

Cascaded Displays: Spatiotemporal Superresolution using Offset Pixel Layers

Felix Heide Douglas Lanman Dikpal Reddy Jan Kautz Kari Pulli David Luebke
NVIDIA Research

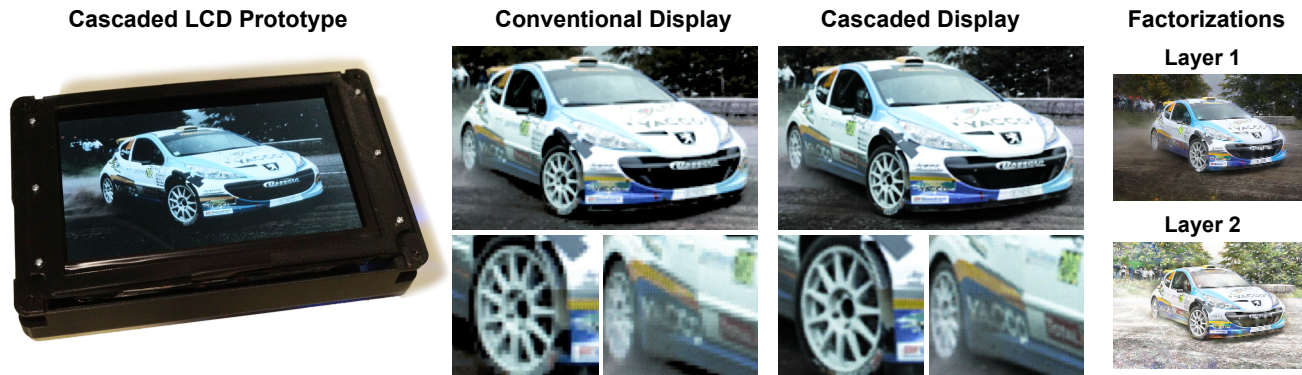


Figure 1: A pair of off-the-shelf LCDs were modified to create a cascaded display: placed in direct contact with a fixed lateral offset. In comparison to any single LCD used in their construction, cascaded displays can quadruple spatial resolution by presenting attenuation layer patterns that are optimized, in real time, using non-negative matrix factorization. (Motorsport image courtesy Aurélien Vialatte.)

Abstract

We demonstrate that layered spatial light modulators (SLMs), subject to fixed lateral displacements and refreshed at staggered intervals, can synthesize images with greater spatiotemporal resolution than that afforded by any single SLM used in their construction. Dubbed *cascaded displays*, such architectures enable superresolution flat panel displays (e.g., using thin stacks of liquid crystal displays (LCDs)) and digital projectors (e.g., relaying the image of one SLM onto another). We introduce a comprehensive optimization framework, leveraging non-negative matrix and tensor factorization, that decomposes target images and videos into multi-layered, time-multiplexed attenuation patterns—offering a flexible trade-off between apparent image brightness, spatial resolution, and refresh rate. Through this analysis, we develop a real-time dual-layer factorization method that quadruples spatial resolution and doubles refresh rate. Compared to prior superresolution displays, cascaded displays place fewer restrictions on the hardware, offering thin designs without moving parts or the necessity of temporal multiplexing. Furthermore, cascaded displays are the first use of multi-layer displays to increase apparent temporal resolution. We validate these concepts using two custom-built prototypes: a dual-layer LCD and a dual-modulation liquid crystal on silicon (LCoS) projector, with the former emphasizing head-mounted display (HMD) applications.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display Algorithms

Keywords: superresolution, multi-layer displays, dual modulation displays, non-negative matrix and tensor factorization

Links: [DL](#) [PDF](#) [WEB](#) [VIDEO](#)

1 Introduction

The development of higher-resolution displays is of central importance to the display industry. Leading mobile displays recently transitioned from pixel densities of less than 50 pixels per cm (ppcm) and now approach 150 ppcm. Similarly, the consumer electronics industry is beginning to offer “4K ultra-high definition (UHD)” displays, having a horizontal resolution approaching 4,000 pixels, as the successor to high-definition television (HDTV). Furthermore, 8K UHD standards already exist for enhanced digital cinema. Achieving such high-resolution displays currently hinges on advances that enable spatial light modulators with increased pixel counts.

Beyond these larger market trends, several emerging display technologies necessitate even greater resolutions than 4K/8K UHD standards will provide. For example, wide-field-of-view head-mounted displays (HMDs), such as the Oculus Rift, incorporate high-pixel-density mobile displays. Such displays already approach or exceed the resolution of the human eye when viewed at the distance of a phone or tablet computer; however, they appear pixelated when viewed through magnifying HMD optics, which dramatically expand the field of view. Similarly, glasses-free 3D displays, including parallax barrier [Ives 1903] and integral imaging [Lippmann 1908] designs, require an order of magnitude higher resolution than today’s displays. At present, HMDs and glasses-free 3D displays remain niche technologies and are less likely to drive the development of higher-resolution displays than the existing applications, hindering their advancement and commercial adoption.

We propose an alternative to achieve displays with increased spatial resolution using current-generation light-attenuating spatial light modulator (SLM) technology, including liquid crystal displays (LCDs), digital micromirror devices (DMDs), and liquid crystal on silicon (LCoS) displays. Rather than directly pursuing the “brute force” solution of increasing the addressable pixel count, we propose *cascaded displays* as an alternate construction: stacking two or more SLMs on top of one another, subject to a lateral offset of half a pixel or less along each axis. As shown in Figures 1 and 2, lateral offsets are necessary so that each pixel on one layer modulates multiple pixels on another; in this manner, the intensity of each *subpixel fragment*—defined by the geometric intersection of a pixel on one display layer with one on another layer—can be controlled, thereby increasing the effective display resolution. In this paper, we

	time multiplexing	multiple displays	moving parts	thin	pixel fill	image/video quality	temporal superres.
vibrating display elements [Allen and Ulichney 2005; Berthouzoz and Fattal 2012b]	yes	no	yes	yes	low	moderate→high	no
superimposed projections [Jaynes and Ramakrishnan 2003]	no	yes	no	no	low	moderate→high	no
smooth pursuit eye motion [Didyk et al. 2010; Berthouzoz and Fattal 2012a]	yes	no	no	yes	low	low (single direction)	no
optical pixel shift and overlay [Sajadi et al. 2013]	no	no	no	no	low	moderate	no
optical pixel sharing [Sajadi et al. 2012]	yes	yes	no	no	high	moderate→high	no
dual-modulation projectors [Kusakabe et al. 2008; Kusakabe et al. 2009]	no	yes	no	no	high	high	no
cascaded displays	optional	yes	no	yes	high	high	yes

Table 1: Comparison of superresolution displays. Cascaded displays are distinguished as the first system to support temporal superresolution.

describe how to factorize target images into multi-layer attenuation patterns, demonstrating that cascaded displays may operate as “compressive displays”: utilizing fewer independently-addressable pixels than apparent in the displayed image. We further demonstrate that similar methods may be adopted to increase the temporal resolution of stacks of two or more SLMs, refreshed in staggered intervals.

Cascaded displays offer several distinct advantages relative to prior superresolution displays: achieving thin form factors, without moving parts, and using recently-introduced, computationally-efficient factorization methods to enable interactive content. Most significantly, we demonstrate an operation mode that eliminates the need for temporal multiplexing of factorized imagery; as a result, videos can be presented without the appearance of artifacts characteristic of prior methods or the requirement for high-refresh-rate displays.

1.1 Contributions

Our primary technical contributions include the following:

- We demonstrate that a laterally-offset stack of two or more SLMs can synthesize a spatially-superresolved image. We further demonstrate that presentation of time-multiplexed attenuation patterns minimizes the appearance of artifacts, but is optional, since high fidelity is achieved without multiplexing.
- We demonstrate that a stack of two or more SLMs, refreshed at staggered intervals, can synthesize a video with an effective refresh rate that exceeds that of a single display layer by a factor equal to the number of layers. We further demonstrate that optically averaging neighboring pixels minimizes artifacts.
- We provide a comprehensive optimization framework, based on non-negative matrix and tensor factorization. Multiple algorithms are compared, revealing the benefits of weighted rank-1 residue iterations over prior multiplicative update rules.
- We derive and implement a real-time, GPU-accelerated cascaded display algorithm that eliminates the need for temporal multiplexing, while preserving superresolution image fidelity.
- We construct a pair of cascaded display prototypes: a dual-layer LCD screen, supporting direct-view and head-mounted display applications, and a dual-layer LCoS projector.

