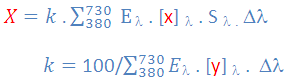
**CIE Tristimulus and Chromaticity Coordinates:**

Our visual system detects light using the light-sensitive organs that transmit generated impulses to the brain. The incident light is focused by cornea and the eye's lens to form an image being projected onto the retina. The retina is a light-sensitive layer at the back of the eye. The nerve cells in the retina that respond to light stimuli are called photoreceptors. These are classified into two types based on their shape: rods and cones. The area of the retina around the optical axis is designed for high resolution perception with a high concentration of cones. At the central point, each cone is connected by one optic nerve fiber to one point on the cortex of the brain. The cones are responsible for color vision. Observers with normal color vision have three different types of cones, commonly called **S**, **M**, and **L**. These abbreviated forms correspond to short, medium, and long wavelengths that each particular type of cones is sensitive to. Spectral sensitivities of the cones peak at 420 nm, 530 nm, and 560 nm, respectively. Rod receptors are more sensitive to light, but a large number of them are connected together before the sum of signals is sent to the brain. Thus they are good for sensing brightness and movement. In normal human observers, the spectral sensitivities of the three types of cones are linearly independent. The cones, as do any light detectors, integrate the light incident on them. This process reduces the entire spectrum of incident light to three signals, one for each type of cone, resulting in what is called trichromacy. The importance of trichromacy for the color imaging is that we can simulate almost any color by using just three primary colors - red, green, and blue.

The trichromatic nature of human vision has been mathematically formulated by [CIE](http://www.cie.co.at/cie/) (The Commission Internationale de l'Éclairage) to provide tristimulus values X, Y, and Z. CIEXYZ tristimulus values are thus fundamental measure of color. For the average human observer, under the same illuminant, stimuli with the same CIEXYZ coordinates are perceptually identical, meaning that the same color is perceived by an observer. This also means that the same color perceived under the same illuminant (and the same standard observer) has the same tristimulus values. From the measured color samples one would see that X, Y, and Z values approximately correspond to the red, green, and blue colors, respectively. In order to calculate tristimulus values, object spectra, spectral power distribution (SPD) of the illuminant, and the color matching functions are multiplied and summed over a range of wavelengths (usually from 380 nm to 730 nm in 10 nm intervals).

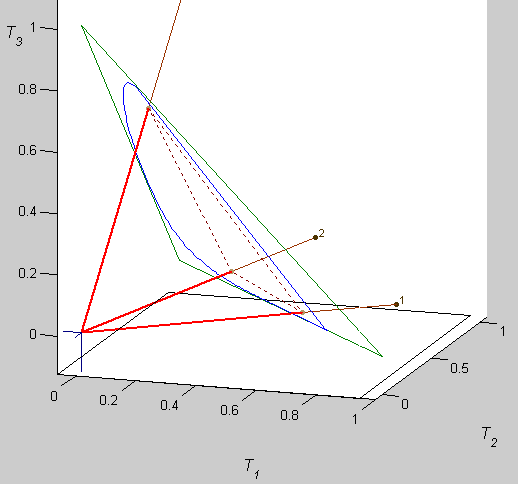
(I)

Equation **(I)** shows how the **X**-tristimulus component of the CIEXYZ triplet is obtained. **Y** and **Z** components are calculated similarly using the corresponding color-matching functions [**y**] and [**z**]. **S**(λ) denotes the object spectrum (as reflectance, transmittance, or radiance), [**x**] is the color-matching function for the CIE 1931 or 1964 standard observer (tabulated), **E**(λ) is the relative spectral power distribution of a reference illuminant (also tabulated), and ***k*** is a normalizing constant. The original definition of the tristimulus values involved integration over the visible range of spectrum (**I**). However, as the analytical expressions for the color-matching functions do not exist, integration was replaced by summation over the wavelength intervals. For light sources and displays, S(λ) is given in quantities such as spectral radiance. If S(λ) is given in absolute units and k = 683 lm/W is chosen, **Y** yields an absolute photometric quantity called luminance. The only measurement data needed to calculate CIE tristimulus is the spectral power density (SPD) of the emission source (monitor) or reflectance spectra from a reflective object (e.g., print). In case of reflectance, measurements have to be done under a defined reference illuminant (frequently D50 - see the next section for more details). Processing of emission or reflectance spectra in a form of SPD requires some reference "white", e.g., spectrum of a white patch on your monitor, the ambient light, or the white ceramic target that is measured during spectrophotometer calibration. For the emission data, an equi-energy "Illuminant E" is chosen (spectral power density is equal to 1 at any wavelength). Strictly speaking, there is no reference illuminant **E** for the emissive case as the source (display) provides the light energy itself. For more details on the emission spectra -> XYZ calculations (r2XYZ), the following Excel spreadsheet is available as a reference (download [here)](http://www.marcelpatek.com/download/r2XYZ_PM.xls). Processing of up to 153 spectral data with the subsequent Bradford transformation and LAB calculation is implemented in the VBA macro mode. Another Excel spreadsheet is available from the Bruce Lindbloom's web site (see [Spectral Calculator Spreadsheet](http://www.brucelindbloom.com/SpectCalcSpreadsheets.html)). Calculations of the CIE tristimulus values for reflective materials are thoroughly described by [Pascale](http://www.babelcolor.com/download/A%20review%20of%20RGB%20color%20spaces.pdf) and [Sharma](http://www.marcelpatek.com/color.html#Sharma86). Good news is that almost any spectrophotometer or colorimeter with a decent software will provide the XYZ values at the output.

**Color-matching functions:** also referred to as the CIE Standard Observer are intended to represent an average color vision of the normal human observer. Using three primary colors (R,G,B), the CIE performed experiments resulting in [tables](http://www.cvrl.org/cmfs.htm) of values determining how much of each of the primaries is needed to match colors of a reference spectrum. The corresponding graphical results are called the color-matching functions (CMF). The interpretation of these graphs is that they show how much of each primary (called the tristimulus value) is needed to match unit of intensity of a single wavelength of light. Following example shows how color-matching functions are used in practice. Let's calculate tristimulus values for a monochromatic laser light at 694 nm. We will use the equation (**I**) described above. The corresponding color-matching functions (CIE 1931 2deg) at 694 nm are 0.016987, 0.006138, and 0.000000, respectively. Calculated tristimulus values XYZ are then: 276.7, 100.0, 0.0. The way we arrived to these numbers is following: each value of the color-matching functions at 694 nm was multiplied by 1 (**S694**), then again by 1 (**E**) for the illuminant E (hypothetical equal-energy spectrum), and then again by 1 for Δλ. Resulting values were normalized for the Y-component (division by 0.006138 and multiplication by 100). The corresponding (x,y) chromaticity values are x = 0.735 and y = 0.265 (calculated from eq. (**II**) discussed later). This would be the "purest", most saturated red that most people could see. Although the responsivity at 700 nm is already very low, even the light from some near-infrared laser diodes at wavelengths beyond 750 nm can be seen if that light is sufficiently intense. A wavelenght of about 760 nm is considered the red limit of the naked eye.

The CIEXYZ system can be thought of as a 3-dimensional color space with orthogonal axes (**Fig. 1**). Corners of the green triangle in **Fig. 1** have coordinates of XYZ primaries ([1,0,0], [0,1,0], [1,0,0]) and lie in the plane defined as X + Y + Z = 1. The CIEXYZ primaries are defined arbitrarily and are clearly outside the gamut of all real colors. As such, they are invisible. No color of light or surface can reproduce them as they are not physically realizable. However, this is of no consequence, because the system is computational rather than visual. XYZ mixing triangle completely contains the chromaticity space of all real, visible colors, so all colors can be described as the positive mixture of the XYZ imaginary lights. Another feature of the CIE XYZ primaries is that equal values of X, Y, and Z produce white. When tristimulus values of visible colors (i.e. the color matching functions) are projected onto the green unit plane, trace of the horseshoe shape becomes apparent (in blue). This so called "horseshoe" diagram is essentially a projection of 3-D XYZ values onto 2-D xy plane.

**Figure 1:** CIEXYZ to xy transformation

[[](http://www.marcelpatek.com/images/blank.html)**From XYZ to xy system:** Projection of 3-D XYZ values onto 2-D xy plane](http://www.marcelpatek.com/images/blank.html)

Contour of such projection represents boundary of the human gamut (colors that we can see). Red lines in **Fig. 1** intersect the unit plane at primaries of any arbitrary gamut within the full gamut of all visible colors. Unfortunately, CIEXYZ values do not give any obvious representation of color and thus some other representations of tristimulus system are needed. One such system is the CIE 1931 xy-chromaticity diagram. The chromaticity coordinates x, y, and z are obtained by taking the ratios of the tristimulus values to their sum.

x = X/(X + Y + Z)

y = Y/(X + Y + Z)(II)

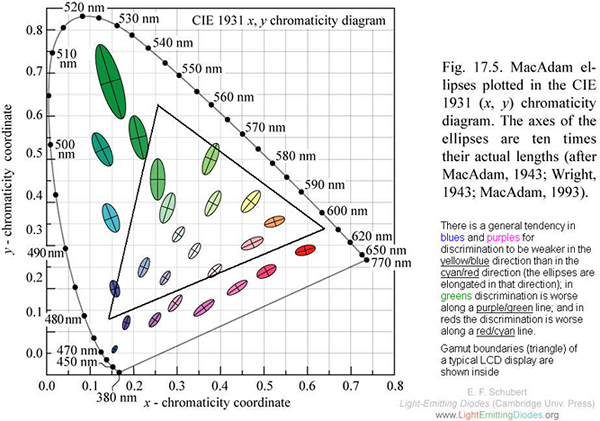
z = Z/(X + Y + Z)

and thus X = x/y . Y,  Z = z/y . Y   and  x + y + z = 1

The chromaticity coordinates represent the relative amounts of the three stimuli (X, Y, Z) required to obtain any color that we can see. As a consequence of its definition, the xy-system does not provide any information about the luminance. As mentioned earlier, the luminance is fully described by the **Y**-component of the XYZ triplet. Thus complete description of any color is given by the triplet xyY. Furthermore, as the **z** component bears no additional information (z = 1 - x - y), it is often omitted. The xy values of colors can be plotted in a useful graph mentioned above as the CIE (x,y)-chromaticity diagram (**Fig. 2**). There is a long and quite informative thread on XYZ->RGB transformation on [luminous-lanscape forum](http://luminous-landscape.com/forum/index.php?showtopic=37695).

