

## modern color models

## color vision

The first six pages have examined the fundamental aspects of color perception — the **trichromatic mechanism**, the three **colormaking attributes** that describe color sensations, **the geometry of color** (including the hue circle, opponent functions and response compression), the many **basic forms of color**, and the complex effects of **context, contrast & adaptation**, and finally the principles of **additive & subtractive color mixing**. This page describes the development of modern color models and discusses their relationship to these color fundamentals.

A **color model** represents the logical or perceptual relationships among colors of lights or surfaces. From the modern perspective, a color model must meet the following **four requirements**:

- (1) a **color specification** that analyzes every light or surface color into a mixture of fundamental attributes (such as **"primary" colors**, trichromatic cone responses or **tristimulus values**, or **unique hues**);
- (2) a **geometrical framework** that locates all the possible colors in relation to each other and to the fundamental attributes;
- (3) a **unique color identifier** or color notation (now usually the numerical value of the three **colormaking attributes** — brightness/lightness, hue and hue purity) for every possible color; and
- (4) a definition of **physical exemplars**, specific mixtures of lights or paints, that recreate the measured color perception when viewed within a standard surround under standard lighting conditions.

There are two kinds of modern color models. A **color order system** is based on a geometrical or enumerative framework that provides the color notation. The fundamental attributes used in the color specification are "pure" pigments or ideal colors; exemplars are manufactured as pigment recipes in specific media. These systems are common in manufacturing applications based on visual color comparisons under fixed (standard) viewing conditions.

A **color appearance model** is based on a geometrical framework whose dimensions can be adjusted to represent the color changes caused by different illuminants, illuminance levels and surround contrasts. The color specification is based on **color matching functions** measured by spectrophotometer; these measurements produce the color notation and identify matching physical exemplars.

Contemporary artists' **color wheels** are color order systems that show only the hue/chroma relationships as defined in a more comprehensive color model.

There are many color models currently in use by scientists and designers, but I will describe six, including **CIELAB** — one of the simplest and most practical from an artist's point of view — and **CIECAM02**, probably the most complex model in current use.

the evolution of  
color models

Swedish Natural  
Color System

Munsell color system

OSA uniform color scales

CIELUV  
uniform color space

CIELAB  
uniform color space

CIECAM  
color appearance model



## the evolution of color models

We are so accustomed to our technologically enhanced, fashion nuanced and market driven color environment — our modern "color culture" — that it is surprising to realize how differently people experienced and thought about color just two centuries ago. There were no modern pigments, dyes or neon signs to dazzle human eyes. Color was a mysterious quality for laypersons and naturalists alike, and it was a difficult task simply to find reliable ways to attach a name to a specific color.

**Color Naming.** The earliest color labels are **embedded in common speech**. Every language has words to name colors, though in many primitive cultures only a handful of words are available, and the words make very gross distinctions between light and dark, or warm and cool colors. Thus the *Black Sea* is not black, but dark blue; its color name has been handed down from a time when the color ideas *green*, *blue*, *dark* and *black* were denoted by the same word. This doesn't mean that primitive speakers could not see the differences between those colors, just that the differences were not conceptualized as abstract color categories.

As human knowledge and technology developed in areas such as painting, dyeing, weaving, ceramics, cosmetics, medical diagnosis, mineralogy, botany, horticulture, entomology and chemistry, more precise color comparisons across different kinds of materials became useful. This led to the development of **color naming systems**. The earliest and most obvious approach was to anchor color names on a distinctively colored flower, fruit, mineral or organic compound (dye) — and so we have the colors *lemon*, *primrose*, *saffron*, *amber*, *gold*, *orange*, *vermillion*, *rose*, *ruby*, *burgundy*, *carmine*, *violet*, *sapphire*, *turquoise*, *emerald*, *leaf green*, *ochre*, *sepia*, *indigo*, *ivory*, *ebony* and so on. The world itself was a reference book of color.

During the 16th and 17th centuries, naturalists and their rich patrons collected samples of sea shells, corals, flowers, gems, minerals, ceramics and insects as a **cabinet of curiosities**, and these artifacts were sometimes used to identify and preserve unusual color exemplars from the natural world. As a result, dyers, printers and naturalists (especially botanists and entomologists, collectors of plants and insects) were among the most active color researchers during the 18th and early 19th centuries.

However, early color enthusiasts had no assurance that they had sampled the full range of possible colors — they did not understand the **physiological limits** of color vision, and therefore could not be sure that a new insect, coral, flower or gem might not reveal a hue that had never been seen before.

This uncertainty was swept away in 1704 when **Isaac Newton** published his *Opticks*, the first scientific analysis of color. Among the many stunning breakthroughs in this book was Newton's claim to have identified *all fundamental colors* — and by extension all possible mixed colors. Any color could be reduced either to a pure spectral hue or to a geometrically defined *mixture of spectral hues*. Newton symbolized this intellectual closure as a closed circle: his **hue circle**. So ended doubt about the *range* of possible colors.

However, the practical problem remained that precise spectral mixtures were difficult to achieve until the mid 19th century, object colors were typically duller and more complex than the

colors of mixed spectral lights, and all those complex colors still had to be identified with a name and captured as a physical exemplar. In fact, the last two requirements were the most important and the easiest to fulfill. So naturalists had to derive their color identifications and color labels from a **color atlas**, a tradition solution that extends from A.G. Werner's *Nomenclature of Colors* (1774) up to Robert Ridgeway's *Color standards and color nomenclature* (1912). These provided a series of hand painted colored patches as a visual standard for each color, with the color's name and, in earlier versions, a list of flowers, insects or minerals that exemplified the color in nature. Similar tools are still used today in horticulture, agriculture and food processing.

**Color Mixture.** If *naming* the color was incidental to using it, then colors could be reduced to a recipe or proportional mixture of a few standard colorants or "primary" colors. The entrepreneurial printer **Jakob Christoffel Le Blon** first applied this approach to the industrial task of color printing, by showing that color images could be created from three primary or "primitive" colored inks, usually supplemented by a fourth black ink, one ink printed over another with separate, carefully aligned mezzotint plates (a technique developed earlier in woodblock printing of book illustrations). Multicolor printing systems similar to LeBlon's, though not commercially successful, continued up to the end of the 18th century; the multicolor printing method was revived when *chromolithography* was invented in the 1860's, which did not become widely used until the 1890's. But from that time until today color printing systems based on standardized ink colors and mixture recipes have been universally used by commercial printers for mass reproduction of color images. Modern color mixing systems include the four ink **process (CMYK) color system** (developed in 1934), the 14 ink **Pantone™ color system**, and the six ink **Hexachrome™ color system**, to name a few.

The Pantone and other printing systems are based on standardized printed color samples, each color reproducible as a carefully proportioned mixture of a limited number of "primary" inks. These are really "color cook books" tied to specific technologies of color reproduction, and they require the graphic artist to visually select from a limited (though sometimes very large) index of printed color samples. These books provide no way to visualize how all the colors fit together, in the same way that a city phone directory gives you no idea of how the city looks on a map.

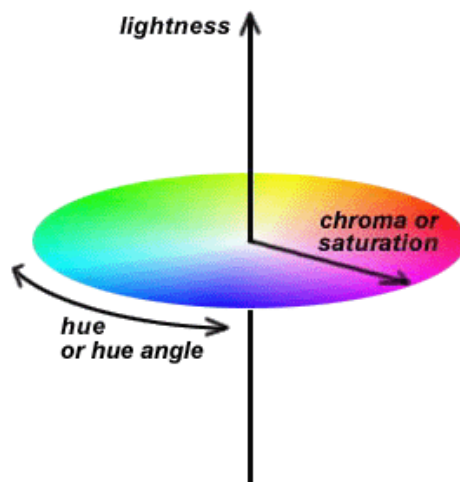
**Color Models.** This need for a *geometrical framework* or map of color brings us to modern **color order systems**, which put every color in its place and specify colors as combinations of fundamental color attributes. Arguably the first color order system was proposed by the German astronomer **Tobias Mayer** as a university lecture in 1758. Mayer offered his system, which was not published until 1775, as a concept, but he described how to analyze all colors as "primary" color mixtures, showed lightness as a completely separate dimension of color from hue, and demonstrated how the system could be used to identify or compare object colors, such as the colors of pigments. Mayer relied on visual color matching to meet the first requirement of a modern color model (stimulus measurement), but he showed how a consistent geometrical framework could provide unique numerical labels for colors.

The first color model to adopt the three dimensional framework that has become standard in color science was the *Farbenkugel* or "Color Sphere" proposed by the German romantic painter **Philipp Otto Runge** (1777-1810) in 1810 (right). The sphere contains the white to black **value scale** as its polar axis, and Newton's complete **hue circle** of pure or maximally intense hues as the equator. Runge's hue circle is also one of the first to divide the hue circle into "primary", secondary and **tertiary hues**. As Runge proudly claimed:

*It will be impossible to think of any nuance produced by a mixture of the five elements (blue, yellow, red, white and black) and not contained in this framework; nor can the whole system be represented by any other correct and complete figure. And since each nuance is given its correct relation to all the pure elements as well as to all mixtures, this sphere must be considered a universal chart, enabling anybody to orient himself as to the overall context of all colors.*

**Michel-Eugène Chevreul** followed with his own "color hemisphere" in 1839, and abstract color systems became one of the minor achievements of 19th century European scientists and the artists who studied their works.

**Color Model Assumptions.** The geometrical framework used in modern color models consists of the three **colormaking attributes** arranged to define a three dimensional space. The figure below shows how this is done. The vertical black to white dimension is the *lightness* or *value* of a color; the circumference of the horizontal disk perpendicular to the lightness dimension is the *hue* of a color; and the lateral distance or radius on this disk, from the center outward, is the *chroma* (approximately the *saturation*) of a color.

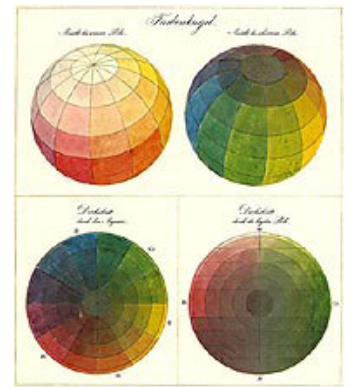


**the geometry of colormaking attributes in modern color models**

Notice that this arrangement allows hues to change gradually from one to the next, but defines chroma and lightness as rulers or *ratio scales* starting from a definite zero point — black for lightness, and gray for saturation.

Color models differ in the specific assumptions that are used to define the color geometry and color labels and the specific measurement methods used to link the geometry to color stimuli and color exemplars. My selection of color models is intended to highlight these differences:

- **Geometrical Uniformity** - The perceptual facts of color are tremendously complex, and many early models emphasized a



**one of the first modern color models: Philipp Otto Runge's "color sphere" (1810)**

top row: exterior views of the white and black "poles"; lower row: horizontal and vertical slices through the sphere

geometrical or logical framework that summarized color in a simplified way. These systems use visual color matching to measure colors, appeal to the *phenomenology* or subjective relationships among colors, and emphasize a framework for verbal color communication. The best modern example is the **Swedish NCS** color model.

- **Perceptual Uniformity** - Here the goal is to create a color geometry where the units of measurement stand for equal apparent color differences on all the colormaking dimensions. Perceived color differences determine the color geometry and exemplification as physical color samples. The classic modern example is the **Munsell Color Order System**; very different recent examples include the **OSA uniform color scales** and the **CIE 1976 UCS** (uniform chromaticity scales) diagram.
- **Trichromatic Colorimetry** - Increasingly, modern color models are linked to the electronic measurement of spectral emission or reflectance profiles, expressed as a mixture of three **imaginary "primary" lights** called the **XYZ tristimulus values**. This standardizes the "viewing eye" used to compare colors and allows automated color matching. The modern examples of this approach are **CIELUV** and **CIELAB**.
- **Color Appearance** - The earliest color models only described colors that were viewed or compared in a **standardized viewing situation**. In the past few decades, more ambitious color models have been developed that adjust the color space to represent the effects of the visual context (the illuminant, surround and background colors) and extreme viewing conditions (small or large color samples, dim or bright illumination). This raises them to the status of **color appearance models** — color models that predict a wide range of real world color experiences. The current example of this approach is **CIECAM02**.

An authoritative summary of the development and empirical foundations of modern color models is Rolf G. Kuehni's **Color Space and Its Divisions: Color Order from Antiquity to the Present** (Wiley Interscience, 2003). For an interactive tutorial on historical color models, see J. Frans Gerritsen's **Evolution in Color** (Schiffer, 1988).

The best technical overview of modern color order systems, uniform color spaces and the newest color appearance models and prototype *image color appearance models* (ICAMs) designed for color reproduction technologies is Mark Fairchild's **Color Appearance Models (2nd ed.)** (Addison Wesley, 2005); see also the updates and information at Fairchild's **web site**.

There's a delightful menagerie of historical color models, based in part on Gerritsen's book, collected at the **Colorcube web site**. The most comprehensive historical review of color theories, with ample illustrations of each color model, is provided at the Colorsystem **Virtual Color Museum**. For an archive of historical color model illustrations, see the page by **Dr. Hans Irtel**.

## Swedish Natural Color System ▲

Two different European approaches to color emerged by the middle of the 19th century. One relied on a *quantitative description* of color via the methods of **psychophysics**, a tradition founded by **James Clerk Maxwell** and **Hermann Helmholtz**. This approach originated in the work of **Isaac**

**Newton** and laid the foundations for the color system of **Albert Munsell** and the many **CIE** colorimetric color models.

The alternative approach, which became dominant in European philosophy and visual arts, emphasized the *logical description* of color experience. Color was idealized as "pure" sensations, organized as verbal or geometrical models of color relationships. The Swedish **Natural Color System (NCS)**, proposed by Tryggve Johansson in the 1930's and eventually realized by Sven Hesselgren and Anders Hård at the Scandinavian Colour Institute in the 1960's, is a modern color model that exemplifies this approach. The NCS is currently a color standard in several Scandinavian and EU countries.

The experimental pedigree of the NCS goes back to Ewald Hering's *Das natürliche Farbsystem* (1905), which describes colors as the result of **opponent color processing** and the mixture of **four unique hues** plus black and white. Hering's was a *phenomenological* or subjective approach to color: he tried to describe the universal patterns or laws of our color experience without specifying what kind of external stimulus or internal color receptors might produce it. This phenomenological approach can be traced from Hering back to the Romantic German philosophers Arthur Schopenhauer, Georg W.F. Hegel and ultimately to **J.W. von Goethe** — although those authors and their philosophical concerns were of little practical use to the artists of their time.

**Four Components of NCS.** This phenomenological or "in the head" approach to color modeling is based on four components.

The first is a single, universal color specification formula:

$$\text{color} = \mathbf{C}(\mathbf{u}_1 + \mathbf{u}_2) + \mathbf{S} + \mathbf{W}$$

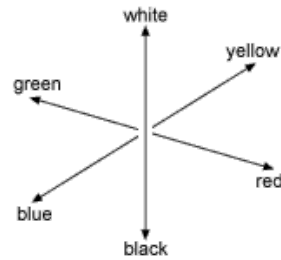
where **C** represents the apparent proportion of pure hue (chromaticness) in the color, **S** (German *Schwartz*) is the apparent proportion of black in the color, and **W** (German *Weiss*) is the apparent proportion of white. The chromatic ingredient consists either of a single *unique hue*, or more commonly of the mixture of two adjacent unique hues, shown as  $(\mathbf{u}_1 + \mathbf{u}_2)$ . Hering called any color mixed with either white or black (or both) a **veiled color**, and by these mixtures accounted for all dull, diluted and near neutral colors.

The basic formula has a long pedigree. The Western painting tradition used the technique of **broken color** to adjust pure pigment (high chroma) colors by mixture with white and/or black paints, a method described by Cennini, Alberti and Leonardo. The same mixture geometry is used in Michel-Eugène Chevreul's **color hemisphere**. Hering's color specification was also used by the German chemist and Nobel laureate **Wilhelm Ostwald** (1853-1932) in his *Mathetische Farbenlehre* (Quantitative Chromatics) published in 1917 and exemplified by the *Farbkörper* (Color Atlas) published in 1919. Ostwald's system was very influential among the artists and designers of *De Stijl* and the Bauhaus.

The second component of the phenomenological approach to color is a restricted version of the universal **color geometry**. The four Hering unique hues are arranged to create four **cardinal points** around the hue circle, as shown in the figure. Following Hering's original concept, the y/b and r/g **opponent functions** define two dimensions at right angles within the hue circle. Antagonistic unique hues, which are not allowed to mix in



Hering's theory, are on opposite ends of the two dimensions. The third opponent dimension defined by white and black is placed perpendicular to the first two.



### three color dimensions defined by the hering colors

The third component consists of two **color mixing rules**:

(1) Only unique hues on different opponent dimensions can be mixed; unique hues at opposite ends of the same dimension cannot be mixed. (That is, we are not allowed to make "yellowish blue" or "reddish green" mixtures.)

(2) White and black can mix with each other, and with any hue mixture, in any proportions; this provides the full range of **shades, tones and tints** for any hue.

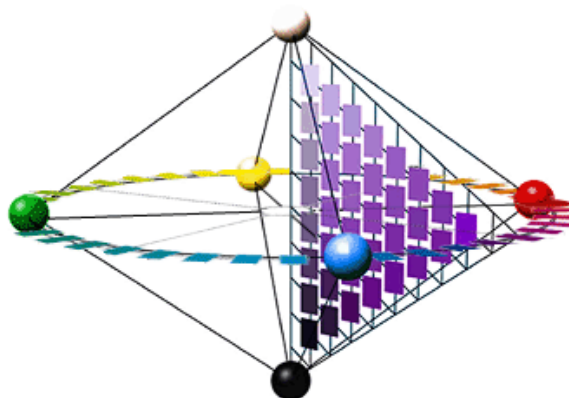
The fourth and last component is a **fixed sum** for the color specification:

$$C(u_1 + u_2) + S + W = 100$$

in the same way that mixtures in Tobias Mayer's **color hexahedron** must sum to 1. This forces the color geometry into a triangular (conical, hexahedral) form.