1.2 Overview of Limitations

We share the limitations of other multi-layer displays, including decreased brightness due to scattering, absorption, and interreflection, increased cost and complexity, the need for optical engineering to eliminate moiré patterns [Bell et al. 2008], synchronization challenges, and the need for achieving and maintaining precise alignment. Implementing practical cascaded LCDs for flat panel applications further requires confronting three challenges: limited pixel aperture ratios (fill factors), fixed color filter array (CFA) designs, and parallax due to a physical gap between modulation layers. Solutions for each of these challenges are discussed in Section 4. In contrast, cascaded projectors offer a more direct path to adoption; for example, LCoS microdisplays have high aperture ratios and reflectivity, dispense with CFAs in favor of field-sequential color (FSC), and eliminate physical gaps between layers through the use of relay optics. Regardless of the hardware embodiment, cascaded displays require additional computational resources, potentially mitigating

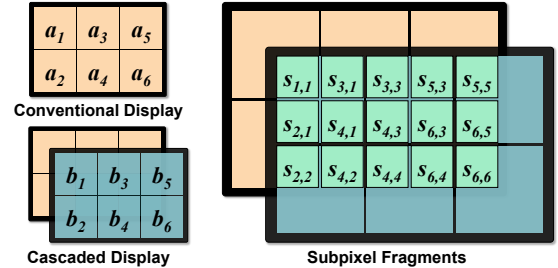


Figure 2: Creation of subpixel fragments by cascaded displays. (Left) A cascaded display is constructed by layering a shifting pair of conventional displays. (Right) The overlap of offset pixel layers creates an array of subpixel fragments (green). The emittance of subpixel fragment $s_{i,j}$ is given by $a_i b_j$, following Equation 1.

power benefits and increasing latency. As described in Section 3.2, we address this by introducing an efficient optimization algorithm based on weighted rank-1 residue iteration (WRR1) [Ho 2008].

2 Related Work

The following section briefly reviews related work, expanding on the comparison of superresolution display methods presented in Table 1.

2.1 Superresolution Displays

As described by Baker et al. [2002], superresolution imaging algorithms recover a high-resolution image (or video) from low-resolution images (or videos) with varying perspectives. Superresolution imaging requires solving an ill-posed inverse problem: the high-resolution source is unknown. Methods differ based on the prior assumptions made regarding the imaging process. For example, Ben-Ezra et al. [2004] eliminate camera motion uncertainty by using piezoelectric actuators to control sensor displacement.

Application of superresolution imaging concepts to displays is a relatively recent development. Unlike superresolution imaging, the target image content is known *a priori*. The construction of the display device may exploit spatial or temporal multiplexing to increase the effective number of addressable pixels; as a result, a decomposition problem must be solved to determine the optimal control of the display components to maximize the perceived resolution, subject to physical constraints (e.g., limited dynamic range, restricted color gamut, and prohibition of negative emittances).

Allen and Ulichney [2005] describe one of the earliest superresolution display systems; their “wobulation” method doubles the addressed resolution for front-projection displays incorporating a single high-speed DMD. Functioning in a manner akin to the “jittered camera” of Ben-Ezra et al. [2004], a piezoelectrically-actuated mirror displaces the projected image by half a pixel, both horizontally and vertically. Since DMDs can be addressed faster than the critical flicker fusion threshold [Hart 1987], two shifted images can be rapidly projected, so that the viewer perceives their **additive superposition**. As with the jittered camera, the superresolution factor increases as the pixel aperture ratio decreases [Baker and Kanade 2002]. Performance is further limited by motion blur introduced

during the optical scanning process. More recently, Berthouzoz et al. [2012b] extend wobulation to flat panel displays, using an eccentric rotating mass (ERM) vibration motor applied to an LCD.

Similar superresolution display concepts have been developed for digital projectors. Rather than presenting a time-multiplexed sequence of shifted, low-resolution images, projector arrays can be used to display the displaced image set simultaneously. Such “superimposed projection” systems have been demonstrated by Jaynes and Ramakrishnan [2003], Damera-Venkata and Chang [2009], and Aliaga et al. [2012]. As with all projector arrays, superimposed projections require precise radiometric and geometric calibration, as well as temporal synchronization. Mitigating these issues, Sajadi et al. [2013] describe a single-projector superresolution method wherein multiple offset images are created by an array of lenses within the projector optics. Unlike superimposed projectors, these images must be identical, resulting in limited image quality.

Wobulation and other temporally-multiplexed methods introduce artifacts when used to superresolve videos due to unknown gaze motion. Eye movement alters the desired alignment between subsequent frames, as projected on the retina. If the gaze can be estimated, then superresolution can be achieved along the eye motion trajectory, as demonstrated by Didyk et al. [2010] and Berthouzoz et al. [2012a].

All of the superresolution displays discussed thus far implement the same core concept: additive (temporal) superposition of shifted low-resolution images. As with image superresolution, such designs benefit from low pixel aperture ratios—diverging from industry trends to increase aperture ratios. In contrast, cascaded displays create a **multiplicative superposition**: synthesizing higher spatial frequencies by the (simultaneous) interference of shifted light-attenuating displays with large aperture ratios. In closely-related work, Sajadi et al. [2012] introduce “optical pixel sharing (OPS)”. Similar to cascaded displays, OPS uses two spatial light modulators in series. We emphasize that OPS is the first approach to exploit dual modulation projectors for superresolution by depicting an *edge-enhanced image* using a two-frame decomposition: the first frame presents a high-resolution, sparse edge image, whereas the second frame presents a low-resolution non-edge image. Unlike cascaded displays, OPS requires an element be placed between the display layers (e.g., an array of lenses or a randomized refractive surface); correspondingly, existing OPS implementations do not allow thin form factors. As analyzed in Section 5, OPS reproduces imagery with decreased brightness and decreased peak signal-to-noise ratio (PSNR) compared to cascaded displays employing similar layer architectures.

2.2 Dual-Modulation Displays

Dual-modulation displays are routinely applied to achieve high-dynamic-range (HDR) display. Seetzen et al. [2004] and Pavlovych and Stuerzlinger [2005] implement HDR projectors by modulating the output of a digital projector using large flat panel LCDs. In closely-related work, Kusakabe et al. [2008; 2009] describe a high-dynamic-range and high-resolution projector. In their system, a three-chip LCoS projector emits a low-resolution chrominance image, which is subsequently projected onto another, higher-resolution LCoS chip to achieve luminance modulation. Unlike this design, cascaded displays exceed the native resolution of any single SLM contained in their construction and offer a means to further extend the performance of emerging dual-modulation displays.

2.3 Multi-Layer 3D Displays

Displays with two or more SLMs are also incorporated in glasses-free 3D displays. Lanman et al. [2010] demonstrate that content-adaptive parallax barriers can be used with dual-layer LCDs to create

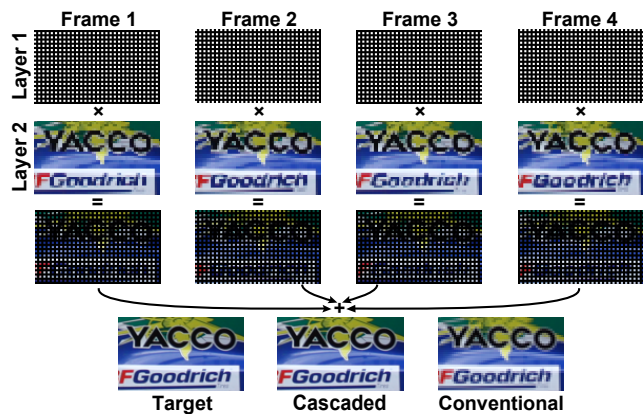


Figure 3: Heuristic factorization for spatial superresolution. (First and Second Rows) Following Section 3.2, a time-multiplexed sequence of shifted pinhole grids are displayed on the bottom layer (first row), together with aliased patterns on the top layer (second row). (Third Row) Each bottom-layer pixel illuminates the corners of four top-layer pixels. (Fourth Row) When the four frames are presented at a rate exceeding the flicker fusion threshold, the viewer perceives an image with four times the number of pixels in any layer. (Note that, unless the backlight brightness is increased, the cascaded display will appear dimmer than a conventional display.)

brighter, higher-resolution 3D displays. Gotoda [2010; 2011], Wetstein et al. [2011; 2012], and Lanman et al. [2011] generalize this concept to three or more layers. In related work, Heide et al. [2014] achieve superresolution using such multi-layer displays by requiring a diffuser to be placed slightly in front of the LCD stack. Relative to cascaded displays, this construction requires a thicker form factor, a greater number of time-multiplexed frames, and achieves a reduced superresolution factor. Rather than presenting multi-view imagery, our goal is to enhance spatiotemporal resolution. Operating cascaded displays to achieve superresolution places fewer practical restrictions: no physical gap is required between the layers, enabling thinner form factors, and significantly fewer time-multiplexed frames are necessary to eliminate image artifacts. We emphasize that cascaded displays are the first use of multi-layer displays to increase apparent temporal resolution; in addition, our introduction of the weighted rank-1 residue iteration (WRR1) offers a means to improve factorization algorithms for existing multi-layer 3D displays.

