Color Pass-Through via Camera-Display Coupling

Ruikang Li[1], Molin Li[2], Jiarui Wu[1], Zhe Wei[3], Pengpeng Liu[3], and Tianfan Xue[1,†]

[1] CUHK MMLab
[2] Zhejiang University
[3] Central Media Technology Institute, Huawei

Abstract. When a real-world scene is captured by a smartphone camera and viewed on its screen, the displayed image often differs noticeably from the original scene in color, brightness, and contrast. This gap persists despite substantial advances in both modern cameras and displays. A key reason is that most pipelines factor the high-dimensional capture-to-display process into two separately calibrated camera and display stages, and then connect them through low-dimensional color transforms, leading to information bottlenecks and inevitable error accumulation. To address this systemic challenge, we propose **Color Pass-Through**, an end-to-end learned framework that operates directly on captured images. Our key insight is to treat the camera and display as a coupled system rather than calibrating them in isolation. Coupling the camera and display yields two practical advantages: (1) it brings the entire real-world scenes to the display via end-to-end optimization, and (2) it allows efficient one-step calibration for each distinct observer via complete capture-to-display path. We validate **Color Pass-Through** using both digital and human observers. Compared with representative baselines, our method achieves an average gain of $+\mathbf{2.0}$ points on a 5-point user-study and more than $\mathbf{2}\times$ improvement on quantitative metrics, demonstrating improved reproduction of the perceived color of the original scene. See project page: `https://lyricccco.github.io/color-pass-through/`

Keywords: Computational Photography · Color Pass-Through

1 Introduction

Photography seeks to faithfully reproduce a real scene's colors when showing them on a display. In practice, however, what we see in the real world often differs from what a smartphone screen presents. As shown in Fig. 1 (a), capturing a scene with a smartphone and viewing it on the screen can introduce shifts in both chromaticity and lightness: the dolls' hues drift, and the displayed image may appear overly bright, desaturated, or washed out compared with the original scene. Historically, this perceptual gap was exacerbated by limited sensor dynamic range and narrow-gamut, low-bit panels. Although modern sensors and high-quality displays mitigate this gap, a noticeable discrepancy remains

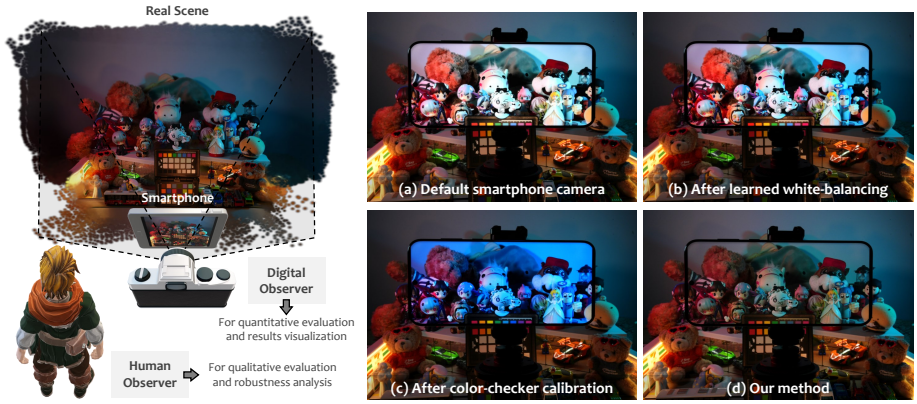

Fig. 1: Color Pass-Through. Capturing a real scene under multiple unknown light sources and re-displaying it on a smartphone screen often introduces perceptual color inconsistencies (a). Learned multi-illuminant auto white balance reduces illumination-induced color casts (b), while per-scene color-checker calibration maps the image into a standard color space (c); but neither reproduces the same color. In contrast, our method better preserves the perceived scene colors for both digital and human observers (d).

under standard sensor calibration and downstream image post-processing. However, even strong post-processing baselines do not close this gap. Learned multi-illuminant auto white-balance can reduce illumination-induced color casts [2] (Fig. 1 (b)), and per-scene color-checker calibration can map the captured image into a standard color space [51] (Fig. 1 (c)). Fundamentally, these methods rely on constrained assumptions about illumination or standard color space, and therefore can not provide a consistent guarantee of faithful color reproduction for a specific camera–display pair across diverse, unconstrained real-world scenes. This gap is even more problematic in immersive systems: see-through views in VR headsets (e.g., Vision Pro) can deviate significantly from the naked eye in color and contrast, undermining comfort and presence [9,22], and virtual try-on mirrors in consumer AR displays have similar fidelity issues [54].

To understand why existing solutions struggle to close this color gap, we revisit the capture-to-display imaging pipeline. Standardized International Color Consortium (ICC) workflows [34] rely on a three-channel color representation as a tractable intermediate (Fig. 2 (a)). In doing so, they decompose the process into two separately calibrated stages: *camera calibration*, mapping scene radiance to "RGB", and *display calibration*, mapping "RGB" to emitted radiance. This separation accumulates error. More fundamentally, camera measurements are inherently three-dimensional, whereas real-world radiance is high-dimensional, so calibration alone cannot overcome this intrinsic information bottleneck.

To close this perceptual color gap, we propose to treat camera and display as a single, coupled system, bridged by an end-to-end learnable neural projector (Fig. 2 (b)). Instead of calibrating each device to a reference color and composing

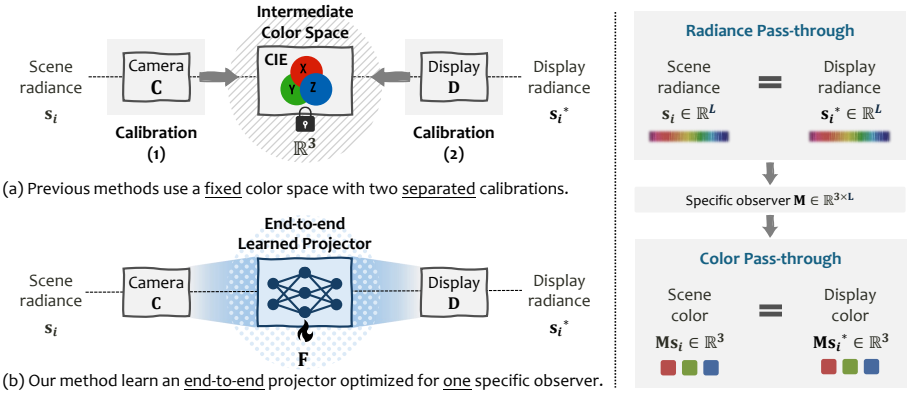

Fig. 2: Comparison of Methods and Objectives for Color Reproduction.

the separate mappings, we directly map displayed colors to the original scene colors for a specific observer. This end-to-end formulation reduces error accumulation and lessens the impact of information bottleneck, yielding markedly smaller mismatches in both color and brightness (Fig. 1 (d)). While this may appear to increase the calibration burden, since each camera–display pair must be characterized, it is well suited to pass-through use cases, where users typically view the captured scene on the same phone or immersive device. We therefore calibrate per camera–display pair rather than for every possible combination.

We represent the coupled camera–display pair as an unknown end-to-end non-linear projector (a device-specific mapping), learned with a lightweight pixel-wise neural network. To acquire training data, we introduce a re-capture protocol: each digital RGB sample is rendered on the target display and re-imaged by the paired camera, and the resulting measurements supervise the camera–display projector. We further study dataset design and network architectures, identifying an efficient configuration that learns the projector with a compact model, enabling practical camera–display projection with fast inference.

