CONJ: Dual-Pair-Based 1+2 Decomposition with Residual Invariance

Bungaku Yokota

February 9, 2026

abstract

Many color spaces and image/video processing pipelines adopt a 1+2 structure that decomposes a three-dimensional color signal into a one-dimensional principal axis and a two-dimensional residual. However, when nonlinear operations such as gamma correction, tone mapping, or roll-off are applied along the principal-axis direction, the residual component (chroma) generally varies as a by-product of the principal-axis manipulation, resulting in chroma shift (crosstalk).

In this paper, we treat the working space as an arbitrary three-dimensional real vector space V. Given a pair $(u, \ell) \in V \times V^*$ satisfying the normalization $\ell(u) = 1$ (a normalized dual pair), we introduce the $1 + 2$ decomposition

$$x = \ell(x)u + r(x), \qquad r(x) := x - \ell(x)u \in \ker \ell,$$

that is, $V = \operatorname{span}(u) \oplus \ker \ell$. Along this decomposition, we define the map

$$T_\varphi(x) = \varphi(\ell(x))u + r(x),$$

which applies an arbitrary scalar nonlinearity φ only to the principal-axis coordinate $\ell(x)$ (the CONJ transform), and show that for the projection $P(x) := x - \ell(x)u$,

$$P(T_\varphi(x)) = P(x)$$

holds for any φ. Thus, from only the choice of the decomposition (the principal axis and the residual subspace) together with the normalization, we obtain a design condition that keeps the residual geometrically invariant under arbitrary nonlinear principal-axis operations.

Numerical experiments on uniform random samples and real photographs confirm that conventional pipelines (e.g., Y′ CbCr operations after gamma correction, or lightness operations in CIELAB) exhibit crosstalk on the order of 10^{-2} for uniform random inputs and 10^{-3} for real images, whereas the proposed method reduces it to the rounding-error level of double-precision floating-point arithmetic.

The proposed framework assumes neither a specific reference stimulus space (e.g., XYZ) nor a specific implementation pipeline. For an arbitrary three-channel working space, it provides a design principle that separates tone operations as a one-dimensional mapping along the principal axis while preserving the two-dimensional residual component pointwise. It does not, however, directly

address issues that depend on gamut constraints, clipping, quantization, or other implementation-dependent effects over line segments; rather, its scope is to strongly guarantee geometric consistency (independence between the principal axis and the residual) at the tone-operation stage.

1 Introduction

In digital image/video processing and coding, a $1+2$ structure that decomposes a three-dimensional color signal into a one-dimensional luminance component and a two-dimensional chroma component is widely used. Typical examples include YCbCr-type color-difference representations and the separation of lightness and chromatic components in perceptual spaces such as CIELAB [1–3]. In practical pipelines, however, nonlinear operations such as gamma correction, tone mapping, roll-off, and HDR/SDR mapping are unavoidable [4], and the chroma component often varies as a consequence of luminance-side operations (chroma shift), which becomes a recurring issue [5, 6]. A variety of post-tone-mapping color-correction and saturation-adjustment methods have been proposed, yet achieving both color-appearance preservation and luminance compression typically requires manual tuning and involves trade-offs [5, 6].

In this paper, we regard chroma shift not as an inherent limitation of a particular color space, but as a consequence of the *order of operations*—namely, *at which stage a $1+2$ split is performed and to which component the nonlinearity is applied.* Even when the underlying signal is the same three-dimensional space V, the difference between

- applying a nonlinearity only to the principal-axis scalar after splitting ($S \to N1$), and
- applying a nonlinearity to the three channels first and splitting afterwards ($N3 \to S$)

essentially determines whether crosstalk between luminance and chroma occurs.

The core of this work is as follows. We treat the working space as an arbitrary three-dimensional real vector space V and fix a pair $(u, \ell) \in V \times V^*$ satisfying the normalization $\ell(u) = 1$. We then define the residual

$$r(x) := x - \ell(x)u, \qquad r(x) \in \ker \ell,$$

and apply an arbitrary scalar nonlinearity φ only to the principal-axis scalar $\ell(x)$ via the map

$$T_\varphi(x) := \varphi(\ell(x))u + r(x),$$

which we call the *CONJ transform*. For the projection $P(x) := x - \ell(x)u$, we have

$$P(T_\varphi(x)) = P(x)$$

identically, so that the residual (chroma) is exactly invariant and does not depend on the specific form of φ. In what follows, we refer to $r(x)$ as the *residual (residual component)* and $\ker \ell$ as the *residual subspace.*

The contributions of this paper can be summarized as follows.

1. We introduce a $1+2$ decomposition based on a normalized dual pair satisfying $\ell(u) = 1$, and define the CONJ transform T_φ as the composition of this split and a one-axis nonlinearity.

2. We show that the residual invariance $P(T_\varphi(x)) = P(x)$ does not depend on the specific form of φ, and provide a design condition that structurally eliminates crosstalk under nonlinear principal-axis operations.

3. Using uniform random samples and real photographs, we confirm that, compared with conventional pipelines (which decompose and manipulate signals in a nonlinear domain), the CONJ-type pipeline reduces the observed residual variation down to the rounding-error level of double-precision computation.

■Scope and limitations. The CONJ transform studied in this paper formulates tone operations on a three-channel signal as a one-dimensional mapping along a principal axis, with the goal of preserving the two-dimensional residual component (chroma) exactly at each point. Accordingly, the proposed framework does not directly resolve issues that depend on gamut constraints, clipping, quantization, or implementation-dependent effects over entire line segments such as boundary intersections and endpoint behavior. These are design problems orthogonal to the present focus. The scope of this paper is therefore limited to providing a design principle that strongly guarantees the geometric consistency of luminance–chroma separation at the tone-operation stage.

2 Proposed Method: CONJ Transform

2.1 Signal space and assumptions

In this paper, we treat pixel values in a working color space as triples of real numbers. The target is an arbitrary three-channel signal, which we regard as an element of a three-dimensional real vector space V. We do not assume a specific choice of basis (e.g., RGB, XYZ, or Lab), nor do we assume a particular inner product. The only structure required is the minimal one: V *is a three-dimensional linear space.*

The goal of this section is to choose, within V,

- a vector $u \in V$ representing a principal axis (to be treated as a luminance/neutral direction), and
- a linear functional $\ell \in V^* := \mathrm{Hom}(V, \mathbb{R})$ that reads out the scalar coordinate along that principal axis,

and to use the pair (u, ℓ) to split the signal into a one-dimensional principal-axis component and a two-dimensional residual component, on top of which we define the CONJ transform. A central point is that, prior to discussing *where* a nonlinearity is placed, the decisive choice is *how* the principal axis u and the readout ℓ (i.e., the $1 + 2$ decomposition) are selected.