3 Cascaded Dual-Layer Displays

This section describes how to operate cascaded displays containing two SLM layers. Section 3.1 introduces an idealized image formation model. Section 3.2 establishes that cascaded dual-layer displays can quadruple the spatial resolution, without introducing any reconstruction artifacts, using a heuristic four-frame factorization. Section 3.2 further introduces a constrained least-squares optimization framework, allowing a flexible trade-off between spatial resolution, effective frame rate, image quality, and brightness. Sections 3.3 and 3.4 generalize these concepts to enable temporal superresolution and to leverage three or more SLM layers, respectively.

3.1 Modeling Cascaded Dual-Layer Displays

Consider an idealized dual-layer display, comprising a pair of spatial light modulators placed in direct contact in front of a uniform backlight and containing a uniform array of pixels with individually-addressable transmissivity at a fixed refresh rate. We observe that spatial superresolution only works with a lateral offset between the layers, as otherwise the pixels on the top layer would directly overlap

Algorithm 1 Weighted Rank-1 Residue Iterations (WRI)

```

1: Initialize  $\mathbf{A}$  and  $\mathbf{B}$ 
2: repeat
3:   for  $k = 1$  to  $K$  do
4:      $\mathbf{R}_k = \mathbf{T} - \sum_{i \neq k} \mathbf{a}_i \mathbf{b}_i^T$            ▷ Evaluate rank-1 residue.
5:      $\mathbf{a}_k \leftarrow \left[ \frac{[(\mathbf{W} \circ \mathbf{R}_k) \mathbf{b}_k]_+}{\mathbf{W}(\mathbf{b}_k \circ \mathbf{b}_k)} \right]_+$    ▷ Update column  $k$  of  $\mathbf{A}$ .
6:      $\mathbf{b}_k \leftarrow \left[ \frac{[(\mathbf{W} \circ \mathbf{R}_k)^T \mathbf{a}_k]_+}{\mathbf{W}^T(\mathbf{a}_k \circ \mathbf{a}_k)} \right]_+$    ▷ Update column  $k$  of  $\mathbf{B}$ .
7:   end for
8: until Stopping condition
  
```

those on the bottom layer. As shown in Figure 2, when the top layer is laterally shifted by half a pixel, both horizontally and vertically, then the pixel centers of the top layer coincide with the pixel corners of the bottom layer. As depicted, this configuration creates a uniform array of *subpixel fragments* defined by the overlap of pixels on the bottom layer with those on the top. Most significantly, there exist four times as many subpixel fragments as pixels on a single layer—establishing the capacity to quadruple the spatial resolution.

We assume a dual-layer display has a bottom layer with N pixels and a top layer with M pixels. Furthermore, we assume K time-multiplexed frames are presented to the viewer at a rate above the critical flicker fusion threshold, such that their temporal average is perceived. As with the systems proposed by Wetzstein et al. [2012] and Sajadi et al. [2012], we emphasize that temporal multiplexing increases the degrees of freedom available to reduce image artifacts, as analyzed in Figure 9. As shown in Figure 2, the emissivity of pixel i in the bottom layer, for frame k , is denoted $a_i^{(k)}$, such that $0 \leq a_i^{(k)} \leq 1$. Similarly, $b_j^{(k)}$ denotes the transmissivity of the pixel j of the top layer, for frame k , such that $0 \leq b_j^{(k)} \leq 1$. As a result, the emissivity of each subpixel fragment is given by $s_{i,j}$, such that

$$s_{i,j} = w_{i,j} \left(\sum_{k=1}^K a_i^{(k)} b_j^{(k)} \right), \quad (1)$$

where $w_{i,j}$ denotes the overlap of pixel i and pixel j . This expression implies that dual-layer image formation can be concisely expressed using matrix multiplication:

$$\mathbf{S} = \mathbf{W} \circ (\mathbf{A}\mathbf{B}^T), \quad (2)$$

where \circ denotes the Hadamard (elementwise) matrix product, \mathbf{A} is an $N \times K$ matrix, whose columns contain bottom layer pixel emissivities during frame k , \mathbf{B} is an $M \times K$ matrix, whose columns contain the top-layer pixel transmissivities during frame k , \mathbf{W} is an $N \times M$ sparse weight matrix, containing the pairwise overlaps, and \mathbf{S} is a sparse $N \times M$ matrix containing the subpixel fragment emissivities—it can be non-zero only where pixel i and pixel j overlap.

We emphasize the generality of the image formation model given by Equations 1 and 2. Dissimilar spatial light modulators can be represented, including panels with differing pixel pitches. Furthermore, relative lateral translations and in-plane rotations of the two layers can be encoded by an appropriate choice of the weight matrix \mathbf{W} . As described in Section 4, this model can be practically applied to existing flat panel displays (e.g., LCD panels containing color filter arrays and limited pixel aperture ratios) and digital projectors (e.g., those containing LCD, LCoS, or DMD spatial light modulators).

3.2 Spatial Superresolution

Heuristic Lossless Factorization Achieving spatial superresolution requires factorizing a target high-resolution image—sampled

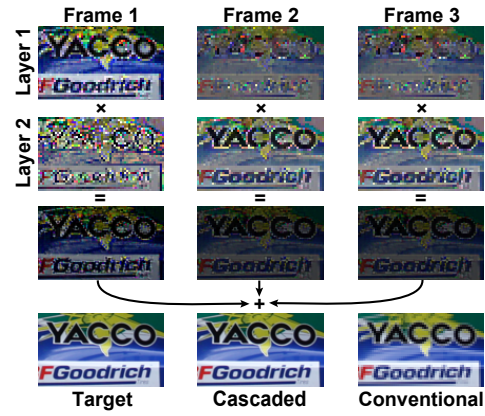


Figure 4: Optimized factorization for spatial superresolution. (First and Second Rows) Following Section 3.2, Algorithm 1 provides the optimal three-frame dual-layer factorization of the target image. (Third Row) The partial reconstructions presented during each frame. (Fourth Row) When the three frames are presented at a rate greater than the critical flicker fusion threshold, the viewer perceives a superresolved image with four times the number of pixels. (Note that, unless the backlight brightness is increased, the cascaded display will appear dimmer than a conventional display. Increasing the brightness scaling factor β also compensates for absorption losses, albeit at the cost of decreased reconstruction fidelity.)

and rearranged as a sparse matrix $\mathbf{W} \circ \mathbf{T}$ containing subpixel fragment values analogously to \mathbf{S} —into a series of time-multiplexed attenuation pattern pairs (i.e., columns of \mathbf{A} and \mathbf{B} to be displayed across the two layers). There is a simple heuristic factorization capable of *losslessly* reconstructing a spatially-superresolved target image using four time-multiplexed attenuation layer pairs (assuming that both layers have the same pixel structure and the lateral shift is half a pixel along both axes). As shown in Figure 3, during the first frame, the bottom layer depicts a pinhole grid, where only the first pixel in each 2×2 pixel block is illuminated. Each top-layer pixel is assigned the transmittance of the corresponding target subpixel fragment. Note that only one quarter of the target subpixel fragments will be reconstructed when a given pinhole grid is displayed on the bottom layer. As a result, four time-multiplexed layer pairs are required, comprising four shifted pinhole grids.

Although no artifacts are present in the reconstructed images, heuristic factorizations appear with one quarter the brightness as a conventional single-layer display, since each subpixel fragment is only visible during one of four frames. For this reason, we develop a formal optimization criterion for spatial superresolution; this allows for less than four time-multiplexed frames, or even a single frame, significantly enhancing the practicality of cascaded displays.