As mentioned above, when we convert and plot the (x, y) chromaticity coordinates of the pure wavelengths of the visible spectrum, the resulting points all fall on a horseshoe-shaped line. This line is known as the spectrum locus. By definition, since all visible colors are composed of mixtures of these pure wavelengths, all visible colors must be located within the boundary formed by this curve. The line formed by connecting the endpoints of the horseshoe is called the "purple line" or "purple boundary." Colors on this line are composed of mixtures of pure 380nm (violet) and 770nm (red) light. It is essential to realize that chromaticity coordinates alone do not tell us what colors the eyes see. They just mean that two 3-D surfaces with the same (x, y, Y) coordinates in a given illuminant, or two lights with the same values of (X, Y, Z) appear identical. Also such diagram is visually non uniform, meaning that colors with an equal perceptual difference are not equally spaced in this color system. Note the colored ellipses (known as MacAdam ellipses) in **Figure 2**.

**Figure 2:** CIE 1931 xy-chromaticity diagram

[[](http://www.marcelpatek.com/images/MacAdam1.jpg)**MacAdam Ellipses:** Demonstrate poor perceptual uniformity of the CIE xy diagram. *E.F.Schubert, www.LightEmittingDiodes.org*](http://www.marcelpatek.com/images/MacAdam1.jpg)

If one plots small color differences that are just noticeable in various directions in the CIE (xy) diagram, the differences fall into an ellipsoid instead into a circle. This means that two colors can be very far apart in, e.g., the green region of the diagram before they appear to change color. In contrast, in the blue region, the same perceptual difference occurs over a much smaller distance. The realization that constant chromaticity differences would not yield constant perceived color differences led to development of multitude of color spaces, one of which has a dominant place in CIE colorimetry - CIELAB color space. More about it later.

Another, and more practical way to arrive to the horshoe diagram is to convert color matching functions to chromaticity (x,y) coordinates (eq. **II**) and plot the values in Excel spreadsheet. In this example, tristimulus value **X** is equal to the component [**x**] of color matching function. Analogously, **Y** and **Z** components equal to [**y**] and [**z**], respectively. It is important not to mix the color matching functions for different observer ([CIE 1931, 2o, CIE 1964, 10o](http://www.cvrl.org/cmfs.htm)). For details, see the formulas in "cie-xy" tab of this [spreadsheet](http://www.marcelpatek.com/download/BlackBody2.xls). Charts in **Fig. 5** were created using VB macros included with this spreadsheet.

**[(top)↑](http://www.marcelpatek.com/color.html" \l "top) Light Sources and Illuminants:**

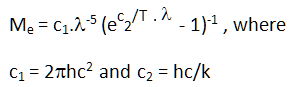
Terms such as reflectance and spectral radiance are used to describe the optical properties of materials and light sources. Reflectance (a measure of reflected light) depends on many parameters such as thickness of the sample, surface, angle of light incidence, spectral composition of the radiant source, and polarization effects. Both light sources and surfaces are described by their spectral power distribution (SPD) that contains all the basic physical data about the light and serves as the starting point for quantitative analyses of color. From the SPD, both the luminance and the chromaticity can be derived to describe color in the CIE system. For the reflective objects, the tristimulus color specification (CIEXYZ) is dependent on the viewing illuminant described by its SPD. Color of an emissive light source is also described by its unique CIEXYZ values. In contrast to light sources, reflective surfaces do not have intrinsic color although tristimulus XYZ values can be ascribed to the light reflected from them under a specific illuminant.

The CIE standards committee made a clear distinction between the terms illuminant and source. Source refers to a physical emitter of light, such as lamp, monitor, or the sun. Illuminant refers to a spectral power distribution, not necessarily provided directly by a source. As mentioned in the previous section, any color of physical light can be expressed by the chromaticity coordinate (x, y) on the CIE 1931 (x, y) chromaticity diagram (**Fig. 2**).

**Blackbody radiator:**

Technically, any object at any temperature above the absolute zero (0 K) will emit and absorb electromagnetic radiation to some extent. The intensity and frequency distribution of the radiation depends on the detailed structure of the body. A black body is a **theoretical object** that is capable of emitting and absorbing all frequencies of electromagnetic radiation uniformly (resulting in continuous spectra). A good approximation to a black body is a pinhole in an empty container maintained at a constant temperature. Such bodies do not emit light, and therefore appear black if their temperatures are low enough so as not to be self-luminous (black body is completely black at temperature of absolute zero while still having a ground state vibrational energy a.k.a. zero-point energy). At high temperatures (> 1,500 K) an appreciable proportion of the radiation is in the visible region of the spectrum.

As a result of established equilibrium (by emitting and re-absorbing radiation energy), emission spectrum of the black body is determined only by its temperature. In practice no material has been found to absorb all incoming radiation. The equation describing the specific intensity emitted by a "black body" at a given wavelength as a function of temperature is known as Planck's law for the black body radiation. The mathematical function describing the shape is called the Planck function (eq. **III**).

(III)

Other parameters are: λ = wavelength (m), T = temperature (K), h = Planck constant (6.62607E-34 J.s), c = velocity of light (299,792,458 m/s), and k = Boltzmann constant (1.38065E-23 J/K). Resulting spectral radiant exitance per unit wavelength interval Me(λ) is in units of W . m-2 . nm-1. Normalizing values of Me(λ) to unity (dimensionless) at λ=560 nm leads to Relative radiant power, which is often used in many graphical representations of spectral distribution (**Fig. 3**). Planckian blackbody calculation is demonstrated in this [spreadsheet](http://www.marcelpatek.com/download/BlackBody2.xls) together with calculation of tristimulus XYZ values for blackbody radiator.

| **Figure 3: Relative spectral distributions of radiant exitance of blackbody.** |
| --- |
| Blackbody1 |

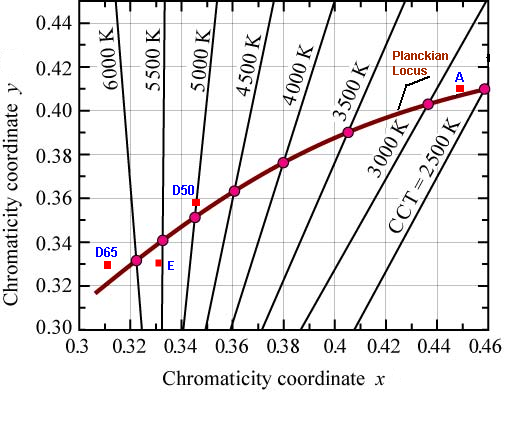
Regions of invisible radiation are shown in black. the above referenced spreadsheet allows for calculations of relative radiant power (dimensionless) and total radiant exitance (W . m-2) at a given temperature. Inspect formulas in the corresponding cells to uncover math behind graphs in **Fig. 3**.

| **Figure 4: Daylight (D65, D50) and blackbody radiators** |
| --- |
| Blackbody |

Deatails of continuous spectrum of blackbody light in comparison to two standard CIE daylight illuminantsis are shown in **Fig. 4** (thicker black lines). Since the temperature of a blackbody radiator describes its complete spectral power distribution and thereby its color, it is commonly referred to as the color temperature of the black body and as such it can be expressed by the chromaticity coordinates (x, y) on the CIE 1931 (x, y) chromaticity diagram, as shown in **Fig. 5**. The trace of the chromaticity coordinates of a blackbody plotted near the center of the diagram is the so-called Planckian locus (**Fig. 5**). The colors around the Planckian locus from about 2,500 to 20,000 K can be regarded as the source [white](http://en.wikipedia.org/wiki/White_point): 2,500 K being reddish white and 20,000 K being bluish white. The point labeled "A" (illuminant A) is the typical color of an incandescent lamp (CCT 2,856K) and illuminants [D50 and D65](http://en.wikipedia.org/wiki/Standard_illuminant) are the typical colors of day light, as standardized by the CIE. The color shift along the Planckian locus (warm to cool) is visually acceptable for general lighting, while color shift away from the Planckian locus (greenish or purplish) is perceived as unnatural. Strictly speaking, color temperature cannot be used for color chromaticity coordinates (x, y) off the Planckian locus, in which case the **correlated color temperature** (CCT) is used. CCT is the temperature of the black body whose perceived color most resembles that of the light source in question. Due to the perceptual non uniformity of the CIE (x, y) diagram, the iso-CCT lines are not perpendicular to the Planckian locus on the (x, y) diagram (**Fig. 6**).

| **Figure 5: CIE 1931 (x,y)-Chromaticity Diagram** |
| --- |
| Blackbody1.xlsm |

**Figure 6:** Correlated color temperature isotherms

[[](http://www.marcelpatek.com/images/planckian.png)**Planckian Locus:** Lines of constant correlated color temperature in the xy-chromaticity diagram. *E.F.Schubert, www.LightEmittingDiodes.org*](http://www.marcelpatek.com/images/planckian.png)

To calculate CCT, another improved chromaticity diagram (Uniform-chromaticity scale CIE 1960 (u-v) diagram - now superseded by CIE 1976 (u'-v') diagram) is used, where the iso-CCT lines are perpendicular to the Planckian locus.

As mentioned earlier, the CIE has defined several standard illuminants for the use in colorimetry. To represent different phases of daylight, a continuum of daylight illuminants has been defined by the CIE, which are uniquely specified in terms of their correlated color temperature (CCT). The **D65** and **D50** illuminant spectra are two daylight illuminants commonly used in colorimetry, which have CCTs of 6,504 K and 5,003 K, respectively (**Fig.3**). Spectral power distribution of the D65 source has clearly more of the blue component of the visible spectrum while the D50 source would appear reddish. The CIE **Illuminant A** represents a blackbody radiator at a temperature of 2,856 K and closely approximates the spectra of incandescent lamps. For links and references, go [here](http://www.marcelpatek.com/color.html#blackbody).

