Supplementary Material:

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

Gordon Wetzstein Douglas Lanman Matthew Hirsch Ramesh Raskar
MIT Media Lab

In this document we provide additional discussion and results in support of the primary text. Supplementary Appendix A presents an extended introduction to nonnegative tensor factorization (NTF), with additional details on our application of weighted NTF. Supplementary Appendix B includes additional descriptions and illustrations to establish an intuition into how NTF decomposes a given light field for a specific tensor display architecture. The depth of field of tensor displays is analyzed in Supplementary Appendix C. Photographs documenting the prototype construction are provided in Supplementary Appendix D. We conclude with an extended description of our GPU-based NTF solver, including pseudocode, in Supplementary Appendix E.

A Additional Details on Nonnegative Matrix and Tensor Factorization

This section provides a brief introduction to nonnegative matrix and tensor factorization (NMF/NTF) using standard notation [Cichocki et al. 2009; Kolda and Bader 2009]. Our goal is to show, in an intuitive way, how to derive the multiplicative update rules for weighted NTF, as a generalization of weighted NMF using multilinear algebra.

Notation	Interpretation
a	scalar
$\mathbf{a}$	vector
$\mathbf{A}$	matrix
$\mathcal{X}$	tensor
$\mathbf{X}_{(i)}$	matricization (unfolding) of tensor $\mathcal{X}$ along mode i
$\mathcal{X} \times_i \mathbf{A} = \mathbf{A}\mathbf{X}_{(i)}$	tensor-matrix product along mode i
$\mathbf{a} \circ \mathbf{b}$	vector outer product
$\mathbf{A} \circledast \mathbf{B}$	Hadamard matrix product (elementwise product)
$\mathbf{A} \oslash \mathbf{B}$	Hadamard matrix division (elementwise division)
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of two matrices $\mathbf{A}$, $\mathbf{B}$
$\mathbf{A}_{\otimes} = \mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(1)}$	Kronecker product of N matrices $\mathbf{A}^{(N)}, \ldots, \mathbf{A}^{(1)}$
$\mathbf{A}^n_{\otimes} = \mathbf{A}^{(N)} \otimes \cdots \otimes \mathbf{A}^{(n+1)} \otimes \mathbf{A}^{(n-1)} \otimes \cdots \otimes \mathbf{A}^{(1)}$	Kronecker product of N–1 matrices $\mathbf{A}^{(N)}, \ldots, \mathbf{A}^{(1)}$, skipping $\mathbf{A}^{(n)}$
$\mathbf{A} \odot \mathbf{B}$	Khatri-Rao product of two matrices
$\mathbf{A}_{\odot} = \mathbf{A}^{(N)} \odot \cdots \odot \mathbf{A}^{(1)}$	Khatri-Rao product of N matrices $\mathbf{A}^{(N)}, \ldots, \mathbf{A}^{(1)}$
$\mathbf{A}^n_{\odot} = \mathbf{A}^{(N)} \odot \cdots \odot \mathbf{A}^{(n+1)} \odot \mathbf{A}^{(n-1)} \odot \cdots \odot \mathbf{A}^{(1)}$	Khatri-Rao product of N–1 matrices $\mathbf{A}^{(N)}, \ldots, \mathbf{A}^{(1)}$, skipping $\mathbf{A}^{(n)}$

Table 1: *Overview of tensor notation and operators.*

A.1 Nonnegative Matrix Factorization

Nonnegative matrix factorization is a well-established approach to decompose a matrix into a sum of nonnegative rank-one matrices. The decomposition problem is not convex, therefore solutions are not straightforward. We review the problem and possible solutions. For a detailed treatise of this subject, please refer to Cichocki et al. [2009].

The problem can be stated as that of decomposing a matrix $\mathbf{X}$ into a sum of rank-one matrices:

$$\mathbf{X} \approx \sum_{k=1}^{K} \mathbf{a}_k \circ \mathbf{b}_k = \mathbf{A}\mathbf{B}^T,$$

where $\mathbf{X} \in \mathbb{R}^{I \times J} : x_{ij} \geq 0$, $\mathbf{A} \in \mathbb{R}^{I \times K} : a_{ij} \geq 0$, and $\mathbf{B} \in \mathbb{R}^{J \times K} : b_{ij} \geq 0$. A sum of rank-one matrices results in a rank-K approximation of the matrix $\mathbf{X}$.

A least-squares solution to the problem optimizes the following objective function:

$$\begin{array}{ll} \text{minimize} & \left\|\mathbf{X} - \mathbf{A}\mathbf{B}^T\right\|_F^2 \\ \text{subject to} & \mathbf{A}, \mathbf{B} \geq 0, \end{array}$$

where the squared Frobenius norm of a matrix is given as $\|X\|_F^2 = \sum_{ij} x_{ij}^2$.

A.1.1 Multiplicative Update Rules

A detailed overview of different approaches to nonnegative matrix factorization can be found in Berry et al. [2006] and Cichocki et al. [2009]. The most popular way of computing the factorization, as proposed by Lee and Seung [1999], initializes the unknown values with random noise and then iteratively updates the matrices $\mathbf{A}$ and $\mathbf{B}$ independently. The subproblem of fixing one and solving for the other is convex; an alternating least-squares (ALS) update scheme can be performed using gradient descent. The gradient descent approach requires the derivative of the objective function with respect to the variables and a step length. Lee and Seung [1999] demonstrate that, by choosing the step length appropriately, the following multiplicative update rules can be applied.

$$\begin{array}{l} \mathbf{A} \leftarrow \mathbf{A} \circledast (\mathbf{X}\mathbf{B}) \oslash \left(\left(\mathbf{A}\mathbf{B}^T\right)\mathbf{B}\right), \\ \mathbf{B} \leftarrow \mathbf{B} \circledast \left(\mathbf{A}^T\mathbf{X}\right) \oslash \left(\mathbf{A}^T\left(\mathbf{A}\mathbf{B}^T\right)\right) \end{array}$$

Starting from an initial guess that contains only positive values and assuming that the data matrix $\mathbf{X}$ is nonnegative, these update rules are guaranteed to keep $\mathbf{A}$ and $\mathbf{B}$ positive throughout the iterative process. In practice, a small value is added to the divisor so as to avoid division by zero.

A.1.2 Weighted Nonnegative Matrix Factorization

As derived by Blondel et al. [2008], the multiplicative update rules can be modified to include weights for each matrix element x_{ij}, such that

$$\begin{array}{l} \mathbf{A} \leftarrow \mathbf{A} \circledast ((\mathbf{W} \circledast \mathbf{X})\,\mathbf{B}) \oslash \left(\left(\mathbf{W} \circledast \left(\mathbf{A}\mathbf{B}^T\right)\right)\mathbf{B}\right), \\ \mathbf{B} \leftarrow \mathbf{B} \circledast \left(\mathbf{A}^T\left(\mathbf{W} \circledast \mathbf{X}\right)\right) \oslash \left(\mathbf{A}^T\left(\left(\mathbf{A}\mathbf{B}^T\right) \circledast \mathbf{W}\right)\right), \end{array}$$

where $\mathbf{W}$ is a weight matrix of the same dimensions as $\mathbf{X}$.

A.2 Nonnegative Tensor Factorization

Comprehensive reviews on the subject can be found in Cichocki et al. [2009] and Kolda and Bader [2009]. Generally, nonnegative tensor factorization is a multilinear extension of NMF to higher-order tensors.

A.2.1 Nonnegative Tensor Factorization for Third-Order Tensors

The basic objective of NTF is the decomposition of a tensor $\mathcal{X}$ into a sum of rank-one tensors, such that

$$\mathcal{X} \approx \sum_{k=1}^{K} \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!] .$$

To compute the optimal low-rank factorization of a given tensor, the following objective must be minimized:

$$\begin{array}{ll} \text{minimize} & \|\mathcal{X} - [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]\|_F^2 \\ \text{subject to} & \mathbf{A}, \mathbf{B}, \mathbf{C} \geq 0. \end{array}$$

As proposed by Welling and Weber [2001] and Mørup et al. [2006], the multiplicative update rules commonly used in NMF [Lee and Seung 1999] can be extended to third-order tensors as

$$\begin{aligned} \mathbf{A} &\leftarrow \mathbf{A} \circledast \left(\mathbf{X}_{(1)} \left(\mathbf{C} \odot \mathbf{B}\right)\right) \oslash \left(\mathbf{A} \left(\mathbf{C} \odot \mathbf{B}\right)^T \left(\mathbf{C} \odot \mathbf{B}\right)\right) \\ \mathbf{B} &\leftarrow \mathbf{B} \circledast \left(\mathbf{X}_{(2)} \left(\mathbf{C} \odot \mathbf{A}\right)\right) \oslash \left(\mathbf{B} \left(\mathbf{C} \odot \mathbf{A}\right)^T \left(\mathbf{C} \odot \mathbf{A}\right)\right) \\ \mathbf{C} &\leftarrow \mathbf{C} \circledast \left(\mathbf{X}_{(3)} \left(\mathbf{B} \odot \mathbf{A}\right)\right) \oslash \left(\mathbf{C} \left(\mathbf{B} \odot \mathbf{A}\right)^T \left(\mathbf{B} \odot \mathbf{A}\right)\right) . \end{aligned}$$

As with NMF, a small value is added to the divisor so as to avoid division by zero.