Optimized Compressive Factorization By application of Equation 2, we observe that optimal dual-layer factorizations are provided by solving the following constrained least-squares problem:

$$\arg \min_{\{0 \preceq \mathbf{A} \preceq \mathbf{1}, 0 \preceq \mathbf{B} \preceq \mathbf{1}\}} \frac{1}{2} \|\mathbf{W} \circ (\beta \mathbf{T} - \mathbf{A}\mathbf{B}^T)\|_2^2, \quad (3)$$

where \preceq is the elementwise matrix inequality operator. Note that the *brightness scaling factor* $0 < \beta \leq 1$ is required to allow solutions that reduce the luminance of the perceived image, relative to the target image (e.g., as observed with the heuristic four-frame factorization). If the upper bounds on \mathbf{A} and \mathbf{B} are ignored, then Equation 3 corresponds to **weighted non-negative matrix factorization (WNMF)**. As a result, any weighted NMF algorithm can be applied to achieve spatial superresolution, with the pixel values clamped to the feasible

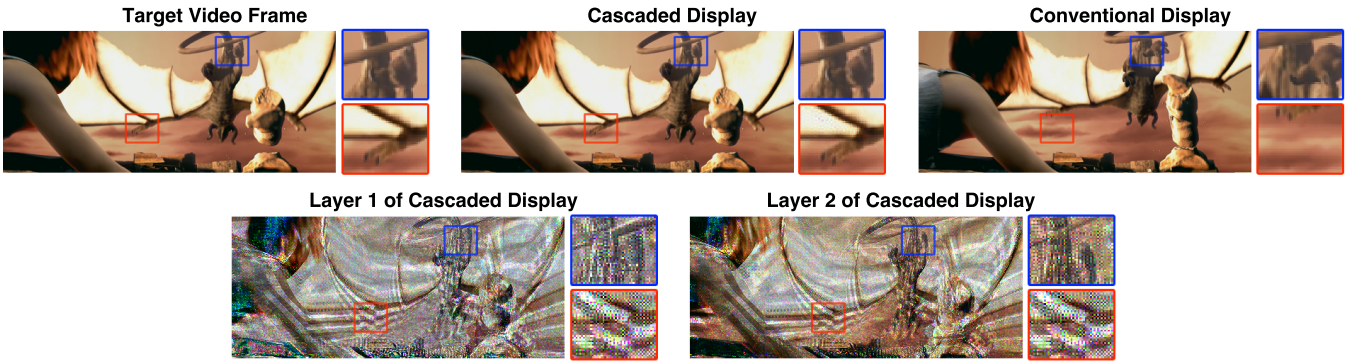


Figure 5: Temporal superresolution results using a cascaded dual-layer display. In this example, the display layers refresh in a staggered fashion and are assumed to be mechanically aligned. (Top Left) A single frame from the target video (which has twice the refresh rate as the display layers). (Top Middle) Equations 6 and 7 are used to factorize the target video. The reconstruction of the target frame shows minimal artifacts, after blurring by a uniform 2×2 -pixel spatial blur kernel. (Top Right) A conventional display refreshed at half the rate of the target video. During this frame, the conventional display lags behind the target video and cascaded display for the depicted frame. (Bottom) Note that high-frequency details are spatially averaged before being perceived by the viewer (e.g., by a diffuser or by defocusing projection optics). (Sintel movie still copyright Blender Foundation | www.sintel.org.)

range after each iteration. For instance, the following *multiplicative update rules* proposed by Blondel et al. [2008] are commonly used.

$$\mathbf{A} \leftarrow \mathbf{A} \circ \frac{(\mathbf{W} \circ (\beta \mathbf{T})) \mathbf{B}}{(\mathbf{W} \circ (\mathbf{A} \mathbf{B}^T)) \mathbf{B}} \quad \mathbf{B} \leftarrow \mathbf{B} \circ \frac{\mathbf{A}^T (\mathbf{W} \circ (\beta \mathbf{T}))}{\mathbf{A}^T (\mathbf{W} \circ (\mathbf{A} \mathbf{B}^T))} \quad (4)$$

Note that the double lines (\Leftarrow) denote Hadamard (elementwise) matrix division. Similar multiplicative update rules have been applied to prior multi-layer 3D displays, as reviewed in Section 2.3; however, the computational efficiency and convergence properties have been superseded by recent algorithms, as surveyed by Ho [2008] (which is a good starting point for a reader not already familiar with NMF). In Supplementary Appendix A.1, we evaluate the performance of these alternatives for cascaded displays, concluding that weighted rank-1 residue iterations (WRI) are more robust and efficient than existing multiplicative update rules. As adapted from Ho [2008], WRI is specified in Algorithm 1, with \mathbf{x}_j denoting column j of a matrix \mathbf{X} and $[\mathbf{x}_j]_+$ denoting projection onto the positive orthant, such that element i of $[\mathbf{x}_j]_+$ is given by $\max(0, x_{i,j})$. Figure 4 shows spatial superresolution results obtained using the WRI algorithm. In this example, the layers are initialized with uniformly-distributed random values for all frames. In comparison to the heuristic factorization, both layers contain content-dependent features.

We emphasize that Equations 2 and 3 cast image formation by dual-layer cascaded displays as a matrix factorization problem, such that the factorization rank equals the number of time-multiplexed frames. Hence, WNMF-based factorization allows a flexible trade-off between the reconstruction accuracy, the number of time-multiplexed frames, and the brightness of the reconstructed image. Section 5 provides an extended analysis of this tradespace.

Real-time Rank-1 Factorization As we shall see later on, cascaded displays achieve very high-quality results even without temporal multiplexing, which is one of their key benefits. As discussed above, eliminating temporal multiplexing is equivalent to displaying a rank-1 factorization. In Supplementary Appendix A.1, we demonstrate that WRI is the most efficient method for solving this rank-1 factorization, achieving real-time frame rates for high-definition (HD) target frames (a variant of alternating least squares for solving NMF, which we propose in Supplementary Appendix A.3, is also quite efficient). As described in Section 4, this observation is the key to enabling real-time applications (e.g., our GPU-based implementation of fast rank-1 factorization is necessary for interactive operation of the cascaded head-mounted display).

3.3 Spatiotemporal Superresolution

Cascaded displays enhance spatial resolution by layering pairs of spatially-offset, temporally-averaged displays. We observe, by analogy, that cascaded displays can also enhance temporal resolution by layering multiple temporally-offset, spatially-averaged displays. For instance, temporally offsetting two 50 Hz display panels—by staggering their refreshes by 10 milliseconds—synthesizes a 100 Hz display. Note that, for spatial superresolution, temporal multiplexing generally enhances the reconstruction fidelity (but proves optional in practice); similarly, for temporal superresolution, spatial averaging reduces reconstruction artifacts by increasing the degrees of freedom afforded by dual-layer displays with staggered refreshes. In practice, spatial averaging can be achieved by introducing a diffusing optical element on top of a flat panel cascaded display (e.g., a dual-layer LCD) or by defocusing a projector employing cascaded displays.

Similar to Equation 3, we propose the following objective function to determine optimal factorizations for temporal superresolution:

$$\arg \min_{\{0 \leq \mathbf{A} \leq 1, 0 \leq \mathbf{B} \leq 1\}} \frac{1}{2} \|\mathbf{W} \circ (\beta \mathbf{T} - \mathbf{C} \mathbf{P}_1 \mathbf{A} \mathbf{B}^T \mathbf{P}_2)\|_2^2, \quad (5)$$

Here, \mathbf{A} is a length- FN column vector, containing the bottom-layer pixel emissivities, concatenated over F video frames; similarly, \mathbf{B} is a length- FM column vector, containing the top-layer pixel transmissivities, concatenated over F video frames. The permutation matrices $\{\mathbf{P}_1, \mathbf{P}_2\}$ reorder the reconstructed subpixel fragments $\mathbf{S} = \mathbf{A} \mathbf{B}^T$ such that the first F columns of the product $\mathbf{P}_1 \mathbf{A} \mathbf{B}^T \mathbf{P}_2$ contain the length- NM subpixel fragments, corresponding to the superresolved image displayed during the corresponding frame. Spatial averaging is represented as the $FN \times FN$ convolution matrix \mathbf{C} , which low-pass filters the columns of $\mathbf{P}_1 \mathbf{A} \mathbf{B}^T \mathbf{P}_2$. Once again, \mathbf{W} is a sparse weight matrix, containing the pairwise overlaps across space and time. Finally, $\mathbf{W} \circ \mathbf{T}$ denotes the subpixel fragments for the target temporally-superresolved video. Note that we do not time-multiplex each target frame over K factorization frames, as the goal is to increase frame rate, not spatial fidelity.

We emphasize that joint spatial and temporal superresolution is directly supported by the objective function presented in Equation 5. The weight matrix \mathbf{W} subsumes temporal as well as spatial overlaps; hence, it is sufficient to set the weight matrix elements accordingly.

