FactoredSweeper: Optical See-Through Display Integrating Light Attenuation and Addition with Single Spatial Light Modulator

Yuichi Hiroi* Cluster Metaverse Lab
Takefumi Hiraki† Cluster Metaverse Lab
Yuta Itoh‡ The University of Tokyo

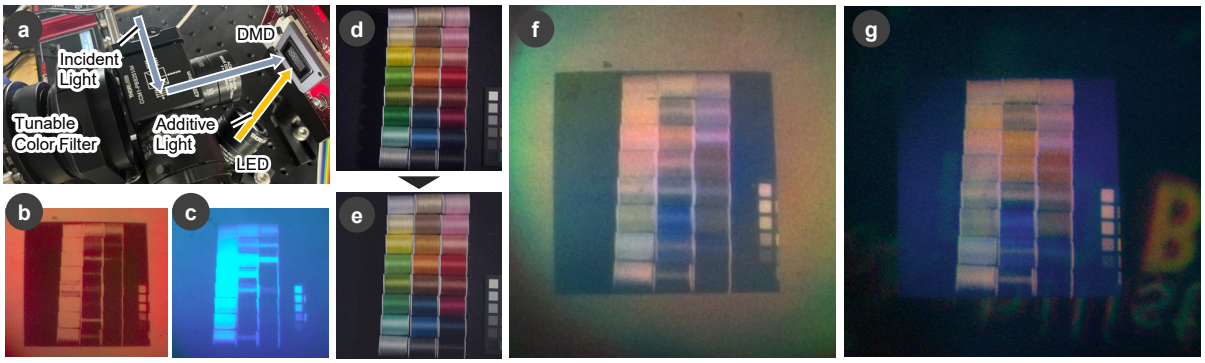


Figure 1: (a) Our FactoredSweeper prototype, which integrates color subtraction and addition of incident light into a single SLM by synchronizing a white LED, a DMD, and a polarizing tunable color filter. (b, c) Imperceptible time-multiplexed viewpoint images in (b) light subtraction and (c) addition modes. Humans perceive integrated color images through high-speed time-multiplexing. (d) Target multispectral image and (e) simulated rendering optimized by perceptually driven non-negative matrix factorization of time-multiplexing schedule and DMD images. (f, g) Viewpoint images with (f) white surface light source background and (g) real object background, demonstrating the system's ability to display both dark and light colors while compensating for the background spectrum.

Abstract

Light Attenuation Displays (LADs), a subset of Optical See-Through Head-Mounted Displays (OST-HMDs), enable image display in bright environments by filtering incident light at the pixel level. Although recent methods have proposed single-DMD light attenuation, they do not consider additive color display and background compensation, limiting their applicability in real-world scenarios. We present FactoredSweeper, a single digital micromirror device (DMD) system that incorporates both light attenuation and addition. By synchronizing the DMD, color filter, and light source, our system generates an additive virtual image, light attenuation, and occlusion through time multiplexing. To produce the target image while compensating for the background, we optimize time-multiplexed binary DMD patterns and LED/color filter schedules using perceptually-driven non-negative matrix factorization. Simulations and prototypes demonstrate that our integrated attenuation-addition single-SLM system achieves superior dynamic range and perceptual image quality compared to conventional occlusion-capable OST-HMDs using grayscale occlusion masks.

Keywords: Light attenuation display, see-through display, augmented reality, time-multiplexing, perceptual-driven matrix factorization

1 Introduction

Head-mounted displays (HMDs) are essential for the seamless integration of real and virtual environments in augmented reality (AR), enabling natural interaction with digital content. Optical see-through HMDs (OST-HMDs) provide direct access to rich light information from the real world while overlaying digital content, unlike video see-through HMDs that view the world through a camera feed. However, the additive approach used by OST-HMDs to blend virtual content using optical elements such as beam combiners results in limited depth cues and a narrow dynamic range due to the inability to occlude real-world light.

Occlusion-capable OST-HMDs (OCOST-HMDs) [29] address this issue by incorporating spatial light modulators (SLMs), such as digital micromirror devices (DMDs) or liquid crystal on silicon (LCoS), as occlusion masks to control incident light at the pixel level. However, OCOST-HMDs typically require two SLMs for simultaneous light addition and occlusion, complicating the optical system (Fig. 2, a).

To reduce the form factor, single-DMD OCOST-HMDs have been proposed, generating both an occlusion mask and a digital image by optimizing time-multiplexed DMD and LED patterns to approximate a target AR scene (Fig. 2, b) [25, 28, 30]. Although these approaches optimize the trade-off between occlusion mask accuracy and digital image color fidelity, the monochrome occlusion mask limits color reproduction. However, OCOST-HMDs have limited flexibility in color reproduction due to their inherent characteristic of attenuating all light equally with the occlusion mask and then compensating for color using an additive display.

Light Attenuation Displays (LADs) [24, 26], a type of OCOST-HMD, provide multi-color occlusion using SLMs as pixel-wise color filters, similar to stained glass. However, LADs have not been combined with additive displays to produce images brighter than the incident light. A recent single-SLM LAD design synchronizes a uniform color filter with a DMD to optimize time-multiplexing for displaying target images darker than reality [17]. We consider that integrating a time-multiplexing schedule with LED patterns into this single-SLM LAD would enable both addition and subtraction of color images using a single SLM, resulting in an OST-HMD with higher color fidelity and dynamic range.

In this paper, we propose FactoredSweeper, a single-SLM LAD

*e-mail: y.hiroi@cluster.mu
†e-mail: t.hiraki@cluster.mu
‡e-mail: yuta.itoh@iii.u-tokyo.ac.jp

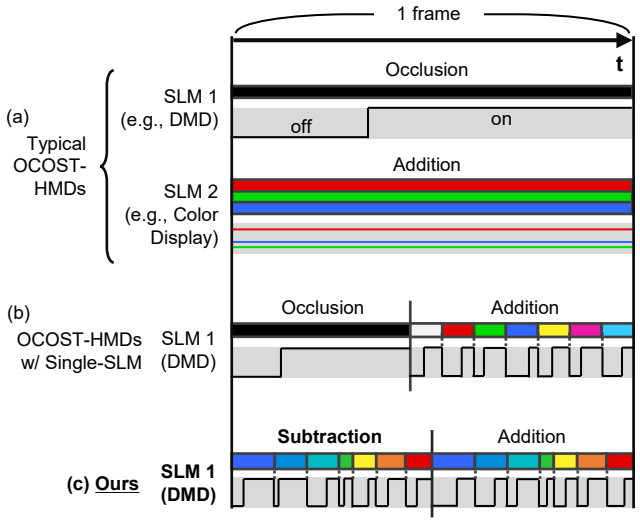


Figure 2: Per-pixel comparison of our system with other OCOST-HMDs. Gray area graphs represent SLM states. (a) Conventional OCOST-HMDs [29]: two SLMs for occlusion and addition. (b) Single-SLM OCOST-HMDs [25, 28, 30]: synchronized DMD and color light sources for monochromatic occlusion and addition. (c) Our system: synchronized DMD, white light source, and color filter for color subtraction and addition.

that achieves both color subtraction and addition by synchronizing a white LED, DMD, and non-pixel color filters. By only requiring the addition of a single white LED to the existing single-SLM optical system, our system maintains a compact form factor while improving color reproduction and background compensation compared to conventional occlusion methods in single-SLM configurations.

To realize FactoredSweeper, we propose a see-through optical design that enables switching between incident light subtraction and LED-driven virtual image addition (Fig. 2, c). To display the target image while compensating for its background, we also formulate the time-multiplexed image formation and optimize the schedule of each device and DMD images using an extended form of perceptually-driven non-negative matrix factorization (PNMF) [27]. Finally, we demonstrate a proof-of-concept prototype that displays color images in bright environments with color compensation for see-through views. Our main contributions include:

- Introducing FactoredSweeper, a single-SLM, variable-intensity LAD that enables both color addition and subtraction by synchronizing a DMD, LED, and color filters.
- Developing perception-driven factorization algorithms for optimizing DMD and color filter states to render target virtual images with background compensation.
- Implementing a proof-of-concept prototype that displays color images in both bright and dark environments with see-through color compensation.
- Evaluating the dynamic range and image quality of simulations and prototypes, discussing limitations and future research directions.