**The NCS Color Model.** These four design components interconnect the six Hering colors as 13 **elementary scales** — a single gray scale (white to black), and three chromatic scales for each hue (e.g., red to white, red to black, and red to gray) — that measure off mixtures among the six cardinal colors.

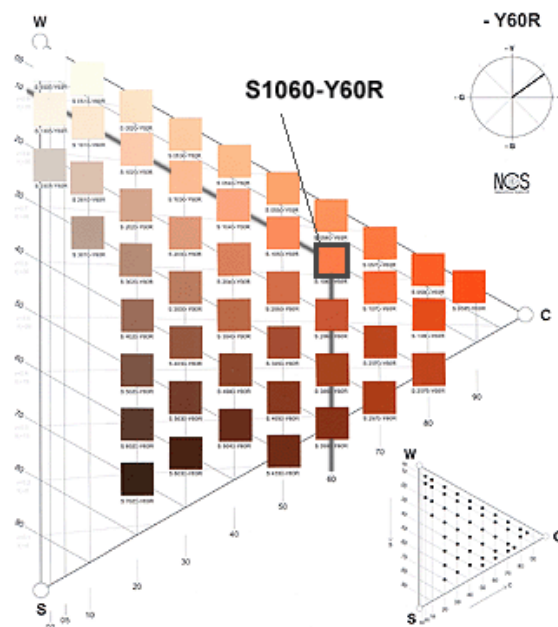
The basic mixing formula **color = C + S + W = 100** dictates that the model is most conveniently represented as vertical sections, called **hue triangles**, that include one pure hue or hue mixture, pure white and pure black as the three corners of an equilateral triangle. This is equivalent the geometry of two cones joined at their base that was used in Ostwald's system, as shown below for the 13 elementary scales and a purple (mixed red and blue) hue section.



**geometrical framework of the swedish NCS color model**

The NCS is exemplified as the **NCS color atlas**, a large book of standard color swatches. The atlas contains 40 of these hue triangles at equal intervals around the hue circle — nine steps (increments of 10% mixture) between each pair of adjacent unique hues. The atlas also includes a representation of the complete hue circle at maximum chroma, a lightness scale from white to black, and several "off white" color samples as a help to specifying pastel colors.

Each hue triangle displays a wide selection color samples made from mixtures of the pure hue with black (**S**) and/or white (**W**). Lines of equal black mixture are shown as diagonal rows parallel to the top edge of the triangle; blackness increases from the top edge to the bottom corner. Lines of equal whiteness are shown as diagonal rows parallel to the bottom edge; whiteness increases from the bottom edge to the top corner. As a result, gradations in chroma (**C**) are represented by vertical columns through the triangle, from the left corner (the maximum chroma), which contains no white or black, to the gray scale along the vertical edge. Color samples for the pure hues, white and black are omitted due to the limitations of available pigments, and very dark or near neutral samples are usually omitted to reduce production costs and visual clutter.



a page from the **NCS color atlas 96** (standard version)

The NCS notation or color labeling defines the visual appearance of the color and shows how to find it in the **NCS color atlas 96** (1996), represented by the sample page above. The tonal mixture is stated before the hue. The first two digits give the percentage of black in the color, followed by the two digit percentage of pure color. (The percentage of white is omitted, as it is equal to the other two percentages subtracted from 100.) Next, hue mixtures are notated with letters for the two unique hues contributing to the mixture, bracketing the percentage of the second hue in the mixture: **Y60R** represents a red orange mixture of 60% red and (by subtraction) 40% yellow. Pure hues are notated simply **R**, **Y**, **G** or **B**. Thus, the full notation **S 1060-Y60R** represents (1) a pure hue that is mixed from 60% red and 40% yellow, and (2) a color that is mixed from 10% black and 60% pure hue (and 30% white). Achromatic (gray) tones are indicated simply with the proportion of black, followed by **N** for neutral: **S 1500-N** represents a light gray containing 15% black (and 85% white).



A significant feature of the NCS is that all the pure color standards are *imaginary*: the point of comparison for a pure white or pure yellow is visualized in the mind, and color mixtures are judged against these ideal standards. (This is why there are no color swatches at the points **C**, **S** or **W** in the sample hue triangle above: no pigment can match these pure, ideal colors.) This was consistent with Hering's original phenomenological color definitions, but it was also validated by experiments that showed people were just as accurate judging color mixtures from ideal (imaginary) color standards as from physical color samples.

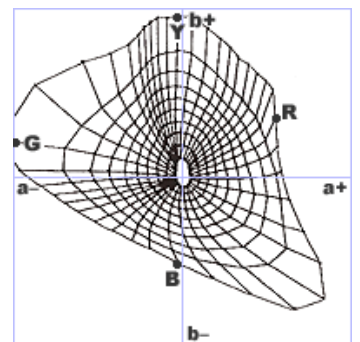
**Using and Interpreting the NCS.** The NCS was created primarily for communicating about color in everyday and design applications, because its simple color geometry and direct relationship to visual color judgments **makes color easy to describe**. The only "measurement tool" necessary is a person with normal color vision, training in the NCS notation conventions and access to an NCS color atlas to use as a reference. (As with Munsell, viewing conditions must be carefully standardized if color judgments are to be comparable.)

A designer or architect who wants to communicate a specific color standard simply estimates the mixture of Hering colors that creates it, finds the matching color in the NCS color atlas, and determines the color's NCS color notation. A textile mill or construction company then uses the NCS notation to identify the color sample they should match in fabric or wall paint. The major drawback to this method is that most people find it very difficult to identify dull or dark colors (browns, tans or maroons in particular) in terms of saturated unique hue mixtures.

The geometrical simplicity of the NCS introduces some significant distortions or complications into the *perceptual* color space. Because the four Hering colors are **not equally spaced** around a perceptual color circle (as shown for example on the **CIECAM  $a_c b_c$**  plane), the hue differences around the NCS color circle are not perceptually equal, either. Differences between the yellow green hue triangles are perceptually very small, while hue differences between the blue green and purple hue triangles are quite large. This means fine color discriminations are possible for yellow green colors, while the blue green and purple colors require some guesswork. In addition, the chroma spacing is unequal from one hue to the next, and from neutral to intense colors within any hue (right).

In addition, for most parts of the color circle, opposite colors are not "mix to gray" **visual complements**, because the unique hues themselves are not visual complements. The visual complement of unique red is not unique green, but a blue green located at roughly **G40B**; the visual complement of unique blue is not unique yellow, but a deep yellow at roughly **Y30R**. Finally, the artist's **mixing complements** are also different from the NCS complements. And while the NCS establishes new complementary relationships, there is no evidence that these are more (or less) pleasing or effective than the complementaries defined in other ways.

Most perplexing, the "pure" hues at maximum chroma **vary significantly in lightness** around the color circle: an intense yellow is much lighter than an intense blue, but in NCS they are all notated as containing no black and no white, and are placed at the same halfway point on the vertical white/black dimension. This means that the white/black mixture does not consistently describe the actual lightness of colors, and all the mixture



lines of constant NCS hue and chroma plotted on the CIELAB  $a^*b^*$  plane

after Kuehni (2003)

intervals (columns and diagonal rows in the hue triangle) do not define comparable steps across different hues. For a light valued yellow hue, the 10 gradations in **W** (between yellow and white) are perceptually quite small, while the 10 gradations in **S** (between yellow and black) are quite large. For a dark valued blue violet, the reverse is true. This is entirely consistent with the Hering definition of black as a *color* rather than as *the absence of light*, but it makes the NCS noncomparable with other color systems based on lightness.

The benefit of this apparently arbitrary relationship is that *different hues* at the same location within their respective hue triangles — that is, hues identified with the same whiteness and blackness numbers (**S1060** for the orange in the illustration above) — have the same **nuance**, and nuance is an effective way to equate colors, as described in the page on **color harmonies**.

Most color order systems that build on a simple subjective or geometrical framework have similar drawbacks. These are not serious as long as the models are used only as a way to label and communicate about individual color samples, but the NCS **must not be used to model perceived color differences**. And these drawbacks may even be an advantage in color design, for example when equating the apparent whiteness of various hues in pastel color schemes.

In general, however, the phenomenological and geometrical approach did not generate much interest among American and British research psychologists, who turned instead to psychophysical methods of color scaling.

For more information about the NCS and related color products, see the NCS **web site**.

## Munsell Color System ▲

The **Munsell Color System** is arguably the first modern color order system — based on the three colormaking attributes, and implemented through careful color measurement. Conceived in the 1890's by the American artist and educator Albert H. Munsell (1858-1918), it was described as a theoretical color model in 1905 and exemplified as a 15 page atlas of color samples in 1915. (This was extensively revised and republished in 1929 as the *Munsell Book of Color*, comprising 20 hue pages that each contain about 20 color samples.) The Munsell system was extensively revised or "renotated" in the early 1940's, when it was adopted as the standard color reference system in the USA. It remains one of the most popular and widely used color order systems.

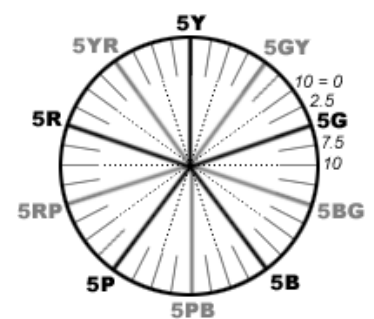
Munsell originally conceived of his system as a method to describe color for artists and teach color to children. He was aware of the European, geometrically regular color models of Chevreul, Wundt and others, and at first conceived (and patented) his system as a color sphere. But with a typically American empiricism, he convinced himself through analysis of paint samples that color space was not geometrically regular. Instead, Munsell sought to build his model on **equal perceived color differences** on each colormaking attribute, linked in some cases to measurable changes in color stimulus composition, in a freely branching geometry called a *color tree*.

Munsell was inspired by and heavily relied on the expertise of the American physicist and color researcher **Ogden Rood**, who

showed Munsell how to use a **color top** to analyze complex colors into additive color mixtures of "primary" colors, and to develop perceptually equal scales of gray or color chroma based on proportional mixtures of black and white. With this method Munsell was able to simulate **incremental steps in chroma** (proportional mixtures with gray) or hue (by proportional mixtures of "primary" colors, or analogous colors such as red and yellow), and to identify **visual complementary colors** (saturated hues that additively mix to a pure gray). Over a decade of patient, incremental work, Munsell analyzed reflective or surface colors into their component lightness, hue and chroma, published in his atlas as equal apparent differences on each of these colormaking attributes.

**Geometry of the Munsell System.** The backbone of the Munsell system, the "trunk" of the color tree, is a vertical dimension of lightness or **value**, which was developed through human perceptual judgments of equal differences in lightness across visual (color top) mixtures of white and black paints. The Munsell **value scale** ranges from pure black (0) to pure white (10); each step is divided into decimal increments, resulting in a 100 step lightness scale.

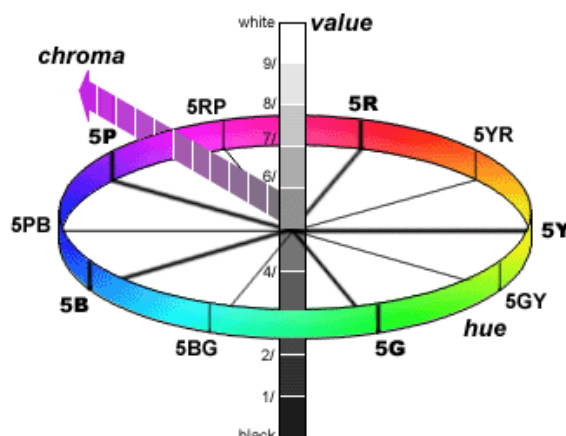
At the middle value level is a **hue circle** defined by **five equally spaced hue dimensions** — red (R), yellow (Y), green (G), blue (B) and purple (P) in clockwise order (yellow at the top, red on the left). These **principal colors** are separated by five mixture hues between them — yellow red (YR), green yellow (GY), blue green (BG), purple blue (PB), and red purple (RP) — each the **visual complementary color** of the principal hue directly opposite it. These principal and mixture hues divide the hue circle into ten equal hue segments, and each hue segment is again divided by ten (counted in the clockwise direction), resulting in a 100 step hue circle. The "standard" or central example of each color segment is located at step 5 (diagram, right).



the 40 standard munsell hue circle divisions

Finally, each hue was calibrated in equal **chroma** steps from zero (pure gray) to the maximum color intensity in the paints or inks used for each hue; in the idealized (aim) colors, chroma intervals are calculated out to the **optimal color** limits.

The approximately cylindrical framework of the Munsell system (below), which includes a value scale limited by the luminance factor of surfaces and chroma steps bounded by the optimal color limits for each hue. The dimensions are not otherwise connected, which corresponds to the separate perceptual measurement procedures used to define them.



conceptual framework of the munsell color order system (1905)

Colors are "named" through the standard notation sequence *hue value/chroma*. Thus, the pigment **vermilion** would be notated as **8.5R 5.5/12** — a hue of 8.5 in the **R** segment, at value 5.5 and chroma 12. Any physically realizable surface color (paint, ink or dye) can be located within the Munsell framework by using decimal increments on the hue, value and chroma dimensions. (Munsell apparently adapted his color terms from the French *valeur* and *chromatisme*, which he picked up while studying art in Paris in the 1880's.)

So how does one measure color with the Munsell system? At first (and even today) through human color judgments. The Munsell colors are standardized as carefully prepared color samples or paint chips, presented on separate pages of a reference catalog, the **Munsell Book of Color**. Each page contains color samples of a single hue, arranged in a two dimensional grid defined by value and chroma. Colors are identified by placing them next to the atlas samples until the nearest color match is found. Decimal color values are inferred by locating a sample color between two existing Munsell chips. These visual judgments are accurate only if the comparison color sample is as large as the atlas samples, and all colors are viewed on a gray background under the same daylight or incandescent illumination (usually, illuminant C).

However, conversion formulas and lookup tables published by the ASTM are now available to produce a Munsell equivalent for any spectrophotometric (CIE XYZ or Yxy) color specification. These allow translation to or from Munsell and any other modern color system also on the CIE standards.

**The Munsell Renotations.** In the 1920's some of the 1915 Munsell samples were first measured with a spectrophotometer, and these were mapped into the CIE **2° Yxy color space** shortly after it was developed in 1931.

The early measurements were made by the USA National Bureau of Standards, which took an interest in developing the Munsell system as a color standard. In the 1930's the Optical Society of America also began a technical study of the system. The entire 1929 Book of Color was measured spectrophotometrically in 1935 and published as **Yxy** values in 1940. These revealed major discrepancies between the 1915 and 1929 color samples, and major irregularities in the spacing of color samples on the **Yxy** chromaticity diagram, particularly in the chroma intervals but also in the lines of equal hue.

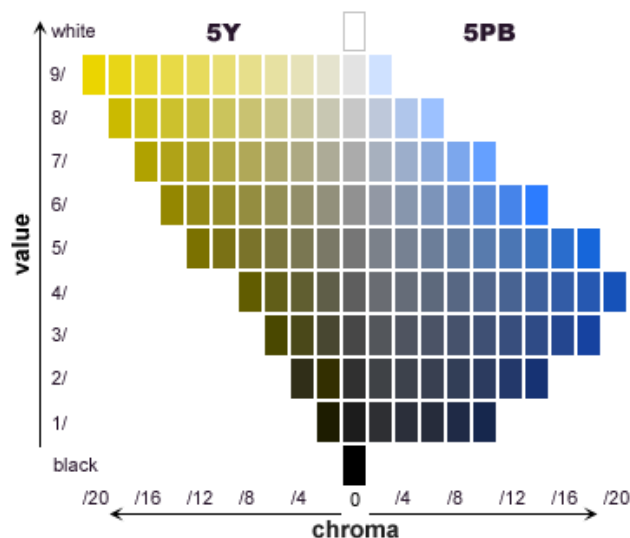
The OSA committee primarily focused on defining the Munsell value scale as a mathematical formula, based on visual judgments of gray samples by a large sample of viewers, and on judgmentally adjusting or "smoothing" the contours of Munsell hue and chroma on 1931 **xy chromaticity diagrams**. These calculations were used to revise the value scale, to equalize the hue and chroma intervals, to extend chroma intervals out to the optimal color limits, and to redefine material color specifications. This **renotated system** was published in 1943 as specific **Yxy** values for 2700 unique color locations, as a mathematical formula to calculate Munsell value from **XYZ** tristimulus values, and as several 1931 xy chromaticity charts used for the visual estimation of Munsell hue and chroma values from xy values. Subsequent editions of the Book of Color are blended to match these 1943 targets. The **renotated system is the standard specification** of Munsell colors used today.

Subsequently, the OSA committee working on uniform color scales (**discussed below**) revisited the renotation data, generally to examine the effects of imposing a genuinely *three* dimensional definition of color differences, but particularly to examine the effects of **response compression** on the chroma of different hues. The data for this **re-renotation system** were published in 1967 for more than 2900 color samples, and represent a major revision to the Munsell standards. It is not in general use.

Today all the Munsell colors are available as idealized **tristimulus values** or *aim colors*, and **Yxy** color locations or color spectrophotometric data can be translated into the Munsell notation entirely by computer software. Thus the color system designed for teaching color to children has ascended to the realm of digital exemplification and industrial color standards.

**The Asymmetry of Color Space.** Munsell and his successors worked hard to produce color samples that were **perceptually equidistant** or equally different from neighboring colors on the individual dimensions of value, chroma and hue. Thus, the difference between a chroma of 2 and a chroma of 4 is intended to appear to the viewer as large as the chroma difference between 12 and 14. However, these **measurement units are not comparable across dimensions**; the principle of equal perceived distance applies *within* each dimension of the model separately, not between any combination of color attributes. This is obvious from the fact that 10 units encompass the entire range of values (from black to white), but only 10% of the total range of hues (for example, from yellow to orange) and, for most hues, only half the total range of chroma.