Despite this learned projector, visible color casts still persist for a target human observer under complex illumination. This residual mismatch arises from camera metamerism. We therefore derive objectives that transition from idealized *radiance pass-through* to *color pass-through* (Fig. 2, right-side), revealing that the dominant discrepancy lies in the camera's metameric-black subspace: spectral components invisible to the camera can still affect the observer's perceived color after display, i.e., colors appear identical to the camera may looks totally different to human observers. This limitation is fundamental, because camera spectral sensitivities generally differ from human visual responses.

Empirically, we find that the camera-null space exhibit low intrinsic dimensionality in practice and can be well approximated by a single dominant component, which also keeps the observer-specific correction low-dimensional. We therefore compensate the residual color cast using a learned predictor together with a single observer-specific calibration coefficient $\in \mathbb{R}^{3\times 1}$ applied in one step.

Contributions. This paper propose an end-to-end optimizing system that couples a camera–display pair to achieve color pass-through for a specific observer.

- *Learned Camera-Display Projection.* Our core contribution is an efficient pixel-wise neural projector that models the end-to-end mapping of a coupled camera–display, along with a practical re-capture protocol for data collection.
- *Camera-Null Color Correction.* We identify the dominant residual error as lying in the camera's metameric-black (camera-null) subspace, and estimated its main component through a learned predictor with a one-step, observer-specific calibration coefficient to compensate the remaining color cast.
- *Experimental Validation.* We validate our method with both objective measurements using a DSLR (a fixed digital observer used to produce all quantitative results and figures) and subjective evaluation via user studies, demonstrating robust color pass-through across diverse scenes and illuminants.

2 Related Work

Traditional color reproduction pipelines (e.g., ICC workflows) rely on an intermediate device-independent reference space such as CIE XYZ to mediate device-to-device transforms. In this paradigm, capture devices (e.g., cameras and scanners) map sensor measurements into the reference space, while output devices (e.g., displays and printers) apply an inverse mapping to device-dependent signals to reproduce the intended colors. This design yields two decoupled calibration stages: *camera calibration* and *display calibration*. We focus on the lines of work most relevant to our setting, and refer readers to [23,30] for broader background.

Computational color constancy. A large body of work in *camera calibration* addresses the correction of scene illumination in photos, assuming that perceptual colors remain consistent under different lighting conditions [13,27,28]. This problem is commonly known as *color constancy* or *white balance*. Traditional methods rely on hand-crafted assumptions [14, 24, 53], whereas recent learning-based approaches estimate illuminants directly from data [1, 4, 10–12, 33, 37], forming the field of *computational color constancy*. However, due to the high-dimensional nature of light, a global linear 3×3 transform is insufficient for accurate camera color constancy [18, 25, 36]. Recent work addresses this limitation by either extending single-illuminant estimation to multi-illuminant settings [3, 38, 39, 49] or moving from RGB alignment to spectral modeling [20, 40, 42]. These methods improve camera-side color correction, but do not model downstream display reproduction. As a result, they cannot guarantee that the corrected image, once shown on a specific display, will reproduce the perceived appearance of the original scene for a target observer.

Display color management. Display calibration typically follows ICC workflows, which transform device-dependent signals through a device-independent reference space. Chromatic adaptation transforms (e.g., Bradford or von Kries)

compensate for illuminant differences [32], while color appearance models such as CIECAM02 further account for perceptual viewing conditions [46, 58]. Display characterization itself is often modeled using parametric transfer functions, such as Gain–Offset–Gamma (GoG) models [19, 55], which describe the mapping between device input signals and emitted radiance. Modern devices further adapt these transformations to ambient lighting conditions (e.g., Apple's True Tone [6, 17, 48]). These methods characterize display-side reproduction under viewing conditions, but ignore the camera capture process, leaving the coupled capture-to-display problem unaddressed.

3 Derivation of Color-Accurate Pass-Through

We first introduce a theoretical model for *color pass-through* via a camera $\mathbf{C}$ and a display $\mathbf{D}$, as perceived by a target observer $\mathbf{M}$. Directly shown the captured image on the display may introduce perceptual gap between actual scene color and display color. To mitigate this gap, we apply a correction mapping $\mathbf{F}$ to the captured image before display, counteracting the distortions (Fig. 2 left-side).

3.1 Preliminary: Idealized Radiance Pass-Through

An idealized case of color reproduction is *radiance pass-through*: for every scene point i, the corresponding display radiance $\mathbf{s}_i^*$ should precisely reproduce the scene radiance $\mathbf{s}_i$, that is $\mathbf{s}_i^* = \mathbf{s}_i$. This guarantees that *any* observer with spectral sensitivities $\mathbf{M}$ perceives identical colors, as $\mathbf{M}\mathbf{s}_i^* = \mathbf{M}\mathbf{s}_i$ holds for $\forall\ \mathbf{M}$.

However, this ideal radiance pass-through is in general impossible, as the scene radiance is an high-dimension information, but cameras only record a 3-dimensional color. The coupled camera–display acts as an autoencoder: it encodes a high-dimensional radiance into low-dimensional colors and decodes it back, inevitably discarding spectral information during the transfer process.

One potential solution is to modify the captured image such that the display radiance matches the scene radiance. However, this is generally impossible. We model a camera by its spectral sensitivities $\mathbf{C} \in \mathbb{R}^{3\times L}$ and a display by its spectral primaries $\mathbf{D} \in \mathbb{R}^{L\times 3}$. Let $\mathbf{s}_i$, $\mathbf{s}_i^* \in \mathbb{R}^L$ denote the discretized scene and reproduced radiance at pixel i, sampled at L wavelengths, 3 be the number of color channels. To achieve radiance pass-through (no gap between scene and display radiance), we need to design a correction function matrix $\mathbf{F} \in \mathbb{R}^{3\times 3}$ that applied to the captured image, such that:

$$\mathbf{s}_i \;\equiv\; \mathbf{s}_i^* \;\equiv\; \mathbf{DFC}\mathbf{s}_i, \quad \forall i, \tag{1}$$

For Eq. 1 to hold for all $\mathbf{s}_i$, the composite operator $\mathbf{DFC} \in \mathbb{R}^{L\times L}$ must be the identity. However, this is generally impossible: since the correction $\mathbf{F} \in \mathbb{R}^{3\times 3}$, the rank of $\mathbf{DFC}$ is at most 3, and therefore it cannot equal the identity in $\mathbb{R}^{L\times L}$ when $L \gg 3$. This rank argument formalizes a fundamental limitation: with only finitely many channels, an arbitrary radiance cannot be reconstructed precisely. We therefore target a weaker version of pass-through, discussed below.

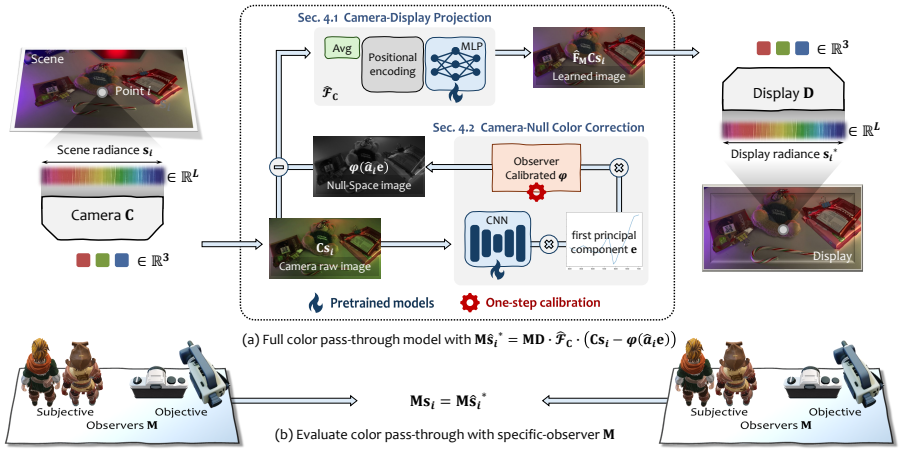

Fig. 3: Overall System of Our Color Pass-Through. (a) Full color pass-through model corresponding to Eq. (11); (b) Objective function and evaluation criteria.