2 Related Work

2.1 OCOST-HMDs and LADs

OCOST-HMDs enhance the dynamic range of OST-HMDs by attenuating incident light [9] and providing depth cues [6] to virtual images. To selectively occlude incident light, an SLM is inserted as an occlusion mask between the optical combiner and the eye [36, 37]. However, this introduces a focal depth discrepancy between the mask and virtual image, resulting in a blurred (soft-edge) mask. Itoh et al. proposed to compensate soft-edge masks using scene camera captures [23], while it requires complex calibration.

Various hard-edge OCOST-HMDs have been proposed using relay optics to match the focal depth of the virtual image and mask [3, 4, 13, 14, 29, 45, 48, 53, 54, 57]. While hard-edge occlusion enhances perceptual reality, it typically requires two SLMs (one for addition, one for occlusion) and focusing optics, resulting in bulky designs. Yamaguchi and Takaki minimized form factor with a light-field OCOST-HMD [55], but severely limited the viewing angle to 4.3 °. Recent varifocal OCOST-HMDs [15, 19, 44] aim to mitigate vergence-accommodation conflict (VAC) [21], the inconsistency between eye convergence and focal depth in HMDs, but require multiple SLMs and complex mechanics/optics, increasing form factor.

Single-SLM OCOST-HMDs have been proposed to reduce form factor by time-multiplexing the virtual image and real scene using a DMD [25, 28, 30]. Krajancich et al. optimized the DMD/LED time-multiplexing schedule using matrix factorization to display color-compensated images [30], but with highly tilted incident light. Ju et al. proposed a DMD-based OCOST-HMD with a pupil-matched optical system [25], but without see-through color compensation. Our system is inspired by the former for optimization and the latter for optics. Other single-SLM OCOST-HMD approaches include varifocal designs using Pancharatnam-Berry lenses and phase-only SLMs [16]. and photochromic occlusion masks formed by ultraviolet projection via wavelength-selective holographic optical elements (HOEs) and DMDs [41].

LADs attenuate incident light using multispectral spatial color filters [24, 26], potentially improving dynamic range, power consumption, and color reproduction compared to conventional OCOST-HMDs. A recent single-SLM LAD synchronizes a DMD with uniform color filters [17]. We extend this approach by time-multiplexing LED control to enable additive virtual images alongside color subtraction. Moreover, we formulate and optimize spatially-varying incident light modulation to form images, reproduce arbitrary colors, and compensate for see-through views at the multispectral level.

2.2 Factored Displays

Many computational displays factorize images into non-negative matrices to reproduce desired virtual image properties [38], such as improved resolution [20, 33], dynamic range [5, 20, 46], varifocal and light field reproduction [20, 33, 47, 51, 52], and extended color gamut [27]. In HMDs, Huang et al. implemented a light-field VR-HMD using tensor factorization [22], and Mercier et al. extended this to eye tracking with interactive matrix factorization for multifocal displays [40]. Recently, Ebner et al. applied non-negative matrix decomposition to an off-axis varifocal display combining a direct view display and an OST-HMD [8].

The single-SLM OCOST-HMD by Krajancich et al. [30] uses factorization to mix the occlusion mask and virtual images for desired image display. We extend their approach to multispectral color filters, including background compensation. We further extend the perceptual factorization [27] in multispectral projectors to LAD for perceptual color rendering. Although our formulation yields a nonlinear and discontinuous matrix factorization, we demonstrate sufficient optimization using automatic differentiation and gradient descent.

2.3 Background Compensations of OST-HMDs

To mitigate the narrow dynamic range of virtual images in OST-HMDs, studies have investigated background detection and compensation without occlusion masks. Colorimetry [50] and radiance [31] approaches have been proposed as compensation metrics, but are limited to physical properties without considering human perception.

Fukiage et al. developed a metric integrating the visibility of image features (spatial frequency, color, etc.) through psychophysical experiments and applied it to alpha blending [10, 11]. Zhang et al. incorporated simultaneous color induction, where surrounding color differences affect target color perception, into their model to improve OST-HMD alpha-blending [57].

While OCOST-HMDs have proposed the use of occlusion masks for background compensation [18, 23, 30], our method is the first to achieve this in LADs. In addition, for the first time, we introduce the non-linearity of human perception into the background compensation process for OST-HMDs using perceptual color difference.

3 FactoredSweeper

3.1 Optical Design

Figure 3 illustrates the schematic of our FactoredSweeper system, which enables both addition and subtraction of incident light on a pixel-by-pixel basis by synchronizing a color filter, a white-light source, and a DMD. We first describe the overall optical path of the see-through view (Fig. 3a) and then detail the switching between each mode (Fig. 3b-d).

3.1.1 System Overview

Our system realizes a see-through configuration using a polarizing beam splitter (PBS) to fold the optical path back and sequentially introduce light into the DMD and the color filter. A PBS splits unpolarized incident light into p- and s-polarized light, with the p-polarized light passing through the PBS and the s-polarized light being reflected.

The DMD, an array of high-density micromirrors, can individually flip mirrors between two states at < 20 kHz. Each micromirror has a binary $\pm 12^\circ$ ON/OFF state, but time-multiplexing beyond human perception can create grayscale intensities, and synchronization with colored light sources enables color display. In our system, the DMD is angled such that ON pixels reflect light perpendicular to the system, while OFF pixels tilt 24 $^\circ$ towards the LED. This configuration allows switching between see-through and light-adding modes.

An achromatic $\lambda/4$ plate and lens are placed before the DMD. The lens focuses light from infinity onto the DMD, while the achromatic $\lambda/4$ plate converts s-polarized light to p-polarized light for transmission through the PBS. This waveplate should be achromatic to minimize light attenuation.

Tunable color filters, such as a Lyot [35] or Solc [49] filter with tunable liquid crystal retarders, are used to transmit light of specific wavelengths in the visible band. These filters must be fast-responding to synchronize with other equipment. Note that this folded optical design, based on [17], maintains image magnification but shifts the viewpoint along the depth axis.

3.1.2 Switching Display Mode

We describe how to add and subtract colors to and from the background while switching between multiple modes in this system.

Color Filter Mode When the micromirror is turned on, incoming light is reflected vertically and the pixel is in see-through mode regardless of LED status. Fig. 3a shows the optical path: (1) unpolarized incident light becomes s-polarized through a 0° linear polarizer, reflects off the PBS, and enters the $\lambda/4$ plate; (2) the light becomes circularly polarized, reflects off the DMD, and passes through the $\lambda/4$ plate again; (3) the light becomes p-polarized, passes through the PBS, and reaches the color filter; (4) the filter emits a mixture of p- and s-polarized light; (5) only s-polarized light reflects off the PBS and reaches the eye. As a result, the eye sees colored light as the product of the incident light and the color filter's spectral distribution (Fig. 3b).

Occlusion Mode When both LED and DMD are off, the system is in occlusion mode (Fig. 3c). Incident light reflects towards the LED and does not reach the eye, modulating the light intensity.

Addition Mode When the LED is on and DMD is off, LED light reaches the eye instead of the see-through view, displaying a virtual image (Fig. 3d). White LED light passes through the polarized color filter, and synchronization with the DMD displays a color image.

3.2 Image Formation

Next, we describe the image formation in FactoredSweeper. Fig. 4 shows time-multiplexing at a given pixel in our system. The image is formed by time-multiplexing M color filters $\psi_m(\lambda)$ for each LED on/off pattern, resulting $2M$ slots per unit time. The LED is off for $m = 1 \dots M$ and on for $m = M+1 \dots 2M$. Let g_m be the slot duration per unit time and $h_m(x)$ be the time ratio that the DMD is on in each duration, both constrained to [0, 1]. $r(\lambda, x)$ represents the incident light spectrum at position x, and $l(\lambda)$ be the spectrum of the LED.