Like Rood, Munsell originally conceived of his color model as a sphere. But the empirical measurements that have been at the heart of the Munsell system from the beginning compelled Munsell to an important realization: the **natural color space is highly irregular** when it is represented without geometrical preconceptions. As a result Munsell allowed for unequal dimensions of chroma at different levels of lightness and across different hues. Two "pages" from the Munsell system at the complementary hues **5Y** and **5PB** show this clearly.



**two pages from the munsell book of color (1929)**

These hues are directly opposite each other on the Munsell hue circle, so they show a vertical "slice" through the Munsell color solid (sometimes called a *color tree*). This shows what is obvious

when you think about it: yellow at its highest chroma has a much lighter value than does blue at its highest chroma.

These lopsided relationships between value and chroma **affect nearly all hues** around the color circle. As a result, **no simple geometrical form accurately represents perceptual color space**. All color models based on triangles, circles, squares, pyramids, cones, spheres, cubes or cylinders must (and do) grossly distort perceived color relationships. Munsell was the first researcher to incorporate this basic color fact into a color model.

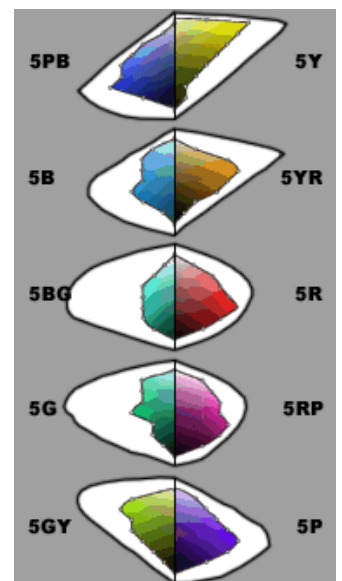
Munsell's color studies and the subsequent renotation research revealed that the background or surround color significantly affects the response compression (contrast) in lightness or chroma; in effect, lightness and chroma *are* contrast perceptions of brightness and colorfulness. In addition, the Munsell workers documented poor chroma or lightness discrimination in very intense or very light valued colors, and uneven perception of hue differences, especially as hue differences become very large. By its focus on *perceptual uniformity* in the spacing of material color samples, the Munsell research uncovered the many *geometrical complexities* in color space.

**The Limited Gamut of Color Atlases.** The continued use of human color judgment with the Munsell or NCS color atlases may create the impression that all surface colors are contained in these references. This is false, and leads to the issue of **color gamut** as it applies to *any* hardcopy (printed or painted) or softcopy (computer monitor) displays of a color space.

The range of colors represented in a Munsell atlas is limited by the gamut of paints or inks used to create the color samples. In fact, two different versions of Munsell color samples, in "glossy" and "matte" preparations, are available. The glossy colors have a slightly larger gamut or range of chroma.

But even the glossy colors are incomplete. The figure at right shows the range of surface colors represented by the Munsell color solid as irregular colored shapes, enclosed within a white envelope that represents the **optimal color stimuli** or maximum chroma for surface colors of each hue. (Colors with a higher color intensity than the white boundary will appear fluorescent or self luminous.) The Munsell system represents the range of colors that can be achieved by high quality acrylic paints, or inks printed on highly reflective white paper; the range of colors possible with watercolor pigments is significantly smaller.

Comparison of the Munsell color range with the optimal color envelope shows that the yellow, orange, red and magenta (5RP) pigments do a fairly good job of filling out the potential color space, while modern lightfast **blue** (5B), **blue green** (5BG) and **green** (5G) pigments reach less than half the chroma potential for their hues. Thus, the Munsell color atlas, and in fact any color atlas in any color system, exemplifies an incomplete and skewed subset of the total asymmetrical color space. Note, for example, that the **G** dimension of the Munsell atlas is smaller than the **RP** dimension, although in the perceptual surface color space just the reverse is true.



**maximum chroma of munsell color samples**

*compared to the limits of surface color chroma; after Kuehni (2003) and Perales, Mora et al. (2004)*

## OSA uniform color scales ▲

As we've seen, the Munsell system contains two significant problems: (1) a variety of discrepancies were found in the



perceptual spacing of colors, depending on their location in the color space, and (2) the *quantitative* difference between colors could only be defined on a single color attribute (lightness, chroma or hue) at a time. These problems were not resolved in the Munsell renotation, and the Munsell system remained an inconvenient and uncertain basis for quantifying **perceived color differences** that had to be defined on all three colormaking attributes.

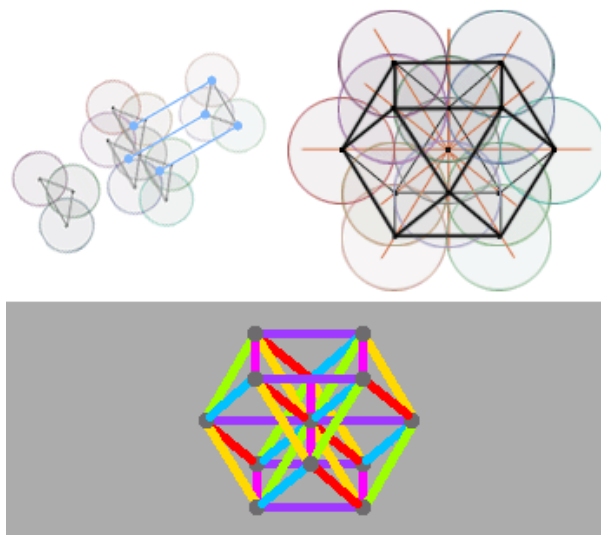
The subsequent development of the **colorimetric system** did not resolve matters, as the standard **chromaticity diagram** also badly distorted perceived color differences — the chromaticity differences between greens are much too large, and differences between blues are much too small.

Methods to quantify perceived color differences were very important in the implementation of color measurement, for purposes as diverse as photographic dye patents and governmental standards, and for manufacturing color control in industries as dissimilar as textiles, food products and automobiles. These requirements stimulated an effort to develop a truly uniform color space, from the ground up, as a set of *uniform color scales* across a limited value range of moderately saturated colors. This task was taken up by a committee under the auspices of the Optical Society of America (OSA), in collaboration with the USA National Bureau of Standards, and led by Deane Judd and David MacAdam. The final UCS aim color tristimulus values, color scales, recommendations for their use and conclusions regarding the human color space, were published in the 1970's as the Optical Society of America's **Uniform Color Scales (UCS)**.

**Rhombohedral Lattice.** The first issue was the specification of the color geometry. The decision was made to shift from the "wheel and spoke" arrangement of Munsell to a rigorously cartesian or three dimensional set of dimensions, as in the CIE **Yxy color space**.

In addition, this color geometry should support equivalent measurement of perceived color differences in all directions (including directions in which all three colormaking attribute change), which meant that color differences along the diagonal dimensions in the cartesian space should be equivalent to color differences in any direction parallel to one of the three defining dimensions.

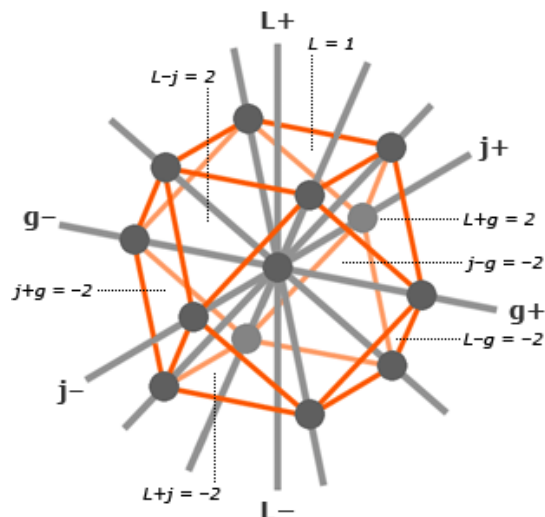
The OSA committee determined that the metric solution to both problems was to use a **rhombohedral lattice**, based on the cuboctohedron, as the organizing geometry for locating aim colors within the system. This is difficult to represent as a single image, but the logic behind it is straightforward and builds on the idea of a *regular stacking*.



### rhombohedral lattice structure of OSA UCS

(left) a triangle of three spheres fitted into a hexagon of seven spheres, with three more spheres fitted from below; (right) the resulting cuboctahedron with edges of equal length (black), linked by six scales passing through a central color; (bottom) the six sets of parallel color scales identified by color

The densest possible stacking of three spheres in two dimensions is with the center of the spheres forming an equilateral triangle (as billiard balls are stacked in the rack). The densest possible stacking in three dimensions is pyramidal: a single sphere is fitted into the gap between three spheres below, as oranges are stacked in a grocery bin or carbon atoms are stacked in a diamond crystal. The cuboctahedron is the geometric solid that defines this stacking. Starting from a single sphere that marks a neutral, mid valued gray, we continue this stacking in all directions out to the perceptual limits of the color space, creating a lattice of color "atoms" of color in this stacking arrangement.



### cuboctahedral geometry of the osa uniform color scales

lightness is measured on the  $L+/L-$  dimension, yellow/blue on the  $j+/j-$  dimension and red/green on the  $g+/g-$  dimension; rhombohedral stacking of color "atoms" defines a total of seven unique cleavage planes through the color space (labeled), with additional two cleavage planes (not shown) defined perpendicular to the  $j$  and  $g$  dimensions

As in a diamond crystal, the stacking aligns the UCS aim colors (the "atoms" in the lattice) into nine planar color groupings or **cleavage planes** (diagram, above), and **six perceptual scales** defined by the intersections between any two cleavage planes.

The three cleavage planes perpendicular to the three measurement dimensions stack the target colors in a square lattice; all other planes are stacked in a hexagonal lattice. This creates a rigid measurement framework within a three dimensional (euclidean) color space, intended to reveal for color scientists any irregularities in the color space as discrepancies in color distances calculated across the different scales or cleavage planes. In addition, the various color scales and color planes were suggested to artists as dimensions of harmonious color variation.

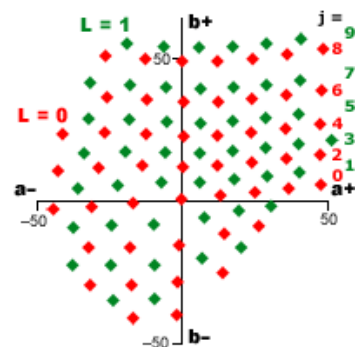
The lattice naturally suggests four measurement benchmarks: (1) the interval between any two adjacent color samples represents a **just noticeable difference** between the colors; (2) the smallest intervals or steps between colors must be perceptually equal in all directions and all parts of the color space; (3) the perceptual uncertainty or disagreement between observers in identifying a specific color must approximately form a **sphere** around the average color specification; and (4) the spheres around all color samples will be the same size — not the ellipses of different sizes and shapes that appear in the CIE **chromaticity diagram**. In fact, both the color stimuli originally used to measure the perceptual color space, and the OSA UCS aim colors (the standard tabular exemplification of the system), were spaced by about 20 jnd's, but in theory this should only change the perceived distance between color samples, not the geometrical regularity of the color scales. Unfortunately, as the OSA committee discovered, this assumption turns out to be false.

**The OSA Uniform Color Space.** The UCS exemplification projects the rhombohedral lattice into the color space so that a cleavage plane of square stacking defines the hue/chroma planes at constant lightness; each square defines a measurement interval of two units. These planes are staggered or shifted in relation to each other, so that chromaticity coordinates are even or odd numbered on alternating lightness levels. The lightness dimension is perpendicular to these planes.

The three perpendicular dimensions of the **uniform color space** are the lightness dimension **L** and two chromatic dimensions, **j** (for *jaune*, approximately a y/b opponent dimension) and **g** (for *green*, approximately a g/r opponent dimension). The mid valued gray that is the anchor of the system has a value of 0 on all dimensions; colors with negative values are darker, bluer or redder.

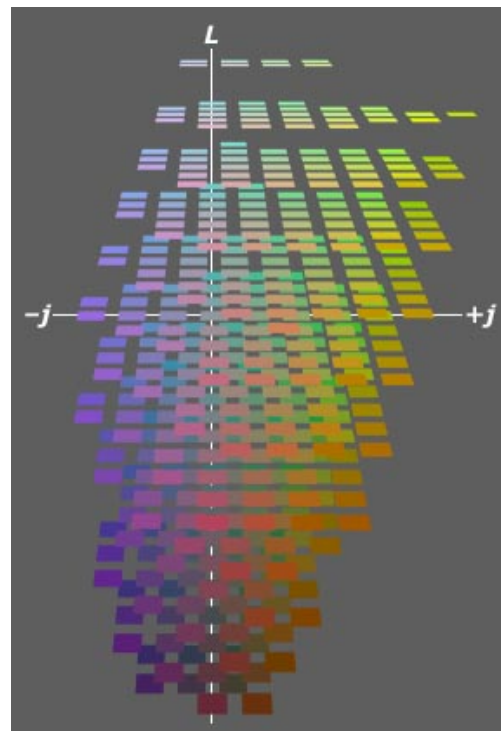
The UCS dimensions are computed from the stimulus CIE XYZ tristimulus values for the **10° standard observer**, as viewed under the **D65** illuminant against a neutral background of 30% reflectance (light gray). The UCS dimension formulas are actually calculated using the luminance **Y** and the standardized **x<sub>10</sub>y<sub>10</sub>** coordinates of the chromaticity plane. In other words, despite the intricate cuboctahedral lattice, the OSA UCS is a continuous color space, and any material color specified under **D65** can be located within it.

Unfortunately the OSA UCS exemplification as 558 color samples is no longer manufactured or available. To redress this in part, I've posted in **this directory** large format image files of the standard OSA full step aim colors, displayed as 111 cleavage planes with the correct rhombohedral spacing between color samples.



**staggered square tiling of hue planes in OSA UCS**

*two adjacent lightness levels shown on the CIECAM **ab** plane, with numbering of **j+** samples*



### the OSA UCS color solid

*conceptual side view of the OSA UCS whole step aim colors, with the vertical **L** scale expanded for visual clarity; adapted from **Yxy** tristimulus values in Table I(6.6.4) of Wyszecki & Stiles (1982), which reverse the orientation of the *g* dimension*

The development of the OSA UCS sprawled across three decades and many color scaling experiments, drew on dozens of previous color scaling studies, as well as data from the Munsell renotation and the CIE **Yxy** system, and began with psychophysical studies of limited color judgment tasks, such as equal hue spacing at constant lightness and chroma, or equal chroma spacing at constant lightness and hue. In effect it was a conceptual integration of the best color research up to that time, with empirical testing of the major conclusions and solutions.

These studies led to a selection of 43 hexagonally arranged, perceptually equally spaced color samples at a constant Munsell value of 6. I reproduce below the 48 OSA aim colors for  $L=1$  as approximate equivalents to these color stimuli. As the example shows, the range of colors studied is relatively muted: the most intense color sample (the orange at top right) has a CIELAB chroma of 74.



### OSA UCS constant lightness plane

OSA UCS aim colors for  $L = 1$  (CIELAB  $L^* = 63$  to  $68$ ) under equal energy illuminant;  $j = 9$  to  $-5$  and  $g = -9$  to  $5$ , with neutral indicated by white dot; data from Wyszecki & Stiles (1982)

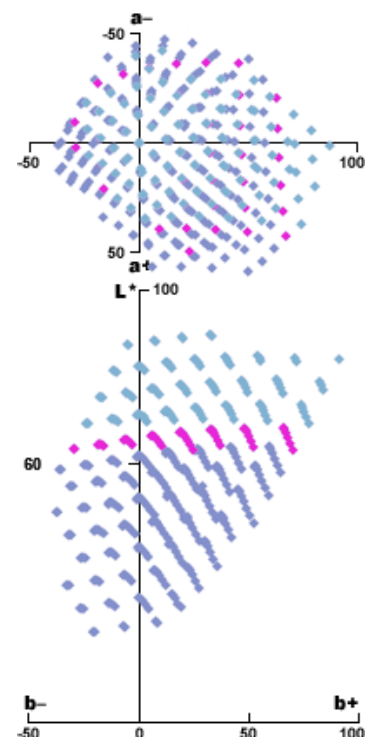
Several groups of viewers examined two pairs of colors sampled at random from this set, and were asked to choose the pair that showed the *larger* color difference; by statistically combining these color judgments across all viewers, the perceived differences between all colors could be given an average numerical value. These values were then used to check formulas proposed to estimate perceived color differences, based on the **Ljg** values of two colors. Subsequent studies using revised or limited color samples were done to extend the results to different values and to adjust for the **Helmholtz-Kohlrausch effect**. The final rhombohedral model of 500 color samples and 58 pastel samples was built on these foundations.

The diagram (right) shows the 500 OSA UCS "full step" aim colors projected into the **CIELAB** color space. The average distance between aim colors on the CIELAB  $L$ ,  $a^*$  or  $b^*$  dimensions is about  $\Delta E_{ab} = 12$  — that is, about 12 jnd's. The image shows the sampling is fairly complete (when compared for example to a similar representation of **watercolor pigments**) and is generally regular, with a tendency for CIELAB chroma spacing to grow larger in yellow and green colors. (The slight downward bowing visible in the color rows is a compensation for the Helmholtz-Kohlrausch effect.)

The various diagonal scales or cleavage planes through the color space were intended to identify palettes of complex color change that might be **useful to graphic artists and color designers**. They turned out to be wildly unpopular for that purpose. The legacy of the OSA UCS is as a research landmark and a uniform tiling of the color space that is useful to assess color models or imaging systems.

**Color Space Is Noneuclidean.** Despite the extensive and patient research program and the elegant color model it produced, the OSA committee had already in 1967 come to a dispiriting conclusion:

*We will affirm to the world that no regular rhombohedral lattice sampling of color space, with a fixed background, can exist; we*



OSA UCS aim colors in the CIELAB color space

will produce the best approximation to such a lattice for a neutral value 6 background that we can design.

In plain language: **it is impossible to represent uniform color differences as a three dimensional color model.** The human color space is **noneuclidean**, which means a three dimensional framework cannot describe it. Any color difference formula or color model that represents color in only three dimensions (such as lightness, hue and hue purity) must produce a distorted or inconsistent picture of color differences and color relationships.