3.2 Color Pass-Through

While exact *radiance pass-through* is generally impossible, we instead target a more practical goal: *color pass-through*: the display radiance need not match the scene radiance exactly; rather, it should appear similar to an observer (Fig. 2 right-side). Here the observer can be either a human or a three-channel camera, both of which possess limited color perceptron. We name this *color pass-through*.

Formally, let $\mathbf{M} \in \mathbb{R}^{3\times L}$ denote the spectral sensitivities of an observer. To achieve color pass-through, the observed display color $\mathbf{M}\mathbf{s}_i \in \mathbb{R}^3$, for a given scene point should equal the observed scene color $\mathbf{M}\mathbf{s}_i^* \in \mathbb{R}^3$. Here we still introduce a correction $\mathbf{F}_\mathbf{M}$ on captured image to enforce perceptually equivalence:

$$\mathbf{M}\mathbf{s}_i \equiv \mathbf{M}\mathbf{s}_i^* \equiv \mathbf{M}\mathbf{D}\mathbf{F}_\mathbf{M}\mathbf{C}\mathbf{s}_i, \quad \forall i \quad \Rightarrow \quad \mathbf{F}_\mathbf{M} = (\mathbf{M}\mathbf{D})^\dagger\mathbf{M}\mathbf{C}^\dagger, \tag{2}$$

where $(\cdot)^\dagger$ denotes the Moore–Penrose pseudoinverse. Here $\mathbf{M}\mathbf{D} \in \mathbb{R}^{3\times 3}$ and $\mathbf{M}\mathbf{C}^\dagger \in \mathbb{R}^{3\times L}$. In practice, directly instantiating the correction $\mathbf{F}_\mathbf{M} \in \mathbb{R}^{3\times L}$ is difficult because the sensitivity $\mathbf{M}$ is an unknown high-dimentional matrix and varies across observers. Consequently, $(\mathbf{M}\mathbf{D})^\dagger$ cannot be reliably calibrated. We therefore propose to approximate $\mathbf{F}_\mathbf{M}$ through decomposition described below.

4 Learning Color Pass-Through

Since it is hard to directly learn an observer-specifc $\mathbf{F}_\mathbf{M}$, we derive a robust, practical solution by decomposing it into two objectives: (i) a *camera–display projection* term $\mathcal{F}_\mathbf{C}$, which enforces pass-through in the camera measurement space, and (ii) a *camera-null color correction* term δ_i, which compensates observer-dependent deviations lies in the spectral space that are invisible to the camera.

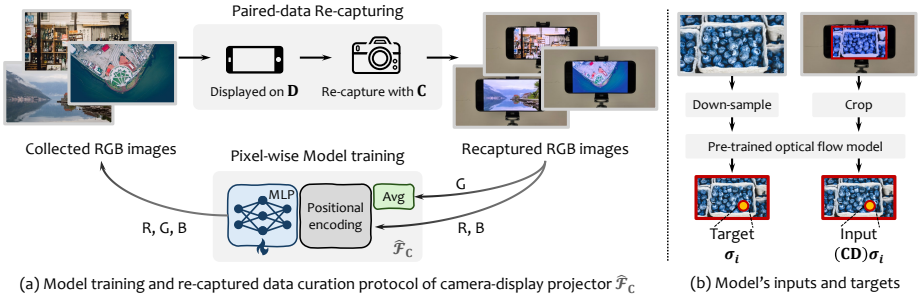

Fig. 4: Learning $\mathcal{F}_{\mathbf{C}}$ from Recaptured Data. A pixel-wise model is trained to approximate $\mathcal{F}_{\mathbf{C}}$ via (a) a re-capture protocol and (b) pixel-aligned input–target pairs.

This decomposition turns $\mathbf{F_M}$ into two learnable predictors that can be trained and optimized independently and then combined end-to-end at inference time to calibrate a coefficient $\varphi \in \mathbb{R}^{3\times 1}$ within one step for a specific observer, as illustrated in Fig. 3. The mathematical derivation and implementation details of each module are described below, leading to the final formulation in Eq. (11).

4.1 Camera-Display Projection

In this section, we simplify the problem by replacing the actual human observer $\mathbf{M}$ with the camera $\mathbf{C}$, i.e. $\mathbf{M} := \mathbf{C}$, where we assume a same-model camera as the observer. Under this condition, Eq. (2) simplifies to a special case in which the observer-specific objective $\mathbf{F_M}$ collapses to a camera-specific mapping $\mathbf{F_C}$:

$$\mathbf{C}\mathbf{s}_i \;\equiv\; \mathbf{C}\mathbf{s}_i^* \;\equiv\; \mathbf{C}\mathbf{D}\mathbf{F}_{\mathbf{C}}\mathbf{C}\mathbf{s}_i, \quad \forall i \quad \Rightarrow \quad \mathbf{F_C} \;=\; (\mathbf{CD})^{\dagger}, \tag{3}$$

This yields a simple solution for color pass-through, $\mathbf{F_C} = (\mathbf{CD})^{\dagger} \in \mathbb{R}^{3\times 3}$, which we refer to as the *camera–display projector*. It forms a central component of our model and, as we show later, can be learned from data via a simple training objective. The resulting *camera–display projection* is defined as $\mathbf{P}_{\mathrm{C}} := \mathbf{D}\mathbf{F}_{\mathrm{C}}\mathbf{C} \in \mathbb{R}^{L\times L}$ as it is an idempotent projection satisfying $\mathbf{P}_{\mathrm{C}}^2 = \mathbf{P}_{\mathrm{C}}$, which can map any scene radiance $\mathbf{s}_i$ to a *camera metamer* $\mathbf{s}_i^*$ that preserves the camera color.

In principle, one could estimate $\mathbf{F_C}$ via classical spectral calibration (e.g., we can measure the camera $\mathbf{C}$ with a monochromator and the display $\mathbf{D}$ with a spectrometer). However, such calibration may not be accurate as commercial displays exhibit substantial non-linearities in their default modes (e.g., gamma and tone mapping) that cannot be fully disabled. Consequently, a single affine transform is insufficient to accurately model camera-display projector $\mathbf{F_C}$.

This motivates a data-driven alternative. Instead of enforcing linearity, we treat $\mathbf{F_C} = (\mathbf{CD})^{\dagger}$ as an unknown (potentially non-linear) operator $\mathcal{F}_{\mathbf{C}}(\cdot)$ and learn it from *re-captured* pairs (Fig. 4 (a)). In the forward capture-and-display process, scene radiance passes through the camera $\mathbf{C}$ and display $\mathbf{D}$ sequentially, forming the operator $(\mathbf{CD})$. To approximate its inverse, we construct a reverse

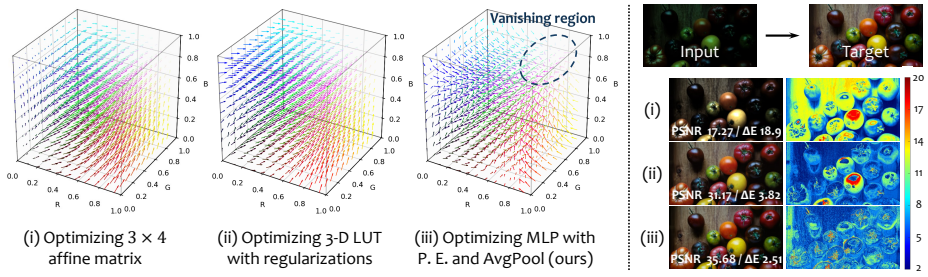

Fig. 5: Effectiveness of the Learned Model for $\widehat{\mathcal{F}}_{\mathbf{C}}(\cdot;\theta)$. In contrast to an affine fit (i) or a learned 3D LUT [57] (ii), our model (iii): captures (a) higher-frequency details and (b) achieves better performance. Notably, the learned projector exhibits a "vanishing" region where multiple camera color collapse to the same displayed color.

process: digital images are first rendered on the display and then re-captured by the camera, implicitly modeling the pseudo-inverse of $\mathbf{CD}$.