From Fig. 4, when the LED is on ($m = M+1 \dots 2M$), the light from LED $l(\lambda)$ reaches the eye if the DMD is off, while $r(\lambda)$ reaches the eye if the DMD is on. Then the spectrum of the light $i(\lambda, x)$ passing through the system in unit-time can be formulated as

$$i(\lambda,x) = \sum_{m=1}^{M} g_m r(\lambda,x)\psi_m(\lambda)h_m(x) \tag{1}$$
$$+ \sum_{m=M+1}^{2M} g_m \psi_{m-M}(\lambda)\{l(\lambda)(1-h_m(x)) + r(\lambda,x)h_m(x)\}.$$

We discretize the wavelength of $i(\lambda, x)$ by dividing λ by L and x by the total number of pixels N. Then, the intensity of light $i_{l,n}$ at λ_l $(l = 1 \dots L)$ passing through pixel x_n $(n = 1 \dots N)$ is

$$\begin{aligned} i_{l,n} &= r_{l,n} \sum_{m=1}^{M} \psi_{l,m} g_m h_{m,n} \\ &\quad + \sum_{m=M+1}^{2M} \psi_{l,m-M} g_m \{l_l(1-h_{m,n}) + r_{l,n} h_{m,n}\}. \end{aligned} \tag{2}$$

To rewrite Eq. (2) as matrix operation, we define matrices as

$$\Psi = [\psi_{l,m}] \in \mathbb{R}^{L\times M},\ \Lambda_d = \mathrm{diag}([l_1 \cdots l_L]^\mathrm{T}),$$
$$\Psi_1 = [\Psi|\Psi] \in \mathbb{R}^{L\times 2M},\ \Psi_2 = [\mathbf{0}_{L\times M}|\Lambda_d\Psi] \in \mathbb{R}^{L\times 2M},$$
$$\mathbf{g} = [g_1 \cdots g_{2M}]^\mathrm{T} \text{ s.t. } \sum_{m=1}^{2M} g_m = 1 \text{ and } 0 \le g_m \le 1, \quad \mathbf{G} = \mathrm{diag}(\mathbf{g}),$$
$$\mathbf{R} = [r_{l,n}] \in \mathbb{R}^{L\times N}, \quad \mathbf{H} = [h_{m,n}] \in \mathbb{R}^{N\times 2M} \quad \text{s.t. } 0 \le h_{m,n} \le 1,$$

where Ψ_1, Ψ_2 are discrete color filter spectral matrices associated with the DMD and not, respectively; $\mathbf{G}$ is a diagonal matrix that encodes the duration of each slot $\mathbf{g}$; $\mathbf{H}$ encodes the DMD states for N pixels in each slot; and $\mathbf{R}$ represents the discretized spectral distribution of incident light at each pixel.

By rewriting Eq. (2) using these matrices, we obtain

$$\mathbf{I} = \mathbf{R} \odot (\Psi_1 \mathbf{G}\mathbf{H}^\mathrm{T}) + \Psi_2 \mathbf{G}(\mathbf{1}_{N\times 2M} - \mathbf{H})^\mathrm{T}, \tag{3}$$

where $\mathbf{I} \in \mathbb{R}^{L\times N}$ is a multispectral image with L spectral bands and N pixels.

To model the color image perceived by a human observer, we consider the spectral sensitivity of retinal receptor cells, generally modeled by the CIE 1931 color-matching functions $\mathbf{p}_{(\bar{x},\bar{y},\bar{z})}(\lambda) = [\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)]^\mathrm{T}$ Using $\mathbf{p}_{(\bar{x},\bar{y},\bar{z})}(\lambda)$, the spectral distribution $i(\lambda, x)$ is converted to CIE XYZ color space $\mathbf{i}_{XYZ} \in \mathbb{R}^3$ as

$$\mathbf{i}_{XYZ} = \int \mathbf{p}_{(\bar{x},\bar{y},\bar{z})}(\lambda) i(\lambda, x) \mathrm{d}\lambda. \tag{4}$$

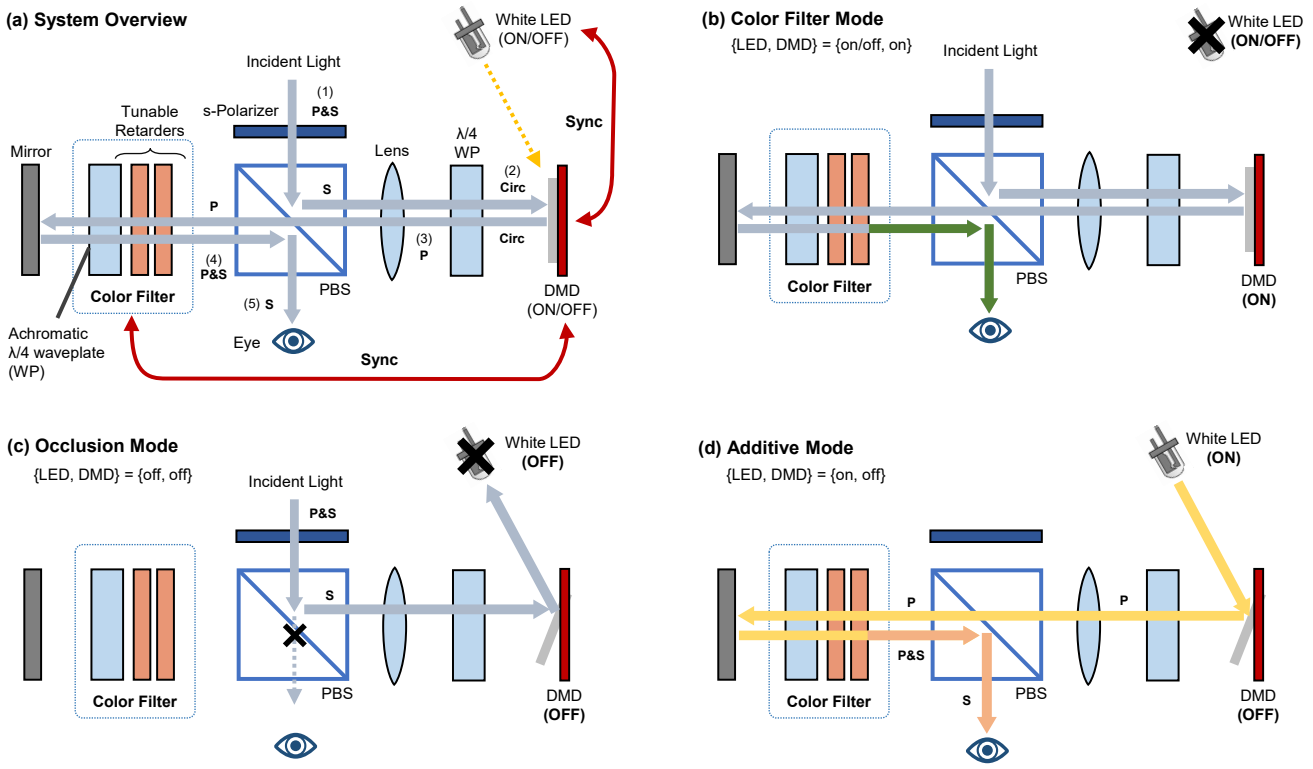


Figure 3: (a) FactoredSweeper system overview. Gray arrows show real-world light path, and dashed yellow arrow shows LED light path. Color filter, DMD, and LED are synchronously controlled. (b) In color filter mode (DMD on), real-world light is colored through the filter, regardless of the LED status. (c) In occlusion mode (both LED and DMD are off), no incident light reaches the eye. (d) In additive mode (LED is on and DMD is off): real-world light is blocked and colored LED light reaches the eye.

Let $\mathbf{P}_{XYZ} \in \mathbb{R}^{3\times L}$ be the matrix discretizing $\mathbf{p}_{(\bar{x},\bar{y},\bar{z})}(\lambda)$ with respect to λ. Integrating Eq. (3) and Eq. (4), we have

$$\mathbf{I}_{XYZ} = \mathbf{P}_{XYZ}(\mathbf{R} \odot (\Psi_1 \mathbf{G}\mathbf{H}^{\mathrm{T}}) + \Psi_2 \mathbf{G}(\mathbf{1}_{N\times 2M} - \mathbf{H})^{\mathrm{T}}) \tag{5}$$

where $\mathbf{I}_{XYZ} \in \mathbb{R}^{3\times N}$ is $\mathbf{i}_{XYZ}$ at all pixels N.

3.3 Perceptually-optimal Image Generation

This section describes the optimization of the time-multiplexed schedule $\mathbf{g}$ and the DMD display image $\mathbf{H}$ from the target color image $\mathbf{I}_{XYZ}$, considering Eq. (5) for image formation in this system.

In multispectral DLP, Kauver et al. proposed perceptually-driven non-negative matrix factorization (PNMF) [27], which optimizes $\mathbf{g}$ and $\mathbf{H}$ when mixing multiple light sources with different spectra to achieve perceptually close colors between the target and projected images. Inspired by their method, we optimize the non-negative matrices $\mathbf{g}$ and $\mathbf{H}$ based on perceptually accurate color differences in the CIE L*a*b* color system.