A.2.2 Weighted Nonnegative Tensor Factorization for Third-Order Tensors

Acar et al. [2011] analyze the problem of weighted NTF under the aspect of missing data. In their application, weighted NTF is developed to interpolate and extrapolate missing elements from the existing ones using a low-rank factorization. In the case of weighted NTF for multilayer light field displays, zeros-weights correspond to light rays outside the desired field of view as opposed to missing data. Although not directly considered in our optimization, the work of Acar et al. [2011] suggests that weighted NTF provides low-rank interpolation and extrapolations for these light rays, thereby providing graceful image degradation outside the specified field of view. Further investigation of the behaviour of unspecified rays in the target light field tensor are desirable in the future. Disregarding missing elements, as is common practice [Blondel et al. 2008], leads to the following simple and efficient multiplicative update rules for third-order tensors:

$$\begin{aligned} \mathbf{A} &\leftarrow \mathbf{A} \circledast \left(\left(\mathbf{W}_{(1)} \circledast \mathbf{X}_{(1)}\right) \left(\mathbf{C} \odot \mathbf{B}\right)\right) \oslash \left(\left(\mathbf{W}_{(1)} \circledast \left(\mathbf{A} \left(\mathbf{C} \odot \mathbf{B}\right)^T\right)\right) \left(\mathbf{C} \odot \mathbf{B}\right)\right) \\ \mathbf{B} &\leftarrow \mathbf{B} \circledast \left(\left(\mathbf{W}_{(2)} \circledast \mathbf{X}_{(2)}\right) \left(\mathbf{C} \odot \mathbf{A}\right)\right) \oslash \left(\left(\mathbf{W}_{(2)} \circledast \left(\mathbf{B} \left(\mathbf{C} \odot \mathbf{A}\right)^T\right)\right) \left(\mathbf{C} \odot \mathbf{A}\right)\right) \\ \mathbf{C} &\leftarrow \mathbf{C} \circledast \left(\left(\mathbf{W}_{(3)} \circledast \mathbf{X}_{(3)}\right) \left(\mathbf{B} \odot \mathbf{A}\right)\right) \oslash \left(\left(\mathbf{W}_{(3)} \circledast \left(\mathbf{C} \left(\mathbf{B} \odot \mathbf{A}\right)^T\right)\right) \left(\mathbf{B} \odot \mathbf{A}\right)\right) . \end{aligned}$$

Here, $\mathcal{W}$ is a (binary) weight tensor with the same dimensions as $\mathcal{X}$. We emphasize that this formulation is a direct generalization of weighted NMF [Blondel et al. 2008] to higher-order tensors.

A.2.3 Nonnegative Tensor Factorization for N$^{\text{th}}$-Order Tensors

The extension of NTF from third-order tensors to N$^{\text{th}}$-order tensors is straightforward. Tensor $\mathcal{X}$ is approximated by a sum of K nonnegative rank-one tensors, such that

$$\mathcal{X} \approx \sum_{k=1}^{K} \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \cdots \circ \mathbf{a}_k^{(N)} = \mathcal{J} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)} = \left[\!\left[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)} \right]\!\right].$$

The corresponding update rules, following Mørup et al. [2006], are given by the following expression.

$$\mathbf{A}^{(n)} \leftarrow \mathbf{A}^{(n)} \circledast \left(\mathbf{X}_{(n)} \mathbf{A}_{\odot}^{n}\right) \oslash \left(\mathbf{A}^{(n)} \; \mathbf{A}_{\odot}^{n\,T} \mathbf{A}_{\odot}^{n}\right)$$

Including the weight tensor provides the following general update rules for weighted NTF.

$$\mathbf{A}^{(n)} \leftarrow \mathbf{A}^{(n)} \circledast \left(\left(\mathbf{W}_{(n)} \circledast \mathbf{X}_{(n)}\right) \mathbf{A}_{\odot}^{n}\right) \oslash \left(\left(\mathbf{W}_{(n)} \circledast \left(\mathbf{A}^{(n)} \; \mathbf{A}_{\odot}^{n\,T}\right)\right) \mathbf{A}_{\odot}^{n}\right)$$

B Understanding Tensor Displays

Tensor displays comprise a family of display architectures that includes many possible implementations. The characteristic feature shared by all tensor display incarnations is that multiple light-attenuating optical elements are combined in a way such that each ray in a target light field intersects each optical element at most once. Light-attenuating elements are usually arranged in layers which can be composed of any of the following: angularly-invariant spatial light modulators, purely directional modulators, and spatio-angular modulators. A low-resolution light field backlight, for instance, implemented by a lenslet array on top of an LCD, is one type of spatio-angular modulator. In this section we use a series of examples to provide an intuition for how nonnegative matrix factorization and nonnegative tensor factorization decompose a given light field for a specific tensor display implementation. In Section B.1, we illustrate the tensor space spanned by different display types in detail, whereas Section B.2 demonstrates NTF decompositions for a variety of display implementations and compares them to decompositions computed with alternative methods proposed in the literature.

B.1 Light Field Tensors

The tensor space spanned by a tensor display with N optical elements, such as layers, is of dimension N. As observed in Figure S.1, the light field only occupies a low-dimensional manifold within the tensor space. The shape of the manifold depends on a particular tensor display configuration and is shown for a three-layer display as well as for a dual-layer configuration with an additional directional backlight. A weighted nonnegative tensor decomposition, following Section A.2.2, has non-zero values only on the low-dimensional manifold created by the light field in tensor space. These visualizations illustrate the tensor space for different displays in an intuitive manner.

B.2 Light Field Tensor Factorization

The following subsections show nonnegative tensor factorizations for a variety of tensor display implementations.

B.2.1 Single Layer and Purely Directional Backlight

Figure S.2 illustrates an intuitive case: a single, high-resolution layer, for instance an LCD, is combined with a purely angular light source. This could be a large lens directly behind the layer. The angular resolution of the backlight is assumed to be the same as the target light field (Fig. S.2, top row), in this case 3×3. Given the target light field and the physical setup, a nonnegative tensor factorization can then be performed for any desired or feasible number of temporally-multiplexed frames. The naïve solution would be to use nine frames and illuminate a single backlight direction at a time, showing the corresponding light field view on the front layer. We demonstrate in Figure S.2 (rows 6 and 7) that NTF converges to the naïve solution if nine time frames are available. Doing so would, however, require a synchronized LCD and backlight to run at a minimum of 540 Hz, assuming a flicker fusion rate of the human visual system of 60 Hz. A lower frame rate may be required by the available hardware, which does not have an obvious heuristic solution. Nonnegative tensor factorization handles these cases naturally and provides the optimal decompositions, in a least-squared error sense (Fig. S.2, rows 2–5).

B.2.2 Single Layer and Low-Resolution Light Field Backlight

Figure S.3 evaluates the performance of a single light-attenuating layer combined with a low-resolution backlight. The backlight is simulated with four spatial resolutions, all lower than the layer resolution. PSNRs of the reconstructions are given in the insets. As shown, a low-resolution directional backlight combined with a high-resolution layer, such as an LCD, can achieve high image quality by temporally multiplexing only a few frames.

