

# Mixed-primary Factorization for Dual-frame Computational Displays

FU-CHUNG HUANG, NVIDIA

DAWID PAJĄK, Light

JONGHYUN KIM, JAN KAUTZ, and DAVID LUEBKE, NVIDIA

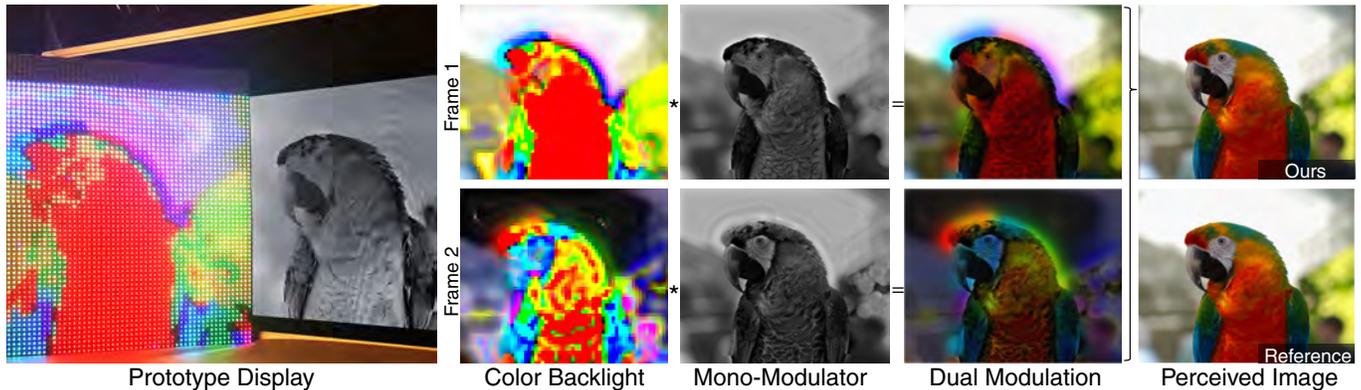


Fig. 1. The proposed display architecture. Our prototype uses a monochromatic LCD to attenuate a low-resolution color backlight generated by an LED array. Our two-frame factorization defines the color image as a pair of products of a low resolution color backlight with a high-resolution monochrome modulator. The final image (top right, simulated) is integrated through temporal (our prototype) or spatial multiplexing. Image courtesy of Derrick Coetzee.

Increasing resolution and dynamic range of digital color displays is challenging with designs confined by cost and power specifications. This necessitates modern displays to trade-off spatial and temporal resolution for color reproduction capability. In this work we explore the idea of joint hardware and algorithm design to balance such trade-offs. We introduce a system that uses content-adaptive and compressive factorizations to reproduce colors. Each target frame is factorized into two products of high-resolution monochromatic and low-resolution color images, which then get integrated through temporal or spatial multiplexing. As our framework minimizes the error in colorimetric space, the perceived color rendition is high, and thanks to GPU acceleration, the results are generated in real-time. We evaluate our system with a LCD prototype that uses LED backlight array and temporal multiplexing to reproduce color images. Our approach enables high effective resolution and dynamic range without increasing power consumption. We also demonstrate low-cost extensions to hyperspectral and light-field imaging, which are possible due to compressive nature of our system.

CCS Concepts: • **Hardware** → **Displays and imagers**;

Additional Key Words and Phrases: color optimization, LCD, LED array

## ACM Reference format:

Fu-Chung Huang, Dawid Pająk, Jonghyun Kim, Jan Kautz, and David Luebke. 2017. Mixed-primary Factorization for Dual-frame Computational Displays. *ACM Trans. Graph.* 36, 4, Article 1 (July 2017), 13 pages.

DOI: <http://dx.doi.org/10.1145/3072959.3073654>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2017 ACM. 0730-0301/2017/7-ART1 \$15.00

DOI: <http://dx.doi.org/10.1145/3072959.3073654>

## 1 INTRODUCTION

Modern display design revolves around two sets of conflicting goals. On the one hand, there is a strong drive towards higher spatial resolution, dynamic range and color fidelity. This is evident in current high-definition television sets which reproduce images at 4K resolution (3840x2160) with high contrast and maximum brightness exceeding  $500\text{cd}/\text{m}^2$ . On the other hand, displays need to be energy efficient, low-cost and long-lasting. Systematic updates to display digital interfaces, such as VESA DisplayPort (DP), provide insight into trade-offs the industry made to satisfy the above goals. One example is Digital Stream Compression (DSC) in DP1.4. To reduce cost and power consumption of bandwidth intensive 8K (7680x4320) displays, VESA introduced a mandatory lossy compression directly in the display link. Lossy compression degrades the image quality and addresses only the transmission part of the problem: after decompression the actual data bandwidth required to drive the hardware electronics – writing the exact value of every pixel every frame – is still huge.

Most of the existing display technologies compromise some of the above goals to excel in others. The LCD design, for example, with its simple construction – a combination of a uniform backlight and a color filter array – yields low-cost at the expense of reduced spatial resolution and light efficiency. About two thirds of the light is absorbed by the color filters alone, which for HDR displays translates to a waste of hundreds of Watts. Alternative display designs have been proposed, but they all come with their own trade-offs. Field-sequential color displays are more power-efficient, but require very fast driving rates [Mori et al. 1999] which increases their cost and complexity. More recent OLED displays cannot produce high luminosity without risking early burn-out [Tsujiura 2012].

In this work, we explore the idea of joint design of display algorithms and optical hardware in order to balance their performance, energy efficiency and costs. Inspired by recent advances in the computational display community [Masia et al. 2013], e.g. HDR Displays [Seetzen et al. 2004], Cascaded Displays [Heide et al. 2014], Tensor Displays [Wetzstein et al. 2012], and Multi-primary projectors [Kauvar et al. 2015], we propose a perceptually-motivated factorization that decomposes the target image into two pairs of low-resolution chromatic and high-resolution achromatic images (see Fig. 1). More importantly, the proposed decomposition combined with off-the-shelf display hardware shows promise for possible real-world applications and strikes a balance between hardware complexity and performance. In particular, our technical contributions are the following:

- We show that natural images can be efficiently factorized into a pair of modulations of low-resolution color backlight and high-resolution monochromatic attenuation.
- We develop a content-adaptive two-frame factorization framework that supports multiple compression schemes and non-linear error metrics.
- We implement the framework on the GPU and demonstrate that the linear factorization runs in real-time, while the highly non-linear perceptual variant is 2-3 orders of magnitude faster than the previous work.
- We construct a low-cost LCD display prototype using off-the-shelf components. The display design supports multi-spectral reproduction on a flat-panel form factor. The monochromatic modulator increases light efficiency and allows for lower display power consumption. Additionally, we show extensions to light-field and high dynamic range display applications.

## 2 RELATED WORK

All color displays use a limited set of primary colors (*primaries*) to reproduce a full-color image. The most common approaches rely on the principles of additive color mixtures [Silverstein 2005], where colors are a result of optical super-position or spatial- and/or temporal- multiplexing of three primaries. Systems based on optical superposition produce the highest image quality, but their size, relative complexity and high cost inhibit widespread use. Cost and complexity can be reduced by interleaving primary color samples spatially [Bayer 1976; Schroeder 1948], which, due to the low-pass characteristics of our visual system, combine in the retina to form a full-color percept. This idea defined all modern CRT/LCD/OLED displays, and while elegant, comes with a significant trade-off – using pixels for color synthesis reduces spatial resolution. Moreover, due to varying eye sensitivity to color primaries, the color pixel mosaic produces a fixed-noise pattern, which is particularly disturbing in applications requiring high pixel density, such as head-mounted displays (HMD). Field-sequential color (FSC) displays [Frankenstein and von Jaworski 1904; Polumordvinov 1899] mix primaries in time and therefore avoid artifacts inherent to spatial interleaving. However, showing color frames sequentially introduces new problems. Because individual frames have significantly different luminance values and our visual system is sensitive to luminance change over time,

displaying those frames at speeds below critical flicker frequency produces visible flicker [Simonson and Brozek 1952]. Another problem occurs when the viewer’s retina is in motion with respect to the display. The time-varying color components do not project onto the same locations in the retina, which for high-contrast features results in *color break-up* or *rainbow* artifacts. To prevent those, RGB FSC displays need to run at 360–480 frames-per-second, which imposes high bandwidth requirements and complicates the hardware design.

*Displays using two color primaries.* In 1906, George Albert Smith invented a motion picture process, *Kinemacolor*, that used a projector equipped with red and green rotating color filters. This prototypical field-sequential display was capable of color reproduction, but due to severe color break-up artifacts (projection at 32FPS) and poor image quality, the system enjoyed limited commercial success [Kindem 1981]. Several decades later Land [1959] rediscovered the potential of the approach in his famous two-color projection experiments, that produced a full range of color appearances despite not having the three primaries required by conventional colorimetric theory. Instead of superimposing red, green and blue projections of the image, he suggests to only combine red and green channels, but with green projected as white light. The result, while not as saturated as the three-color projection, is colorful and apparently contains the (missing) blue. This phenomena can be explained by simultaneous contrast and chromatic adaptation; its strength depends on the image content, but can go up if memory color is applied [Fairchild 2013].

