Web-Accessible Color Palette Metrics

Huu Khanh Nguyen

***Abstract*—The design of accessible color systems for digital interfaces currently lacks unified quantitative standards, leading to subjective and manual workflows. This document introduces a quantitative framework for evaluating sequential monochromatic color palettes by defining five principal evaluation dimensions: Contrast Efficiency, lightness linearity, chroma smoothness, hue stability, and spacing uniformity. The framework adopts the CIELAB color space for its interpretability and its role as the reference color space for computing perceptual differences using CIEDE2000. By benchmarking eleven industry-leading design systems, the framework objectively differentiates palette quality, with composite scores ranging from 59.01 to 88.69, identifying IBM Carbon and Adobe Spectrum as high-performing benchmarks. This work establishes a quantitative foundation for systematic palette evaluation and comparative analysis.**

Keywords: color ramp; colormap; monochromatic; color accessibility; color palette; design systems

An open-source reference implementation and benchmarks are available at
https://github.com/chromametry/chromametry

I. Introduction

Designing color systems for digital interfaces requires balancing uniformity with accessibility constraints. In practice, the absence of a unified quantitative evaluation framework leads to several key limitations:

- **High subjectivity:** Color palette quality assessment remains largely perceptual, lacking objective measurement indicators for comparing different systems.
- **Automation barriers:** Without standardized metrics, the generation of systematic color scales (M hues $\times$ N steps) remains predominantly manual. Existing interpolation methods struggle to preserve linearity and accessibility in the absence of quantitative validation criteria.
- **Combinatorial contrast testing:** Verifying WCAG-compliant contrast across all color pairs is operationally inefficient, often requiring exhaustive inspection during both design and deployment stages.
- **Poor maintainability and scalability:** Color systems without quantitative foundations face difficulties when extending palette ranges or performing synchronized updates while preserving brand-specific characteristics.

This work aims to shift color system design from a heuristic-driven process toward a metrics-based methodology, establishing a quantitative reference for systematic palette evaluation and comparison.

II. Definitions and Standards

A. Color Palette Structure

A palette P comprises n monochromatic families:

$$P = \{S_1, S_2, \ldots, S_n\}.$$

Each color family S_j is a *monochromatic ramp* consisting of N lightness steps, generated from a single base color. Formally,

$$S_j = \{c_{j,0}, c_{j,1}, \ldots, c_{j,N-1}\}.$$

Colors within each family are arranged in a strictly monotonic order of perceived lightness, forming a one-dimensional lightness ramp. Within S_j, there exists a unique base color $C_{\text{base},j} \in S_j$ satisfying

$$C_{\text{base},j} = \arg\max_{c_{j,i} \in S_j} C^*(c_{j,i}).$$

This color possesses the maximum chroma in the range and serves as the representative hue for the entire family. By definition, S_j is assumed to be the result of a fixed-hue interpolation where only lightness and its dependent chroma vary. This assumption ensures consistent reference quantities $(L_{\text{base}}, h_{\text{base}}, C_{\text{base}})$ for all subsequent metrics.

B. WCAG Contrast Ratio

According to WCAG 2.1 [6], the contrast ratio is defined based on relative luminance $Y \in [0, 1]$:

$$CR(c_1, c_2) = \frac{Y_1 + 0.05}{Y_2 + 0.05}, \quad Y_1 \geq Y_2.$$

Relative luminance is calculated as

$$Y = 0.2126\, R_{\text{linear}} + 0.7152\, G_{\text{linear}} + 0.0722\, B_{\text{linear}}.$$

Standard contrast ratio thresholds include:

- **3:1** for UI components and large text.
- **4.5:1** for body text (AA standard).
- **7:1** for enhanced readability (AAA).

This framework identifies the **4.5:1 threshold** as the primary benchmark for comparative evaluation.

C. Contrast Span

The *Contrast Span* K of a hue family is the minimum index distance that guarantees a contrast ratio ≥ 4.5:1 for all color pairs:

$$K = \min \{k \in \mathbb{N} : CR(c_i, c_{i+k}) \geq 4.5,\ \forall i \in [0, N-k-1]\}.$$

For multi-family palettes, the global span is defined as

$$K_{\text{palette}} = \max_{j=1}^{n} K_j.$$

This metric enables predictable accessibility: designers can ensure compliant contrast by selecting colors separated by at least K steps without requiring runtime calculations. As illustrated in Table I, a fixed span guarantees WCAG-compliant contrast across all valid index pairs.

TABLE I
ILLUSTRATION USING $N = 12$ WITH CONTRAST SPAN $K = 6$.