However, our problem is more complex than that of Kauver et al.; in particular, our problem involves (1) crosstalk term $\mathbf{P}_{XYZ}\Psi_2\mathbf{G}(\mathbf{1}_{N\times 2M} - \mathbf{H})^{\mathrm{T}}$ between the incident light and the light source, and (2) spatially varying real light $\mathbf{R}$ as part of the incident light. While [27] explicitly formulated the differential form of the objective function and optimized with the Newton method, we confirmed that our more complex objective function can be optimized using automatic differentiation [1] and gradient descent optimization [34], which have been developed in recent years. With their help, we performed our PNMF optimization using CIEDE2000 color differences that were difficult to obtain explicitly in differential form.

We first explain the design of the objective function in CIE XYZ. For simplicity, we introduce an intermediate variable $\mathbf{X}(\mathbf{g},\mathbf{H}) \in \mathbb{R}^{3\times N}$ as

$$\mathbf{X}(\mathbf{g},\mathbf{H}) = \mathbf{P}_{XYZ}(\mathbf{R} \odot (\Psi_1 \mathbf{G}\mathbf{H}^{\mathrm{T}}) + \Psi_2 \mathbf{G}(\mathbf{1}_{N\times 2M} - \mathbf{H})^{\mathrm{T}}). \tag{6}$$

Based on Eq. (5) and (6), we define the objective function that minimizes the residual between the target and reproduced colors in the XYZ color space as

$$\begin{aligned} &\underset{\{\mathbf{g},\mathbf{H}\}}{\text{minimize}} \; \|\beta \mathbf{I}_{XYZ} - \mathbf{X}(\mathbf{g},\mathbf{H})\|_F^2 \\ &\text{s.t.} \sum_{m=1}^{2M} g_m = 1, \;\; 0 \le g_m, h_{m,n} \le 1; \;\; \forall m,n. \end{aligned} \tag{7}$$

where $0 \le \beta$ is a hyperparameter determining the brightness of the reproduced color relative to the target color.

To remove the linear constraint on the sum of g_m, we use an intermediate vector $\Gamma = [\gamma_1 \dots \gamma_{2M}]^{\mathrm{T}}$ and redefine $\mathbf{g}(\Gamma)$ as a softmax function of Γ:

$$g_m(\Gamma) = \frac{\exp(\gamma_m)}{\sum_{k=1}^{2M} \exp(\gamma_k)}. \tag{8}$$

By Eq. (8), Eq. (7) is modified to optimize Γ instead of $\mathbf{g}$:

$$\underset{\{\Gamma,\mathbf{H}\}}{\text{minimize}} \; \|\beta \mathbf{I}_{XYZ} - \mathbf{X}(\mathbf{g}(\Gamma),\mathbf{H})\|_F^2, \;\; \text{s.t.} \;\; 0 \le h_{m,n} \le 1; \;\; \forall m,n. \tag{9}$$

In the following, we denote $\mathbf{X}(\mathbf{g}(\Gamma),\mathbf{H}) = \mathbf{X}(\Gamma,\mathbf{H})$ if there is no description.

We then reformulate the objective function using perceptual color difference metrics. After converting the objective function (Eq. (9))

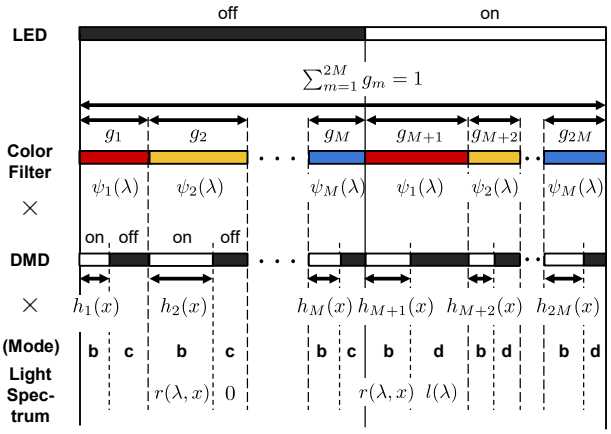


Figure 4: Time-multiplexed LED, color filter, and DMD schedule at a pixel within a unit time. The "mode" row represents the modes in Fig. 3. For M color filters, the unit time is divided into $2M$ slots based on LED binary states. Within each slot, the DMD binary state is switched.

from the CIE XYZ to CIE L*a*b*, we optimize it to account for the perceptual color metrics defined by CIE.

The function $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ that converts color $\mathbf{x}_{XYZ} = [x_X, x_Y, x_Z]^{\mathrm{T}}$ in CIE XYZ to color $\mathbf{x}_{Lab} = [x_{L^*}, x_{a^*}, x_{b^*}]^{\mathrm{T}}$ in CIE L*a*b* can be written as:

$$\mathbf{x}_{Lab} = \varphi(\mathbf{x}_{XYZ}) = \begin{bmatrix} 116\phi\left(x_Y/w_Y\right) - 16 \\ 500(\phi\left(x_X/w_X\right) - \phi\left(x_Y/w_Y\right)) \\ 200(\phi\left(x_Y/w_Y\right) - \phi\left(x_Z/w_Z\right)) \end{bmatrix} \tag{10}$$

where $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\phi(x) = \begin{cases} x^{1/3} & \text{if } x > (6/29)^3, \\ (1/3)(29/6)^2 x + (4/29) & \text{otherwise.} \end{cases} \tag{11}$$

CIE76 ΔE_{76} is the first color difference formula defined and calculated as the Euclidean distance in CIEL*a*b* coordinates. With ΔE_{76}, we rewrite Eq. (9) as

$$\begin{aligned} & \underset{\{\Gamma,\mathbf{H}\}}{\text{minimize}} \, \|\Delta E_{76}(\beta \mathbf{I}_{Lab}, \varphi(\mathbf{X}(\Gamma,\mathbf{H})))\|_F^2 \\ = \; & \underset{\{\Gamma,\mathbf{H}\}}{\text{minimize}} \, \|\beta \mathbf{I}_{Lab} - \varphi(\mathbf{X}(\Gamma,\mathbf{H}))\|_F^2, \\ & \text{s.t. } 0 \le h_{m,n} \le 1; \; \forall m,n. \end{aligned} \tag{12}$$

where $\mathbf{I}_{Lab} \in \mathbb{R}^{3\times N}$ is the target image in CIE L*a*b*.

CIE76 ΔE_{76} overestimates color differences in highly saturated regions. Subsequent improvements, such as CIE94 ΔE_{94} and CIEDE2000 ΔE_{00} (see Supp. Sec. 1.1 and 1.2), account for perceptual nonlinearity, blue color discrimination, hue rotation, and saturation and luminance corrections. Using ΔE_{00}, the perceptually-driven objective function for time-multiplexing matrix optimization is described as

$$\begin{aligned} \underset{\{\Gamma,\mathbf{H}\}}{\text{minimize}} \quad & \|\Delta E_{00}(\beta \mathbf{I}_{Lab}, \varphi(\mathbf{X}(\Gamma,\mathbf{H})))\|_F^2 \\ \text{s.t.} \quad & 0 \le h_{m,n} \le 1; \; \forall m,n. \end{aligned} \tag{13}$$

Note that as with [17], the see-through color is affected by individual filters, but we have optimized the duration of $\mathbf{g}$ to achieve uniform transmittance across all wavelengths. This reduces the overall transmittance without changing the color.

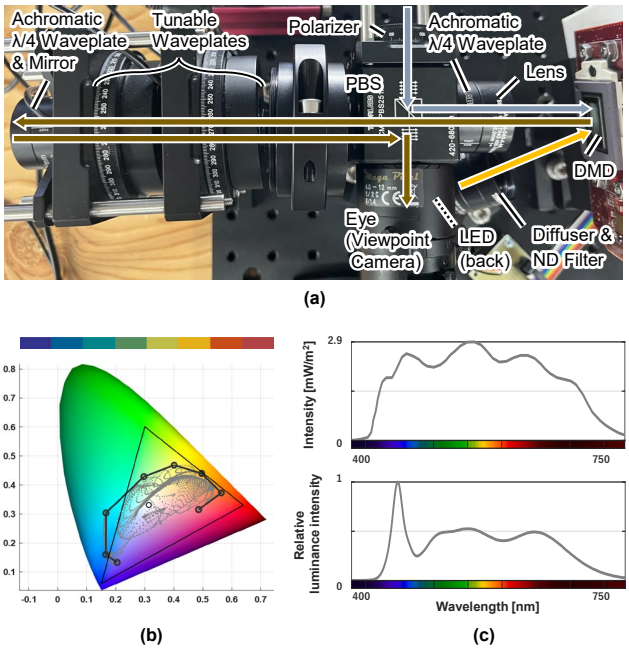


Figure 5: (a) Experimental hardware setup, schematic identical to Fig. 3a. (b, top) Selected primary colors ($M = 8$) (b, bottom) The primary colors (black line) and all color filter distribution (gray, 0.08 V steps) in xy chromaticity diagram. (c, top) White LED spectrum (blue) with wavelength (x-axis) and irradiance [mW/m²] (y-axis) measured at DMD through ND filter. (c, bottom) Reference light spectrum for white background. Since this LED can change its intensity in 40 steps from 100 mW to 4000 mW, the figure shows the relative intensity.

