Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting

Gordon Wetzstein Douglas Lanman Matthew Hirsch Ramesh Raskar
MIT Media Lab

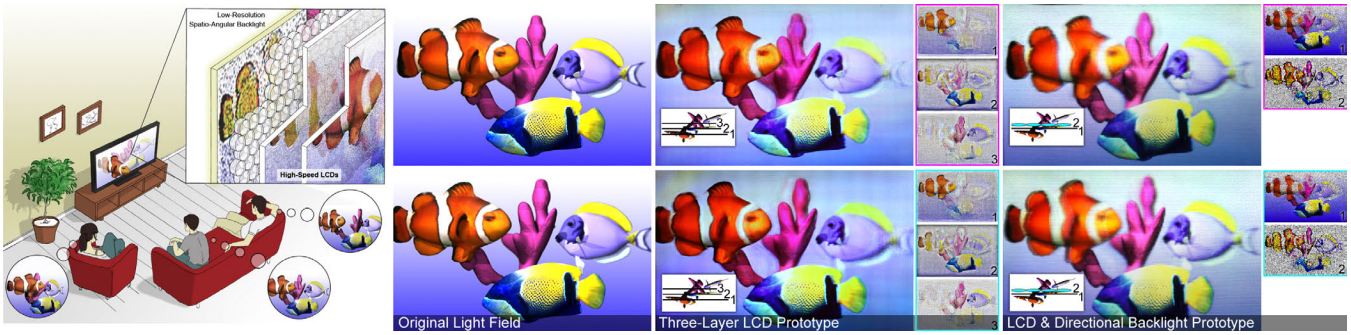

Figure 1: *Wide field of view glasses-free 3D display using tensor displays. (Left) We introduce a new family of light field displays, dubbed tensor displays, comprised of stacks of light-attenuating layers (e.g., multilayer LCDs). Rapid temporal modulation of the layers is exploited, in concert with directional backlighting, to allow large separations between viewers. (Right) From left to right: target light field view, photograph of three-layer LCD with uniform backlighting, and photograph of single LCD with directional backlighting. Layers are shown to the right of each photograph. The upper and lower rows depict perspectives seen to the left and to the right of the display, respectively.*

Abstract

We introduce tensor displays: a family of compressive light field displays comprising all architectures employing a stack of time-multiplexed, light-attenuating layers illuminated by uniform or directional backlighting (i.e., any low-resolution light field emitter). We show that the light field emitted by an N-layer, M-frame tensor display can be represented by an N^{th}-order, rank-M tensor. Using this representation we introduce a unified optimization framework, based on nonnegative tensor factorization (NTF), encompassing all tensor display architectures. This framework is the first to allow joint multilayer, multiframe light field decompositions, significantly reducing artifacts observed with prior multilayer-only and multiframe-only decompositions; it is also the first optimization method for designs combining multiple layers with directional backlighting. We verify the benefits and limitations of tensor displays by constructing a prototype using modified LCD panels and a custom integral imaging backlight. Our efficient, GPU-based NTF implementation enables interactive applications. Through simulations and experiments we show that tensor displays reveal practical architectures with greater depths of field, wider fields of view, and thinner form factors, compared to prior automultiscopic displays.

Keywords: light fields, automultiscopic 3D displays, multilayer LCDs, directional backlighting, nonnegative tensor factorization

Links: DL PDF WEB VIDEO DATA CODE

1 Introduction

Consumer stereoscopic displays have been enabled by the introduction of high-speed LCDs and inexpensive shutter glasses. Although adoption is gradual, these displays are further supported by an expanding content stream, including theatrical and sports productions, interactive entertainment, and stereoscopic cameras. Manufacturers are beginning to offer automultiscopic (glasses-free) 3D displays, primarily based on the century-old concepts of parallax barriers [Ives 1903] and integral imaging [Lippmann 1908]. Early products are restricted to mobile devices for which the limitations of these methods, particularly narrow fields of view and reduced spatial resolution, do not preclude adoption. Researchers are advancing a wide variety of technologies to address these limitations, spanning volumetric to holographic methods. Yet, with the widespread use of mobile devices, research must increasingly focus on solutions that preserve the thin form factors, low power consumption, and high resolution expected from modern display technologies.

We are inspired to address the limitations of existing automultiscopic displays by taking advantage of three emerging display technologies: multilayer panels, high-speed temporal modulation, and directional backlighting. As early as 1920, Louis Lumière stacked backlit films to approximate the appearance of extended objects [Lumière 1920]. More recently, researchers have reinvestigated layered displays, considering the benefits of dynamic LCDs and optimization of the layer patterns. Research is divided into approaches using a single pair of temporally-modulated layers vs. approaches using three or more static layers. With high-speed temporal modulation, an observer perceives the time average of a multiframe sequence of patterns displayed across the layers. While enhancing performance relative to conventional methods, multilayer decompositions increase display thickness and exhibit artifacts that cannot be eliminated by increasing layers. Similarly, multiframe decompositions require high-speed displays to mitigate artifacts, yet result in dimmer images. In both cases, the field of view is typically restricted to a narrow region. Yet, the recent emergence of directional backlighting, consisting of a fast-switching display paired with a rear-illuminating light guide, has enabled the pro-

jection of autostereoscopic imagery. Thus, we aim for a hybrid approach combining the advantages of multilayer and multiframe decompositions, together with directional backlighting, to achieve wider fields of view with thin form factor display architectures.

We introduce a unified optimization framework combining the benefits of multiple layers and temporal modulation, enabled by a new tensor representation for light field displays. Using this representation we extend nonnegative tensor factorization (NTF) to optimize prior multilayer and multiframe displays, while providing the first joint multilayer, multiframe decompositions. With the combined degrees of freedom from additional layers and frames, the resulting *tensor displays* exhibit increased image fidelity compared to multilayer-only and multiframe-only decompositions. Our framework further supports directional backlighting (i.e., any low-resolution light field emitter). Our prototype tensor display (see Figures 1 and 2) demonstrates that directional backlighting achieves a wide field of view, while reducing the need for additional layers and frames, yielding a thin, power-efficient, high-resolution light field display well suited for mobile and home theater applications.

1.1 Contributions

Our primary contribution is to introduce and characterize tensor displays. Additional technical contributions include the following:

- We show that any light field emitted by an N-layer, M-frame tensor display is represented by an N^{th}-order, rank-M tensor. The light field tensor is decomposed as a sum of M rank-1 tensors, each corresponding to the outer product of N masks representing the transmittance of each layer for each frame.
- We extend our tensor representation to support emerging displays incorporating directional backlighting consisting of any low-resolution light field emitter placed behind the layers.
- We introduce a unified optimization framework for tensor displays using nonnegative tensor factorization (NTF). It is the first to allow multilayer, multiframe decompositions and to combine benefits of multiple layers and directional backlights.
- We implement a reconfigurable tensor display prototype using modified LCD panels and a custom integral imaging backlight. We also implement an efficient, GPU-based NTF solver.
- We show, via simulation and experiment, that tensor displays achieve greater depths of field, wider fields of view, and thinner form factors compared to prior automultiscopic displays.

1.2 Overview of Benefits and Limitations

The primary benefit of tensor displays is to combine advantages of multiple layers, high-speed temporal modulation, and directional backlighting to depict brighter images with fewer artifacts, over a greater depth of field and a wider field of view, than prior approaches. Our framework enables trade-offs between image fidelity, resolution, brightness, and display complexity. These benefits primarily stem from our development of *compressive display modes*, wherein low-rank tensor approximation efficiently exploits correlations between neighboring views to synthesize an emitted light field with an apparent number of views exceeding the number of available frames. In contrast, prior *direct display modes* assign a single view to each frame, limiting resolution and brightness. Furthermore, such direct display modes have only been proposed for a subset of tensor display architectures. Similar to other multilayer LCDs, we support a full resolution 2D mode by rendering all but one of the layers transparent. We identify single-layer directional backlight displays as a practical solution for glasses-free 3D display, requiring minimal components and projecting bright images with a wide field of view and a thin form factor.

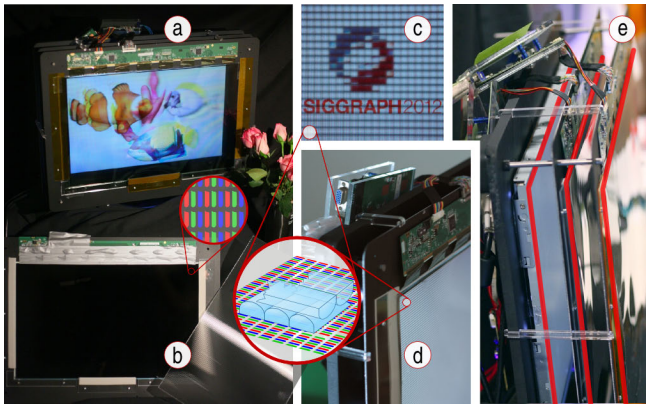

Figure 2: *The tensor display prototype. (a) The prototype configured as a three-layer display, photographed outside the optimized viewing zone so layer patterns are individually visible. (b) An LCD layer mounted on an aluminum frame (left) and a lenticular sheet (right). (c) The directional backlight, consisting of two crossed lenticular sheets on top of the rear LCD (inset). High-resolution text is shown on an LCD layer suspended in front of the directional backlight. (d) The single-layer directional backlight configuration. (e) The three-layer configuration, with layers highlighted in red.*

For designs with more than one layer, tensor displays exhibit moiré, color-channel crosstalk, and decreased brightness. Inclusion of multiple layers or directional backlighting increases cost and complexity, introduces scattering and interreflections, and requires accurate mechanical alignment. Nonnegative tensor factorization (NTF) requires significant computational resources, particularly for the multiplicative update rules used in our implementation—currently requiring a GPU-based solver to approach interactive refresh rates. Unlike conventional parallax barriers and integral imaging, viewing zones do not repeat periodically. Outside of the central viewing zone and, to a lesser degree, between target viewing positions, the solution is unconstrained and artifacts are observed.