#	Indices	0	1	2	3	4	5	6	7	8	9	10	11
1	span 0–6	[0	1	2	3	4	5	6]	7	8	9	10	11
2	span 1–7	0	[1	2	3	4	5	6	7]	8	9	10	11
3	span 2–8	0	1	[2	3	4	5	6	7	8]	9	10	11
4	span 3–9	0	1	2	[3	4	5	6	7	8	9]	10	11
5	span 4–10	0	1	2	3	[4	5	6	7	8	9	10]	11
6	span 5–11	0	1	2	3	4	[5	6	7	8	9	10	11]

III. EVALUATION METRICS

This section defines five quantitative metrics for evaluating sequential monochromatic color palettes. Three metrics assess the internal behavior of individual CIELAB components: lightness linearity, chroma smoothness, and hue stability, capturing how each perceptual dimension evolves along a color ramp. A fourth metric evaluates perceptual spacing uniformity using the CIEDE2000 color difference, quantifying the consistency of stepwise transitions. The final metric measures contrast efficiency by analyzing the accessibility span required to satisfy the WCAG 4.5:1 contrast ratio. Together, these five metrics characterize both the perceptual structure and accessibility performance of a color palette.

A. Contrast Efficiency (η)

Contrast Efficiency evaluates the economy of index separation required to satisfy accessibility constraints. To establish an ideal upper bound for accessibility, this framework references a **neutral gray ramp** ($a^* = 0, b^* = 0$).

The choice of an achromatic scale is motivated by its inherent stability: neutral colors provide the most direct mapping between relative luminance Y and perceived lightness L^*, as they are unaffected by the Helmholtz–Kohlrausch effect—where high chroma increases perceived brightness despite constant luminance [3]. Furthermore, neutral scales occupy the achromatic axis of the color gamut, minimizing non-linearities induced by gamut mapping constraints common in highly saturated hues.

In this ideal achromatic case, L^* is related to relative luminance $Y \in [0, 1]$ by the CIE 1976 transform [1]:

$$L^*(Y) = \begin{cases} 116\, Y^{1/3} - 16, & Y > \left(\frac{6}{29}\right)^3 \\ 116\left(\frac{29}{6}\right)^2 Y + 16, & Y \leq \left(\frac{6}{29}\right)^3 \end{cases}$$

Applying the WCAG 4.5:1 boundary conditions:

- For $Y_{\min} = 0$ (black background), the required $Y_{\text{text}} = 0.175 \Rightarrow L^* \approx 48.9$.
- For $Y_{\max} = 1$ (white background), the required $Y_{\text{text}} \approx 0.183 \Rightarrow L^* \approx 49.9$.

Accordingly, the conservative lightness threshold is defined as

$$L^*_{\text{target}} \approx 49.9.$$

Since the CIELAB lightness axis spans $[0, 100]$, this corresponds to a normalized lightness ratio

$$\lambda = \frac{L^*_{\text{target}}}{100} \approx 0.50.$$

TABLE II
REFERENCE VALUES: K_{IDEAL} (DESIGN TARGET) AND D_{IDEAL} (METRIC CALCULATION)

Steps (N)	Formula	K_{ideal}	D_{ideal}
10	$\lceil 0.50 \times 9 \rceil$	5	0.450
11	$\lceil 0.50 \times 10 \rceil$	5	0.455
12	$\lceil 0.50 \times 11 \rceil$	6	0.458
13	$\lceil 0.50 \times 12 \rceil$	6	0.462
14	$\lceil 0.50 \times 13 \rceil$	7	0.464
15	$\lceil 0.50 \times 14 \rceil$	7	0.467
16	$\lceil 0.50 \times 15 \rceil$	8	0.469
17	$\lceil 0.50 \times 16 \rceil$	8	0.471
18	$\lceil 0.50 \times 17 \rceil$	9	0.472

Although the exact luminance-derived ratio yields $\lambda \approx 0.49$, this work adopts a rounded value $\lambda = 0.50$ to avoid false numerical precision and to maintain symmetry along the CIELAB lightness axis.

For a palette with N discrete steps, the **ideal density** is defined as:

$$D_{\text{ideal}} = \lambda \cdot \frac{N - 1}{N}$$

For practical design guidance, the corresponding **ideal contrast span** is:

$$K_{\text{ideal}} = \lceil \lambda \cdot (N - 1) \rceil$$

Note that D_{ideal} is used for metric calculation to ensure smooth behavior, while K_{ideal} serves as a discrete design target. For example, with $N = 12$:

- $D_{\text{ideal}} = 0.5 \times \frac{11}{12} \approx 0.458$
- $K_{\text{ideal}} = \lceil 5.5 \rceil = 6$

This formulation ensures smooth metric behavior across different palette sizes and eliminates the aliasing effects inherent in ceiling-based calculations. As $N \to \infty$, $D_{\text{ideal}} \to \lambda = 0.5$.

Let D be the actual contrast density, where K is the measured contrast span. Contrast Efficiency η is defined by:

$$\eta = \begin{cases} 1, & D \leq D_{\text{ideal}} \\ 0, & D \geq 1.0 \\ 1 - \dfrac{D - D_{\text{ideal}}}{1.0 - D_{\text{ideal}}}, & \text{otherwise} \end{cases}$$

B. Lightness Linearity ($\mathcal{L}$)

The **Lightness Linearity** metric evaluates the degree to which *perceptual* lightness progression in a palette follows a linear trend with step index. Unlike pure L^*, lightness is affected by the Helmholtz–Kohlrausch effect [3], where high-chroma colors are perceived as lighter than achromatic colors with the same L^*.