2.2 Principal axis and residual plane: $1 + 2$ decomposition and dual-pair normalization

Choose a nonzero vector $u \in V$, which will be treated as the *principal axis* (luminance direction). Next, introduce a linear functional $\ell : V \to \mathbb{R}$ to read out the scalar component along the principal axis. At the implementation level, this corresponds to a weighted sum of the three channels. For example, in

an RGB coordinate system, with coefficients (ℓ_R, ℓ_G, ℓ_B), it can be written as

$$\ell(x) = \ell_R x_R + \ell_G x_G + \ell_B x_B. \tag{2.1}$$

To fix the scale ambiguity of (u, ℓ), we adopt the following *dual-pair normalization*:

$$\ell(u) = 1. \tag{2.2}$$

For a general pair (u, ℓ) with $\ell(u) \neq 0$, one can normalize it to satisfy (2.2), for example by $u \leftarrow u/\ell(u)$. Hereafter, unless stated otherwise, we assume that (u, ℓ) satisfies the dual-pair normalization (2.2) (and we keep the same notation after normalization).

Under this condition, any signal $x \in V$ can be decomposed as

$$x = \ell(x)\, u + \big(x - \ell(x)u\big). \tag{2.3}$$

The first term is the component along the principal axis, and the second term is the remainder. Define

$$Y(x) := \ell(x), \tag{2.4}$$

$$r(x) := x - Y(x)u. \tag{2.5}$$

We call $Y(x)$ the principal-axis scalar (hereafter, the *luminance scalar*) and $r(x)$ the *residual component*. In what follows, we refer to $r(x)$ as the *residual (residual component)* and to $\ker \ell$ as the *residual subspace*. In particular, when $\dim V = 3$, $\ker \ell$ is a two-dimensional plane, so for intuition we also call it the *residual plane*.

For the residual, by linearity and the normalization (2.2),

$$\ell\big(r(x)\big) = \ell\big(x - \ell(x)u\big) = \ell(x) - \ell(x)\ell(u) = 0 \tag{2.6}$$

holds, so $r(x)$ lies in the kernel of ℓ,

$$\ker \ell := \{v \in V \mid \ell(v) = 0\}.$$

Therefore, any x can be represented as the sum of

- the point $Y(x)u$ on the line $\mathbb{R}u$, and
- the point $r(x)$ on the plane $\ker \ell$.

Moreover, the representation is unique because $\mathbb{R}u \cap \ker \ell = \{0\}$ (indeed, if $v = au$ and $\ell(v) = 0$, then $0 = \ell(v) = a\ell(u) = a$). Thus we obtain the direct-sum decomposition

$$V = \mathbb{R}u \oplus \ker \ell \tag{2.7}$$

into a one-dimensional principal axis and a two-dimensional residual plane (a more general discussion is summarized in Appendix AppendixA).

Hereafter, by the "$1+2$ structure" we mean the structure based on the direct-sum decomposition (2.7) induced by the pair (u, ℓ) together with the dual-pair normalization (2.2).

2.3 Definition of the CONJ transform

Next, on top of this $1+2$ structure, we define a transform that applies an arbitrary nonlinearity to the luminance scalar while leaving the residual component unchanged. We call this the *CONJ Transform* (CONJ transform, hereafter simply the CONJ transform).

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a real-valued function of one variable. It represents a tone curve that maps a luminance scalar Y to another scalar $\varphi(Y)$, and may be any nonlinearity including gamma correction, tone mapping, and roll-off [4, 7].

Fix a pair (u, ℓ) satisfying the dual-pair normalization (2.2). For an input signal $x \in V$, the CONJ transform T_φ is defined by the following procedure:

1. Use the luminance readout ℓ to compute the input luminance

$$Y = \ell(x). \tag{2.8}$$

2. Subtract the contribution along the principal axis to obtain the residual component

$$r(x) = x - Yu, \tag{2.9}$$

   and keep it unchanged.

3. Apply the tone curve φ to Y to obtain the new luminance $\varphi(Y)$.

4. Recompose by placing $\varphi(Y)$ along the principal axis and adding back the original residual:

$$x' = \varphi(Y)\, u + r(x). \tag{2.10}$$

Summarizing the procedure, the CONJ transform can be written as

$$T_\varphi(x) := \varphi(\ell(x))\, u + \big(x - \ell(x)u\big). \tag{2.11}$$

That is, we first decompose the input into the luminance scalar $Y(x)$ and the residual vector $r(x)$, deform only the luminance scalar by an arbitrary one-dimensional nonlinearity φ, and then recompose while keeping the residual vector fixed.

2.4 Invariance of the residual component (proposition and proof)

The key property of the CONJ transform is that, as long as the tone curve φ is applied as a one-dimensional function, the residual component is invariant *under ideal real arithmetic*, regardless of the specific form of φ. In practice, finite-precision arithmetic and operations such as clipping may restrict the region in which this invariance effectively holds.

We present a concise proof of this invariance. Under the dual-pair normalization (2.2), define linear maps $Q, P : V \to V$ by

$$Qx := u\,\ell(x), \qquad P := I - Q,$$

where I denotes the identity map. Then $Px = x - \ell(x)u = r(x)$, and P acts as a projection onto the residual plane $\ker \ell$. Indeed, the normalization (2.2) implies $Pu = u - \ell(u)u = 0$, and also $P^2 = P$ holds (see Appendix AppendixA for details).

Proposition 2.1 (Residual invariance)**.** *For any pair* (u, ℓ) *satisfying the dual-pair normalization* (2.2) *and any function* $\varphi : \mathbb{R} \to \mathbb{R}$*, the identity*

$$P\big(T_\varphi(x)\big) = P(x)$$

holds for all $x \in V$*. That is, the CONJ transform preserves the residual component* $r(x) = Px$*.*

Proof. By the definition (2.11), we can write $T_\varphi(x) = \varphi(\ell(x))u + Px$. Therefore,

$$P\big(T_\varphi(x)\big) = P\big(\varphi(\ell(x))u + Px\big) = \varphi(\ell(x)) \, Pu + P^2 x = 0 + Px = P(x),$$

where we used $Pu = 0$ and $P^2 = P$. $\qquad\square$

Importantly, this property does not depend on the specific form of φ; it follows solely from the $1 + 2$ decomposition induced by (u, ℓ) and the dual-pair normalization (2.2). Therefore, once the principal axis and the residual plane are fixed, any choice of one-dimensional nonlinearity preserves the residual component under ideal real arithmetic.

In actual implementations, the range over which exact invariance can be guaranteed is finite due to finite-precision arithmetic, clipping, and dynamic-range constraints. In this paper, we treat this range as the *region where residual invariance effectively holds* and evaluate the property through numerical experiments.

3   A CONJ-oriented view of existing $1 + 2$ color spaces

3.1   Axes of comparison (decomposition and the order of nonlinearity)

In this section, we organize representative $1 + 2$-structured color spaces using a minimal set of axes from the viewpoint of CONJ. Our focus is not on historical developments or implementation details of each space, but on how the relationship between the principal axis and the residual plane is affected by the presence and ordering of nonlinear operations.

As representative examples, we consider YCbCr (BT.601/BT.709) [8, 9], CIELAB [2], IPT [10], ICtCp [7, 11], Oklab [12], and LogLuv [13].