*Content-adaptive color primaries.* The dependency between color-reproduction quality and input data suggests selecting color primaries on a per-frame basis. The resulting color gamut encompasses the input data tightly, which improves the color reproduction quality [Bergquist and Wennstam 2006]. As primaries approach each other in terms of brightness/energy, color break-up artifacts are attenuated, but not eliminated [Ou-Yang and Huang 2007]. Interestingly, decreasing the number of adaptive primaries to two reduces the temporal artifacts further while producing a modest degradation in color reproduction quality. This is because color distribution of natural images [Pouli et al. 2010] has low rank and often forms a set of *color lines* [Omer and Werman 2004] that can be represented with two dimensional projections. Such projections are coarse approximations of the color in the entire image, but become more accurate when computed locally [Fattal 2014; Liu et al. 2010; Tai et al. 2008]. Similarly, Cheng et al. [2009] combine two-field sequential display with *local* RGB LED backlight to reduce color break-up and projection rates. Their decomposition, similar to ours, computes two backlight and two modulation frames. However, our primaries are fully content-adaptive, while theirs adapt only to local brightness with color hue limited to R+B(magenta) and G+B(cyan). This approach does not guarantee full-color reproduction, and leads to large reconstruction errors in the missing blue channel.

*Multi-primary and hyperspectral displays.* In recent years, there has been a growing interest in multi-primary and hyperspectral displays [Teragawa et al. 2012]. Ajito et al. [2000] show that a 6-primary projection can greatly expand the color gamut. Multiple projection systems [Mohan et al. 2008; Rice et al. 2012, 2007] take it a step

further and accurately reproduce hyperspectral content. However, most require a combination of Digital-Mirror-Device (DMD) and wavelength decoupling device, which implies large form factor and high power consumption. Similar to Kauvar et al. [2015], our display is content adaptive and supports multispectral content, but can be implemented in a thin and lightweight form factor using off-the-shelf consumer LCDs.

*Hybrid spatio-temporal color synthesis.* The work by Silverstein [2005] shows that combining spatial and temporal multiplexing yields an attractive in-between solution. They propose a two-field sequential display with a uniform backlight switchable between yellow and blue. Illuminated with yellow light, the cyan/magenta checkerboard modulation layer produces red and green checkerboard frames, while blue backlight yields a uniform blue frame. Collectively they get integrated into a full-color image. This decomposition also reduces perceived color break-up as the red-green frame contains both high luminance and high contrast components and alternates in time with low luminance blue frame. While an improvement over classic spatial color synthesis displays, their design has lower spatial resolution and optical throughput when compared to a regular FSC display. More recently, Langendijk [2007] proposed a similar approach, but with a different set of wide-band per-pixel filters (green and magenta interleaved horizontally) and switchable uniform backlight (yellow and cyan). Combined with temporal synthesis, their configuration effectively produces four per pixel primaries and outperforms spatial synthesis displays in terms of brightness and resolution.

*Power saving display pipelines.* Power efficient displays are particularly important in a context of mobile devices. The improvements to LED and OLED materials [Forrest 2003] and dimming backlight circuitry [Narra and Zinger 2004] led to huge savings in power consumption. Recent work increase these further with gaze aware dimming [Moshnyaga and Morikawa 2005], content-adaptive local dimming [Seetzen et al. 2004] or global dimming with concurrent brightness and contrast scaling [Cheng et al. 2004]. The content-based power optimizations are further explored by saliency preserving dimming [Chen et al. 2014], color remapping [Chen et al. 2016; Chuang et al. 2009], and extensions to power-optimizing rendering [Wang et al. 2016].

*GPU texture compression.* Reconstruction of an image from two high-frequency modulations of low-frequency color images is common to many existing GPU texture compression methods. DXTC [Iourcha et al. 1999], for example, allows for a lossy compression of up to 6:1. The reduced memory/bandwidth requirements enable the use of higher resolution textures in the same memory budget, and largely mitigates the drop in color/texture quality. Because our decomposition algorithm can produce data compatible with the above encoding scheme, the output modulation layers can also be stored at lower precision to reduce bandwidth requirements. This allows us to transmit at higher resolutions or higher frame rates without increasing the nominal bandwidth. Notably, in contrast to texture compression methods, which require extra digital logic and hardware support for image decoding, our images are decoded in

“analog” domain – partially by the display hardware and partially by the human visual system.

In summary, we propose a display design based on FSC technology and content-adaptive selection of local color primaries. In contrast to prior work, we maintain all the advantages of FSC displays while addressing problems inherent to this technology – high refresh rate and bandwidth requirements.

### 3 COLOR FACTORIZATION

The emissive spectral distribution of a RGB color display is given by the irradiance:

$$e(x, \lambda) = \sum_{k=1}^3 i_k(x) f_k(\lambda), \quad (1)$$

where the image  $i$  contains 3 channels  $k$ , each multiplied with its corresponding spectral distribution color light source  $f_k$ . To model the perceived image for a human eye, under the “standard” observer model, the International Commission on Illumination CIE 1931 standard defines the perceived image  $i^{xyz}$  as a projection onto the 3 color-matching spectral basis functions  $\psi^{\bar{x}\bar{y}\bar{z}}(\lambda)$ :

$$\begin{aligned} i^{xyz}(x) &= \int e(x, \lambda) \psi^{\bar{x}\bar{y}\bar{z}}(\lambda) d\lambda \\ &= \sum_{k=1}^3 i_k(x) \int f_k(\lambda) \psi^{\bar{x}\bar{y}\bar{z}}(\lambda) d\lambda. \end{aligned} \quad (2)$$

Discretizing Equation 2, we can further factorize image  $\mathbf{I} \in N \times 3$  of  $N$  pixels into a more flexible representation using mixed color primaries and its corresponding modulations:

$$\mathbf{I}^{xyz} = \mathbf{I}^{gb} \mathbf{F} \Psi = \tilde{\mathbf{M}} \tilde{\mathbf{P}} \mathbf{F} \Psi, \quad (3)$$

where the matrix  $\Psi \in L \times 3$  encodes the spectral color matching function  $\psi^{\bar{x}\bar{y}\bar{z}}(\lambda)$  at  $L$  discrete wavelengths, the original spectral distribution of the three (RGB) light sources  $\mathbf{F} \in 3 \times L$  are blended by the primary-mixing matrix  $\tilde{\mathbf{P}} \in 3 \times 3$  such that the multiplication  $\tilde{\mathbf{P}} \mathbf{F}$  forms new bases for the *mixed-primary displays*, and the modulation matrix  $\tilde{\mathbf{M}} \in N \times 3$  represents the new coordinates of the pixels on the new mixed-primary axes. While an identity primary-mixing matrix  $\tilde{\mathbf{P}}$  gives a standard RGB space, it not necessarily the most efficient representation for a natural image: its pixels are not scattered uniformly in the RGB cube, but rather form a collection of *color lines* [Fattal 2014; Omer and Werman 2004] that can be efficiently represented via low-rank planar projection in RGB color-space.

*Content-adaptive global low-rank approximation.* Finding a content-dependent primary-mixing matrix  $\mathbf{P}$  instead of using the RGB primary matrix  $\tilde{\mathbf{P}}$  to capture the intrinsic image statistics allows for a more succinct representation of the color lines, and can potentially reduce the required data bandwidth and hardware cost. Principal Component Analysis (PCA) [Johnson et al. 2010] provides a formal tool to analyze the color variations between pixels: the principal components represent the primary-mixing axes with decreasing color variances. This gives us an elegant way of removing the primaries that contribute little to the color reproduction. Mathematically, we can fit, with minimal errors, the color pixels and lines embedded in a plane in RGB space by factorizing the data matrix

$I^{rgb}$  into a low-rank approximation such that the new modulation  $\mathbf{M} \in^{N \times 2}$  and the new primary-mixing  $\mathbf{P} \in^{2 \times 3}$  minimize the error with respect to the displayed image:

$$\begin{aligned} \arg \min_{\mathbf{M}, \mathbf{P}} \quad & \|\mathbf{I}^{xyz} - \mathbf{M}\mathbf{P}\Phi\|^2 \\ \text{subject to} \quad & 0 \leq \mathbf{M}_{i,k}, \mathbf{P}_{k,j} \leq 1, \end{aligned} \quad (4)$$

where  $\Phi = \mathbf{F}\Psi$  is the RGB-to-XYZ transform, and the non-negativity constraint enforces physically realizable pixel states. To enable transformations to other linear and non-linear color-spaces, we define the residual in CIEXYZ as opposed to display native RGB color-space. This step is optional and since the transformation is linear and strictly positive, it does not impact the quality of our results. Non-negative Matrix Factorization (NMF) [Berry et al. 2006] is commonly applied to computational displays; solutions can be efficiently found via an iterative solver [Blondel et al. 2008].

While the two mixed-primary axes found in Equation 4 span a best-fitting 2D plane to the majority of the 3D color pixels, there are certain regions that are difficult to fit and that are hard to ignore for a human observer, as shown in Figure 5, second column from the right.

*Locally low-rank approximation.* While natural images might not be globally low-rank, the reflectance and light transport imply that a locally low-rank approximation is possible [Mahajan et al. 2007]; an example can be found in the supplement, Sec.A. Since the chromatic spatial sensitivity bandwidth is also much lower than that in the luminance channel [Mullen 1985], we split the image into smaller blocks and replace the global backlight with a low-resolution color backlight array, and the fitting result is greatly enhanced by solving the NMF locally.

We achieve local fitting by dividing the image of resolution  $N$  into  $B$  blocks where  $B \ll N$  and growing the corresponding primary-mixing matrix  $B$ -times taller such that  $\mathbf{P} \in^{2B \times 3}$ . The local factorization problem can still be formulated in a similar and consistent manner:

$$\begin{aligned} \arg \min_{\mathbf{M}, \mathbf{P}} \quad & \|\mathbf{I}^{xyz} - \mathbf{M}\mathbf{P}\Phi\|^2 \\ \text{subject to} \quad & 0 \leq \mathbf{M}_{i,k}, \mathbf{P}_{k,j} \leq 1, \end{aligned} \quad (5)$$

where the binary permutation matrix  $\Pi \in^{2 \times 2B}$  reorders the modulation  $\mathbf{M}$  into a repeated block-diagonal structure such that each block multiplies its corresponding local mixed-primaries  $\mathbf{P}$ ; the locally factorized solution is shown in Figure 1.

