both masks to exhibit non-binary opacities and to rapidly vary in time. In Sec. 3 we review Layered 3D, considering a second generalization wherein the display architecture contains more than two light-attenuating layers. In Sec. 4 we review Polarization Fields, demonstrating optically-efficient architectures and computationally-efficient algorithms for automultiscopic display using multi-layer LCDs. In Sec. 5 we derive an upper bound on the depth of field for all multi-layer attenuation-based displays, formally establishing their potential to significantly outperform conventional parallax barrier designs. We conclude by discussing future directions for research in Sec. 6.

2. HIGH-RANK 3D (HR3D): TIME-MULTIPLEXED TWO-LAYER DISPLAYS

In this section two-layer LCDs are analyzed as general spatial light modulators that act in concert to recreate a light field by attenuating rays emitted by the backlight. It is shown that any fixed pair of masks only creates a rank-1 approximation of a light field. Higher-rank approximations are achieved with time multiplexing. We optimize 3D display with two-layer LCDs using a matrix approximation framework. This leads to content-adaptive parallax barriers allowing brighter displays with increased refresh rates.

2.1 Overview

A general parallax barrier display, containing two mask layers and a backlight, can be analyzed as a light field display device. The following analysis adopts an *absolute two-plane parameterization* of the 4D light field. As shown in Fig. 1, an emitted ray is parameterized by the coordinates of its intersection with each mask layer. Thus, the ray (u, v, s, t) intersects the rear mask at the point (u, v) and the front mask at the point (s, t) , with both mask coordinate systems having an origin in the top-left corner.

In a practical automultiscopic display one is primarily concerned with the projection of optical rays within a narrow cone perpendicular to the display surface (see Fig. 1), since most viewers will be located directly in front of the device. The distinct images viewable within this region are referred to as the “central views” projected by the display. As a result, a *relative two-plane parameterization* proves more convenient to define a target light

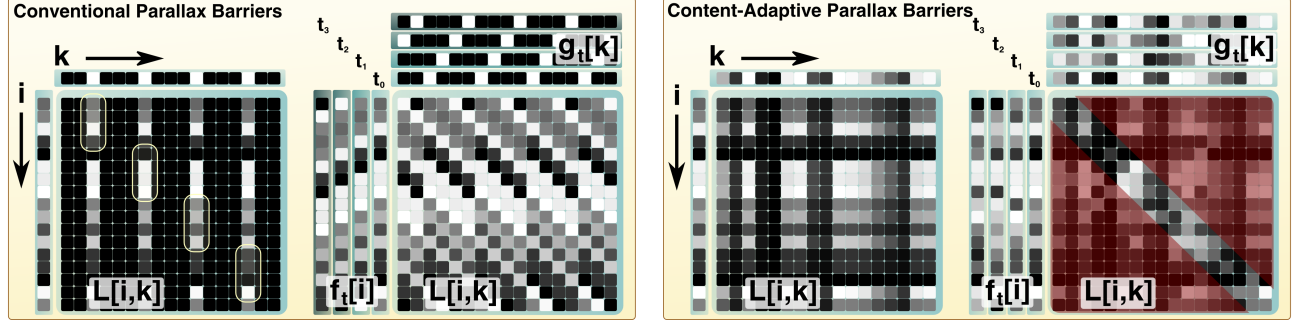


Figure 2. Rank constraints for parallax barriers. (Left) Conventional parallax barriers approximate the light field matrix (center) as the outer product of mask vectors (above and to the left). The resulting rank-1 approximation accurately reproduces the circled elements (corresponding to the central views). Note that most columns are not reconstructed, reducing display resolution and brightness. Periodic replicas of the central views are created outside the circled regions. (Middle Left) Time-shifted parallax barriers achieve higher-rank reconstructions by integrating a series of rank-1 approximations, each created by a single translated mask pair. (Middle Right) Content-adaptive parallax barriers increase display brightness by allowing both masks to exhibit non-binary opacities. Here a rank-1 approximation is demonstrated using a single mask pair. (Right) Rank- M approximations are achieved using temporal multiplexing of M content-adaptive parallax barriers via Eqn 2. In practice, the light field will be full rank without enforcing periodic replication (as created by conventional parallax barriers). As a result, we do not constrain rays (shown in red) outside the central view in Eqn. 3.

field; in this parameterization, an emitted ray is defined by the coordinates (u, v, a, b) , where (u, v) remains the point of intersection with the rear plane and (a, b) denotes the relative offset of the second point of intersection such that $(a, b) = (s - u, t - v)$. Note that a 2D slice of the 4D light field, for a fixed value of (a, b) , corresponds to a skewed orthographic view (formally an oblique projection).

2.2 Optimization Procedure

In the following analysis, we consider 2D light fields and 1D masks, with a straightforward generalization to 4D light fields and 2D masks. A general pair of optical attenuation functions, $f_m(\xi)$ and $g_m(\xi)$, is defined with the absolute parameterization, where m denotes the index into a sequence of M such masks pairs. These functions correspond to the rear and front masks, respectively. Following Lanman et al.⁶, the emitted 2D light field $l(u, a)$, averaged over the M -frame sequence, is given by the following expression:

$$l(u, a) = \frac{1}{M} \sum_{m=1}^M f_m(u - (d_f/d_r)a) g_m(u + (d_g/d_r)a), \quad (1)$$

where d_r is the distance of the s -axis from the u -axis and d_f and d_g are the distances of the rear and front masks from the u -axis, respectively. In practice, the masks and the emitted light field are discrete functions. The discrete pixel indices are denoted as (i, j, k, l) , corresponding to the continuous coordinates (u, v, s, t) , such that the discretized light field matrix is $\mathbf{L}$ and the sampled mask vectors are $\mathbf{f}_m$ and $\mathbf{g}_m$. Under this model, the emitted 2D light field matrix $\mathbf{L}$ is given by the following expression.

$$\mathbf{L} = \frac{1}{M} \sum_{m=1}^M \mathbf{f}_m \otimes \mathbf{g}_m = \frac{1}{M} \sum_{m=1}^M \mathbf{f}_m \mathbf{g}_m^T \quad (2)$$

From Eqn. 2 it is clear that a fixed pair of 1D masks can only produce a rank-1 approximation of any given 2D light field matrix. Using a sequence of M masks pairs, a rank- M approximation can be achieved (see Fig. 2). This provides an important insight into the inherent limitations of all two-layer attenuation-based displays, including conventional parallax barriers and time-shifted parallax barriers⁹; while two-layer displays using fixed mask pairs produce rank-deficient approximations, perceptually-acceptable approximations can be obtained using conventional parallax barriers, at the cost of decreasing the achievable spatial resolution and image brightness.