We summarize the structure of each space using the following three symbols:

- S: a $1 + 2$ decomposition (splitting into a principal-axis scalar and a residual component). In the notation of this paper, this corresponds to choosing (u, ℓ) and obtaining $Y = \ell(x)$ and $r(x) = x - Yu$.
- $N1$: a one-dimensional nonlinearity acting only on the principal-axis scalar Y, i.e., $Y \mapsto \varphi(Y)$, with no direct action on the other components.
- $N3$: a three-dimensional nonlinearity acting on all three components of the working space. Typical examples include gamma correction and cube-root-type compression functions.

The column "Residual plane" indicates which of the following applies to the residual invariance (the property stated in Section 2 that the principal-axis operation does not change the position on the residual plane):

Table1: Key properties of 1+2 spaces from the CONJ viewpoint (order of decomposition and nonlinearity / residual plane). S: 1+2 decomposition, $N1$: one-dimensional nonlinearity on the principal-axis scalar, $N3$: nonlinearity on three components.

Space	Working space V	Procedure	Residual plane
YCbCr (BT.601/709)	$R'G'B'$	$N3 \to S$	$\times$
CIELAB	XYZ	$N3 \to S$	$\sim$
IPT / ICtCp	LMS-family	$N3 \to S$	$\sim$
Oklab	approximate LMS-family	$N3 \to S$	$\sim$
LogLuv	XYZ	$S \to N1$	$\checkmark$ (reinterpretation)
CONJ	arbitrary V	$S \to N1$	$\checkmark$

$\checkmark$: holds exactly, $\sim$: intended as a separation but not guaranteed exactly / holds only approximately, $\times$: not guaranteed in general.

Table 1 provides an overview.

3.2 $S \to N1$: applying a one-dimensional nonlinearity after decomposition

As shown in Section 2, once we choose a principal axis u and a readout ℓ on a working space V, adopt the dual-pair normalization $\ell(u) = 1$, and thereby fix the decomposition $V = \mathbb{R}u \oplus \ker \ell$, applying $N1$ only to the principal-axis scalar $Y = \ell(x)$ yields residual-plane invariance under ideal real arithmetic. In this sense, $S \to N1$ is a structural condition for "nonlinearly deforming only the principal axis while keeping the residual plane fixed."

■CONJ   CONJ explicitly realizes this $S \to N1$ pattern on an arbitrary working space V. That is, it first fixes the decomposition S and then applies $N1$ to the principal-axis scalar. This ordering places crosstalk on the side where it can be excluded by design.

■LogLuv (positioning as a reinterpretation)   LogLuv [13] takes linear XYZ as the working space, applies logarithmic compression (either $\log Y$ or an approximation thereof) to a luminance readout $Y = \ell_Y(x)$, and represents chromaticity by a two-dimensional parameter normalized by Y. From the procedural viewpoint, it can therefore be organized as an $S \to N1$ type: a decomposition S followed by a one-dimensional nonlinearity $N1$ on the principal-axis scalar. Here, "isomorphic (reinterpretation)" means that, if we allow a change of coordinates (a basis change) on the residual plane, the formulation can be rewritten into the CONJ canonical form under the normalization $\ell(u) = 1$,

$$T_\varphi(x) = \varphi(\ell(x))u + (I - u\ell)x.$$

This does not claim that LogLuv was designed with CONJ in mind.

3.3 $N3 \to S$: constructing a $1+2$ form after a three-dimensional nonlinearity

CIELAB [2], IPT/ICtCp [7, 10, 11], and Oklab [12] typically apply a three-dimensional nonlinearity $N3$ to all components of the working space first and then aggregate the result into a $1+2$ form ($N3 \to S$). These spaces are primarily designed for perceptual convenience (e.g., handling of brightness and color differences), and in practice provide quantities that behave like a principal axis and chromatic differences. However, residual-plane invariance in the sense of Section 2 is not structurally guaranteed in general.

Accordingly, when one manipulates only the principal-axis coordinate (e.g., L^* or L) and applies the inverse transform, how well the position on the residual plane is preserved depends on the specific nonlinearity, matrix design choices, dynamic range, and implementation aspects such as rounding. Thus it cannot be claimed to be "exactly invariant" (denoted by $\sim$ in Table 1). From our viewpoint, this difference is best understood as a difference in design objectives.

3.4 Decomposition on $R'G'B'$ as the working space: positioning of YCbCr

YCbCr takes $R'G'B'$ (gamma-encoded RGB) as the working space and constructs a luminance-like quantity and color-difference components by linear combinations [8, 9]. Although it appears to be in a $1+2$ form, because gamma correction acts on all three components prior to the decomposition, independence between principal-axis operations and residual operations (residual-plane invariance) is not guaranteed in general. In this paper, we treat this structure as a representative example of $N3 \to S$.

3.5 Summary: design degrees of freedom emphasized in this paper

The discussion above indicates that even when a representation appears to have a $1+2$ structure, the treatment of the residual plane changes essentially depending on whether the nonlinearity is placed before or after S. Our claim is that, prior to the specific *shape* of a tone curve, what is decisive is how one chooses the principal axis u and the readout ℓ on a working space V, and whether one adopts the dual-pair normalization $\ell(u) = 1$ to fix S.

In the next section, we illustrate this dual-pair normalization with concrete examples (e.g., scene-linear sRGB) and describe how fixing (u, ℓ) translates into an implementation.

4 Dual-pair normalization and concretization

4.1 Dual-pair normalization (general form)

Let V be a three-dimensional real vector space and let its algebraic dual be $V^* := \mathrm{Hom}(V, \mathbb{R})$. A pair consisting of a principal-axis vector $u \in V$ and a linear functional $\ell \in V^*$ corresponds to specifying the $1+2$ decomposition (the principal-axis scalar and the residual plane) discussed in Section 2.

To fix the scale freedom, we adopt the *dual-pair normalization*

$$\ell(u) = 1. \tag{4.1}$$

For a general pair (u, ℓ) with $\ell(u) \neq 0$, one can enforce (4.1) by rescaling, for example,

$$u \leftarrow \frac{u}{\ell(u)} \qquad \text{or} \qquad \ell \leftarrow \frac{\ell}{\ell(u)}.$$

(Since scaling ℓ leaves $\ker \ell$ invariant, the geometry of the residual plane does not change.) Hereafter, also in this section, we assume (4.1) unless stated otherwise, and we keep the same notation (u, ℓ) after normalization.

4.2 Direct-sum decomposition and projection after normalization (minimal form for implementation)

Under the dual-pair normalization (4.1), define for any $x \in V$

$$Y(x) := \ell(x), \qquad r(x) := x - \ell(x)u. \tag{4.2}$$

Then

$$\ell(r(x)) = \ell(x) - \ell(x)\ell(u) = 0$$

implies $r(x) \in \ker \ell$. Hence we obtain the direct-sum decomposition

$$V = \mathbb{R}u \oplus \ker \ell, \tag{4.3}$$

and the representation $x = Y(x)u + r(x)$ is unique.

Define the rank-one map and the associated projection by

$$Qx := u\,\ell(x), \qquad P := I - Q \tag{4.4}$$

where I denotes the identity map. From (4.1), we have

$$Pu = u - \ell(u)u = 0, \qquad P^2 = P,$$

so P acts as a projection onto $\ker \ell$.