2 Related Work

Glasses-free 3D Displays have been studied for more than a century, with early works including those of Ives [1903] and Lippmann [1908]. Such methods trade spatial resolution for angular resolution. With advances in computation and display technology, researchers have integrated viewer tracking [Perlin et al. 2000; Peterka et al. 2008], image compression [Matusik and Pfister 2004], electronically-switchable displays [Jacobs et al. 2003], and temporal multiplexing [Kim et al. 2007]. By exploiting temporal multiplexing with dual-layer displays, in combination with low-rank matrix factorizations, these architectures can be optimized in terms of image fidelity, brightness, and frame rate [Lanman et al. 2010]. We are the first to explore similar concepts for multilayer architectures using multilinear algebra; our framework also facilitates efficient image synthesis for emerging directional backlights.

Recent multilayer 3D displays have been demonstrated to achieve high-fidelity light field synthesis. However, such displays are restricted either to static images [Wetzstein et al. 2011; Holroyd et al. 2011] or to shallow depths of field and narrow fields of view [Putilin et al. 2001; Gotoda 2010; Gotoda 2011; Lanman et al. 2011]. In contrast, we introduce a tensor framework that allows both multiple stacked display layers and temporal multiplexing of the displayed content. With a pair of dynamic, full-color display prototypes we demonstrate how depth of field and field of view can be significantly improved in comparison to previous automultiscopic displays.

Other 3D display design paradigms include volumetric displays [Favalora 2005] and stacks of light-emitting, rather than light-attenuating, layers [Akeley et al. 2004]. Volumetric devices usually require mechanically moving parts [Cossairt et al. 2007; Jones et al. 2007] or time-multiplexed diffusers [Sullivan 2003]. The majority of such volumetric displays can only depict 3D content that is confined within the physical device enclosure, excluding the light field displays proposed by Cossairt et al. and Jones et al. In contrast to the additive image formation model inherent to most volumetric displays, our optical designs exploit multiplicative light attenuation to allow synthesized 3D objects to extend outside the enclosure. Furthermore, tensor displays support specularities, occlusions, and global illumination effects, without requiring moving parts.

Directional Backlights are an emerging trend in display technology. The combination of a fast-switching LCD and a rear-illuminating light guide allows stereoscopic [Travis 1990; Toyooka et al. 2001; Chu et al. 2005; Chien and Shieh 2006; Brott and Schultz 2010] and multiscopic image synthesis [Mather et al. 2009; Travis et al. 2009]. Stolle et al. [2008] and Kwon and Choi [2012] implement multidirectional backlighting using lenslet arrays. We employ a similar design in one of our prototypes: combining a low-resolution lenslet-based backlight with a high-resolution LCD. In contrast to prior work employing time-sequential directional backlights, we decompose a target light field into a low-rank tensor approximation, increasing brightness and allowing more views to be generated than available frames. Furthermore, our multilinear framework is the first to support arbitrary combinations of directional backlights and multiple light-attenuating display layers.

Factorizations and Graphics have a long tradition together. Tensor factorizations of multilinear visual data, including textures [Vasilescu and Terzopoulos 2004], BRDFs [Bilgili et al. 2011], and BTFs and volume sequences [Wang et al. 2005], have applications in modeling and image-based rendering. Low-rank matrix factorizations have been used for efficient representations of subsurface scattering [Peers et al. 2006] and intuitive editing of spatially-varying BRDFs [Lawrence et al. 2006].

3 Tensor Displays

This section presents a unifying framework for depicting arbitrary light fields using tensor displays. First, we introduce a tensor representation for multilayer displays illuminated by a uniform backlight. The light field emitted by an N-layer, M-frame display is represented by a sparse set of non-zero elements restricted to a plane within an N^{th}-order, rank-M tensor. Second, we show that this tensor representation allows for optimal decomposition of a light field into time-multiplexed, light-attenuating layers using non-negative tensor factorization (NTF). Third, we demonstrate that our tensor representation also allows optimization of multilayer displays illuminated by a directional backlight. We conclude by interpreting the structure of tensor display decompositions.

3.1 Representing Multilayer Displays with Tensors

As shown in Figure 3, tensor displays consist of a stack of N light-attenuating layers illuminated by either a conventional uniform backlight or a directional backlight. For full generality, we assume that display layers support synchronized, high-speed temporal modulation, such that an observer perceives the time average of an M-frame multilayer mask sequence. We consider 2D light fields and 1D layers in the following analysis, with the extension to 4D light fields and 2D layers covered in Section 3.2. A relative two-plane light field parameterization $l(x, v)$ is adopted, shown in Figure 3, where v denotes the point of intersection of the ray (x, v)

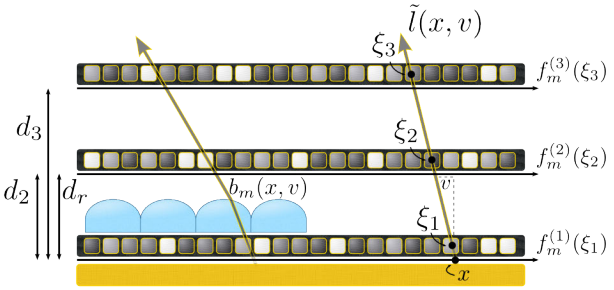

Figure 3: *Tensor display coordinates. A stack of N light-attenuating layers is illuminated by a uniform or directional backlight (here depicted as a lenslet array affixed to the rear display).*

with a plane located a distance d_r from the x-axis, expressed relative to x [Chai et al. 2000; Durand et al. 2005]. In the following analysis we assume familiarity with multilinear algebra, particularly tensor notation; consult Kolda and Bader [2009] for a review.

3.1.1 Static Multilayer Displays

Consider a fixed stack of N light-attenuating layers (i.e., one that does not support temporal variation of the mask patterns). When illuminated by a uniform backlight with unit radiance, the emitted light field $\tilde{l}(x, v)$ is given by the following expression:

$$\tilde{l}(x, v; N) = \prod_{n=1}^{N} f^{(n)}(x + (d_n/d_r)v), \tag{1}$$

where $f^{(n)}(\xi_n) \in [0, 1]$ is the transmittance at the point ξ_n of layer n, separated a distance d_n from the x-axis. Consider a three-layer configuration, with the transmittances for the rear, middle, and front layers given by $f(\xi_1)$, $g(\xi_2)$, and $h(\xi_3)$, respectively. Equation 1 gives the following expression for the emitted light field.

$$\tilde{l}(x, v) = f(\xi_1)\, g(\xi_2)\, h(\xi_3), \text{ for } \xi_n = x + (d_n/d_r)v \tag{2}$$

We observe that the emitted light field $\tilde{l}(x, v)$ can be represented as the restriction of the function

$$\tilde{t}(\xi_1, \xi_2, \xi_3) = f(\xi_1)\, g(\xi_2)\, h(\xi_3), \tag{3}$$

defined in the three-dimensional Euclidean space $\mathbb{R}^3$ spanned by $\{\xi_1, \xi_2, \xi_3\}$, to the two-dimensional subspace defined by the equation $\alpha\xi_1 + \beta\xi_2 + \gamma\xi_3 = 0$, with

$$\alpha = d_3 - d_2, \quad \beta = d_1 - d_3, \quad \gamma = d_2 - d_1. \tag{4}$$

Thus, as shown at the top of Figure 4, elements of the emitted light field $\tilde{l}(x, v)$ are restricted to the plane corresponding to Equation 4.

For the general case with $N > 3$ layers, the emitted light field $\tilde{l}(x, v)$ can also be represented as the restriction of the function $\tilde{t}(\xi_1, \xi_2, \ldots, \xi_N) = \prod_{n=1}^{N} f^{(n)}(\xi_n)$, defined on $\mathbb{R}^N$, to a plane.

In practice, each layer has discrete pixels with constant transmittances rather than continuously-varying opacities. As a result, we tabulate the transmittance $f_{i_n}^{(n)}$ at each pixel i_n within the vector $\mathbf{f}^{(n)}$. As shown in Figure 3, each light field ray (x, v) can be equivalently parameterized by the corresponding points of intersection $\{\xi_1, \xi_2, \ldots, \xi_N\}$ with each layer. For a three-layer display with discrete pixels, the intensity of the emitted light field $\tilde{l}(\xi_1, \xi_2, \xi_3)$ is approximated by the product $f_i g_j h_k$, where $\{i, j, k\}$ denote the pixel indices nearest to the points of intersection $\{\xi_1, \xi_2, \xi_3\}$. With

this parameterization we observe that Equation 3 can be represented in discrete coordinates as a 3rd-order, rank-1 tensor $\tilde{\mathcal{T}}$, given by

$$\tilde{\mathcal{T}} = \mathbf{f} \circ \mathbf{g} \circ \mathbf{h}, \text{ such that } \tilde{t}_{ijk} = f_i g_j h_k, \tag{5}$$

where $\circ$ is the vector outer product. Note that only a subset of tensor elements $\tilde{t}_{ijk}$ correspond to valid light field rays; most tensor elements correspond to "non-physical" rays (i.e., ones that spontaneously change position or direction after passing through a layer). To address this limitation of our tensor representation, we further define a sparse, binary-valued weight tensor $\mathcal{W}$ such that the emitted light field tensor $\tilde{\mathcal{L}}$ is given by the following expression:

$$\tilde{\mathcal{L}} = \mathcal{W} \circledast \tilde{\mathcal{T}}, \text{for } w_{ijk} = \begin{cases} 1 & \text{if } \{i,j,k\} \text{ gives a light field ray,} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where $\circledast$ is the Hadamard (elementwise) product. Following Figure 4, non-zero elements of $\tilde{\mathcal{L}}$ are close to the plane defined by Equation 4. We conclude that tensors provide sparse, memory-efficient representations for static N-layer displays; as described in Section 5.1.2, only the non-zero elements of $\tilde{\mathcal{L}}$ must be stored.