**Summary:**

* Black body is a theoretical object that when heated to a high temperature will emit visible light. Spectrum of this light can be theoretically predicted from the Planck's law.
* Chromaticity (color) of the light emitted over a range of temperatures will be perceived as a "white" (bluish, yellowish or reddish). Perceived "white" color is commonly referred to as the color temperature of the black body.
* The trace of the chromaticity coordinates of a blackbody radiator plotted inside the CIE (1931) xy chromaticity diagram is called Planckian locus.
* Correlated color temperature (CCT) is the temperature of the black body whose perceived color most resembles that of the light source in question. Chromaticities falling on the Planckian locus always have true color temperature while chromaticities near the locus have only correlated color temperature.
* Chromaticity of a light source cannot be described by CCT only. Information on how far off the locus our chromaticity values are is as important as the CCT.
* In most cases, where a "color temperature" is stated for some light source, it is almost always the CCT. Also for the CIE defined chromaticities of standard illuminants/sources, reference should always be to a CCT since these sources are not true blackbody radiators.
* In practice, monitor calibration can be always targeted at either Planckian or CCT temperatures. We just need to select chromaticity of the desired "white".

**Chromatic Adaptation:**

Let's start the topic of chromatic adaptation with a few examples. Suppose we have a white piece of paper that we look at under three different lightning conditions: midday, late sunset and at home under incandescent light. While all three light sources (illuminants) have different spectra and appearance (bluish, yellowish and yellow-reddish, respectively), color of the white paper that we perceive is very similar, i.e. white. This experiment works best if one looks only at the piece of paper without any background. Similarly, perceived differences in a tint of monitor that is calibrated to D50 (yellowish) versus D65 (bluish) will be less apparent in about 10 seconds (though it may take a few minutes to completely adjust). The effect is more pronounced when monitor is observed without any external reference object (painted walls, gray or black frames of your monitor). This is because our visual system adapts to the illuminant (the chromaticity of the illumination source) by shifting the appearance of all colors to restore the neutral appearance of a white surface. This phenomenon is known as the color constancy of human vision.

One of the mechanisms for color constancy is chromatic adaptation which can be defined as the ability of the human visual system to adjust to changes in the color of illumination. Preceding two examples suggest that chromatic adaptation is a result of sensory adaptation as well as cognitive behavior. The cognitive component of chromatic adaptation process is also known as discounting the illuminant and depends on observer's knowledge of scene content.

From the digital photography stand point, more general conclusion could be formulated such that any color space transformation is dependent on the source and destination illuminants. In other words, to achieve the same visual appearance of the original image under different display conditions (such as a computer monitor or a light booth), the captured image tristimulus values have to be transformed to take into account the light source of the display viewing conditions. Such transformations are called chromatic adaptation transforms (CATs). Thus, applying a chromatic adaptation transform to the tristimulus values (X,Y,Z) of a color under one adapting light source predicts tristimulus values (X',Y',Z') of the corresponding color under another adapting light source. Corresponding color refers to two stimuli that might have different XYZ tristimulus values, but because of chromatic adaptation to both illumination sources, they might appear to match. To be completely clear, CAT does not predict the colorimetric (XYZ) values of an object when illuminated by the destination source!! It is all about appearance of the corresponding color, not the absolute match of tristimulus values. Basic CIE tristimulus colorimetry is not designed to predict matches across different viewing conditions. Chromatic adaptation transforms thus extend basic tristimulus colorimetry. However, it is important to recognize that chromatic adaptation can only deal with small variations in the color of the illuminant. For huge changes in the color of the light source, we would surely see a change in the color of the object.

Common white-point chromaticities for a computer monitor are D50, D65, or D93 (CRTs). The "native white point" of common LCD monitors is around 6,000 K. Photographic printed outputs are usually evaluated under a standard illuminant of D50 (light boxes). The profile connection space (PCS) as defined by International Color Consortium ([ICC](http://www.color.org/index.xalter)) is D50, resulting in each device profile mapping image colors from source specific illuminant (e.g. CCT 6,000 K or D65 for monitors) to D50 illuminant.

There are numerous methods described in color textbooks developed for the chromatic adaptation. Many such transforms are based on the underlying model of chromatic adaptation proposed by von Kries. This model states that the individual sensors in the human eye are independent of the other and each type is adapted according to its adaptation function. Experiments indicate that a significant part of chromatic adaptation appears to take place in the photoreceptors, either as a change in the individual sensitivity curves of the L (long wavelengths), M (medium wavelengths) and S (short wavelengths) cones, or in the response of the retinal secondary cells to the cone outputs. To model this process in a meaningful way, it is necessary to transform from CIE XYZ tristimulus values into LMS cone responses. The cone responses under two illuminants can be described by a linear transform.

Mathematically the von Kries model can be written as:

[(top)↑](http://www.marcelpatek.com/color.html#top)

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/vonKries.gif | (IV) |

where coefficients at the diagonal of the 3 x 3 matrix (a, e, i) are the independent gain control factors. How these gain control coefficients are calculated is the key aspect to most CAT models. When transformation (**IV**) is performed in CIEXYZ space or RGB space (as opposed to the LMS cone space) it is referred to as an XYZ scaling or a "wrong von Kries" transformation (**V**). Symbols **D** and **S** denote a pair of **d**estination and **s**ource illuminants and the diagonal matrix is composed of ratios of source and destination illuminants. Such a simple scaling by diagonal matrix can be also applied to the RGB -> XYZ transfer matrix **C** (see below eq. [XI](http://www.marcelpatek.com/color.html#XI)) to ensure the white point preservation.

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/scaleXYZ.gif | (V) |

A more general model of chromatic adaptation (based on the von Kries model **IV**) is the so called generalized linear model (**VI**). In this case chromatic adaptation is once again controlled by independent gain factors (diagonal matrix **D**), but this time, **D** operates not on the tristimulus values but on a linear combination of vision system's sensors (cone responsivities):

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/linear.gif | (VI) |

**M** is a 3x3 matrix that transfers the source tristimulus values to three cone responsivities Ls, Ms, Ss as per **IV**. This source cone responsivity is then converted to destination (Ld, Md, Sd) by a diagonal matrix **D** containing the destination-to-source ratios of the cone responsivities (such as Ld/Ls, Md/Ms, Sd/Ss). Coefficients of the diagonal matrix **D** are determined from the tristimulus values of the source and destination illuminants (denoted by W) and the matrix **M** as follows:

[ Lw,s Mw,s Sw,s ]T = M \* [ Xw,s Yw,s Zw,s ]T and [ Lw,d Mw,d Sw,d ]T = M \* [ Xw,d Yw,d Zw,d ]T

By dividing the destination and source responses, we obtain: D = diag( Lw,d/Lw,s , Mw,d/Mw,s, Sw,d/Sw,s )

Last step is to convert the destination cone responsivities to the new tristimulus values:

[ XD YD ZD ]T = M-1 \* [ L d M d S d ]T

[Finlayson et al.](http://www.marcelpatek.com/color.html#finlayson) have shown that matrix M can be used to transform either the cone or tristimulus values into a suitable RGB space. In this case, the diagonal matrix D is composed of the white points of two illuminants in the RGB space.

From the practical point of view, the most commonly used transformation is the so called Bradford chromatic adaptation transform. Analogously to the above described CAT, it consists of a 3 x 3 linear transform from XYZ tristimulus values to an optimized set of Ls, Ms, Ss responses. These cone response values are then scaled to the Ld, Md, Sd values for the illuminant (or chosen white point) using a simple von Kries normalization. Specific feature of the Bradford transformation is the adaptation-dependent exponential function applied to the short wavelength cone signal (blue channel). The red (L) and green (M) domains remain strictly linear. However, in many applications, blue channel non-linearity is neglected. The Matlab code for both versions of the Bradford CAT is shown below (bradford.m). For simple spreadsheet form, [download this](http://www.marcelpatek.com/download/Bradford.xls) Excel file. Practical use of the Bradford transform involves multiplication of the transform matrix by the source tristimulus values as in eq. (**VI**).

% bradford.m

% calculates Bradford matrix from source to destination illuminant

Mcat = [0.8951 0.2664 -0.1614;

-0.7502 1.7135 0.0367;

0.0389 -0.0685 1.0296];

% destination white (here D50)

D50 = [96.4220 100.0000 82.5210];

% enter tristimulus of the source white (e.g., monitor)

DXX = [95.105 100.000 108.548];

sxx = Mcat \* DXX'; % source cone responses

d = Mcat \* D50'; % destination cone responses

tau = (sxx(3)/d(3))^0.0834; % S-cone non-linear correction

diag = [d(1)/sxx(1) 0 0; % diagonal linear matrix

0 d(2)/sxx(2) 0;

0 0 d(3)/sxx(3)];

diag\_nl = [d(1)/sxx(1) 0 0;

0 d(2)/sxx(2) 0;

0 0 (d(3)/sxx(3)^tau)\*sxx(3)^(tau-1)]; % non linear coeff in S

Mbfd = inv(Mcat) \* diag \* Mcat; % "linear" Bradford CAT matrix

Mbfd\_nl = inv(Mcat) \* diag\_nl \* Mcat; % "non-linear" Bradford CAT matrix

disp ('Bradford Matrix:' )

disp (Mbfd)

disp ('Bradford Matrix nonlinear:' )

disp (Mbfd\_nl)

Bradford transform is implemented as the only CAT in Adobe Photoshop.

**Summary:**

* Colors change their appearance upon changes in the color of the light source.
* Chromatic adaptation is the ability of the human visual system to discount the color of the illumination and to approximately preserve the appearance of an object.
* Chromatic adaptation transforms (CATS) seek to best model how the same color appears under two different lightning conditions (illuminants). CAT is a method for computing the corresponding color under a reference illuminant for a stimulus defined under a test illuminant.
* Due to its good performance, the Bradford CAT is recommended for appearance transformations from one illuminant to another.
* Such transformations are needed when comparing colorimetric data obtained at or referenced to different illuminants (light sources). Examples may include comparison of color patches on monitor having different color temperature, referencing colors to a standard illuminant (D50), or matching the monitor and light box light sources.