### 3.1 Hardware and Perceptual Consideration

Realizing the operations described in the previous section involves computing a low resolution image shown on an array of RGB color backlights, with each individual backlight modulated by a local grayscale modulation pattern. The multiplication of the mixed-primary backlight array and the achromatic modulation layer, shown in Figure 1, constitutes a frame (one rank); to achieve a high fidelity approximation, it requires adding up two frames through spatial superposition or temporal multiplexing.

*Physical artifacts.* The temporal multiplexing scheme requires that both the backlight and the modulation layer refresh at high speeds. Although the LED backlight is capable of updating at 120Hz

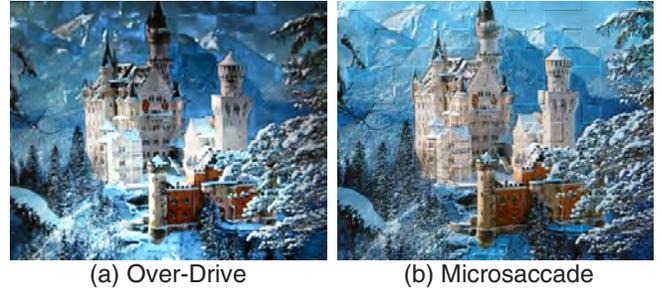


Fig. 2. Prototype problems due to transmissive modulations. (a) Photograph of nonlinear color blending due to over-driving liquid crystals. (b) Sharp block boundaries perceived during microsaccade.

almost instantly, using an LCD as the transmissive modulation layer presents challenges: the refresh rate is limited by the physical inertia of liquid crystals as they twist [Kälántár et al. 2006]. Voltage overdrive is commonly used to accelerate the refresh rate, but it also introduces nonlinear intensity responses [Ruckmongathan 2014]. The nonlinearity produces discontinuity artifacts between adjacent blocks, as shown in Figure 2(a): In areas of uniform color, only the first mixed-primary basis is needed to represent the area's color, and the second primary is free to be any color and is likely to vary between neighboring blocks. The different nonlinear mappings of the liquid crystal between blocks then result in different temporal summations of the frames that cause discontinuities.

One solution used in most high speed or 3D-vision displays to solve the nonlinear intensity change is through temporal impulse, a.k.a. low-persistence [Klompener 2006], usually in the form of a short backlight flash or an active shutter. The backlight or the shutter is turned off until the LC molecules are settled, and the duty cycle or the persistence is usually turned on for less than 50% of the 8ms for a 120Hz display. Although low-persistence solves the above artifacts as perceived by a static camera, we find it introduces new artifacts disturbing to a human observer.

*Perceptual artifacts.* The human eye constantly performs small and involuntary eye movements, known as microsaccades, to avoid overcharging the same retinal cones or to prevent retinal images from adaptation fading [Martinez-Conde et al. 2013]. However, during a microsaccade or smooth pursuit, temporal multiplexing can update the frames and pixels near block boundaries will add to its neighboring blocks. While the effect can sometimes be used to enhance resolution [Berthouzoz and Fattal 2012; Didyk et al. 2010], an incorrect integration can also generate grid artifacts around the block boundaries [Pan et al. 2005], as shown in Figure 2(b).

Large saccadic eye movement yields another perceptual problem with low-persistence displays. Extending the foveal vision, saccades help the mapping of a broader 3D world: the brain needs to localize and register features before and after saccade. Normally, during the quick eye movement, motion blur and saccadic suppression raise the detection threshold by the brain, so we become blind during this period even though photons continue to land on the retina [Bridgeman et al. 1975]. However, flashing or strobing of display content during saccade allows high frequency differences between frames to interfere with the registration process and causes *visual*

*instability* [Melcher 2011]: after saccade, we unconsciously notice movement of the image, even when it remains static.

### 3.2 Optimization with constraints

Constraining our optimization framework can address the above problems. The physical artifacts are due to the unconstrained free primary bases in flat areas; we found that regularizing the two modulation frames to be similar whenever possible also encourages the two primary frames to agree on the rank-1 basis.

Perceptual artifacts, on the other hand, require smoothing out edges between neighboring blocks, which simultaneously reduces visual instability artifacts and differences between neighboring local color primaries. We empirically found having a Gaussian kernel with standard deviation larger than the block size eliminates the perceptual artifacts in our prototypes. We insert a diffusing layer between the light source and the modulation, and rewrite the new permutation matrix  $\Pi_{\mathcal{N}}$  by including the (normalized Gaussian) diffusion kernel  $\mathcal{N}$  into Equation 5. The example result shown in Figure 1 is free from physical artifacts and visually stable. We improve the optimization framework by including both backlight diffusion and regularization of the modulation frames, and obtain the following:

$$\begin{aligned} \arg \min_{\mathbf{M}, \mathbf{P}} \quad & \|\mathbf{I}^{xyz} - \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi\|^2 + \gamma\|\mathbf{M}_{k_1} - \mathbf{M}_{k_2}\|^2 \\ \text{subject to} \quad & 0 \leq \mathbf{M}_{i,k}, \mathbf{P}_{k,j} \leq 1, \end{aligned} \quad (6)$$

where  $\gamma$  is a regularization constant to encourage the two modulation frames  $\mathbf{M}_{k_1}$  and  $\mathbf{M}_{k_2}$  to be as similar as possible. Similar to the factorization done by Heide et al.[2014], the solution to this optimization can be obtained efficiently using the Alternating Least Squares (ALS) iterative method: the update is given as follows:

$$\begin{aligned} \mathbf{M}^{i+1} & \leftarrow \left( \mathbf{M}^i + \frac{\mathbf{R}(\Pi_{\mathcal{N}}\mathbf{P}^i\Phi)^T + \gamma(\mathbf{M}_k^i - \mathbf{M}_k^i)}{\Pi_{\mathcal{N}}\mathbf{P}^i\Phi(\Pi_{\mathcal{N}}\mathbf{P}^i\Phi)^T} \right)_+ \\ \mathbf{P}^{i+1} & \leftarrow \left( \mathbf{P}^i + \frac{(\mathbf{M}^i\Pi_{\mathcal{N}})^T\mathbf{R}\Phi^{-1}}{(\mathbf{M}^i\Pi_{\mathcal{N}})^T\mathbf{M}^i\Pi_{\mathcal{N}}} \right)_+, \end{aligned} \quad (7)$$

where  $\mathbf{R} = (\mathbf{M}^i\Pi_{\mathcal{N}}\mathbf{P}^i\Phi - \mathbf{I}^{xyz})$  is the reconstruction residual, and  $(\cdot)_+$  accepts solutions that are positive and smaller than one. Note that explicitly evaluating the permutation matrix and solving the update globally is expensive and impractical for real-time display system. In Section 4, we show an efficient matrix-free solver that can be implemented in GPU.

### 3.3 Optimizing for Multispectral Displays

Display development traditionally has been focused on higher resolution and power efficiency. This has changed recently, with the rise of interest in displaying multispectral data with wider color gamuts. Packaging multiple color filters within a subpixel mosaic further reduces spatial resolution and light, thus multispectral flat-panel displays are rarely seen with more than 4 primaries (e.g., Sharp's Quattron<sup>1</sup>); temporal multiplexing schemes like projection systems are more suitable to hyperspectral imagery, but require multiple frames using fast DMD arrays or LCoS to temporally multiplex more

than three frames; the optical design makes it harder to apply to flat-panel displays.

By using locally mixed primaries with a *multispectral* LED array, multispectral images can be shown with minimal change to the aforementioned architecture. We just need to replace the  $\phi^{-1}$  in Eq. 7 with its Moore-Penrose pseudoinverse  $\phi^+$  to properly back-project from CIEXYZ to hyperspectral primaries. While minimizing the  $L_2$  error norm in Equation 6 yields satisfactory results in CIEXYZ space, human perception is neither linear nor uniform in the spectral space of radiometrically linear light sources, especially when considering wide color gamuts. We further extend the local factorization framework with perceptual metrics.

### 3.4 Perceptual Optimization

Human color perception is nonlinear in the trichromatic signal space: Hering's Opponent theory suggests that the signal received by the LMS cones is translated into luminance and opposing red-green and blue-yellow channels in the later stages of vision [Baumann 1992]. While there are perceptually uniform standards to model this behavior, more advanced metrics account for chromatic contrast sensitivity function at different spatial frequencies (S-CIELab [Zhang and Wandell 1997]), Crispening effect (CIEDE2000 [Sharma et al. 2005]), or Visual Masking [Ferwerda et al. 1997]. Although these metrics are powerful predictors of human perception, their formulations make them difficult to be included in an optimizer. We found the distance metrics in CIELab and in IPT space [Fairchild and Johnson 2004] produce satisfactory results and are easy to implement.

Directly solving the nonlinearly transformed Equation 6 in CIELab or IPT space is not straightforward, as the NMF is already itself nonlinear. We employ a simple decomposition technique by splitting the optimization into subproblems [Boyd et al. 2011] with the help of an intermediate variable  $\mathbf{T} = \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi$  and a scaled dual variable  $\mathbf{U}$ :

$$\begin{aligned} \mathbf{T} & \leftarrow \arg \min_{\mathbf{T}} \|\mathbf{I}^{Lab} - \Theta(\mathbf{T})\|^2 + \frac{\rho}{2}\|\mathbf{T} - \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi + \mathbf{U}\|^2 \\ \mathbf{M}, \mathbf{P} & \leftarrow \arg \min_{\mathbf{M}, \mathbf{P}} \rho\|\mathbf{I}^{xyz} - \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi\|^2 + \gamma\|\mathbf{M}_1 - \mathbf{M}_2\|^2 + \\ & \quad \frac{\rho}{2}\|\mathbf{T} - \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi + \mathbf{U}\|^2 \\ \mathbf{U} & \leftarrow \mathbf{U} + (\mathbf{T} - \mathbf{M}\Pi_{\mathcal{N}}\mathbf{P}\Phi). \end{aligned} \quad (8)$$

Although the subproblems are still nontrivial, they can be linearized using the Gauss-Newton method and alternating least squares, and we leave the derivation based on Alternating Direction Method of Multipliers (ADMM) to the supplement, Sec.C. Then, the final algorithm is highly parallel and can be implemented on the GPU; we will describe the details in Section 4.