4 Implementation

4.1 Hardware Setup

Figure 5 (a) shows our hardware setup. We implemented a reflective Solc filter-based LAD following [17] and added a multi-band white LED beside the DMD.

For the reflective Solc filter, we used two Thorlabs LCC1421-A tunable retarders, which can change phase retardation $\psi(\lambda, V)$ by varying the input voltage V. Each tunable waveplate is mounted on a Thorlabs LCRM2/M rotary mount with $\pm 11.25^\circ$ rotation and connected to a Thorlabs KLC101 liquid crystal controller. The output voltage of the controller (0 to ± 25 Vrms) is set by an external 0-5 VDC signal at the "MOD IN" connector.

We determined the primary color filters $\psi_m(\lambda)$ according to [17]. We placed a white reference light (Thorlabs SOLIS-3C) in front of the system and measured the spectrum passing through with a spectrometer (UPRTek MK350N Premium) while varying the retarder voltages from 0-8 V in 0.08 V increments. Then these measured color filters were divided by the measured spectrum by the spectral distribution of the light source. After converting this color filter to an xy chromaticity diagram, a filter with $M = 8$ is selected to maximize the convex hull. As a result, we obtain the primary color filters as shown in Fig. 5 b.

We used a Texas Instruments DLP LightCrafter 6500 (1920×1080 pixels) for the DMD and controller. We also used the white LED (Toyota Gosei ENB01-NHSD7-F1) to reproduce natural sunlight. We used this LED as a 3x5 LED array as a surface light with a Thorlabs DG10-220-MD diffuser. An achromatic lens (Thorlabs AC254-045-A) focuses the LED light on the DMD, and an ND filter (Thorlabs NE13A) maintains brightness balance with the reference light (Fig. 5c top). This entire assembly was mounted on a Thorlabs TRB1/M ball-socket mount to illuminate the DMD at an angle.

A uniform white reference light (Thorlabs SOLIS-3C) projected

onto diffusion paper served as the background for optimization and measurement. Its brightness is adjustable in 40 steps, and its spectrum was pre-measured with a spectrometer (Fig. 5c bottom).

Other components were from Thorlabs: LPVISE100-A for linear polarizers, AQWP10M-580 for the achromatic $\lambda/4$ waveplates, PBS251 for the PBS, BB1-E02 for the mirror, AC508-075-A for the lens in front of the Solc filter, and AC254-045-A for the lens in front of the DMD. We used a color camera (Basler a2A2590-60ucBAS, resolution 2592 $\times$ 1944 pixels) with a varifocal lens (TAMRON M12VM412) as a user-perspective camera.

4.2 Synchronous Control of Devices

To function as a high-speed time-division multiplexed display for light attenuation and addition, the DMD, the Solc filter, and the white LED require synchronization.

The previous LAD [17] only aimed at attenuating light, leveraging the DMD controller's ability to output $2^3 = 8$ patterns synchronously with the DMD for 3ch (RGB) LED control. This allowed the DMD controller to control the Solc filter via microcontroller and display eight primary colors. However, our system requires twice as many control patterns for color attenuation and addition. Using the existing framework would limit us to four primary colors, reducing the color gamut.

Therefore, we developed a board and software framework where the DMD controller sends a trigger signal to a microcontroller (PJRC Teensy 4.0), which controls the Solc filter and illumination LED. The microcontroller listens for the trigger signal from the DMD controller and detects the rising edge of the signal to determine the state change of DMD. Then, the microcontroller references internal settings, and controls the Solc filter by changing the LC controller voltages via a DAC (MCP4725A0T-E/CH, Microchip Technology) to display corresponding primary colors. It also controls the white LED's ON/OFF state according to these settings.

We defined a control program structure consisting of the setting name, LC controller voltage values, and LED ON/OFF state for primary color display and color addition/subtraction. Control patterns are expressed as an array of these structures, switched by an index array defining the invocation order based on the trigger signal.

This framework allows developers to use projector development software (DLPC900 GUI, Texas Instruments) for DMD pattern switching and trigger signal setup, then configure synchronous control sequences for the DMD, Solc filter, and LED by writing control pattern structure arrays and call index sequences in the microcontroller program. This simple extension of projector development software can be used comfortably by researchers and engineers experienced in DLP projector design, increasing the number of synchronously controlled patterns from 8 to several dozen, enabling color representation with our system.

Throughout the process, 40 ms (25 fps) was set as the frame unit time, with hardware controlled for time-multiplexing within that range. Optimized DMD images were downsampled to 4-bit color depth (16 shades) to ensure DLP minimum exposure times did not exceed the optimized schedule's minimum slot.

4.3 Datasets and PNMF Optimization

We used the CAVE dataset [42], a 31-band, 512×512 px multispectral dataset acquired in 10 nm increments from 400-700 nm across 32 scenes. We used the same 11 scenes as [27] for target images.

As a preprocessing step, the spectral distribution of each multispectral image was interpolated in 1 nm steps for all pixels and converted to CIEXYZ $\mathbf{I}_{XYZ}$. CIEXYZ must be normalized so that the maximum value is 100, so we normalized each multispectral image by considering the white area of the color chart as $Y = 100$. We also normalized the CIEXYZ of the light source we used, since the measured spectrum and the LED spectrum were converted to XYZ and normalized so that the brighter Y value was $Y = 100$. All calculations were performed under D65 illumination, CIE 1931 2° color observation function, and the Rec. 709 standard white point.

We implemented the PNMF optimization of Eq. (13) using PyTorch with the AdamW optimizer (lr=1e-2, 1000 iterations) on an Intel i7-13700F 2.10 GHz processor with 64 GB RAM and an NVIDIA GeForce RTX 3070.

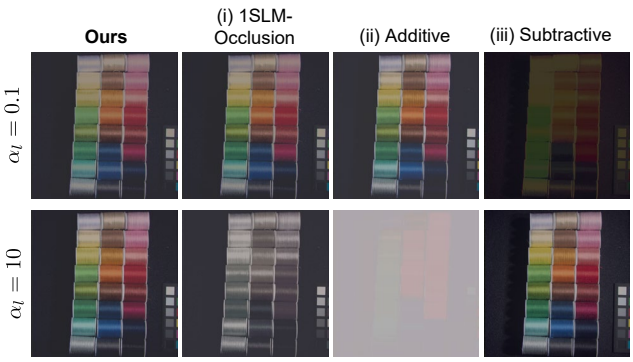


Figure 6: Examples of images optimized by each method by changing the brightness ratio α_l between backlight and LED light in simulation. (top) $\alpha_l = 0.1$ and (bottom) $\alpha_l = 10$.

5 Simulation

Before building the actual prototype, we evaluated the performance of our system by simulation, comparing it to (i) single-SLM OCOST-HMD, (ii) addition-only conventional OST-HMD, and (iii) subtraction-only single-SLM LAD with different background brightness. See Supp. Sec. 2 for optimization details of each comparison method.

5.1 Reproduced Image Quality with Wide Dynamic Range

Our system, which integrates both light addition and subtraction, can adaptively display images in both bright and dark backgrounds by optimizing the addition and subtraction ratio based on the background light brightness. To evaluate this feature, we simulated reproduced images at different background brightnesses by varying the real white light intensity $r(\lambda)$ for $\alpha_l r(\lambda)$ on logarithmic scales: $\alpha_l = 10^{-1}, 10^{-0.5}, 10^{0}, 10^{0.5}, 10^{1}$.

Fig. 6 shows an example of the optimized result image for each method. When $\alpha_l = 0.1$ (real light darker than LED), the subtraction-only image is almost black in (iii) the subtraction-only case. Also when $\alpha_l = 10$ (real light brighter than LED), (ii) the addition-only image is almost invisible due to whitening. Also in (i) single-SLM occlusion, colors are incorrectly reproduced because most of the time-multiplexing schedule is spent occluding real light in grayscale. In contrast, our method can reproduce stable and faithful images regardless of the reality being dark or bright compared to LED light.