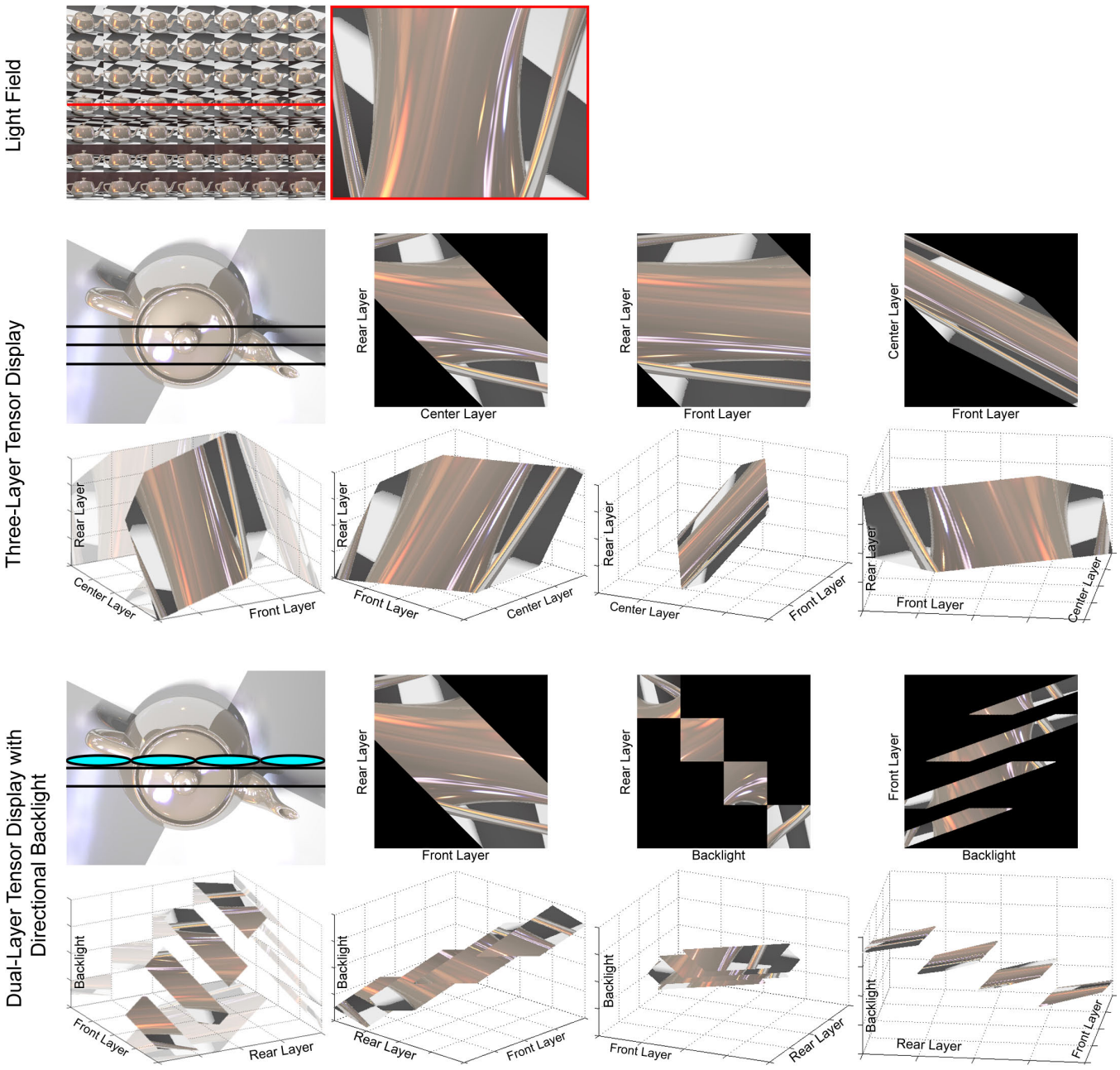

Figure S.1: *Tensor display visualization for a three-layer implementation and a dual-layer display with a directional backlight. For illustrative purposes, the light field is a 1D slice (upper right) of a full 4D light field (upper left). While the pairwise layer parameterizations in the three-layer case span individual matrices (row 2), the tensor itself spans a higher-dimensional space with the light field embedded in a two-dimensional manifold within that tensor space (row 3). A similar effect can be observed for the dual-layer and backlight configuration (row 5); the lower-dimensional manifold within the tensor space differs from the three-layer case. Elementwise parameterizations are shown in row 4. A layer held against the directional backlight creates blockwise-independent matrix components (row 4, center right), whereas a gap between the two optical elements creates a shear in the light field (row 4, right).*

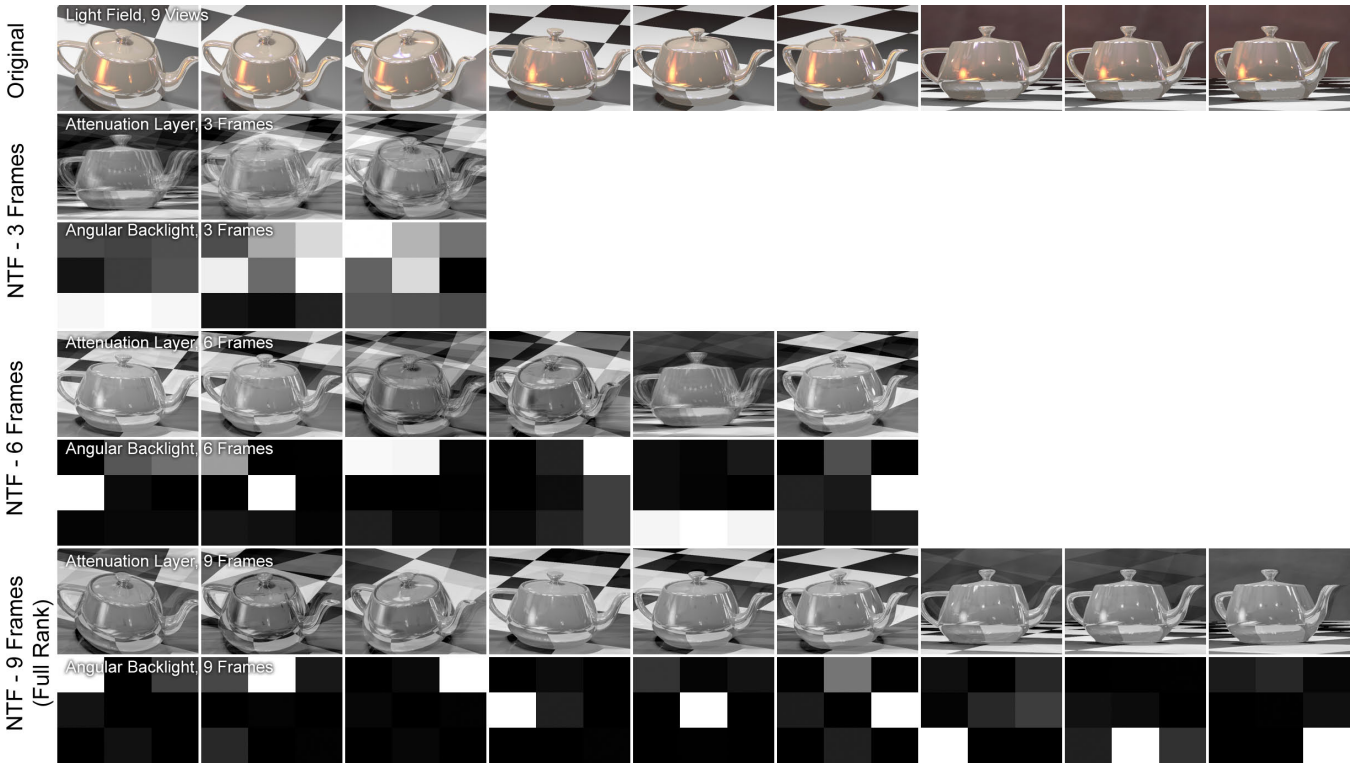

Figure S.2: *Original light field with* 3×3 *views (top row) and decompositions for a high-resolution layer directly on top of a purely directional backlight. This kind of backlight corresponds to a single large lens directly behind an LCD with another spatial light modulator (SLM) mounted at the focal length of the lens; the secondary SLM has a resolution of* 3×3*, corresponding to the angular resolution of the light field. Decompositions for both high-resolution LCD and low resolution angular backlight are shown for three time-multiplexed frames (rows 2 and 3), six frames (rows 4 and 5), and nine frames (rows 6 and 7). The brightness for all decompositions is scaled by the inverse of the number of frames. As seen in the lower two rows, NTF converges toward the obvious solution: turning on each direction of the backlight sequentially over time with the LCD showing the corresponding view of the light field. NTF, however, generalizes the factorization problem to an arbitrary number of frames and different brightness tradeoffs. For the case of rank-deficient decompositions (rows 2–5), the views and corresponding backlight directions are automatically grouped into the set of structurally similar views that result in the optimal image quality.*

B.2.3 Dual Layer Factorization

Dual-layer automultiscopic display architectures have been driven using nonnegative matrix factorization (NMF) in prior work [Lanman et al. 2010]. As demonstrated in Supplementary Appendix A, nonnegative tensor factorization (NTF) mathematically reduces to NMF for the special case of dual-layer displays, because the spanned tensor is just a matrix. Therefore, NTF produces identical layer decompositions as NMF for this special display configuration. The NTF framework, however, generalizes to multilayer architectures as well as combined multilayer and directional backlight configurations.