Specifically, we construct training data by (i) sampling RGB images as supervision $\sigma_i \in \mathbb{R}^3$, (ii) rendering on $\mathbf{D}$, and (iii) re-capturing the displayed images with $\mathbf{C}$ to obtain $(\mathbf{CD})\sigma_i \in \mathbb{R}^3$. After pixel-wise alignment using an optical-flow model [52] (Fig. 4 (b)), we train a pixel-wise network to explicitly learn the mapping $(\mathbf{CD})\sigma_i \to \sigma_i$, yielding a non-linear approximation of $(\mathbf{CD})^{\dagger}$.

We introduce two key architectural refinements that enable a neural surrogate to represent the practically complex projector $\mathcal{F}_{\mathbf{C}}(\cdot)$. First, following learned color-transfer models [16, 41],we use a lightweight multilayer perceptron (MLP) parameterized by θ–two fully connected layers with hidden width 128–augmented with positional encoding to better preserve high-frequency variations. Second, we apply a simple but effective AvgPool layer to the green channel of the input. The motivation is that demosaicing introduces spatial interpolation: each RGB triplet is partially synthesized from neighboring sensor samples. Although pixel-shift cameras could provide per-pixel tri-stimulus measurements [44], we instead adopt this minimal preprocessing step, which we find consistently improves accuracy, detailed in supplementary. With these refinements, we adopt a learned neural projector $\widehat{\mathcal{F}}_{\mathbf{C}}(\cdot;\theta)$ and optimize θ by minimizing the following objective:

$$\theta^* \;=\; \arg\min_{\theta} \sum_i \left\| \sigma_i - \widehat{\mathcal{F}}_{\mathbf{C}}\big((\mathbf{CD})\sigma_i;\theta\big) \right\|_1 , \tag{4}$$

Compared to alternative learned fits, our model captures substantially higher-frequency details (Fig. 5 (a)) and yields better quantitative results (Fig. 5 (b)). However, the learned camera–display projector guarantees color pass-through only when camera itself is the observer (Fig. 6 (a)).When the observer differs from the camera—e.g., a DSLR—$\widehat{\mathcal{F}}_{\mathbf{C}}$ alone leads to noticeable color shifts in the image displayed on the phone (Fig. 6 (b)), we address this in the next section.

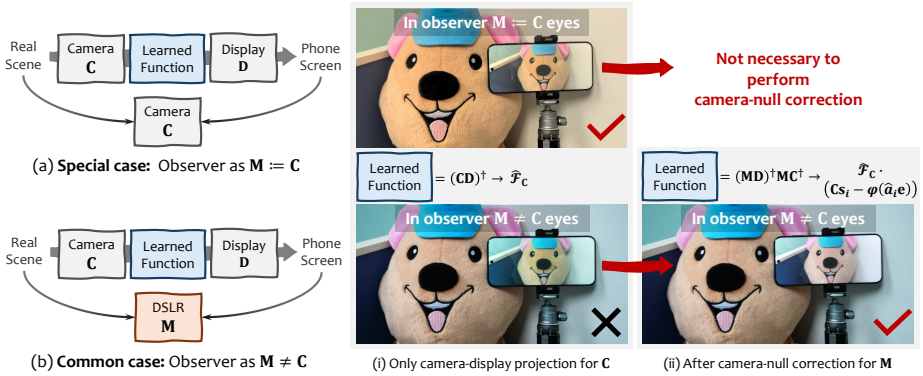

Fig. 6: Effectiveness of the Two Learned Components for Color Pass-Through. (a) When using another camera of the same model as the digital observer: (i) the learned camera–display projection alone is sufficient to achieve pass-through, (ii) no observer calibration for $\mathbf{M}$ is required; (b) When observer is different from camera (e.g. a DSLR): (i) the camera–display projection alone is insufficient, (ii) adding the camera-null correction for a target observer $\mathbf{M}$ yields consistent color reproduction.

4.2 Camera-Null Color Correction

While the learned projector $\widehat{\mathcal{F}}_{\mathbf{C}}$ (approximating $(\mathbf{CD})^{\dagger}$) reliably enforces color pass-through *in the camera measurement space*, our ultimate goal is pass-through for a target observer $\mathbf{M}$. We therefore analyze the conditions under which the camera-aligned projector transfers to other observers and when it fails.

Our motivation is empirical. Consider two smartphones of the same model: one acts as the camera and the other as the observer (Fig. 6 (a)). When we compare the real scene to the image displayed on phone through an identical phone (a digital observer), the camera–display projector alone reproduces matching colors from phone's viewpoint. Yet the color of the displayed image on phone can still deviate from the real scene perceptually for other viewers. In practice, we often observe a faint tinted "color mask" over the screen. To make this discrepancy explicit, we introduce a DSLR as a proxy 3-channel observer (Fig. 6 (b)).

To eliminate this "color mask", we introduce *camera-null color correction*. We estimate the residual color cast using a single observer-specific calibration coefficient and compensate for it to achieve color pass-through for $\mathbf{M}$.

First, we derive the origin of the residual color cast. For clarity of the derivation, we temporarily treat $\mathbf{F}_{\mathbf{C}}$ as linear and apply the correction at the *input* to the projector rather than at its output. The reason is practical: in real systems $\mathbf{F}_{\mathbf{C}}$ is intrinsically non-linear (see Fig. 5), making compensation *after* this mapping more entangled and less stable. Applying the correction *before* the projector instead yields a cleaner and more robust adjustment. Therefore, we introduce a correction term $\delta_i \in \mathbb{R}^3$ at the input of $\mathbf{F}_{\mathbf{C}}$ and relate the observer-specific solution in Eq. (2) to the camera-specific solution in Eq. (3), yielding:

$$\mathbf{F}_{\mathbf{M}}\,\mathbf{C}\mathbf{s}_i \;=\; \mathbf{F}_{\mathbf{C}}\big(\mathbf{C}\mathbf{s}_i - \delta_i\big) \quad \Rightarrow \quad \delta_i \;:=\; \mathbf{C}\big(\mathbf{I} - \mathbf{P}_{\mathbf{M}}\big)\mathbf{s}_i, \tag{5}$$

where $\mathbf{P_M} = \mathbf{D}(\mathbf{MD})^{\dagger}\mathbf{M}$ is an idempotent projection sharing similar properties with $\mathbf{P}_{\mathrm{C}}$. Eq. (5) reveals two key regimes characterizing the correction term δ_i:

(1) when $\delta_i = 0$ for all observers $\mathbf{M}$. Sufficient condition for Eq. (5) to hold is:

$$\mathrm{null}(\mathbf{C}) \subseteq \mathrm{null}(\mathbf{M}), \tag{6}$$

which states that any spectrum invisible to the camera is also invisible to the observer. We refer to this as the *Luther–Ives condition under camera–display projection*. It is closely related in spirit to the classical Luther–Ives condition for colorimetric capture ($\mathbf{M} = \mathbf{TC} \Leftrightarrow \mathrm{null}(\mathbf{M}) = \mathrm{null}(\mathbf{C})$), but extends its condition. We provide the proof in the supplementary and suggest a hardware-driven solution that enforce $\delta_i = 0$. However, Eq. (6) generally does not hold in common settings. We therefore seek an optimization-based learning method to estimate an observer-specific approximation of δ_i, leading to the second regime.