**Transformation of RGB Primaries:**

With many imaging devices being additive systems of red, green and blue colors, RGB primaries are well suited for colorimetric calculations. Those devices include, for example, digital cameras, scanners, projectors, and LCD/CRT displays. Moreover, since CIE color-matching functions are linearly related to RGB primaries, there is a simple relationship between RGB and CIEXYZ coordinates. In the 3-D coordinate space, one conveniently uses matrix algebra to do all the necessary conversions (for basics, look here - [Introduction to Matrix Algebra from Autar K. Kaw](http://numericalmethods.eng.usf.edu/matrixalgebrabook/frmMatrixDL.asp) : e-book and also download [this great Excel add-in](http://digilander.libero.it/foxes/)).

Let's start with few assumptions commonly used in colorimetry, namely proportionality and additivity of the color stimuli that are used in additive color mixing.

**Proportionality** **(scalability):** means that the radiant power (Lλ) of the color stimulus (think of color intensity) is changed proportionally by multiplying the maximum stimulus value by a positive factor (α): Lλ= α . Lλ,max. Vaguely speaking, the color intensity is linearly dependent on the maximum intensity via a scaling coefficient. When analyzing primary ramps, ramp chromaticities should plot as a single point.

**Additivity** of color stimuli means that a mixture of colors is simply sum of its components: Lλ,mix = Lλ,red + Lλ,green + Lλ,blue. In other words, the light produced is an additive combination of the light produced by each red, green, and blue primary.

Finally, combining the proportionality and additivity, a linear mixing equation results:  
Lλ,mix = r . Lλ,red,max + g . Lλ,green,max + b . Lλ,blue,max   
where r,g,b are relative amounts of the respective light.

Now, since the light adds, so must the tristimulus values. This ultimately leads to the expression for tristimulus values and RGB components of common digital devices:  
X = R . Xr,max + G . Xg,max + B . Xb,max  
Y = R . Yr,max + G . Yg,max + B . Yb,max (T-1)  
Z = R . Zr,max + G . Zg,max + B . Zb,max

R,G,B are normalized (R,G,B ∈ <0, 1>) and linearized (sometimes called "gamma corrected") R'G'B' input values that are used to calculate the corresponding normalized XYZ values. Xr,max is the tristimulus X for the red channel at maximum radiant output (R = 255). Xg,max and Xb,max are the corresponding values for the green and blue channels, respectively.

Now, we have just shown that the digital device RGB coordinates can be transformed to visual tristimuli using a system of linear equations **(T-1)**. In this system, R, G, and B are independent variables and X, Y, and Z are the dependent variables. Following the matrix notation, transformation **(VII)** is generally used to calculate the XYZ coordinates of pixels displayed on the monitor :

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/3x3mat.gif | (VII) |

Equation **(VII)** essentially performs a linear operation that converts between two 3-dimensional spaces, in our case RGB and XYZ (see the side panel for graphical representation of such transformations).

In even more general form, [XYZ]T = [3x3 matrix] x [RGB]T (superscript "T" denotes the transpose of the matrix). The middle [3x3] matrix (so far unknown) is derived from chromaticity coordinates of the RGB primaries and coordinates of the adapted white point. For example, for sRGB (D65) color space, the matrix is following:

[(top)↑](http://www.marcelpatek.com/color.html#top)

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/RGB2XYZm.gif | (VIII) |

Note that the sum of the row coefficients in **eq. (VIII)** equals to the XYZ coordinates of the white illuminant (in this case D65). This follows from the **eq. (VII)** for a special case of R=G=B=1, that is, a white light. To calculate just the luminance (Y), the following simplified equation can be used:

**Y** = [Yr,max Yg,max Yb,max] \* [RGB]T  (vector of the Y values comes from the second row of the [3x3] matrix).

In case that we know the XYZ coordinates and need to calculate the coresponding RGB values, the [3x3] matrix has to be inverted and the following equation **(VIII-inv)** used.

|  |  |
| --- | --- |
| xyz2rgb | (VIII-inv) |

To practice on a real example, let's calculate the R'G'B' values of the red patch of the GretagMacbeth ColorChecker. We will assume the sRGB color space and use the provided D65 L\*a\*b\* values. The L\*a\*b\* values for this particular patch (L\* = 40.554, a\* = 49.972, b\* = 25.45) were reported by the GretagMacbeth and are also part of the [Imatest](http://www.marcelpatek.com/imatest.html) software. First, we need to obtain the XYZ values from the L\*a\*b\* values. In this case we need the D65 XYZ values (because the sRGB color space is D65) and thus no [chromatic adaptation](http://www.marcelpatek.com/color.html#CAT) will be needed. Conversion from the L\*a\*b\* values to the XYZ can look rather complicated, but can be easily achieved by using a spreadsheet.

% Lab to XYZ (Matlab style) # for L\* > 8.0   
X = Xn\*(((L\*+16)/116) + (a\*/500))^3;  
Y = Yn\*((L\*+16)/116)^3;  
Z = Zn\*(((L\*+16)/116) - (b\*/200))^3;

In this case, the Xn,Yn,Zn values correspond to the standard illuminant of the chosen color space (sRGB D65)

Xn= 0.95047  
Yn= 1.00  
Zn= 1.08883

In this example, we will calculate the blue component (Z) of the red ColoChecker patch:

Z = 1.08883 \* (((40.554 +16)/116) - ((25.45/200))^3 = 0.0509

The corresponding red and green values are: X = 0.1927, Y = 0.1159. The complete XYZ matrix is then:

[0.1927   
0.1159   
0.0509]

And now the matrix multiplication. From the matrix algebra, it is known that two matrices **A** and **B** can be multiplied if the number of columns of **A** is equal to the number of rows of **B**. Number of columns in eq. **VIII-inv** is 3 and so is number of rows in the preceding XYZ matrix. Thus we can multiply both matrices. In this part, we will calculate the R component of the RGB triplet:

R = 0.1927\*3.2406 + 0.1159\*(-1.5372) + 0.0509\*(-0.4986) = 0.4210

G = 0.3280 and B = 0.0409 are calculated similarly.

Now, we are almost done. However, since we used a linear transformation of XYZ values, we obtained the linear RGB values. The linear RGB values need to be transformed to nonlinear (gamma encoded) values via the gamma correction, having a gamma 2.4-1 , an offset 0.055, and a slope of 12.92 (standard sRGB encoding). Finally the sR'G'B' values are scaled to 8-bit integers by multiplication with 255. We will demonstrate the calculations by a pseudo-code:

if RGB(i) <=0.004045   
R'G'B'(i) = 12.92\*RGB(i)\*255;  
else  
R'G'B'(i) = (1.055\*RGB(i)^(1/2.4)-0.055)\*255;

For the red component: R = (1.055\*0.4210^(1/2.4)-0.055)\*255 = 174

G = 51 and B = 57.  
In conclusion, the L\*a\*b\* values of the red patch (L\* = 40.554, a\* = 49.972, b\* = 25.45) have the correspong R'G'B' values of 174, 51, 57 in the sRGB color space.

Now the challenge comes. When characterizing a display (as an example), we would not necessarily be dealing with standard radiation source of a known color temperature, not mentioning the uncertainty in values of the RGB primaries of our display and its *gamma* correction. Hence, the above sRGB 3x3 matrix (**VIII**) won't be any good for our work.

Well, how do we get the right matrix coefficients and how can we use the above equations? Reasonable point to start is to assume complete spectral [proportionality](http://www.marcelpatek.com/color.html#proportionality) of our monitor. This means that if each RGB channel has scalable spectra, its chromacities will not vary with the lightness level. If this assumption holds up, the use of eq. **(VII)** is justified - as long as the [additivity](http://www.marcelpatek.com/color.html#additivity) is also preserved. We will make the latter assumption as well. This additionally means that the tristimulus values for the display white are at least very close to the summed tristimulus values of the full-on red, green, and blue primaries.

Now, to get the corresponding XYZ coordinates of primaries, a simple measurement of red, green, blue, and white patches on our display will be required. Why the white patch? Besides checking the additivity, we also have to determine the white point of our display. This is important for correct CAT transform from monitor white to any standard state of defined source, say D50 (which is conveniently the ICC standard for profile connection space ). For this purpose, we will apply Bradford chromatic adaptation transform (CAT) on our measured XYZ data.

There is one more step to check. Due to the various errors connected to measurements, [least square estimation](http://www.marcelpatek.com/monitor.html#character) of the matrix coefficients, or quality of the display circuitry, there is a good chance that our [3 x 3] matrix is not white point adapted (i.e., the additivity rule does not hold well). This means that RGB to XYZ conversion (for R=G=B=1) would not result in expected coordinates of the white at D50 (0.964212, 1.00000, 0.825188). In such a case, correction to the XYZ primaries in the matrix has to be done. Only then the sum of the row coefficients will be equal to the XYZ coordinates of the white illuminant (in this case D50). Just to clarify - this is not a chromatic adaptation transform - that one is done on XSYSZS/XDYDZD values and corresponds to an appearance transform ([see above](http://www.marcelpatek.com/color.html#CAT)). Required white point adaptation is purely mathematical scaling by a diagonal matrix that again ensures correct tristimulus values of the destination "white" and the corresponding neutral gray. We have already seen one example of XYZ scaling (a "wrong von Kries") in transformation (**IV**) in the previous paragraph. It may not be immediately apparent, but multiplication by a diagonal matrix results in a matrix where each column is multiplied by the same coefficient from the diagonal. In our specific case of RGB to XYZ transfer matrix, tristimulus triplet defining each of the RGB channels is multiplied by a coefficient such that the sum of rows will correspond to [XYZ]T vector of the white point (**IX**). In general, consequences of XYZ scaling by a factor are following: i) x,y chromaticity values do not change (as per (**II**)), ii) CCT of the white point will not change, iii) L\*a\*b\* and RGB values are, however changed. Since XYZ to L\*a\*b\* conversion formula contains a white point scaling, for R=G=B=1, L\* = 100, a\* = b\* = 0.