We generalize the concept of temporal multiplexing for parallax barriers by considering all possible mask pairs with arbitrary opacities. To achieve this goal, we formulate light field display using a time-multiplexed,

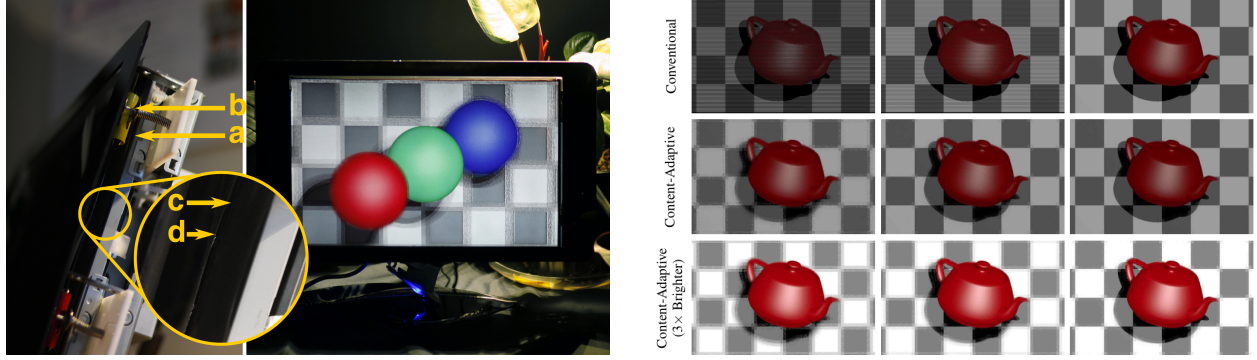


Figure 3. HR3D: High-Rank 3D display. (Left) Prototype automultiscopic display using dual-stacked LCDs, showing: (a) rear LCD with backlight, (b) spacer, (c) front LCD, and (d) replacement polarizing sheet. (Right) Content-adaptive barriers are compared to time-shifted barriers⁹, with the exposure normalized so the relative image brightness is consistent with observation. Reconstructions with three, six, and nine time-multiplexed mask pairs are shown along the columns from the left to right, respectively. While content-adaptive barriers produce some high-frequency artifacts, even with nine mask pairs, they can compress the light field with higher PSNR than conventional barriers. As shown along the bottom row, adaptation also allows the brightness to be increased with minimal degradation in image fidelity.

two-layer attenuation-based display as a constrained weighted least-squares optimization problem. We define the matrices $\mathbf{F}$ and $\mathbf{G}$ such that column m of $\mathbf{F}$ and row m of $\mathbf{G}$ denotes the rear and front masks displayed during frame m , respectively. We seek a *content-adaptive* light field factorization $\tilde{\mathbf{L}} = \mathbf{F}\mathbf{G}$ that minimizes the weighted Euclidean distance to the target light field $\mathbf{L}$, under the necessary non-negativity constraints, such that

$$\arg \min_{\mathbf{F}, \mathbf{G}} \|\mathbf{W} \circ (\mathbf{L} - \mathbf{F}\mathbf{G})\|^2, \text{ for } 0 \leq \mathbf{F}, \mathbf{G} \leq 1. \quad (3)$$

Unlike conventional barriers, the field of view can be adapted to one or more viewers by specifying elements of the weight matrix $\mathbf{W}$ (i.e., the Euclidean norm will be minimized where $\mathbf{W}$ is large). The weight matrix plays a crucial role, ensuring a low-rank approximation can obtain high reconstruction accuracy by artificially reducing the rank of the target light field. General 4D light fields are handled by reordering as 2D matrices, with 2D masks reordered as vectors, allowing a similar matrix approximation scheme to be applied.

Equation 3 can be solved using non-negative matrix factorization (NMF). Prior numerical methods include the multiplicative update rule.¹⁰ We use the weighted update introduced by Blondel et al.¹¹ Initial estimates $\{\mathbf{F}, \mathbf{G}\}$ are refined as follows.

$$\mathbf{F} \leftarrow \mathbf{F} \circ \frac{(\mathbf{W} \circ \mathbf{L})\mathbf{G}^\top}{(\mathbf{W} \circ (\mathbf{F}\mathbf{G}))\mathbf{G}^\top} \quad \mathbf{G} \leftarrow \mathbf{G} \circ \frac{\mathbf{F}^\top(\mathbf{W} \circ \mathbf{L})}{\mathbf{F}^\top(\mathbf{W} \circ (\mathbf{F}\mathbf{G}))} \quad (4)$$

In our implementation of Eqn. 4, the masks are initialized with random values uniformly distributed on $[0, 1]$; alternative strategies, including seeding with conventional parallax barriers, did not yield reconstructions with reduced errors or increased transmission. After each iteration the mask elements are truncated to the range $[0, 1]$. In conclusion, we propose the resulting non-negative, content-adaptive parallax barriers as a generalization of traditional parallax barrier displays, in which images displayed on both layers are jointly optimized, independently for each target automultiscopic video frame.