(2) when $\delta_i \neq 0$ for a specific observer $\mathbf{M}$. According to Eq. (5), using $\mathbf{F_C}$ requires estimating δ_i and subtracting it from the camera color $\mathbf{Cs}_i$. To analyze δ_i explicitly, we propose to decompose $\mathbf{s}_i$ by project it with projector $\mathbf{P_C}$:

$$\mathbf{s}_i = \mathbf{r}_i + \mathbf{n}_i, \qquad \begin{cases} \mathbf{r}_i := \mathbf{P}_{\mathrm{C}}\mathbf{s}_i & \in \mathrm{Range}(\mathbf{D}), \\ \mathbf{n}_i := \mathbf{s}_i - \mathbf{r}_i & \in \mathrm{Null}(\mathbf{C}), \end{cases} \tag{7}$$

Geometrically, Eq. (7) corresponds to an oblique projection of $\mathbf{s}_i$ onto range($\mathbf{D}$) along null($\mathbf{C}$). A key consequence is $\mathbf{C}(\mathbf{I}-\mathbf{P_M})\mathbf{r}_i = \mathbf{0}$, see the supplementary for a proof. Therefore, the correction term δ_i of Eq. (5) simplifies to $\mathbf{C}(\mathbf{I}-\mathbf{P_M})\mathbf{n}_i$, which shows that the correction term originates entirely from the *metameric black* (i.e., invisible) space to the camera $\mathbf{C}$, since $\mathbf{n}_i \in \mathrm{null}(\mathbf{C})$ and thus $\mathbf{Cn}_i = \mathbf{0}$.

In practice, the only available measurement is the camera color $\mathbf{Cs}_i \in \mathbb{R}^3$, yet the desired correction δ_i depends on the unobserved component $\mathbf{n}_i$. A naïve solution would be to treat $\mathbf{Cs}_i$ as a prior and attempt to reconstruct $\mathbf{n}_i \in \mathbb{R}^L$ (or even the full spectrum). Even if such reconstruction were feasible, a second bottleneck remains: computing δ_i requires the observer-related operator $\mathbf{C}(\mathbf{I} - \mathbf{P_M}) \in \mathbb{R}^{3\times L}$, which is impractical to calibrate since it has L dimensionality.

To make $\mathbf{C}(\mathbf{I}-\mathbf{P_M})$ calibratable, we use the classic observations that natural reflectance and radiance spectra exhibit low intrinsic dimensionality [45,47]. We therefore assume that the camera-null component $\mathbf{n}_i$ also lies approximately in a low-dimensional subspace and model it with a PCA basis, we get:

$$\mathbf{n}_i \approx \sum_{k=1}^{K} a_i^{(k)}\, \mathbf{e}^{(k)} \qquad K \ll L, \tag{8}$$

where K is the number of PCA bases. We find that the first principal component explains approximately 93% of the variance on a large radiance set (see supplement)), motivating our practical rank-1 approximation with $K = 1$: $\mathbf{n}_i \approx \hat{a}_i\mathbf{e} \in \mathbb{R}^{1\times L}$ and consequently the correction color δ_i from Eq. (5) becomes:

$$\delta_i = \mathbf{C}(\mathbf{I} - \mathbf{P_M}) \cdot \mathbf{n}_i \approx \varphi \cdot (a_i\mathbf{e}), \tag{9}$$

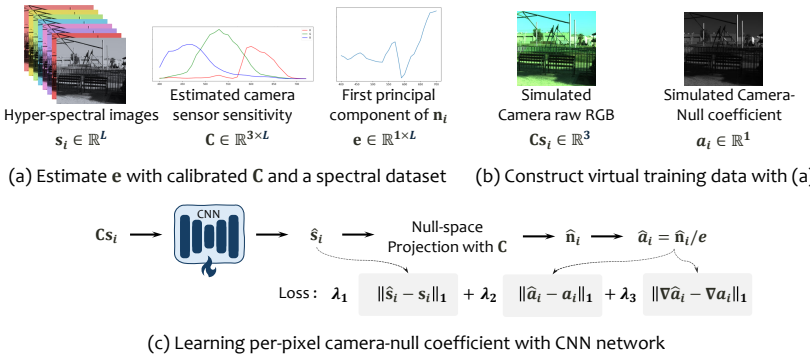

Fig. 7: Learning Camera-null Color Coefficient a_i from Hyper-spectral Data.

Here, all three terms can be estimated. The first principle component $\mathbf{e}$ can be pre-computed from data; the scalar a_i is a *camera-null color coefficient* at pixel i can be estimated via a learned predictor, and φ is a *calibration coefficient* only relates to observer $\mathbf{M}$ and display $\mathbf{D}$. This parameterization makes calibration practical, since $\varphi \in \mathbb{R}^{3\times 1}$ contains only 3 unknown variables. Moreover, later in Fig. 13, we show φ remains nearly stable across varying in-the-wild scenes.

Specifically, to learn a predictor for a_i, we proceed in three stages, as shown in Fig. 7: (a) we estimate the camera spectral sensitivities $\mathbf{C}$ via a ColorChecker calibration procedure, following the robust method of [35]; using the calibrated $\mathbf{C}$, we then compute a PCA basis for the camera-null component from hyper-spectral datasets (treated as $\mathbf{s}_i$) and retain the dominant component as $\mathbf{e}$; (b) we construct training inputs as simulated camera raw measurements $\mathbf{Cs}_i$, with the learning target being a_i; and (c) we estimate the per-pixel coefficient $\hat{a}_i$ with an efficient spectral reconstruction network [15] by minimizing the loss function:

$$\mathcal{L} = \sum_i \Big[\lambda_1 \|\hat{\mathbf{s}}_i - \mathbf{s}_i\|_1 + \lambda_2 \|\hat{a}_i - a_i\|_1 + \lambda_3 \|\nabla\hat{a}_i - \nabla a_i\|_1\Big]. \tag{10}$$

where ∇ denotes the total-variation operator applied to the coefficient a_i, encouraging spatial smoothness. In addition to supervising $\hat{a}_i$ directly, we also retain the spectral reconstruction loss $\|\hat{\mathbf{s}}_i - \mathbf{s}_i\|_1$. Later in Tab. 1b, we show spectral reconstruction supervision also improves the estimated coefficient $\hat{a}_i$ accuracy.

In summary, combining the observer-specific color pass-through correction of Eq. (2) with the learned camera–display projector $\widehat{\mathcal{F}}_{\mathbf{C}}$ (Eq. (5)) and the learned camera-null correction $\widehat{\delta}_i = \hat{a}_i\mathbf{e}$ (Eq. (9)), our full Color Pass-Through model is:

$$\underbrace{\mathbf{M}\mathbf{s}_i}_{\text{scene color}} \approx \underbrace{\mathbf{M}\widehat{\mathbf{s}}_i^*}_{\text{display color}} = \mathbf{M}\mathbf{D} \cdot \underbrace{\widehat{\mathcal{F}}_{\mathbf{C}}}_{\textbf{learned}} \cdot \Big(\underbrace{\mathbf{C}\mathbf{s}_i}_{\text{camera color}} - \underbrace{\varphi}_{\textbf{calibrate}} \cdot \underbrace{(\hat{a}_i\mathbf{e})}_{\textbf{learned}} \Big) \tag{11}$$

(a) Evaluation of Learned Camera-Display Projector $\widehat{\mathcal{F}}_{\mathbf{C}}$.