B.2.4 Multilayer Tensor Factorization

With Figure S.4, we want to build an intuition for NTF-based multilayer decompositions. As illustrated in these examples, the low spatial frequencies in the decomposed layers are comparable to the tomographic solution. This acts similarly to a 3D geometry slicing operator for Lambertian objects on the layers. Multiframe decompositions computed with our tensor framework additionally contain high-frequency variations in image regions exhibiting mo-

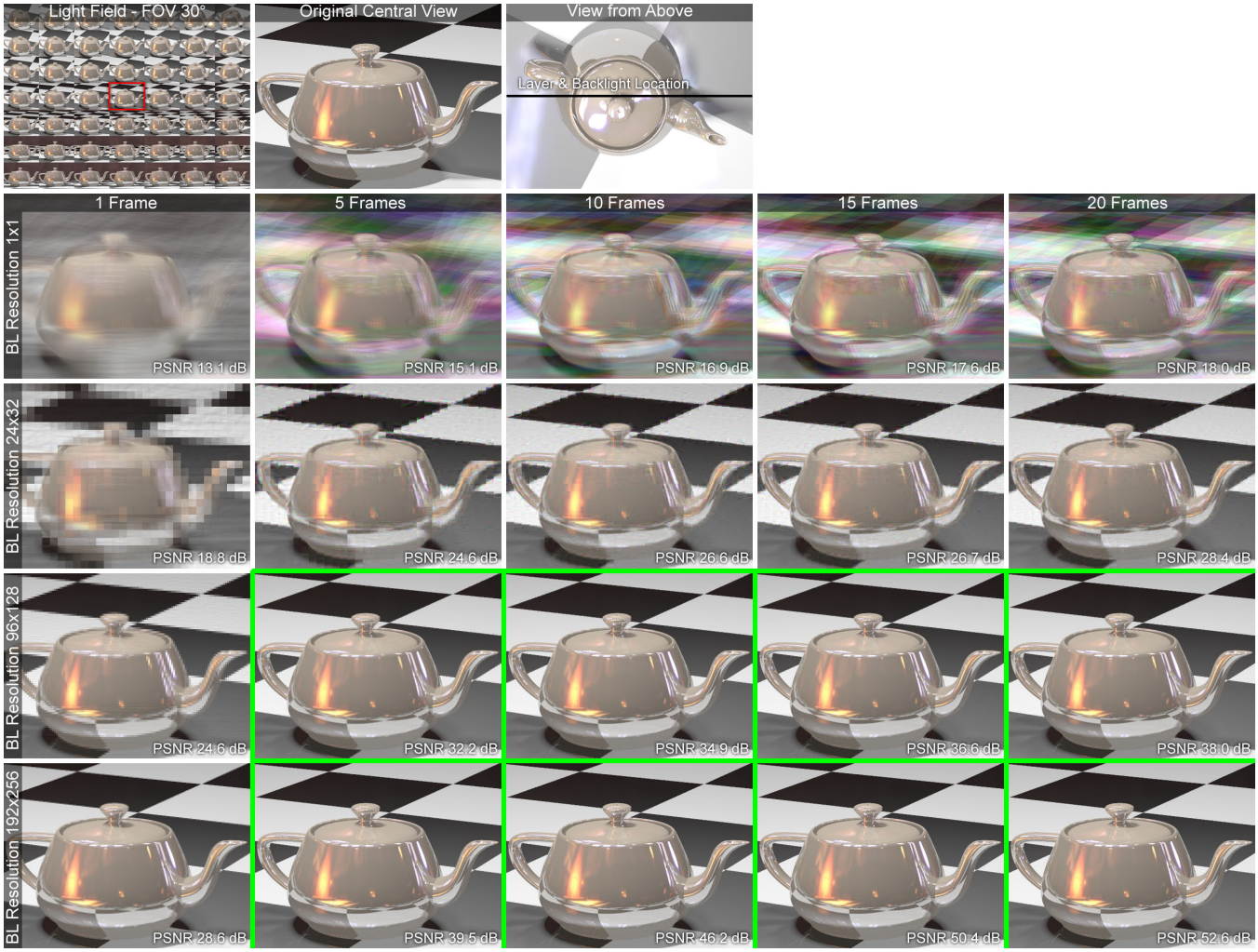

Figure S.3: *Simulated reconstructions of the central view for a light field covering a field of view of 30° with a varying number of frames and different spatial resolutions of the backlight. The spatial resolution of the layer is* 512×384*; the backlight is simulated to have (rows, from bottom) a spatial resolution of a factor of 2, 4, and 16 times lower than the layer resolution, as well as no spatial resolution at all (row two). As illustrated by the green boxes, a backlight with a spatial resolution of 1/4-1/8 of that of the layer can achieve high-quality reconstructions for only a few temporally-multiplexed frames.*

tion parallax. Although these high-frequencies could be perceived as noise, they actually contain the information that increases the 3D image quality for temporally-multiplexed tensor displays. With these experiments, we confirm that multilayer decompositions computed with nonnegative tensor factorization are structurally similar to the tomographic case if no temporal multiplexing is used, but combine the advantages of multiple layers with temporal multiplexing for all other cases.

In Figures S.5 and S.6 we analyze the behavior of NTF with respect to the number of update iterations and the rank; these results compare photographs of our three-layer prototype. A minimum of 50 iterations is generally necessary to ensure high image fidelity, but about 6–12 time-multiplexed frames achieve a high image quality even for the challenging teapot scene exhibiting a large depth of field. Figure S.7 demonstrates how light fields with uncorrelated views, such as Arabic numerals, can be successfully synthesized using the proposed low-rank tensor factorization.

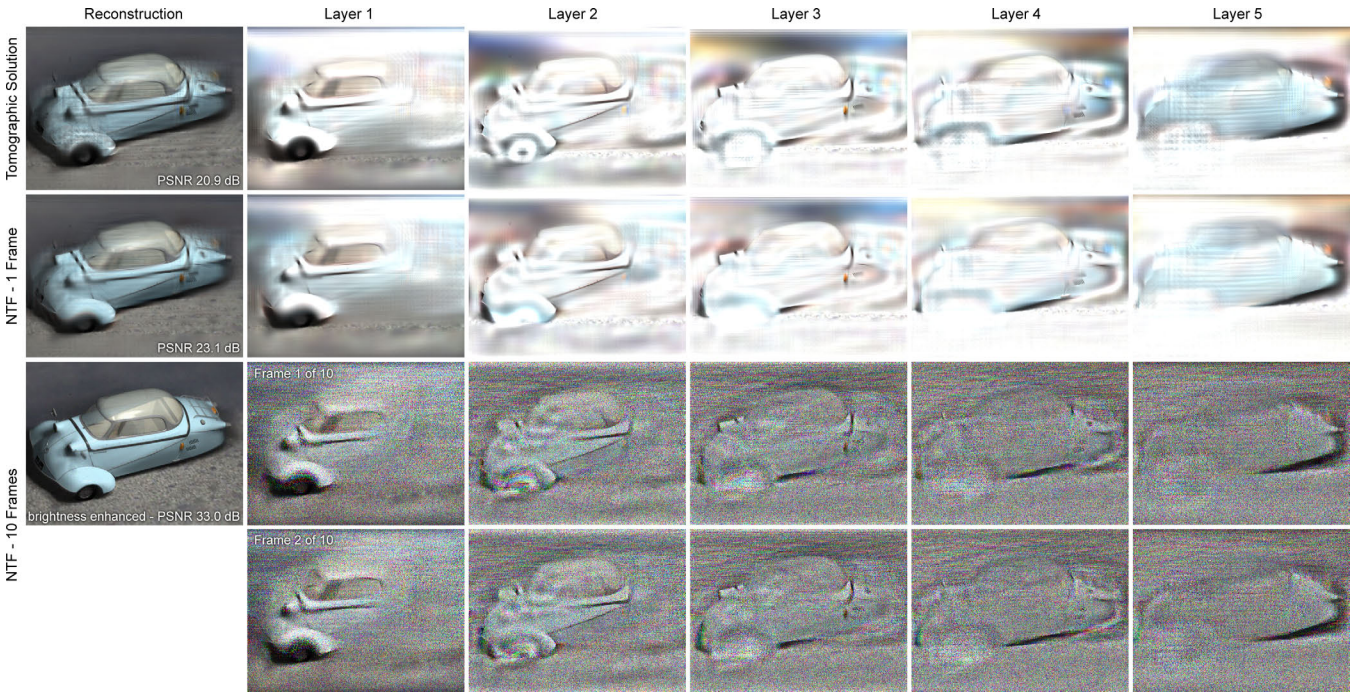

Figure S.4: *Multilayer decompositions. A tomographic five-layer decomposition (upper row) is intuitive, because it acts similar to a 3D geometry slicing operator for diffuse objects inside the physical display enclosure. For global illumination effects and objects outside the layers, however, the decompositions are more complicated. A nonnegative tensor factorization for the same optical configuration, without any time multiplexing, is shown in row two. The decompositions show a close similarity to the tomographic solution. The difference between the two is that the tomographic solution is computed in log-space, resulting in a linear problem which can be solved efficiently, but with biased errors. As seen in column one, specular highlights are slightly blurred and artifacts resulting from the parallax between different viewpoints are more pronounced. By adding temporal multiplexing, as shown in the lower two rows, the achieved quality can be significantly improved. The decompositions themselves still resemble a slicing operator in the lower frequencies, but what is perceived as temporally-varying high-frequency noise (lower two rows, columns two to six) actually contains the information necessary to improve the resulting 3D image quality. Note that any multiframe decompositions computed with NTF represent a tradeoff between PSNR and brightness; the latter is enhanced for the simulated reconstruction in row three.*

B.2.5 Multilayer and Purely Directional Backlight

In addition to the multilayer-only decompositions, we show decompositions for a dual-layer display with an additional, purely directional backlight in Figure S.8. This setup resembles dual-layer configurations explored in [Lanman et al. 2010], but generalizes to include an additional directional backlight. The layer decompositions exhibit high spatial frequencies, as analyzed in [Lanman et al. 2010], whereas the directional backlight adds more degrees of freedom that increase the field of view and depth of field of the tensor display as compared to a dual-layer configuration (see Section 4 of the primary text).