## 4 IMPLEMENTATION

We implement the temporal multiplexing scheme with an LCD and an LED array, as shown in Figure 1 and Figure 3 with the individual frame shown. We also implement an optical spatial superposition as a projector using two monochromatic Spatial Light Modulators (SLMs) and two low resolution color LCDs; the projector prototype is detailed in the supplement, Sec.B.

<sup>1</sup><https://en.wikipedia.org/wiki/Quattron>

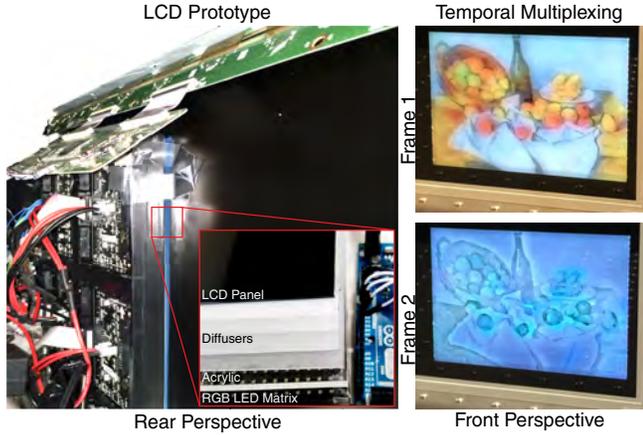


Fig. 3. Prototypes implementing the temporal multiplexing scheme, on the left, using a LCD panel and an array of color LED backlights. On the right, we show two temporal multiplexing frames.

*Hardware prototype.* We use an ASUS VG236H LCD display running at 120Hz with a corrected display-gamma (= 2.45), showing grayscale images as the modulation layer. The backlights are four  $32 \times 32$  RGB LED matrices with 4mm pitch (Adafruit Product ID 607), they also refresh at 120Hz and are driven by four Arduino MEGA boards, which provide 4 bits Pulse Width Modulation in each color channel. With the Arduino controller, we are able to lower the duty cycle of the backlight persistence to 46%, and the LEDs are refreshed with vertical scanning allowing enough front modulation LCD settlement. The pixels-to-block ratio between the LCD modulation and the LED backlight is  $15^2 : 1$ . The spacing between the modulation layer and the backlight is 11mm, and the diffusion is achieved by stacking 4 Edmund Optics 120 Grit Ground Glass Diffusers; we measured a Gaussian diffusion kernel with 5.3mm standard deviation (20 pixels on the modulation layer). The final display area is about 14 inches diagonally with a resolution of  $960 \times 960$  with a measured dynamic range of 3100:1. On the left of Figure 3 is the backlight module behind the LCD panel, and the inset shows the diffusing layers.

*Computer system.* The system is driven by a PC running on an Intel core-i7 3.4GHz CPU with 16GB of RAM. The displays are driven by an NVIDIA Quadro K6000 graphics card; we use the 3D-vision stereo signal to synchronize vertical refresh of the modulation and backlight.

*GPU implementation of the optimization.* Optimizing the factorization problem involves solving two subproblems in Equation 8: The first step solves for the intermediate variable  $T$ , and then factorizes it into the modulations  $M$  and the mixed-primaries  $P$ . The two subproblems are implemented on the GPU using CUDA.

Although the first subproblem is nonlinear in both CIELab and IPT space, it can be linearized using the Gauss-Newton iterative method, as shown in Algorithm 1. For each pixel  $j$ , we simply apply a 3-by-3 per-pixel transform to the residual vector  $r_j$ . The transform is derived from the Jacobian  $J_{t_j}$  to the first subproblem, and we show

#### Algorithm 1 Perceptual Optimization in GPU

```

1: procedure GAUSS-NEWTON ITERATION
2:    $r_j \leftarrow (i_j^{Lab} - \Theta(t_j)) + \frac{\rho}{2} (t_j - i_j^{rec} + u_j)$ 
3:    $t_j \leftarrow t_j - (J_{t_j}^T J_{t_j})^{-1} J_{t_j} r_j$ 
4: end procedure
5: procedure MODULATION UPDATE(Pixel  $j$ )
6:    $r_j^k \leftarrow \rho (i_j^{rec} - i_j^{xyz}) + \frac{\rho}{2} (i_j^{rec} - t_j + u_j)$ 
7:    $w_j^k \leftarrow \Phi \left( \sum_{n \in NeighborLEDs(j)} N_{\sigma}^{n,j} p_n^k \right)$ 
8:    $m_j^k \leftarrow (m_j^k - (r_j^{kT} w_j^k + \gamma(m_j^k - m_j^k)) / (w_j^{kT} w_j^k))_+$ 
9: end procedure
10: procedure MIXED-PRIMARY UPDATE(Backlight  $n$ )
11:   for all  $j \in Cover(n)$  do
12:      $r_j^k \leftarrow \rho (i_j^{rec} - i_j^{xyz}) + \frac{\rho}{2} (i_j^{rec} - t_j + u_j)$ 
13:      $w_j^k \leftarrow N_{\sigma}^{n,j} m_j^k$ 
14:   end for
15:    $p_n^k \leftarrow (p_n^k - \sum_j (\Phi^{-1}(w_j^k r_j^k) / \rho (w_j^k)^2))_+$ 
16: end procedure
17: procedure AUGMENTED NMF ITERATION
18:   MODULATION UPDATE( $M_0$ )
19:   MIXED-PRIMARY UPDATE( $P_0$ )
20:   MODULATION UPDATE( $M_1$ )
21:   MIXED-PRIMARY UPDATE( $P_1$ )
22: end procedure
23: procedure ADMM ITERATION
24:   repeat
25:     GAUSS-NEWTON ITERATION
26:   until  $T$  converges or iteration criteria is met
27:   repeat
28:     AUGMENTED NMF ITERATION
29:   until  $M$  and  $P$  converge or iteration criteria is met
30: end procedure

```

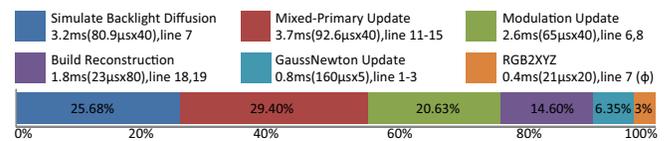


Fig. 4. Performance break down of one ADMM iteration, which takes 12.5ms; the individual operation time, kernel time, and invocation counts are also given. The dominant cost comes from the direct diffusion process (line 7, dark blue) and the inverse diffusion for residual aggregation (line 11-13, red) in the PRIMARY UPDATE; they together account for 55% of the time; the overall performance is 2-3 orders of magnitude faster than prior work.

the analytical solution to the inversion  $(J_{t_j}^T J_{t_j})^{-1}$  in the supplement, Sec.C.1.

The second subproblem in the optimization is an augmented NMF that can be solved in a manner similar to the update rules in Equation 7. However, the local block-permutation matrix with Gaussian kernel  $\Pi_N$  is expensive to build. To enable an efficient

GPU implementation, the local block NMF is cast into matrix-free per-pixel kernel updates. The following MODULATION UPDATE and MIX-PRIMARY UPDATE kernels in Algorithm 1 states the procedures to complete one iteration of the NMF. In the modulation update, the summation of the weighting  $w_j^k$  accounts for the 49 neighboring backlights, and can be precomputed as a diffused backlight image in another independent kernel pass. The diffusion  $\mathcal{N}_\sigma^{n,j}$  is a normalized Gaussian with the distance measured from the center of the backlight  $n$  to the pixel  $j$ . In the mixed-primary update, a particular backlight  $p_n^k$  requires summing the residual over all pixels  $j$  under the LED's coverage, which can be thousands per backlight unit. Instead, we distribute the coverage computation to more threads, and sum them up through standard parallel reduction on the GPU.

For video and time-critical applications, we typically run 20 to 50 iterations of the NMF algorithm, which takes 0.59ms per iteration on a NVIDIA Titan X GPU, and we use the results from the previous frame as warm-start to achieve realtime performance and temporal coherence, as shown in the supplementary video. To accelerate the computation, we implement separable kernel filtering and prefix sum on GPU, achieving  $O(N)$  complexity for a 2D image of  $N$  pixels; the performance of using our more complex backlight system is still 650 times faster than that of the implementation by Kauvar et al.[2015].

For static image requiring high color fidelity, as shown in our result Figure 5, we use the perceptual ADMM optimization. The choices of the constants  $\rho$  and  $\gamma$  need to be consistent with respect to the scale in CIELab or IPT space. For our CIEXYZ optimizer, we choose  $\rho = 1.5 \times 10^5$  and  $\gamma = 0.25 \times 10^4$ . For IPT, we choose  $\rho = 25$ . The ADMM algorithm takes 4 to 6 global iterations to achieve a reasonable solution, and it slowly converges for another 15 to 17 ADMM iterations. Each ADMM iteration requires 5 iterations of the Gauss-Newton solver and followed by 20 iterations of the NMF solver. The time consumed by each ADMM iteration is 12.5ms with its breakdown shown in Figure 4, and it takes 300ms in total to solve for an image compared to several hours of computation in prior work.