Methods	Params (K)	Runtime (ms)	PSNR ↑	ΔE_{mean} ↓	ΔE_{p95} ↓	STRESS ↓
IA-3DLUT [57]	593.0	339.92	26.69	5.66	9.22	9.22
NiLUT [21]	33.9	31.16	31.03	3.59	9.39	5.02
CSRNet [31]	36.5	253.78	31.12	3.59	9.34	5.00
Ours	30.8	41.30	32.13	3.34	8.81	4.69

(b) Evaluation of Learned $\hat{a}_i$

λ_1	λ_2	λ_3	PSNR($\hat{a}_i, a_i$)↑	PSNR($\hat{\mathbf{s}}_i, \mathbf{s}_i$)↑
✓			28.27	14.94
	✓	✓	29.42	32.69
✓	✓		29.00	31.49
✓	✓	✓	29.68	31.95

Table 1: Evaluate performance of learned models for camera–display projector $\widehat{\mathcal{F}}_{\mathbf{C}}$ and camera–null color coefficients $\hat{a}_i$. We report quantitative metrics, including PSNR [29], ΔE_{mean} [50], $\Delta E * 95$ (95th percentile), and STRESS [26], as well as runtime on 2K-resolution inputs measured on an NVIDIA RTX 4090 GPU.

After pre-training the two predictors $\widehat{\mathcal{F}}_{\mathbf{C}}$ and $\hat{a}_i\mathbf{e}$ (see Fig. 4 and Fig. 7), we perform a one-time calibration for a target observer to estimate $\varphi \in \mathbb{R}^{3\times 1}$. Once calibrated, φ is kept fixed for all subsequent evaluations.

5 RESULTS

We first evaluate the fitting quality of the two learned predictors, then compare the full color pass-through model against several baselines, and finally present ablations to validate our design choices and method's robustness. Additional experiments are reported in the supplementary material.

All experiments are conducted using two identical smartphones (HUAWEI Pura 70 Pro or Xiaomi 17 Pro Max): one acts as the camera **C** and the other as the display **D**; a DSLR camera (Sony ILCE-7M4) is used as the digital observer **M** to produce all quantitative results and figures. We set the phone screen to maximum brightness to maximize the usable color gamut and capture linear RAW images in the phone's Pro mode to minimize extra in-camera processing. For indoor multi-illuminant tests, we use Ulanzi VL49RGB lights in HSI mode, whose dedicated RGB LEDs provide spectrally peaky illumination.

5.1 Fitting Quality on Two Learned Components

Evaluation on Learned Camera-Display Projector $\widehat{\mathcal{F}}_{\mathbf{C}}$. We train $\widehat{\mathcal{F}}_{\mathbf{C}}$ using the DIV2K dataset [5], denoised by [43], and perform recapture process to obtain 800 pairs for training and 100 for testing. We report evaluation in Tab. 1a and compare our learned model with existing pixel-wise color transfer methods.

Evaluation on Learned Camera-Null Color Coefficient $\hat{a}_i$. We train $\hat{a}_i$ using hyper-spectral datasets (ARAD-1K [8], CAVE [56] and ICVL [7]) with $L = 31$ spectral channels. In total, we use 1000 images for training and 182 for testing. We report evaluation in Tab. 1b under different loss configurations. Based on the setting that achieves the highest accuracy for $\hat{a}_i$, we preserve λ_1, λ_2 and adopt $\lambda_1, \lambda_2, \lambda_3 = (1, 0.01, 1)$ in Eq. (10) for all subsequent experiments.

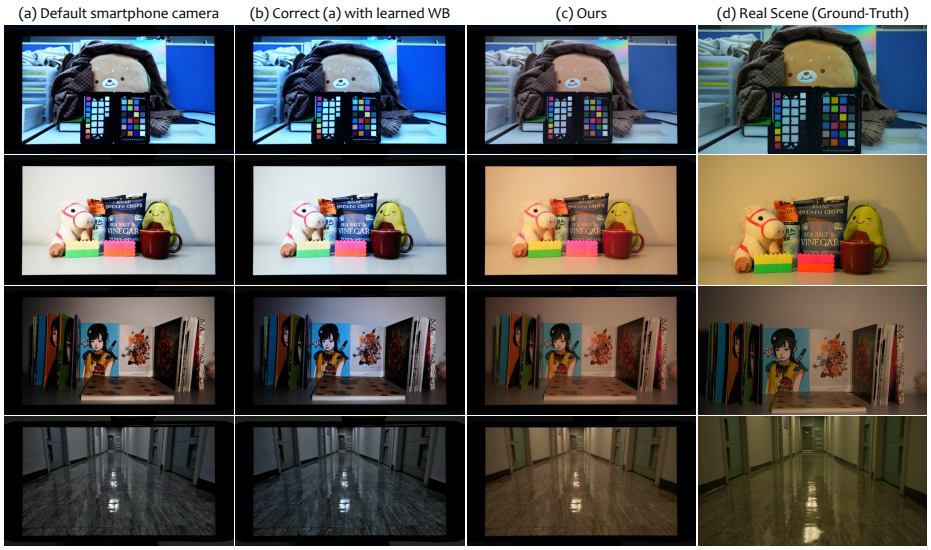

Fig. 8: Separated-view Comparison. The left three columns show images observed through a phone display, while the rightmost column directly observes the real scene. Our method preserves more consistent colors than other baselines.

Table 2: Quantitative comparisons for color pass-through using the 24 ColorChecker patches under unseen, diverse illumination. Numbers in gray indicate results with brightness aligned to our method; our method still achieves the best performance.

Methods	PSNR ↑		ΔE_{mean} ↓		STRESS ↓	
	Huawei	Xiaomi	Huawei	Xiaomi	Huawei	Xiaomi
Default smartphone camera	13.78 (15.80)	14.61 (15.93)	14.62 (12.31)	13.49 (11.54)	26.23 (27.58)	25.07 (25.27)
Color-checker calibration	15.02 (15.23)	16.36 (16.85)	18.49 (18.40)	15.32 (15.09)	38.85 (38.56)	36.42 (36.04)
Multi-illuminants Auto-WB	12.84 (12.95)	13.92 (14.41)	17.84 (17.37)	17.08 (16.26)	31.12 (30.48)	30.05 (29.66)
Ours w/o camera-null correction	27.32	27.84	6.49	5.97	17.27	16.39
Ours	28.65	29.10	5.18	4.79	17.48	16.12

5.2 Evaluating Full Color Pass-Through Model

Compared Methods. We compare against three strong and practical baselines that cover the standard camera-side and ICC-style correction pipelines. (a) *Default smartphone camera:* the image produced by the smartphone's commercial ISP, including its proprietary white balance, tone mapping, and color tuning. This represents a highly optimized commercial pipeline. (b) *Multi-illuminant Auto-WB:* a state-of-the-art learned white-balance method [2] applied to the default smartphone output to correct spatially varying illumination-induced color casts. (c) *ColorChecker calibration:* an instrumented calibration baseline using a ColorChecker Passport [51]. We estimate an ICC/profile-based color correction from the ColorChecker measurements and render the corrected image in Photoshop before display, making this a strong calibration baseline and an approximate upper bound for standard per-scene color correction workflows.

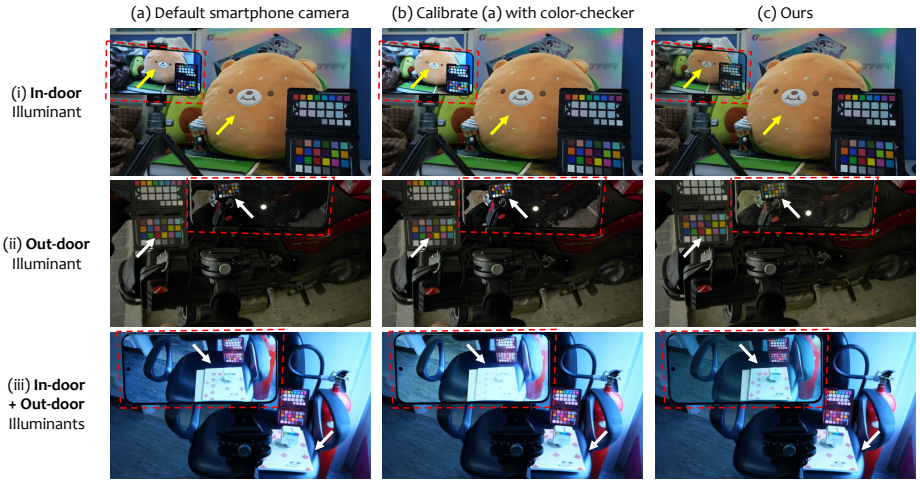

Fig. 9: Direct in-scene Comparison. Each column shows a smartphone screen with the processed image hovering over the real scene. Our method reproduces closer colors.