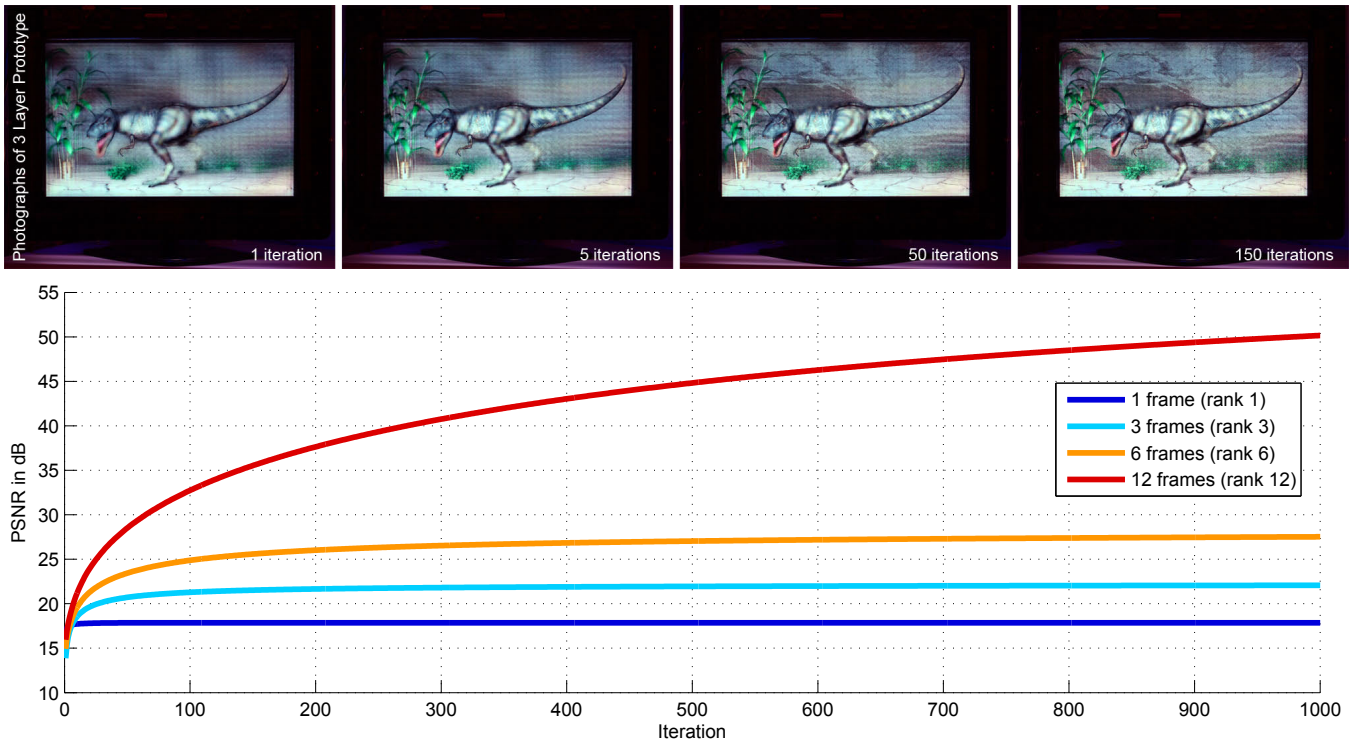

Figure S.5: *Convergence rate of multiplicative update rules. Top: four photographs of our three layer prototype showing an increasing number of NTF iterations for a rank 6 light field. At least 50 iterations (center right) are necessary to produce high-quality 3D images. Using only a few iterations (top left two photographs) result in blurred reconstructions, whereas a larger number of iterations (top right) do not significantly improve quality and represent an increased overhead in processing times. Bottom: plots showing convergence rates of the multiplicative update rules simulating the above experiments using 1, 3, 6, and 12 frames, respectively. With an increasing number of unknowns (i.e., larger numbers of frames), more iterations are required to converge.*

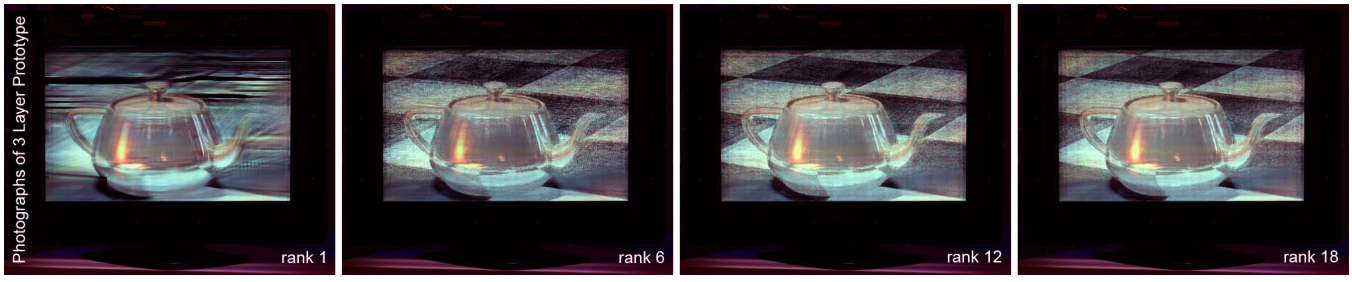

Figure S.6: *Rank analysis. Four photographs of our three layer prototype showing an increasing rank of the light field tensor. Without any time multiplexing (left photograph), low image quality is achieved for this scene due to the large depth of field. Low-rank approximations using 6 (center left) and 12 (center right) time-multiplexed frames create a visually appealing approximation of the light field. Higher-rank factorizations (right photograph) do not improve image quality significantly, demonstrating that light field tensors are inherently of low rank.*

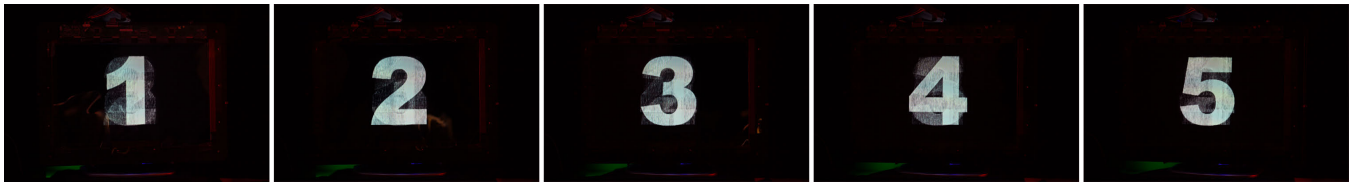

Figure S.7: *Synthesis of uncorrelated views. Five photographs of the three-layer prototype showing the performance of a rank 12 factorization for a light field comprising uncorrelated views (i.e., Arabic numerals in this example).*