3.1.2 Time-Multiplexed Multilayer Displays

Following Section 1, static multilayer displays have finite degrees of freedom. Artifacts, resulting from limited depths of field and fields of view, persist in the emitted light field, as observed by Gotoda [2010; 2011] and Wetzstein et al. [2011]. The degrees of freedom must be increased to mitigate artifacts, typically observed as blur. We propose exploiting rapid temporal modulation, such that the observer perceives the average of an M-frame sequence. Generalizing Equation 1, the emitted light field $\tilde{l}(x, v)$ is given by

$$\tilde{l}(x, v; N, M) = \frac{1}{M} \sum_{m=1}^{M} \prod_{n=1}^{N} f_m^{(n)}(x + (d_n/d_r)v), \tag{7}$$

where $f_m^{(n)}(\xi_n)$ is the transmittance at the point ξ_n of layer n during frame m. Let columns of the matrix $\mathbf{F}^{(n)} = [\mathbf{f}_1^{(n)}\ \mathbf{f}_2^{(n)} \cdots \mathbf{f}_M^{(n)}]$ define the sequence of M masks displayed on layer n. For a three-layer display, Equation 7 can be represented in discrete coordinates as a 3rd-order, rank-M tensor $\tilde{\mathcal{T}}$ given by

$$\tilde{\mathcal{T}} = [\![\mathbf{F}, \mathbf{G}, \mathbf{H}]\!] \equiv \frac{1}{M} \sum_{m=1}^{M} \mathbf{f}_m \circ \mathbf{g}_m \circ \mathbf{h}_m, \tag{8}$$

where matrices enclosed by double square brackets correspond to the *CP decomposition* of a tensor into a sum of rank-1 tensors [Cichocki et al. 2009]. The CP decomposition is equivalent to CANDECOMP (canonical decomposition) and PARAFAC (parallel factors), with elements of the tensor $\tilde{\mathcal{T}}$ given by $\tilde{t}_{ijk} = \frac{1}{M} \sum_{m=1}^{M} f_{im} g_{jm} h_{km}$ [Kolda and Bader 2009]. For the general case with N light-attenuating layers and M time-multiplexed frames, we observe that the emitted light field can be represented as an N^{th}-order, rank-M tensor $\tilde{\mathcal{T}} = [\![\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \ldots, \mathbf{F}^{(N)}]\!]$.

3.2 Synthesizing Light Fields

Light field synthesis with time-multiplexed, multilayer displays requires decomposing a target light field $l(x, v)$ into an M-frame sequence of N transmittance functions $f_m^{(n)}(\xi_n)$. This can be formulated as the following constrained nonlinear least squares problem:

$$\underset{\{f_m^{(n)}(\xi_n)\}}{\arg\min} \int_{\mathcal{V}} \int_{\mathcal{X}} \Big(l(x, v) - \tilde{l}(x, v) \Big)^2 dx dv, \text{for } 0 \le f_m^{(n)}(\xi_n) \le 1, \tag{9}$$

where $\tilde{l}(x, v)$ is the emitted light field, given by Equation 7, and $\mathcal{X}$ and $\mathcal{V}$ denote the intervals $[x_{\min}, x_{\max}]$ and $[v_{\min}, v_{\max}]$.

The tensor representation introduced in Section 3.1 provides an efficient means for solving Equation 9. Using this representation for a three-layer configuration with discrete coordinates, the objective function is expressed as

$$\underset{\mathbf{F},\mathbf{G},\mathbf{H}}{\arg\min} \left\| \mathcal{L} - \mathcal{W} \circledast [\![\mathbf{F}, \mathbf{G}, \mathbf{H}]\!] \right\|^2, \text{ for } 0 \le \mathbf{F}, \mathbf{G}, \mathbf{H} \le 1, \tag{10}$$

where $\mathcal{L}$ is the target light field tensor, obtained by assigning the target light field $l(x, v)$ to the plane defined by Equation 4, and $\|\mathcal{X}\|^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} x_{ijk}^2$ is the squared tensor norm of $\mathcal{X}$. We observe that this expression can be solved by applying weighted nonnegative tensor factorization (NTF). Following Cichocki et al. [2009], a broad set of procedures have emerged for the solution of NTF problems. In this paper we use multiplicative update rules that extend the weighted nonnegative matrix factorization (NMF) procedure proposed by Blondel et al. [2008] to higher-order tensors. For a three-layer display, these update rules have the following forms:

$$\mathbf{F} \leftarrow \mathbf{F} \circledast \left(\frac{(\mathbf{W}_{(1)} \circledast \mathbf{L}_{(1)})(\mathbf{H} \odot \mathbf{G})}{(\mathbf{W}_{(1)} \circledast (\mathbf{F}(\mathbf{H} \odot \mathbf{G})^{\mathsf{T}}))(\mathbf{H} \odot \mathbf{G})} \right) \tag{11}$$

$$\mathbf{G} \leftarrow \mathbf{G} \circledast \left(\frac{(\mathbf{W}_{(2)} \circledast \mathbf{L}_{(2)})(\mathbf{H} \odot \mathbf{F})}{(\mathbf{W}_{(2)} \circledast (\mathbf{G}(\mathbf{H} \odot \mathbf{F})^{\mathsf{T}}))(\mathbf{H} \odot \mathbf{F})} \right) \tag{12}$$

$$\mathbf{H} \leftarrow \mathbf{H} \circledast \left(\frac{(\mathbf{W}_{(3)} \circledast \mathbf{L}_{(3)})(\mathbf{G} \odot \mathbf{F})}{(\mathbf{W}_{(3)} \circledast (\mathbf{H}(\mathbf{G} \odot \mathbf{F})^{\mathsf{T}}))(\mathbf{G} \odot \mathbf{F})} \right) \tag{13}$$

In these expressions $\odot$ is the Khatri-Rao product, defined for a pair of matrices $\mathbf{A} \in \mathbb{R}^{I \times K}$ and $\mathbf{B} \in \mathbb{R}^{J \times K}$, such that

$$\mathbf{A} \odot \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1 \quad \mathbf{a}_2 \otimes \mathbf{b}_2 \quad \cdots \quad \mathbf{a}_K \otimes \mathbf{b}_K], \tag{14}$$

where $\otimes$ is the Kronecker product and $\mathbf{a}_i$ and $\mathbf{b}_j$ denote the i^{th} and j^{th} columns of $\mathbf{A}$ and $\mathbf{B}$, respectively. These update equations also make use of the tensor matricization (unfolding) operation, defined such that $\mathbf{X}_{(n)}$ arranges the mode-n fibers of $\mathcal{X}$ to be columns of the resulting matrix. We observe, for two layers, these weighted NTF update rules reduce to the weighted NMF update rules used by Blondel et al. [2008] and Lanman et al. [2010]. Consult Supplementary Appendix A for additional details on Equations 11–13.

For the general case with N light-attenuating layers and M frames, we observe that Equation 10 has the following form:

$$\underset{\{\mathbf{F}^{(n)}\}}{\arg\min} \left\| \mathcal{L} - \mathcal{W} \circledast \tilde{\mathcal{T}} \right\|^2, \text{ for } 0 \le \mathbf{F}^{(n)} \le 1, \tag{15}$$

where $\tilde{\mathcal{T}} = [\![\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \ldots, \mathbf{F}^{(N)}]\!]$. Similarly, the update rules are generalized such that

$$\mathbf{F}^{(n)} \leftarrow \mathbf{F}^{(n)} \circledast \left(\frac{(\mathbf{W}_{(n)} \circledast \mathbf{L}_{(n)})\mathbf{F}_{\odot}^{n}}{(\mathbf{W}_{(n)} \circledast (\mathbf{F}^{(n)}(\mathbf{F}_{\odot}^{n})^{\mathsf{T}}))\mathbf{F}_{\odot}^{n}} \right), \tag{16}$$

where $\mathbf{F}_{\odot}^{n}$ is defined by the following expression:

$$\mathbf{F}_{\odot}^{n} \equiv \mathbf{F}^{(N)} \odot \cdots \odot \mathbf{F}^{(n+1)} \odot \mathbf{F}^{(n-1)} \odot \cdots \odot \mathbf{F}^{(1)}. \tag{17}$$

4D light fields and 2D layers require vectorizing the 2D layer transmittances, giving a similar set of transmittance vectors $\{\mathbf{f}_m^{(n)}\}$. Following standard practice [Cichocki et al. 2009], values are clamped to the feasible range after each iteration of Equation 16.

In summary, our tensor representation allows for the decomposition of a target light field into a set of time-multiplexed, light-attenuating layers. As described in Section 5.1, the multiplicative update rules allow an efficient, GPU-based implementation that achieves interactive refresh rates with multilayer LCDs.