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/3x3xdiag.gif | (IX) |

Let's first see how we obtain this diagonal matrix. As we recall from the first paragraph, XYZ values can be transformed into x,y,z chromaticity values using eq. **(II)**. Substituting into eq. **(VII)**, we get a new equation **(X)**. Now, if you are wonderings how the second row of the 3x3 matrix in **(X)** came about, here is the trick: **Yr** = yr \* **Yr**/yr. Now, let's take out the Yi/yi and make them equal to Ti (i = r,g,b). In the resulting eq. **(XI)**, colorimetric RGB primaries **(C)** are related to tristimulus values via additional "corrective" matrix **(T)**. This relationship is in agreement with the formula recommended by Xerox Corp. (as per "Color Encoding Standard", C-3 (1989)).

|  |  |
| --- | --- |
| wp1.png |  |

Coefficients (xr, yr, zr), (xg, yg, zg), and (xb, yb, zb) in **(XI)** are the chromaticity coordinates of the red, green, and blue primaries, respectively. New parameters Tr, Tg, and Tb are the proportional constants for the corresponding primaries under the adapted white point (that's the correction we need). Equation **(XI)** can be solved for [T] using a known condition of a gray-balanced and normalized RGB system (i.e., for R=G=B=1, a reference white point - in our case D50 - should be produced). Normalized tristimulus values of the D50 white point [W] = [0.964212, 1, 0.825188]T are used together with known coordinates of primaries to calculate [T] = [C]-1 [W] from eq. **(XI)**.

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/wp2.gif | (XII) |

Resulting T values are substituted back into eq. **(XI)** with the known chromaticity coordinates of primaries. Now, we know [C] and [T] that can be combined into one transfer matrix [M]. Final transformation is then of the form of:

[XYZ]T= [Mwp] x [RGB]T

To help you with the above calculations of matrix [Mwp], here is the Excel [spreadsheet for download](http://www.marcelpatek.com/download/XYZ33.xls). Below is the Matlab code.

% WP\_adapt.m

% Reads manually provided 3x3 matrix and

% returns RGB2XYZ conversion matrix adapted to a white point

clear all

Tm = [48.2793 32.1731 14.4443;

25.9368 65.8478 10.5637;

1.1492 8.1183 74.8447]'; % enter the 3x3 matrix, transpose;

W=[0.964212; 1; 0.825188]; % XYZ of the white point (here D50)

% Build xyz chromaticity matrix

xyz=[Tm(1)/sum(Tm(1,:)), Tm(1,2)/sum(Tm(1,:)),Tm(1,3)/sum(Tm(1,:));

Tm(2)/sum(Tm(2,:)), Tm(2,2)/sum(Tm(2,:)),Tm(2,3)/sum(Tm(2,:));

Tm(3)/sum(Tm(3,:)), Tm(3,2)/sum(Tm(3,:)),Tm(3,3)/sum(Tm(3,:))]';

T = inv(xyz) \* W; % Proportional constants for R,G,B

Tc = [T(1) 0 0;0 T(2) 0;0 0 T(3)]; % 3x3 Matrix of proportional constants

MD50 = (xyz \* Tc); %Final RGB2XYZ matrix for the white point Tm

format short

%-----------Display---------%

disp ('Original matrix: ')

disp (Tm')

disp ('M adapted to D50 (MD50): ')

disp (MD50.\*100)

Unfortunately, there is one more complication in the matrix transformation. All the above equations work fine when radiant output at R=G=B=0 is equal to zero. For non zero output at the black, a correction has to be made to the matrix to account for this "flare" component. This phenomenon is quite common for LCD monitors ([link](http://www.marcelpatek.com/LCD.html#BC)) and cannot be avoided as the liquid crystals pass some light through even at the "switched-off" state. Typical [3x3] matrix is then extended to a [3x4] matrix by adding the black-level radiant input Xk, Yk, and Zk (**XII**). This form of the matrix ensures that values for the black output are added to the resulting XYZ values at the dark levels while the bright values will be corrected proportionally by only a fraction of the black values. Values at the highest output (R=G=B=1) will not be corrected at all.

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/3x4mat.gif | (XIII) |

In order to calculate radiometric scalars (linearized RGB values) from the measure XYZ values and RGB primaries, the inverse black corrected matrix should be used (**XIV**).

|  |  |
| --- | --- |
| http://www.marcelpatek.com/images/3x4inv_mat.gif | (XIV) |

Note the last two equations as we will use them extensively in display characterization examples

### Gamma:

Several authoritative and very insightful papers and web sites were written on the subject of **gamma** ([ref](http://www.marcelpatek.com/gamma.html#gammaref)). However, definitions and interpretations still remain a subject of continuing discussion. Instead of going into the merits and controversials of gamma encoding/decoding/compensation, I will focus on its principles, application, and measurement.

To at least introduce the topic, let's start with simple empirical observations that will lead us to more intuitive meaning of the gamma. Our everyday experience points to the fact that human senses, such as hearing, sight, smell, and heat respond and adapt to extremely wide ranges of stimuli. To cope with such a large dynamic range, our hearing system, eyes, nose, and sensory neurons in skin evaluate the corresponding stimuli on nonlinear (approximately logarithmic) scale. Simply, when we double the light intensity, we don't see the light as twice as bright! The magnitude of a physical stimulus and its perceived intensity or strength follows the [Stevens' power law](http://en.wikipedia.org/wiki/Stevens%27_power_law&link=1) (**Fig. 14**). Stevens' formulation is widely considered to supersede the famous [Weber-Fechner law](http://en.wikipedia.org/wiki/Weber-Fechner_law&link=1) on the basis that it describes a wider range of sensations. Mathematically, the Stevens' law is formulated as follows:

R = k . (S-S0)p

Logarithm of both sides of this equation leads to a linear relationship between log(S-S0) and log R, with slope of the line determined by **p**. **Figure 14** illustrates the qualitative relationship between stimulus of the human sensor system (S) and the response that we feel (R). Dashed line (p =1) would correspond to the linear dependence.

| **Figure 14: Stimulus intensity vs. objective response** |
| --- |
| Stevens |

For example, the way we feel heaviness of an object is characterized by a high value of **p**, reflected in a steep curve. In other words, once a stimulus is strong enough to sense "heaviness" of the object, the perception of "heaviness" rapidly becomes stronger as the stimulus becomes stronger. The other sensory responses shown have lower **p** value, which means that they can cover much wider ranges of stimulus intensity (dynamic range). Perception of light brightness has particularly low power exponent, which indicates that our vision system adapts to a wide dynamic ranges of the light. We will see the coefficient 0.33 (1/3) later in the context of perceptually uniform lightness scale as defined in the CIELAB color space. The sensitivity of the human eye is influenced by the average light level to which the eye is exposed.

Representative result of noticeable luminance-difference measurements performed in the range of typical LCD luminances is shown in **Fig. 15**. The quantity log(ΔL/L) is plotted against the log(L) where L is luminance in cd/m2 (correlates with relative brightness). Some theoretical approaches predict a three-segmented curve, with slopes of -1 (the linear range), -0.5 (the square root range), and 0 (the Weber range). The same data on the linear scale is shown in **Fig. 16**. For luminance levels of interest (0.1 - 200 cd/m2), the ratio of ΔL/L has approximately constant value of 0.01 (so called **Weber-Fechner fraction**). Fit to a line with the slope of ~0.01 is remarkably good. In this region, the just noticeable difference in luminance (JND ~ ΔL) then follows the linear equation: ΔL = 0.01 \* L. This relationship just confirms the empirical observation that we are more sensitive to luminance changes in the darker levels (at L = 100, JND = 1 while at L = 10, JND = 0.1). This observation will be mentioned later when gamma encoding is discussed. It is important to note that sensitivity of human vision varies with the light level and that in order for us to discriminate between two close luminance levels, one has to be fully adapted to the surround luminance. To be able to distinguish shades of gray in darker areas, our eyes have to be completely adapted to a lower intensity environment (globally as well as locally). For additional information on this topic, refer to the paper "Gamma and its disguises" of [Charles Poynton](http://www.marcelpatek.com/gamma.html#disg) and see the section "Why gamma" at [Norman Koren's web pages.](http://www.normankoren.com/makingfineprints1A.html&link=1)

| **Figure 15: Log-Log scale** |
| --- |
| wyszecki.xls |

| **Figure 16: Linear scale** |
| --- |
| Stevens |

In contrast to the human senses, linear data representation is widely used for computer generated imagery, mostly because models of illumination, white-point conversion, RGB color conversions, and color matching are defined linearly. Furthermore, the image sensors (CCD or CMOS) are linear devices that respond proportionally to the number of photons that interact with their electrode structure.

#### Quantization:

At the outset of the gamma discussion, the topic of quantization will be briefly discussed. In order to render captured scenes faithfully, digital devices have to have the ability to adequately sample (quantize) dynamic ranges of typical everyday scenes. Digital imaging systems with usually a limited number of bits (typically 8) must be designed so that there is enough precision in the low levels (shadows) to avoid visible contouring (our eyes are quite sensitive to transitions in shadow areas - **Fig. 16**). This can be achieved by applying a logarithmic or gamma-law "correction" to the linear data (logarithmic quantization has often been recommended but rarely implemented). Such transformations usually proceeds in two steps. First, the image is converted by analog-to-digital converter to, e.g., 12 bits in uniform steps (4,096 steps). Second, the output signal is then reduced to 8 bits by a nonlinear transformation that leads to compression of high level bins. When working with a typical LCD display that has about a 300:1 contrast ratio, our eyes can distinguish about half of the luminance levels. For 8-bit grayscale gradient, there would be theoretically 8.2 f-stops (exposure zones, log2300 = 8.2). For all 256 levels, there would be about 256/8.2 = ~ 31 distinguishable levels per a zone. As a result of interference from higher luminance levels (flare, bright environment), our capacity to distinguish levels in darker areas is compromised. Coming back to the [Weber-Fechner law](http://www.marcelpatek.com/gamma.html#webfech), the theoretical number of distinguishable levels per exposure zone is equal to log(2)/log(1.01) = 70. Number two relates to the exposure zone definition (half or double of the light intensity) and 1.01 is the 1% JND difference coming from the Weber-Fechner law. Again, this number of levels assumes complete adaptation to relatively narrow range of luminances around the evaluated neighboring levels.