With this notation, the CONJ transform defined in Section 2 is

$$T_\varphi(x) = \varphi(\ell(x))\,u + \big(x - \ell(x)u\big), \tag{4.5}$$

but for implementation it is often convenient to use the following equivalent *minimal form*:

$$T_\varphi(x) = x + \big(\varphi(\ell(x)) - \ell(x)\big)\,u. \tag{4.6}$$

That is, one can implement the transform by (1) computing $Y = \ell(x)$, (2) applying the one-dimensional mapping $Y \mapsto \varphi(Y)$, and (3) adding the difference $\Delta = \varphi(Y) - Y$ along the principal-axis direction. Residual-plane invariance under ideal real arithmetic has already been shown as a proposition in Section 2. In practice, however, finite-precision arithmetic and operations such as clipping may restrict the region where this property effectively holds.

4.3 Concrete example: normalization and a coordinate realization in scene-linear sRGB

As a concrete example, consider the working space to be the three-dimensional space of scene-linear sRGB, $V_{\mathrm{sRGB}} = \mathbb{R}^3_{\mathrm{RGB}}$ [3]. We choose the principal axis as the achromatic direction

$$u_s = (1, 1, 1)^\top.$$

Define the luminance readout using weights $w = (w_R, w_G, w_B)^\top$ by

$$\ell_s(x) = w^\top x = w_R x_R + w_G x_G + w_B x_B.$$

With a typical choice

$$w = (0.2126,\ 0.7152,\ 0.0722)^\top,$$

we have $w_R + w_G + w_B = 1$, hence $\ell_s(u_s) = 1$ holds and the dual-pair normalization $\ell(u) = 1$ is already satisfied (a representative set of Rec. 709 coefficients) [9]. (When the coefficients do not sum to 1, one can always normalize them by the procedure in the previous subsection.)

A coordinate realization of the residual plane $\ker \ell_s$ is not unique (there is freedom in choosing a basis on $\ker \ell_s$). For example, take two independent vectors satisfying $\ell_s(v) = 0$:

$$v_1 = \left(1,\ -\frac{w_R}{w_G},\ 0\right)^\top, \qquad v_2 = \left(1,\ 0,\ -\frac{w_R}{w_B}\right)^\top.$$

One immediately verifies that $w^\top v_1 = w^\top v_2 = 0$. Then

$$B := [\, u_s\ v_1\ v_2 \,] \in \mathbb{R}^{3\times 3}$$

is invertible, and for $y = (Y, C_1, C_2)^\top$ we can write

$$x = B\, y = Y\, u_s + C_1 v_1 + C_2 v_2.$$

Conversely, letting $A := B^{-1}$ yields

$$y = Ax.$$

Since $v_1, v_2 \in \ker \ell_s$ and $\ell_s(u_s) = 1$, the first row of A must coincide with ℓ_s (i.e., $w^\top$), which guarantees that

$$Y = \ell_s(x).$$

Accordingly, one may implement CONJ as

$$(Y, C_1, C_2)^\top = Ax, \quad Y' = \varphi(Y), \quad x' = B(Y', C_1, C_2)^\top,$$

or, using the minimal form (4.6), as

$$x' = x + (\varphi(\ell_s(x)) - \ell_s(x))\, u_s.$$

While the choice of coordinates on the residual plane (i.e., the choice of v_1, v_2) can affect numerical stability and scaling, residual-plane invariance itself is determined by (u, ℓ) and the normalization and does not depend on the particular coordinate realization.

4.4 An implementation example via a 3×3 matrix representation (coordinate realization)

In this subsection, we present an explicit implementation form of the CONJ transform as a pair of 3×3 matrices (a pre-transform and a post-transform) under the dual-pair normalization $\ell(u) = 1$. This is equivalent to the definition in Section 2 and is useful when one wishes to make the coordinates on the residual plane explicit.

■(Step 1) Choose a basis for the residual plane   Choose two independent vectors $v_1, v_2 \in V$ that span $\ker \ell$ (i.e., $\ell(v_1) = \ell(v_2) = 0$). This choice has degrees of freedom; it suffices that v_1 and v_2 are linearly independent and that extreme numerical scaling is avoided.

■(Step 2) Basis matrix and pre/post transforms   Form the matrix whose columns are u, v_1, v_2,

$$B := [\, u \; v_1 \; v_2 \,] \in \mathbb{R}^{3 \times 3}.$$

If u, v_1, v_2 are linearly independent, then B is invertible and we define

$$A := B^{-1}.$$

Then for any $x \in V$ we have

$$y = (Y, C_1, C_2)^\top := Ax, \qquad x = By = Yu + C_1 v_1 + C_2 v_2.$$

Moreover, from $\ell(u) = 1$ and $\ell(v_1) = \ell(v_2) = 0$ we obtain

$$\ell(x) = \ell(By) = Y,$$

so Y coincides with the readout by ℓ. Therefore, the first row of A coincides with ℓ (in its coordinate representation).

■(Step 3) Matrix form of CONJ   Given a tone curve $\varphi : \mathbb{R} \to \mathbb{R}$, define

$$y' := (\varphi(Y),\, C_1,\, C_2)^\top$$

and output

$$x' := By'.$$

Equivalently,

$$x' = B \begin{pmatrix} \varphi(Y) \\ C_1 \\ C_2 \end{pmatrix}, \qquad (Y, C_1, C_2)^\top = Ax.$$

This is equivalent to the definition (2.11) in Section 2, and in matrix form it can be implemented as "apply a one-dimensional nonlinearity to the first component while keeping the second and third components unchanged."

Symbol	Meaning
V	Color vector space (e.g., $\mathbb{R}^3$)
x	Sample (pixel) vector (RGB/XYZ, etc.)
u	Principal-axis vector (achromatic axis, luminance axis, etc.)
ℓ	Readout functional (principal-axis scalar; e.g., $\ell(x) = c^\top x$)
$\ell(u) = 1$	Dual-pair normalization (consistency between axis and readout)
$P(x)$	Residual projection: $P(x) = x - \ell(x)u \in \ker \ell$
$r(x)$	Residual: $r(x) = P(x)$
ϕ	1D tone function (e.g., $\phi(s) = \max(s, 0)^\alpha$)
α	Exponent parameter of the tone function
γ	Gamma correction (nonlinearity on the legacy side; reduced model)
F	Mapping representing an entire pipeline (Legacy/CONJ)
D	Evaluation sample set (random samples or a pixel set)
$C(F; D)$	Crosstalk metric (Eq. (5.1))
$\Sigma,\ \Delta\Sigma$	Residual covariance and its change (Eq. (5.2))
$\|\cdot\|_2,\ \|\cdot\|_F$	ℓ_2 norm and Frobenius norm

■Implementation remarks Constructing $A = B^{-1}$ explicitly is useful when one wants to expose the residual-plane coordinates (C_1, C_2) externally or when integrating into an existing linear-transform pipeline. On the other hand, if computational cost is the primary concern, the minimal form

$$x' = x + \big(\varphi(\ell(x)) - \ell(x)\big)\, u$$

(see (4.6) in the previous subsection) is more direct. In either implementation, finite precision and clipping may restrict the region where invariance effectively holds.