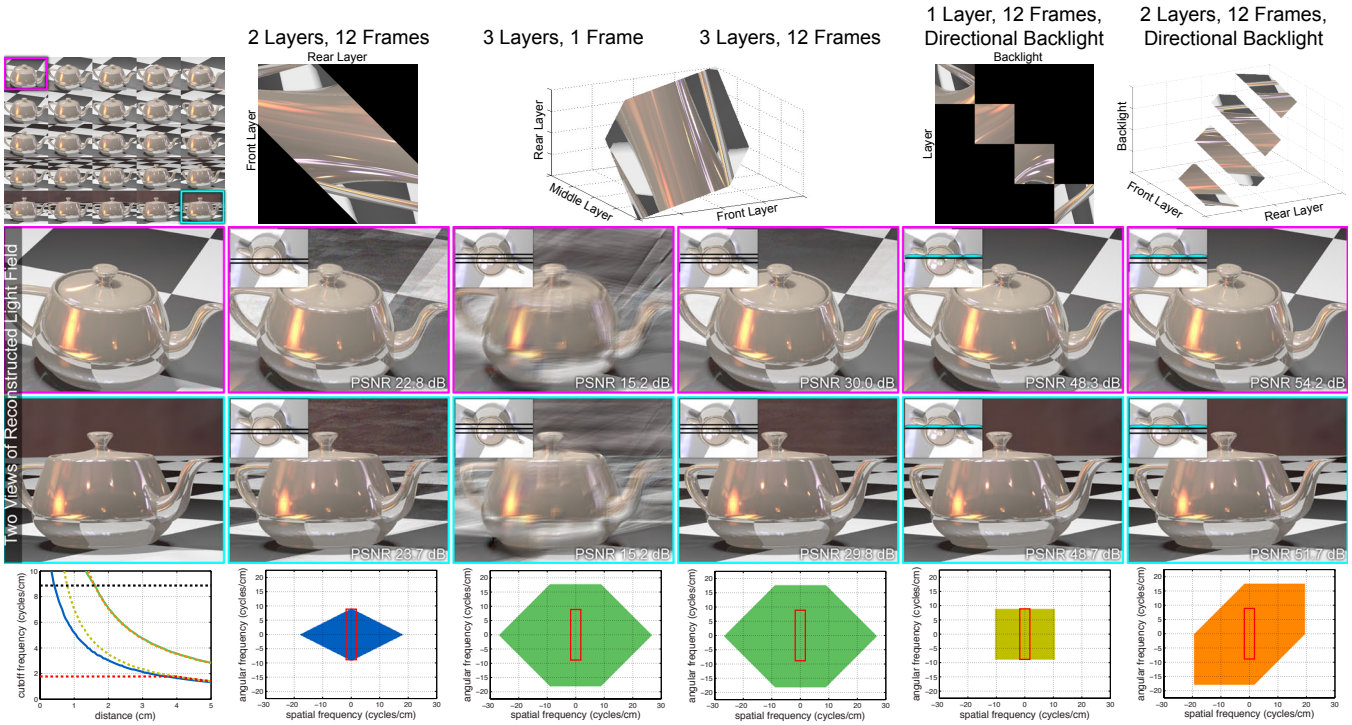

Figure 4: *Overview of tensor displays. (First Row, Left) A target light field for a teapot, rendered as 5×5 views with a 20° field of view. (First Row, Right) Visualizations of the light field, as restricted to the plane within the display tensor $\mathcal{T}$ given by Equation 4. Five architectures are compared from left to right: two-layer, 12-frame display, static three-layer display, three-layer, 12-frame tensor display, and single-layer and two-layer tensor displays using directional backlights with 12 frames (spatial backlight resolution is a quarter that of each layer). (Second and Third Rows) Two reconstructed views using each display. Note that time-multiplexing, as allowed by tensor displays, significantly reduces artifacts observed with the static three-layer configuration. (Fourth Row, Left) Upper bound on depths of field (similar to Figure 6). (Fourth Row, Right) Upper bound on the spatio-angular bandwidth for each display, as described in Section 4.1. These results demonstrate increased depth of field for tensor displays, relative to prior work, as indicated by reduced artifacts for the checkerboard and reflections in the teapot.*

3.3 Incorporating Directional Backlighting

As shown in the fourth column of Figure 4, time multiplexing significantly reduces artifacts observed with multilayer displays, as quantified by the peak signal-to-noise ratio (PSNR). Yet, such displays are still restricted to relatively narrow fields of view (i.e., $\lesssim 20°$). Expanding the field of view requires further increasing the refresh rate—a solution that may be precluded by the underlying display hardware. In this section we propose an alternate approach for achieving wider fields of view: replacing conventional uniform backlighting with time-multiplexed directional backlighting.

A directional backlight is equivalent to a low-resolution light field display. In this analysis we assume the directional backlight has significantly lower spatial resolution, but equivalent angular resolution and field of view, as compared to the target light field $l(x, v)$. Thus, our goal is to primarily enhance the spatial resolution by covering a low-resolution light field display with an N-layer stack of light-attenuating layers. Generalizing Equation 7, the light field emitted by such a display architecture is given by the following expression:

$$\tilde{l}(x, v) = \frac{1}{M} \sum_{m=1}^{M} b_m(x, v) \prod_{n=1}^{N} f_m^{(n)}(x + (d_n/d_r)v), \qquad (18)$$

where $b_m(x, v)$ denotes the light field emitted by the backlight during frame m. Let $\mathbf{B}$ denote the discrete backlight light field, such that b_{as} corresponds to pixel s of view a. The backlight light field can be equivalently represented as a vector $\mathbf{b}$, defined as follows.

$$\mathbf{b} = [\mathbf{b}_1^\mathsf{T}\ \mathbf{b}_2^\mathsf{T}\ \cdots\ \mathbf{b}_S^\mathsf{T}]^\mathsf{T}, \text{ for } \mathbf{b}_s = [b_{1s}\ b_{2s}\ \cdots\ b_{As}]^\mathsf{T} \qquad (19)$$

Using this parameterization, Equation 18 can be represented in discrete coordinates as an N+1-order, rank-M tensor $\tilde{\mathcal{T}}$, given by

$$\tilde{\mathcal{T}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{b}_m \circ \mathbf{f}_m^{(1)} \circ \mathbf{f}_m^{(2)} \circ \cdots \circ \mathbf{f}_m^{(N)}, \qquad (20)$$

where tensor element $\tilde{t}_{ij_1j_2\cdots j_N} = \frac{1}{M}\sum_{m=1}^{M} b_{im} \prod_{n=1}^{N} f_{j_n m}^{(n)}$. Since Equations 8 and 20 are similar, NTF can also be applied to optimize multilayer displays with directional backlighting.

As shown in Figure 4, directional backlighting allows multilayer displays to achieve wide fields of view, even with a single high-speed, light-attenuating layer. In summary, our tensor representation for multilayer displays provides a computationally-efficient optimization scheme encompassing a wide variety of display architectures. While providing the first method for joint multilayer, multiframe decompositions, this framework also naturally extends to emerging directional backlighting. In the following sections we further analyze the theoretical and practical benefits of display architectures supported by the tensor display framework.

3.4 Interpreting Tensor Display Decompositions

Tensor displays exploit the additional degrees of freedom arising from multiple layers and frames to achieve high-fidelity light field reconstructions. The benefits of joint multilayer, multiframe decompositions are demonstrated in Figure 4. However, these results do not provide intuition into the underlying structure of the decomposed layers. What spatial and temporal modulation patterns give rise to accurate reconstructions? We examine the decompositions

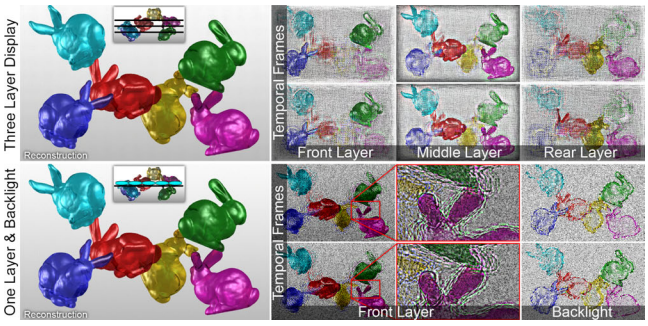

Figure 5: *Interpreting tensor display decompositions. Reconstruction and decomposition results are compared for a three-layer display with uniform backlighting (top) and a single-layer display using a directional backlight (bottom). The structures of the multilayer, multiframe decompositions are discussed in Section 3.4.*

for two architectures: a three-layer display with uniform backlighting and a single-layer display with directional backlighting.

Multilayer decompositions are shown at the top of Figure 5. We observe that objects close to the display appear sectioned across layers. The green bunny maps primarily to the front layer, with residual details assigned to other layers. Similar sectioning behaviors have been observed with multilayer-only decompositions, including those of Gotoda [2010] and Wetzstein et al. [2011]. Unlike these works, our joint multilayer, multiframe decompositions produce additional time-varying, high-frequency patterns that appear across all layers and resemble content-adaptive parallax barriers [Lanman et al. 2010].

Decompositions for a single-layer display with directional backlighting are shown at the bottom of Figure 5. We observe that the front layer contains the view-independent portions of the scene, with flowing, slit-like patterns appearing around regions with view-dependent features. The directional backlight is primarily comprised of view-dependent features, such as objects extending from the physical display enclosure (e.g., the green bunny).

Tensor display decompositions exhibit predictable structures, whose arrangement arise from the specific display configuration. A natural direction for future work is to more closely assess these structures for promising architectures, such as the single layer with directional backlighting, in the hope that heuristically-defined methods may achieve similar fidelity with reduced computation.