1) Equivalent Achromatic Lightness: To compensate for this effect, we use **Equivalent Achromatic Lightness (EAL)** according to High et al. [3]:

$$L_{\text{EAL}} = L^* + \big(f_{BY}(h) + f_R(h)\big)\, C^*$$

where chroma and hue in CIELAB space are determined by

$$C = \sqrt{a^{*2} + b^{*2}}$$

$$h = \operatorname{atan2}(b^*, a^*)$$

The hue-dependent correction functions are given by

$$f_{BY}(h) = 0.1644 \left| \sin\left(\frac{h - 90^\circ}{2} \right) \right| + 0.0603$$

$$f_R(h) = \begin{cases} 0.1307\,|\cos(h)| + 0.0060, & h \in [0^\circ, 90^\circ] \cup [270^\circ, 360^\circ], \\ 0, & \text{otherwise.} \end{cases}$$

2) Linear Regression by Step Index: For a hue family of N colors ordered by **monotonic lightness progression** (either strictly increasing or strictly decreasing), let $L_{\mathrm{EAL},i}$ denote the EAL value at step i. A linear model is fitted using **ordinary least squares**:

$$\hat{L}_{\mathrm{EAL}}(i) = \alpha i + \beta, \qquad i = 0, \ldots, N-1.$$

Here, the sign of α implicitly captures the direction of progression: $\alpha > 0$ corresponds to increasing lightness, while $\alpha < 0$ corresponds to decreasing lightness.

The root mean square error is computed as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} \left(L_{\mathrm{EAL},i} - \hat{L}_{\mathrm{EAL}}(i) \right)^2}.$$

3) Normalization by Fitted Range: To normalize error independently of absolute lightness magnitude and progression direction, error is normalized by the **extremal values of the fitted line**:

$$L^{\text{fit}}_{\min} = \min\big(\hat{L}_{\mathrm{EAL}}(0),\ \hat{L}_{\mathrm{EAL}}(N-1)\big).$$

$$L^{\text{fit}}_{\max} = \max\big(\hat{L}_{\mathrm{EAL}}(0),\ \hat{L}_{\mathrm{EAL}}(N-1)\big).$$

At each step i, the maximum allowable deviation within this fitted range is

$$\epsilon_i^{\max} = \max\left(\left| \hat{L}_{\mathrm{EAL}}(i) - L^{\text{fit}}_{\min} \right|, \left| L^{\text{fit}}_{\max} - \hat{L}_{\mathrm{EAL}}(i) \right| \right).$$

From this, the maximum normalized error is defined as

$$RMSE_{\max} = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} \left(\epsilon_i^{\max} \right)^2}.$$

This normalization constrains the metric to the envelope of the fitted linear model and is invariant to whether the palette progresses from dark-to-light or light-to-dark, ensuring score stability independent of absolute position or direction on the L^* axis.

4) Metric Definition: The **Lightness Linearity** metric is defined by

$$\mathcal{L} = \max\left(0,\ 1 - \frac{RMSE}{RMSE_{\max}} \right).$$

The value $\mathcal{L} \in [0, 1]$ approaches 1 when lightness progression closely follows a linear trend; strong local fluctuations or non-linearity will decrease the score.

In degenerate cases where the fitted line magnitude is negligible (palette nearly flat in lightness), the metric is conventionally set to $\mathcal{L} = 1$.

5) Interpretation: This metric evaluates **consistency of lightness progression** rather than purely geometric L^*. Using EAL allows $\mathcal{L}$ to more accurately reflect users' visual perception, especially for high-chroma palettes where the Helmholtz–Kohlrausch effect plays a significant role.

C. Chroma Smoothness ($\mathcal{S}_C$)

1) Theoretical Background: Zeileis et al. (2009) Power Function Model [4]: Zeileis et al. [4] propose a power-function-based parameterization to control the rate of chroma and lightness variation along a sequential color scale, improving contrast distribution. Using a *normalized continuous position* $t \in [0, 1]$, the model is defined as

$$\begin{cases} H(t) = H_2 - t\,(H_1 - H_2), \\ C(t) = C_{\max} - t^{p_1}\,(C_{\max} - C_{\min}), \\ L(t) = L_{\max} - t^{p_2}\,(L_{\max} - L_{\min}), \end{cases} \tag{1}$$

where H_1, H_2 are hue values at the two ends of the scale; $C_{\max}, C_{\min}$ and $L_{\max}, L_{\min}$ are chroma and lightness bounds; and $p_1, p_2 > 0$ control the curvature of chroma and lightness variation.