5 Numerical experiments: comparison with conventional pipelines

In this section, we validate by numerical experiments that the **choice of the dual pair (u, ℓ) defining the principal axis and the residual subspace, together with the normalization condition** $\ell(u) = 1$, decisively determines chroma invariance under the ordering of one-axis nonlinear operations—namely, whether we apply *nonlinearity→decomposition* or *decomposition→nonlinearity*.

■Reader's guide (purpose of this section) The comparison in this section is not intended to reproduce strictly standard-compliant pipelines. Rather, it is a **reduced model** designed to visualize how the difference in ordering—placing a nonlinearity before the $1 + 2$ decomposition (nonlinearity→decomposition) versus after it (decomposition→nonlinearity)—appears as crosstalk in the residual component when mapped back to the scene-linear layer. Notations such as sRGB/XYZ/BT.709 are used to represent typical definitions commonly encountered in practice.

Notation

5.1 Setup: reference spaces and normalization of the dual pair

All evaluations are performed in the scene-linear layer. We first generate linear-sRGB samples $x \in [0,1]^3$ and obtain XYZ samples by a linear transform:

$$x_{\mathrm{XYZ}} = M_{\mathrm{sRGB} \to \mathrm{XYZ}}\, x_{\mathrm{sRGB}}.$$

In the implementation, small negative values caused by numerical error are clipped to 0. The sRGB$\leftrightarrow$XYZ conversion used here follows the standard definition with the D65 white point [1, 3].

■**Normalization on the XYZ side**   On $V_{\mathrm{XYZ}} = \mathbb{R}^3$, we set the principal axis to coincide with the physical luminance Y:

$$u_Y := (0,1,0), \qquad \ell_Y(x) := x_Y.$$

Then $\ell_Y(u_Y) = 1$ holds.

■**Normalization on the sRGB side**   On the sRGB space $V_{\mathrm{sRGB}} = \mathbb{R}^3$, we define ℓ_s as the Y component after mapping linear sRGB to XYZ:

$$\ell_s(x) := \big(M_{\mathrm{sRGB} \to \mathrm{XYZ}}\, x\big)_Y.$$

We define the principal axis u_s by pulling back the XYZ-side axis u_Y to the sRGB side:

$$u_s := M_{\mathrm{XYZ} \to \mathrm{sRGB}}\, u_Y.$$

With these definitions, the normalization

$$\ell_s(u_s) = 1$$

is satisfied automatically.

5.2 Metrics

We define the projection (residual extraction) by

$$P(x) := x - \ell(x)u \in \ker \ell.$$

■**Crosstalk metric**   For a sample set D and a self-map $F : V \to V$, we define

$$C(F; D) := \frac{1}{|D|} \sum_{x \in D} \|P(F(x)) - P(x)\|_2^2. \tag{5.1}$$

The condition $C(F; D) = 0$ means that the residual is exactly invariant on D.

■**Change in chroma covariance**   Let the covariance of the residual $r(x) = P(x)$ be

$$\Sigma_{\mathrm{chroma}}(D) := \frac{1}{|D|} \sum_{x \in D} (r(x) - \bar{r})(r(x) - \bar{r})^\top, \quad \bar{r} := \frac{1}{|D|} \sum_{x \in D} r(x).$$

We measure the change by

$$\Delta\Sigma(F) := \big\|\Sigma_{\mathrm{chroma}}(F(D)) - \Sigma_{\mathrm{chroma}}(D)\big\|_F. \tag{5.2}$$

5.3   Experiment 1: sRGB side (legacy vs. CONJ)

We take the reference space to be V_{sRGB} and use (u_s, ℓ_s) defined above. The sample set is

$$D_{\mathrm{sRGB}} := \{x_i\}_{i=1}^{N}, \quad x_i \sim \mathrm{Unif}([0,1]^3).$$

■Legacy pipeline (reduced model)   As a representative example of applying a nonlinearity first and then performing a $1 + 2$ decomposition, we use the composite map

$$F_{\mathrm{legacy}}^{(s)} := \gamma^{-1} \circ T_{Y'CbCr}^{-1} \circ T_{Y'} \circ T_{Y'CbCr} \circ \gamma.$$

Here,

$$\gamma(x) = x^{1/\gamma}, \quad \gamma^{-1}(x) = x^{\gamma},$$

$T_{Y'CbCr}$ is a simplified $R'G'B' \leftrightarrow Y'CbCr$ transform based on BT.709 coefficients, and $T_{Y'}$ is

$$Y' \mapsto \mathrm{clip}(Y', 0, 1)^{\alpha}.$$

■Remark (reduction in the legacy model)   Our $R'G'B' \leftrightarrow Y'CbCr$ transform is a **simplified model** meant to represent the order mismatch, and is not intended to reproduce transmission-accurate implementations in standards, including limited/full range handling, offsets, or quantization. See Appendix B for the explicit formulas used.

■CONJ pipeline   In CONJ, we perform the decomposition in the scene-linear layer and apply the nonlinearity only to the principal-axis component:

$$F_{\mathrm{conj}}^{(s)}(x) = \phi(\ell_s(x))u_s + \big(x - \ell_s(x)u_s\big), \quad \phi(s) = \max(s, 0)^{\alpha}.$$

5.4   Experiment 2: XYZ/Lab side (Lab processing vs. CONJ on XYZ)

We map the same samples to XYZ and form

$$D_{\mathrm{XYZ}} := \{M_{\mathrm{sRGB}\to\mathrm{XYZ}}x_i\}_{i=1}^{N}.$$

The evaluation is performed on scene-linear XYZ, using (u_Y, ℓ_Y) as the dual pair.

■CIELAB-based legacy procedure   We compare against the following mapping:

$$F_{\mathrm{Lab\text{-}legacy}} := T_{\mathrm{Lab}\to\mathrm{XYZ}}^{-1} \circ T_{L^*} \circ T_{\mathrm{XYZ}\to\mathrm{Lab}}.$$

Here, T_{L^*} is defined by

$$L^* \mapsto 100 \cdot \mathrm{clip}\left(\frac{L^*}{100}, 0, 1\right)^{\alpha},$$

while a^*, b^* are preserved. In this experiment, after applying the L^* operation in Lab space, we map back to scene-linear XYZ and then evaluate the residual.

Table2: Evaluation results for uniformly random samples ($N = 100\,000$, $\alpha = 0.8$, $\gamma = 2.2$).

Mapping F	$C(F; D)$	$\Delta\Sigma(F)$
sRGB: Legacy	4.62×10^{-2}	3.27×10^{-2}
sRGB: CONJ	4.38×10^{-32}	2.28×10^{-17}
XYZ/Lab: Legacy	1.25×10^{-2}	1.42×10^{-2}
XYZ: CONJ	0	0

■CONJ on XYZ

$$F_{\text{XYZ-conj}}(x) = \phi(\ell_Y(x))u_Y + \big(x - \ell_Y(x)u_Y\big), \quad \phi(s) = \max(s, 0)^{\alpha}.$$

5.5 Results

On both the sRGB side and the XYZ/Lab side, the legacy ordering (nonlinearity→decomposition) leaves residual variations on the order of 10^{-2} when evaluated after mapping back to the scene-linear layer. In contrast, CONJ—where the decomposition is performed first under the dual-pair normalization $\ell(u) = 1$—reduces residual variation to the level of machine precision in double-precision arithmetic.