2.3 Implementation and Results

As shown in Fig. 3, a two-layer LCD was constructed using a pair of 1680×1050 Viewsonic FuHzion VX2265wm 120 Hz LCD panels. The panels have a pixel pitch of $282 \mu\text{m}$ and are separated by 1.5 cm. Thus, for a typical light field with an angular resolution $N_h \times N_v$ of 5×3 views, the prototype supports an $11^\circ \times 7^\circ$ field of view; a viewer sees correct imagery when moving within a frustum with similar apex angles. The rear layer is an unmodified panel, whereas the front layer is a spatial light modulator (SLM) fashioned by removing the backlight



Figure 4. Layered 3D: Inexpensive, glasses-free light field display using volumetric attenuators. (Left) A stack of spatial light modulators (e.g., printed masks) recreates a target light field (here for a car) when illuminated by a backlight. (Right) The target light field is shown in the upper left, together with the optimal five-layer decomposition, obtained with iterative tomographic reconstruction. (Middle) Oblique projections for a viewer standing to the top left (magenta) and bottom right (cyan). Corresponding views of the target light field and five-layer prototype are shown on the left and right, respectively. Such attenuation-based 3D displays allow accurate, high-resolution depiction of motion parallax, occlusion, translucency, and specularities, being exhibited by the trunk, the fender, the window, and the roof of the car, respectively.

from a second panel. The front polarizing diffuser and rear polarizing film are removed. The front polarizing diffuser is replaced with a transparent polarizer, restoring the spatial light modulation capability of the panel. Without such modifications, the polarizers in the front panel completely attenuate light polarized by the rear panel. Eliminating the redundant rear polarizer of the front panel increases light transmission. The LCD panels are driven separately via DVI links from a dual-head NVIDIA Quadro FX 570 display adapter, automatically synchronizing the display refreshes.

As with any 3D display, a viewer is concerned with resolution (both spatial and angular), brightness, refresh rate, and reconstruction error. As shown in Fig. 3, simulations are used to compare the performance of content-adaptive parallax barriers to time-shifted parallax barriers.⁹ Two primary benefits result from content-adaptive parallax barriers: increased display brightness and increased display refresh rate. We conclude that content adaptation, as allowed by the HR3D display architecture, reveals a novel trade-off between automultiscopic display brightness, refresh rate, and reconstruction error. Additional details are presented in Lanman et al.⁶

3. LAYERED 3D: STATIC MULTI-LAYER DISPLAYS

3.1 Overview

In this section multi-layer attenuators (e.g., stacks of transparencies or LCDs) are analyzed as general spatial light modulators that act in concert to recreate a light field by attenuating rays emitted by a uniform backlight. Since arbitrary oblique views may be inconsistent with any single attenuator, iterative tomographic reconstruction minimizes the difference between the emitted and target light fields, subject to physical constraints on attenuation. As multi-layer generalizations of conventional parallax barriers, such displays are shown, both by theory and experiment, to exceed the performance of existing two-layer architectures. For 3D display, spatial resolution, depth of field, and brightness are increased, compared to conventional parallax barriers.

3.2 Optimization Procedure

Similar to Sec. 2.2, we consider 2D light fields and 1D masks in the following analysis, with a straightforward generalization to 4D light fields and 2D masks. An N -layer stack of optical attenuation functions (i.e., masks) is defined using the absolute light field parameterization, such that the mask on layer n is given by the transmittance function $f^{(n)}(\xi)$. Following Wetzstein et al.⁷, the emitted 2D light field $l(u, a)$ is given by the following expression:

$$l(u, a) = \prod_{n=1}^N f^{(n)}(u + (d_n/d_r)a), \quad (5)$$

where d_n denotes the distance of layer n from the u -axis. Taking the logarithm gives the following additive model:

$$\bar{l}(u, a) \triangleq \sum_{n=1}^N \ln f^{(n)}(u + (d_n/d_r)a) = - \sum_{n=1}^N \alpha^{(n)}(u + (d_n/d_r)a), \quad (6)$$

where $\alpha^{(n)}(\xi) = -\ln f^{(n)}(\xi)$ is the absorbance function for layer n . As before, we consider a discrete light field matrix $\bar{\mathbf{L}}$ (containing the logarithm of the target intensities for each ray) and discrete attenuation vectors $\alpha^{(n)}$ for each layer n . Furthermore, we vectorize $\bar{\mathbf{L}}$ such that an element of $\bar{\mathbf{l}}$ denotes the logarithm of an individual ray. With this notation, we formulate light field display using a multi-layer attenuation-based display as the following constrained linear least-squares optimization problem.

$$\arg \min_{\alpha} \|\bar{\mathbf{l}} + \mathbf{P}\alpha\|^2, \text{ for } \alpha \geq 0 \quad (7)$$

Here, $\mathbf{P}$ is the *projection matrix* such that the total attenuation along ray i is given by the summation $\sum_{j=1}^J P_{ij}\alpha_j$. In practice, we use a sparse, constrained, large-scale trust region method¹² to solve Eqn. 7. Limited layer contrast is incorporated as a constraint.

3.3 Implementation and Results

As shown in Fig. 4, our Layered 3D display prototype consists of five layers separated by clear acrylic sheets, each 0.3175 cm thick. Layers are printed on transparencies at 300 dots per inch using an Epson Stylus 2200 inkjet printer with six color primaries. In practice, interreflections between layers do not lead to visual artifacts. Similarly, reflections on the display surface are minimized by dimming ambient lighting (an anti-reflection coating could be applied if needed). Moiré is not observed with our layered fabrication, likely due to ink naturally blending neighboring pixels. We ignore scattering and diffraction due to the ink pigments, although both likely place an upper limit on the achievable spatio-angular resolution. As shown in Fig. 4 the resulting prototype demonstrates accurate automultiscopic display using tomographic image synthesis.

Capabilities of multi-layer attenuation-based displays are comparable to volumetric and automultiscopic displays, particularly parallax barriers and integral imaging. While volumetric displays faithfully reproduce perceptual depth cues, most cannot represent objects beyond the display enclosure; many require moving assemblies and often cannot depict opaque scenes. Yet, the inexpensive alternatives of parallax barriers and integral imaging cannot reproduce high-resolution images, even in the display plane. Our Layered 3D display architecture presents a unique set of capabilities within this space. Foremost, depth of field can exceed that of conventional automultiscopic displays (see Sec. 5), allowing representation of objects floating beyond the display enclosure. Within or near the display, full-resolution depiction is routinely achieved (see Fig. 4). Multi-layer displays approach the brightness of integral imaging, with additional layers further enhancing contrast. Finally, similar to volumetric displays, accommodation is preserved for objects within the display, with sufficiently-dense layering. Additional details are presented in Wetzstein et al.⁷

4. POLARIZATION FIELDS: DYNAMIC MULTI-LAYER DISPLAYS

In this section we extend Layered 3D to multi-layer LCD panels to achieve dynamic light field display. The resulting Polarization Field display architecture is composed of a stacked set of liquid crystal panels with a single pair of crossed linear polarizers (see Fig. 5). Each layer is modeled as a spatially-controllable polarization rotator, as opposed to a conventional spatial light modulator that directly attenuates light (e.g., in contrast to the optical models applied with HR3D and Layered 3D). Color display is achieved using field sequential color illumination with monochromatic LCDs, mitigating severe attenuation and moiré occurring with layered color filter arrays. We further demonstrate interactive refresh rates by adopting the simultaneous algebraic reconstruction technique (SART) algorithm to tomographically solve for the optimal spatially-varying polarization state rotations.