2) Derivative Discontinuity and Kink Artifacts: Although the single power-function model guarantees monotonicity, applying it to sequential palettes with *maximum chroma near the center* typically requires piecewise definitions (dark-to-peak and peak-to-light). The chroma derivative is

$$\frac{dC}{dt} = -p_1\, t^{p_1 - 1}\,(C_{\max} - C_{\min}). \tag{2}$$

When parameters differ across the two segments, the first derivative becomes discontinuous at the peak position t_{peak}:

$$\lim_{t \to t^-_{\text{peak}}} \frac{dC}{dt} \neq \lim_{t \to t^+_{\text{peak}}} \frac{dC}{dt}. \tag{3}$$

This lack of C^1 continuity produces a sharp chroma cusp, which can induce visible artifacts (Mach bands) due to the human visual system's sensitivity to first-derivative luminance changes [8]. This motivates the use of *monotonic cubic splines* [2], which allow enforcing a zero derivative at the chroma peak while preserving global monotonicity.

3) Chroma in CIELAB Space: Chroma in CIELAB space is computed as

$$C^* = \sqrt{a^{*2} + b^{*2}}. \tag{4}$$

While C^* is not perceptually uniform, it provides a device-independent chroma magnitude suitable for relative comparison when properly normalized.

4) Reference Chroma Standard (C^*_{ref}): To stabilize chroma magnitude across palettes, chroma is first normalized by the maximum chroma within the palette and subsequently rescaled into a fixed reference frame derived from the sRGB gamut extremum in CIELAB space.

Let C^*_i denote the CIELAB chroma of step i, and define

$$C^*_{\max} = \max_i C^*_i. \tag{5}$$

The rescaled chroma used for smoothness evaluation is

$$\tilde{C}_i = \frac{C^*_i}{C^*_{\max}} \cdot C^*_{\text{ref}}. \tag{6}$$

Evaluating primary and secondary vertices of the sRGB cube shows that the maximum chroma occurs at the blue primary, yielding

$$C^*_{\text{ref}} \approx 133.8. \tag{7}$$

5) Ideal Chroma Trajectory: An ideal chroma trajectory is assumed to increase monotonically from the palette start, reach a single maximum, and then decrease monotonically toward the end. For a discrete palette of N colors indexed by $i = 0, \ldots, N-1$, the ideal trajectory $C_{\text{ideal}}(i)$ is constructed using a *monotonic cubic spline* [2] passing through three anchors:

- $(0, C_0)$ — start chroma,
- $(i_{\text{peak}}, \tilde{C}_{\max})$ — maximum chroma,
- $(N-1, C_{N-1})$ — end chroma.

This avoids polynomial oscillation (Runge's phenomenon) [9] and enforces smooth, monotonic chroma variation.

6) Smoothness Metric Definition: Deviation from the ideal trajectory is measured using root mean square error:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\left(\tilde{C}_i - C_{\text{ideal}}(i)\right)^2}. \tag{8}$$

The maximum deviation envelope is defined as

$$\epsilon_{\max,i} = \max\Big(C_{\text{ideal}}(i) - \tilde{C}_{\min}, \tilde{C}_{\max} - C_{\text{ideal}}(i)\Big), \tag{9}$$

where $\tilde{C}_{\min} = \min_i \tilde{C}_i$ and $\tilde{C}_{\max} = \max_i \tilde{C}_i$.

The maximum achievable error is

$$RMSE_{\max} = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\epsilon^2_{\max,i}}. \tag{10}$$

The *Chroma Smoothness* metric is defined as

$$\mathcal{S}_C = \max\left(0,\ 1 - \frac{RMSE}{RMSE_{\max}}\right). \tag{11}$$

7) Interpretation: $\mathcal{S}_C \approx 1$ indicates smooth chroma progression closely matching the ideal trajectory, while $\mathcal{S}_C \to 0$ reflects strong chroma oscillation or derivative discontinuities. For achromatic or near-achromatic palettes where $C^*_{\max} \approx 0$, chroma smoothness is defined as maximal, reflecting the absence of chroma variation.

D. Hue Stability ($\mathcal{H}$)

The **Hue Stability** metric evaluates the consistency of hue in a monochromatic color scale as lightness varies. A palette is considered stable if hue fluctuates minimally around a reference color rather than drifting toward other hues.

1) Hue Representation and Periodicity Handling: Hue in CIELAB space is determined by

$$h_i = \operatorname{atan2}(b^*_i, a^*_i). \tag{12}$$

Since hue is a periodic angular quantity, the hue sequence is **unwrapped** to remove discontinuities at $360°$ and ensure continuity:

$$h^{\text{unwrap}}_i = h^{\text{unwrap}}_{i-1} + \Delta h_i, \qquad \Delta h_i \in (-180°, 180°]. \tag{13}$$

This unwrapping ensures that hue deviations reflect actual shifts rather than artifacts from angular periodicity.