Unlike Equation 3, to our knowledge Equation 5 cannot be solved using any existing optimization algorithm. However, this expres-



Figure 6: Cascaded LCD screen prototype. The custom-fabricated enclosure, interface electronics, IMU, magnifying optics, and modified panels are shown on the left. The assembled prototype supports both direct viewing and head-mounted display using the lens attachment. The two photographs on the right were taken through the magnifying optics to illustrate the HMD mode (using the real-time rank-1 factorization in Section 4.1). Note the improved legibility of text using the cascaded LCD, in comparison to a conventional (low-resolution) display.

sion has a similar form to WNMF problem statements. As a result, we follow the construction presented in Chapter 6 of the thesis by Ho [2008], which provides a detailed derivation of the weighted multiplicative update rules in Equation 4. We obtain the following update rules for implementing temporal superresolution using cascaded dual-layer displays (see Supplementary Appendix A.2).

$$\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\mathbf{P}_1^T \mathbf{C}^T (\mathbf{W} \circ (\beta \mathbf{T})) \mathbf{P}_2^T \mathbf{B}}{\mathbf{P}_1^T \mathbf{C}^T (\mathbf{W} \circ (\mathbf{C} \mathbf{P}_1 \mathbf{A} \mathbf{B}^T \mathbf{P}_2)) \mathbf{P}_2^T \mathbf{B}} \quad (6)$$

$$\mathbf{B} \leftarrow \mathbf{B} \circ \frac{\mathbf{A}^T \mathbf{P}_1^T \mathbf{C}^T (\mathbf{W} \circ (\beta \mathbf{T})) \mathbf{P}_2^T}{\mathbf{A}^T \mathbf{P}_1^T \mathbf{C}^T (\mathbf{W} \circ (\mathbf{C} \mathbf{P}_1 \mathbf{A} \mathbf{B}^T \mathbf{P}_2)) \mathbf{P}_2^T} \quad (7)$$

For simplicity, we specify multiplicative update rules for spatiotemporal superresolution; however, as discussed in Supplementary Appendix A.2, the WRRI algorithm can be similarly adapted. In practice, we never construct the matrices $\{\mathbf{C}, \mathbf{P}_1, \mathbf{P}_2\}$. Instead, given an implementation for the update rules of Equation 4, between the iterations we apply a spatial blur to the current estimate $\mathbf{A} \mathbf{B}^T$. Reconstruction results are shown in Figure 5. All layers and frames are initialized to uniformly-distributed random values. The entire video is factorized simultaneously. For longer videos, a sliding window of frames can be factorized, constraining the first frames in each window to equal the last frames in the previous window [Heide et al. 2013]. As demonstrated in Figure 5, a uniform 2×2 blur kernel proves sufficient; however, as with rank-1 spatial superresolution, Equations 6 and 7 support spatiotemporal superresolution without any optical blurring, albeit with the introduction of reconstruction artifacts. Spatiotemporal superresolution performance, for several sample videos, is demonstrated in the supplementary video.

3.4 Multi-Layer Superresolution

So far, we have only described dual-layer cascaded displays. We now generalize to multiple layers. Without loss of generality, we assume that all layers have identical square pixels. Given L layers, we propose offsetting each layer by $1/L$ pixels with respect to the previous layer. The resulting cascaded display then has L^2 times as many subpixel fragments as any single layer. The factorization proceeds in a similar manner; however, the matrices are now tensors and a weighted non-negative tensor factorization (WNMF) is required. Specific update rules and simulated multi-layer, multi-frame factorizations are reported in Supplementary Appendix C.

4 Implementation

This section describes our software implementation (Section 4.1) and the construction of three hardware prototypes: a dual-layer LCD (Section 4.2), a projector containing a pair of LCoS microdisplays (Section 4.3), and multi-layer stacks of printed films (Section 4.4).

4.1 Software

Factorization Algorithms The multiplicative update rules (Equation 4) and the WRRI method (Algorithm 1) were implemented in Matlab for spatial superresolution with dual-layer displays. Only the multiplicative update rules were programmed for more than three layers and for temporal superresolution. In all cases, arbitrary numbers of frames (i.e., factorization ranks) are supported. The fast rank-1 solver was implemented using CUDA to leverage GPU acceleration (source code is provided in Supplementary Appendix B). All factorizations were performed on an Intel 3.2 GHz Intel Core i7 workstation with 8 GB of RAM and an NVIDIA Quadro K5000. We report that the fast rank-1 solver maintains the native 60 Hz refresh rate, including overhead for rendering scenes and applying post-processing fragment shaders (e.g., in the HMD demonstration). Run times are tabulated in Supplementary Appendix A.1.

Calibration Tools Operation of cascaded displays requires knowledge of the physical configuration of the display layers and their radiometric characteristics (i.e., to compute the pixel overlaps encoded in \mathbf{W} in Equation 2). While we seek precise layer offsets, this can only be approximated in our prototypes. Misalignment is corrected by warping the image displayed on the second layer to align with the image displayed on the first layer. Two photographs are required to estimate this warp. In each photograph, a checkerboard is displayed on one layer, while the remaining layer is set to be fully transparent or fully reflective. Scattered data interpolation estimates the warping function that projects photographed first-layer checkerboard corners into the coordinate system of the image displayed on the second layer. The second-layer checkerboard (or any other image) is warped to align with the first-layer checkerboard. In addition, radiometric characteristics are measured by photographing flat field images; these curves are inverted such that each display is operated in a linear radiometric fashion. We emphasize that the geometric and radiometric calibration is used to rectify the captured images and correct vignetting—allowing direct comparison to predicted results.

4.2 Cascaded LCD

Construction As shown in Figure 6 and in the supplementary video, we modified a pair of off-the-shelf LCD panels to create a mobile-form-factor cascaded LCD prototype—choosing smaller panels in order to demonstrate both direct-view and head-mounted display (HMD) applications. Mobile displays with HDMI inputs are relatively uncommon. As a result, our prototype is built using a pair of 7-inch HannStar HSD070PWW1-B00 LCD panels and accompanying interface boards taken from the “HDMI 4 Pi” product sold by Adafruit Industries (originally intended for use with the Raspberry Pi single-board computer). We operate each panel at the native resolution of 1280×800 pixels and with a 60 Hz refresh rate.

A custom enclosure was fabricated using a Dimension 1200es 3D printer, comprising a base plate, a rear case, a top frame, and a lens attachment for HMD use (see Figure 6). The interface boards were screwed to the bottom of the base plate and an unmodified LCD was affixed to the top. Mobile displays are highly integrated; careful disassembly was required to separate the backlight and brightness enhancing films from the second LCD. The bare liquid crystal panel was affixed to the base plate, held in direct contact with the first LCD at a fixed lateral offset. As assembled, the front polarizer on the bottom LCD is crossed with the rear polarizer on the top LCD. Rather than remove the polarizers from the thin glass panels, we placed quarter-wave retarder film between the two (American Polarizers APQW92-004-PC-140NMHE): rotating the polarization state to restore the operation of the top LCD.

Direct-View Results As shown in Figure 6, the cascaded LCD supports direct viewing from a distance, as with a mobile phone or tablet computer. All spatial superresolution results, for this prototype and the others, were captured using a Canon EOS 7D camera with a 50 mm f/1.8 lens. Temporal superresolution results, included in the supplementary video, use a Point Grey Flea3 camera with a Fujinon 2.8–8 mm varifocal lens. Due to the gap between the LCD modulation layers, the lateral offset will appear to shift depending on viewer location. The calibration procedure in Section 4.1 compensates for this parallax (with a tripod-mounted camera placed directly in front of the prototype). (Section 6 discusses manufacturing solutions for minimizing this gap.) The full set of results, captured for a variety of target images including natural images, text, and resolution test patterns, can be viewed using the supplementary data archive.

Similar to Sajadi et al. [2012], we display layer patterns at lower resolution than the native panel resolution, allowing direct comparison to “ground truth” superresolved images. For results in the supplementary data archive, we clarify that sets labeled with heights of 300 or 600 pixels correspond, respectively, to treating each 4×4 or 2×2 pixel block of the LCDs as a single addressable pixel. As discussed in Section 6.1, using larger pixel blocks has the added benefit of mitigating artifacts due to the color filter arrays (CFAs).