4 Analysis

This section analyzes the performance of tensor displays, focusing on the quantitative benefits of additional layers, additional frames, and directional backlighting. First, we derive the upper bound on the depth of field for any tensor display. This allows comparison of alternative display architectures. The upper bound also provides antialiasing prefilters for each design. Second, we assess the interdependence of display design and decomposition algorithm parameters, documenting their influence on reconstructed image fidelity. Extended derivations of the depth of field expressions are provided in Supplementary Appendix C.

4.1 Depth of Field

The performance of an automultiscopic display can be quantified by its depth of field: an expression for the maximum spatial frequency $\omega_{\xi_{\max}}$ that can be depicted in a plane oriented parallel to the screen and separated by a distance d_o. As described by Zwicker et al. [2006], this expression is derived using a frequency-domain analysis of the emitted light field $\tilde{l}(x, v)$. Taking the 2D Fourier transform of Equation 18 yields the following expression for the emitted light field spectrum $\hat{l}(\omega_x, \omega_v)$:

$$\hat{l}(\omega_x,\omega_v)=\frac{1}{M}\sum_{m=1}^{M}\hat{b}_m(\omega_x,\omega_v)*\left[\mathop{\circledast}_{n=1}^{N}\hat{f}_m^{(n)}(\omega_x)\delta(\omega_v-(d_n/d_r)\omega_x)\right], \quad (21)$$

where ω_x and ω_v are the spatial and angular frequencies, $*$ denotes convolution, and the repeated convolution operator is defined as

$$\mathop{\circledast}_{n=1}^{N}\hat{f}_m^{(n)}(\omega_x,\omega_v) \equiv \hat{f}_m^{(1)}(\omega_x,\omega_v) * \cdots * \hat{f}_m^{(N)}(\omega_x,\omega_v). \quad (22)$$

For uniform backlighting, the backlight spectrum $\hat{b}_m(\omega_x, \omega_v) = \delta(\omega_x, \omega_v)$, the Dirac delta function, reducing Equation 21 to the expression derived for multilayer displays by Wetzstein et al. [2011].

The spectral support of a tensor display is the region of non-zero values in the emitted light field spectrum, for all possible layer masks and backlight illumination patterns. Following Chai et al. [2000], the spectral support for the light field reflected by a diffuse surface is the line $\omega_v = (d_o/d_r)\omega_x$. Intersecting this line with the spectral support for a given display provides a geometric construction for the upper bound on the depth of field. For example, the emitted light field spectrum for a parallax barrier or integral imaging display is non-zero only for $|\omega_x| \leq 1/(2\Delta x)$ and $|\omega_v| \leq 1/(2\Delta v)$ (e.g., the red boxes shown in Figure 4), where Δx and Δv are the spatial and angular sampling rates, respectively. In practice, the spatial sampling rate Δx is the spacing between barrier slits/pinholes or lenslets. The geometric construction yields the following expression for the depth of field:

$$\omega_{\xi_{\max}}(d_o) = \begin{cases} \frac{1}{2\Delta x} & \text{for } |d_o| \leq d_r\left(\frac{\Delta x}{\Delta v}\right), \\ \frac{d_r}{2|d_o|\Delta v} & \text{otherwise,} \end{cases} \quad (23)$$

where $\Delta v = (2d_r/A)\tan(\alpha/2)$ with A views and field of view α.

The geometric construction provides an upper bound on the depth of field for any tensor display architecture. Consider a two-layer display with uniform backlighting, with the layers separated by a distance Δd and $\omega_0 = 1/(2p)$ denoting the maximum spatial frequency for each layer with pixel pitch p. Equation 21 defines the light field spectrum, where $d_1 = -\Delta d/2$ and $d_2 = \Delta d/2$. As shown in Figure 4, a diamond-shaped region bounds the spectral support for any two-layer display. The spatial cutoff frequency $\omega_{\xi_{\max}}$ is again found by intersecting the line $\omega_v = (d_o/d_r)\omega_x$ with the boundary of the spectral support, yielding the following upper bound on the depth of field for any two-layer display.

$$\omega_{\xi_{\max}}(d_o) = \left(\frac{2\Delta d}{\Delta d + 2|d_0|}\right)\omega_0 \quad (24)$$

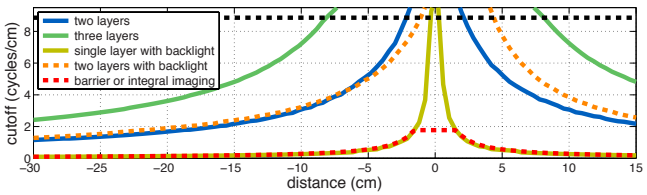

Figure 6: *Comparison of upper bounds on depth of field for parallax barriers and integral imaging (red), two-layer (blue) and three-layer (green) displays with uniform backlighting, and single-layer (yellow) and two-layer (orange) displays with directional backlighting. The dashed black line denotes the spatial cutoff frequency for each layer. Display parameters correspond to the prototypes described in Section 5.2.*

In Section 5 we compare two tensor display architectures: a three-layer display with uniform backlighting vs. a single-layer display with directional backlighting. Using the previously described geometric construction, the depth of field for a three-layer display with uniform backlighting and equally-spaced layers is given by

$$\omega_{\xi_{\max}}(d_o) = \begin{cases} \left(\frac{3\Delta d}{\Delta d + |d_0|}\right)\omega_0 & \text{for } |d_o| \leq 2\Delta d, \\ \left(\frac{2\Delta d}{|d_o|}\right)\omega_0 & \text{otherwise,} \end{cases} \tag{25}$$

where Equation 21 is again applied to find the spectral support, with $d_1 = -\Delta d$, $d_2 = 0$, and $d_3 = \Delta d$. As shown in the fourth row of Figure 4, the spectral support for a three-layer display exceeds that of a similar parallax barrier or integral imaging display, leading to the increased depth of field observed in Figure 6.

As described in Section 3.3, incorporating directional backlighting can significantly expand the field of view. The depth of field for a single-layer display using directional backlighting is obtained by a similar geometric construction. We assume the directional backlight implements a low-resolution light field display, such that $\hat{b}_m(\omega_x, \omega_v)$ has non-zero support for $|\omega_x| \leq 1/(2\Delta x)$ and $|\omega_v| \leq 1/(2\Delta v)$. This yields the following depth of field expression:

$$\omega_{\xi_{\max}}(d_o) = \begin{cases} \frac{1}{2\Delta x} + \omega_0 & \text{for } |d_o| \leq d_r\left(\frac{\Delta x}{\Delta v + 2\Delta x \Delta v \omega_0}\right), \\ \frac{d_r}{2|d_o|\Delta v} & \text{otherwise,} \end{cases} \tag{26}$$

where ω_0 again denotes the spatial cutoff frequency for the layer. As shown in Figure 6, the addition of a single light-attenuating layer significantly increases the spatial resolution for a conventional parallax barrier or integral imaging display, particularly near the display surface. However, far from the display, the depth of field is identical to these conventional automultiscopic displays.

Our analysis indicates a promising application for tensor displays: increased depth of field can be achieved by covering any low-resolution light field display with time-multiplexed, light-attenuating layers. In this analysis, we assume continuously-varying layer transmittances; a promising research direction is to characterize the upper bound with discrete pixels. However, with our analysis, we observe that static and time-multiplexed tensor displays have identical spectral supports (i.e., averaging over an M-frame sequence does not alter the support via Equation 21). Yet, as depicted in the second and third rows of Figure 4, time multiplexing significantly reduces artifacts. We attribute this to the additional degrees of freedom allowed with time multiplexing. While the upper bound may be identical, in practice it cannot be achieved with static methods, motivating tensor displays for joint multilayer, multiframe decompositions capable of approaching the upper bound.

4.2 Design Trade-Offs

One of the main benefits of tensor displays is to open a design trade space not accessible to prior automultiscopic displays. Existing multilayer-only or multiframe-only decompositions require many layers or prohibitively high frame rates, limiting their practicality using current LCD technology. However, with joint multilayer, multiframe decompositions, display designers can explore the interdependence of the number of layers, the number of frames, and the image brightness. In this section we demonstrate that tensor displays using relatively few layers and frames achieve higher-fidelity reconstructions than prior methods, in a manner supported by current LCD technology. We also show that tensor displays achieve wide fields of view, as required for multiviewer scenarios.

We employ PSNR to quantify the difference between reconstructed views and the target light field. We expect perceptual error metrics

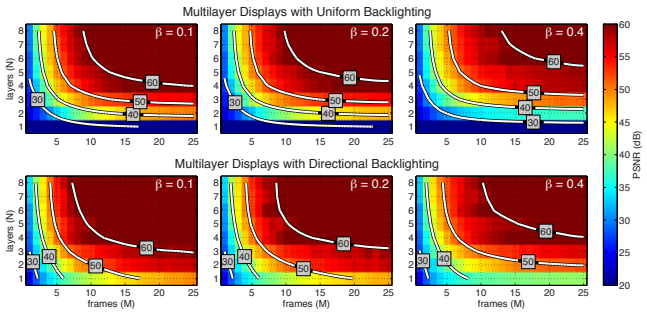

Figure 7: *Design trade-offs for tensor displays. Peak signal-to-noise ratio (PSNR), as a function of the number of frames M, number of layers N, and brightness β, evaluated for the teapot scene and the display parameters in Section 4.2. (Top) Results with uniform backlighting. (Bottom) Results with directional backlighting.*

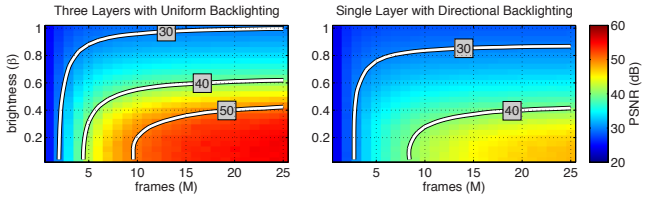

Figure 8: *Optimizing the tensor display prototypes. PNSR is evaluated, as a function of the number of frames M and brightness β, for the teapot scene and the display parameters in Section 4.2.*

to better predict subjective assessments; unfortunately, multiview perceptual metrics remain an open research topic. We consider a fixed set of uniformly-spaced viewpoints during optimization. As shown in the supplementary video, providing closely-spaced target views sufficiently constrains the decompositions so minimal artifacts are perceived at intermediate viewpoints.