The first and perhaps most important reason why this is so was identified by Deane Judd in 1968 as **hue superimportance**; the effect might better be described as *chroma compression*. It appears as a fundamental discrepancy between measures of just noticeable difference in chroma and in hue. The diameter of any hue circle at constant chroma is always perceptually shorter than the circumference across all hues as calculated from that chroma "radius". This discrepancy is quite large: if the ratio of the circumference to the diameter of a euclidean circle is roughly 3.14 to 1, the same ratio in a hue circle is greater than 6 to 1.

Judd represented this fact as a "pleated disk" (diagram, top right): chroma is measured outward along the folds in the pleats, while hue is measured zigzag along the edge of the fan, which has roughly double the length of a flat circumference. Alternately, if perceived hue and chroma differences are represented on a plane of equal lightness, then chroma differences within the same hue must be elongated relative to hue differences within the same chroma, producing elliptical discrimination ellipses. In addition, these ellipses must become larger (chroma discrimination becomes weaker) at higher chroma values (diagram, bottom right). The OSA committee concluded that both lightness and chroma had to be adjusted by a cube root or **exponential compression** to correct for this discrepancy.

The second problem involves **chromatic contrast**, in particular **crispening**, which makes small color differences, or differences judged against a similar background, appear relatively larger. Most of us are familiar with the "reading magnifiers" that enlarge words when rested on a printed page. Crispening acts like a magnifying bubble inside the color space, enlarging the apparent contrast between similar colors on lightness, chroma and hue. As a result, all **small color differences are magnified in comparison to large color differences**, and "small distance" color difference formulas tend to overestimate the perceived color difference between colors that are very different.

Just as a reading magnifier can be shifted around a page, the crispening locus can migrate around the color space, depending on the color and lightness of the background (surround) used to judge colors, the chromaticity of the illuminant. To control this, color vision experiments and color atlas viewing standards use a mid valued **gray surround** in color presentations, so that crispening occurs as an increased chroma sensitivity around near neutral colors.

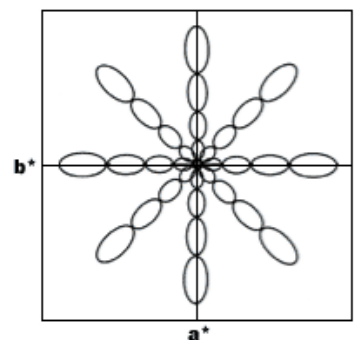
The third problem is that hue superimportance is not equal in all directions: **chroma compression is strongest in yellow and green hues** and weakest in red, purple and blue hues. It also changes across different values, becoming more extreme in lighter valued (more luminous) colors.

under illuminant D65; layer  
**OSA L = 1** shown in magenta;  
the **g** dimension is reversed in  
all CIE color models



**the hue circle as a pleated disk**

*the circumference is more than six times the length of the diameter;  
after Judd (1968)*



**hue/chroma discrimination ellipses in CIELAB**

*from Kuehni (2003)*



The final problem, related to the previous one, is the **unequal spacing of hues**, which is explored on a **previous page**. Some hues produce a greater number of perceptible color gradations than others; in particular, color discriminations are generally more acute for warm hues than for green hues.

The two fundamental problems in color models uncovered by the OSA UCS research strangely parallel the two **primary color paradoxes**:

- The first is that *all color models we can see cannot represent colors relationships accurately*, and color models that represent color relationships accurately cannot be seen (physically constructed or represented as an image).
- The second is that *all color difference formulas are arbitrary*, in the sense that they must include many judgmental corrections for specific discrepancies, can only provide difference estimates that are approximately correct for small (or large) color differences, and provide difference estimates whose accuracy depends on where in the color space the color difference is calculated.

The primary color paradoxes can be explained as resulting from the overlap in three cone sensitivity curves. The color model paradoxes can be generally explained by the fact that **the perceptual color space changes continually** under the influence of many **contextual factors**. It's not that a changing metric cannot be expressed as a three dimensional color model; it's that the changes within a perceptual space vary depending on the color stimuli, color difference, contrast backgrounds, illuminance levels and light sources that are being used to assess the color difference.

A detailed and very useful explanation of the OSA-UCS scales and their application in color design issues, by Danny Pascale, is available from **BabelColor.com**.

## CIELUV uniform color space ▲

The previous section described the efforts by the Optical Society of America to build a truly uniform color space in three dimensions based on rigorous psychometric and geometrical foundations, and the discovery that such a goal is **fundamentally impossible**. Now we turn to a second strand of color research, which had a more limited aim in view: to discover a **uniform color difference formula** that could be used in specific practical applications.

**Color Matching and Color Differences.** Since the early 1920's, a major goal of **colorimetry** was the prediction of an **unrelated color** response through the photometric measurement of a physical color stimulus. The approach that seemed most efficient was to measure the **spectral power distribution** of the light emitted or reflected by the color stimulus, translate this into photometric units by means of a **luminous efficiency function** or a standard set of **color matching functions**, then use these quantities to estimate the subjective brightness or chromaticity of the stimulus when it is viewed as an isolated patch of color.

However, the **XYZ** tristimulus values that came out of color matching tasks literally only **predict color identity**, that is, whether two colors will *appear the same*. The whole problem of how much two colors will *appear different* was left

undetermined. The **CIE xy chromaticity diagram** does represent the *relative* difference between two colors — that one color sensation is redder, or brighter, or duller than another, if both colors are measured in the same way. But the desired next step was to subtract one set of **XYZ** tristimulus values from another, and use the numerical difference to **predict the perceived color difference** between them.

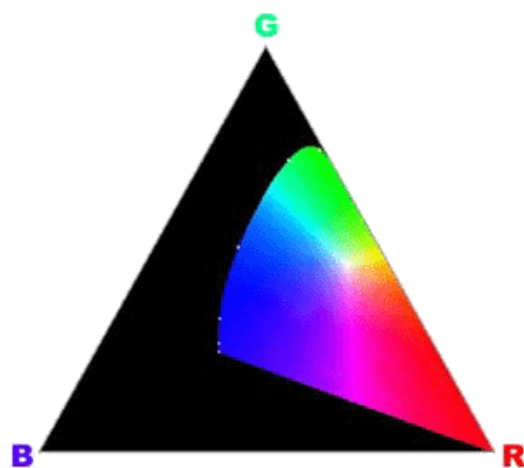
This estimation procedure did not work well. A major problem was that the CIE xy chromaticity diagram is not a *uniform chromaticity scales* diagram or **UCS**. In a true UCS an equal chromaticity distance is an equal perceptual difference, no matter which color attributes are measured. But two colors separated by an equal **Yxy** (or **XYZ**) difference can appear nearly identical (if the colors are green or light valued) or very different (if the colors are violet or dark valued). They cannot consistently predict whether a color difference will be *unnoticeable or obvious* to an average viewer. This was a major obstacle to using colorimetry in the industrial control of dyes (textiles and photography), inks (printing and packaging) and paints (automotive and architectural colors) — especially after inexpensive and reliable spectrophotometers became available in the 1950's.

As the measurement standards were already in place, the easiest solution was a hack. The original chromaticity diagram could be mathematically stretched or compressed until the *numerical* difference between the tristimulus values of two colors approximated the *perceived* difference between the colors. (These differences would be applied only to similar colors: large color differences were not important to quality control problems.) These mathematical distortions are called **projective transformations**, and they are identical to the warping or "keystoning" that occurs when a slide is projected onto a curved surface or onto a wall at an oblique angle (right). These transforms don't change any of the foundation procedures of colorimetry or the "straight line" color mixing relationships of the chromaticity diagram, just the color differences calculated from them.

Projective transformations were proposed independently by Deane Judd and David MacAdam in the 1930's; Judd's version is shown below. It radically compresses the "green" part of the chromaticity space while expanding the "blue" and "red". The white point is located close to the R-G mixing line, because the **L** and **M** cones dominate the perception of whiteness (compare with my **cone chromaticity diagrams**).



"keystoning" as an example of a projective transformation



an early projective transformation

*the CIE 1931 chromaticity diagram within a Maxwell triangle defined by three imaginary primary lights that produce an approximately uniform color space (after Judd, 1935)*

Projective transformations were only one of several solutions to the color difference problem. By the 1960's anywhere from 13 to 20 different formulas were being used for different applications in England and the USA. To corral this diversity, the **Commission Internationale de l'Éclairage** (CIE), an international clearinghouse for color research at universities and research laboratories, adopted in 1976 two color difference formulas that seemed most promising for further work: **CIELUV**, described here, and **CIELAB**.

**The CIELUV Dimensions.** CIELUV is a much revised and expanded version of the projective transformation proposed by David MacAdam in 1937 and adapted as a color difference space by Günter Wyszecki in 1964. It builds on two **uniform chromaticity scales** (UCS), denoted **u'** and **v'**, derived from the standard **XYZ** tristimulus values.

Recall that a basic problem in colorimetry is devising a measure of chromaticity (hue and hue purity) that is not affected by changes in the luminance (brightness or lightness) of the color. The CIE **xy chromaticity diagram** accomplishes this by *normalizing* the **XYZ** values, or dividing each score by the **sum of XYZ values**:

$$\begin{aligned}x &= X / (X + Y + Z) \\y &= Y / (X + Y + Z)\end{aligned}$$

The CIE UCS adopts the same normalizing strategy for the color's tristimulus values, but with a specific reweighting of the **X<sub>c</sub>**, **Y<sub>c</sub>** and **Z<sub>c</sub>** scores that changes the shape and orientation of the chromaticity diagram, and the location of the **white point** inside it:

$$\begin{aligned}u'_c &= 4X_c / (X_c + 15Y_c + 3Z_c) \\v'_c &= 9Y_c / (X_c + 15Y_c + 3Z_c)\end{aligned}$$

The **u'** and **v'** scales alter the dimensions of the chromaticity space but preserve the trichromatic color geometry. That is, in the CIE UCS and CIELUV, **Grassmann's Laws** of color mixture are still valid: all colors are defined by additive mixtures of three imaginary "primary" lights, and therefore **color mixtures are described by straight mixing lines**.

The next problem is representing the luminance or intensity of the illumination and **chromatic adaptation** to its color. To incorporate these effects, the luminance factor **Y<sub>c</sub>** is used to expand the dimensions of the **u'<sub>c</sub>v'<sub>c</sub>** diagram to approximate the increase in colorfulness that occurs with increased illuminance, and to shift the location of the white point to center it on the chromaticity of the light source, defined by a second set of tristimulus values (**X<sub>w</sub>Y<sub>w</sub>Z<sub>w</sub>**). This yields the *relative* brightness dimension **L\*** and the *relative* chromaticity dimensions **u\*** and **v\*** (**u'<sub>w</sub>** and **v'<sub>w</sub>** are computed from **X<sub>w</sub>Y<sub>w</sub>Z<sub>w</sub>** as shown above):

$$\begin{aligned}L^* &= 116*(Y_c/Y_w)^{1/3} - 16 \\u^* &= 13L^*(u'_c - u'_w) \\v^* &= 13L^*(v'_c - v'_w)\end{aligned}$$

These adjustments center the **u\*** and **v\*** dimensions on the achromatic point at the origin of a **u\*v\*** cartesian plot, insert a **cube root compression** into the brightness contrast, and (by

the factor **13L**) mimic the **Hunt effect** which causes colors to increase in colorfulness as their luminance increases. These calculations apply to colors perceived as lights, so extremely dark colors — with luminance ratios ( $Y_c/Y_w$ ) less than 0.01 — are out of bounds.

At last we get a definition of **chroma** across different levels of brightness as the euclidean chromaticity distance between the color and the white point:

$$C_{uv} = [u^{*2} + v^{*2}]^{1/2}$$

... we get a simple numerical definition of hue as the **hue angle** measured from the white point in relation to the  $u^*$  dimension:

$$h_{uv} = \arctan[v^*/u^*]$$

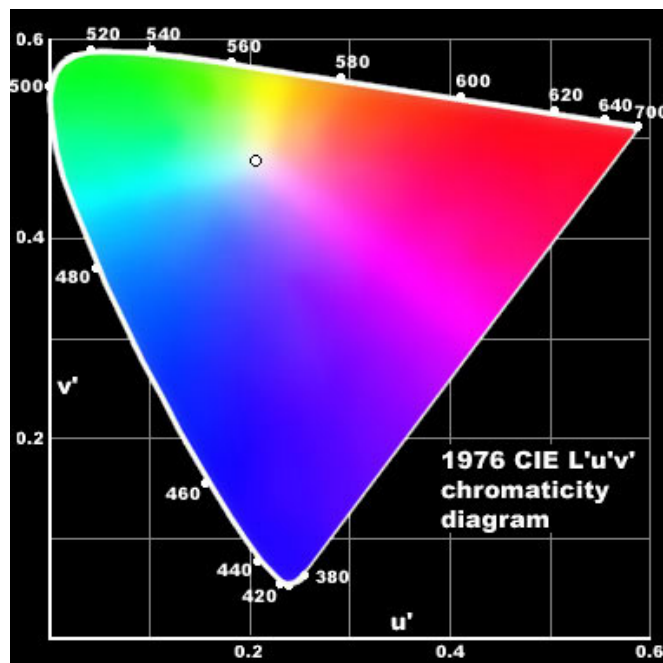
... and we end up with a version of the color difference score that was the origin of the model,  $\Delta E_{uv}$ , as the euclidean distance between the  $L^*$ ,  $u^*$  and  $v^*$  values of two *similar* colors:

$$\Delta E_{uv} = [(L^*_1 - L^*_2)^2 + (u^*_1 - u^*_2)^2 + (v^*_1 - v^*_2)^2]^{1/2}$$

where "similar" means colors that are separated by a  $\Delta E_{uv}$  value of 10 or less.

**The Uniform Color Space.** The result is the 1976 CIE  $L^*u^*v^*$  uniform color space or **CIELUV**, in which the brightness and chromaticity difference between two colors is proportional to their *perceived difference* when viewed side by side in a gray field at moderate brightness.

CIELUV provides the same chromaticity dimensions as the  $u^*v^*$  **uniform chromaticity scales** diagram (the CIE 1976 UCS), shown below. This represents equal chromaticity differences between lights of equal (unspecified) luminance as approximately equal euclidean distances on the chromaticity plane.

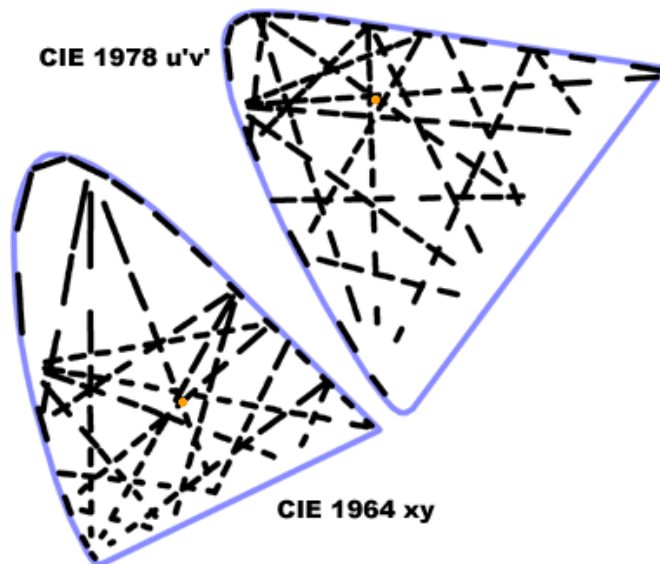


**CIE 1976 UCS diagram**

*computed from 10° XYZ values under an equal energy illuminant;  
the diagram shows all possible colors as they appear in lights,  
bounded by the pure spectral hues; colors within the diagram only  
symbolize the approximate location of colors*

The top edge of the color surface spans the hues of monochromatic or single wavelength light measured from 750 (700) to 520 nm; the left diagonal edge completes the spectrum from 520 nm to the short wavelength visual limit at around 380 nm. The right diagonal edge (called *the purple line*) represents the extraspectral crimson, magenta and purple hues that are mixtures of "red" and "violet" light.

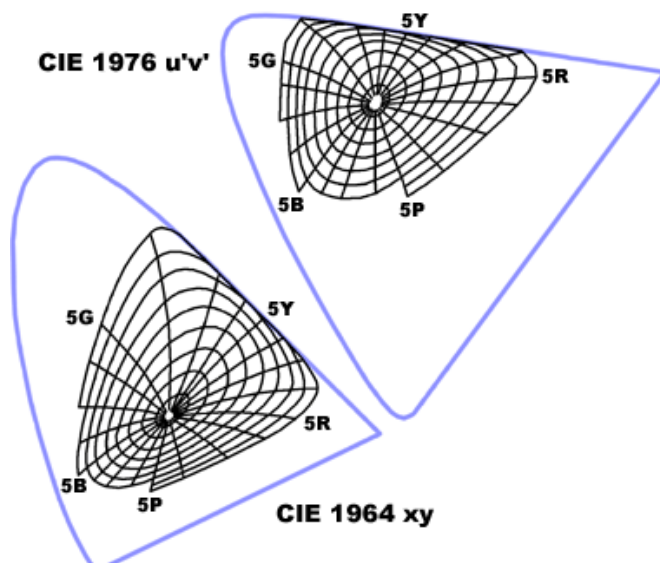
The diagram below shows the effect of the **u'v'** scaling on equal chromaticity differences measured in different directions around the chromaticity surface. The equalizing effect is especially noticeable by comparing the color differences around the spectrum locus, and the color distances from "blue" toward "green" and "blue" toward "red".



#### comparison of equal perceptual distances

*equal perceptual distances, drawn at 10 times actual size, in CIE xy and u'v' chromaticity diagrams; after Wright (1969) and Hunt (2003)*

Throughout the 20th century, the Munsell system figured prominently in efforts to identify color scaling problems or confirm that a color difference formula accurately described color differences. The diagram below shows the relative improvement in the spacing of Munsell lines of equal hue and chroma for surface colors of value 6.