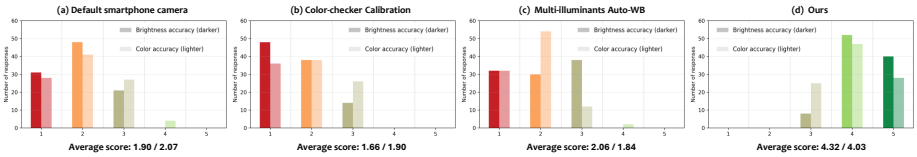

Fig. 10: User study on brightness and color accuracy. Participants rated similarity to the reference on a 5-point Likert scale (1 for least similar, 5 for most similar).

Objective Evaluation. We first calibrate the digital observer (DSLR) via a one-time grid search for the calibration coefficient $\varphi \in \mathbb{R}^{3\times1}$ using 24 ColorChecker patches under natural illumination. Each entry of φ is searched in $[0.025, 0.075]$ with step 0.0125, and the φ that minimizes the error is selected.

We then evaluate our method and all baselines using the same ColorChecker under a diverse set of illuminations: ten correlated color temperatures ranging from 2500 K to 9000 K, and five randomly sampled RGB-LED illuminant colors, quantitative results over the 24 patches are summarized in Tab. 2. Moreover, qualitative comparisons under diverse real scenes are shown in Fig. 8 and Fig. 9.

Subjective Evaluation. We conduct user studies with ten human observers to assess perceptual color pass-through. A one-time grid search generates candidate calibration coefficients $\varphi \in \mathbb{R}^{3\times1}$, from which each participant selects the best φ in one scene. After calibration, participants evaluate each method across ten unseen scenes by rating *brightness accuracy* and *color accuracy*, measuring how well the displayed image matches the real scene, as summarized in Fig. 10.

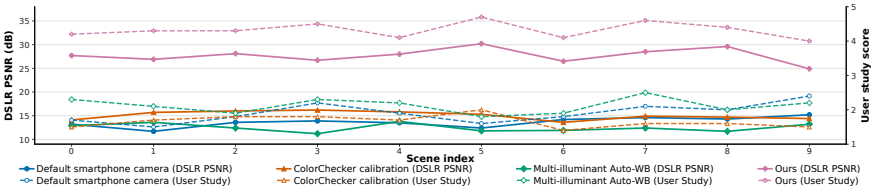

Fig. 11: Preference comparison between DSLR-based and human-subject evaluations.

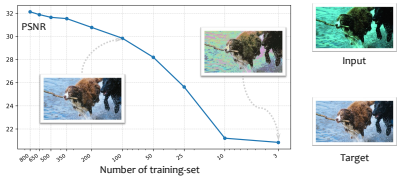

Fig. 12: Learning efficiency of $\widehat{\mathcal{F}}_{\mathbf{C}}$.

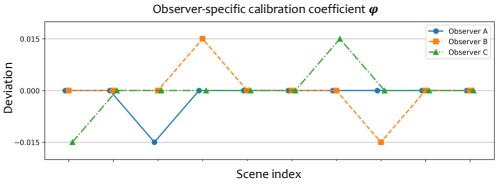

Fig. 13: Robustness of calibrated φ.

5.3 Ablations.

Validity of the DSLR observer proxy. We use a DSLR camera as a digital proxy observer for quantitative evaluation. This does not assume that the DSLR is identical to human vision; rather, it serves as a controlled and repeatable three-channel observer, consistent with human's three-cone color responses. As shown in Fig. 11, DSLR-based preferences follow trends consistent with human-subject judgments across diverse scenes, while often providing a stricter evaluation.

Learning Efficiency of $\widehat{\mathcal{F}}_{\mathbf{C}}$. We test the learned $\widehat{\mathcal{F}}_{\mathbf{C}}$ with limited training data. We progressively subsample the training set and find that even with 100 images for training, the results still preserves similar colors, as shown in Fig. 12.

Robustness of Calibrated $\boldsymbol{\varphi}$. To assess whether the calibrated vector φ remains stable across illuminants for the same observer $\mathbf{M}$, we present a user study by perturbing each entry of φ by ± 0.015, and ask the participant whether any of these alternatives better match the real scene. We visualize the selections of three representative observers across 10 scenes in Fig. 13. For most scenes, observers keep their original φ without switching, indicating cross-illuminant stability.

6 Conclusion

We presented **Color Pass-Through**, an end-to-end learned correction applied to captured images via a coupled camera–display pair. After optimizing two predictors, we combine them at inference time and perform a one-time coefficient calibration for a target observer. This enables color pass-through across diverse scenes and illuminants, as confirmed by both evaluation metrics and user studies.

Acknowledgements

This work was supported in part by the Research Grants Council of Hong Kong under the Early Career Scheme (ECS), Grant No. 24209224.

References

1. Afifi, M., Barron, J.T., LeGendre, C., Tsai, Y.T., Bleibel, F.: Cross-camera convolutional color constancy. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1981–1990 (2021)
2. Afifi, M., Brubaker, M.A., Brown, M.S.: Auto white-balance correction for mixed-illuminant scenes. In: IEEE Winter Conference on Applications of Computer Vision (WACV) (2022)
3. Afifi, M., Brubaker, M.A., Brown, M.S.: Auto white-balance correction for mixed-illuminant scenes. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 1210–1219 (2022)
4. Afifi, M., Zhao, L., Punnappurath, A., Abdelsalam, M.A., Zhang, R., Brown, M.S.: Time-aware auto white balance in mobile photography. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5038–5047 (2025)
5. Agustsson, E., Timofte, R.: Ntire 2017 challenge on single image super-resolution: Dataset and study (July 2017)
6. Apple Inc.: Color correction based on perceptual criteria and ambient light chromaticity (2023)
7. Arad, B., Ben-Shahar, O.: Sparse recovery of hyperspectral signal from natural rgb images. In: European conference on computer vision. pp. 19–34. Springer (2016)
8. Arad, B., Timofte, R., Yahel, R., Morag, N., Bernat, A., Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., et al.: Ntire 2022 spectral recovery challenge and data set. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 863–881 (2022)
9. Bailenson, J.N., Beams, B., Brown, J., DeVeaux, C., Han, E., Queiroz, A.C., Ratan, R., Santoso, M., Srirangarajan, T., Tao, Y., et al.: Seeing the world through digital prisms: Psychological implications of passthrough video usage in mixed reality (2024)
10. Barron, J.T.: Convolutional color constancy. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 379–387 (2015)
11. Barron, J.T., Tsai, Y.T.: Fast fourier color constancy. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 886–894 (2017)
12. Bianco, S., Cusano, C., Schettini, R.: Color constancy using cnns. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 81–89 (2015)
13. Brainard, D.H., Freeman, W.T.: Bayesian color constancy. Journal of the optical Society of America A **14**(7), 1393–1411 (1997)
14. Buchsbaum, G.: A spatial processor model for object colour perception. Journal of the Franklin institute **310**(1), 1–26 (1980)
15. Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., Pfister, H., Timofte, R., Van Gool, L.: Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 745–755 (2022)