2) Reference Color and Hue Deviation: The reference hue h_{base} is chosen as the hue of the **color with maximum chroma** in the palette, as this typically represents the dominant hue identity.

The angular distance between each step and the reference color is defined as

$$d_i = \min\left(|h_i - h_{\text{base}}|,\ 360° - |h_i - h_{\text{base}}|\right), \tag{14}$$

where $d_i \in [0, 180°]$ represents the hue deviation at step i.

3) Normalization by Worst-Case Linear Drift: For normalization, a conservative worst-case scenario is defined in which hue drifts linearly from the reference hue toward its complementary hue. Although hue values are unwrapped to ensure numerical continuity, hue distance remains bounded by $180°$, which represents the maximum distinguishable angular deviation in circular hue space.

The maximum deviation at step i in this scenario is

$$d^{\max}_i = 180° \cdot \frac{i}{N-1}. \tag{15}$$

This value serves as the **theoretical upper bound** for hue drift at each position.

4) Hue Stability Metric Definition: The root mean square error of hue deviation is computed as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} d^2_i}. \tag{16}$$

The corresponding worst-case error is

$$RMSE_{\max} = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} (d^{\max}_i)^2}. \tag{17}$$

The **Hue Stability** metric is defined by

$$\mathcal{H} = \max\left(0,\ 1 - \frac{RMSE}{RMSE_{\max}}\right). \tag{18}$$

5) Interpretation:

- $\mathcal{H} \approx 1$: Hue remains stable around the reference color throughout the palette.
- $\mathcal{H} \to 0$: Strong hue drift approaching the worst-case linear scenario.

This metric is independent of lightness and chroma, measuring only the **geometric stability of hue** in color space.

E. Spacing Uniformity ($\mathcal{U}$)

Spacing between adjacent color steps is measured using the **CIEDE2000 color-difference metric** [5]. For a sequential color scale of N ordered steps $\{c_0, c_1, \ldots, c_{N-1}\}$, differences are computed only between consecutive entries:

$$\delta_i = \Delta E_{00}(c_{i-1}, c_i), \qquad i = 1, \ldots, N-1. \tag{19}$$

An ideally spaced color scale exhibits approximately constant perceptual increments, i.e., $\delta_i \approx \delta_j$ for all i, j. To quantify relative dispersion of these increments independently of their absolute magnitude, spacing uniformity is evaluated using the **coefficient of variation (CV)**:

$$CV = \frac{\sigma(\{\delta_i\})}{\mu(\{\delta_i\})}, \tag{20}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the mean and standard deviation of the set $\{\delta_i\}$, respectively. As a dimensionless quantity, CV provides a scale-invariant measure of non-uniformity.

Since CV is unbounded above and inversely related to quality, it is mapped to a bounded score with intuitive directionality using the following monotonic transform:

$$\mathcal{U} = \frac{1}{1 + CV}. \tag{21}$$

This definition ensures $\mathcal{U} \in (0, 1]$, with $\mathcal{U} = 1$ corresponding to perfectly uniform spacing ($CV = 0$), and $\mathcal{U}$ decreasing monotonically as dispersion increases.

A palette with uniform sampling along the color ramp therefore achieves $\mathcal{U}$ values close to unity.

IV. Composite Quality Score

To aggregate multiple quality metrics into a single scalar score without allowing strong dimensions to compensate for weak ones, the **geometric mean** is used instead of the arithmetic mean. The composite score is defined as

$$\text{SCORE} = 100 \cdot \left(\prod_{k=1}^{5} (M_k + \varepsilon) \right)^{1/5}, \tag{22}$$

where

$$M_k \in \{\eta, \mathcal{L}, \mathcal{S}_C, \mathcal{H}, \mathcal{U}\}$$

denote the five normalized component metrics, each bounded in $[0, 1]$, and $\varepsilon = 10^{-6}$ is a small constant introduced solely to ensure numerical stability when any component metric approaches zero.

The use of the geometric mean has several desirable properties:

- **Strong penalty for imbalance**: a single poor metric significantly reduces the overall score, preventing compensation by unrelated dimensions.
- **Encouragement of uniform quality** across accessibility, perceptual uniformity, and structural consistency criteria.
- **Scale invariance** with respect to the component metrics, preserving their relative contributions.

Since all component metrics are bounded in $[0, 1]$, the resulting SCORE is guaranteed to lie in the interval $[0, 100]$, enabling direct and intuitive comparison across different palette generation methods.

Accordingly, SCORE reflects overall palette completeness rather than excellence in any single isolated attribute.

V. Experimental Evaluation

This section reports quantitative evaluation results for **11 widely used design systems**, serving as a benchmark for the proposed metric framework when applied to real-world monochromatic color ramps.

An interactive online report containing all benchmark data and visualizations is available at:

> https://chromametry.github.io/chromametry/benchmarks/monochromatic

A. Quantitative Analysis Results

Table III summarizes the measured quality metrics across all evaluated design systems. Clear differentiation is observed across both individual metric dimensions and composite scores, indicating substantial variation in how existing design systems balance accessibility, uniformity, and structural consistency.