To quantitatively analyze image quality, we converted optimized XYZ images to RGB for 11 multispectral images and evaluated them using image quality metrics (PSNR, SSIM, and LPIPS [56]). Fig. 7 shows the image metrics for each method for different α_l, which confirms that our method produces higher-quality images in both bright and dark environments than (ii) additive-only or (iii) subtractive-only systems. Compared to (i) single-SLM OCOST-HMD, both show high metrics with a wide dynamic range, but our method is quantitatively superior, especially in bright environments.

5.2 Optimization by Different Color Metrics

Next, we evaluated the effect of color differences used in the objective function on optimization convergence and image quality. We fixed $\alpha_l = 1$ and optimized Eq. (9) with four loss types: XYZ, $\Delta E_{76}, \Delta E_{94}$ and ΔE_{00}.

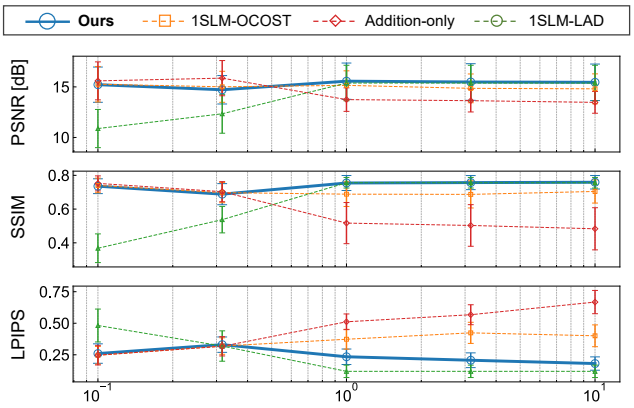


Figure 7: Comparison of PSNR, SSIM, and LPIPS when optimized with four different methods at different α_l. Higher PSNR and SSIM are better, and lower LPIPS is better.

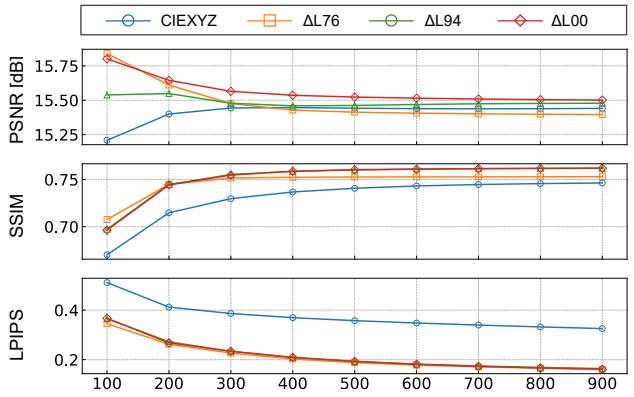


Figure 8: Average convergence trajectories of the image quality metrics (PSNR, SSIM and LPIPS) when we use the loss function as CIEXYZ, ΔE_{76}, ΔE_{94}, and ΔE_{00}. The x-axis is training iteration.

Fig. 8 shows the trajectory of each image quality metric when optimizing with different losses. From the result, optimization with CIEXYZ has the worst convergence and image quality. There is no overall difference between ΔE_{94} and ΔE_{00}; ΔE_{76} is inferior to ΔE_{94} and ΔE_{00} in PSNR and SSIM. Although ΔE_{00} theoretically improves perceptual accuracy over ΔE_{94} in grayscale and highly saturated blue regions, the difference is less significant for typical daily colors. Moreover, the luminance-matched target and optimized images (Fig. 6) likely result in small grayscale differences, leading to similar overall performance between ΔE_{94} and ΔE_{00}.

Average computation times for XYZ, ΔE_{76}, ΔE_{94}, and ΔE_{00} losses after 200 iterations were 11.0 s, 13.4 s, 14.4 s, 17.2 s, and after 1000 iterations were 49.3 s, 59.8 s, 64.3 s, and 77.2 s. Considering ΔE_{00} optimization took $\times 1.2$ longer than ΔE_{94} at 1000 iterations with only small image quality differences, using ΔE_{94} is sufficient for the optimization of this system.

Note that even with ΔE_{76} as the objective function, real-time processing has not been achieved. Future prospects for real-time processing are discussed in Sec. 7.

5.3 Background Compensation

Finally, we compared compensation performance when the background image is spatially variable. We selected a "cloth" scene with highly saturated red and complex texture as the background and a "thread spools" scene with multiple colors as the target. We fixed $\alpha_l = 1$ and optimized for 1000 iterations.

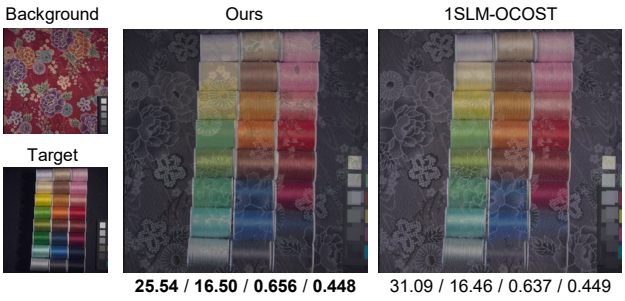


Figure 9: Simulation of background compensation for our system and single-SLM OCOST-HMD. The values below the image are the image perception metrics: ΔE_{00} per # of pixels / PSLM / SSIM / LPIPS.

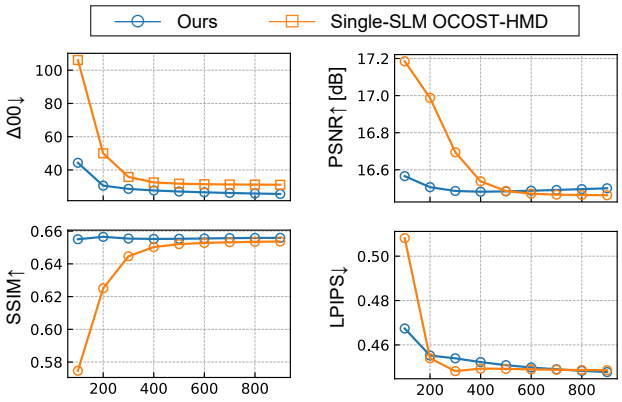


Figure 10: Converge trajectories of the image quality matrics, including (top-left) ΔE_{00} perceptional color difference per # of pixels, (top-right) PSNR, (bottom-left) SSIM, and (bottom-right) LPIPS. The arrows on y-axis indicate whether each metric is good when high or low.

Fig. 9 shows the results. Our method can compensate for the background with almost the same qualitative appearance as single-SLM OCOST-HMD, especially in suppressing texture patterns.

For quantitative evaluation, Fig. 10 shows the convergence trajectory of image quality metrics, and Fig. 9 (bottom) shows the final metrics. Our method is superior to the comparators on all metrics, albeit by a small margin. In particular, our method excels in perceptual color difference ΔE_{00} from the target image, indicating better compensation for background patterns.

6 Demonstration

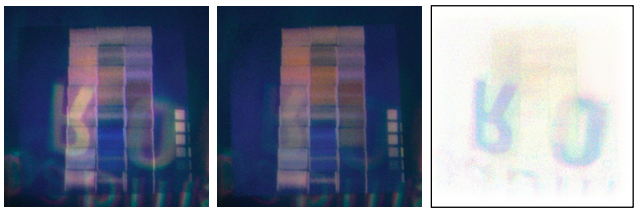
Using the optimized time-multiplexing schedule and DMD images, we displayed images on our proof-of-concept prototype.

6.1 Image Reproduction with White Background

We first demonstrated image reproducibility on a white background. A reference light was placed directly in front of our prototype, and DMD images and schedule were optimized for $\alpha_l = 10^{-0.5} = 0.32$. The reference light brightness was adjusted during measurement with a spectrometer to approach this α_l relative to the LED light.

Note that color differences from the optimized image exist due to difficulties in perfectly adjusting the viewpoint camera's white balance. Therefore, the discussion in this paper is mainly qualitative. However, using a multispectral camera with controllable exposure time as the viewpoint camera could enable more quantitative analysis based on actual spectral distributions in the future (Sec. 7).