In particular, the results $C(F; D) = 0$ and $\Delta\Sigma(F) = 0$ for CONJ on XYZ numerically reflect the fact that the direct-sum decomposition induced by $\ell_Y(u_Y) = 1$ is algebraically exact.

6 Validation on a real photograph

6.1 Input and preprocessing (unifying to linear light)

The input is an 8-bit RGB image (JPEG) encoded in sRGB, normalized to $[0, 1]$. Since the $1 + 2$ decomposition (principal axis + residual) in this paper is defined and evaluated in **linear light**, we apply the inverse sRGB EOTF E^{-1} to the input x_{srgb} and obtain

$$x_{\text{lin}} := E^{-1}(x_{\text{srgb}}). \tag{6.1}$$

(Note) E^{-1} denotes the inverse sRGB EOTF applied per channel; the explicit formula is given in Appendix B.

6.2 Dual pair (definition of the principal axis and the residual)

We define the readout $\ell(x) = c^{\top}x$ using the linear luminance coefficients of Rec. 709 / sRGB,

$$c := (0.2126,\ 0.7152,\ 0.0722).$$

For the achromatic axis we use

$$u := (1, 1, 1),$$

and assume it has been normalized so that $\ell(u) = 1$. Then each pixel x is uniquely decomposed as

$$x = \ell(x)\,u + r(x), \qquad r(x) := x - \ell(x)\,u \in \ker \ell. \tag{6.2}$$

6.3 One-axis tone function

We use the one-axis tone function

$$\varphi_\alpha(s) := \mathrm{clip}(s, 0, 1)^\alpha, \tag{6.3}$$

with the default $\alpha = 0.8$.

6.4 Two pipelines to be compared

For a real image, we compare the following two pipelines (all evaluations are performed in linear light).

■(A) CONJ (decompose in linear light → apply a nonlinearity only to the principal axis → recompose)

$$x_{\mathrm{lin}} := E^{-1}(x_{\mathrm{srgb}}), \tag{6.4}$$

$$Y_{\mathrm{lin}} := \ell(x_{\mathrm{lin}}), \qquad r_{\mathrm{lin}} := x_{\mathrm{lin}} - Y_{\mathrm{lin}}u, \tag{6.5}$$

$$x_{\mathrm{conj,lin}} := \varphi_\alpha(Y_{\mathrm{lin}})\,u + r_{\mathrm{lin}}. \tag{6.6}$$

■(B) Legacy (decompose in a nonlinear (sRGB) space → apply a principal-axis nonlinearity → recompose)
As a reduced model representing the order mismatch of "performing the decomposition in a nonlinear space," we define the legacy pipeline by

$$Y_{\mathrm{gam}} := \ell(x_{\mathrm{srgb}}), \qquad r_{\mathrm{gam}} := x_{\mathrm{srgb}} - Y_{\mathrm{gam}}u, \tag{6.7}$$

$$x_{\mathrm{legacy,srgb}} := \mathrm{clip}\big(\varphi_\alpha(Y_{\mathrm{gam}})\,u + r_{\mathrm{gam}}, 0, 1\big), \tag{6.8}$$

$$x_{\mathrm{legacy,lin}} := E^{-1}(x_{\mathrm{legacy,srgb}}), \tag{6.9}$$

where the clipping enforces the range $[0, 1]$.

6.5 Evaluation metrics (linear light)

Let the residuals in linear light be

$$r_{\mathrm{orig}} := r(x_{\mathrm{lin}}), \quad r_{\mathrm{conj}} := r(x_{\mathrm{conj,lin}}), \quad r_{\mathrm{legacy}} := r(x_{\mathrm{legacy,lin}}),$$

where $r(x) = x - \ell(x)u$.

■Crosstalk magnitude For the pixel set D_{img}, we use

$$C_{\mathrm{img}} := \frac{1}{|D_{\mathrm{img}}|} \sum_{p \in D_{\mathrm{img}}} \|r_{\mathrm{out}}(p) - r_{\mathrm{orig}}(p)\|_2^2, \tag{6.10}$$

with $\mathrm{out} \in \{\mathrm{conj}, \mathrm{legacy}\}$.

Table3: Crosstalk evaluation on a real photograph (`YKT_3336`, $\alpha = 0.8$), evaluated in linear light.

	Legacy	CONJ
C_{img}	6.9615×10^{-4}	8.2927×10^{-33}
$\Delta\Sigma_{\mathrm{img}}$ (Fro)	3.5171×10^{-3}	9.4495×10^{-18}

■Change in chroma covariance   Using the centered covariance of residuals

$$\Sigma(r) := \mathbb{E}\big[(r - \mathbb{E}[r])(r - \mathbb{E}[r])^\top\big],$$

we evaluate

$$\Delta\Sigma_{\mathrm{img}} := \|\Sigma(r_{\mathrm{out}}) - \Sigma(r_{\mathrm{orig}})\|_F \tag{6.11}$$

with the Frobenius norm.

6.6   Results

Table 3 shows the results for the image `YKT_3336.jpg` with $\alpha = 0.8$. Due to residual invariance, CONJ suppresses C_{img} and $\Delta\Sigma_{\mathrm{img}}$ down to the rounding-error level of double-precision arithmetic, whereas the legacy pipeline exhibits observable residual variations caused by the order mismatch of performing the decomposition in a nonlinear space.

6.7   Remarks (implementation details)

This experiment uses an 8-bit JPEG as input and therefore includes quantization error at the input stage. The legacy pipeline also includes clipping to $[0, 1]$, so the measured values include the effect of out-of-gamut handling. Nevertheless, the comparison focuses on the ordering difference—whether the decomposition is performed in linear light or in sRGB space—when applying the same 1D tone function, and the evaluation is conducted in linear light so as to remain consistent with the definition of the $1 + 2$ decomposition.

7   Conclusion

In this paper, we introduced CONJ as a minimal framework for treating, in a consistent manner on the scene-linear (linear-light) domain, the widely recurring $1 + 2$ structure in color spaces, namely a decomposition into a one-dimensional principal axis and a two-dimensional residual. CONJ is built from a direct-sum decomposition induced by a dual pair (u, ℓ) and a one-axis nonlinear operator.

The point of this paper is not to argue the superiority of any specific existing color space or coding scheme. Rather, we emphasize that the **choice of the pair (u, ℓ) determining the principal axis and the residual, together with the normalization condition**

$$\ell(u) = 1, \tag{7.1}$$

structurally determines the independence between principal-axis processing and residual processing (i.e., crosstalk suppression).