Several trends emerge from these results. Design systems that tightly control contrast span and lightness progression (e.g., IBM Carbon and Adobe Spectrum) achieve consistently high scores across most dimensions. In contrast, systems with large span values or uneven spacing exhibit reduced contrast efficiency and spacing uniformity, which significantly impacts the composite score due to the use of geometric aggregation.

> **Note.** Design systems such as Bootstrap, Google Material 3, Apple Human Interface Guidelines, and Fluent UI are excluded from this evaluation, as they primarily define discrete semantic color tokens rather than algorithmically constructed sequential color ramps.

B. Example: A Typical Report

The contrast span value K is taken directly from measured *Span*. Observed density D is determined by

$$D = \frac{K}{N-1}, \tag{23}$$

where N is the number of palette steps.

This table shows that design systems achieving high efficiency all have low density, while palettes with excessively wide span lead to wasted lightness space and are strongly penalized by metric η.

TABLE III
BENCHMARK RESULTS OF POPULAR DESIGN SYSTEMS EVALUATED USING THE PROPOSED METRIC FRAMEWORK.

Design System	Ramps	Steps	K	η	$\mathcal{L}$	$\mathcal{S}_C$	$\mathcal{H}$	$\mathcal{U}$	SCORE
Adobe Spectrum	10	18	9	0.947	0.9333	0.8786	0.9138	0.7722	88.67
IBM Carbon	12	12	6	0.923	0.9303	0.8688	0.9252	0.7919	88.62
U.S. Web Design System	25	12	6	0.923	0.9359	0.8096	0.9380	0.7997	87.90
Salesforce Lightning 2	13	14	7	0.933	0.9187	0.8464	0.9372	0.7107	86.47
GitHub Primer Brand	13	12	6	0.923	0.9243	0.8405	0.9408	0.6841	85.67
Atlassian	9	14	8	0.800	0.8964	0.9094	0.9465	0.7129	84.86
Tailwind CSS	18	13	8	0.789	0.8705	0.8565	0.9147	0.6780	81.74
Ant Design	12	12	9	0.711	0.8586	0.8734	0.9276	0.6550	79.81
Material UI	19	12	11	0.565	0.7967	0.7861	0.9239	0.5500	70.95
Radix UI	16	13	10	0.543	0.7979	0.7679	0.9481	0.5207	69.67
Shopify Polaris	12	17	15	0.356	0.7281	0.6892	0.9223	0.4667	59.86

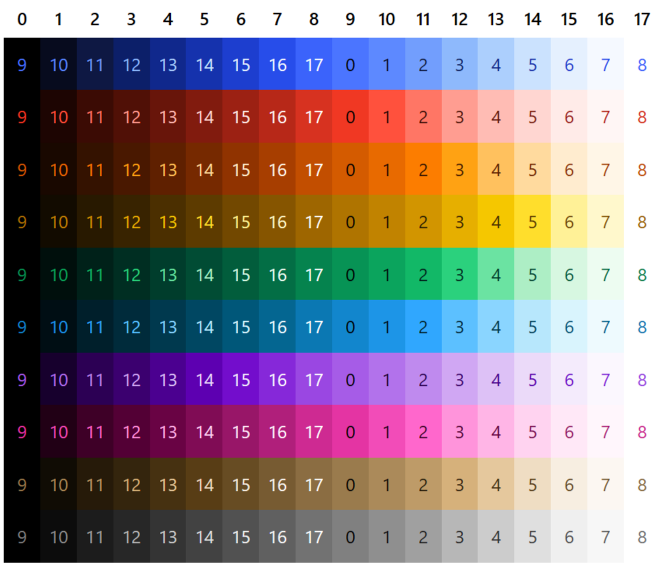
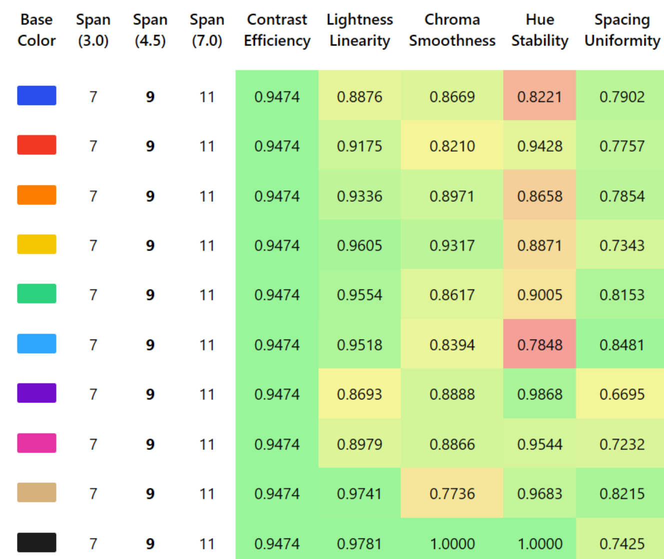


Fig. 1. Adobe Spectrum Color Palette. Cell value = Background Index ±9 steps (WCAG 4.5:1).