## 5 RESULTS AND EVALUATION

In Figure 5, we show our CIELab optimization results and photographs from the LCD display prototype, and compare against the content-adaptive global two-primary display, as described in Section 3, the Two-Field Scheme [Cheng et al. 2009], and the reference image using Field-sequential Color, where  $\mathbf{P}$  is identity in Equation 3. The selection covers a wide range of images, and more examples can be found in the supplement, Sec.K.

The content-adaptive global two-primary method chooses the two mixed-primaries as global backlights for the two frames. However, due to the lack of local adaptation, certain colors are missing, as illustrated throughout. In contrast, the two-field scheme preserves more local color details. While their scheme, just like ours, uses an array of LED backlights, their decomposition method constrains the hue of two-frame primaries to magenta/cyan variants, which significantly degrades the fidelity of the blue channel. In the supplementary material we show that our method, with fully adaptive

	Metrics	Global 2-primary	Cheng [2009]	Ours
Dining Hall	PSNR	34.53	19.12	37.77
	HDRVDP2	99.25	92.18	99.75
	SSIM	0.9978	0.9938	0.9990
	$\Delta E_{00}$	4.00	13.74	2.49
Cameleon	PSNR	23.58	21.00	33.58
	HDRVDP2	90.65	93.72	99.13
	SSIM	0.9978	0.9932	0.9976
	$\Delta E_{00}$	9.95	8.35	2.46
Painting	PSNR	34.58	17.77	34.01
	HDRVDP2	98.96	95.64	99.64
	SSIM	0.9966	0.9875	0.9973
	$\Delta E_{00}$	4.98	18.22	4.16
Parrot	PSNR	25.35	22.99	38.64
	HDRVDP2	92.77	91.76	98.82
	SSIM	0.9736	0.9951	0.9959
	$\Delta E_{00}$	7.24	7.82	1.11

Table 1. Quantitative image reproduction quality results.

selection of primaries and iterative optimization, further improves on the results of the two-field scheme.

Although our algorithm generally produces pleasing results, we notice errors when cluttered multi-colored features are much smaller than the backlight block size, which is a fundamental limitation of any compressive type display. Nevertheless, the apparent reconstruction errors are small, as indicated by the SSIM map. This is because the modulation layer to a large extent preserves the detail and luma changes, which our visual system is more sensitive to.

*Color metrics and perceptual optimization.* In Table 1, we compare the LCD prototype results with common image error metrics like Peak Signal-to-Noise Ratio (PSNR), perceptually-based metrics like HDR-VDP2 [Mantiuk et al. 2011], Structural Similarity Index (SSIM) [Wang et al. 2004], and perceptual color metric CIE  $\Delta E_{00}$  [Sharma et al. 2005]; in the supplement, Sec.K, we also expand our comparison with Feature Similarity Index (FSIM) [Zhang et al. 2011], CIE  $\Delta E_{76}$ , and Spatial-CIELab [Zhang and Wandell 1997]. With the perceptual color metric CIE  $\Delta E_{00}$ , we found Georges Seurat's painting has particular high reconstruction error among all test examples: since the painting uses dots with primary contrast colors to rearrange the desired color, known as chromoluminarism.

While these numbers give a global evaluation of the image quality, we find that the error maps give more visually meaningful information to the high error region, as shown on the left-most column in Figure 5.

In Figure 6, we compare the results obtained, in the same amount of computation time, by optimizing in the linear sRGB and CIEXYZ, and the perceptual CIELab and IPT spaces. While linear sRGB optimizer generally works well, it is not as flexible as those transformed through CIEXYZ, and does not supports multi-primary displays. We observed that optimization in linear CIEXYZ space does not produce image as rich in color as that in CIELab space. Finally, although the IPT optimizer is easy to implement, it is not on par with the CIELab optimizer.

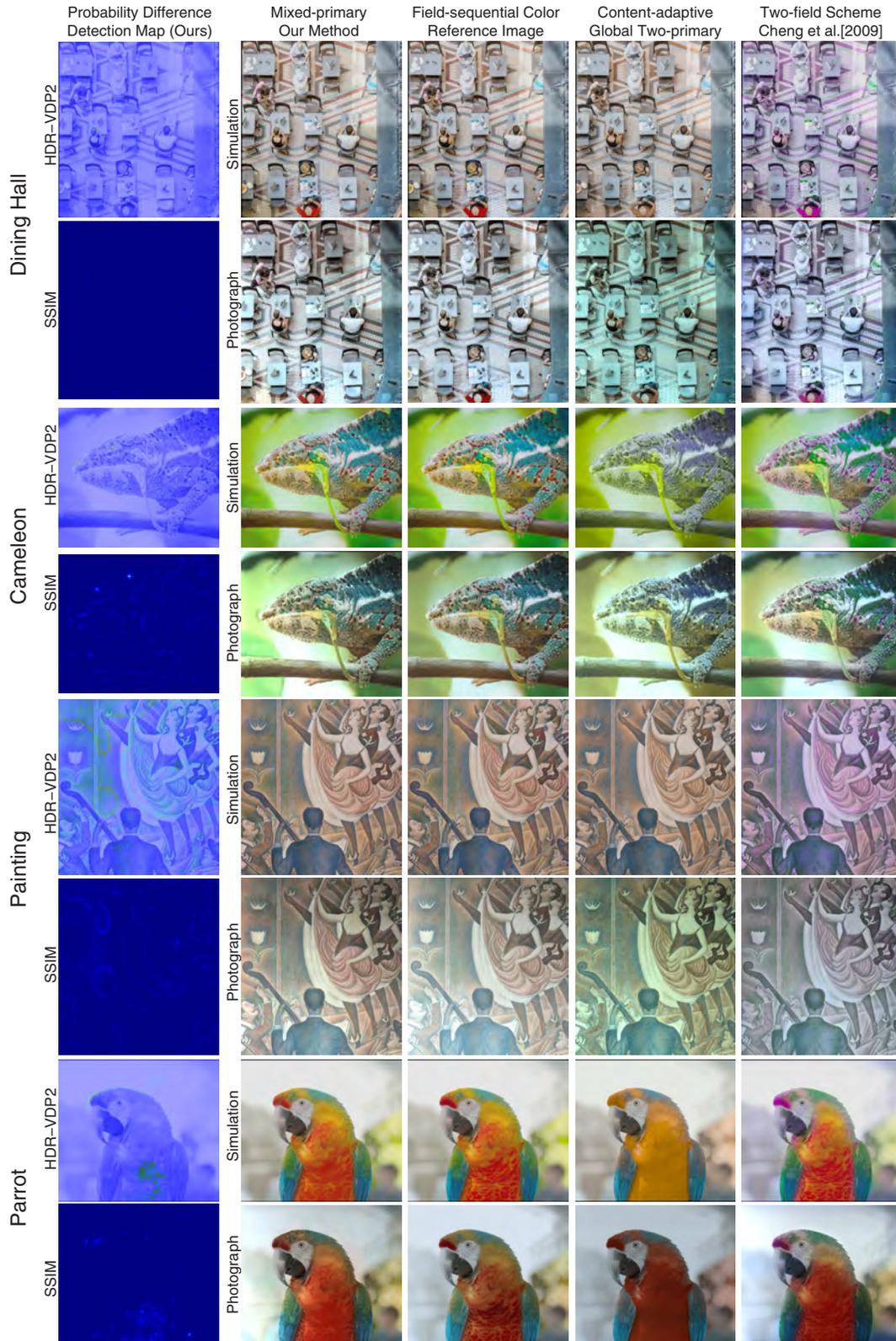


Fig. 5. Simulated results and photograph from the LCD prototype, captured with Canon T3i and 18-135mm lens at ISO100/F11/2s with linear RAW output. Quantitative results are shown in Table 1. Images courtesy of Jiamin Bai, Flickr user Kuhnmi, Georges Seurat, and Derrick Coetzee. ACM Transactions on Graphics, Vol. 36, No. 4, Article 1. Publication date: July 2017.

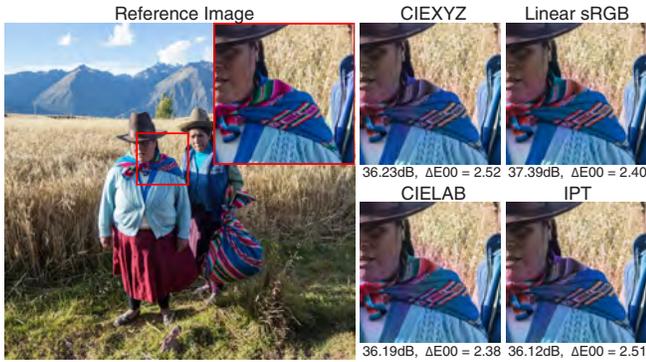


Fig. 6. Comparing simulated results optimized in different color spaces. Optimizations done in the linear spaces give higher PSNR, but the perceptual optimizers give better  $\Delta E_{00}$  values. Although the linear sRGB optimizer gives good result, its framework in practice does not allow the matching of the color filter spectrum in Eq. 2. Image courtesy of Jiamin Bai.