4.2.1 Interdependence of Layers, Frames, and Brightness

Display designers seek to maximize image fidelity (e.g., PSNR) as a function of device complexity (i.e., the number of layers and frames). Consider optimizing multilayer designs with uniform backlighting. The design trade space is shown in Figure 7. The teapot scene is decomposed for a field of view $\alpha = 20^\circ \times 20^\circ$, spatial resolution of 160×100 pixels, 3×3 views, and layer separation $\Delta d = 4.0$ cm. Note that these display parameters differ from those for Figure 4, where the layers are separated by only 8 mm. These simulations verify a key benefit of tensor displays: increasing the number of frames allows the number of layers to be decreased (for a given PSNR). These simulations also reveal the dependence on the brightness scale $\beta \in [0, 1]$ applied to the target light field; specifically, we modify Equation 15 to yield the following objective function supporting a trade-off between image brightness and fidelity.

$$\underset{\{\mathbf{F}^{(n)}\}}{\arg\min} \left\| \beta \mathcal{L} - \mathcal{W} \circledast \tilde{\mathcal{T}} \right\|^2, \text{ for } 0 \leq \mathbf{F}^{(n)} \leq 1 \tag{27}$$

We observe that decreasing brightness generally yields higher-fidelity reconstructions for the same number of layers and frames.

The trade space for multilayer displays with brightness $\beta = 0.2$ is shown in the center of the top row of Figure 7. We observe that static decompositions (i.e., $M = 1$) cannot exceed 30 dB, even with as many as eight layers. To achieve 40 dB with eight layers, two frames are required. However, note the trade-off between layer complexity and refresh rate along the 40 dB curve. Using six frames, only three layers are required, with more frames providing

marginal benefits. Thus, with tensor displays, designers can exploit high-speed displays to reduce device complexity, minimizing the number of layers to achieve a certain image fidelity.

Adding a directional backlight alters the design trade space, as shown at the bottom of Figure 7 for a directional backlight with 47×29 lenslets. We observe that two frames are still required to reach 40 dB using eight layers. However, only a single layer is now required using eight frames. For this example, the directional backlight effectively reduces the number of required layers by one. This underscores the practical benefits of the tensor display framework, which is the first to combine the benefits of multilayer decompositions, time-multiplexing, and directional backlighting.

Tensor displays encompass a broad set of architectures. In Section 5.2, we configure the prototype to demonstrate two designs: three layers with uniform backlighting and a single layer with directional backlighting. The design trade spaces are shown in Figure 8. For three layers, four frames are required to achieve 40 dB. With additional frames, brightness can be significantly increased (up to $\beta \approx 0.6$). To our knowledge, this is the first automultiscopic display demonstrating such trade-offs between display refresh rate and brightness, providing additional motivation for developing high-speed spatial light modulators. Similarly, with directional backlighting, a minimum of eight frames are required to achieve 40 dB. We confirm predicted PSNR trends in Section 5.2.

4.2.2 Increasing Field of View

Conventional automultiscopic displays, including parallax barriers and integral imaging, exhibit a set of periodically-repeating viewing zones. In contrast, recent computationally-optimized multilayer and multiframe displays generally exhibit a set of non-repeating viewing zones; while yielding extended depths of field, greater resolution, and increased brightness, viewers are typically limited to a field of view of $\alpha \lesssim 20^\circ$. As shown in Figure 9, tensor displays support wider fields of view, while retaining the benefits of computational optimization. A field of view of $\alpha = 50^\circ \times 20^\circ$ is achieved, for a light field with 9×3 views, using either five layers and uniform backlighting or a single layer and directional backlighting. We observe that prior multilayer-only and multiframe-only decompositions lack sufficient degrees of freedom to achieve high-PSNR reconstructions for this scenario. Differences between predicted and observed depths of field and PSNR, as shown in Figures 6–8 and Figure 9, are due to differing fields of view in these experiments.

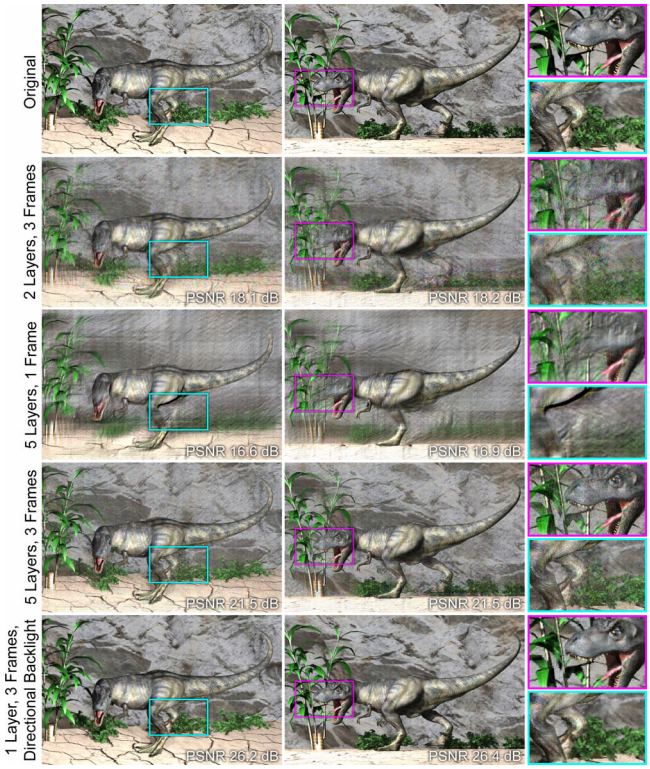

Figure 9: *Tensor displays achieve wider fields of view than prior multilayer displays. This example assumes a field of view of $\alpha = 50^\circ \times 20^\circ$ and three frames. We observe that tensor displays, using five layers with uniform backlighting (fourth row) or a single layer and directional backlighting (fifth row), minimize artifacts compared to multiframe-only (second row) and multilayer-only (third row) decompositions.*

5 Implementation and Assessment

This section describes the tensor display prototype and assesses its performance. We first review the prototype hardware and software implementation. Afterwards, we evaluate the performance for two prototype configurations: a three-layer LCD with uniform backlighting and a single LCD with directional backlighting.

5.1 Implementation

5.1.1 Hardware

We built a reconfigurable tensor display prototype capable of implementing two-layer and three-layer architectures with uniform or directional backlighting (see Figure 2). The layers are constructed using three modified Viewsonic VX2268wm 120 Hz LCD panels. The front and rear polarizing films are removed from the front two LCDs, and the stack is interleaved with alternating crossed linear polarizers. Aluminum brackets added to the rear panel allow lenslet arrays to be affixed for operation as a directional backlight. A rectangular lenslet array is approximated using two crossed lenticular sheets, purchased from Micro Lens Technology, Inc. The corrugated surfaces of the sheets are held in direct contact, minimizing astigmatic aberrations [Bader et al. 1997]. The directional backlight supports varying spatio-angular resolution trade-offs using 10, 15, and 20 lenses per inch (LPI) lenticular sheets. We observe that the sheets are birefringent due to stresses introduced during manufacturing. In directional backlighting modes, an additional polarizing film is placed after the lenslet arrays, restoring the linear polarization state before rays impinge on the next LCD in the stack.

We implemented offline and online solvers based on Equation 16. Computation is divided between CPUs, for the offline solver, and GPUs for the online solver. The offline solver is run on an Intel Core i5 workstation with 10 GB of RAM. The online solver is run on an Intel Core i7 workstation with 6 GB of RAM and an external Nvidia QuadroPlex 7000 graphics unit containing two Quadro GPUs and a G-Sync card. This provides four frame-synchronous DVI outputs capable of driving the LCDs at 120 Hz.

5.1.2 Software

Target light fields are rendered using POV-Ray or, for interactive applications, using OpenGL. Rendered light fields have a spatial resolution of 840×525 pixels (i.e., half the resolution of LCDs used in the prototype) and an angular resolution of 5×5 views.

We implemented nonnegative tensor factorization (NTF) using the multiplicative update rules from Section 3.2. An offline, Matlab-based solver is used for simulations. Decomposing a target light field into a six-frame sequence for three layers takes approximately

30 minutes using 50 updates. Color channels are processed independently. An online, GPU-accelerated solver is implemented in OpenGL and Cg. Our update rules can be cast as additive combinations of the logarithms of the layer transmittances. Using this representation, the update rules are mapped to standard operations of the graphics pipeline, including projective texture mapping, accumulation buffers, floating point framebuffers, and perspective rendering. These operations are not only computationally efficient, but also memory-efficient, as only the non-zero tensor elements need to be stored and processed. For interactive applications we exploit temporal coherence between decompositions, seeding each frame with the prior result, as shown in the supplementary video. In Supplementary Appendix E we provide pseudocode and additional details for the GPU-accelerated solver.