Head-Mounted Display Results The cascaded LCD converts into a head-mounted display by attaching the lens assembly. Similar to the design of Olson et al. [2011] and the Oculus Rift, this assembly holds a pair of aspheric magnifying lenses (ElectroOptix, Inc. AHM-5). The lenses are separated from the top LCD by slightly less than their 5.1 cm focal length in order to synthesize a magnified, erect virtual image appearing near “optical infinity”. Head tracking is supported through the use of an inertial measurement unit (IMU) attached to the rear case (Hillcrest Laboratories FSM-9). The supplementary video includes experimental results recorded through the HMD optics. In all HMD results we use the GPU-accelerated fast WRRI solver described in Section 4.1; as demonstrated, this implementation is able to maintain the native 60 Hz refresh, including the time required to render the OpenGL scene, apply a GLSL fragment shader to warp the imagery to compensate for spherical and chromatic aberrations [Watson and Hodges 1995], and to factorize the resulting target image. Unlike direct viewing, an HMD allows a limited range of viewing angles—reducing the influence of viewer parallax and facilitating practical applications of cascaded LCDs.

4.3 Cascaded LCoS Projector

Construction While LCDs dominate consumer applications, superresolution may also prove an attractive option to meet 8K UHD cinematic projection standards. To verify this use case, we also constructed a cascaded LCoS projector. A pair of AAXA P3 pico projectors were disassembled, providing an affordable source for two Syndiant SYL2061 LCoS microdisplays, interface electronics,

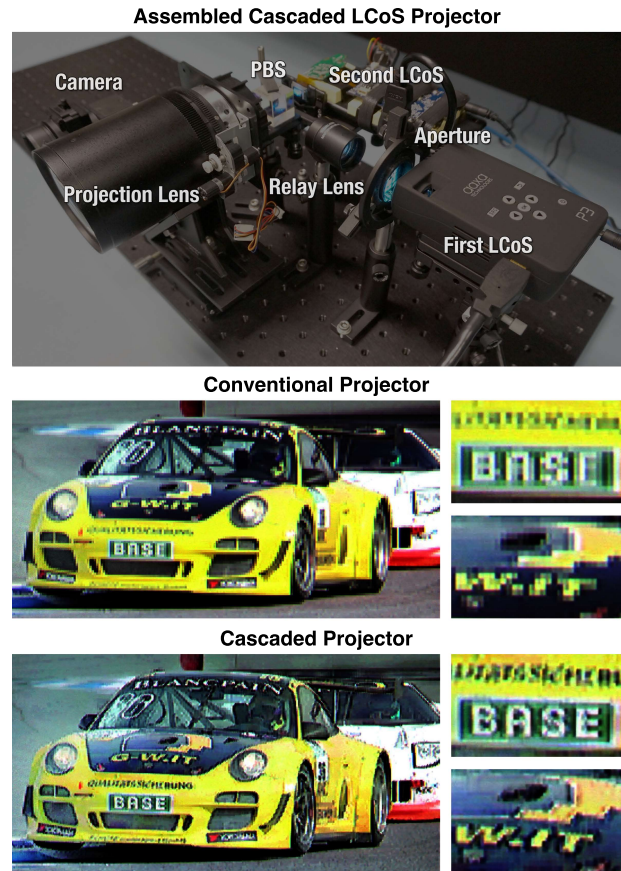


Figure 7: Cascaded LCoS projector. (Top) The assembled prototype. (Middle) Photograph of an image projected using a conventional (low-resolution) LCoS projector. (Bottom) Photograph of the cascaded LCoS projector result. Note the improved legibility of text. (Motorsport image courtesy Wikipedia user “KaPhiMoritz”.)

and a field sequential color (FSC) LED illumination engine. These displays were operated at their native resolution of 1024×600 pixels and at a refresh rate of 60 Hz. (The microdisplay specifications indicate that LCoS is particularly well-suited for practical cascaded displays: with an aperture ratio of 95.8% and reflectivity of 70%.)

As shown in Figure 7 and in the supplementary video, the projection lens was removed from the first projector, which was then mounted to an optical breadboard. The case, battery, polarizing beam splitter (PBS) plate, and illumination engine were removed from the second projector, exposing a bare LCoS panel and interface board, which were also mounted to the breadboard. A relay lens (Thor Labs MAP10100100-A) was used to achieve dual modulation—projecting the image of the first LCoS onto the second with unit magnification. A PBS cube (Thor Labs PBS251) was positioned between the relay lens and second LCoS, replacing the original PBS plate. The dual-modulated image was projected onto a screen surface using projection optics repurposed from a larger Sanyo PLC-XP18N projector (since the original projection lens has an insufficient back focal distance to accommodate the PBS cube).

Our cascaded LCoS design is similar to that described by Kusakabe et al. [2008; 2009]; however, we advance that off-axis positioning of the LCoS panels is necessary to prevent multiple reflections. If the two LCoS panels are perpendicular to, and centered along, the optical axis of the relay lens, then light can be reflected back to the first LCoS from the PBS cube, leading to experimentally-observed aberrations. Laterally shifting the LCoS panels away from the optical axis eliminates these artifacts. As shown in Figure 7, an aperture is

placed in front of the first LCoS to prevent any reflected light—now offset from the optical axis—from continuing to propagate.

Results Experimental results are summarized in the supplementary video and data archive. Artifacts are primarily due to color channel crosstalk: the two LCoS timing controllers are unsynchronized. As a result, the active LED, synchronized with a certain channel of the first LCoS image, may illuminate different color channels in second LCoS image. For commercial implementations, timing issues could be resolved with purpose-built components.

4.4 Cascaded Printed Films

We also experimented with a simpler architecture: printed semi-transparent color films. These films can be reproduced using the patterns provided with the supplementary material. Only single-frame (i.e., rank-1) factorizations can be presented with static films.

5 Analysis

In this section we quantitatively analyze cascaded display performance, both in simulation and by experiment. Section 5.1 assesses design trade-offs. Section 5.2 compares the performance of our system with prior superresolution displays.

5.1 Performance of Cascaded Display Factorizations

Spatial Superresolution Tradespace Solutions of Equation 3 offer a display designer a flexible trade-off between apparent image brightness, spatial resolution, and refresh rate, as captured by the dimming factor β , the resolution of the target image $\mathbf{W} \circ \mathbf{T}$, and the factorization rank K , respectively. The peak signal-to-noise ratio (PSNR) was recorded as these parameters were varied (averaged over the set of target images in the supplementary data archive). As plotted in Figure 9, high-PSNR reconstructions are obtained with a dimming factor of 0.25 and four frames—as to be expected, since the heuristic factorization in Section 3 exactly reconstructs the target image in this case. Note that three-frame factorizations closely approach the performance achieved with four frames. Most significantly, Figure 3 reveals a key insight: spatial superresolution—with a PSNR exceeding 30 dB—can be achieved at the native display refresh rate, without reducing the apparent brightness.

Temporal Superresolution Tradespace Solutions of Equation 5 also offer a flexible trade-off between brightness, resolution, and refresh rate; however, unlike spatial superresolution designs, architectures intended for spatiotemporal superresolution may include an optical blurring element (characterized by the point spread function embedded in the convolution matrix \mathbf{C}). As assessed in Supplementary Appendix E, factorizations with 2×2 -pixel uniform blur kernels prove sufficient to render high-PSNR reconstructions for a variety of target videos; however, as with the limited degrees of freedom afforded with low-rank factorizations for spatial superresolution, we experimentally observe that certain videos can be effectively superresolved without added blur. As a result, experiments in the supplementary video do not incorporate diffusing elements.

5.2 Comparison of Superresolution Display Systems

Display Alternatives Figure 10 compares image patches depicted with various superresolution displays. Table 2 provides quantitative comparisons using PSNR and structural similarity (SSIM) index [Wang et al. 2004] measurements. Target images are spatially superresolved by a factor of four (i.e., twice as many pixels along each axis). Following Section 2, we subsume several prior works under a general *additive superresolution display* model: presenting

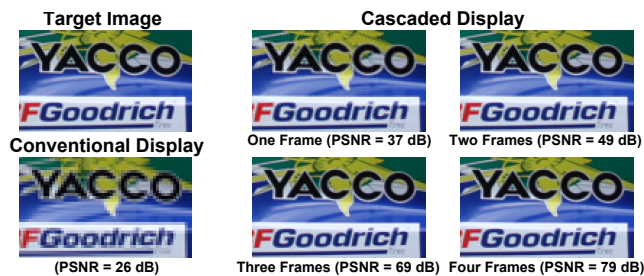


Figure 8: Spatial superresolution with varying numbers of frames. The target image from Figure 4 was factorized using the WRRI algorithm. High-PSNR reconstruction is achieved with fewer than four frames: the single-frame decomposition significantly enhances spatial resolution in comparison to a conventional single-layer display.