Fig. 1 (f) shows viewpoint images through our prototype. From the images, our system can display both lighter and darker colors



(a) w/o compensation (b) w/ compensation (c) difference

Figure 11: (a) viewpoint camera image without compensation, (b) viewpoint camera image with optimization including background compensation, and (c) difference between (a) and (b). The absolute values of the differences are inverted for visualization.

than the background. However, saturated colors (green in the left column and red thread spools in the right column) were less saturated than in the simulation results. One reason for this difference between simulation and measurement is stray light when the LED is on, discussed in detail in Sec. 7.

6.2 Background Compensation

We next evaluated background compensation on our prototype. The background image $\mathbf{R}$ should ideally be acquired multi-spectrally by a coaxially installed multispectral camera. However, due to practical difficulties in obtaining a multispectral camera and form factor increase, we took color images directly from the viewpoint camera and used a deep learning-based approach [2] to estimate the multispectral distribution instead. The estimated spectral distribution was normalized so that the maximum luminance relative to the LED luminance was $\alpha_l = 10^{-0.5}$. Note that the real environment colors cannot be perfectly reproduced due to viewpoint camera color balance matching issues. Consequently, this multispectral estimation will be an approximation containing differences from the true spectrum.

Fig. 1 (g) and Fig. 11 show the viewpoint image with background. This figure demonstrates the ability of our method to selectively compensate for bright background regions when displaying images.

7 Discussion and Future Work

Based on our proof-of-concept prototype results, we discuss the remaining issues and future research directions for FactoredSweeper. Our hardware design, particularly the reflective Solc filter configuration, follows [17]. Therefore, discussions on hardware miniaturization and improvements to the color filter itself are deferred to that paper. Instead, this section mainly focuses on algorithm-related topics.

Faster PNMF Optimization While our optimization accounts for complex metrics and background compensation, running faster than previous work, it still does not operate in real-time. The main bottleneck is the computational complexity of the Hadamard product with the large matrix $\mathbf{R} \in \mathbb{R}^{L\times N}$ ($L = 301$, $N = 512^2$). As discussed in Sec. 6.2, defining the real-world background $\mathbf{R}$ multispectrally also requires capturing the scene with a multispectral camera. Therefore, approximating $\mathbf{R}$ in three-primary color spaces (e.g., XYZ, RGB) is crucial for hardware configuration and real-time performance. Although our implementation demonstrates optimization using a simple gradient method, matrix operations like Alternating Direction Method of Multipliers (ADMM) [12] with GPU support [43] may provide faster convergence for NMF optimization. Also, weighting objects based on attention level [30] or applying a convolutional neural network with an attention mechanism [7, 39] could address both speed and quality.

Stray Light Simulation The measured image approximates the target image's color and brightness, but does not match the simulation. Particularly, saturated colors were difficult to reproduce due to stray light when the LED is on. Our model (Fig. 4) assumes no mixing of LED light $l(\lambda)$ and real light $r(\lambda, x)$ when both LED and DMD are on. However, in reality, $l(\lambda)$ inevitably enters the DMD due to its tilt, and stray light should be formulated as $r(\lambda, x) + \alpha_{\text{stray}} l(\lambda)$ using the weight α_{stray}. Consequently, highly saturated colors ($h_k(x) = 0$ for filter $M+1 \le k \le 2M$ but $h_{m\neq k}(x) = 1$) experience reduced saturation. Since the DMD's slight tilt causes spatial variation in micromirror position x, accurate measurement requires calibrating spatial stray light contribution $\alpha_{\text{stray}}(x)$ at the viewpoint.

Chromatic Aberration Suppression The long polarized color filter sections and non-specialized doublet lenses increase chromatic aberration, affecting the final image appearance. Miniaturizing the reflective Solc filter could suppress chromatic aberration and improve the final appearance.

Trade-off between Occlusion and Color Reproduction Our system adaptively optimizes the time ratio between light addition and subtraction based on the background, similar to Krajancich et al. [30], which considers a balance between background compensation and color reproduction. On the other hand, a manual method of determining the ratio of subtraction to addition, as in Ju et al. [25]. By manually adjusting the ratio of subtraction to addition, it is possible to increase the darkness of the black, but the intensity of the additive image is reduced. This trade-off can be mitigated by increasing the LED light intensity.

Adaptive Dynamic Range We conducted experiments with the fixed parameter β determining the maximum luminance of the reproduced image. However, in a real environment, dynamically adjusting the dynamic range of the reproduced image based on incident light intensity and situation is desirable. Future research should include selecting the optimal brightness of the reproduced image according to the brightness of the background.

Varying LED Brightness Varying LED brightness can effectively display grayscale images within a reduced time frame, and our hardware enables this through pulse-width modulation. However, it requires a more complex formulation and optimization for both LED brightness and DMD control. Furthermore, varying LED brightness may make flicker more noticeable. Therefore, these factors must be considered when introducing LED brightness control.

Accommodating Diverse Perceptual Characteristics The proposed system may not effectively enhance the visibility of virtual images for people with color vision deficiencies or flicker sensitivity. Methods to compensate for these individual visual characteristics [32] to make backgrounds and displayed images easier to see are promising future research directions.

8 Conclusion

We introduced FactoredSweeper, a compact LAD that integrates light subtraction and attenuation in a single SLM by synchronizing the DMD, non-pixel polarized color filter, and LED. To realize FactoredSweeper, we proposed an optical system that switches between a color-filtered see-through view and a virtual image, formulated the perceived image, and developed a PNMF-based optimization algorithm. Simulations confirm that our system displays images with better color and brightness reproduction than conventional methods across a wide background dynamic range. Moreover, we implemented a proof-of-concept prototype demonstrating bright and dark pixels against a background and discussed remaining issues and research directions by comparing simulations and the prototype.

Our system is a step toward expanding OST-HMD expressiveness, but hardware and optimization algorithms can still be improved. We hope this study encourages researchers to explore further LAD design possibilities and realize multispectral virtual visual appearances indistinguishable from the real world using OST-HMDs.

ACKNOWLEDGMENTS

This project was partially supported by JST FOREST Grant Number JPMJFR206E, JST ASPIRE Grant Number JPMJAP2327, and JSPS KAKENHI Grant Number JP22KJ0753, JP23KJ16920, and JP23H04328 Japan.

REFERENCES

[1] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. Automatic differentiation in machine learning: a survey. *J. Mach. Learn. Res.*, 18(1):5595–5637, jan 2017. 4

[2] Y. Cai, J. Lin, Z. Lin, H. Wang, Y. Zhang, H. Pfister, R. Timofte, and L. V. Gool. Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In *IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2022. 8

[3] O. Cakmakci, Y. Ha, and J. P. Rolland. A compact optical see-through head-worn display with occlusion support. In *3rd IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, pages 16–25, 2004. 2

[4] M. Chae, K. Bang, Y. Jo, C. Yoo, and B. Lee. Occlusion-capable see-through display without the screen-door effect using a photochromic mask. *Opt. Letters*, 46(18):4554–4557, 2021. 2

[5] J.-H. R. Chang, B. V. K. Vijaya Kumar, and A. C. Sankaranarayanan. 2^{16} shades of gray: high bit-depth projection using light intensity control. *Opt. Express*, 24(24):27937–27950, Nov. 2016. 2

[6] J. E. Cutting and P. M. Vishton. Chapter 3 - perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth*. In W. Epstein and S. Rogers, editors, *Perception of Space and Motion*, Handbook of Perception and Cognition, pages 69–117. Academic Press, San Diego, 1995. 2

[7] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations (ICLR)*, 2021. 8

[8] C. Ebner, P. Mohr, T. Langlotz, Y. Peng, D. Schmalstieg, G. Wetzstein, and D. Kalkofen. Off-axis layered displays: Hybrid direct-view/near-eye mixed reality with focus cues. *IEEE Trans. Vis. Comput. Graph.*, 29(5):2816–2825, 2023. 2

[9] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Exploring the limitations of environment lighting on optical See-Through Head-Mounted displays. In *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*, number Article 9 in SUI '20, pages 1–8, New York, NY, USA, Oct. 2020. Association for Computing Machinery. 2

[10] T. Fukiage and T. Oishi. A content-adaptive visibility predictor for perceptually optimized image blending. *ACM Trans. Appl. Percept.*, 20(1), jan 2023. 3

[11] T. Fukiage, T. Oishi, and K. Ikeuchi. Visibility-based blending for real-time applications. In *2014 IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, pages 63–72, 2014. 3

[12] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers and Mathematics with Applications*, 2(1):17–40, 1976. 8