Figure 5. Polarization Fields: Dynamic light field display using multi-layer LCDs. (Left) We construct an optically-efficient polarization field display by covering a stack of liquid crystal panels with crossed linear polarizers. Each layer functions as a polarization field rotator, rather than as a conventional optical attenuator. (Right, Top) A target light field. (Right, Bottom) Light fields are displayed, at interactive refresh rates, by tomographically solving for the optimal rotations to be applied at each layer. (Middle) A pair of simulated views is compared to corresponding photographs of the prototype on the left and right, respectively. Inset regions denote the relative position with respect to the display layers, shown as black lines, demonstrating objects can extend beyond the display surface.

4.1 Overview

Constructing polarization field displays requires an accurate characterization of the optical properties of LCDs. The transformation of polarized light due to passage through layered materials is modeled by the Jones calculus.¹³ Orthogonal components of the electric field are represented as a complex-valued Jones vector. The optical action of a given element (e.g., a birefringent layer or polarizing film) is represented by a Jones matrix, with the product of this matrix and a Jones vector encoding the polarization state transformation. Yeh and Gu¹⁴ formally characterize the polarization properties of LCDs, providing analytic Jones matrices for common technologies, including twisted nematic (TN), vertical alignment (VA), and in-plane switching (IPS) panels. In this paper we consider a unifying, but simplified, Jones matrix model, wherein LCDs are approximated as spatially-controllable polarization rotators. A more detailed Jones matrix model, appropriate for the IPS panels used in our prototype, is presented in Lanman et al.⁸

4.2 Optimization Procedure

Consider a pair of crossed linear polarizers enclosing a single liquid crystal cell; under the polarization rotator model, the transmitted intensity I is given by Malus' law:

$$I = I_0 \sin^2(\theta), \quad (8)$$

where I_0 is the intensity after passing through the first polarizer and θ is the angle of polarization after passing through the liquid crystal cell, defined relative to the axis of the first polarizer. Similar to Secs. 2.2 and 3.2, we consider 2D light fields and 1D masks in the following analysis, with a straightforward generalization to 4D light fields and 2D masks. Following Lanman et al.⁸, the emitted 2D light field $l(u, a)$ is given by:

$$l(u, a) = \sin^2(\theta(u, a)) = \sin^2 \left(\sum_{n=1}^N \phi^{(n)}(u + (d_n/d_r)a) \right), \quad (9)$$

where $\phi^{(n)}(\xi)$ denotes the spatially-varying polarization state rotation induced at point ξ along layer n . Under this model, ray (u, a) intersects the N layers, accumulating incremental rotations at each intersection, such that emitted polarization field $\tilde{\theta}(u, a)$ approximates the target polarization field $\theta(u, a)$, given by

$$\theta(u, a) = \pm \sin^{-1} \left(\sqrt{l(u, a)} \right) \bmod \pi. \quad (10)$$

Under these assumptions, the principal value of the arcsine ranges over $[0, \pi/2]$. Note, with full generality, the target polarization field is multi-valued and periodic, since a rotation of $\pm\theta \bmod \pi$ radians will produce an

identical intensity by application of Malus' law. In our implementation we only consider the principal value of this expression, limiting the target polarization field to $\theta(u, a) \in [0, \pi/2]$. Yet, if this restriction is lifted, additional degrees of freedom appear accessible.

Similar to Layered 3D, light field display using multi-layer LCDs can be formulated as a constrained linear least-square optimization problem. We consider a discrete parameterization for which the emitted polarization field is represented as a column vector $\tilde{\theta}$ with I elements, each of which corresponds to the angle of polarization for a specific light field ray. Similarly, the polarization state rotations are represented as a column vector ϕ with J elements, each of which corresponds to a specific display pixel in a given layer. Under this parameterization, Eqn. 9 yields a linear model such that

$$\tilde{\theta}_i = \sum_{j=1}^J P_{ij} \phi_j, \quad (11)$$

where $\tilde{\theta}_i$ and ϕ_j denote ray i and pixel j of $\tilde{\theta}$ and ϕ , respectively. An element P_{ij} of the projection matrix $\mathbf{P}$ is given by the normalized area of overlap between pixel j and ray i , occupying a finite region determined by the sample spacing. Using this notation, an optimal set of polarization state rotations is found by solving the following optimization problem:

$$\arg \min_{\phi} \|\theta - \mathbf{P}\phi\|^2, \text{ for } \phi_{min} \leq \phi \leq \phi_{max}, \quad (12)$$

where each layer can apply a rotation ranging over $[\phi_{min}, \phi_{max}]$.

Similar to Wetzstein et al.⁷, Eqn. 12 can be solved using a sparse, constrained, large-scale trust region method.¹² However, we observe that this problem can be solved more efficiently by adapting the simultaneous algebraic reconstruction technique (SART). As proposed by Andersen and Kak¹⁵ and further described by Kak and Slaney¹⁶, SART provides an iterative solution wherein the estimate of ϕ is updated such that:

$$\phi \leftarrow \phi + \mathbf{v} \circ (\mathbf{P}^T (\mathbf{w} \circ (\theta - \mathbf{P}\phi))), \quad (13)$$

where $\circ$ denotes the Hadamard product (element-wise multiplication) and the $\mathbf{w}$ and $\mathbf{v}$ vectors are given by

$$w_i = \frac{1}{\sum_{j=1}^J P_{ij}} \quad \text{and} \quad v_j = \frac{1}{\sum_{i=1}^I P_{ij}}. \quad (14)$$

After each iteration, additional constraints on ϕ are enforced by clamping the result to the feasible rotation range. Building upon the Kaczmarz method for solving linear systems of equations¹⁷, SART is shown to rapidly converge to a solution approaching the fidelity of that produced by alternative iterative methods, including trust region and conjugate gradient descent techniques.¹⁶