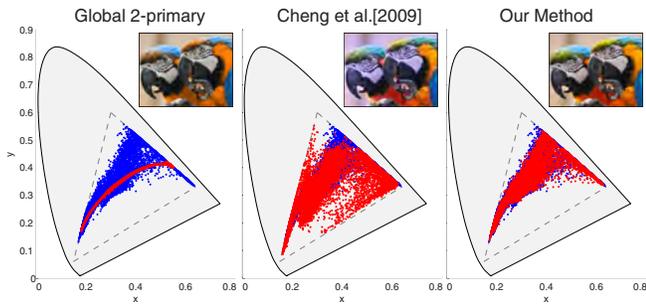


Fig. 7. Comparing color gamut result. The pixels from the reference image are plotted in blue dots. **(Left)** Since the global two-primary method uses only two optimized mixed-primary colors, the reconstructed points are concentrated along a line. **(Center)** The two-field scheme uses an array of color LED backlights so their color gamut is much expanded. However, without proper content adaptive optimization, their method has a problem in reproducing yellow and incorrectly shifts toward blue/cyan. **(Right)** Our method optimizes the use of the color backlights array so the gamut closely agrees with the reference.

**Color gamut.** Beyond comparing the numerical errors, we also evaluate how effective the localized mixed-primary optimization preserves the color gamut of the original image. On the left of Figure 7, we show the pixels optimized using the content-adaptive global two-primary method in red and the original color pixels in blue. Since the global two mixed-primary frames represent two end points, the color gamut is collapsed onto a narrow curve. The two-field scheme [Cheng et al. 2009] (Fig. 7, center) uses an array of color LED backlights, allowing many local curves to span the color gamut, but these curves are not well optimized to fit the content, so the error is large; our method (Fig. 7, right) optimizes the primary-mixing backlights so that these individual local curves represent the original colors well. Finally, our method also reconstructs colors well for extreme cases like showing a color palette with wide color variations, as seen in Figure 8, where we also evaluate the reconstruction quality at different frequencies.

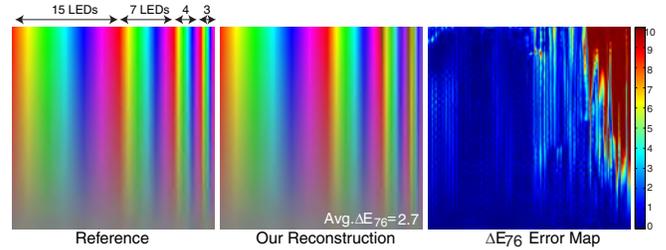


Fig. 8. Reconstruction of the hue wheel at different frequency and saturation levels. We test the limit of chromatic reconstruction at 4 frequency bands:  $2\pi$ ,  $4\pi$ ,  $8\pi$ , and  $16\pi$  of the hue wheel. The first  $2\pi$  portion uses 15 color LEDs, and the reconstruction has low  $\Delta E_{76}$  errors, as shown on the error map. As the frequency goes to higher, the error ramps up quickly; more details in the supplement, Sec.F.



Fig. 9. The simulated spatio-temporal hybrid approach reproduces the complete color gamut by showing red and green colors in the first frame and blue in the second frame. Since the reds and greens are spatially interleaved (checkerboard), the maximum spatial resolution is not as high as in our mixed-primary approach. Image courtesy of Flickr user kuhnmi.

**Comparison with a spatio-temporal hybrid decomposition.** Spatio-temporal multiplexing provides an alternative to spatial and temporal color synthesis that preserves color gamut well. In Fig. 9 we demonstrate the performance of [Silverstein 2005] against our method. The hybrid approach has good color reproduction and better spatial resolution when compared to the traditional RGB CFA approach, but when compared to our method, the achieved image resolution and light efficiency are still lower due to the use of the cyan-magenta CFA.

**Comparison with Adaptive Spectral Projection.** Enabling hyperspectral displays is fundamentally challenging, as discussed in Section 3.3. With temporal multiplexing, even three-frames displays still requires proper content adaptation, as shown by the Adaptive Spectral Projection (ASP) technique [Kauvar et al. 2015]. Here, we use a two-frame LCD prototype with a multispectral LEDs (same spectral distributions as those used by Kauvar et al.) array as the backlight, and compare with their rank-2 (two-frames) results in simulation in Figure 10. Note that the errors using ASP are significantly higher with such low rank reconstruction, as compared to ours; the errors will reduce drastically when more ranks become available, as shown in the supplement, Sec.I.

## 5.1 Hardware Prototypes

Practical implementations are still subject to physical constraints imposed by the underlying hardware. In the temporal multiplexing LCD prototype, the speed of turning the liquid crystals has a great

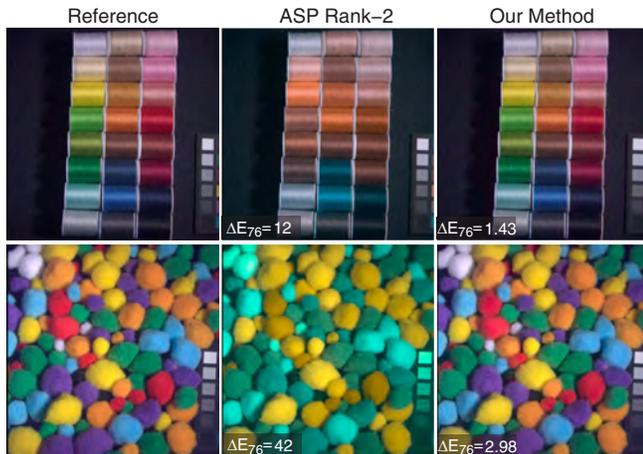


Fig. 10. Comparing simulated Adaptive Spectral Projection (ASP) (middle) with only rank-2 (two frames) reconstruction on a multispectral dataset to our method (right), which has significantly lower perceptual error. As shown in the supplement, Sec.I, ASP improves notably with a higher frame rate budget. Dataset from Columbia University Multispectral Image Database.

impact on the color reproduction. The slower the liquid crystal turns from a full white in one modulation frame to a full black in the other frame, the less color saturation the display can achieve. The slow turning also causes the hue shifting toward the primary in the first frame for any mixed-primary display. In our case, it can be seen in the Dining Hall scene that the central blue menu is not as saturated in the photograph. The hue shift is particularly visible in the global two-primary method with the first Dining Hall and the last Parrot examples, and our smooth constraint in the optimization solves the problem.

*Low-persistence and duty cycle.* Lowering the persistence of the temporal multiplexing is one way to improve the saturation problem. However, shortening the duty cycle while maintaining the bit-depth for the color backlights requires fast micro-controller or FPGA: this is beyond the scope of the paper. Another problem with the low duty cycle is visual instability. While increasing the diffusion alleviates the perceptual artifact, the diffusion also damages the image quality; adding more backlights can solve the visual instability problem if the modulation LCD panel has low speed.

*Power efficiency, bandwidth, and compression.* The advantage of locally mixed-primary displays is great power efficiency, as shown in Figure 11. In a color filter array mosaic display, the maximum brightness throughput is only one third of the backlight. The optimized color backlight allows better light efficiency without using spatial color filters by directly showing local color variations in the backlight. Note that our dual modulation approach allows optimizing for a more transparent modulation layer and reduced backlight energy output without degrading image quality by adding a small penalty term  $-\epsilon M_k$  to Equation 6; our display in the example figure requires about 70% less lighting intensity and thus provides 3x power saving to traditional LCD displays. We also compared with methods using local dimming: HDR display [Seetzen et al. 2004]

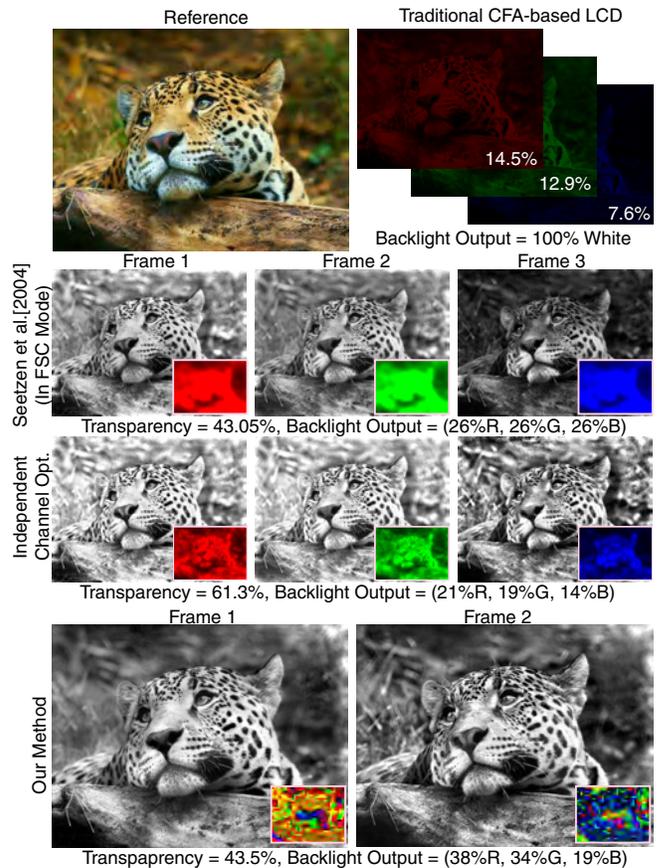


Fig. 11. Comparing the simulated light efficiency and energy requirement. Using a Color Filter Array (top right) filters out 66% of the light per channel, and a lot of energy is wasted. Seetzen et al.[2004] replace the backlight with white LEDs array allowing local dimming, but CFA is still used and thus inefficient in light (backlight output = 78% white). Modifying the operation to run in 3-frame Field-Sequential Color without color filters (second row) or optimizing individual frame (third row) allow significant power saving (3x-6x from the original CFA mode) but introduces more color breakup. Our locally mixed-primary display (bottom row) also allows more light to pass through the modulation layer at reduced color breakup; the optimization can be modified to allow more transparent modulations and decreased brightness in the backlight by adding a small penalty term ( $\epsilon = 0.0005$ ), achieving more than 3x power saving compared to the traditional CFA approach.

running in field-sequential color and in a modified mode that optimizes each channel independently, as shown in Figure 11. The FSC method generates same color backlight blobs by sharing the same luminance, and the modified method optimizes better to individual chromatic features; Both methods allow to save power (26% and 18% respectively), but also both suffer from color breakup.