[13] C. Gao, Y. Lin, and H. Hua. Occlusion capable optical see-through head-mounted display using freeform optics. In *11th IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, pages 281–282, 2012. 2

[14] C. Gao, Y. Lin, and H. Hua. Optical see-through head-mounted display with occlusion capability. In *Proc. SPIE*, volume 8735, pages 87350F–1:9, 2013. 2

[15] T. Hamasaki and Y. Itoh. Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask. *IEEE Trans. Vis. Comput. Graph*, 25(5):1961–1969, 2019. 2

[16] W. Han, J.-W. Lee, J.-Y. Shin, M.-H. Choi, H.-R. Kim, and J.-H. Park. Varifocal occlusion in an optical see-through near-eye display with a single phase-only liquid crystal on silicon. *Photon. Res.*, 12(4):833–853, Apr 2024. 2

[17] Y. Hiroi, T. Hiraki, and Y. Itoh. Stainedsweeper: Compact, variable-intensity light-attenuation display with sweeping tunable retarders. *IEEE Trans. Vis. Comput. Graph.*, preprint:1–11, 2024. 1, 2, 3, 5, 6, 8

[18] Y. Hiroi, Y. Itoh, T. Hamasaki, and M. Sugimoto. Adaptivisor: Assisting eye adaptation via occlusive optical see-through head-mounted displays. In *8th Augmented Human International Conference*, pages 9:1–9:9, New York, NY, USA, 2017. ACM. 3

[19] Y. Hiroi, T. Kaminokado, S. Ono, and Y. Itoh. Focal surface occlusion. *Opt. Express*, 29(22):36581–36597, Oct 2021. 2

[20] M. Hirsch, G. Wetzstein, and R. Raskar. A compressive light field projection system. *ACM Trans. Graph.*, 33(4), jul 2014. 2

[21] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008. 2

[22] F.-C. Huang, K. Chen, and G. Wetzstein. The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Trans. Graph.*, 34(4):60, 2015. 2

[23] Y. Itoh, T. Hamasaki, and M. Sugimoto. Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator. *IEEE Trans. Vis. Comput. Graph*, 23(11):2463–2473, 2017. 2, 3

[24] Y. Itoh, T. Langlotz, D. Iwai, K. Kiyokawa, and T. Amano. Light attenuation display: Subtractive see-through near-eye display via spatial color filtering. *IEEE Trans. Vis. Comput. Graph*, 25(5):1951–1960, May 2019. 1, 2

[25] Y.-G. Ju, M.-H. Choi, P. Liu, B. Hellman, T. L. Lee, Y. Takashima, and J.-H. Park. Occlusion-capable optical-see-through near-eye display using a single digital micromirror device. *Opt. Lett.*, 45(13):3361–3364, July 2020. 1, 2, 8

[26] T. Kaminokado, Y. Hiroi, and Y. Itoh. StainedView: Variable-Intensity Light-Attenuation display with cascaded spatial color filtering for improved color fidelity. *IEEE Trans. Vis. Comput. Graph.*, 26(12):3576–3586, Dec. 2020. 1, 2

[27] I. Kauvar, S. J. Yang, L. Shi, I. McDowall, and G. Wetzstein. Adaptive color display via perceptually-driven factored spectral projection. *ACM Trans. Graph.*, 34(6), nov 2015. 2, 4, 6

[28] K. Kim., D. Heo., and J. Hahn. Occlusion-capable head-mounted display. In *Proceedings of the 7th International Conference on Photonics, Optics and Laser Technology - Volume 1: PHOTOPTICS,*, pages 299–302. INSTICC, SciTePress, 2019. 1, 2

[29] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *2nd IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, page 133, 2003. 1, 2

[30] B. Krajancich, N. Padmanaban, and G. Wetzstein. Factored occlusion: Single spatial light modulator occlusion-capable optical see-through augmented reality display. *IEEE Trans. Vis. Comput. Graph*, 26(5):1871–1879, 2020. 1, 2, 3, 8

[31] T. Langlotz, M. Cook, and H. Regenbrecht. Real-time radiometric compensation for optical see-through head-mounted displays. *IEEE Trans. Vis. Comput. Graph*, 22(11):2385–2394, Nov. 2016. 2

[32] T. Langlotz, J. Sutton, S. Zollmann, Y. Itoh, and H. Regenbrecht. Chromaglasses: Computational glasses for compensating colour blindness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 390:1–390:12, New York, NY, USA, 2018. ACM. 8

[33] D. Lanman, F. Heide, D. Reddy, J. Kautz, K. Pulli, and D. Luebke. Cascaded displays: spatiotemporal superresolution using offset pixel layers. In *ACM SIGGRAPH 2014 Emerging Technologies*, SIGGRAPH '14, New York, NY, USA, 2014. Association for Computing Machinery. 2

[34] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. 4

[35] B. Lyot. Optical apparatus with wide field using interference of polarized light. *CR Acad. Sci.(Paris)*, 197(1593), 1933. 3

[36] A. Maimone and H. Fuchs. Computational augmented reality eyeglasses. In *12th IEEE Intl. Symp. on Mixed and Augmented Reality (ISMAR)*, pages 29–38, 2013. 2

[37] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs. Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources. In *ACM SIGGRAPH 2014 Emerging Technologies*, page 20, 2014. 2

[38] B. Masia, G. Wetzstein, P. Didyk, and D. Gutierrez. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *CG*, 37(8):1012–1038, 2013. 2

[39] X. Meng, N. Wang, F. Shao, and S. Li. Vision transformer for pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–11, 2022. 8

[40] O. Mercier, Y. Sulai, K. Mackenzie, M. Zannoli, J. Hillis, D. Nowrouzezahrai, and D. Lanman. Fast gaze-contingent optimal decompositions for multifocal displays. *ACM Trans. Graph.*, 36(6), nov 2017. 2

[41] Ooi, Hiroi, and Itoh. A compact photochromic occlusion capable see-through display with holographic lenses. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, volume 0, pages 237–242, Mar. 2023. 2

[42] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral imaging using multiplexed illumination. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007. 6

[43] J. Platoš, P. Gajdoš, P. Krömer, and V. Snášel. Non-negative matrix factorization on gpu. In F. Zavoral, J. Yaghob, P. Pichappan, and E. El-Qawasmeh, editors, *Networked Digital Technologies*, pages 21–30, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. 8

[44] K. Rathinavel, G. Wetzstein, and H. Fuchs. Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics. *IEEE Trans. Vis. Comput. Graph.*, 25(11):3125–3134, 2019. 2

[45] Santos, Pedro, Gierlinger, Thomas, Machui, O. Stork, , and André. The daylight blocking optical stereo see-through hmd. In *Workshop on Immersive projection technologies/Emerging display technologies*, page 4. ACM, 2008. 2

[46] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. *ACM Trans. Graph.*, 23(3):760–768, aug 2004. 2

[47] K. Takahashi, Y. Kobayashi, and T. Fujii. From focal stack to tensor light-field display. *IEEE Transactions on Image Processing*, 27(9):4571–4584, 2018. 2

[48] T. Uchida, K. Sato, and S. Inokuchi. An optical see-through mr display with digital micro-mirror device. *Trans. of the Virtual Reality Society of Japan*, 7(2), 2002. (in Japanese). 2

[49] I. Šolc. Birefringent chain filters. *J. Opt. Soc. Am.*, 55(6):621–625, Jun 1965. 3

[50] C. Weiland, A.-K. Braun, and W. Heiden. Colorimetric and photometric compensation for optical see-through displays. In C. Stephanidis, editor, *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, pages 603–612, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. 2

[51] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3d: tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4), jul 2011. 2

[52] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.*, 31(4), jul 2012. 2

[53] A. Wilson and H. Hua. Design and demonstration of a vari-focal optical see-through head-mounted display using freeform alvarez lenses. *Optics express*, 27(11):15627–15637, 2019. 2

[54] A. Wilson and H. Hua. Design of a pupil-matched occlusion-capable optical see-through wearable display. *IEEE Trans. Vis. Comput. Graph*, 28(12):4113–4126, 2022. 2

[55] Y. Yamaguchi and Y. Takaki. See-through integral imaging display with background occlusion capability. *Applied Optics*, 55(3):A144–A149, 2016. 2

[56] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018. 6

[57] Y. Zhang, X. Hu, K. Kiyokawa, and X. Yang. Add-on occlusion: Turning Off-the-Shelf optical see-through head-mounted displays occlusion-capable. *IEEE Trans. Vis. Comput. Graph.*, PP, Feb. 2023. 2, 3