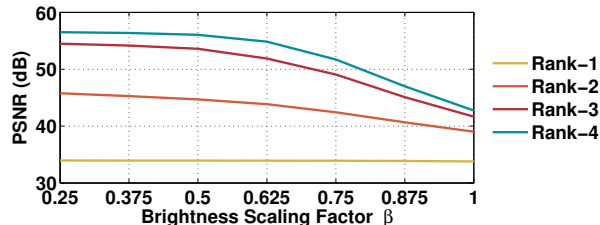


Figure 9: Tradespace analysis of spatial superresolution with respect to factorization rank and the brightness scaling factor β .

a set of superimposed, shifted low-resolution images (e.g., vibrating displays and superimposed projections). We further assume ideal conditions such that no motion blur is introduced, which would further degrade image quality for vibrating displays.

Optical pixel sharing [Sajadi et al. 2012] is simulated using source code provided by the original authors. This implementation requires specifying two tuning parameters: the edge threshold and the smoothing coefficient. We use grid search to optimize these parameters—independently for each target image—to maximize the PSNR or the SSIM index. Optimizing the smoothing coefficient was not explored in the original publication; while this two-dimensional search is computationally intensive (on the order of minutes per image), it maximizes OPS performance in the reported comparisons. In practice, ensemble-averaged tuning parameters must be used, increasing reconstruction artifacts (see Supplementary Appendix E). We emphasize that cascaded displays do not require optimizing any such tuning parameters, further facilitating real-time applications.

The spatial light modulators used in each of these display alternatives may have variable pixel aperture ratios. As previously observed, limited aperture ratios translate to improved image quality for additive superresolution displays; yet, as evaluated by Majumder [2005] and Sajadi et al. [2012], spatial superresolution from additive superpositions is practically hindered due to the engineering challenges associated with limiting aperture ratios—particularly for superimposed projections. Furthermore, industry trends are pushing ever-higher aperture ratios (e.g., LCoS microdisplays and power-efficient LCDs). As a result, we assume a 100% aperture ratio in all comparisons.

Several observations can be made from the visual comparisons and PSNR table. Foremost, we find that, for these examples, single-frame cascaded display factorizations closely approach or outperform all other methods utilizing two time-multiplexed frames. These PSNR advantages translate to visible reductions in artifacts: notice the enhancement of the text in Figure 10. We also find that cascaded displays significantly outperform optical pixel sharing (OPS), which relies on a similar dual-modulation architecture containing relay optics. Simulations of additive superresolution also appear to outperform OPS, despite Sajadi et al. [2012] finding the opposite (for vibrating displays). This discrepancy is due to the fact that



Figure 10: Visual comparison of superresolution displays. Image patches are reproduced with simulations of three different superresolution displays: additive superresolution using two frames, optical pixel sharing (OPS) using two frames with per-image PSNR- and SSIM-optimized edge thresholds and smoothing coefficients (see Section 5.2), and cascaded displays using one or two frames. Notice the enhancement relative to a conventional (low-resolution) display. Consult the supplementary data archive for the full set of comparisons. (Source images courtesy Yoshikazu Hara (first and second rows), Wen-Yan King (third row), JJ Harrison (fourth row), and Wilfredo R. Rodriguez H. (fifth row).)

we assume no motion blur in our additive simulations; in practice, OPS should achieve superior image quality. Finally, we find that two-frame cascaded display factorizations outperform all other two-frame factorizations by a significant margin and even four-frame additive superresolution. This highlights the benefits of the *compressive* capabilities enabled by our matrix-factorization-based approach.

Modulation Transfer Function (MTF) Analysis This section expands on our PSNR analysis by comparing the modulation transfer functions (MTFs) characterizing each superresolution display alternative: specifying the contrast of spatially-superresolved images, as a function of spatial frequency. The MTF of a display can be measured using a variety of test patterns, including natural image sets, spatial frequency chirps, and slanted edges. We adopt a chirped *zone plate* pattern given by $(1 + \cos(cr^2))/2$, where $r = \sqrt{x^2 + y^2}$, $\{x, y\} \in [-\pi, \pi]$, and c controls the maximum spatial frequency.

Figure 11 compares the (theoretical) MTFs exhibited by each display alternative. MTF analysis confirms the earlier observations made regarding the relative performance of each approach; furthermore, it reveals that single-frame cascaded displays effectively quadruple the spatial resolution (doubling it along each image dimension)—albeit with artifacts introduced by compression—maintaining greater than

70% contrast for the highest superresolved frequencies. Reflecting the conclusions drawn from Figures 8 and 9, Figure 11 also shows that the MTFs for two-frame and three-frame factorizations are nearly identical, indicating that most practical applications of cascaded display would require no more than a pair of time-multiplexed frames. Figure 12 shows the measured MTF from our cascaded LCD prototype for 1 and 2 frame factorizations. While the MTF is lower than predicted in simulation, it offers a clear improvement over a conventional display. Consult Supplementary Appendix E for further MTF evaluation using natural images and slanted edges.

6 Discussion

6.1 Addressing Hardware Limitations

Cascaded displays require all but one layer to contain light-attenuating SLMs; hence, for thin form factors, this leaves LCDs as the only widely-available SLM alternative. LCDs are a mature technology, however their modern implementation poses additional challenges for cascaded displays, including: fixed color filter array (CFA) designs, limited pixel aperture ratios, and parallax due to a physical gap between display layers. We address each issue in turn.

	conventional	additive (two frames)	additive (four frames)	OPS	cascaded (one frame)	cascaded (two frames)	cascaded (four frames)
"motorcycle" (dB)	27.42	33.12	38.52	38.96	33.78	45.51	74.97
"drift" (dB)	28.06	36.44	44.98	36.51	36.61	51.28	87.37
all images (dB)	27.71	33.64	41.55	35.85	34.01	47.45	82.32
"motorcycle" (SSIM)	2.641	2.924	2.985	2.869	2.934	2.997	3.000
"drift" (SSIM)	2.637	2.935	2.995	2.857	2.938	2.998	3.000
all images (SSIM)	2.580	2.886	2.983	2.850	2.893	2.995	3.000

Table 2: Peak signal-to-noise (PSNR) and structural similarity (SSIM) index (sum over all three color channels) for two target images. Similar to Figure 10, three alternatives are compared: additive superresolution displays using either two or four frames, optical pixel sharing (OPS) using two frames, and cascaded displays using one, two, or four frames. See Supplementary Appendix E for extended PSNR and SSIM evaluations. (Motorcycle and motorsport images courtesy Yoshikazu Hara and Aurélien Vialatte, respectively.)

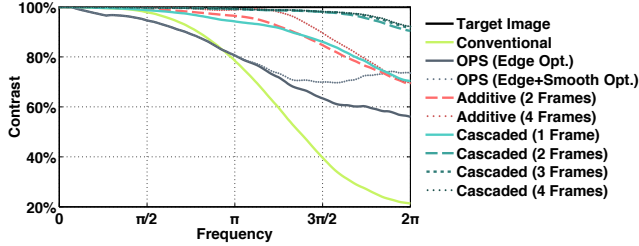


Figure 11: Simulated comparison of the MTF for display alternatives. Single-frame cascaded displays effectively quadruple spatial resolution and perform on par with two-frame additive displays.

Customizing Color Filter Arrays (CFAs) LCDs primarily achieve color display using CFAs, typically comprising a periodic array of vertically-aligned red, green, and blue bandpass filters. As a result, cascaded LCDs can double the vertical resolution by vertically offsetting the display layers; however, increasing the horizontal resolution proves problematic without modifying the CFA. One solution is to substitute a single, custom CFA that individually filters the subpixel fragments, rather than the (larger) layer pixels. Another solution is to replace the usual red-green-blue CFA with a cyan-yellow-magenta CFA. The superposition of a pair of such CFAs, with the appropriate lateral displacement, yields superresolved red, green, and blue subpixel fragments (see Supplementary Appendix D). We emphasize, however, that emerging high-speed LCDs may eliminate the need for CFAs, instead using field sequential color (FSC) illumination, similar to the cascaded LCoS projector.

Increasing Pixel Aperture Ratios The pixel aperture ratio quantifies the portion of an LCD pixel that actively shutters light. The remaining opaque area contains thin-film transistors and control elements. As illustrated in Figure 1, cascaded displays benefit from high aperture ratios, since the superposition of offset pixel layers reduces the effective aperture ratio. Industry developments are fostering higher aperture ratios primarily to reduce power consumption. As a result, LCD technologies that increase the aperture ratio are well-suited for cascaded displays, including LTPS (low-temperature polysilicon) and IGZO (indium-gallium-zinc-oxide). As an alternative to modifying LCD pixel structures, we observe that a tailored diffusing film may be placed on top of an LCD to increase the effective pixel aperture ratio. Woo et al. [2011] describe a similar application to OLED panels: demonstrating that light emitted by a pixel spreads into neighboring opaque areas by subsurface scattering.