Separate threads are used to decouple the decomposition from the display routines. Decompositions are evaluated in an asynchronous thread, updating layer patterns as they become available. This ensures that all display layers can be continuously refreshed at 120 Hz, without waiting for updated decompositions. Using the prototype hardware, we achieve up to 10 multiplicative updates per second for as many as 12 frames. Light fields with reduced spatial or angular resolution can be decomposed and displayed at interactive refresh rates, as shown in the supplementary video. All experiments using the prototype display employ the GPU-accelerated solver.

5.2 Assessment

5.2.1 Three-Layer LCD with Uniform Backlighting

As shown in Figure 2, the prototype was configured as a three-layer LCD with uniform backlighting. Acrylic spacers separated each panel by $\Delta d = 4.0$ cm. The target light field was rendered with a field of view of $\alpha = 20^\circ \times 20^\circ$ and brightness $\beta = 0.2$ (see Section 4.2). Photographs of the central view, seen directly in front of the prototype, are shown along the center column of Figure 10. Each light field was decomposed using twelve frames. The camera exposure was set to 100 ms, simulating a 720 Hz display for a human observer (i.e., for a 60 Hz flicker fusion threshold). We observe that fine details are preserved (e.g., the fish scales and specular highlights on the teapot) and occlusion cues are correctly rendered (e.g., between the bunnies). See the supplementary video for demonstrations of smooth horizontal and vertical motion parallax.

Experiments with the prototype provide insights into practical engineering issues. Foremost, we found that accurate mechanical alignment is crucial. As shown in Figure 5, decomposed layers exhibit high-frequency patterns that must be properly aligned. Accurate alignment was ensured by displaying perspective images of a crosshair array on each layer. A camera was placed at the desired viewer position (e.g., directly in front of the display at a distance of 2 m) and the patterns were shifted until alignment was obtained. We also found that radiometric calibration is necessary, including measuring the black levels and gamma values. The former are incorporated as constraints in the update rules, while the latter are addressed by applying gamma correction at runtime. We attribute remaining variations in color and intensity to differences in the LCD color gamut, color filter cross-talk, moiré due to stacking multiple layers, and angular color variation common to high-speed LCDs.

5.2.2 Single LCD with Directional Backlighting

As shown in Figure 2, the prototype was also configured as a single LCD with a directional backlight. The backlight was fashioned using crossed 10 LPI lenticular sheets, yielding a field of view of $\alpha = 48^\circ \times 48^\circ$ and backlight resolution of 187×117 lenslets. The front LCD was separated by $\Delta d = 8.5$ mm from the middle of the lenticular sheets. Remaining system parameters were identical to the three-layer prototype. Photographs of the central view are shown along the right column of Figure 10. We observe the crossed lenticular sheets produce strong absorption along lens boundaries. In a commercial implementation, lenslet arrays could be manufactured with minimal absorption. Alternatively, edge-lit directional backlighting could eliminate this artifact (see Section 2). As demonstrated in Figure 11, adding an LCD in front of a low-resolution directional backlight increases the spatial resolution for virtual objects appearing on the display surface (e.g., the logo) and for objects extending in depth (e.g., the fish tail on the right). While resolution can be enhanced at the display surface without time multiplexing, enhancement for extended scenes can only be achieved with time multiplexing, as facilitated by our tensor framework.

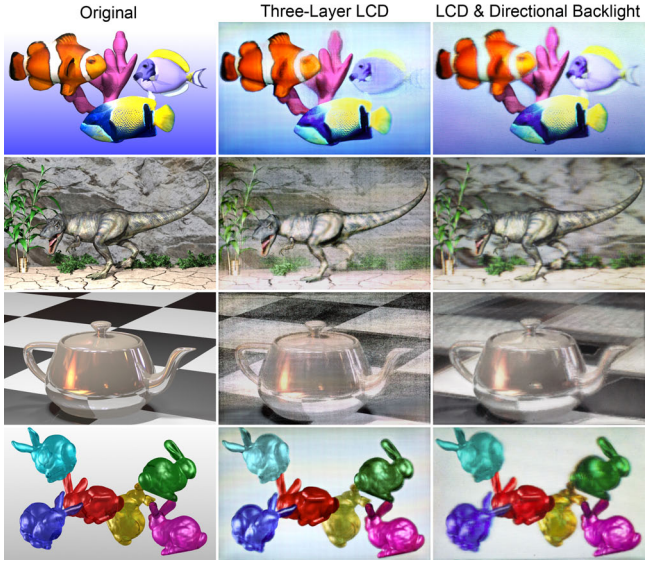

Figure 10: *Experimental results using the tensor display prototype. Central views of four scenes are shown for the input light fields (left column), photographs of the three-layer LCD (center column), and the single LCD with directional backlighting (right column).*

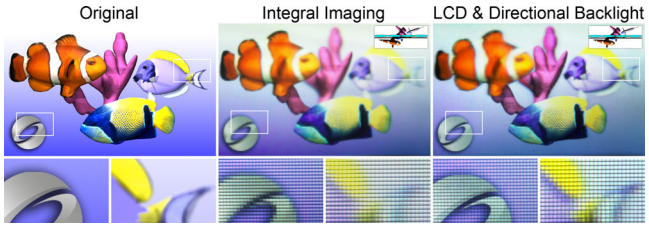

Figure 11: *Enhancing integral imaging with tensor displays. While integral imaging, here implemented with a lenslet array affixed to an LCD, achieves a convincing 3D effect, spatial resolution is significantly reduced (center). Adding an LCD in front of the low-resolution backlight and exploiting temporal multiplexing using our tensor framework increases the spatial resolution, not only on the physical layers, but also outside the hardware enclosure (right).*

6 Discussion

6.1 Benefits and Limitations

We extend our discussion from Section 1.2 in light of experiences with the prototype. Tensor displays combine, for the first time, the advantages of multiple layers, high-speed temporal modulation, and directional backlighting within a compressive optimization frame-

work. Our development of tensor displays is timely, as each of these technologies is an emerging trend in display design. Our tensor framework opens a large design trade space that was inaccessible using prior automultiscopic displays. With our framework, designers can maximize image fidelity, brightness, and field of view, depending on the number of layers and maximum refresh rate allowed by the design constraints and display technology, respectively.

We identify single layer LCDs with directional backlighting as a particularly promising design. As shown in Figure 9, such displays support a wide field of view with relatively few frames (i.e., as few as three in our simulations and experiments). Thus, provided with 180 Hz LCDs, this design achieves the stated goal of a thin form factor, wide field of view, bright automultiscopic display with an effective refresh rate of 60 Hz. We also identify joint multilayer, multiframe decompositions as an effective tool for optimizing multilayer displays with uniform backlighting. Such displays have the added benefit of a tunable field of view; unlike directional backlighting, which in our implementation has a fixed field of view, this design allows viewing zones to adapt to the location of viewers.

The prototype reveals several limitations inherent to layered architectures, including moiré, color-channel crosstalk, interreflections, misalignment, and dimming due to layered color filter arrays. Many of these issues can be resolved with additional optical engineering. Moiré, interreflections, and misalignment can be mitigated using holographic diffusers, antireflective coatings, and rigid enclosures, respectively. A direct solution to crosstalk is to alter the transmission profiles of the color filters; however, this approach will further decrease brightness. Instead, field sequential color could be applied (i.e., using a backlight that sequentially strobes each color), albeit by placing additional demands on the refresh rate. A promising direction for future work is to consider whether color filters are necessary for each layer. Decompositions could be performed assuming monochromatic panels interspersed with a few color filters.

6.2 Future Work

The weight tensor applied in Equation 16 allows decompositions to be tuned to the positions of viewers. In our implementation we consider a field of view centered about the display's surface normal; however, if head or eye tracking was available, the weight matrix could be altered to only project automultiscopic imagery aligned to each viewer. Between viewers, the emitted light field would be unconstrained, possibly allowing for higher-fidelity, brighter imagery.

Our image formation model, given by Equation 18, could be generalized. We consider time-multiplexed, light-attenuating layers over a uniform light source, potentially with one lenslet array between the first and second layers. In general, layers could be composed of both light-attenuating and light-emitting materials. In addition, refractive elements could be placed at any point (e.g., a Fresnel lens in front of the display to extend the depth of field [Gotoda 2011]).

We address limitations of automultiscopic displays with high spatial resolution and low angular resolution. For accommodation and convergence cues, additional views are required. Our framework supports more views, if provided with sufficiently high-speed displays. For example, digital microshutters (DMS) [Steyn et al. 2010] are one promising option, capable of achieving 1,440 Hz refresh rates, allowing 24 frames with an effective refresh rate of 60 Hz.

We apply least-squares optimization; however, perceptual error metrics will likely allow further reductions in complexity (i.e., fewer layers and frames). Such error metrics will likely involve nonlinear objectives and require refined optimization schemes.

We advocate for multilinear optimization as a practical tool for compressive light field synthesis using tensor displays. We employ a frequency-domain analysis to assess the capabilities of tensor displays. A promising direction for future work is to apply our tensor framework not only for optimization, but also for formal analysis.

7 Conclusion

Automultiscopic displays have not found widespread consumer adoption. While compelling multiview content is first necessary, long-standing limitations must be conclusively resolved. Any viable solution must preserve the thin form factors, low power consumption, and high resolution of modern displays. The construction must rely on near-term, mass-market technology. We identify three key trends: multilayer panels, high refresh rates, and directional backlighting. Tensor displays provide the first framework combining the advantages of these technologies. This framework applies the principles of emerging *computational displays*, wherein the display architecture and encoding algorithm are jointly optimized to maximize optical and computational efficiency. By exploring the space of tensor displays, we arrive at a simple, but compelling, architecture: a single LCD with a directional backlight. This design achieves a wide field of view and large depth of field with a thin form factor using efficient multiplicative updates. By integrating emerging technologies, we hope tensor displays provide a foundation for evaluating competing architectures and inspire others to optimize all degrees of freedom afforded by any given design.