#### comparison of munsell hue and chroma

*lines of equal Munsell chroma (steps of 2) and hue (steps of 5) at value 6 in CIE xy and u'v' chromaticity diagrams; after Wyszecki &*

*Stiles (1982) and Kuehni (2003). Note that all CIE systems reverse the ordering of Munsell hues.*

The improvement is especially noticeable in the more circular shape of the chroma contours around the white point, in the more equal spacing of green hues, and in the balanced chroma intervals around the green, blue and purple sides of the chromaticity space. However, significant discrepancies (chroma compression or expansion) appear in the yellow green to red parts of the diagram, with the greatest compression approximately along the yellow (5Y) hue dimension.

**Current Use of CIELUV.** The CIE 1976  $u'v'$  diagram is the "least wrong" chromaticity diagram in general use, and is a huge improvement over the distorted 1931 or 1964 diagrams which still lurch through the color literature. It gets the relative sizes of the green, blue and red areas of color space about right, as well as the **perceived hue purity** of colors, including spectral hues, as the distance from the white point. It should always be used to display basic chromaticity data.

CIELUV has been a popular color space for lighting, video and photographic applications which require measurement of lights, light filters or optical systems, because CIELUV represents the additive mixture of two lights as a **straight line** across the color space. (R.W.G Hunt's authoritative 2004 text, *The Reproduction of Color, 6th edition*, consistently relies on it.) CIELUV is quite useful to predict the appearance or mixture of colors produced by luminous displays — such as computer monitors, television screens, or transparencies projected onto a uniformly reflecting surface — where the brightness of colors is a constant proportion of the luminance of the source illumination (there is no "external" light source), although CIELUV is also used to specify the color of gems or optical glasses illuminated by a standard "white" light source, and the color of photographic dyes and filters.

Considered in terms of the **four requirements** of a modern color model, CIELUV (and the **Yxy** standard observer system on which it is based) represent a major departure. This isn't that the *color specification* is exclusively in terms of continuous numerical values, because that was innovated centuries ago by **Isaac Newton** and **Tobias Mayer**. The innovation is that there is **no color exemplification**. All color measurements are made electronically, using a **spectrophotometer** that measures the separate **XYZ** values directly and then calculates the coordinate values for any standard color system. The exemplification is implicit in all physical color stimuli, whether lights or surfaces, that have identical or similar color coordinates.

However, CIELUV suffers from a major drawback: the  $\Delta E_{u^*v^*}$  color difference formula yields a **poor prediction of color differences**, especially for surface colors. In fact, when a variety of color models are tested for their ability to predict surface color differences, CIELUV is almost always at the bottom of the heap.

CIELUV does even worse at predicting **corresponding colors**, or colors that match under different colors of illumination (illuminants). The problem is that CIELUV represents relative color differences in relation to the white point as a *difference* or simple subtraction (e.g.,  $u'_c - u'_w$ ), rather than as a *ratio* or proportion. This means that if the white point (the illuminant) is changed, or is strongly chromatic, all colors are shifted in a parallel direction by a constant amount (as if printed on a



moveable transparency), regardless of where they are located in the chromaticity space.

If colors are shifted from yellow toward blue, the compressed chroma intervals in the yellow green to orange part of the **u\*v\*** chromaticity diagram do not change, so CIELUV grossly overestimates chroma differences wherever these boundaries are relocated. If colors are shifted from blue to yellow (for example, the color of a gem viewed under tungsten rather than halogen light), CIELUV thrusts the predicted color locations completely outside the spectrum locus, which is perceptually impossible.

As corresponding color predictions are a minimal requirement for modern color appearance models, CIELUV has been discarded as the basis for further work.

## CIELAB uniform color space ▲

The second color difference formula adopted by the CIE in 1976 (along with **CIELUV**) was the *CIE L\*a\*b\* uniform color space* or **CIELAB**. CIELAB is based on a color difference formula developed by E.Q. Adams and D. Nickerson in the 1930's to measure color fading in textiles. Like CIELUV, it is an attempt to hack the **XYZ** tristimulus values to produce a reliable method for calculating **color difference scores**. However, CIELAB is based on a very different transformation strategy.

**The CIELAB Dimensions.** The CIELAB formula comprises four transformation steps, which are intuitive when examined separately.

Recall that the **color geometry** common to all modern color models consists of a brightness/lightness dimension that is perpendicular to (separate from) the chromaticity plane or hue circle. This geometry is implicit in the **XYZ tristimulus values**: the **Y** value represents the **photopic luminosity function** of the specified color. We start with the color's **Y** tristimulus value (**Y<sub>c</sub>**) as the measure of its luminance or luminance factor (labeled **L**, as in **CIELUV**), and use the **X<sub>c</sub>** and **Z<sub>c</sub>** values as measures of chromaticity (hue and hue purity).

The **X** value increases primarily as the **L** cone response to a color becomes relatively large, so it is the "red" anchor for an **r/g** dimension (labeled **a**); the **Z** value increases with **S** cone response, so it is given a negative sign to represent the complementary "yellow" of a **y/b** (**b**) dimension:

$$\begin{aligned} L &= Y_c \\ a &= X_c \\ b &= -Z_c \end{aligned}$$

The first transformation step is an adjustment of the tristimulus values to represent a **chromatic adaptation** to the illuminant (color of illumination). The correct method (a **von Kries transform**) divides the *cone excitations* produced by the color area by the cone excitations produced by the illuminant. CIELAB instead uses the "wrong" ratio between the *tristimulus values* **X<sub>w</sub>Y<sub>w</sub>Z<sub>w</sub>** of the illumination. These are a standard both for the luminance upper limit represented by a perfectly reflecting "white" surface (**Y<sub>w</sub>** is always equal to 100) and for the chromaticity inherent in the light itself (in the values of **X<sub>w</sub>** and **Z<sub>w</sub>** relative to 100).

$$\begin{aligned}L &= Y_c/Y_w \\a &= X_c/X_w \\b &= -Z_c/Z_w.\end{aligned}$$

The tristimulus values of a perfectly white or **equal energy illuminant** are defined as  $X_w = Y_w = Z_w = 100$ . An illuminant with a perceptible "warm" (yellow or red) chromaticity has values of  $X_w$  greater than 100 and values of  $Z_w$  less than 100; "cool" or blue illuminants have low  $X_w$  and high  $Z_w$ . The adaptation ratios reduce or increase the values of  $X_c$  and  $Z_c$  in proportion.

In the second step, CIELAB applies an **exponential compression** to all three adaptation ratios. This inflates the perceived chromaticity and lightness differences among dark colors, and inflates the chromaticity of yellows relative to blues. At the time CIELAB was published, different compression fractions had been tried in color difference formulas, and the *cube root* seemed to perform about as well as any:

$$\begin{aligned}L &= (Y_c/Y_w)^{1/3} \\a &= (X_c/X_w)^{1/3} \\b &= (Y_c/Y_w)^{1/3}.\end{aligned}$$

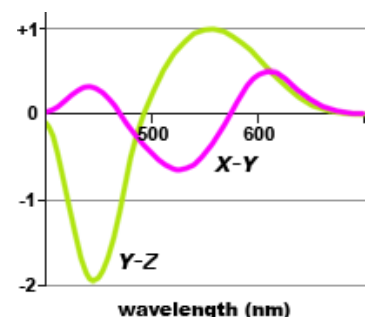
Although the **Y** tristimulus value represents the luminance of the color area, *all three* tristimulus values contain similar luminance information, because the **XYZ** values all get larger as the surface reflectance increases from dark to light. As a result, the **XYZ** values are highly correlated in actual color measurements, from 0.52 [**Y** and **Z**] or 0.56 [**X** and **Z**] to 0.88 [**X** and **Y**]. At the same time, the **Y** tristimulus value contains unique chromaticity information, because the **XYZ** curves all peak at **different parts of the spectrum**.

So the third transformation step separates the chromaticity information from the luminance information contained in all three values. CIELUV does this by *normalizing* the luminance information (**dividing all values** by the **XYZ** sum). CIELAB does it by *subtracting* one value from another:

$$\begin{aligned}L &= (Y_c/Y_w)^{1/3} \\a &= (X_c/X_w)^{1/3} - (Y_c/Y_w)^{1/3} \\b &= (Y_c/Y_w)^{1/3} - (Z_c/Z_w)^{1/3}.\end{aligned}$$

This creates two undulating curves for the difference values of spectral hues (diagram, right).

An explanation: when applied to quantities of the same thing, such as dollars or gallons, subtraction just identifies the quantity of the difference. When applied to quantities of *dissimilar* things measured within a similar range of values, **subtraction purges any information the two measures have in common** (in this case, information about the luminance) and leaves as a single number *all the unique information* in both (the combined information about hue and hue purity). As the  $X_c Y_c Z_c$  values are already standardized on the white point (by the adaptation transform), the subtraction also centers the values of **a** and **b** at zero for a surface color that exactly matches the chromaticity of the illuminant: a pale yellow surface under a pale yellow light will appear to be a "pure" white.



**opponent functions defined by tristimulus differences X-Y and Y-Z**

Finally, the **L**, **a** and **b** dimensions are expanded by different factors so that a unit scale value on any combination of CIELAB dimensions represents an approximately equal and just noticeable perceptual difference. As the **Y-Z** curve (**b** dimension) has a larger amplitude than the **X-Y** curve (**a** dimension, diagram above right), it is multiplied by a smaller amount. The slope and intercept of the **L** function is also adjusted to match the units of Munsell value:

$$\begin{aligned}L^* &= 116 * (Y_c / Y_w)^{1/3} - 16 \\a^* &= 500 * [(X_c / X_w)^{1/3} - (Y_c / Y_w)^{1/3}] \\b^* &= 200 * [(Y_c / Y_w)^{1/3} - (Z_c / Z_w)^{1/3}].\end{aligned}$$

These formulas take an alternate form, not shown here, for extremely dark colors (where any adaptation ratio is less than 0.009). For details see the references at the end of this section.

CIELAB yields a reliable correlate (relative measurement) of **chroma** as the euclidean distance between the chromaticity of the color and the achromatic point:

$$C_{ab} = [a^{*2} + b^{*2}]^{1/2}$$

... and a numerical definition of hue as the **hue angle**:

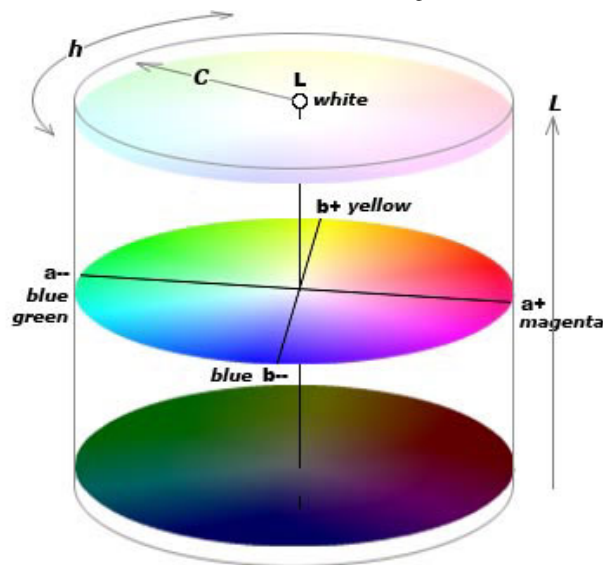
$$h_{ab} = \arctan[b^*/a^*]$$

... and its own color difference score  $\Delta E_{ab}$ , obtained as the euclidean distance between the **L\***, **a\*** and **b\*** values of two similar colors:

$$\Delta E_{ab} = [(L^*_1 - L^*_2)^2 + (a^*_1 - a^*_2)^2 + (b^*_1 - b^*_2)^2]^{1/2}$$

where "similar" means colors that are separated by a  $\Delta E_{ab}$  value of around 10 or less. As a rule of thumb, 10 units of CIELAB lightness exactly matches 1 unit of **Munsell** value, and 10 units of CIELAB chroma approximately matches 2 units of Munsell chroma. The average  $\Delta E_{ab}$  for the first chroma step in Munsell aim colors across all hues and values is about 5.8. Note that  $\Delta E_{ab}$  is approximately in units of just noticeable difference, so color differences at or below 1 are generally not visible.

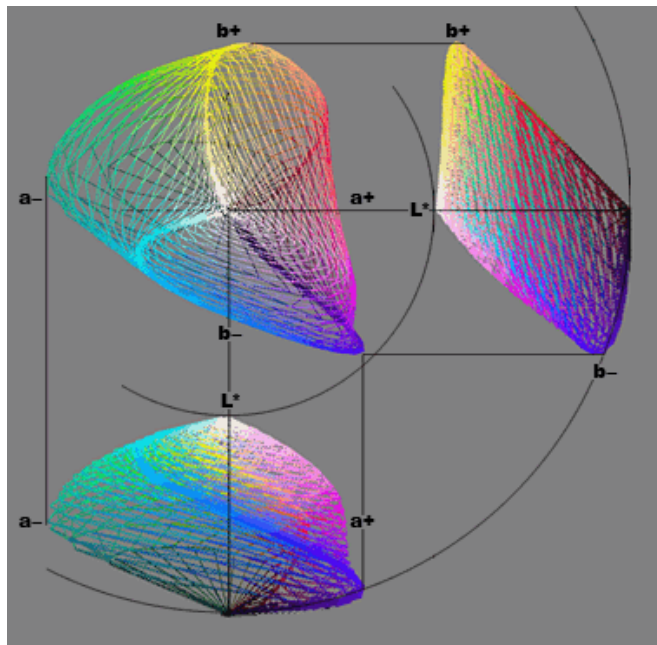
**The Uniform Color Space.** The CIELAB transformations yield approximately perpendicular dimensions that are often represented as a **cylindrical geometry**. That is, the **L\*** dimension is the axis of the cylinder, through the achromatic point; **h<sub>ab</sub>** represents the hue angle or circumferential direction of the color, and **C<sub>ab</sub>** the distance of the color from the cylinder axis.



### geometry of the CIELAB color model

The diagram is somewhat misleading. In practice, the color difference (euclidean distance) score is based on the cartesian ( $a^*b^*$ ) coordinates and not cylindrical radians or degrees; and the distribution of surface colors in CIELAB tapers toward the achromatic center at very high or low lightness.

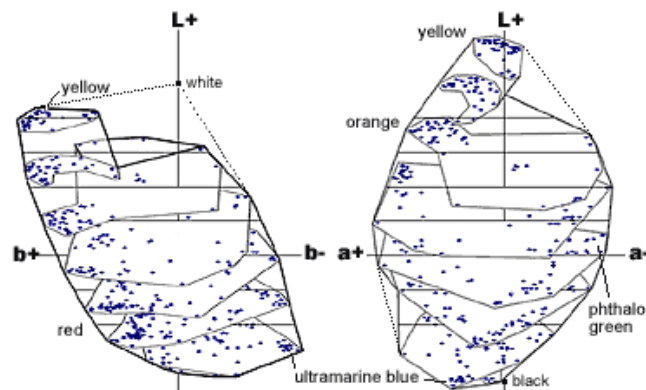
The point is that CIELAB color relationships are not distorted by forcing them into a *conceptual geometry*, such as the ipsative hue triangles of NCS. Color measurements are limited vertically by the luminance factor of surfaces, and the chroma extent is limited by the domain of real surface colors or lights, but these are boundaries imposed by color perception, not by arbitrary geometry or logic.



### surface of optimal color stimuli in CIELAB

composite from a *wireframe animation* © 2007 Bruce Lindbloom

The illustration below shows the location of 700 commercial watercolor paints in the CIELAB space, with color icons grouped onto separate lightness planes for clarity. These define a roughly ellipsoidal or "football shaped" color domain inside the CIELAB color space.

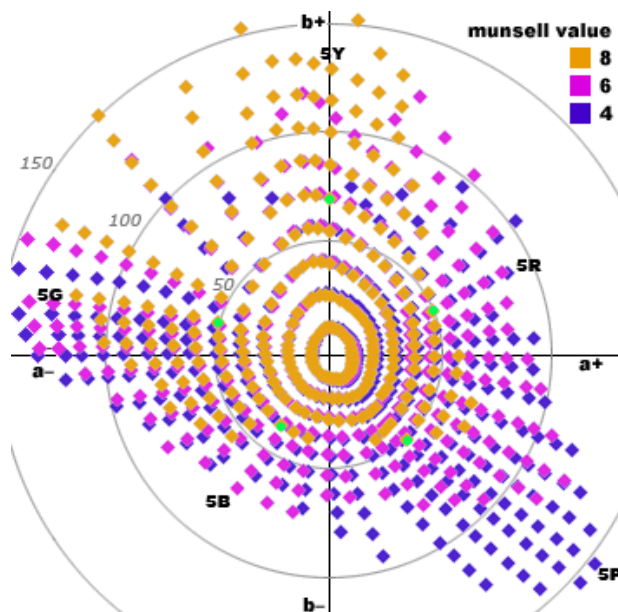


### 700 watercolor paints in the CIELAB color space

based on spectrophotometric measurements reported in the **guide to watercolor pigments**; red (**a+**) placed on left to correspond to artist's color wheels

Note the similarity in the side view of the pigment distribution (at left) and the chroma limits of the yellow/blue color samples in the **Munsell color pages** discussed above. Compare also to the outline of **OSA UCS aim colors** projected into the CIELAB space. (For a "top down" view of pigment locations, see the **CIELAB a\*b\* plane**.)

How closely does CIELAB approximate a *uniform color space*? If we set aside the small problem that a UCS in three dimensions is **impossible**, then CIELAB does remarkably well, given its origins as a hack. A standard basis for comparison — the distribution of Munsell aim colors at Munsell values 4, 6 and 8, out to the limits of surface color chroma — is shown below.