Under this normalization, any $x \in V$ admits the unique decomposition

$$x = \ell(x)\, u + r, \qquad r \in \ker \ell, \tag{7.2}$$

and hence $V = \mathrm{span}(u) \oplus \ker \ell$ holds. Based on this decomposition, defining the CONJ operator that applies a one-axis nonlinearity φ only to the principal scalar,

$$T_\varphi(x) = \varphi(\ell(x))\, u + \big(x - \ell(x)u\big), \tag{7.3}$$

we showed that, under ideal real-arithmetic, the residual r is algebraically invariant, independent of the specific form of φ.

Representative color spaces such as sRGB, YCbCr, CIELAB, IPT, ICtCp, and Oklab exhibit a $1 + 2$ structure in appearance, but many of them are designed in the form "apply a nonlinearity first, then introduce a $1 + 2$ decomposition," i.e., the $N3 \to S$ type. In this case, when one pulls the signal back to the scene-linear domain and evaluates the residual, residual invariance in the CONJ sense is not guaranteed in general. We organized this point structurally.

In numerical experiments on scene-linear sRGB and scene-linear XYZ, we compared a legacy pipeline (nonlinearity $\to$ decomposition) with a CONJ-type pipeline (decomposition $\to$ one-axis nonlinearity). While the legacy pipeline exhibited residual variations on the order of 10^{-2}, the CONJ-type pipeline reduced them down to the rounding-error level of double-precision floating-point arithmetic.

Moreover, in the validation on a real photograph, when applying the same one-axis tone operation, CONJ reduced C_{img} and $\Delta\Sigma_{\mathrm{img}}$ to the rounding-error level, confirming that contamination into the residual component is strongly suppressed. In contrast, the reduced legacy procedure that performs the decomposition in a nonlinear (sRGB) space exhibited remaining residual variation when evaluated on the scene-linear domain. This is consistent with our organization that the ordering between decomposition and one-axis nonlinearity governs the amount of crosstalk.

The residual invariance established in this paper is an algebraic property of (7.3). In practical implementations, however, its validity can be limited by finite-precision arithmetic, quantization (e.g., 3D LUTs), and operations involving clipping or out-of-gamut handling. Accordingly, an important next step is to quantify, under implementation constraints, "to what extent the residual can be treated as effectively invariant," and to examine its relationship to coding efficiency and perceptual difference metrics.

In summary, CONJ provides a minimal and general template for organizing and redesigning color-space design and coding operations in the scene-linear domain, in that it enables **(i) preserving the linear structure of the principal axis and the residual plane, while (ii) introducing nonlinear degrees of freedom only along the principal axis.**

■Code availability.   A reference implementation and experiment scripts corresponding to this preprint are available in the following GitHub repository: `https://github.com/goldkiss2010-ai/conj-preprint`.

8   Outlook: discussion points for follow-up research

Within the scope of this paper, we showed that CONJ, defined under the dual-pair normalization $\ell(u) = 1$, preserves the residual plane under ideal real-arithmetic, and that its numerical crosstalk behavior differs substantially from representative legacy procedures (Section 5). Here we briefly list discussion points that are suitable to be separated as follow-up work.

8.1   "Effective invariance" under finite precision and quantization

While residual invariance is an algebraic property, finite precision dominates in implementations. It is therefore useful to quantify the "region where invariance holds in practice" as a function of:

- floating-point precision (float64/float32/fp16) and fixed-point implementations,
- 3D LUT quantization (e.g., $17^3, 33^3, 65^3$) and interpolation (trilinear/tetrahedral, etc.),
- the placement of rounding within a pipeline (pre-transform / φ / post-transform).

As metrics, the $C(F; D)$ and $\Delta\Sigma(F)$ defined in Section 5 can be re-measured under each quantization condition for direct comparison.

8.2   Operating conditions including clipping and out-of-gamut handling

Because residual invariance relies on the structure on $\ker \ell$, clipping and out-of-gamut handling become major sources of violation. It is desirable to separate and evaluate:

- no clipping (ideal) vs. clipping at each stage (operational),
- differences among out-of-gamut strategies (simple clipping / projection / gamut mapping),
- the coupling between φ and range management in HDR (large dynamic range).

This point can be organized as a trade-off between the design goal "preserve the residual plane" and the projection to safe display/coding domains.

8.3   Systematizing the choice/estimation of (u, ℓ)

This paper focused on properties given a fixed (u, ℓ), but in practice the key question is "which (u, ℓ) is desirable." Natural directions include:

- standard-based fixed choices (e.g., Y in XYZ, linear luminance based on standard matrices),
- data-driven choices (e.g., local PCA/KLT-type selection, or objective-based optimization),
- conditional choices (scene-dependent / exposure-region-dependent) and suppressing discontinuities due to switching.

Evaluation can be formulated as multi-objective optimization, incorporating not only crosstalk but also amplification of quantization error and clipping robustness.

8.4 Connecting to coding and editing pipelines

The separation in CONJ—"a 1D nonlinearity on the principal scalar" and "preservation of the residual plane"—is convenient as a design unit in coding and editing. Possible topics include:

- principal-axis formulation of tone mapping / exposure adjustment and stable preservation of chromatic difference components,
- the possibility of reducing crosstalk by redesigning the decomposition order for existing $1 + 2$ representations (e.g., YCbCr),
- linking rate control / quantization to residual statistics (e.g., $\Delta\Sigma$).

These are best viewed not as extensions of the theoretical claim itself, but as engineering studies that measure the attainable gains under operational constraints.

AppendixA $1 + 2$ Decomposition and One-Axis Nonlinear Operators from a Dual Pair

A.1 Setup and notation

Let V be a real vector space and $V^* := \mathrm{Hom}(V, \mathbb{R})$ its algebraic dual. We write the evaluation pairing as

$$\langle \ell, x \rangle := \ell(x) \qquad (\ell \in V^*,\ x \in V).$$

A.2 A dual pair and the associated projections

Definition AppendixA.1 ((Normalized) dual pair and associated projections)**.** *Let $u \in V$ and $\ell \in V^*$ satisfy*

$$\langle \ell, u \rangle = 1.$$

Then we call (u, ℓ) a (normalized) dual pair. Define linear maps $Q, P : V \to V$ by

$$Q(x) := \langle \ell, x \rangle\, u, \qquad P := I - Q,$$

where I denotes the identity map.

Lemma AppendixA.2 (Rank-one projections and the $1 + 2$ decomposition)**.** *For the maps Q and P defined above, the following identities hold:*

$$Q^2 = Q, \quad P^2 = P, \quad PQ = QP = 0.$$

Moreover,

$$\mathrm{Im}\,Q = \mathrm{span}(u), \qquad \ker Q = \ker \ell, \qquad \mathrm{Im}\,P = \ker \ell.$$

Consequently, the direct-sum decomposition

$$V = \mathrm{span}(u) \oplus \ker \ell$$

holds, and every $x \in V$ can be written uniquely as

$$x = \underbrace{\langle \ell, x \rangle}_{=:Y(x)} u + \underbrace{P(x)}_{=:r(x)}, \qquad r(x) \in \ker \ell.$$

Proof. For any $x \in V$,

$$Q^2(x) = Q(\langle \ell, x \rangle u) = \langle \ell, x \rangle \langle \ell, u \rangle u = Q(x),$$

hence $Q^2 = Q$. Since $P = I - Q$, we have $P^2 = (I - Q)^2 = I - 2Q + Q^2 = I - Q = P$. Also, $PQ = P(I - P) = P - P^2 = 0$, and similarly $QP = (I - P)P = 0$.