Acknowledgements

We thank the reviewers for their insightful feedback and recognize the support of the Camera Culture group. We also thank the MIT Media Lab sponsors and NVIDIA Research. Satoshi Miyania and Nam Hye Yeon provided the underwater models. Gordon Wetzstein was supported by the DARPA SCENICC program. Douglas Lanman was supported by NSF Grant IIS-1116452 and by the DARPA MOSAIC program. Ramesh Raskar was supported by an Alfred P. Sloan Research Fellowship and a DARPA Young Faculty Award.

References

AKELEY, K., WATT, S. J., GIRSHICK, A. R., AND BANKS, M. S. 2004. A stereo display prototype with multiple focal distances. *ACM Trans. Graph. (SIGGRAPH) 23*, 804–813.

BADER, G., OTT, P., LUEDER, E., AND SCHMID, V. 1997. Hybrid shape recognition system with microlens array processor and direct optical input. In *SPIE Optical Pattern Recognition VIII*, vol. 3073, 277–287.

BILGILI, A., OZTURK, A., AND KURT, M. 2011. A general BRDF representation based on tensor decomposition. *Computer Graphics Forum 30*, 8, 2427–2439.

BLONDEL, V., HO, N.-D., AND VAN DOOREN, P. 2008. Weighted nonnegative matrix factorization and face feature extraction. In *Image and Vision Computing*, 1–17.

BROTT, R., AND SCHULTZ, J. 2010. Directional backlight lightguide considerations for full resolution autostereoscopic 3D displays. *SID Digest*, 218–221.

CHAI, J.-X., TONG, X., CHAN, S.-C., AND SHUM, H.-Y. 2000. Plenoptic sampling. In *ACM SIGGRAPH*, 307–318.

CHIEN, K.-W., AND SHIEH, H.-P. D. 2006. Time-multiplexed three-dimensional displays based on directional backlights with fast-switching liquid-crystal displays. *Applied Optics 45*, 13, 3106–3110.

CHU, Y. M., CHIEN, K. W., SHIEH, H. P. D., CHANG, J. M., A. HU, Y. C. S., AND YANG, V. 2005. 3D mobile display based on dual-directional light guides with a fast-switching liquid-crystal panel. *J. Soc. Inf. Display 13*, 10, 875–879.

CICHOCKI, A., ZDUNEK, R., PHAN, A. H., AND ICHI AMARI, S. 2009. *Nonnegative Matrix and Tensor Factorizations*. Wiley.

COSSAIRT, O. S., NAPOLI, J., HILL, S. L., DORVAL, R. K., AND FAVALORA, G. E. 2007. Occlusion-capable multiview volumetric three-dimensional display. *Applied Optics 46*, 8, 1244–1250.

DURAND, F., HOLZSCHUCH, N., SOLER, C., CHAN, E., AND SILLION, F. X. 2005. A frequency analysis of light transport. *ACM Trans. Graph. (SIGGRAPH) 24*, 3, 1115–1126.

FAVALORA, G. E. 2005. Volumetric 3D displays and application infrastructure. *IEEE Computer 38*, 37–44.

GOTODA, H. 2010. A multilayer liquid crystal display for autostereoscopic 3D viewing. In *SPIE Stereoscopic Displays and Applications XXI*, vol. 7524, 1–8.

GOTODA, H. 2011. Reduction of image blurring in an autostereoscopic multilayer liquid crystal display. In *SPIE Stereoscopic Displays and Applications XXII*, vol. 7863, 1–7.

HOLROYD, M., BARAN, I., LAWRENCE, J., AND MATUSIK, W. 2011. Computing and fabricating multilayer models. *ACM Trans. Graph. (SIGGRAPH Asia) 30*, 187:1–187:8.

IVES, F. E., 1903. Parallax stereogram and process of making same. U.S. Patent 725,567.

JACOBS, A., MATHER, J., WINLOW, R., MONTGOMERY, D., JONES, G., WILLIS, M., TILLIN, M., HILL, L., KHAZOVA, M., STEVENSON, H., AND BOURHILL, G. 2003. 2D/3D switchable displays. *Sharp Technical Journal*, 4, 1–5.

JONES, A., MCDOWALL, I., YAMADA, H., BOLAS, M., AND DEBEVEC, P. 2007. Rendering for an interactive 360° light field display. *ACM Trans. Graph. (SIGGRAPH) 26*, 40:1–40:10.

KIM, Y., KIM, J., KANG, J.-M., JUNG, J.-H., CHOI, H., AND LEE, B. 2007. 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.

KOLDA, T. G., AND BADER, B. W. 2009. Tensor decompositions and applications. *SIAM Review 51*, 3, 455–500.

KWON, H., AND CHOI, H.-J. 2012. A time-sequential multiview autostereoscopic display without resolution loss using a multi-directional backlight unit and an LCD panel. In *SPIE Stereoscopic Displays and Applications XXIII*, vol. 8288, 1–6.

LANMAN, D., HIRSCH, M., KIM, Y., AND RASKAR, R. 2010. Content-adaptive parallax barriers: Optimizing dual-layer 3D displays using low-rank light field factorization. *ACM Trans. Graph. (SIGGRAPH Asia) 29*, 163:1–163:10.

LANMAN, D., WETZSTEIN, G., HIRSCH, M., HEIDRICH, W., AND RASKAR, R. 2011. Polarization fields: Dynamic light field display using multi-layer LCDs. *ACM Trans. Graph. (SIGGRAPH Asia) 3*, 1–9.

LAWRENCE, J., BEN-ARTZI, A., DECORO, C., MATUSIK, W., PFISTER, H., RAMAMOORTHI, R., AND RUSINKIEWICZ, S. 2006. Inverse shade trees for non-parametric material representation and editing. *ACM Trans. Graph. (SIGGRAPH) 25*, 3.

LIPPMANN, G. 1908. Épreuves réversibles donnant la sensation du relief. *Journal of Physics 7*, 4, 821–825.

LUMIÈRE, L. 1920. Représentation photographique d'un solide dans l'espace. photo-stéréo-synthèse. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 891–896.

MATHER, J., BARRATT, N., KEAN, D. U., WALTON, E. J., AND BOURHILL, G., 2009. Directional backlight, a multiple view display and a multi-direction display. U.S. Patent Application 11/814,383.

MATUSIK, W., AND PFISTER, H. 2004. 3D TV: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Trans. Graph. (SIGGRAPH) 23*, 814–824.

PEERS, P., VOM BERGE, K., MATUSIK, W., RAMAMOORTHI, R., LAWRENCE, J., RUSINKIEWICZ, S., AND DUTRÉ, P. 2006. A compact factored representation of heterogeneous subsurface scattering. *ACM Trans. Graph. (SIGGRAPH) 25*, 3, 746–753.

PERLIN, K., PAXIA, S., AND KOLLIN, J. S. 2000. An autostereoscopic display. In *ACM SIGGRAPH*, 319–326.

PETERKA, T., KOOIMA, R. L., SANDIN, D. J., JOHNSON, A., LEIGH, J., AND DEFANTI, T. A. 2008. Advances in the Dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system. *IEEE TVCG 14*, 3, 487–499.

PUTILIN, A. N., LUKIANITSA, A. A., AND KANASHIN, K. 2001. Stereodisplay with neural network image processing. In *SPIE Advanced Display Technologies*, vol. 4511, 245–250.

STEYN, J., BROSNIHAN, T., FIJOL, J., GANDHI, J., HAGOOD, N., HALFMAN, M., LEWIS, S., PAYNE, R., AND WU, J. 2010. A MEMS digital microshutter (DMS) for low-power high brightness displays. In *Optical MEMS and Nanophotonics*, 73–74.

STOLLE, H., OLAYA, J.-C., BUSCHBECK, S., SAHM, H., AND SCHWERDTNER, A. 2008. Technical solutions for a full-resolution autostereoscopic 2D/3D display technology. In *Proc. SPIE*, 1–12.

SULLIVAN, A. 2003. A solid-state multi-planar volumetric display. In *SID Digest*, vol. 32, 207–211.

TOYOOKA, K., MIYASHITA, T., AND UCHIDA, T. 2001. The 3D display using field-sequential LCD with light direction controlling backlight. *SID Digest*, 177–180.

TRAVIS, A., LARGE, T., EMERTON, N., AND BATHICHE, S. 2009. Collimated light from a waveguide for a display backlight. *Optics Express 17*, 22, 19714–19719.

TRAVIS, A. R. L. 1990. Autostereoscopic 3-D display. *Applied Optics 29*, 29, 4341–4342.

VASILESCU, M. A. O., AND TERZOPOULOS, D. 2004. TensorTextures: Multilinear image-based rendering. *ACM Trans. Graph. (SIGGRAPH) 23*, 336–342.

WANG, H., WU, Q., SHI, L., YU, Y., AND AHUJA, N. 2005. Out-of-core tensor approximation of multi-dimensional matrices of visual data. *ACM Trans. Graph. (SIGGRAPH) 24*, 527–535.

WETZSTEIN, G., LANMAN, D., HEIDRICH, W., AND RASKAR, R. 2011. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph. (SIGGRAPH) 30*, 1–11.

ZWICKER, M., MATUSIK, W., DURAND, F., AND PFISTER, H. 2006. Antialiasing for automultiscopic 3D displays. In *EGSR*.