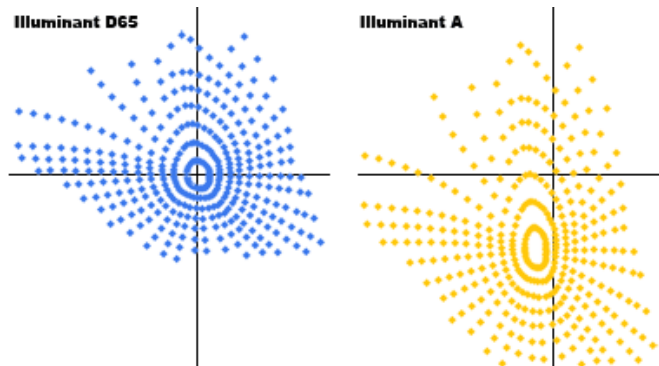
### munsell aim colors on the CIELAB a\*b\* plane

under illuminant D65; aim colors at 2.5 hue intervals and /2 chroma intervals for values 4, 6 and 8 out to the optimal color limits; constant CIELAB chroma indicated by gray circles. Note that all CIE systems reverse the ordering of Munsell hues.

Ideally, this diagram would look like radial spokes within concentric circles. (The "wedge" shaped outline results from the **optimal color** limits — see, for example, **this diagram** — and is not a fault per se.) The exaggerated spacing of Munsell chroma into the yellow and yellow green hues, the displacement in the lines of constant hue as lightness increases (especially in the blue greens), the curving lines of constant violet and green hue, and the wide gaps in the hue spacing of green colors, are all primarily or partly due to irregularities in CIELAB (although

some are actually problems in Munsell). The distribution of **OSA aim colors** makes a much more favorable impression of uniformity. CIELAB should always be used to represent surface colors, and not **CIELUV**.

Because CIELAB includes an adjustment for chromatic adaptation, the location of colors will shift, depending on the chromaticity and luminance specified for the achromatic point, as shown below.



#### corresponding colors in CIELAB

*Munsell aim colors at value  $V = 6$  represented under "daylight" illuminant D65 (with a correlated color temperature of 6500°K, left) or "incandescent" illuminant A (2860°K, right)*

These illuminant shifted color locations or **corresponding colors** are not very accurate. Nearly all the change is produced along the **b** dimension. The overall color predictions — for example, under illuminant A deep blue of chroma /14 will appear gray, and the gray will appear a deep yellow of chroma /8 — are approximately right, though much better results are possible by using a **different adaptation** to the illuminant.

CIELAB can justifiably be called an *opponent color space*, because its chromatic dimensions are defined as the contrast between specific opposing hues, analogous of the opponent processes hypothesized by Hering. In fact, the **X-Y** and **Y-Z contrast dimensions** have a very similar shape and peak wavelengths to the Hurvich and Jameson **opponent functions**. On these grounds, the dimensions are usually interpreted as:

**$L^*$  = white ( $L+$ ) vs. black ( $L=0$ )**  
 **$a^*$  = red ( $a+$ ) vs. green ( $a-$ )**  
 **$b^*$  = yellow ( $b+$ ) vs. blue ( $b-$ )**

This interpretation requires three qualifications. First, the zero end of the lightness dimension  **$L^*$**  only represents the absence of light (surface reflectance), and not black as a distinct color sensation, as conceived by Hering and observed in **lightness induction**.

Second, the connection between the  **$a^*b^*$**  dimensions and hue depends on how the connection is defined: the opponent contrasts do not consistently identify four hues as the contrast anchors:

- The *physiological contrast* defined by the **a** or **b** dimensions is found in the dominant wavelength or peak of the positive and negative sections of the **X-Y** and **Y-Z** curves (**above**), interpreted as cone sensitivity curves or transmittance profiles. These identify the contrast hues as magenta ( **$a+$** , a mixture of orange and violet), middle green ( **$a-$** ), middle yellow ( **$b+$** ) and violet ( **$b-$** ).



- Across the total sample of Munsell surface color exemplars, the **a** and **b** dimensions are correlated around  $-0.35$ . Descriptively, this means that the most intense yellow colors (**b+**) tend to be greenish (**a-**), and the most intense violet colors (**b-**) tend to be reddish (**a+**). Perceptually, it implies that the **y/b** and **r/g** **opponent functions** are not independent of each other — that is, a change on one dimension will have no perceptual effect on the other — although this is how they are normally presented.

- The hue of surface colors located at the ends of perpendicular **a\*** and **b\*** dimensions, as **shown here**, are magenta (**a+**), blue green (**a-**), light yellow (**b+**) and cobalt blue (**b-**).

- The Hering unique hues are displaced from the ends of the perpendicular **a\*b\*** dimensions by a large amount, as **shown here**. Unique red and unique green are both shifted about  $20^\circ$  toward yellow, and the separation between unique red and unique blue is more than one third ( $130^\circ$ ) of the hue circumference.

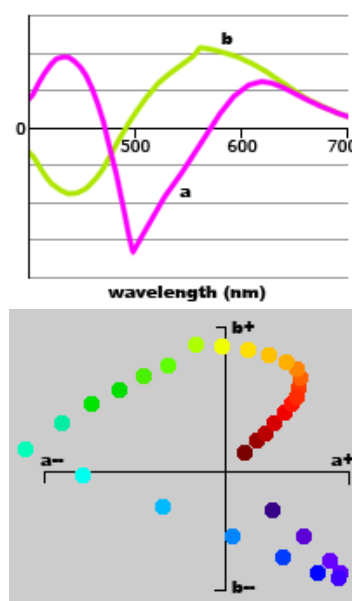
Third, the cube root compression applied to the tristimulus values creates a nasty kink in the **a\*** chromatic dimension (diagram right, top), which significantly departs from the undulating profile of empirically measured opponent functions. This especially stretches out the distance between blue hues with "green" content, and causes the **a+** red to contain too much blue, compressing the distance between yellow and red colors.

Because they are based on difference scores, CIELAB chromaticity dimensions do not redistribute or normalize the luminance information across the chromaticity coordinates: they expunge it. As a result, the CIELAB color metrics cannot be expressed directly as the additive mixture of three "primary" colors of real or imaginary lights. When narrowband optimal colors are represented in this cone opponent space, they produce an erratic, double lobed circuit, with the spectrum ends diving toward the achromatic origin (diagram right, bottom). This is recognizably an upside down image of the **cone excitation** space, pinched and skewed by the cube root transformation. Note also the exaggerated spacing of the blue hues.

CIELAB measures **chroma**, the chromatic content of a color independent of its luminance; the **a\*b\*** plane does not represent saturation. Unlike the straight line, additive mixture between any two colors of light in the CIE UCS (or any chromaticity diagram), color mixtures can take curving or **unpredictable paths** on the CIELAB **a\*b\*** plane. This is not a result of the subtractive mixture of the two physical colorants (paints or dyes) but is basic to the CIELAB space.

**Current Use of CIELAB.** Like CIELUV, color specification in CIELAB is tied entirely to **spectrophotometric measurements** of color; it can't be used routinely without a color measurement device, and there is no standard CIELAB color atlas available. However, the wide use and practicality of the CIELAB system have spawned a number of proprietary color order systems or color atlases based in whole or part on the CIELAB geometry, including the American **Colorcurve System**, the German **RAL Design System** and the English **Eurocolor Atlas**.

Fairchild (2005) reports validation studies that pit several color appearance models against each other, using datasets that measured color matching under illuminant adaptation



**opponent geometry of CIELAB**

(top) opponent functions **a** and **b**;  
(bottom) trace of optimal color stimuli on the **a\*b\*** plane; both displays under equal energy illuminant

(corresponding colors) or preferred contrast and chroma in color image reproductions viewed at different luminance levels. Overall, these studies showed that different models performed well in predicting some types of color judgments but not others. In every case, however, CIELAB typically performed well against more complex models and occasionally was as good as the best; CIELUV was often the worst.

A major problem is that CIELAB uses what is called a **wrong von Kries transform** to model chromatic adaptation. This computes the contrast ratio between a color and the white standard or **illuminant** directly on the **XYZ** tristimulus values, rather than on cone fundamentals. In addition, the cube root compression is not optimal when applied to the tristimulus values, and results in chroma intervals that can be misplaced in some parts of the hue plane: in particular, yellow chroma is not compressed enough, and blue chroma is compressed too much.

These failings aside, CIELAB has become a standard colorimetric space and is one of the most practical and widely implemented color models available. Indeed, several extensions of or revisions to the basic CIELAB framework, including **S-CIELAB** and **RLAB**, have been published to improve its performance and suitability for a wider range of color modeling problems. These extensions are possible because the CIELAB geometry does not impose an arbitrary symmetry on chroma, hue or lightness. The calculation from **XYZ** values to **Lab** values is simple and easily inverted (from **Lab** values back into **XYZ**), and can be replaced by a variety of chromatic adaptation transforms.

The CIELAB color difference formula is widely used for automated color quality control in areas such as textiles, plastics, architectural and automotive coatings, printing, imaging and art materials, and in ASTM documents. The measured XYZ values of color samples from any other color model, including Munsell and NCS, can be located and compared in the CIELAB space, so it has become a convenient benchmark space for the evaluation of color models in general.

CIELAB is a useful color framework for painters who want to understand color perception and color mixing problems in paints. In practical applications, it is a useful framework for linking digital imagery to color printing technology. When working with **Munsell Color System** coordinates, 10 units on the CIELAB lightness dimension **L\*** represent 1 unit on the Munsell value **V**, and 10 units in CIELAB chroma **C<sub>ab</sub>** represent, on average, 2 units of Munsell chroma **C**.

In 1996 I used CIELAB pigment locations as the foundation of my original **artist's color wheel**. However some of the "bad" hue and chroma spacing problems, which I had corrected by hand using the Munsell renotation data, later decided me in favor of the CIECAM02 color relationships for the **current (2006) version** of the color wheel.

My principal source for information about CIELAB and CIECAM is Mark Fairchild's **Color Appearance Models (2nd ed.)** (Addison Wesley, 2005); see also the updates and information at Fairchild's **web site**.

Artists may be interested in the average location of watercolor pigments on the **CIELAB a\*b\* plane**, also available as a **PDF file**.

The **Adobe Technical Guides** have a good overview of color models in general, the CIE series, and the groundbreaking **Munsell color system**.

Daniel Smith has republished their Watercolor Paint Guide, which includes a map of their watercolor paint colors on the CIELAB  $a^*b^*$  plane, as "The Study of Color", available online at their *InkSpot* archive and reprinted occasionally in their catalog.

## CIECAM color appearance model ▲

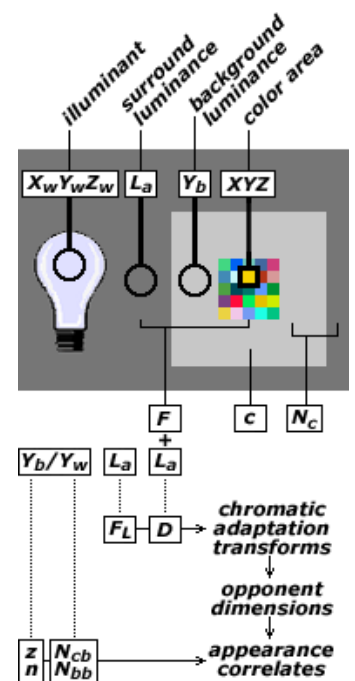
Because they are limited to the trichromatic outputs from the **L**, **M** and **S** cones, both CIELUV and CIELAB only model the color information generated by the photoreceptor response to **small and isolated** color samples. They cannot model many of the contextual effects of color perception, including **luminance adaptation** (surround illuminance levels), **chromatic contrast** or **chromatic assimilation** that arise in real world viewing conditions with heterogeneous, strongly contrasted, or three dimensional color stimuli.

In the three decades since CIELAB was published, several models have been proposed to represent these more complex color phenomena. The most recent is **CIECAM02** (CIE Color Appearance Model 2002, based on CIECAM97s), developed through the research and committee products of Robert Hunt, Yoshinobu Nayatani, Mark Fairchild, Nathan Moroney and others. CIECAM represents a culmination and also an end point in the development of color appearance models.

Detailed instructions for computing and using the latest CIE models are available in the texts by Fairchild and by Hunt (cited below). My goal is to explain how the input information and calculations affect the final color space. All diagrams were generated from Fairchild's spreadsheet version of the CIECAM formulas, cited below.

**The CIECAM02 Dimensions.** The CIECAM calculations are quite complex when compared to **CIELAB**. CIECAM also requires more starting information (refer to graphic at right):

- Colorimetric information: the **XYZ** tristimulus values for the color area being modeled, the **X<sub>w</sub>Y<sub>w</sub>Z<sub>w</sub>** tristimulus values for the white standard or illuminant, and the adaptation luminance **Y<sub>b</sub>** of the adjacent color *background*, equivalent to the standardized (**XYZ**) luminance (where **Y** = 2 for a Munsell value of 1 [near black], **Y** = 20 for a value of 5 [middle gray], and **Y** = 90 for the Munsell white).
- Photometric information: the average **luminance** of the visual environment or *surround* **L<sub>a</sub>**, in nits (cd/m<sup>2</sup>).
- Context parameters: (1) contrast factors **F** and **N<sub>c</sub>**, which both have values of 1.0 when the luminance (reflected illuminance) in the surround matches the illuminance on the color area (documents or prints under workplace lighting), 0.9 when the surround illuminance is dimmed compared to the image (computer displays or overhead transparencies), and 0.8 when the surround illuminance is dark compared to the luminance of the image (cinema, slide or video images); and (2) an exponent **c** that modulates the response compression in image lightness, brightness and chroma caused by background lightness (the **Bartleson Breneman effect**), empirically set at values of 0.69, 0.59 or 0.525 for the same viewing contexts. These parameters are not measured in the scene, but are determined by visual judgment and then read from a graph.



graphical summary of CIECAM input values and parameters

- **Response Compression Factors:** (1) a luminance adaptation factor **D**, which increases toward a limit value of 1 as the color area appears reflective rather than emitting (that is, the surround luminance **L<sub>a</sub>** increases, and/or **F** increases — the luminance of the surround matches the luminance of the image); (2) a power function **F<sub>L</sub>** that models the increase in brightness and colorfulness caused by brighter viewing environments (higher levels of surround luminance **L<sub>a</sub>**); (3) chromatic contrast factors **N<sub>bb</sub>** and **N<sub>cb</sub>**, derived from the ratio **Y<sub>b</sub>/Y<sub>w</sub>**, that are used to calculate the increase in chroma, lightness and brightness caused by darker backgrounds; and (4) exponents **z** and **n**, also derived from **Y<sub>b</sub>/Y<sub>w</sub>**, that define the response compression on lightness and chroma caused by background lightness.

Note that the color area may be part of a larger image, but all adaptation or contrast factors operate on the color area in isolation. CIECAM only represents shifts due to lightness contrast with the background or the effects of illuminance levels or luminance adaptation. It does not model chromatic contrast, chromatic assimilation or effects of spatial frequency between components of an image.

**Chromatic Adaptation.** Much of the development work leading up to CIECAM involved the choice of chromatic adaptation calculations that gave accurate and invertible results.

The equal area **XYZ tristimulus values** remain the standard color specification only by virtue of their broad implementation. They give bad results if used directly to calculate chromatic adaptation (as is done in CIELAB), and equal area **cone fundamentals** must be used instead. In addition, different response functions must be used to model the effects of chromatic adaptation and luminance adaptation. These tasks are combined through an intricate sequence of hacks:

- The tristimulus values of the color area (**X<sub>c</sub>Y<sub>c</sub>Z<sub>c</sub>**) and the illuminant (**X<sub>w</sub>Y<sub>w</sub>Z<sub>w</sub>**) are converted into "sharpened" color matching functions (which include negative values), resulting in *preadaptation values* for the color area (**R<sub>G</sub>B**) and the illuminant (**R<sub>w</sub>G<sub>w</sub>B<sub>w</sub>**), as shown in the **diagram below**. These equalize the relative contribution of the **R** and **G** (**X** and **Y**) values and create negative (supersaturated) values in the **G** curve.
- Adjustments for luminance adaptation (**D**) are made on the color and illuminant **RGB** values, resulting in *post adaptation values* for the color area (**R<sub>c</sub>G<sub>c</sub>B<sub>c</sub>**) and the illuminant. This is the point where the illuminant tristimulus values are used to perform a *von Kries transform* for **chromatic adaptation**, shown below for the **R** value:

$$R_c = [(100D/R_w)+1-D]R$$

The same adjustment is performed on the **G<sub>c</sub>** and **B<sub>c</sub>** values. This step corresponds to the **adaptation ratios** in CIELAB, which are done on the tristimulus values directly and are therefore *wrong von Kries transforms*. The CIECAM formula has zero effect (**R<sub>c</sub> = R**) in cases where (1) colors are viewed under an **equal energy illuminant** (the **R<sub>w</sub>G<sub>w</sub>B<sub>w</sub>** values are all 100), and (2) there is complete adaptation to the illuminant (**D = 1**). Otherwise this transform decreases or increases the cone output to match the chromaticity of the illuminant, in proportion to the

degree that adaptation only partly occurs — that is,  $D < 1$  when the surround is darker than the image, and the difference between image and surround luminance prevents complete adaptation. This becomes especially significant when the surround luminance is below  $300 \text{ cd/m}^2$ .

- The post adaptation  $R_c G_c B_c$  values are converted back into **XYZ** tristimulus values. Parallel transformations are done to retrieve the illuminant **XYZ** values.
- The post adaptation **XYZ** values are converted into a specific type of equal area **cone fundamentals** (Hunt Pointer Estevez fundamentals), denoted  $R'G'B'$ .
- Finally, a response compression is applied to the  $R'G'B'$  cone fundamentals. This compression is approximately hyperbolic with a minimum value near zero and an upward slope that changes in relation to the level of surround luminance  $L_a$ , expressed as the response compression factor  $F_L$  (right). The calculation is shown below for the  $R'$  value:

$$R'_a = \frac{400(F_L R'/100)^{0.42}}{27.13 + (F_L R'/100)^{0.42}} + 0.1$$

The same adjustment is made to the illuminant  $R'_w G'_w B'_w$  values. This transformation has the double effect of amplifying cone response differences (contrast) at very low tristimulus values (the curve has a much steeper slope at low  $R'$  values), and increasing the overall response contrast in brightness and chroma as luminance increases (the maximum value of  $R'_a$  and the overall slope increases). Because wavelengths at the tails of a cone fundamental produce a much lower response (tristimulus value) than wavelengths near the peak, and the transformation increases relative response at low tristimulus values, the transformed cone fundamentals have substantially raised tails and broad, rounded shoulders around the peaks (see **diagram below**). These calculations yield post adaptation, response compressed cone fundamentals denoted  $R'_a G'_a B'_a$  and  $R'_{aw} G'_{aw} B'_{aw}$ .