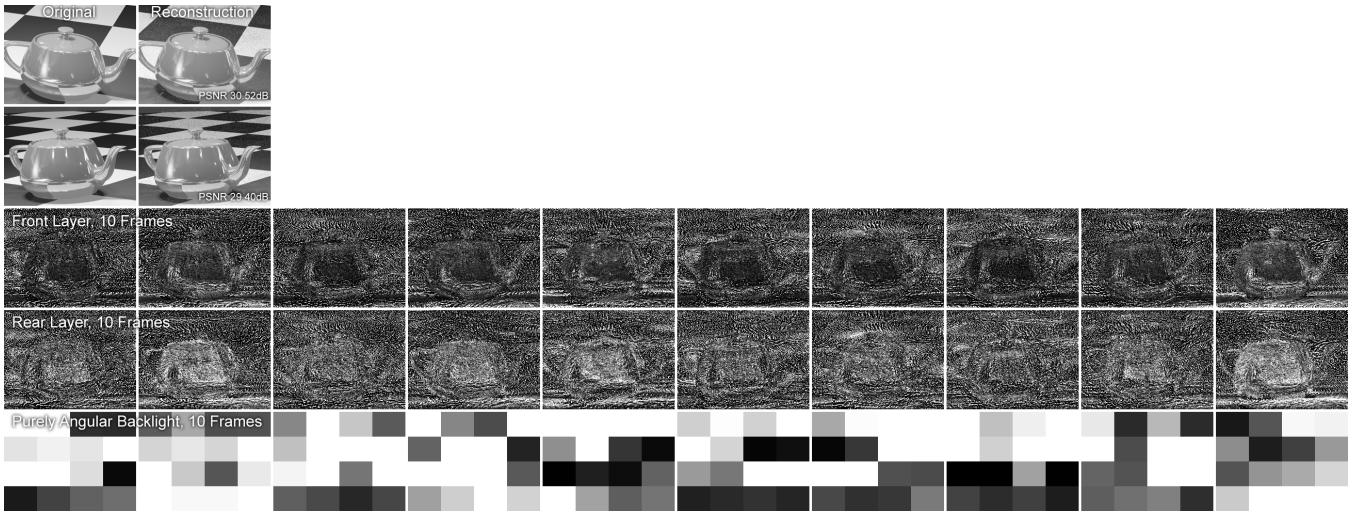

Figure S.8: *Decomposition for dual-layer display containing a purely angular backlight behind the rear layer. The original light field has* 4×4 *views within a field of view of* 30°*. Two of the original views are shown on the upper left with corresponding reconstructions next to them. This data set represents a rank-16 light field, which is decomposed using 10 frames. The two layers are separated by a distance that corresponds to the separation distance for an equivalent parallax barrier display. Each frame is shown for the front layer (row 3), for the rear layer (row 4), and for the angular backlight (bottom row). The layer decompositions resemble what NMF produces for dual-layer setups (see [Lanman et al. 2010]), but adds an angular backlight (see Section B.2.1) for improved depth of field and field of view, as demonstrated in the primary text.*

C Depth of Field for Tensor Displays

This supplementary appendix provides an extended analysis of the depth of field for tensor displays, expanding on Section 4.1 of the primary text. The performance of an automultiscopic display can be quantified by its depth of field: an expression for the maximum spatial frequency that can be depicted in a plane oriented parallel to the screen, as a function of the separation of this plane from the display surface. We apply a frequency-domain analysis of various tensor display architectures to derive analytic expressions for the upper bound on the depth of field, inspired by Zwicker et al. [2006] and Wetzstein et al. [2011]. We begin by assessing conventional automultiscopic displays, including parallax barriers [Ives 1903] and integral imaging [Lippmann 1908]. This analysis is extended to multilayer displays with uniform and directional backlighting in Sections C.2 and C.3, respectively.

C.1 Depth of Field for Conventional Automultiscopic Displays

As described by Zwicker et al. [2006], the *depth of field* of an automultiscopic display is an expression describing the maximum spatial frequency $\omega_{\xi_{\max}}$ that can be depicted, without aliasing, in a virtual plane oriented parallel to, and located a distance d_o from, the display surface. As established in that work, 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 (x, v) produces a light field spectrum $\hat{l}(\omega_x, \omega_v)$ that is non-zero only within a rectangle, where ω_x and ω_v are the spatial and angular frequencies, respectively. As described by Chai et al. [2000] and Durand et al. [2005], the spectrum of a Lambertian surface, located a distance d_o from the display surface, corresponds to the line $\omega_v = (d_o/d_r)\omega_x$ in the frequency domain, where d_r is the distance between the x-axis and v-axis, following Figure 3 in Section 3. The spatial cutoff frequency $\omega_{\xi_{\max}}$ is given by the intersection of this line with the spectral bandwidth of the display.

Consider the emitted light field spectrum for a parallax barrier or integral imaging display. Let Δx denote the spatial sampling rate (i.e., the spacing between barrier slits/pinholes or lenslets. Similarly, let Δv denote the angular sampling rate; for a conventional automultiscopic display with field of view α and A distinct views, the angular sampling rate $\Delta v = (2d_r/A)\tan(\alpha/2)$. Thus, the light field spectrum for a conventional automultiscopic display is non-zero only for $|\omega_x| \leq 1/(2\Delta x)$ and $|\omega_v| \leq 1/(2\Delta v)$. Intersecting the line $\omega_v = (d_o/d_r)\omega_x$ with this rectangular region yields the following expression for the depth of field.

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

This expression supports the following intuition: near the display (i.e., for $|d_o| \leq d_r(\Delta x/\Delta v)$), the maximum spatial frequency that can be depicted in a virtual plane, separated by d_o, is limited by the spacing Δx between slits/pinholes or lenslets. Far from the display, however, the maximum spatial frequency is limited by the angular sampling rate Δv and is independent of the spacing between the slits/pinholes or lenslets.

C.2 Upper Bound on Depth of Field for Multilayer Displays with Uniform Backlighting

The upper bound on the depth of field for a *static* multilayer display with uniform backlighting is similarly assessed by considering the maximum spectral bandwidth, for all possible layer mask patterns. Following Equation 1, the emitted light field $l(x, v)$ of an N-layer display is given by

$$l(x, v) = \prod_{n=1}^{N} f^{(n)}(x + (d_n/d_r)v), \tag{S.2}$$

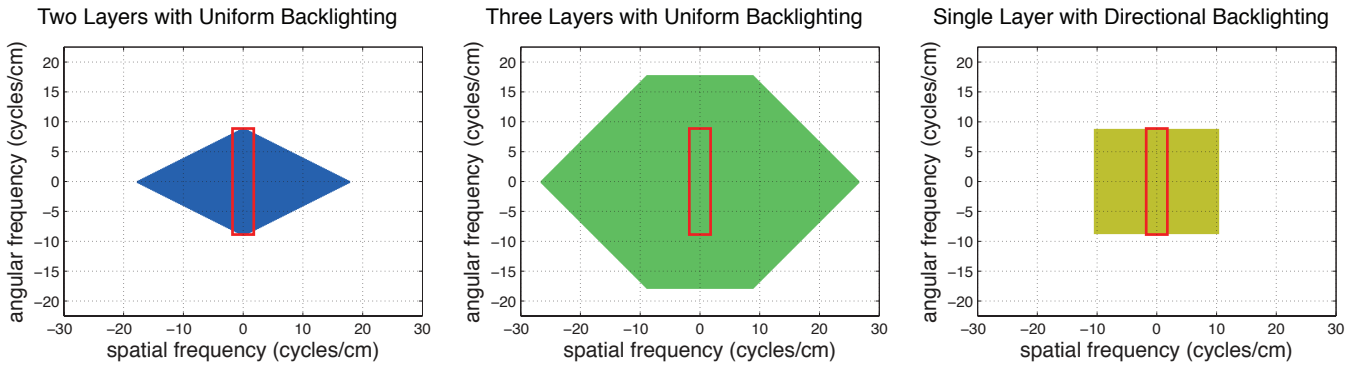

Figure S.9: *Spectral support for multilayer displays with uniform and directional backlighting. The spectral support is illustrated for two-layer and three-layer displays with uniform backlighting, shaded blue and green, on the left and in the middle, respectively. The spectral support for a single-layer display with directional backlighting is shaded yellow on the right. The spectral support for a conventional light field display (e.g., parallax barriers or integral imaging) is denoted by the red box. Display parameters correspond to those in Figure 4 of the primary text.*

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. The 2D Fourier transform of this expression is given by

$$\hat{l}(\omega_x, \omega_v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{n=1}^{N} f^{(n)}(x + (d_n/d_r)v)\, e^{-2\pi j \omega_x x}\, e^{-2\pi j \omega_v v}\, dx\, dv. \tag{S.3}$$

By the convolution property of Fourier transforms [Bracewell 1999], this expression reduces to a *repeated convolution* of the individual mask spectral $\hat{f}^{(n)}(\omega_x, \omega_v)$, such that the light field spectrum $\hat{l}(\omega_x, \omega_v)$ is given by

$$\hat{l}(\omega_x, \omega_v) = \overset{N}{\underset{n=1}{\circledast}} \hat{f}^{(n)}(\omega_x)\, \delta(\omega_v - (d_n/d_r)\omega_x), \tag{S.4}$$

where $*$ denotes convolution and the repeated convolution operator is defined as

$$\overset{N}{\underset{n=1}{\circledast}} \hat{f}^{(n)}(\omega_x, \omega_v) \equiv \hat{f}^{(1)}(\omega_x, \omega_v) * \hat{f}^{(2)}(\omega_x, \omega_v) * \cdots * \hat{f}^{(N)}(\omega_x, \omega_v). \tag{S.5}$$

Note that each layer produces a spectrum $\hat{f}^{(n)}(\omega_x, \omega_v) = \hat{f}^{(n)}(\omega_x)\, \delta(\omega_v - (d_n/d_r)\omega_x)$ that lies along the line $\omega_v = (d_n/d_r)\omega_x$, following Chai at al. [2000]. Since each layer has a finite resolution, the individual layer spectra $\hat{f}^{(n)}(\omega_x)$ may be non-zero only for $|\omega_x| \leq \omega_0 = 1/(2p)$, where p is the pixel pitch.