4.3 Implementation and Results

As shown in Fig. 5, a four-layer LCD panel was constructed by modifying Barco E-2320 PA LCDs; these panels support 1600×1200 8-bit grayscale display at 60 Hz and have an active area of 40.8×30.6 cm. The liquid crystal layer was separated from the case, backlight, and power supply. Polarizing films were removed and the adhesive was dissolved with acetone. The exposed panel, driver boards, and power supply were mounted to a waterjet-cut aluminum frame. Arbitrary layer spacings are supported by translating the frames along rails. Acrylic spacers hold the layers at a fixed spacing of 1.7 cm for all experiments, yielding a total display thickness of 5.1 cm. The prototype is illuminated using an interleaved pair of backlights and controlled by a 3.4 GHz Intel Core i7 workstation with 4 GB of RAM. A four-head NVIDIA Quadro NVS 450 graphics card synchronizes the displays.

Polarization fields accurately depict multiple perspectives of the “Buddha” scene (see Fig. 5). In summary, polarization fields present both an optically and computationally efficient architecture for dynamic light field display using multi-layer LCDs. We observe our adaptation of SART can be similarly applied to attenuation layers, such as those used by Layered 3D, by substituting the logarithm of the emitted light field intensity $\bar{I}$ and the attenuation function α for θ and ϕ in Eqn. 12, respectively; thus, we provide the first implementation for interactive frame rates for both attenuation-based and polarization-rotating multi-layer displays. Additional details are presented in Lanman et al.⁸

5. UPPER BOUND ON DEPTH OF FIELD OF MULTI-LAYER DISPLAYS

Depth of field is a key metric for judging the relative performance of various automultiscopic display technologies. As described by Zwicker et al.¹⁸, the *depth of field* of an automultiscopic display is an expression describing the maximum spatial frequency that can be depicted, without aliasing, in a virtual plane oriented parallel to, and located a known distance from, the display surface. In this section we derive an upper bound on the spatio-angular bandwidth of any multi-layer, attenuation-based display, including those in Secs. 2–4. This upper bound is shown, for two layers, to encompass prior depth of field expressions for parallax barriers and integral imaging. Most significantly, the upper bound indicates multi-layer, attenuation-based displays can significantly increase the spatial resolution for virtual objects located close to the display surface, as compared to current architectures.

5.1 Depth of Field for Conventional Automultiscopic Displays

As established by Zwicker et al.¹⁸, the depth of field can be assessed by analyzing the spectral properties of the displayed light field. For conventional parallax barriers and integral imaging, the discrete sampling of emitted rays (u, a) produces a light field spectrum $\hat{l}(f_u, f_a)$ that is non-zero only within a rectangle. As described by Chai et al.¹⁹ and Durand et al.²⁰, the spectrum of a Lambertian surface, located a distance d_o from the middle of the display, corresponds to the line $f_a = (d_o/d_r)f_u$ in the frequency domain, where d_r is the distance between the u -axis and s -axis. The spatial cutoff frequency is given by the intersection of this line with the spectral bandwidth of the display. Thus, conventional automultiscopic displays have a finite depth of field such that the spatial frequency f_ξ in a plane at d_o must satisfy

$$|f_\xi| \leq \begin{cases} \frac{f_0}{N_a}, & \text{for } |d_o| + (h/2) \leq N_a h \\ \left(\frac{h}{(h/2) + |d_o|} \right) f_0, & \text{otherwise,} \end{cases} \quad (15)$$

where N_a is the number of angular views, h is the thickness of the display, and $f_0 = 1/2p$ is the spatial cutoff frequency of a mask layer with a pixel size of p .

5.2 Upper Bound on Depth of Field for Multi-Layer Displays

The upper bound on the depth of field for a multi-layer display is similarly assessed by considering the maximum spectral bandwidth. Following Eqn. 5, the light field $l(u, a)$ emitted by a multi-layer display is given by

$$l(u, a) = \prod_{n=1}^N f^{(n)}(u + (d_n/d_r)a) \quad (16)$$

where N is the number of layers and $f^{(n)}(\xi)$ is the transmittance of mask n . Note that we assume a uniform backlight and that the target light field is normalized such that $l(u, a) \in (0, 1]$. It is also assumed that the mask layers are uniformly distributed such that $d_n \in [-h/2, h/2]$, where h is the thickness of the display. The light field spectrum $\hat{l}(f_u, f_a)$ is given by the two-dimensional Fourier transform of Eqn. 16, as follows.

$$\hat{l}(f_u, f_a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{n=1}^N f^{(n)}(u + (d_n/d_r)a) e^{-2\pi j f_u u} e^{-2\pi j f_a a} du da \quad (17)$$

By the convolution property of Fourier transforms²¹, this expression reduces to a *repeated convolution* of the individual mask spectra $\hat{f}^{(n)}(f_\xi)$, such that the light field spectrum is given by

$$\hat{l}(f_u, f_a) = \bigotimes_{n=1}^N \hat{f}^{(n)}(f_u) \delta(f_a - (d_n/d_r)f_u) = \hat{f}^{(1)}(f_u) \delta(f_a - (d_1/d_r)f_u) \otimes \cdots \otimes \hat{f}^{(N)}(f_u) \delta(f_a - (d_N/d_r)f_u), \quad (18)$$

where $\bigotimes$ denotes the repeated convolution operator. Note that each mask produces a spectrum $\hat{f}^{(n)}(f_u, f_a)$ that lies along a slanted line, following Chai et al.¹⁹ Since each mask has a finite resolution, the individual spectra may be non-zero only for $|f_u| \leq f_0$. We define the maximum spectral bandwidth of any given mask as follows.

$$\hat{f}^{(n)}(f_u, f_a) \triangleq \begin{cases} \frac{1}{2\epsilon f_0}, & \text{for } |f_u| \leq f_0 \text{ and } |f_a - (d_n/d_r)f_u| \leq \frac{\epsilon}{2} \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