16. Canham, T.D., Tedla, S., Murdoch, M.J., Brown, M.S.: Gain-mlp: Improving hdr gain map encoding via a lightweight mlp. arXiv preprint arXiv:2503.11883 (2025)
17. Chang, S.C., Zhong, J.Z.: Ambient light adaptive displays with paper-like appearance, https://patents.google.com/patent/US20190139512A1/en
18. Cheng, D., Price, B., Cohen, S., Brown, M.S.: Beyond white: Ground truth colors for color constancy correction. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 298–306 (2015)
19. Cho, Y.H., Im, J.H., Ha, Y.H.: Inverse characterization method of alternate gain-offset-gamma model for display devices. Journal of Imaging Science and Technology **50**(2), 139–148 (2006)
20. Cogo, L., Buzzelli, M., Bianco, S., Vazquez-Corral, J., Schettini, R.: Leveraging multispectral sensors for color correction in mobile cameras. arXiv preprint arXiv:2512.08441 (2025)
21. Conde, M.V., Vazquez-Corral, J., Brown, M.S., Timofte, R.: Nilut: Conditional neural implicit 3d lookup tables for image enhancement. vol. 38, pp. 1371–1379 (2024)
22. De Souza, J., Tartz, R.: Visual perception and user satisfaction in video see-through head-mounted displays: a mixed-methods evaluation. Frontiers in Virtual Reality **5**, 1368721 (2024)
23. Fairchild, M.D.: Color appearance models. John Wiley & Sons (2013)
24. Finlayson, G.D., Trezzi, E.: Shades of gray and colour constancy. In: Color and imaging conference. vol. 12, pp. 37–41. Society of Imaging Science and Technology (2004)
25. Finlayson, G.D., Zakizadeh, R.: Reproduction angular error: An improved performance metric for illuminant estimation. perception **310**(1), 1–26 (2014)
26. Finlayson, G.D., Zakizadeh, R.: Reproduction angular error: An improved performance metric for illuminant estimation. Journal of Electronic Imaging **23**(3) (2014)
27. Gehler, P.V., Rother, C., Blake, A., Minka, T., Sharp, T.: Bayesian color constancy revisited. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. IEEE (2008)
28. Gijsenij, A., Gevers, T., Van De Weijer, J.: Computational color constancy: Survey and experiments. IEEE transactions on image processing **20**(9), 2475–2489 (2011)
29. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Pearson, 3 edn. (2008)
30. Green, P.: Color management understanding and using icc profiles, 2010
31. He, J., Liu, Y., Qiao, Y., Dong, C.: Conditional sequential modulation for efficient global image retouching. pp. 679–695. Springer (2020)
32. Hellwig, L., Fairchild, M.D.: Brightness, lightness, colorfulness, and chroma in ciecam02 and cam16. Color Research & Application **47**(5), 1083–1095 (2022)
33. Hu, Y., Wang, B., Lin, S.: Fc4: Fully convolutional color constancy with confidence-weighted pooling. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4085–4094 (2017)
34. International Color Consortium: Specification icc.1:2022 (profile version 4.4.0.0). Tech. rep., International Color Consortium, Reston, VA (2022), https://www.color.org/specification/ICC.1-2022-05.pdf
35. Jiang, J., Liu, D., Gu, J., Süsstrunk, S.: What is the space of spectral sensitivity functions for digital color cameras? In: 2013 IEEE Workshop on Applications of Computer Vision (WACV). pp. 168–179. IEEE (2013)
36. Karaimer, H.C., Brown, M.S.: Improving color reproduction accuracy on cameras. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6440–6449 (2018)

37. Kim, D., Afifi, M., Kim, D., Brown, M.S., Kim, S.J.: Ccmnet: Leveraging calibrated color correction matrices for cross-camera color constancy. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6198–6208 (2025)
38. Kim, D., Kim, J., Nam, S., Lee, D., Lee, Y., Kang, N., Lee, H.E., Yoo, B., Han, J.J., Kim, S.J.: Large scale multi-illuminant (lsmi) dataset for developing white balance algorithm under mixed illumination. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2410–2419 (2021)
39. Kim, D., Kim, J., Yu, J., Kim, S.J.: Attentive illumination decomposition model for multi-illuminant white balancing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 25512–25521 (2024)
40. Koskinen, S., Acar, E., Kämäräinen, J.K.: Single pixel spectral color constancy: S. koskinen et al. International Journal of Computer Vision **132**(2), 287–299 (2024)
41. Le, H.M., Price, B., Cohen, S., Brown, M.S.: Gamutmlp: a lightweight mlp for color loss recovery. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18268–18277 (2023)
42. Li, J., Chen, C., Hu, X., Song, F., Yan, Y., Xiong, Z.: Multi-spectral image color reproduction. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 8400–8409. IEEE (2025)
43. Li, R., Wang, Y., Chen, S., Zhang, F., Gu, J., Xue, T.: Dualdn: Dual-domain denoising via differentiable isp. In: European Conference on Computer Vision. pp. 160–177. Springer (2024)
44. Liu, G., et al.: A super-resolution imaging system based on sub-pixel camera shift. In: Proceedings of SPIE. vol. 10817. SPIE (2018). https://doi.org/10.1117/12.2502377
45. Maloney, L.T.: Evaluation of linear models of surface spectral reflectance with small numbers of parameters. Journal of the Optical Society of America A **3**(10), 1673–1683 (1986)
46. Moroney, N., Fairchild, M., Hunt, R., Li, C.: The ciecam02 color appearance model (2002)
47. Parkkinen, J.P., Hallikainen, J., Jaaskelainen, T.: Characteristic spectra of munsell colors. Journal of the Optical society of America A **6**(2), 318–322 (1989)
48. Qian, Y., Chen, S.C., Hsiang, E.L., Akimoto, H., Lin, C.L., Wu, S.T.: Enhancing a display's sunlight readability with tone mapping. Photonics **11**(6), 578 (2024)
49. Serrano-Lozano, D., Arora, A., Herranz, L., Derpanis, K.G., Brown, M.S., Vazquez-Corral, J.: Revisiting image fusion for multi-illuminant white-balance correction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 8275–8284 (2025)
50. Sharma, G., Wu, W., Dalal, E.N.: The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color Research & Application **30**(1), 21–30 (2005)
51. Sunoj, S., Igathinathane, C., et al.: Color calibration of digital images for agriculture and other applications. Computers and Electronics in Agriculture **155**, 306–318 (2018)
52. Teed, Z., Deng, J.: Raft: Recurrent all-pairs field transforms for optical flow. In: European conference on computer vision. pp. 402–419. Springer (2020)
53. Van De Weijer, J., Gevers, T., Gijsenij, A.: Edge-based color constancy. IEEE Transactions on image processing **16**(9), 2207–2214 (2007)
54. Wang, J., Ma, S., Bayer, K., Zhang, Y., Wang, P., Zhou, B., Nayar, S., Krishnan, G.: Perspective-aligned ar mirror with under-display camera. ACM Trans. Graphic. **43**(6), 1–11 (2024)

55. Wu, J., et al.: Color characterization model for oled displays with improved channel interaction modeling. Color Research & Application (2023)
56. Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S.K.: Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. IEEE transactions on image processing **19**(9), 2241–2253 (2010)
57. Zeng, H., Cai, J., Li, L., Cao, Z., Zhang, L.: Learning image-adaptive 3d lookup tables for high performance photo enhancement in real-time **44**(4), 2058–2073 (2020)
58. Zhao, H., Mainster, M.A.: The effect of chromatic dispersion on pseudophakic optical performance. British journal of ophthalmology **91**(9), 1225–1229 (2007)