In the preceding analysis we have focused on static multilayer displays. Note that, by the linearity property of the Fourier transform, the inclusion of time multiplexing does not alter the maximum spectral support for a given multilayer display configuration. That is, both static and time-multiplexed multilayer displays have identical spectral bandwidths. Yet, as shown in Figure 4, the added degrees of freedom afforded by temporal multiplexing allow time-multiplexed multilayer displays to more closely approach the upper bound on the depth of field.

C.2.1 Two-Layer Displays with Uniform Backlighting

Consider a two-layer display with uniform backlighting, with the layers separated by a distance Δd and $\omega_0 = 1/(2p)$ denoting the spatial cutoff frequency for each layer with pixel pitch p. Equation S.4 yields the following analytic expression for the light field spectrum.

$$\hat{l}(\omega_x, \omega_v) = \hat{f}^{(1)}(\omega_x)\, \delta(\omega_v + (\Delta d/(2d_r))\omega_x) * \hat{f}^{(2)}(\omega_x)\, \delta(\omega_v - (\Delta d/(2d_r))\omega_x) \tag{S.6}$$

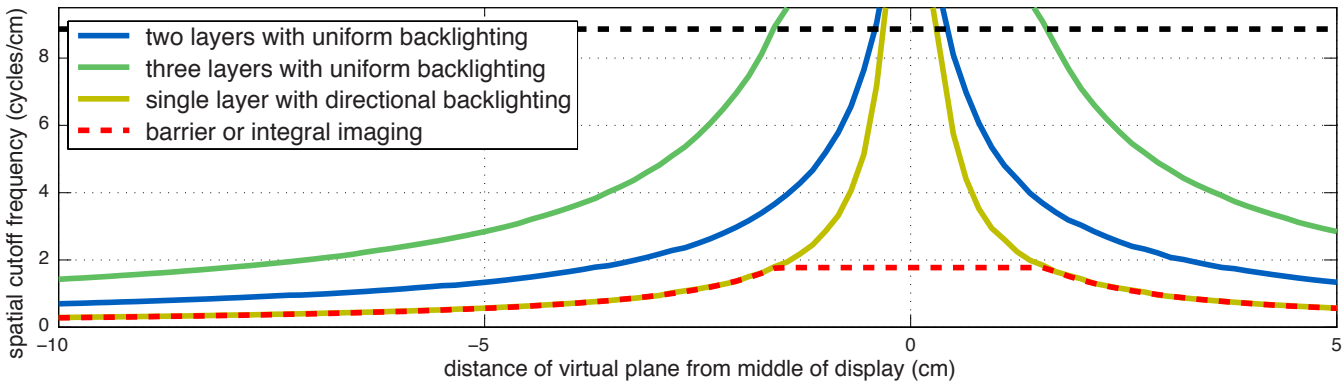

Figure S.10: *Comparison of the upper bound on the depth of field for conventional light field displays (red), two-layer (blue) and three-layer (green) displays with uniform backlighting, and single-layer displays with directional backlighting (yellow). The dashed black line denotes the spatial cutoff frequency ω_0 for each layer. Display parameters correspond to those in Figure 4 of the primary text.*

As shown in Figure S.9, a diamond-shaped region encloses the non-zero spectral support for any two-layer display. This region is bounded by a rectangle, such that $|\omega_x| \leq 2\omega_0$ and $|\omega_v| \leq (\Delta d/d_r)\omega_0$. Following the approach of Zwicker et al. [2006], 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 maximum-achievable spectral support. This geometric construction yields the following upper bound on the depth of field for any two-layer display.

$$\omega_{\xi_{\max}}(d_o) = \left(\frac{2\Delta d}{\Delta d + 2|d_0|}\right)\omega_0 \tag{S.7}$$

Note that this expression is equivalent to that previously derived for two-layer displays by Wetzstein et al. [2011].

C.2.2 Three-Layer Displays with Uniform Backlighting

Equation S.4 yields the following analytic expression for the light field spectrum for any three-layer display, with layers separated by a distance Δd.

$$\hat{l}(\omega_x, \omega_v) = \hat{f}^{(1)}(\omega_x)\,\delta(\omega_v + (\Delta d/d_r)\omega_x) * \hat{f}^{(2)}(\omega_x)\,\delta(\omega_v) * \hat{f}^{(3)}(\omega_x)\,\delta(\omega_v - (\Delta d/d_r)\omega_x) \tag{S.8}$$

As shown in Figure S.9, a hexagonal region encloses the non-zero spectral support for any three-layer display. This region is bounded by a rectangle, such that $|\omega_x| \leq 3\omega_0$ and $|\omega_v| \leq 2(\Delta d/d_r)\omega_0$. Intersecting the line $\omega_v = (d_o/d_r)\omega_x$ with the boundary of the maximum-achievable spectral support again yields the following upper bound on the depth of field.

$$\omega_{\xi_{\max}}(d_o) = \begin{cases} \left(\frac{3\Delta d}{\Delta d + |d_0|}\right)\omega_0 & \text{for } |d_o| \leq 2\Delta d, \\ \left(\frac{2\Delta d}{|d_o|}\right)\omega_0 & \text{otherwise} \end{cases} \tag{S.9}$$

As shown in Figure S.9, the bandwidth for three-layer displays exceeds that for either conventional or two-layer architectures, motivating the development of multilayer displays for extended depth of field. Note that, using the "geometric construction" outlined for two-layer and three-layers displays, one may derive an analytic upper bound for an arbitrary number of layers.

C.3 Upper Bound on Depth of Field for Multilayer Displays with Directional Backlighting

The upper bound on the depth of field for a multilayer display with directional backlighting is assessed, similar to Section C.2, by considering the maximum spectral bandwidth, for all possible layer mask and backlight illumination

patterns. Following Equation 18, the emitted light field $l(x, v)$ of an N-layer display is given by

$$l(x,v) = b(x,v) \prod_{n=1}^{N} f^{(n)}(x + (d_n/d_r)v), \tag{S.10}$$

where $b(x, v)$ denotes the light field emitted by the directional backlight. As described in Section 3.3, we consider directional backlighting to be equivalent to any low-resolution light field display placed behind a stack of light-attenuating layers. As before, the 2D Fourier transform of this expression is given by

$$\hat{l}(\omega_x, \omega_v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(x,v) \prod_{n=1}^{N} f^{(n)}(x + (d_n/d_r)v)\, e^{-2\pi j \omega_x x}\, e^{-2\pi j \omega_v v}\, dx\, dv, \tag{S.11}$$

yielding the following emitted light field spectrum:

$$\hat{l}(\omega_x, \omega_v) = \hat{b}(\omega_x, \omega_v) * \left[\overset{N}{\underset{n=1}{\circledast}} \hat{f}^{(n)}(\omega_x)\, \delta(\omega_v - (d_n/d_r)\omega_x) \right]. \tag{S.12}$$

C.3.1 Single-Layer Displays with Directional Backlighting

Consider a single-layer display with directional backlighting. For example, in Section 5.2.2 we describe a prototype consisting of a single LCD panel placed directly on top of an integral imaging display, comprised of a lenslet array affixed to a second LCD panel. Consistent with Section C.1, we assume that the directional backlight implements a low-resolution light field display, such that $\hat{b}(\omega_x, \omega_v)$ has non-zero support for $|\omega_x| \leq 1/(2\Delta x)$ and $|\omega_v| \leq 1/(2\Delta v)$ (i.e., the red box in Figure S.9). Equation S.12 yields the following analytic expression for the light field spectrum.

$$\hat{l}(\omega_x, \omega_v) = \hat{b}(\omega_x, \omega_v) * \hat{f}(\omega_x)\, \delta(\omega_v) \tag{S.13}$$

Note that the layer spectrum $\hat{f}(\omega_x, \omega_v) = \hat{f}(\omega_x)\, \delta(\omega_v)$ is constrained to a horizontal line of width $|\omega_x| \leq w_0$. Thus, convolution with the directional backlight spectrum results in an extended rectangular spectral support exceeding that of a conventional light field display; as shown in Figure S.9, the rectangular region is given by $|\omega_x| \leq 1/(2\Delta x) + \omega_0$ and $|\omega_v| \leq 1/(2\Delta v)$. Note that placing a light-attenuating layer directly on top of a conventional light field display only increases the spatial resolution; the angular resolution of the underlying low-resolution light field display is preserved. Intersecting the line $\omega_v = (d_o/d_r)\omega_x$ with the boundary of the maximum-achievable spectral support yields the following upper bound on the depth of field.

$$\omega_{\xi_{\max}}(d_o) = \begin{cases} \frac{1}{2\Delta x} + \omega_0 & \text{for } |d_o| \leq d_r \left(\frac{\Delta x}{\Delta v + 2\Delta x \Delta v \omega_0} \right), \\ \frac{d_r}{2|d_o|\Delta v} & \text{otherwise} \end{cases} \tag{S.14}$$

As shown in Figure S.10, the addition of a single light-attenuating layer significantly increases the spatial resolution near the display surface, as compared to a conventional parallax barrier or integral imaging display. However, far from the display, the upper bound on the depth of field is identical to these conventional automultiscopic displays.