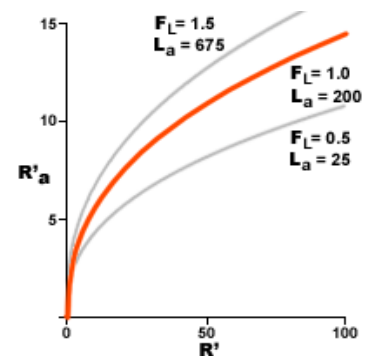
These circuitous calculations result in a "correct" **von Kries transform** for chromatic adaptation, adjusted in relation to the degree of luminance adaptation, and a response compression scaled to surround luminance (luminance adaptation).

**Lightness/Brightness.** Correlates of lightness (**J**) and brightness (**Q**) are computed from the post adaptation, response compressed cone fundamentals  $R'_a G'_a B'_a$ , the background contrast factor  $N_{bb}$ , and the surround luminance compression factor  $F_L$ , as follows:

$$\begin{aligned} A &= [2.0R'_a + 1.0G'_a + 0.05B'_a - 0.305]N_{bb} \\ J &= 100(A/A_w)^{c_z} \\ Q &= (4/c) \sqrt{J/100} (A_w + 4) F_L^{0.25} \end{aligned}$$

$A_w$  is computed from the illuminant  $R'_{aw} G'_{aw} B'_{aw}$  values. **A** is essentially a luminosity function with rounded peak and raised tails that is adjusted by the contrast factors **c**, **z** and  $F_L$  to represent compression effects caused by background contrast and surround luminance.

**Opponent Dimensions.** A set of *preliminary* opponent dimensions are calculated as:



**luminance dependent response compression from  $R'$  to  $R'_a$**

$$\begin{aligned} \mathbf{a} &= [\mathbf{R}'_a + (\mathbf{B}'_a/11)] - (12\mathbf{G}'_a/11) \\ \mathbf{b} &= (\mathbf{R}'_a + \mathbf{G}'_a - 2\mathbf{B}'_a)/9 \end{aligned}$$

Note the similarity between these opponent dimensions and the fundamental dimensions of a *cone excitation diagram*, for example the dimensions **L–M** and **L+M–S** of a trilinear **mixing triangle**. The main differences are in the small amount of **S** (**B'**<sub>a</sub>) output added to the **L** (**R'**<sub>a</sub>) cone output in the **a** contrast, and in the better relative weighting of the cone outputs.

These preliminary **ab** dimensions are used to derive the color's hue angle (**h**), which in turn is used to derive an *eccentricity factor* **e** that adjusts the scaling of the **a** and **b** dimensions to represent the differences in chroma compression that occur around the hue circle.

**Hue Purity Correlates.** Next, chroma (**C**) is computed as the euclidean distance from the origin on the response compressed **ab** dimensions. I present the formula to show how intricate it is:

$$\mathbf{C} = \left[ \frac{(50000/13)\mathbf{N}_{bb}\mathbf{N}_{bc}\mathbf{e}\sqrt{\mathbf{a}^2+\mathbf{b}^2}}{\mathbf{R}'_a+\mathbf{G}'_a+(21/20)\mathbf{B}'_a} \right]^{0.9} \sqrt{\mathbf{J}/100(1.64-0.29^n)^{0.73}}$$

The formula can be parsed as follows: (1) the basic chromaticity distance is defined as the euclidean distance on **a** and **b** between the color and the achromatic point; (2) the chromaticity is multiplied by the eccentricity factor **e**, which produces a different chroma scaling along each hue angle of the **ab** chromaticity plane; (3) the rescaled chroma is multiplied by the chromatic induction factors **N<sub>bb</sub>** and **N<sub>cb</sub>**, which increase chroma as the background becomes darker; (4) the contrast scaled chroma values are **normalized** on the sum of response compressed cone outputs (**R'**<sub>a</sub>**G'**<sub>a</sub>**B'**<sub>a</sub>); and (5) the normalized chroma values are adjusted to compensate for the color's lightness (**J**) and the background contrast exponent (**n**).

Modern chroma metrics in CIECAM and other color models are highly complex, but this is necessary to achieve equal chroma spacing around the hue circle, appropriately normalize the luminance information across an equal chromaticity plane, and fit the correct response compression in relation to background contrast, scene luminance and color lightness. The **cylindrical coordinates** of lightness/brightness, hue angle and chroma are the final form of the CIECAM system.

The cylindrical coordinates are used to compute the remaining appearance attributes saturation (**s**) and colorfulness (**M**). Trigonometric functions are then used to transform the color space into cartesian coordinates **Jab**, whose chromaticity units are in terms of chroma, saturation or colorfulness.

Finally, there is no color difference formula optimized for the CIECAM model: the sequence of color models that began in color difference formulas ended in a model without one. However, very good results have been obtained by using the standard euclidean distance between *similar* colors:

$$\Delta\mathbf{E}_{02} = [(\mathbf{L}_1-\mathbf{L}_2)^2+(\mathbf{a}_{c1}-\mathbf{a}_{c2})^2+(\mathbf{b}_{c1}-\mathbf{b}_{c2})^2]^{1/2}$$

where "similar" means colors that are separated by a **ΔE<sub>02</sub>** value of around 10 or less. The average CIECAM **ΔE<sub>02</sub>** for the first chroma step in Munsell aim colors across all hues and values is about 7.7, slightly larger than in CIELAB. Note that the



$\Delta E_{02}$  is approximately in units of just noticeable difference, so color differences at or below 1 are generally not visible.

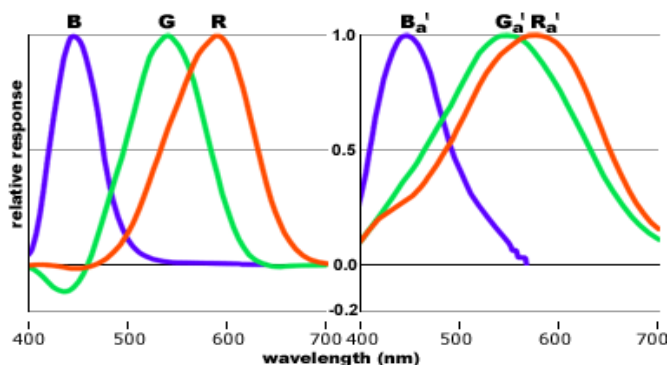
**Graphical Analysis of CIECAM.** It is hardly possible to understand a color model as complex as CIECAM from the calculation stream alone. The effect of the calculations (and parameter values) on color specifications must always be examined graphically.

*Response Compression.* Probably the single most significant feature of CIECAM in comparison to CIELAB and all previous color models is the pervasive use of **exponential functions** to model response compression. CIECAM contains four variables — **c**, **z**, **F<sub>L</sub>** and **D** — used exclusively as exponents or as nonlinear scaling factors. A graphical example of the compressive, "bending" effect of these power transforms is shown in the explanation of the **R'<sub>a</sub>G'<sub>a</sub>B'<sub>a</sub>** cone fundamentals, above.

Response compression is necessary both to model the basic **psychophysical functions** of lightness, brightness, colorfulness, chroma, saturation and hue, and to model the nonlinear changes in these functions caused by changes in the color context (background contrast, surround luminance, chromatic and luminance adaptation). Overall, including the use of trigonometric functions to compute hue angle **h** and the eccentricity factor **e**, nonlinear transforms are used in fourteen of the 22 CIECAM calculation steps. In this regard CIECAM testifies to the **importance of response compression** in all aspects of color perception.

*Chromatic Adaptation Transforms.* All the CIECAM calculations are made with **equal area cone fundamentals**, which correspond to the equal area **XYZ color matching functions**. However they are typically shown in color vision texts as **normalized cone fundamentals** to facilitate visual comparison of the curve shapes, especially in the tails where the adaptation adjustments have the greatest effect.

The diagram shows the pre adaptation, "sharpened" color matching functions **RGB** and the post adaptation, response compressed **R'<sub>a</sub>G'<sub>a</sub>B'<sub>a</sub>** cone fundamentals.



#### chromatic adaptation functions in CIECAM

under illuminant D65; (left) CAT02 **RGB** system of pre adaptation, "sharpened" color matching functions; (right) post adaptation, response compressed cone fundamentals **R'<sub>a</sub>G'<sub>a</sub>B'<sub>a</sub>**

The key features of each set of curves are visually obvious when compared both to the **XYZ** color matching functions and to normalized **LMS** cone fundamentals, linked above.

The **RGB** color matching set introduces negative values in the **R** and **G** curves, which makes them resemble the original **RGB**

**color matching functions**, except that the "out of gamut" hue is not blue green but blue violet. This defines the **B** "primary" as supersaturated and increases the chromaticity distance between the **B** and **G** "primaries".

The **R'<sub>a</sub>G'<sub>a</sub>B'<sub>a</sub>** cone fundamentals resemble the **logarithmic cone fundamentals**, because these curves were **transformed** using a power function. In fact the curves have a roughly triangular shape that opens up hue differences in the tails and significantly improves on the cube root transform used in CIELAB. The separation between **R'<sub>a</sub>** and **G'<sub>a</sub>** is also reduced so far that hue discrimination (spacing) is more specifically affected by changes in **R'<sub>a</sub>** outputs for hues at wavelengths above 580 nm, and by changes in **B'<sub>a</sub>** outputs for all hues below 580 nm.

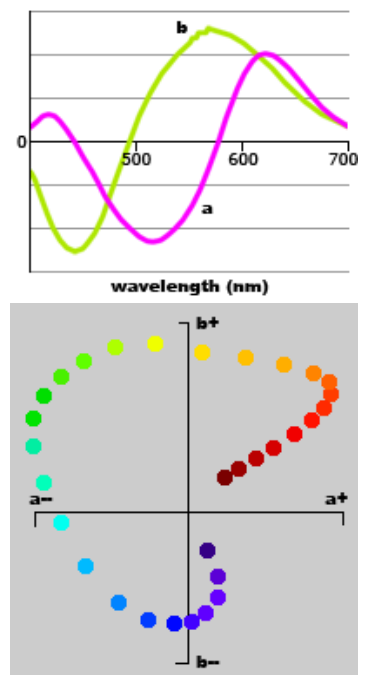
**Opponent Geometry.** The final chroma based (luminance independent) opponent dimensions in CIECAM, **a<sub>c</sub>** and **b<sub>c</sub>** (right, top) closely resemble the Hurvich and Jameson **opponent functions** and the **X–Y** and **Y–Z** dimensions of CIELAB *before* the cube root compression is applied (**above**).

The more intricate chromaticity calculations in CIECAM remove the kink found around 500 nm in the CIELAB **a\*** dimension (see **above**), and produce a less angular, more evenly spaced involute of optimal colors, especially from cyan to violet (right, bottom). This is one of the most important differences between CIELAB and CIECAM. The correlation between the dimensions is also substantially reduced, to  $-0.14$  from the  $-0.35$  in CIELAB — a deviation from perpendicular of about  $8^\circ$ .

The historical XYZ tristimulus values produce some sawtoothing in the **b** curve across yellow green hues, the result of amplifying (via the chromatic adaptation transformations) some coarsely interpolated **Z** values where the **S** cone response is near zero. The sawtoothing is amplified to a significant degree under "yellow" illuminants (such as illuminant A), and will introduce discontinuities in the placement of saturated yellow greens. A new **Z** function derived from the recent **Stockman & Sharpe** estimates of **S** cone sensitivity would eliminate the problem.

Chromaticity differences scaled on chroma (**a<sub>c</sub>** and **b<sub>c</sub>**) are independent of changes in surround luminance **L<sub>a</sub>** or background contrast (**Y<sub>b</sub>/Y<sub>w</sub>**). The dimensions scaled on colorfulness, **a<sub>M</sub>** and **b<sub>M</sub>**, expand as scene luminance **L<sub>a</sub>** increases, to model the **Hunt effect**, and as brightness induction increases (low values of **Y<sub>b</sub>**). CIECAM also provides dimensions scaled on saturation, **a<sub>s</sub>** and **b<sub>s</sub>**, which *contract* under increases in scene luminance or brightness induction; the saturation metric also produces a large "donut hole" gap between the achromatic center and the first Munsell value step, consistent with the **crispening effect**.

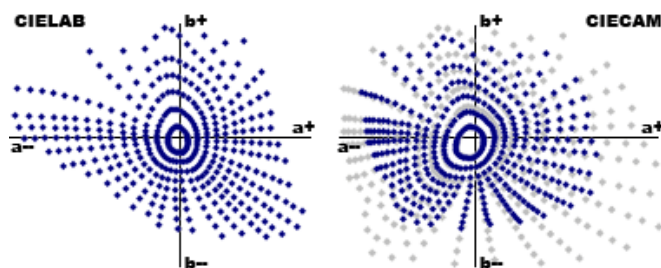
**Chroma Scaling.** As I explain in the discussion of **hue purity**, measures of hue purity (chroma or saturation) have been the most troublesome, imprecise and complex features of modern color models. No exception, the most intricate aspects of CIECAM concern the scaling of colorfulness, chroma and saturation.



**opponent geometry of CIECAM**

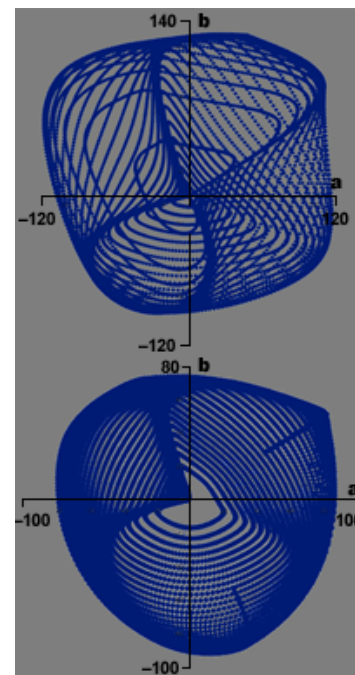
(top) opponent functions **a<sub>c</sub>** and **b<sub>c</sub>** based on chroma scaling; (bottom) trace of monochromatic color stimuli on the **a<sub>c</sub>b<sub>c</sub>** plane; both displays under equal energy illuminant

The diagram below shows the final chroma scaling in CIECAM in comparison to the preliminary CIECAM dimensions and to the chroma scaling in CIELAB. The two diagrams (right) show the distribution of **optimal colors** on the CIECAM  $a_c b_c$  (chroma) and  $a_s b_s$  (saturation) planes. CIECAM achieves good consistency in the scaling of both quantities, and the resulting roughly circular volumes (especially in the saturation metric) indicate that CIECAM chroma and saturation are approximately *relative* measure of hue purity, in which optimal colors provide the **perceptual standard** for maximum chromatic intensity.



**chroma scaling in CIECAM02**

(left) Munsell aim colors for  $V = 5$  in CIELAB; (right) the Munsell aim colors in CIECAM02, showing the chromaticity locations on the preliminary opponent dimensions  $ab$  (gray) and on the eccentricity corrected dimensions (blue); because the dimensions before and after adjustment are measured on different scales, the two views are standardized to have equal chroma at  $C = 8$ .



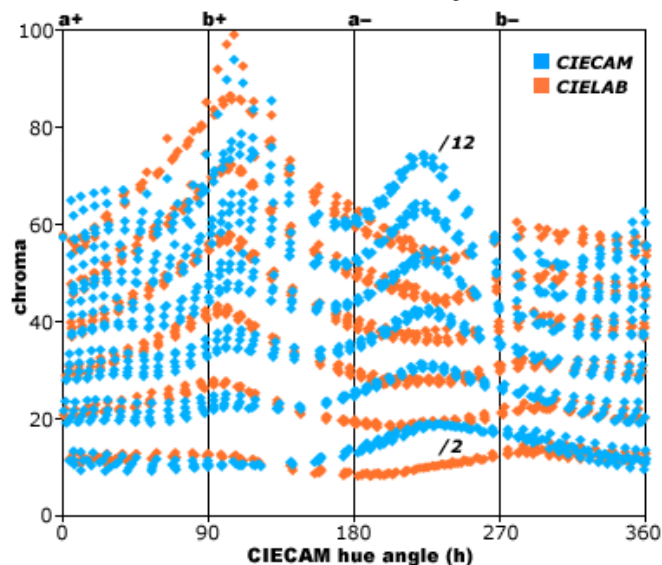
**optimal color boundaries on the CIECAM  $ab$  plane**

(top) chroma metric, (bottom) saturation metric; under EE illuminant, 20% background reflectance, white surface luminance 318 cd/m<sup>2</sup>

In comparison to CIELAB, the final CIECAM chroma values produce a very large contraction in all highly saturated colors, including colors near the optimal color limits that are not usually found in material colors, and an expansion in the chroma of near neutral or desaturated colors. In comparison to the preliminary  $ab$  dimensions, the final CIECAM chroma values produce the largest relative contraction in colors from middle blue through magenta to orange.

During work on the **OSA uniform color scales** in the 1950's it was found that the chromaticities of any opponent dimensions color model must be **scaled differently** for each hue quadrant of the color space. In CIECAM this is achieved (not very satisfactorily) through the eccentricity factor  $e$ , which shifts a circle of constant chroma by about 0.2 toward middle blue (hue angle 245). This most strongly contracts the chroma of deep yellow hues and expands the chroma of blue hues, with no effect on the chroma of greens and purples.

This eccentricity scaling is easiest to see by comparing the CIELAB and CIECAM chroma of dull to moderately saturated Munsell aim colors, as shown below. The increase in blue chroma at around hue angle 245, and the depression of chroma at around hue angle 65 are clearly visible. The adjustments also introduce more variation in chroma values across different lightness levels in yellow, red and purple hues.