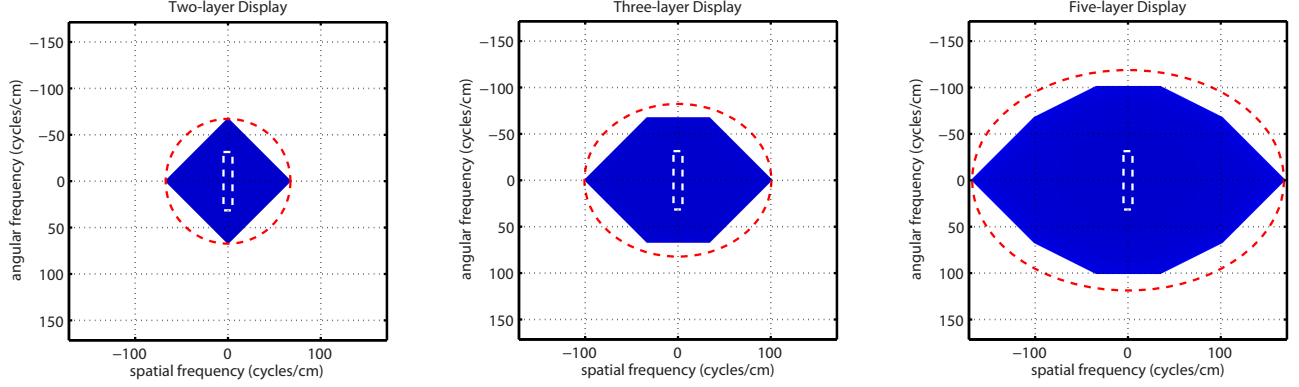


Figure 6. Spectral support for multi-layer displays. The spectral support (shaded blue) is illustrated for two-layer (left), three-layer (middle), and five-layer (right) displays, evaluated using the geometric construction given by Eqn. 18. Note that the shaded area indicates the achievable region of non-zero spectral support. The system parameters correspond with the Layered 3D prototype presented in Sec. 3.3, with the variable $d_r = h/2$. The ellipse corresponding to the upper bound on achievable spatio-angular frequencies, given by Eqn. 28, is denoted by a dashed red line. Note that the spectral support of a multi-layer display exceeds the bandwidth of a conventional, two-layer automultiscopic display with a similar physical extent, shown as a dashed white line and given by Eqn. 15.

This definition corresponds to a thin mask with a spatially-varying opacity characterized by a white noise process (i.e., a uniform spectral energy density).²² Note that, as derived by Chai et al.¹⁹, the spectrum of a mask at d_n will be non-zero along a single line, with slope d_n/d_r , only in the limit as ϵ tends to zero.

5.2.1 Two-Layer Displays

For a fixed number of layers, the depth of field can be algebraically or geometrically constructed directly by substituting Eqn. 19 into Eqn. 18. Note that, for the purposes of deriving an upper bound, the specific form of Eqn. 19 is immaterial—only the limited spectral extent is considered. For example, consider the case of two masks, separated by a distance h . The emitted light field spectrum is given by

$$\hat{l}(f_u, f_a) = \hat{f}^{(1)}(f_u) \delta(f_a - (h/(2d_r))f_u) \otimes \hat{f}^{(2)}(f_u) \delta(f_a + (h/(2d_r))f_u) \quad (20)$$

As shown in Fig. 6, this gives a diamond-shaped region enclosing the non-zero spectral support of any two-layer display. Following the approach of Zwicker et al.¹⁸, the spatial cutoff frequency f_ξ is again found by intersecting the line $f_a = (d_o/d_r)f_u$ with the boundary of the maximum-achievable spectral support. This geometric construction yields the following upper bound on the depth of field for any two-layer, attenuation-based display.

$$|f_\xi| \leq \left(\frac{h}{(h/2) + |d_o|} \right) f_0 \quad (21)$$

Comparing Eqns. 15 and 21 reveals the benefits of non-heuristically-constructed two-layer displays. For a fixed display thickness h , we conclude that conventional parallax barriers and integral imaging achieve the maximum spatial resolution far from the display. However, the resolution close to the display is reduced, by a factor up to N_a (i.e., the number of angular views), compared to the upper bound. Thus, conventional fixed spatio-angular resolution tradeoffs²³ compromise resolution for virtual objects near the display. In comparison, HR3D allows accurate depiction of virtual objects at the full mask resolution near the display, while exhibiting a depth of field similar to conventional methods far from the display.

5.2.2 Multi-Layer Displays

The upper bound for an arbitrary number of layers can be assessed using similar methods. Fig. 6 shows the geometric construction of the spectral support for a three-layer display, composed of masks at $d_1 = -h/2$, $d_2 = 0$,

and $d_3 = h/2$. The resulting depth of field is given by the following expression.

$$|f_\xi| \leq \begin{cases} \left(\frac{3h/2}{(h/2)+|d_o|} \right) f_0, & \text{for } |d_o| \leq h \\ \left(\frac{h}{|d_o|} \right) f_0, & \text{otherwise} \end{cases} \quad (22)$$

Thus, the display bandwidth exceeds conventional two-layer architectures, motivating the development of multi-layer decompositions for enhanced 3D display, such as Layered 3D. Note that one may construct a tight upper bound for any number of layers; however, this construction becomes geometrically cumbersome. Instead, we apply the central limit theorem²⁴ to obtain an approximate expression for the spectral support due to the repeated convolutions in Eqn. 18; as derived by Chaudhury et al.²⁵, the repeated convolution of N_l two-dimensional spectra, each with mean $\mu_k = [0, 0]$ and covariance matrix $\mathbf{\Sigma}_k$, tends to a bivariate Gaussian distribution, such that

$$\hat{l}(f_u, f_a) \approx \left(\frac{1}{2\pi|\mathbf{\Sigma}|^{\frac{1}{2}}} \right) \exp \left(-\frac{1}{2} \mathbf{f}^\top \mathbf{\Sigma}^{-1} \mathbf{f} \right), \text{ for } \mathbf{\Sigma} = \sum_{k=1}^{N_l} \mathbf{\Sigma}_k, \quad (23)$$

where $\mathbf{f} = [f_u, f_a]$ is a given spatio-angular frequency. The cumulative mean spatio-angular frequency $\mu = \mu_k = 0$, since each mask spectrum is symmetric about the origin as defined in Eqn. 19. The covariance matrix for each mask is given by substituting Eqn. 19, such that