**Quantization of digital images:** Process of assigning intensity levels to a digital output is called quantization. During quantization, every floating point value of luminance related to input signal has to be mapped to one of the single integer values accepted by typical visualization devices (monitor, printer, projector). If uniform (linear) quantization steps are used, more than 10 bits (1,024 bins) should be used to maintain details in shadows and smooth transitions\*. Nonlinear quantization techniques provide compression by requiring a smaller sample size (i.e. number of bits) to cover the full range of input than a linear quantization technique. Logarithmic or power function quantization provides more resolution at lower levels with higher levels spaced further apart than the low ones. Thus darker colors are represented in greater detail than they wouldbe in the linearly coded image. This parallels the way in which we perceive differences in lightness as we are more sensitive to absolute changes in shadows than in bright areas.   
Note\*: if we send linearly sampled data through a power function and quantize them, we will lose certain number of levels, depending on the number of bits of input and the number of bits of output. For example if we would send 256 linearly spaced levels (8-bit encoding) through the gamma compression curve of 0.4545 ( = 1/2.2), we would lose 72 levels to give us only 184 to work with. However, if our camera uses 10 or 12 bits to represent the captured linear RAW data, the gamma compression with the same power coefficient of 0.4545 would preserve all 256 levels in 8-bit encoding. See the Bruce Lindbloom's [Level Calculator](http://www.brucelindbloom.com/index.html?LevelsCalculator.html" \t "_blank) for details.

The idea behind nonlinear quantization is depicted in **Figure 17**. The x-axis is normalized digital input such as a photon count passed through the analog-to-digital converter (ADC). Numbers just above are the input values for 8-bit encoding (typical image editing range of 0-255). Black line is the linear response of an arbitrary digital device (camera, scanner sensor). Clearly, the output (on the y-axis) is directly proportional to the input. As an example of logarithmic correction, the blue curve in **Fig. 17** describes transformation from the linear (black) to a logarithmic domain. In order to avoid log troubles at 0, transformation (**I**) was used instead of direct log. **Dc** is the digital count input and **p** is a parameter ∈ (0, ∞).

|  |  |  |  |
| --- | --- | --- | --- |
| *y* = | |  | | --- | | log (1 + p . Dc) | | log (1 + p . maxDc) | |

(I)

The red curve is a "classical" simple gamma-corrected linear mapping. One can see that gamma mapping is steeper in shadows while highlights have nearly identical mapping. If we were quantizing in linear steps, we would be "wasting" bits in bright regions (remember, we can't distinguish much of the brightness level difference in lights anyway) and not having enough in areas where it matters most, i.e. darks. Here is why. For the linear scale, pixels above value of 204 (in 8-bit encoding) would require (1.0-0.8) x 100 = 20% of the available bits (same as pixels below 50). For the nonlinear scale, same regions will claim ~10% ((1.0-0.9) x 100) and 50% of available bits, respectively (orange rectangles). Read the following box and references for more examples and explanations.

**Let's compare outcomes of the linear and gamma encoding schemes:**  
i) we will assume a linear grayscale file in 8-bit encoding (total of 256 levels). Since the output luminance changes linearly with the input luminance, the first exposure zone (from 128-255) will have 128 brightness levels (256/2), the second zone (64-128) 64, and the last three dark zones (no. 6-8) will have the total of 7 levels (4+2+1).  
ii) now for the gamma 2.2 encoded 8-bit grayscale image (e.g., JPG). Taking the half of the normalized luminance input will result in normalized output luminance of 0.5(1/2.2) = 0.7298 for the first exposure zone. This would consume (1-0.7298)\*256 = 69 brightness levels. Second zone would have (0.7298-0.25(1/2.2))\*256 = 50 levels, third 37, fourth 27, ... 20, 14, 10, .. Note that due to the power function applied on the input luminance data, the number of exposure zones is greater than 8. Clearly, the gamma encoding reallocates encoding levels from the upper exposure zones into lower zones to make the distribution of levels much more uniform. This way a serious banding at low levels is avoided. See also, ["Understanding RAW files"](http://www.luminous-landscape.com/tutorials/understanding-series/u-raw-files.shtml) at the Luminous Landscape and [Human vision and tonal levels](http://www.normankoren.com/digital_tonality.html) at Norman Koren's web page.

| **Figure 17: Effect of quantization in a capture device** |
| --- |
| quantize |

Clear benefit of these transformations is that the number of bits required to eliminate contouring of camera images in a computer display can be reduced from the "safe" 10-12 bits for linear data to 8 bits for logarithmic or gamma-corrected data (see the [note above](http://www.marcelpatek.com/gamma.html#quanti)). This has beneficial impact on signal storage and transmission. Also, such nonlinear ("gamma") correction mathematically transforms the physical light intensity into a perceptually-uniform domain. Furthermore, it just happens that operating principle of CRT displays requires a gamma correction close to 1/γ to maintain perceptual and pleasing tone uniformity. So, the first gamma compressed the signal in a capture device, while the second gamma effectively decompressed the corresponding image (**Fig. 18**), keeping it perceptually similar to the original scene. It is the second gamma that we are mostly concerned with when processing and manipulating digital images. **Figure 18** depicts the typical display gamma relationship between digital RGB values (normalized at x-axis) and the luminance output (red curve). Once the input digital counts are linearized, linear relationship between input and luminance becomes apparent (blue line). In general, linearization can be achieved by any of the mathematical models described in the next section or by canceling out the gamma and 1/gamma curves shown in **Figures 17, 18** (red lines). An example of such cancellation is the so called overall system gamma (viewing gamma) that can be also computed by multiplying the camera gamma by the display gamma (γ\_overall = γ\_camera \* 1/γ\_display ~ 1). In reality, this overall gamma is typically in range of 0.96-1.30. The simple camera gamma is equal to 1/2.2 = 0.4545. Display gamma is gamma that we are concerned with when setting the calibration target. For Windows system its value is typically 2.2. It is important to note that gamma only affects middle tones and has negligible effect on dark or white areas.

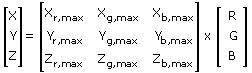
| **Figure 18: OETF curve and its linearized form** |
| --- |
| Sharp |

Unfortunately, the term gamma is used rather ambiguously and it is always necessary to verify how the particular gamma is defined. For display devices which we are most interested in, the correct term for (or meaning of) gamma is the **tone reproduction curve** (TRC). It is also known as the **optoelectronic transfer function** (OETF) and in general, it describes the relationship between the signal sent to a display (from the video frame buffer) and the radiant output produced by each RGB channel. This relationship is usually nonlinear (as discussed above) and must be somehow modeled or predicted in order to characterize any particular display device. Technically, for LCD displays, this relationship is not an easy function, but rather results of curve-fitting to representative input points that are stored in profile or calibration lookup table (LUT). This LUT typically involves 16-bit data mapping. Strictly speaking, it follows that there is no "gamma" for LCD displays as there is no "power function" physics behind the LCD luminance output. While electron guns in CRTs have smooth power-function characteristics along the whole tonal range of each RGB component, the native response of LCD is closer to a sigmoidal shape ([Karim](http://www.marcelpatek.com/gamma.html" \l "karim)). However, LCD displays still have the tone reproduction curve! As for the CRT monitors, such transfer function describes the relationship between the digital input, given by the RGB values, and the luminance produced by each RGB channel. Since images should look very similar regardless of the display type, LCD manufacturers build correction tables into the display circuitry to account for gamma-corrected images. Hence, LCD displays show a response similar to a CRT gamma function !!

**Tristimulus values and radiometric scalars:** We have mentioned the luminance output (**Y**) at several points in the preceding paragraph. This short note explains the math behind measurements of the luminance and the tristimulus values. For more details, visit the [Color section](http://www.marcelpatek.com/color.html) of these web pages. The trichromatic nature of human vision is mathematically formulated by [CIE](http://www.cie.co.at/cie/) (The Commission Internationale de l'Éclairage) to provide tristimulus values X, Y, and Z. CIEXYZ tristimulus values are thus fundamental measure of color - "color coordinates". In this system, the tristimulus value **Y** of a luminous object is known as the luminance. Absolute values of luminance are reported in candela/m2 (cd/m2). By convention, the XYZ values are normalized by a normalizing constant that sets Y = 100 for perfectly reflecting or transmitting samples (e.g. monitor white). For any RGB system (computer displays, multimedia and imaging applications), R'G'B' triplet refers to gamma encoded digital count values while RGB refers to linear RGB values. Note: in most application software (Photoshop, image editors), R'G'B' values are commonly referred to as "RGB". For color calculation, we will use the nomenclature common in the literature, i.e., RGB for linearized values and R'G'B' for the gamma encoded equivalents. The first step in determining the tristimulus values XYZ from R'G'B' values sent to the monitor is to examine the relationship between digital counts (R'G'B'dc) and each channel's scalar (R,G,B). Each of the R'G'B' channels must first be linearized so that the output intensity is linearly related to the input. There are several forms of the linearization function. These include simple gamma (R,G,B = dcγ), GOG, or LUT (see below). The scalar values for each channel and the corresponding R,G,B values can be calculated either from the ramp measurement or from the tristimulus values of the primaries. In the former case: R = Xm/Xmax, G = Ym/Ymax, and B = Zm/Zmax, where index m refers to a measured value and max is the maximum value. The latter option relies on the linearity of the R,G,B and XYZ. From equation **T-2** shown later, [RGB]T = inv(3x3) \* [XYZ]T. Next, coefficients of the linearization function have to be determined, usually using some form of least-squares optimization.

Now, when we have described the nonlinear relationship between the input RGBdc and scalars R', G', and B', tristimulus components XYZ can be calculated as follows:  
X = R . Xr,max + G . Xg,max + B . Xb,max  
Y = R . Yr,max + G . Yg,max + B . Yb,max   
Z = R . Zr,max + G . Zg,max + B . Zb,max(T-1)

These calculations can easily be performed using matrix algebra (see the [Color section](http://www.marcelpatek.com/color.html)).