Another advantage using the mix-primary two frames display is bandwidth saving throughout the display pipeline by an additional 32% since each pixel has only two intensity values instead of three and the added bandwidth from the color backlight is small. The result is also highly compressible; we compare with standard compression algorithms like JPEG2000 in the supplement, Sec.H.

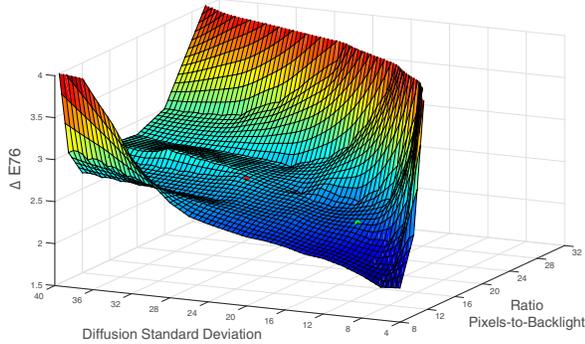


Fig. 12. Effects of local patch size and diffusion width on the image quality. While larger local patch damages the locally low-rank assumption, appropriate diffusion width has bigger impact: standard deviation close to the pixels-to-backlight ratio gives good result; the ratio depends on the desired design of bandwidth. The parameter of the LCD prototype is shown in the red dot, and the projector prototype is the green dot.

## 5.2 Backlight Resolution and Diffusion

As discussed in Section 3.1, to remove the visual instability the diffusion kernel has to be larger than the block size, or the pixels-to-backlight ratio. One important observation from the parameter space is: *the lower the pixels-to-backlight ratio and the smaller the diffusion, the better the image quality*. Although our prototypes are similar in their pixels-to-backlight ratios, their diffusion characteristics differ significantly: kernel width is 120 pixels ( $\sigma = 20$ ) in the LCD prototype vs. 71 pixels ( $\sigma = 11.8$ ) in the projector prototype. In addition to the parameters used in our prototypes, we analyze the parameter space spanned by the diffusion kernel size and the pixel-to-backlight ratio for designing the hardware, and Figure 12 shows the  $\Delta E_{76}$  error of the Dining Hall example. There are two cases of larger error for a given pixels-to-backlight ratio: a diffusion kernel width smaller than or much larger than the backlight block size. To obtain the best results, we need to ask the question: *How small can the diffusion kernel be if we want to avoid any quality loss?* While the obvious answer is to avoid being smaller than the block size, the deciding factor is actually the underlying LCD technology and the visual stability.

*Diffusion and visual instability.* In the section describing perceptual artifacts, we mentioned that the diffusion depends on the visual stability, which is affected by the speed of the LCD technology and the duty cycle of the backlight. In the LCD prototype, we fix the duty cycle (because of the LCD hardware chosen) and evaluate this quality empirically by adding diffusers. A more theoretical analysis on the visual instability subject requires further investigation and is beyond the scope of the paper.

## 6 EXTENSIONS

There are several interesting extensions to our display architecture, e.g. GPU texture compression compatibility, a microlens-based light field display, and an extended dynamic range display; details can be found in the supplement, Sec.H, Sec.G, and Sec.J.

Our optimization and display architecture is largely compatible with block-based GPU texture compression, e.g., DXTC [Iourcha

et al. 1999]. We can use existing DXTC GPU compressors to drive our proposed display architecture, and have both increased spatial resolution and reduced bandwidth. The display design is also suitable to autostereoscopic multiview displays or light field displays. Such displays typically require spatial resolution much higher than the perceived imagery. Although the embedded angular complexity reduces the local color coherence, our algorithm can still produce results faithfully. Finally, the dual modulation design can go beyond standard 8-bit displays to extend the dynamic range. Similar to high-dynamic-range display design by Seetzen et al. [2004], our prototype puts the chroma in the much lower resolution LED backlight as opposed to the front modulation panel, and our display can still show high dynamic range; limitations are discussed in the supplement, Sec.J.

## 7 LIMITATIONS

While our optimization framework minimizes the perceived color reconstruction error, the hardware design of the backlight resolution, the diffuser characteristics, and the underlying LCD technology set the practical limit on the achievable quality. However, the backlights resolution can be optimized with, e.g., color contrast sensitivity function ( $\sim 11$  cycles per degree [Mullen 1985]) given particular viewing conditions. In our prototype, low-persistence used to address the LCD overdrive can cause visual fatigue over extended viewing periods. Although we significantly reduce the requirements for multi-primary colors and improve the light efficiency using just two frames, applications requiring highly accurate colors will not benefit from our design, and the added computational overhead can offset some benefit in power saving from the display. Finally our color and dynamic range reproduction are limited in sharp shadow/highlight transitions due to backlight diffusion and the LED block size.

## 8 FUTURE WORK

The perceptual color distance metrics used by our optimizer are pixel-wise and do not consider spatial processing aspects of human visual system. Extending the model to consider the orientation and spatial frequency [Fairchild and Johnson 2004; Mantiuk et al. 2011; Zhang and Wandell 1997] of the input signal could possibly lead to additional quality gains for complex image patterns. Similarly, the HDR color reproduction can be improved with HDR-enabled versions of CIELab and IPT [Fairchild and Chen 2011]. Unlike in typical HDR display implementations, our colors are produced in the backlight layer. To maximize the image quality of such a setup, we believe that luminance and chrominance contrast sensitivity functions have to be considered jointly. The perception of low-persistent displays requires further study and exploration; in our case, reducing unnecessary high frequencies by introducing strong diffusion in the backlight solved the problem. While low-persistent technology has been used for ultra-low motion blur displays [Fisekovic et al. 2001] or judder reduction in a VR/AR headmount [Abrash 2013], the visual instability artifacts are hard to quantify without a formal psychophysical study. Finally, optimizing locally compressive images with more than three primaries to achieve a higher rank approximation [Teragawa et al. 2012] is highly desirable, and our framework is flexible enough to include it as well.

## 9 CONCLUSION

Saving bandwidth, power and manufacturing costs while enabling higher resolution and dynamic range remains a crucial objective for the next generation displays. To this end, we propose an alternative direction of building color displays based on content-adaptive selection of local color primaries. Our frame decomposition method is well founded mathematically, extensible, and the resulting frame representation is naturally compressible and compatible with existing texture compression formats. We maintain high color reproduction quality and enable local gamut optimization, which improves dynamic range, brightness and power efficiency of the display. With the proposed optimization framework, flexible GPU factorization, and demonstrated LCD and projector prototypes, we hope to contribute a practical technology that improves color displays. We believe that our joint consideration of optics, computations, and human perception is an important step towards more efficient and extensible display architectures.