C.4 Conclusion

The analysis presented in this supplementary appendix 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). 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.

D Constructing Tensor Display Prototypes

Figure S.11: *Prototype construction. Three LCD panels were modified to implement two-layer and three-layer tensor displays. Custom waterjet-cut and laser-cut parts ensured accurate alignment of the display components.*

E Additional Details on the GPU-based NTF Implementation

This section documents our GPU-based implementation of nonnegative tensor factorization for tensor displays. As explained in Section 5.1 of the primary text, all underlying operators for this particular application map well to functions of the fixed graphics pipeline. We implement NTF using OpenGL and a set of CG shaders. Pseudocode is listed on the following pages. We assume the tensor display consists of L light-attenuating layers, each displaying F frames in rapid succession. An optional, low-resolution directional backlight is also supported. Both the original light field and the backlight are assumed to consist of V different views. As the display of the decomposed layers is a time-critical operation, requiring a frame rate that matches the monitor refresh rate, it is implemented in a different thread than the decomposition, which can be run at a lower frame rate.

The pseudocode below documents the main display loop for synchronized rendering of temporally-multiplexed layers and the backlight with monitor refresh rates. This implementation assumes that calibrated interlacing masks are available for each view of the light field. These masks are multiplied by the corresponding rendered view and added together to generate an interlaced image to be displayed behind a lenslet array. The decomposition routines are documented on the following pages, implementing weighted nonnegative tensor factorization, as discussed in Section 3.2 of the primary text and in Section A.2 of this supplement.

Algorithm NTF - Main Display Routines

variables FBO_LAYERS[L][F], FBO_BACKLIGHT[V][F], INTERLACING_MASKS[V], f=0, $bUseBacklight$=true

function mainDisplayLoop()
    // draw layers of current frame
    for all layers l
        set viewport for l
        activate FBO_LAYERS[l][f]
        draw textured 2D quad
    end
    // draw backlight of current frame
    if bUseBacklight
        set viewport for backlight
        activate accumulation buffer
        for all views v
            activate CG_SHADER_MULTIPLY_TWO_TEXTURES(INTERLACING_MASK[v], FBO_BACKLIGHT[v][f])
            draw textured 2D quad
        end
        deactivate accumulation buffer
    end
    // cycle through frames
    f = ($f < F$) ? f+1 : 0;
end

Algorithm NTF - Content-Updating Thread

```
variables FBO_LF[V], FBO_LF_REC[V], FBO_LF_TMP[V], FBO_LAYER_TMP[2], FBO_TMP

function threadDisplayLoop()
    // draw light field
    for all light field views v
        activate FBO_LF[v]
        set perspective v
        drawScene(); // render desired 3D scene (e.g., a teapot)
    end
    // factorize light field using NTF
    NTF();
end

function NTF()
    for all iterations i
        // update the layers
        for all layers l
            // draw current estimate of LF into rec buffers
            drawLightFieldFromLayersRec();
            for all frames f
                // draw layers into LF tmp buffers, but leave out current layer
                drawLightFieldFromLayersTmp(l);
                // compute numerator for multiplicative NTF update
                activate FBO_LAYER_TMP[1]
                activate accumulation buffer
                for all light field views v
                    set perspective i as projective texture matrix
                    activate CG_SHADER_MULT2TEXTURES_AND_PROJECTIVE_TEXMAPTHEM( FBO_LF[v], FBO_LF_TMP[v] )
                    draw 2D quad
                end
                deactivate FBO_LAYER_TMP[1]
                // compute denominator for multiplicative NTF update
                activate FBO_LAYER_TMP[2]
                activate accumulation buffer
                for all light field views v
                    set perspective i as projective texture matrix
                    activate CG_SHADER_MULT2TEXTURES_AND_PROJECTIVE_TEXMAPTHEM( FBO_LF_REC[v], FBO_LF_TMP[v] )
                    draw 2D quad
                end
                deactivate FBO_LAYER_TMP[2]
                // update current layer for current frame
                activate FBO_LAYERS[l][f]
                activate CG_SHADER_MULT2TEXTURES_DIVIDEBYOTHER ( FBO_LAYERS[l][f], FBO_LAYER_TMP[1], FBO_LAYER_TMP[2] )
                draw 2D quad
                deactivate FBO_LAYERS[l][f]
            end
        end
        // update the backlight
        if bUseBacklight
            // draw current estimate of LF into rec buffers
            drawLightFieldFromLayersRec();
            for all frames f
                // draw layers into LF tmp buffers, but leave out backlight
                drawLightFieldFromLayersTmp(-1);
                for all views v
                    set perspective i as projective texture matrix
                    // compute numerator for multiplicative NTF update
                    activate FBO_BL_TMP[1]
                    activate CG_SHADER_MULT2TEXTURES_AND_PROJECTIVE_TEXMAPTHEM( FBO_LF[v], FBO_LF_TMP[v] )
                    draw 2D quad
                    deactivate FBO_BL_TMP[1]
                    // compute denominator for multiplicative NTF update
                    activate FBO_BL_TMP[2]
                    activate CG_SHADER_MULT2TEXTURES_AND_PROJECTIVE_TEXMAPTHEM( FBO_LF_REC[v], FBO_LF_TMP[v] )
                    draw 2D quad
                    deactivate FBO_BL_TMP[2]
                    // downsample the backlight TMP FBOs to backlight resolution by adding up the values
                    downsampleAndAdd(FBO_BL_TMP[1, 2]);
                    // update current backlight view for current frame
                    activate FBO_BL[v][f]
                    activate CG_SHADER_MULT2TEXTURES_DIVIDEBYOTHER ( FBO_BL[v][f], FBO_BL_TMP[1], FBO_BL_TMP[2] )
                    draw 2D quad
                    deactivate FBO_BL[v][f]
                end
            end
        end
    end
end
```

Algorithm NTF - Additional Helper Functions

```
function drawLightFieldFromLayersRec()
    convertAllLayersAndBacklightToLOG();
    for all views v
        for all frames f
            activate FBO_TMP
            activate accumulation buffer
            for all layers l
                draw layer l, textured with FBO_LAYERS[l][f]
            end
            if bUseBacklight
                draw backlight, textured with FBO_BACKLIGHT[v][f]
            end
            deactivate FBO_TMP
            activate FBO_LF_REC[v]
            activate accumulation buffer
            activate CG_SHADER_DRAW_EXPONENTIAL_TEXTUE ( FBO_TMP )
            deactivate FBO_LF_REC[v]
        end
    end
    convertAllLayersAndBacklightFromLOG();
end

function drawLightFieldFromLayersTmp(int leaveOutLayerX)
    convertAllLayersAndBacklightToLOG();
    for all views v
        for all frames f
            activate FBO_TMP
            activate accumulation buffer
            for all layers l
                if leaveOutLayerX!=l
                    draw layer l, textured with FBO_LAYERS[l][f]
                end
            end
            if bUseBacklight && (leaveOutLayerX!=-1)
                draw backlight, textured with FBO_BACKLIGHT[v][f]
            end
            deactivate FBO_TMP
            activate FBO_LF_TMP[v]
            activate accumulation buffer
            activate CG_SHADER_DRAW_EXPONENTIAL_TEXTUE ( FBO_TMP )
            deactivate FBO_LF_TMP[v]
        end
    end
    convertAllLayersAndBacklightFromLOG();
end
```

Supplementary References

ACAR, E., DUNLAVY, D. M., KOLDA, T. G., AND MØRUP, M. 2011. Scalable tensor factorizations for incomplete data. *Chemometrics and Intelligent Laboratory Systems 106*, 1, 41–56.

BERRY, M. W., BROWNE, M., LANGVILLE, A. N., PAUCA, V. P., AND PLEMMONS, R. J. 2006. Algorithms and applications for approximate nonnegative matrix factorization. In *Computational Statistics and Data Analysis*, 155–173.

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.

BRACEWELL, R. 1999. *The Fourier Transform and Its Applications (Third Edition)*. McGraw-Hill.

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

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

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.

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

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

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.

LEE, D. D., AND SEUNG, H. S. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature 401*, 6755, 788–791.

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

MØRUP, M., HANSEN, L. K., PARNAS, J., AND ARNFRED, S. M. 2006. Decomposing the time-frequency representation of EEG using non-negative matrix and multi-way factorization. Tech. rep.

WELLING, M., AND WEBER, M. 2001. Positive tensor factorization. *Pattern Recognition Letters 22*, 1255–1261.

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*.