Minimizing Layer Separation Layering a pair of LCDs may increase the thickness by less than a millimeter (as in our prototype); however, any separation between the layers will manifest as parallax: the composition of the subpixel fragments will then depend on the viewer’s perspective. This limitation can be addressed by manufacturing dual-layer LCDs as a monolithic unit, rather than as separate panels bonded together after fabrication. In such a configuration, redundant elements can be eliminated; foremost, only two polarizers are required: one on each side of the dual-layer LCD, improving

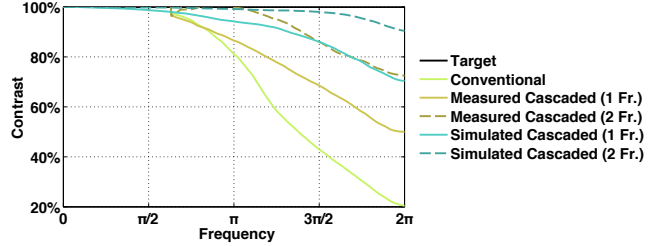


Figure 12: The measured modulation transfer function for our cascaded LCD prototype. The prototype achieves clear superresolution when compared to a conventional display.

image brightness [Wetzstein et al. 2012]. Furthermore, the protective glass alignment layer between the first and second liquid crystal cells can be significantly reduced in thickness, since rigidity of the panel is provided by the outer protective glass layers. We reiterate that HMD applications greatly restrict the range of viewing angles and, as a result, are well-suited for adopting cascaded displays.

6.2 Future Work

The prototypes demonstrate practical applications with LCD and LCoS technologies. A promising avenue for future work is to explore the use of other spatial light modulators, particularly digital micromirror devices (DMDs) for cascaded projectors. Head-mounted displays are beginning to incorporate organic light-emitting diode (OLED) displays to reduce persistence and thereby ameliorate motion blur artifacts. Since cascaded LCDs do not restrict the form of the first SLM, a hybrid design consisting of an LCD-modulated OLED panel may be well-suited for practical HMD applications—limiting persistence, while enabling spatial superresolution.

7 Conclusion

The display industry is rapidly advancing the resolution of existing display technologies to meet the demands of 4K UHD standards. Similarly, mobile displays have recently surpassed the resolution of the human eye at typical viewing distances. However, the need for superresolution displays persists, particularly for emerging head-mounted displays and glasses-free 3D displays. Similarly, 8K UHD standards for cinematic projection will necessitate even higher pixel counts. Cascaded displays present a new approach to meet these demands using today’s display technologies, including LCD panels and LCoS microdisplays. As demonstrated with our prototypes, cascaded dual-layer displays can quadruple the apparent spatial resolution and double the effective refresh rate. Extensions to more than two layers afford even greater increases, albeit with added engineering challenges. Most significantly, we have developed the real-time factorization algorithms necessary to drive cascaded displays—offering practical means to accelerate the development of emerging display applications, supported by a flexible trade-off between spatial resolution, refresh rate, and apparent image brightness.

Acknowledgements

We thank the reviewers for their helpful feedback. We also thank Mushfiqur Rouf for first proposing the application of cascaded display concepts for temporal superresolution with staggered refreshes.

References

- ALIAGA, D. G., YEUNG, Y. H., LAW, A., SAJADI, B., AND MAJUMDER, A. 2012. Fast high-resolution appearance editing using superimposed projections. *ACM Trans. Graph.* 31, 2.
- ALLEN, W., AND ULICHNEY, R. 2005. Wobulation: Doubling the addressed resolution of projection displays. *SID Symposium Digest of Technical Papers* 36, 1, 1514–1517.
- BAKER, S., AND KANADE, T. 2002. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 9, 1167–1183.
- BELL, G. P., CRAIG, R., PAXTON, R., WONG, G., AND GALBRAITH, D. 2008. Beyond flat panels: Multi layer displays with real depth. *SID Digest* 39, 1, 352–355.
- BEN-EZRA, M., ZOMET, A., AND NAYAR, S. K. 2004. Jitter camera: High resolution video from a low resolution detector. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, 135–142.
- BERTHOUSOZ, F., AND FATTAL, R. 2012. Apparent resolution enhancement for motion videos. In *ACM Applied Perception*.
- BERTHOUSOZ, F., AND FATTAL, R. 2012. Resolution enhancement by vibrating displays. *ACM Trans. Graph.* 31, 2, 15:1–15:14.
- 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.
- DAMERA-VENKATA, N., AND CHANG, N. L. 2009. Display supersampling. *ACM Trans. Graph.* 28, 1, 9:1–9:19.
- DIDYK, P., EISEMANN, E., RITSCHER, T., MYRSKOWSKI, K., AND SEIDEL, H.-P. 2010. Apparent display resolution enhancement for moving images. *ACM Trans. Graph.* 29, 4, 113:1–113:8.
- 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.
- HART, W. M. 1987. The temporal responsiveness of vision. In *Adler's Physiology of the Eye*, R. A. Moses and W. M. Hart, Eds. C.V. Moseby Company.
- HEIDE, F., WETZSTEIN, G., RASKAR, R., AND HEIDRICH, W. 2013. Adaptive image synthesis for compressive displays. *ACM Trans. Graph.* 32, 4, 132:1–132:12.
- HEIDE, F., GREGSON, J., WETZSTEIN, G., RASKAR, R., AND HEIDRICH, W. 2014. A compressive multi-mode superresolution display. Technical Report *arXiv:1404.5916*.
- HO, N.-D. 2008. *Nonnegative Matrix Factorization Algorithms and Applications*. PhD thesis, Université catholique de Louvain.
- IVES, F. E., 1903. Parallax stereogram and process of making same. U.S. Patent 725,567.
- JAYNES, C., AND RAMAKRISHNAN, D. 2003. Super-resolution composition in multi-projector displays. In *IEEE Projector-Camera Systems (PROCAMS)*.
- KUSAKABE, Y., KANAZAWA, M., NOJIRI, Y., FURUYA, M., AND YOSHIMURA, M. 2008. A YC-separation-type projector: High dynamic range with double modulation. *Journal of the Society for Information Display* 16, 2, 383–391.
- KUSAKABE, Y., KANAZAWA, M., NOJIRI, Y., FURUYA, M., AND YOSHIMURA, M. 2009. A high-dynamic-range and high-resolution projector with dual modulation. *Proc. SPIE* 7241.
- 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.* 29, 6, 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.* 30, 6, 186:1–186:10.
- LIPPMANN, G. 1908. Épreuves réversibles donnant la sensation du relief. *Journal of Physics* 7, 4, 821–825.
- MAJUMDER, A. 2005. Is spatial super-resolution feasible using overlapping projectors? In *IEEE Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4.
- OLSON, J., KRUM, D., SUMA, E., AND BOLAS, M. 2011. A design for a smartphone-based head mounted display. In *IEEE Virtual Reality*, 233–234.
- PAVLOVYCH, A., AND STUERZLINGER, W. 2005. A high-dynamic range projection system. *Proc. SPIE* 5969.
- SAJADI, B., GOPI, M., AND MAJUMDER, A. 2012. Edge-guided resolution enhancement in projectors via optical pixel sharing. *ACM Trans. Graph.* 31, 4, 79:1–79:122.
- SAJADI, B., QOC-LAI, D., IHLER, A., GOPI, M., AND MAJUMDER, A. 2013. Image enhancement in projectors via optical pixel shift and overlay. In *IEEE International Conference on Computational Photography (ICCP)*, 1–10.
- SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. 2004. High dynamic range display systems. *ACM Trans. Graph.* 23, 3, 760–768.
- WANG, Z., BOVIK, A., SHEIKH, H., AND SIMONCELLI, E. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4, 600–612.
- WATSON, B. A., AND HODGES, L. F. 1995. Using texture maps to correct for optical distortion in head-mounted displays. In *IEEE Virtual Reality Annual International Symposium*, 172–178.
- WETZSTEIN, G., LANMAN, D., HEIDRICH, W., AND RASKAR, R. 2011. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.* 30, 4, 95:1–95:12.
- WETZSTEIN, G., LANMAN, D., HIRSCH, M., AND RASKAR, R. 2012. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.* 31, 4, 80:1–80:11.
- WOO, S., KIM, D.-H., HAN, Y. S., AND CHOI, B.-D. 2011. Full-color LCD microdisplay system based on OLED backlight unit and field-sequential color driving method. *Journal of Photoenergy*.