## REFERENCES

- Michael Abrash. 2013. Why virtual isn't real to your brain: judder. <http://blogs.valvesoftware.com/abrash/why-virtual-isnt-real-to-your-brain-judder/>. (2013).
- Takeyuki Ajito, Takashi Obi, Masahiro Yamaguchi, and Nagaaki Ohyama. 2000. Expanded color gamut reproduced by six-primary projection display. *Proc. SPIE* 3954, 130–137.
- C Baumann. 1992. Ewald Hering's Opponent Colors. History of an idea. *Der Ophthalmologe : Zeitschrift der Deutschen Ophthalmologischen Gesellschaft* 89, 3 (1992), 249–252.
- Bryce E. Bayer. 1976. Color imaging array. U.S. Patent 3971065 A. (1976).
- Johan Bergquist and Carl Wennstam. 2006. Field-Sequential-Colour Display with Adaptive Gamut. *SID Symposium Digest of Technical Papers* 37, 1 (2006), 1594–1597.
- Michael W. Berry, Murray Browne, Amy N. Langville, V. Paul Pauca, and Robert J. Plemmons. 2006. Algorithms and Applications for Approximate Nonnegative Matrix Factorization. In *Computational Statistics and Data Analysis*. 155–173.
- Floraine Berthouzoz and Raanan Fattal. 2012. Resolution Enhancement by Vibrating Displays. *ACM Trans. Graph.* 31, 2, Article 15 (2012).
- V. Blondel, N-D Ho, and P van Dooren. 2008. Weighted Nonnegative Matrix Factorization and Face Feature Extraction. In *Image and Vision Computing*. 1–17.
- Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Found. Trends Mach. Learn.* 3, 1 (2011), 1–122.
- Bruce Bridgeman, Derek Hendry, and Lawrence Stark. 1975. Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research* 15, 6 (1975), 719 – 722.
- Haidong Chen, Ji Wang, Weifeng Chen, Huamin Qu, and Wei Chen. 2014. An image-space energy-saving visualization scheme for OLED displays. *Computers & Graphics* 38 (2014), 61 – 68.
- Weifeng Chen, Wei Chen, Haidong Chen, Zhengfang Zhang, and Huamin Qu. 2016. An Energy-saving Color Scheme for Direct Volume Rendering. *Comput. Graph.* 54, C (Feb. 2016), 57–64.
- Wei-Chung Cheng, Yu Hou, and Massoud Pedram. 2004. Power minimization in a backlit TFT-LCD display by concurrent brightness and contrast scaling. In *Proceedings Design, Automation and Test in Europe Conference and Exhibition*, Vol. 1. 252–257 Vol.1.
- Yu-Kuo Cheng, Yi-Pai Huang, Yi-Ru Cheng, and Han-Ping D. Shieh. 2009. Two-Field Scheme: Spatiotemporal Modulation for Field Sequential Color LCDs. *Journal of Display Technology* 5, 10 (2009).
- Johnson Chuang, Daniel Weiskopf, and Torsten Möller. 2009. Energy Aware Color Sets. *Computer Graphics Forum* 28, 2 (2009), 203–211.
- Piotr Didyk, Elmar Eisemann, Tobias Ritschel, Karol Myszkowski, and Hans-Peter Seidel. 2010. Apparent Display Resolution Enhancement for Moving Images (*SIGGRAPH 2010*). ACM, Article 113, 8 pages.
- Mark Fairchild. 2013. *Color appearance models*. John Wiley & Sons.
- Mark D. Fairchild and Ping-Hsu Chen. 2011. Brightness, lightness, and specifying color in high-dynamic-range scenes and images. *Proc. SPIE* 7867.
- Mark D. Fairchild and Garrett M. Johnson. 2004. The iCAM framework for image appearance, image differences, and image quality. *Journal of Electronic Imaging* 13 (2004), 126–138.
- Raanan Fattal. 2014. Dehazing using Color-Lines. *ACM Transaction on Graphics* 34, 1 (2014), 13:1–13:14.
- James A. Ferwerda, Peter Shirley, Sumanta N. Pattanaik, and Donald P. Greenberg. 1997. A Model of Visual Masking for Computer Graphics. In *SIGGRAPH*. 143–152.
- N. Fisekovic, T. Nauta, H. Cornelissen, and J. Bruinink. 2001. Improved motion-picture quality of AM-LCDs using scanning backlight. In *Asia Display/IDW*. 1637–1640.
- Stephen R. Forrest. 2003. The road to high efficiency organic light emitting devices. *Organic Electronics* 4, 2–3 (2003), 45 – 48.
- A. Frankenstein and Werner von Jaworski. 1904. Verfahren und Vorrichtung zur Fernsichtbarmachung von Bildern und Gegenständen mittels Selenzellen, Dreifarbenfilter, und Zerlegung des Bildes in Punktgruppen durch Spiegelung. German Patent 172376. (1904).
- Felix Heide, Douglas Lanman, Dikpal Reddy, Jan Kautz, Kari Pulli, and David Luebke. 2014. Cascaded Displays: Spatiotemporal Superresolution Using Offset Pixel Layers. *ACM Trans. Graph. (SIGGRAPH)* 33, 4 (2014), 60:1–60:11.
- K.I. Iourcha, K.S. Nayak, and Z. Hong. 1999. System and method for fixed-rate block-based image compression with inferred pixel values. US Patent 5,956,431. (1999).
- Garrett M. Johnson, Xioyan Song, Ethan D. Montag, and Mark D. Fairchild. 2010. Derivation of a Color Space for Image Color Difference Measurement. *Color Research and Application* 35, 6 (2010).
- Kälil Kälantär, Tadashi Kishimoto, Kazuo Sekiya, Tetsuya Miyashita, and Tatsuo Uchida. 2006. Spatio-temporal scanning backlight mode for field-sequential-color optically-compensated-bend liquid-crystal display. *Journal of SID* 14, 2 (2006).
- I. Kauvar, S. Yang, L. Shi, I. McDowall, and G. Wetzstein. 2015. Adaptive Color Display via Perceptually-driven Factored Spectral Projection. *ACM Trans. Graph. (SIGGRAPH Asia)* 34, 6 (2015).
- Gorham Kindem. 1981. The Demise of Kinemacolor: Technological, Legal, Economic, and Aesthetic Problems in Early Color Cinema History. *Cinema Journal* 20, 2 (1981), 3–14.
- Michiel A. Klompenhouwer. 2006. Comparison of LCD Motion Blur Reduction Methods using Temporal Impulse Response and MPRT. *SID Symposium Digest of Technical Papers* 37, 1 (2006).
- Edwin H. Land. 1959. Experiments in Color Vision. *Scientific American* 200, 5 (1959), 84–94.
- Erno H.A. Langendijk. 2007. A novel spectrum-sequential display design with a wide color gamut and reduced color breakup. *Journal of the Society for Information Display* 15, 4 (2007), 261–266.
- Y. H. Liu, Z. Z. Yang, and S. C. Wang. 2010. A novel sequential-color RGB-LED backlight driving system with local dimming control and dynamic bus voltage regulation. *IEEE Transactions on Consumer Electronics* 56, 4 (2010).
- Dhruv Mahajan, Ira Kemelmacher Shlizerman, Ravi Ramamoorthi, and Peter Belhumeur. 2007. A Theory of Locally Low Dimensional Light Transport. *ACM Transaction on Graphics (SIGGRAPH)* 26, 3 (2007), 62:1–62:10.
- Rafal Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. 2011. HDR-VDP-2: A Calibrated Visual Metric for Visibility and Quality Predictions in All Luminance Conditions. *ACM Trans. Graph.* 30, 4 (2011).
- Susana Martinez-Conde, Jorge Otero-Millan, and Stephen L Macknik. 2013. The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nature reviews. Neuroscience* (2013).
- Belen Masia, Gordon Wetzstein, Piotr Didyk, and Diego Gutierrez. 2013. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers & Graphics* 37, 8 (2013), 1012 – 1038.
- David Melcher. 2011. Visual stability. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 366, 1564 (2011), 468–475.
- Ankit Mohan, Ramesh Raskar, and Jack Tumblin. 2008. Agile Spectrum Imaging: Programmable Wavelength Modulation for Cameras and Projectors. *Computer Graphics Forum* 27, 2 (2008), 709–717.
- Mineo Mori, Toyohiko Hatada, Kazuo Ishikawa, Tosio Saishouji, Osamu Wada, Junichi Nakamura, and Nobuyoshi Terashima. 1999. Mechanism of color breakup in field-sequential-color projectors. *Journal of the Society for Information Display* 7, 4 (1999), 257–259.
- V. G. Moshnyaga and E. Morikawa. 2005. LCD display energy reduction by user monitoring. In *2005 International Conference on Computer Design*. 94–97.
- K T Mullen. 1985. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology* 359, 1 (1985), 381–400.
- Prathyusha Narra and D. S. Zinger. 2004. An effective LED dimming approach. In *Conference Record of the 2004 IEEE Industry Applications Conference, 2004. 39th IAS Annual Meeting.*, Vol. 3.
- I. Omer and M. Werman. 2004. Color lines: image specific color representation. *IEEE CVPR* (2004), 946–953.
- Mang Ou-Yang and Shih-Wei Huang. 2007. Design considerations between color gamut and brightness for multi-primary color displays. *Journal of Display Technology* 3, 1 (2007), 71–82.
- Hao Pan, Xiao-Fan Feng, and S. Daly. 2005. LCD motion blur modeling and analysis. In *IEEE ICIP*.
- A. A. Polumordvinov. 1899. Russian Patent 10738. (1899).

- Tania Pouli, Douglas W Cunningham, and Erik Reinhard. 2010. Image statistics and their applications in computer graphics. In *European Computer Graphics Conference and Exhibition*.
- Joseph P. Rice, Steven W. Brown, David W. Allen, Howard W. Yoon, Maritoni Litorja, and Jeesseong C. Hwang. 2012. Hyperspectral image projector applications. *Proc. SPIE* 8254, 82540R–82540R–8.
- Joseph P. Rice, Steven W. Brown, Jorge E. Neira, and Robert R. Bousquet. 2007. A hyperspectral image projector for hyperspectral imagers. *Proc. SPIE* 6565, 65650C–65650C–12.
- Temkar N. Ruckmongathan. 2014. *Addressing Techniques of Liquid Crystal Displays*. Wiley.
- Alfred C. Schroeder. 1948. Color television tube. U.S. Patent 2446791A. (1948).
- Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcovs. 2004. High Dynamic Range Display Systems. *ACM Trans. Graph. (SIGGRAPH)* 23, 3 (2004), 760–768.
- G. Sharma, W. Wu, and E. N. Dalal. 2005. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. *Color research and application* 30, 1 (2005), 21–30.
- Louis D. Silverstein. 2005. STColor: Hybrid Spatial-Temporal Color Synthesis for Enhanced Display Image Quality. In *SID Symposium Digest of Technical Papers*, Vol. 36. 1112–1115.
- Ernst Simonson and Josef Brozek. 1952. Flicker fusion frequency. *Physiological Review* 32, 3 (1952), 349–378.
- Wen-Chih Tai, Chi-Chung Tsai, Shian-Jun Chiou, Chih-Ping Su, Huang-Min Chen, Chia-Lin Liu, and Chi-Neng Mo. 2008. Field Sequential Color LCD-TV Using Multi-Area Control Algorithm. *SID Symposium Digest of Technical Papers* 39, 1 (2008), 1092–1095.
- Masatsugu Teragawa, Akiko Yoshida, Kazuyoshi Yoshiyama, Shinji Nakagawa, Kazunari Tomizawa, and Yasuhiro Yoshida. 2012. Review Paper: Multi-primary-color displays: The latest technologies and their benefits. *Journal of the Soc. for Info. Disp.* 20, 1 (2012).
- Takatoshi Tsujimura. 2012. *OLED Display Fundamentals and Applications*. Wiley.
- Rui Wang, Bowen Yu, Julio Marco, Tianlei Hu, Diego Gutierrez, and Hujun Bao. 2016. Real-time Rendering on a Power Budget. *ACM Trans. Graph.* 35, 4, Article 111 (July 2016), 11 pages.
- Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612.
- G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. 2012. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (SIGGRAPH)* 31, 4 (2012), 1–11.
- Lin Zhang, D. Zhang, Xuanqin Mou, and D. Zhang. 2011. FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Transactions on Image Processing* 20, 8 (2011), 2378–2386.
- X. Zhang and B. A. Wandell. 1997. A spatial extension of CIELAB for digital color-image reproduction. *Journal of the Society for Information Display* 5, 1 (1997), 61–63.