$$\lim_{\epsilon \rightarrow 0} \mathbf{\Sigma}_n = \lim_{\epsilon \rightarrow 0} \begin{bmatrix} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_u^2 \hat{f}^{(n)}(f_u, f_a) df_u df_a & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_u f_a \hat{f}^{(n)}(f_u, f_a) df_u df_a \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_u f_a \hat{f}^{(n)}(f_u, f_a) df_u df_a & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_a^2 \hat{f}^{(n)}(f_u, f_a) df_u df_a \end{bmatrix} = \frac{f_0^2}{3} \begin{bmatrix} 1 & \frac{d_n}{d_r} \\ \frac{d_n}{d_r} & \left(\frac{d_n}{d_r} \right)^2 \end{bmatrix}, \quad (24)$$

where the limit of ϵ tending to zero corresponds to an infinitesimally-thin mask. Thus, the cumulative covariance matrix of the emitted light field spectrum is approximated by the summation

$$\mathbf{\Sigma} = \sum_{n=1}^N \mathbf{\Sigma}_n = \begin{bmatrix} \sigma_{f_u}^2 & 0 \\ 0 & \sigma_{f_a}^2 \end{bmatrix} = \frac{f_0^2}{3} \begin{bmatrix} N & 0 \\ 0 & \left(\frac{N(N+1)}{12(N-1)} \right) \left(\frac{h}{d_r} \right)^2 \end{bmatrix}, \quad (25)$$

for masks uniformly-spaced on the interval $d_n \in [-h/2, h/2]$. Substituting into Eqn. 23 yields the following approximation for the spectrum of an N -layer display.

$$\hat{l}(f_u, f_a) \approx \left(\frac{1}{2\pi\sigma_{f_u}\sigma_{f_a}} \right) \exp \left(-\frac{f_u^2}{2\sigma_{f_u}^2} - \frac{f_a^2}{2\sigma_{f_a}^2} \right) \quad (26)$$

Equation 26 can be used to characterize the upper bound on the depth of field, depending on the number of layers N , the display thickness h , and the mask cutoff frequency $f_0 = 1/2p$. As shown in Fig. 6, the spectral support of a multi-layer display is approximated by a bivariate Gaussian. Note that curves defining spatio-angular frequencies with equal expected modulation energies correspond to ellipses, such that

$$\frac{f_u^2}{\sigma_{f_u}^2} + \frac{f_a^2}{\sigma_{f_a}^2} = \lambda^2, \quad (27)$$

where $\pm\lambda\sigma_{f_u}$ and $\pm\lambda\sigma_{f_a}$ are the points of intersection with the f_u -axis and f_a -axis, respectively. At this point, one can obtain an approximate upper bound on the depth of field, for an arbitrary number of layers, by finding the intersection of the line $f_a = (d_o/d_r)f_u$ with the ellipse corresponding to the highest spatial frequency achievable by an N -layer display. From Eqns. 18 and 19, the repeated convolution of N masks, each extending over $\pm f_0$ along the f_u -axis, will produce non-zero spatial frequencies within the region $|f_u| \leq Nf_0$. Thus, the ellipse with $\lambda = Nf_0/\sigma_{f_u}$ provides the following approximate upper bound on the depth of field for an N -layer display.

$$|f_\xi| \leq Nf_0 \sqrt{\frac{(N+1)h^2}{(N+1)h^2 + 12(N-1)d_o^2}} \quad (28)$$

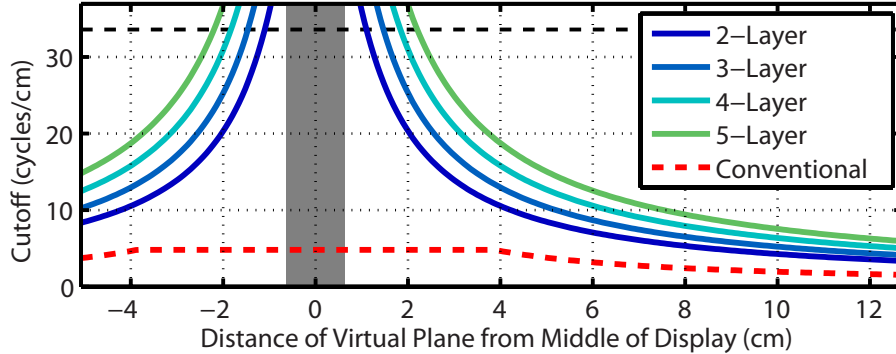


Figure 7. Upper bound on multi-layer depth of field. The spatial cutoff frequency is shown for conventional parallax barriers and integral imaging, using Eqn. 15, and for N -layer displays, using Eqn. 28. Parameters correspond with the Layered 3D prototype presented in Sec. 3.3. Spatial resolution exceeds conventional architectures, particularly near or within the display enclosure (gray shaded region). The maximum spatial resolution of a single mask is denoted by a horizontal black dashed line, indicating full-resolution display is possible within the enclosure.

This expression approximates the depth of field for two-layer and three-layer displays, given exactly by Eqns. 21 and 22, respectively. A comparison of the upper bound for multi-layer vs. conventional automultiscopic displays is shown in Fig. 7. Note that additional layers significantly increase the upper bound on the achievable spatial resolution, which is expected due to repeated convolution of mask spectra via Eqn. 18. In summary, the upper bound indicates the theoretical origin for the increased depth of field observed with HR3D, Layered 3D, and Polarization Fields, as compared to conventional designs making a fixed spatio-angular resolution tradeoff.

6. DISCUSSION AND CONCLUSIONS

Through the HR3D, Layered 3D, and Polarization Field display architectures, we endeavor to establish multi-layer attenuation-based displays as a viable design alternative to conventional two-layer parallax barrier and integral imaging displays. As established in Secs. 2.2, 3.2, and 4.2, light fields can be optimally decomposed into time-multiplexed, multi-layer attenuation patterns by applying formal optimization methods together with a first-principles analysis of the corresponding image formation model for each architecture. This research direction has, to date, uncovered two unexplored generalizations of conventional parallax barriers: time-multiplexed two-layer designs with continuously-varying opacities on each layer (i.e., HR3D) and tomographically-optimized multi-layer architectures (i.e., Layered 3D and Polarization Fields). Most significantly, such designs achieve increased brightness, spatial resolution, and depth of field compared to conventional automultiscopic display architectures. As formally established in Sec. 5, the resolution of such generalized multi-layer displays can be significantly enhanced near the display surface, even exceeding the native display resolution for a single layer.