(T-2)

R,G,B values are thus linearized R'G'B'dc input values that will be used to calculate the corresponding XYZ values. Xr,max is the tristimulus X for the red channel at maximum radiant output (R=G=B=255).

[(top)↑](http://www.marcelpatek.com/gamma.html#top)

#### Linearization:

During the image processing and color space transformations that involve device independent color spaces, a linear relationship between image pixel values specified in software and the luminance has to be established. We already know that monitors (CRT, LCD) will have a nonlinear response. The luminance can be generally modeled using a power function with an exponent, gamma, as in eq. **(II)** (simple gamma). During all these operations, luminance and RGB digital counts (values sent to a monitor) have to be normalized to values between 0 and 1. Again, in order to display image information as linear luminance we need to modify the RGB dc domain (i.e., we have to linearize it and thus remove the gamma encoding). As discussed in the previous paragraph, this need comes from display systems where the camera and displays had different transfer functions (which, unless corrected for, would cause problems with tone reproduction).

Simple gamma correction is given by the following equation:

|  |  |
| --- | --- |
| R,G,B = dcγ | (II) |

Other (and more accurate) models include several parameters and nonlinear curve fitting to a power function. Models such as **GOG** or **GOGO** are successfully used to characterize **CRT** monitors. Briefly, the GOG (gain-offset-gamma) model uses the formula:

|  |  |
| --- | --- |
| O = (a.I+b)γ | (III) |

while GOGO model (model recommended for CRT colorimetry) adds additional offset term:

|  |  |
| --- | --- |
| O = (a.I+b)γ+ c | (IV) |

Another model used in many RGB color-encoding standards (essentially the GOG) is expressed as:

|  |  |
| --- | --- |
| O=[(I+b)/(1+b)]γ | (V) |

and uses only one constant term for the gain and offset. **O** refers to linearized R,G, or B values and **I** is the digital count input normalized to <0-1>. Parameter **a** is an offset ("brightness" on CRTs), **b** is gain ("contrast" on CRTs), **c** is another offset, and **γ** is the power function coefficient. By formally substituting **a** for 1/(1+b) and **c** for b/(1+b) in **eq. III**, we arrive to the **eq. V**. This particular linearization model is used in [GammaCalc](http://www.marcelpatek.com/gammacalc.php) script to fit your experimental data and calculate the corresponding gamma.

To summarize, conversion of R'G'B' image pixel values to the CIEXYZ tri-stimulus values can be achieved via a two stage process.

* Firstly, we need to calculate the relationship between input image pixel values (dc) and the displayed luminous intensity (Y). This relationship is the transfer function, often simplified to gamma. The transfer functions will usually differ for each channel so they are best measured independently. A note on the use of tristimulus values (XYZ) for gamma assessment: while luminance of each channel is described by the corresponding Y-component of the XYZ triplet, calculated gamma is the same regardless of which tristimulus component was used (providing the component was normalized in the range of <0,1>). Thus the red channel gamma can be calculated from normalized **dc** values and the X-component of XYZ, green channel would use Y-component and blue channel the Z-component. This approach is basically building one-dimensional look-up tables where luminance is substituted by a radiometric scalars (R,G,B) according to: R = LUT(dcr) (for the red channel).
* The second stage is to convert between the displayed red, green and blue to the CIE tristimulus values. This is most easily performed by using a matrix transform of the following form: [XYZ]T = [3x3 matrix] \* [RGB]T where X, Y, Z are the desired CIE tri-stimulus values, R, G, B are the RGB values obtained from the transfer functions (now linearized) and the 3x3 matrix contains the measured CIE tri-stimulus values for monitor's three channels at the maximum output.

#### Practical Applications:

In a typical characterization of displays, commonly used technique for monitor gamma assessment uses the so called "log-gamma". This gamma is based on the original gamma definition, i.e., the slope of the "linear part in the tone characteristics obtained by linear regression in the logarithmic domain". Unfortunately, such gamma cannot be defined unambiguously since the slope depends on which part of the curve is chosen to be linear. Besides that, only the simple gamma model (**II**) is assumed, that is: Y = dcγ + kTransformation into logarithmic domain leads to: log(Y) = γ \* log(dc) + k1 which is a linear equation with the slope being equal to gamma (γ). Note that no linearization of R'G'B' values was needed, although the non linearity at the dark are is clearly evident (**Fig. 19**). Overall, this method is still the fastest and relatively accurate way to assess the monitor gamma. To measure grayscale gamma, one usually displays a series of gray patches from RGB=0 to RGB=255 (e.g. in steps of 17) and measures the luminance response as the Yr,g,b component of the tristimulus values (it is the middle Y in the XYZ output). Both R'G'B' and Y values are then normalized to fit within the range of (0-1) by dividing them by 255 and Ymax, respectively. Logarithm is calculated for both series and plotted as y=log(Y/Ymax) against x=log(R'G'B'/255).

| **Figure 19: Example of γ calculation** |
| --- |
| gamma |

Slope of the linear portion of this plot is taken as the overall system gamma (red line in **Figure 19**). If you would rather skip the math part, here is the [spreadsheet](http://www.marcelpatek.com/download/gray_XYZ.xls) that calculates gamma for you in this simple situation. Alternatively, to have log-gamma calculated for you, upload measured values through the [GammaCalc](http://www.marcelpatek.com/gammacalc.php) page.

Unfortunately, as mentioned above, the OETF characteristics of LCDs are such that a single analytic equation cannot be used to accurately describe their general behavior. Consequently, equations such as eqs. (**II-V**) may poorly describe the OETF. However, since all computer-controlled systems include video look-up table, three one-dimensional (1D) look-up tables (one for each channel) can be obtained to define the OETF (see [Display Color Management](http://www.marcelpatek.com/monitor.html)). When the GOG functions are replaced with simple one-dimensional LUTs to characterize the display's electro-optical transfer functions, the characterization performance is excellent.

#### L-star curve:

Some profiling software packages feature several settings for the TRC curve:

* gamma of your choice, typically gamma 2.2 or 1.8
* [sRGB](http://en.wikipedia.org/wiki/SRGB_color_space) TRC (has a linear segment at the dark end and the overall gamma of approximately 2.2), and/or
* the so called L\* (read L-star) curve, which simulates response of the [L\*](http://www.marcelpatek.com/glossary.html#indexL) channel of the CIELAB color space

There is not much one can find on the L\* curve and it seems that some vendors use proprietary algorithms for the L-star calibration curve and display icc profiles. To get at least a qualitative picture of how the L-star calibration curve looks like, **Figure 20** shows the TRCs for both the sRGB and L-star curves in the form of Y(normalized) vs. R'G'B'(normalized) plot - in this section denoted as RGB.

| **Figure 20: sRGB and L-star TRCs (Yn vs. RGBn)** |
| --- |
| Ys vs RGBn |

Relative luminance data (Y) were obtained from the [icclu utility](http://www.marcelpatek.com/argyll.html#icclu) of Argyll CMS for both the sRGB and [L-star](http://www.colormanagement.org/en/workingspaces.html) icc profiles. This utility uses the corresponding icc/icm profiles to output XYZ tristimulus or L\*a\*b\* values for batches of input RGB device values. ColorThink Pro can also do the job, although the output data is rounded to only two decimal places.

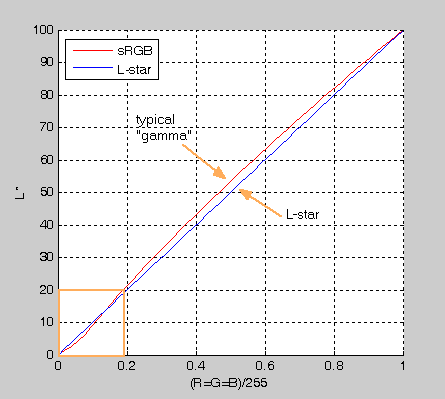
**Note:** It should be stressed at this point that sRGB and L-star icc profiles are just color space profiles, theoretical constructs that have nothing to do with display profiles created during calibration/profiling process. As such, these color space profiles could be considered as ideal examples to demonstrate the characteristics of display gamma and L-star profiles.

Visual inspection of **Fig. 20** suggests that L-star curve brings more contrast (steepness, higher gamma) into the midtones and highlights (RGB 0.5-0.9). We will also see later that value of gamma is not constant and that it varies along the whole tonal range. Curve characteristics in shadows (< 0.25) are not that clear and additional analysis has to be performed. Overall, both TRC curves are quite similar with no distinct features.

An alternative way to arrive to the similar picture is to calculate luminance values (**Y**) for the L-star curve from the corresponding values of the CIELAB L-channel (0-100). However, it should be mentioned that typical L-star curve is not identical to the CIELAB L-channel.

Here is the formula for L\* to Y transformation: **Y** = 100 . [(L\*+16)/116]3 for Y/100 > 0.008856, Y ∈ <0-100> and **Y** = 100 . L\*/903.3 for Y/100 < 0.008856 Another characteristic plot is shown in **Figure 21**. Relationship between calculated CIELAB L\*-values and the RGB values for a 5-step gray ramp is plotted as L\* vs. normalized input RGB. Same ideal sRGB and L-star icc profiles were used, only this time the icclu utility was configured to output L\*a\*b\* values. As one can see, the L-star response is clearly linear in all brightness ranges (blue line). This means that doubling the value of RGB always changes value of L by the factor of two or that by stepping the RGB values by e.g., 10 points will change the L values by a constant increment (in this case by 100/255\*10 = 3.9). Also, since incremental changes in L are perceptually uniform, changes from dark to

**Figure 21:** L-star and ideal sRGB plots

[[](http://www.marcelpatek.com/images/Lstar_gamma.png)**L-star:** LAB Lightness vs. RGB inputs for L-star and ideal sRGB gamma gray ramp](http://www.marcelpatek.com/images/Lstar_gamma.png)

bright values in a synthetic grayscale (RGB form 0-255) will be perceived as smooth and uniform. On the other hand, the sRGB curve (or any other typical calibration gamma curve) results in brighter midtones (red curve). Changes of the same 10 RGB points will be perceptually different in shadows, midtones and highlights. For typical calibration gamma curves (2.2, 1.8, sRGB), the shape of the curve in shadows (orange rectangle) will vary depending on monitor black level and on how the calibration algorithm treats the curve in dark areas. Both the linear and typical gamma curves will usually have non-zero luminance at the black while only typical gamma curves may have a higher contrast in that region. Clear distinguishing feature is a convex character of the typical gamma curve in region from about R'G'B'=80 to R'G'B'=200. Real experimental L-star curves are nearly linear.