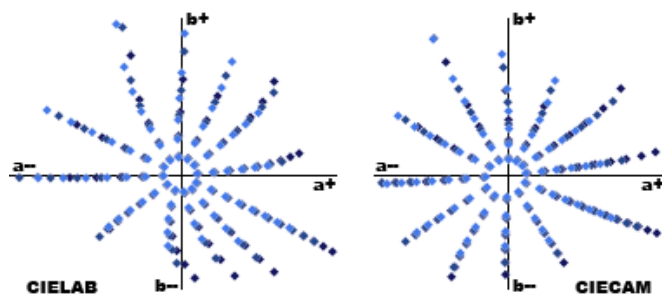


### chroma adjustments by hue angle in CIECAM02

measured chroma of Munsell aim colors for value  $V = 4$  to 8 and chroma  $C = 2$  to 12, in CIELAB and CIECAM02; illuminant D65

Note that, in a color space based on perfectly accurate color samples and a "true" measure of perceived chroma, the lines of points would be perfectly straight and horizontal.

Hue Scaling. Chroma spacing also significantly affects hue spacing. The CIECAM calculations produce a much better spacing and straightness in lines of constant hue. The diagram below shows the location of radial colors, defined in the OSA color space to avoid the flaws in the spacing of Munsell hues, in CIELAB and CIECAM. The CIELAB lines show unequal spacing around the hue circle and strongly curving hue lines in purple hues. The CIECAM values lines are almost exactly equally spaced, the lines of constant hue are much more linear, and the overall distribution is much closer to circular.



### lines of constant hue in CIELAB and CIECAM02

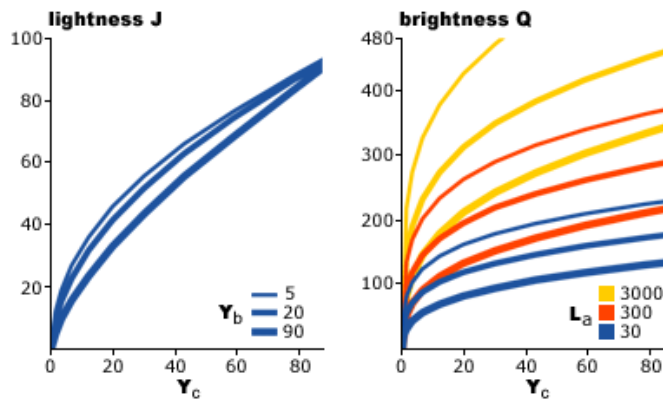
radial sample of OSA UCS color space for three lightness values  $L = -2.3, -0.5$  and  $1.8$ , under illuminant D65 (aim color data from Moroney, 2003)

Brightness/Lightness. The effects of surround luminance and background contrast on both lightness and brightness are fairly straightforward.

First, consider situations in which the scene parameters  $N_c$ ,  $F$  and  $c$  are held constant at "normal" values that represent the viewing of papers or prints under the same illumination as the surround ( $L_a$ ).

In this basic context the environmental illumination can be varied from dim office lighting of about 100 lux ( $L_a = 30$ ) to afternoon daylight illumination of about 10000 lux ( $L_a = 3000$ ), and the documents or prints can be viewed against a background that is either dark gray ( $Y_b = 5$ ), middle gray ( $Y_b =$

20) or white ( $Y_b = 90$ ). The diagram below shows the effect of these variations on lightness ( $J$ ) and brightness ( $Q$ ).



### lightness and brightness in CIECAM02

(left) lightness values  $J$  at three levels of background standardized luminance,  $Y_b$ ; (right) brightness values  $Q$  at three levels of surround luminance  $L_a$  and three levels of background standardized luminance; averaged values across 2800 Munsell colors, equal energy illuminant.

The curves for lightness  $J$  (diagram, left) show the effect of background contrast on the spacing of a gray scale. A white background minimizes the response compression and produces a relatively linear spacing gray values in relation to the color's luminance ( $Y_c$ ). A dark gray background increases the overall response compression, producing more contrast among dark valued colors and less contrast among light valued colors. This is the **Bartleson-Breneman effect**. These changes in the lightness curve are identical across all levels of surround luminance ( $L_a$ ), because relative lightness remains constant across the normal range of illumination.

The curves for brightness  $Q$  (diagram above, right) show the expected effect of increasing surround luminance: as  $L_a$  goes up, so too does the brightness of all colors and the perceived amount of contrast between dark and light colors (the **Stevens effect**). However, background contrast also increases apparent brightness through brightness induction — dark backgrounds make colors appear more luminous (brighter and more colorful). And this "color amplifying" effect of darker backgrounds becomes greater at higher levels of surround luminance.

In a similar way, changes in surround luminance  $L_a$  have no effect on chroma but increase relative colorfulness. Changes in background contrast  $Y_b$  have a small effect on chroma: in the range shown, from 5 to 90, the chroma increases by about 1.5 Munsell steps (as shown below), and produce an even smaller increase in colorfulness.



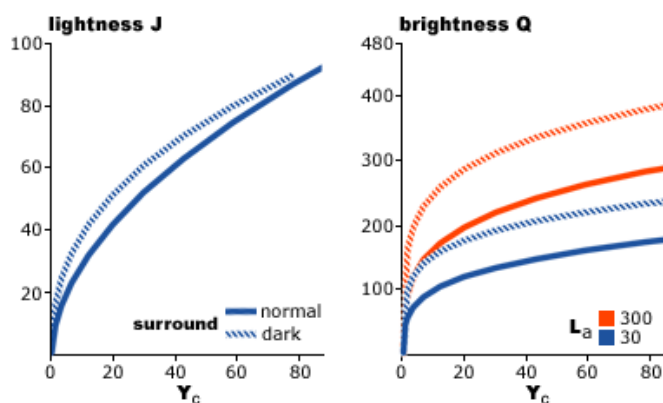


### chroma and background contrast in CIECAM02

(left) Munsell aim colors for  $V = 5$  in CIELAB; (right) the Munsell aim colors in CIECAM02, showing the chromaticity locations on the initial opponent dimensions (gray) and on the eccentricity corrected dimensions (blue); because the dimensions before and after adjustment are measured on different scales, the two views are standardized to have equal chroma at  $C = 8$ .

Surround Contrast. The effect of changing the surround contrast parameters  $c$ ,  $N_c$  and  $F$ , which model the difference in color appearance between reflective documents viewed in bright environments and projected or emitting images viewed in dark environments, is small and uniform within a realistic range of values for the surround luminance ( $L_a$ ).

Changing the luminance value of the surround,  $L_a$  by itself has no effect on the chroma or hue of colors. However it does affect both lightness and brightness and, through brightness, relative colorfulness as well.



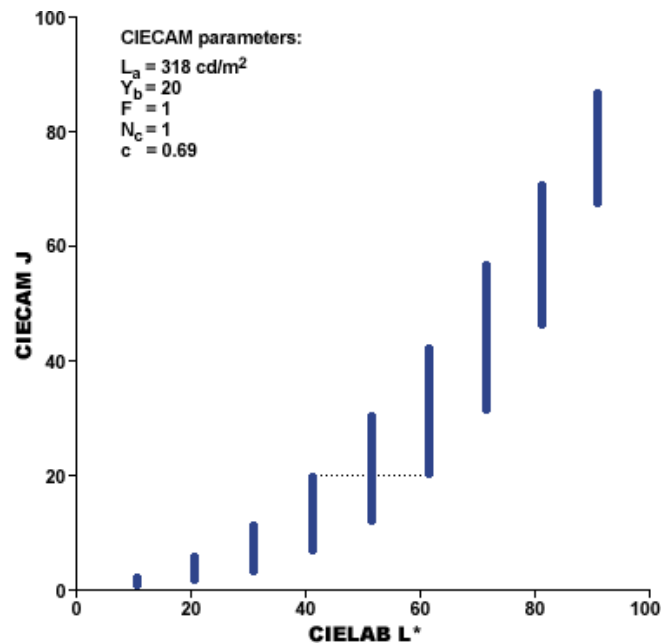
### context and lightness/brightness in CIECAM02

(left) lightness values  $J$  under normal scene contrast ( $N_c = 1.0$ ,  $F = 1.0$ ,  $c = 0.69$ ) and dark scene contrast ( $N_c = 0.8$ ,  $F = 0.8$ ,  $c = 0.525$ ); (right) brightness values  $Q$  at two levels of surround luminance  $L_a$  and two levels of surround contrast; averaged values across 2800 Munsell colors, equal energy illuminant.

The diagram for lightness (left) shows that dark viewing environments decrease somewhat the contrast in light values and increase the contrast in very dark values. Increasing the scene luminance improves the discrimination of light values and reduces contrast in dark values.



The effects on brightness (and by extension on colorfulness) are similar but much larger. Viewing a video display in a dark environment ( $L_a = 30$ ) produces greater brightness than the same display in a well lit environment ( $L_a = 300$ ), but this is offset by the *shape* of the brightness function, which becomes flatter and more sharply bent as surround luminance gets darker, a compression that can be only partly offset by increasing the contrast (gamma) of the display. In fact, an optimal video image appearance usually requires only a slightly dimmed room illumination, provided that the lights do not reflect off the television screen. The newest liquid crystal and plasma displays provide excellent contrast at near normal levels of indoor illumination.



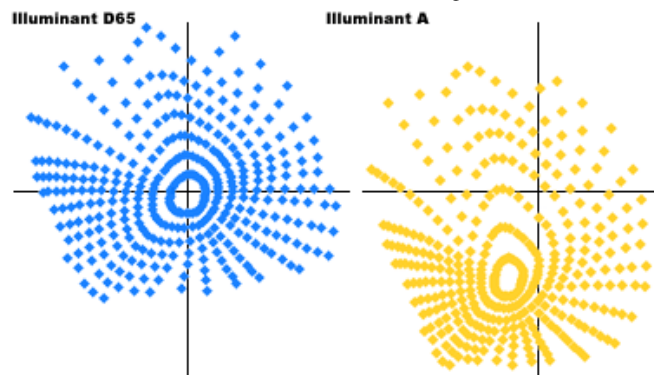
#### comparison of CIELAB $L^*$ with CIECAM $J$

for Munsell aim colors under D65 and CIECAM parameters as shown

Because CIECAM takes into account both the effects of chroma (saturation) and the effects of the display conditions on apparent brightness/lightness, the relationship between CIECAM lightness ( $J$ ) and CIELAB lightness ( $L^*$ ) is complex. The diagram (above) shows that the same CIECAM lightness ( $J = 20$ ) may describe a CIELAB lightness of anywhere from 40 to 60, and that  $J$  is nearly always lower than  $L^*$  for the display parameters shown. Also, CIECAM lightness contrast increases in lighter values, producing a curved function on CIELAB  $L^*$ .

Unfortunately it is not easy to use CIECAM "in reverse" to calculate the best balance between image luminance and surround luminance, because the image luminance values are standardized (limited to a maximum value of 100) by the tristimulus values. Despite the worked examples in the CIE technical reports and current color texts, the luminance assessments that are essential to the CIECAM calculations remain either a judgmental art or anchored in "standard" presentation environments.

**Illuminant Adaptation.** Finally, like both CIELUV and CIELAB, CIECAM can shift the achromatic point of the color space to match changes in the illuminant, which allows prediction of **corresponding colors** under a new illuminant.



### corresponding colors in CIECAM02

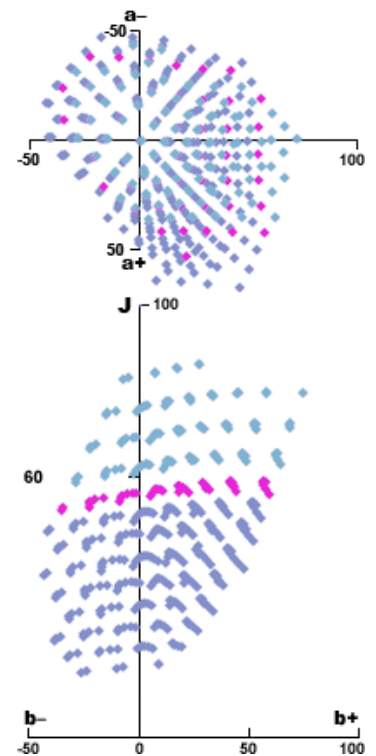
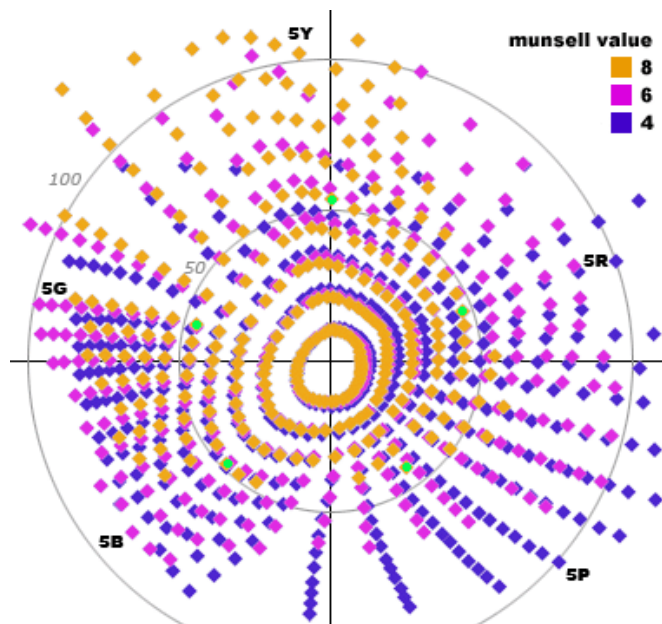
Munsell aim colors at value  $V = 6$  represented under illuminants with a correlated color temperature of 6500°K (daylight D65, left) or 2860°K (incandescent A, right)

The shifts shown here indicate that the Munsell neutral color sample will appear to be a moderately saturated blue in daylight, and a Munsell orange at chroma 8 will appear closest to a neutral gray. Note the improved spacing of colors as the illuminant changes, compared to the stretched out corresponding colors predicted by CIELAB.

**The CIECAM Color Space.** Is CIECAM02 worth the trouble? All available research suggests it is, but much depends on how the evaluation is made.

A comparison of the locations of the Optical Society of America's Uniform Color Scales (OSA-UCS) in CIECAM02 (right) and in CIELAB (shown here) shows many small differences, particularly in the chroma spacing of colors around the hue plane, the downward shift in lightness values, and the more equal chroma spacing across different lightness levels. Cumulatively these amount to significant alterations in the color appearance predictions.

A much sharper contrast is visible in the placement of Munsell aim colors in CIECAM (below), compared to CIELAB (shown here). The strongly curved lines of constant hue in blue violet and purple hues are completely remedied, and there is generally a more stable alignment of hue angles across different lightness levels. The gap in the spacing of green colors (in the upper left quadrant) is a fault in Munsell, not in CIECAM, but I am unsure about the similar gap that appears in blue violet ( $b-$ ) hues.



### OSA UCS aim colors in the CIECAM02 color space

under illuminant D65; layer OSA  
 $L = 1$  shown in magenta;  
 the  $g$  dimension is reversed in  
 all CIE color models

## munsell aim colors on the CIECAM02 $a_c b_c$ plane

*under illuminant D65; aim colors at 2.5 hue intervals and  $/2$  chroma intervals for values 4, 6 and 8 out to the optimal color limits; constant CIECAM chroma indicated by gray circles. Note that all CIE systems reverse the ordering of Munsell hues.*

A specific quirk I uncovered in my technical review is that CIECAM chroma spacing is strongly compressed for colors approaching the optimal limits, which is visible in dark valued violet and blue green hues (above). In contrast, CIELAB maintains a perfectly regular spacing out to the chroma limits. It may be that the optimal limits, which are materially not possible in surface colors without a contrived boost in luminance contrast, are simply outside normal visual experience and therefore irrelevant to the correct spacing of perceptual attributes.

**Future Prospects.** As CIECAM is only a few years old, there is currently no spectrophotometric implementation and no software available, apart from spreadsheets and MATLAB algorithms.

Unlike CIELAB, which can be defined completely and automatically with a spectrophotometer, CIECAM requires the manual (judgmental) specification of parameters for surround luminance, background contrast, and viewing context. These can be set to default values and CIECAM used in the same way as CIELAB, but the potential of the model depends on automating these values.

Fairchild (2005) does not include tests of CIECAM02 per se, but the many development studies of CIECAM, and of related models such as the one developed by Hunt and Luo, indicate that there is no further improvement possible within the current conception of the viewing context. Looking ahead, Fairchild strikes a dour note:

*It appears that the time between CIECAM02 and the next CIE color appearance model will be significantly longer than six years. One reason for this is that this type of model seems to be predicting the available visual data to within experimental uncertainty. ... The cost and difficulty of collecting such data as well as inherent inter-observer variability make it unlikely that significant improvements in the available data will be obtained in the foreseeable future. (p.277)*

He adds that it might be worthwhile just to "start over from scratch" and rebuild colorimetry on more advanced principles based on more recent data.

In fact, there is much more vigorous activity than this view implies. Current work is focused on ICAMs — image color appearance models — that can account for differences in the size of color areas, chromatic contrast, crispening, spreading, and differences in the quality or processing of images. And targeted revisions of color models, like those made on CIELAB, can produce a large number of separate modules or patches that can be swapped in and out of existing color models for specific applications or viewing situations. Future work is likely to be as creatively disorganized as it has been in the past.

My summary of CIECAM97s and CIECAM02 is based on Mark Fairchild's **Color Appearance Models (2nd ed.)** (Wiley: 2005) and R.W.G. Hunt's **Color Reproduction** (Wiley: 2004). An overview of CIECAM02 is available as a **color appearance lecture** and as a **spreadsheet of worked examples** from the Imaging Science Lab at Rochester Institute of Technology. Fairchild also provides an **overview of ICAMs**.

Artists may be interested in the average location of watercolor pigments on the **CIECAM  $a_Cb_C$  plane**, also available as a **PDF file**.

**N E X T : colormaking attributes**



Last revised 08.01.2005 • © 2005 Bruce MacEvoy