Base Color	Span (3.0)	Span (4.5)	Span (7.0)	Contrast Efficiency	Lightness Linearity	Chroma Smoothness	Hue Stability	Spacing Uniformity
	7	9	11	0.9474	0.8876	0.8669	0.8221	0.7902
	7	9	11	0.9474	0.9175	0.8210	0.9428	0.7757
	7	9	11	0.9474	0.9336	0.8971	0.8658	0.7854
	7	9	11	0.9474	0.9605	0.9317	0.8871	0.7343
	7	9	11	0.9474	0.9554	0.8617	0.9005	0.8153
	7	9	11	0.9474	0.9518	0.8394	0.7848	0.8481
	7	9	11	0.9474	0.8693	0.8888	0.9868	0.6695
	7	9	11	0.9474	0.8979	0.8866	0.9544	0.7232
	7	9	11	0.9474	0.9741	0.7736	0.9683	0.8215
	7	9	11	0.9474	0.9781	1.0000	1.0000	0.7425

Fig. 2. Adobe Spectrum Palette Metrics.

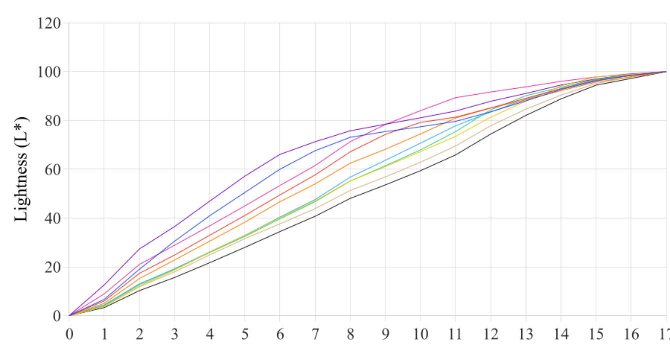


Fig. 3. Adobe Spectrum Lightness Chart.

TABLE IV
EXPERIMENTAL VALUES OF K, D, AND CONTRAST EFFICIENCY η.

Color Palette	N	K	D	η
Adobe Spectrum	18	9	0.500	0.947
IBM Carbon	12	6	0.500	0.923
U.S. Web Design System	12	6	0.500	0.923
Salesforce Lightning 2	14	7	0.500	0.933
GitHub Primer Brand	12	6	0.500	0.923
Atlassian	14	8	0.571	0.800
Tailwind CSS	13	8	0.615	0.789
Ant Design	12	9	0.750	0.711
Material UI	12	11	0.917	0.565
Radix UI	13	10	0.769	0.543
Shopify Polaris	17	15	0.882	0.356

Thus, contrast span K is observed consistently across popular industry palettes, though previously not systematically identified or exploited. With these results, users can select color pairs for background and text with distance K without requiring runtime verification.

C. Discussion

1) High Score Group (Score > 85): IBM Carbon, Adobe Spectrum, and USWDS maintain consistently high indicators across all aspects.

- **Observation:** High Contrast Efficiency ($\eta > 0.9$) and Lightness Linearity ($\mathcal{L} > 0.93$).

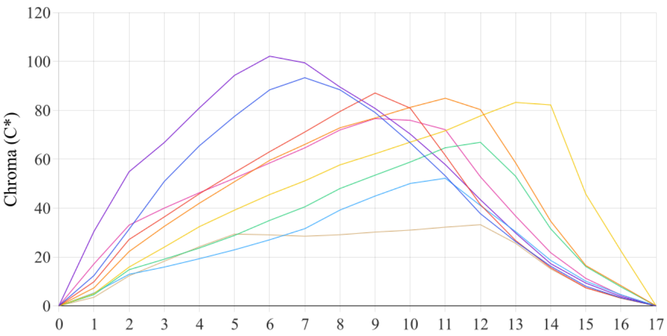


Fig. 4. Adobe Spectrum Chroma Chart.

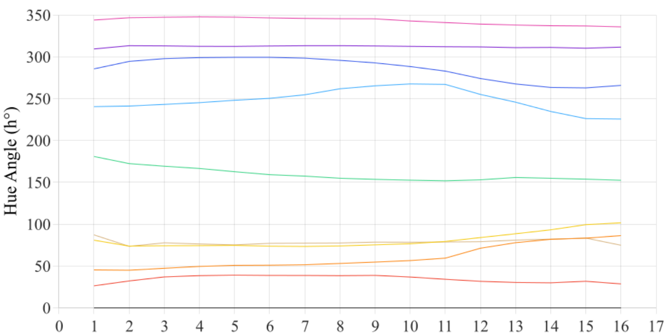


Fig. 5. Adobe Spectrum Hue Chart.