For the images, $\operatorname{Im} Q = \operatorname{span}(u)$ follows from $Q(x) = \langle \ell, x \rangle u$. For the kernel,

$$Q(x) = 0 \iff \langle \ell, x \rangle = 0 \iff x \in \ker \ell,$$

so $\ker Q = \ker \ell$.

Next we show $\operatorname{Im} P = \ker \ell$. If $x \in \ker \ell$, then $Q(x) = 0$ and thus $P(x) = x$, hence $\ker \ell \subset \operatorname{Im} P$. Conversely, for any x,

$$\langle \ell, P(x) \rangle = \langle \ell, x - Q(x) \rangle = \langle \ell, x \rangle - \langle \ell, x \rangle \langle \ell, u \rangle = 0,$$

so $\operatorname{Im} P \subset \ker \ell$. Therefore $\operatorname{Im} P = \ker \ell$.

Finally, for any x we have $x = Q(x) + P(x)$ with $Q(x) \in \operatorname{span}(u)$ and $P(x) \in \ker \ell$. To see that the intersection is trivial, let $v \in \operatorname{span}(u) \cap \ker \ell$. Write $v = au$. Then $0 = \langle \ell, v \rangle = a \langle \ell, u \rangle = a$, hence $a = 0$ and $v = 0$. This proves the direct-sum decomposition and uniqueness of the representation. □

A.3 One-axis nonlinear operators (residual invariance)

Definition AppendixA.3 (One-axis nonlinear operator)**.** *Given a function $\varphi : \mathbb{R} \to \mathbb{R}$, define $T_\varphi : V \to V$ by*

$$T_\varphi(x) := \varphi(\langle \ell, x \rangle)u + P(x).$$

Lemma AppendixA.4 (Residual invariance (projection form))**.** *For the map T_φ defined above, for any $x \in V$,*

$$P(T_\varphi(x)) = P(x)$$

holds.

Proof. Using $P(u) = 0$ and $P^2 = P$,

$$P(T_\varphi(x)) = P\big(\varphi(\langle \ell, x \rangle)u + P(x)\big) = \varphi(\langle \ell, x \rangle)P(u) + P^2(x) = P(x).$$

□

AppendixB   Definitions of the simplified transforms used in the experiments (supplement)

B.1   Simplified $R'G'B' \leftrightarrow Y'CbCr$ (full range)

In the Legacy pipeline, we use a simplified full-range transform based on the BT.709 luma coefficients $(K_r, K_g, K_b) = (0.2126, 0.7152, 0.0722)$ in order to represent the effect of order inconsistency. We do not

model offsets or quantization ranges in the standard.

■Forward transform

$$Y' = K_r R' + K_g G' + K_b B', \quad Cb = \frac{B' - Y'}{2(1 - K_b)}, \quad Cr = \frac{R' - Y'}{2(1 - K_r)}.$$

■Inverse transform

$$R' = Y' + 2(1 - K_r)\, Cr, \quad B' = Y' + 2(1 - K_b)\, Cb, \quad G' = \frac{Y' - K_r R' - K_b B'}{K_g}.$$

B.2 Inverse sRGB EOTF (reconstruction to linear light)

In the real-image experiment, to unify the input $x_{\mathrm{srgb}} \in [0, 1]$ in linear light, we apply the inverse sRGB EOTF component-wise:

$$E^{-1}(v) = \begin{cases} \frac{v}{12.92}, & v \le 0.04045, \\ \left(\frac{v + 0.055}{1.055}\right)^{2.4}, & v > 0.04045. \end{cases}$$

Note. These formulas are simplified forms that omit conditions required for a fully standard-compliant implementation (coding range, quantization, offsets, etc.). They are used in this paper to make the effect of order inconsistency explicit.

References

[1] Commission Internationale de l'Éclairage (CIE). Colorimetry, 4th edition. CIE Publication CIE 15:2018, CIE, 2018.

[2] Commission Internationale de l'Éclairage (CIE). Official recommendations on uniform color spaces, color difference equations and metric color terms. CIE Publication CIE 15, Supplement 2, CIE, 1976.

[3] International Electrotechnical Commission. Multimedia systems and equipment – colour measurement and management – part 2-1: Colour management – default rgb colour space – sRGB. International Standard IEC 61966-2-1:1999, IEC, 1999. (accessed 2025-12-19).

[4] Erik Reinhard, Greg Ward, Sumanta Pattanaik, and Paul Debevec. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting.* Morgan Kaufmann, 2005.

[5] Radosław Mantiuk, Rafał Mantiuk, Anna Tomaszewska, and Wolfgang Heidrich. Color correction for tone mapping. *Computer Graphics Forum*, 28(2):193–202, 2009. doi: 10.1111/j.1467-8659.2009 .01358.x. Proc. Eurographics 2009.

[6] Imran Mehmood, Muhammad Usman Khan, and Mingronnier Luo. Generic color correction for tone mapping operators in high dynamic range imaging. *Optics Express*, 32(16):27849–27866, 2024. doi: 10.1364/OE.527492.

[7] International Telecommunication Union. Recommendation ITU-R BT.2100-2: Image parameter values for high dynamic range television for use in production and international programme exchange.

Technical Report BT.2100-2, ITU-R, July 2018. `https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2100-2-201807-I%21%21PDF-E.pdf` (accessed 2025-12-19).

 [8] International Telecommunication Union. Recommendation ITU-R BT.601-7: Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios. Technical Report BT.601-7, ITU-R, March 2011. `https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-I!!PDF-E.pdf` (accessed 2025-12-19).

 [9] International Telecommunication Union. Recommendation ITU-R BT.709-6: Parameter values for the HDTV standards for production and international programme exchange. Technical Report BT.709-6, ITU-R, June 2015. `https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.709-6-201506-I!!PDF-E.pdf` (accessed 2025-12-19).

[10] Felix Ebner and Mark D. Fairchild. Development and testing of a color space (IPT) with improved hue uniformity. In *Proceedings of the 6th Color Imaging Conference: Color Science, Systems and Applications*, pages 8–13. Society for Imaging Science and Technology (IS&T), 1998. IS&T/SID Color Imaging Conference.

[11] Dolby Laboratories. ICtCp: A new color representation for HDR imaging. White paper, Dolby Laboratories, 2016. Version 0.7.1, `https://professional.dolby.com/siteassets/pdfs/ictcp_dolbywhitepaper_v071.pdf` (accessed 2025-12-19).

[12] Björn Ottosson. Oklab: A perceptual color space for image processing, 2020. `https://bottosson.github.io/posts/oklab/` (accessed 2025-12-19).

[13] Greg Ward Larson. The logluv encoding for full gamut, high dynamic range images. *Journal of Graphics Tools*, 3(1):15–31, 1998. doi: 10.1080/10867651.1998.10487485.