Our multi-layer generalization of parallax barriers opens the door to similar modifications of existing display technologies. Similar to Lanman et al.⁶, it may be possible to exploit temporal modulation to obtain more accurate light field reconstructions using more than two light-attenuating layers. While the upper bound on depth of field indicates a potentially significant gain in spatio-angular resolution (see Fig. 7), factorization methods must first be developed for such dynamically-modulated stacked displays. In contrast, combinations of additive and multiplicative layers may yield similar gains, while also enhancing brightness. Such displays are efficiently modeled with emission-absorption volume rendering.²⁶ Finally, our general image formulation model further facilitates the development of non-planar, volumetric displays with arbitrary curved surfaces.

ACKNOWLEDGMENTS

We recognize the support of the MIT Camera Culture Group and the UBC Imager Laboratory. We also thank the MIT Media Lab sponsors. Douglas Lanman was supported by NSF Grant IIS-1116452 and DARPA Grant HR0011-10-C-0073. Gordon Wetzstein was supported by a UBC Four Year Fellowship. Wolfgang Heidrich was supported under the Dolby Research Chair in Computer Science at UBC. Ramesh Raskar was supported by an Alfred P. Sloan Research Fellowship and a DARPA Young Faculty Award.

REFERENCES

- [1] Ives, F. E., “Parallax stereogram and process of making same.” U.S. Patent 725,567 (1903).
- [2] Lippmann, G., “Épreuves réversibles donnant la sensation du relief,” *Journal of Physics* **7**(4), 821–825 (1908).
- [3] Zwicker, M., Vetro, A., Yea, S., Matusik, W., Pfister, H., and Durand, F., “Resampling, antialiasing, and compression in multiview 3-D displays,” *IEEE Signal Processing Magazine* **24**, 88–96 (November 2007).
- [4] Perlin, K., Paxia, S., and Kollin, J. S., “An autostereoscopic display,” in *[ACM SIGGRAPH]*, 319–326 (2000).
- [5] Peterka, T., Kooima, R. L., Sandin, D. J., Johnson, A., Leigh, J., and DeFanti, T. A., “Advances in the dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system,” *IEEE TVCG* **14**(3), 487–499 (2008).
- [6] Lanman, D., Hirsch, M., Kim, Y., and Raskar, R., “Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization,” *ACM Trans. Graph.* **29**, 163:1–163:10 (2010).
- [7] Wetzstein, G., Lanman, D., Heidrich, W., and Raskar, R., “Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays,” *ACM Trans. Graph.* **30**(4) (2011).
- [8] Lanman, D., Wetzstein, G., Hirsch, M., Heidrich, W., and Raskar, R., “Polarization fields: Dynamic light field display using multi-layer LCDs,” *ACM Trans. Graph.* **30**(6) (2011).
- [9] Kim, Y., Kim, J., Kang, J.-M., Jung, J.-H., Choi, H., and Lee, B., “Point light source integral imaging with improved resolution and viewing angle by the use of electrically movable pinhole array,” *Optics Express* **15**(26), 18253–18267 (2007).
- [10] Lee, D. D. and Seung, H. S., “Learning the parts of objects by non-negative matrix factorization,” *Nature* **401**, 788–791 (1999).
- [11] Blondel, V. D., Ho, N.-D., and van Dooren, P., “Weighted nonnegative matrix factorization and face feature extraction,” *Image and Vision Computing* (2008).
- [12] Coleman, T. and Li, Y., “A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables,” *SIAM Journal on Optimization* **6**(4), 1040–1058 (1996).
- [13] Jones, R. C., “A new calculus for the treatment of optical systems,” *J. Opt. Soc. Am.* **31**(7), 488–493 (1941).
- [14] Yeh, P. and Gu, C., *[Optics of Liquid Crystal Displays]*, John Wiley and Sons (2009).
- [15] Andersen, A. and Kak, A., “Simultaneous Algebraic Reconstruction Technique (SART): A superior implementation of the ART algorithm,” *Ultrasonic Imaging* **6**(1), 81–94 (1984).
- [16] Kak, A. C. and Slaney, M., *[Principles of Computerized Tomographic Imaging]*, Society for Industrial Mathematics (2001).
- [17] Kaczmarz, S., “Angenäherte auflösung von systemen linearer gleichungen,” *Bull. Acad. Pol. Sci. Lett. A* **35**, 335–357 (1937).
- [18] Zwicker, M., Matusik, W., Durand, F., and Pfister, H., “Antialiasing for automultiscopic 3D displays,” in *[Eurographics Symposium on Rendering]*, (2006).
- [19] Chai, J.-X., Tong, X., Chan, S.-C., and Shum, H.-Y., “Plenoptic sampling,” in *[ACM SIGGRAPH]*, 307–318 (2000).
- [20] Durand, F., Holzschuch, N., Soler, C., Chan, E., and Sillion, F. X., “A frequency analysis of light transport,” in *[ACM SIGGRAPH]*, 1115–1126 (2005).
- [21] Bracewell, R., *[The Fourier Transform and Its Applications (Third Edition)]*, McGraw-Hill (1999).
- [22] Haykin, S., *[Communication Systems (Fourth Edition)]*, Wiley (2000).
- [23] Georgiev, T., Zheng, K. C., Curless, B., Salesin, D., Nayar, S., and Intwala, C., “Spatio-angular resolution tradeoff in integral photography,” in *[Eurographics Symposium on Rendering]*, 263–272 (2006).
- [24] Peebles, P., *[Probability, Random Variables, and Random Signal Principles (Fourth Edition)]*, McGraw-Hill (2000).
- [25] Chaudhury, K. N., Muñoz-Barrutia, A., and Unser, M., “Fast space-variant elliptical filtering using box splines,” *IEEE Trans. Image* **19**(9), 2290–2306 (2010).
- [26] Sabella, P., “A rendering algorithm for visualizing 3D scalar fields,” *ACM SIGGRAPH* **22**, 51–58 (1988).