- **Analysis:** Span distance K maintains optimal ratio to total steps N, ensuring high density of WCAG-compliant color pairs.

2) Medium Score Group (75–85): Tailwind CSS and Ant Design show disparities among component indicators.

- **Observation:** High Chroma Smoothness ($\mathcal{S}_C > 0.84$) but lower Contrast Efficiency.
- **Analysis:** Large span relative to N reduces simultaneous usability of color pairs.

3) Low Score Group (< 75): Material UI (v4) and Shopify Polaris record the lowest scores.

- **Observation:** Low Spacing Uniformity ($\mathcal{U} < 0.6$) and Lightness Linearity.
- **Analysis:** Non-uniform step spacing and non-linear lightness progression.

D. Experimental Conclusions

Benchmark data confirms the framework's ability to classify palettes based on physical and mathematical characteristics. Results suggest that using an even number of steps with contrast span $\approx (N-1)/2$ yields favorable usability properties, where every color can serve as background and a corresponding text color always exists satisfying WCAG 4.5:1.

VI. Conclusion

This work introduces a unified quantitative framework for evaluating the quality of sequential monochromatic color palettes, grounded in the CIELAB color space and motivated by both perceptual uniformity and accessibility requirements. Five complementary metrics are proposed, covering contrast efficiency under WCAG 4.5:1 constraints, lightness linearity, chroma smoothness, hue stability, and perceptual spacing uniformity. Each metric captures a distinct structural or perceptual property and is normalized to enable consistent aggregation.

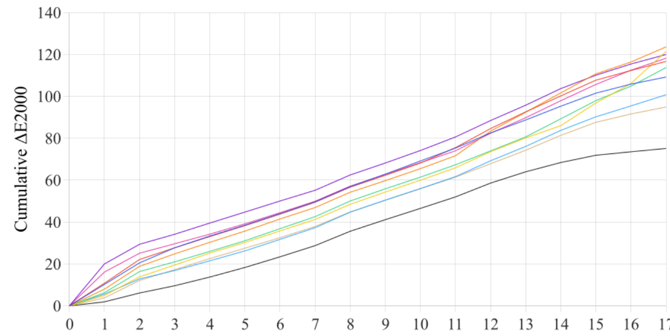


Fig. 6. Adobe Spectrum Cumilatiive DeltaE2000 Chart.

Through large-scale benchmarking of widely used industry design systems, the framework reveals systematic differences in how palettes allocate lightness range, distribute perceptual steps, and preserve hue identity. In particular, the analysis identifies contrast span K as a previously under-formalized but critical parameter, with empirical results indicating that spans close to $(N-1)/2$ maximize accessibility efficiency while preserving usable color density. These findings demonstrate that many high-quality palettes converge toward similar structural ratios, despite differing design origins.

The proposed composite score, based on geometric aggregation, enables objective comparison without allowing compensation between weak and strong dimensions, thereby reflecting overall palette completeness rather than isolated excellence. As a result, the framework supports reproducible benchmarking, automated palette validation, and optimization-driven color system design.

This study focuses on monochromatic sequential color ramps and adopts the CIELAB color space as a unified evaluation domain, ensuring direct compatibility with the CIEDE2000 color-difference metric used for perceptual uniformity assessment. Future work will extend the framework to diverging and multivariate palettes, and investigate how analogous metric formulations may be adapted to alternative perceptually uniform color spaces (e.g., CAM02-UCS or CAM16) while preserving cross-space comparability.

References

[1] ISO/CIE. *Colorimetry — Part 4: CIE 1976 L*a*b* Colour Space*. ISO/CIE 11664-4:2019, 2019.

[2] F. N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246, 1980.

[3] R. High, P. Green, and P. Nussbaum. The Helmholtz–Kohlrausch effect on display-based light colors and simulated substrate colors. *Color Research & Application*, 48(3):341–354, 2023.

[4] A. Zeileis, K. Hornik, and P. Murrell. Escaping RGBland: Selecting colors for statistical graphics. *Computational Statistics & Data Analysis*, 53(9):3259–3270, 2009.

[5] M. R. Luo, G. Cui, and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5):340–350, 2001.

[6] W3C. *Web Content Accessibility Guidelines (WCAG) 2.1*. 2018. Available at https://www.w3.org/TR/WCAG21/.

[7] M. Safdar, G. Cui, Y. J. Kim, and M. R. Luo. Perceptually uniform color space for image signals including high dynamic range and wide gamut. *Optics Express*, 25(13):15131–15151, 2017.

[8] E. Mach. Über die Wirkung der räumlichen Verteilung des Lichtreizes auf die Netzhaut. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Wien*, 52:303–322, 1865.

[9] J. F. Epperson. On the Runge example. *The American Mathematical Monthly*, 94(4):329–341, 1987.

[10] IEC. *Multimedia systems and equipment — Colour measurement and management — Part 2-1: Colour management — Default RGB colour space — sRGB*. IEC 61966-2-1:1999, 1999.