Before we continue with more detailed analysis, some assumptions have to be made. As we have [discussed earlier](http://www.marcelpatek.com/gamma.html#goto1), LCD display has no underlying physics to follow the power law dependence of luminance vs. digital input. However, we also pointed to the fact that manufacturers build corrections into the display circuitry of the LCD panel to approximate the power function dependence. Such corrections may justify use of the gamma concept. Thus, if the power law dependence is adopted, logarithm of Yn vs. RGBn will be a linear function. Indeed, we often see very good linearity at higher luminance values (~ RGB > 160) as shown earlier in **Fig. 19**. Unfortunately, for midtones and shadows, this linear relationship (i.e. constant gamma) breaks down. To ensure high accuracy, examples in this section are based on calibrations of higher end Eizo CG19 display using ColorNavigator. Calibration software adjusts only the monitor 10-bit LUT without adding any corrections to the 8-bit video card LUT.

[(top)↑](http://www.marcelpatek.com/gamma.html#top)

Let's evaluate four different approaches to characterization of TRC curves.

1. First approach assumes that the power law dependence is obeyed for both curves (gamma-based and the L-star) and that gamma can be expressed as a power coefficient in general formula (a\*xγ + c) or as the slope in the log-log plot in any part of the RGB brightness scale (R=G=B=0 -> R=G=B=255).
2. Second approach to TRC analysis is based on polynomial or spline fit to Yn vs. RGBn data points followed by analysis of the fitted function. While this approach seems the most rigorous, it is at the same time the least intuitive and generizable. We would be looking for parameters that uniquely describe the TRC, such as characteristic points and shapes of the n-th derivative of the fitted function (which is still polynomial function). When the same polynomial fit (I used 6th degree polynomial) is done on the log(Yn) vs. log(RGBn) scale, situation is very close to analyzing any log-log diagram such as in **Fig. 19**. Since the log-log plot should be mostly linear (at least in highlights), the tangent line to it at any point gives the best linear approximation to our fit function. Hence the first derivative of the fit function would give us the slope of the linear portion of the fitted curve, i.e. the gamma. This is of course a simplification, though good enough to analyze gamma and L-star curves.
3. Third method is based on comparison against a constant (reference) initial curve such as the one used in gamma encoded sRGB image. This is a reasonable starting point considering that digital cameras frequently transform raw images into the sRGB (or AdobeRGB) color space. Any display TRC curve would then be canceling the encoding gamma curve to provide the resulting TRC. Assuming LUT gamma = 1, resulting TRC would be very close to the real overall TRC curve (approximately linear).
4. Last method illustrates very simple and reliable test for gamma or L-star curves. Plot of Yn vs. RGBn is generated and CIELAB L\*- lightness values are calculated from the measured Yn. Linear regression is performed on the newly created function: L\* = f(RGBn). L-star curves are supposed to simulate response of the CIELAB L\*-channel while the gamma curves are not. Thus fit to the L-star curve would ideally be a line with very small SSE value (sum of squares due to error). On the other hand, gamma curves (after the same Y to L\* transformation) should result in poor linear fit with higher SSE. Analysis of residuals provides additional information on goodness of the fit. This approach is qualitatively depicted in **Fig. 24**.

**Method 1:**

Two types of data were analyzed - experimentally obtained data for both the L-star and the gamma 2.2. calibrated monitor (**Tables 3, 5**) and ideal color space profile data for both types of curves (**Tables 2, 4**). For both the sRGB and L-star profiles, relative luminance data (Y) were obtained from the icclu utility of [Argyll CMS](http://www.marcelpatek.com/argyll.html) . It is the same data and curves as in **Fig. 20**. More specifically, a grayscale RGB ramp (in increments of 5) was used as the input, run through the icc profile to get the corresponding XYZ tristimulus data. Y-component of the tristimulus output and the input RGB data were used in the curve fitting examples. The digital input range (0-255) was divided into smaller subranges as shown in **Tables 2** to **5** (column 1). Data points in each subrange were fitted to a general power function and the power coefficient γ was recorded in column 2. Root mean squared error (the square root of the mean square error or

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2** . Analysis of L-star icc profile curve (power fit, log-log(γ)) | | | | |
| **RGB (8-bit)** | **a\*xγ+c** | **RMSE** | **log-log(γ)** | **RMSE** |
| 0-255 | 2.53 | 3.50E-3 | 2.55 (>210) | 3.13E-4 |
| 200-255 | 2.69 | 3.95E-5 | 2.55 | 3.89E-4 |
| 150-200 | 2.62 | 2.48E-5 | 2.44 | 7.50E-4 |
| 100-150 | 2.49 | 4.67E-5 | 2.27 | 1.80E-3 |
| 50-100 | 2.25 | 8.27E-5 | 1.95 | 6.00E-3 |
| 0-50 | 1.56 | 4.07E-4 | 1.14 | 3.70E-2 |
| 0-25 | 1.00 | 6.39E-5 | 1.00 | 2.40E-3 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3** . Analysis of L-star display curve (power fit, log-log(γ)) | | | | |
| **RGB (8-bit)** | **a\*xγ+c** | **RMSE** | **log-log(γ)** | **RMSE** |
| 0-255 | 2.56 | 5.69E-3 | 2.63 (>210) | 1.20E-3 |
| 200-255 | 3.05 | 7.39E-4 | 2.60 | 1.61E-3 |
| 150-200 | 2.67 | 4.79E-4 | 2.39 | 1.54E-3 |
| 100-150 | 2.44 | 3.20E-4 | 2.18 | 1.75E-3 |
| 50-100 | 2.15 | 1.79E-4 | 1.82 | 4.76E-3 |
| 0-50 | 1.62 | 3.34E-4 | 0.89 | 6.20E-2 |
| 0-25 | 1.16 | 1.06E-4 | 0.69 | 3.36E-2 |

the standard error) is shown in column 3. Linear regression was performed on the log-log scale for the same data points and slope of the fitted line (γ) is listed in column 4. This would correspond to the so called log-log gamma. The corresponding RMSE data are shown in column 5. Gamma values calculated in columns 2 and 4 will obviously be different. It should not matter which definition of gamma we choose as long as we stay consistent. Remember, the log-log(γ) is the definition used in display calibration where linear regression is done in the logarithmic domain.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4** . Analysis of sRGB icc profile curve (power fit, log-log(γ)) | | | | |
| **RGB (8-bit)** | **a\*xγ+c** | **RMSE** | **log-log(γ)** | **RMSE** |
| 0-255 | 2.25 | 1.55E-3 | 2.24 (>160) | 5.94E-4 |
| 200-255 | 2.32 | 1.10E-5 | 2.26 | 1.62E-4 |
| 150-200 | 2.29 | 1.03E-5 | 2.22 | 2.92E-4 |
| 100-150 | 2.25 | 1.80E-5 | 2.15 | 7.70E-4 |
| 50-100 | 2.17 | 2.96E-5 | 2.02 | 3.38E-3 |
| 0-50 | 1.87 | 1.87E-3 | 1.37 | 5.78E-2 |
| 0-25 | 1.07 | 4.80E-4 | 1.14 | 2.86E-2 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 5** . Analysis of display gamma curve (power fit, log-log(γ)) | | | | |
| **RGB (8-bit)** | **a\*xγ+c** | **RMSE** | **log-log(γ)** | **RMSE** |
| 0-255 | 2.26 | 1.53E-3 | 2.21 (>160) | 9.29E-3 |
| 200-255 | 2.32 | 1.36E-5 | 2.23 | 2.41E-4 |
| 150-200 | 2.30 | 7.88E-6 | 2.17 | 3.83E-4 |
| 100-150 | 2.26 | 1.40E-5 | 2.05 | 1.49E-4 |
| 50-100 | 2.17 | 5.12E-5 | 1.76 | 5.72E-3 |
| 0-50 | 1.84 | 1.65E-4 | 0.56 | 6.48E-2 |
| 0-25 | 1.38 | 2.06E-4 | 0.37 | 3.68E-2 |

Following are some observations made from **Tables 2-5**:

1. For the L-star curves, trends in γ values are approximately the same for both the power function and the "log-log" definitions. They decrease gradually with decreasing RGB input. Values in yellow indicate poor fit to the data. Log-log scale shows particularly poor linear fit in shadows.
2. L-star curves have variable γ across the whole RGB input range. The log-log(γ) fit is about 2.5-2.6 when maximum linear part is considered (RGB > 210). In RGB range of <50-100>, the L-star gamma is about 1.8-1.9, in RGB <100-150> ~ 2.2-2.3, in RGB <150-200> ~ 2.4, and RGB <200-255> ~ 2.6. Closer to the black point, log-log(γ) still decreases with linear fit getting very poor.
3. For the "classical' gamma based curves, trends in γ values are very similar for both the power function and the "log-log" definitions. Log-log(γ) falls off faster for the display curves.
4. The log-log(γ) of gamma curves is about 2.2 when maximum linear part is considered (RGB > 160). The log-log(γ) stays around 2.0-2.2 from highlights to midtones. Closer to the black point, γ decreases with log-log linear fit getting again very poor.

**Method 2:**

To further assess differences between the two calibration curves, relative luminance values (Y ∈ <0-100>) were either measured or calculated using the icclu CMS utility. Polynomial fit was done on the data (linear and log-log scale) followed by the plot and data analysis. In general, the first derivative of the fitted function (logYn vs. RGB or logYn vs. logRGBn ) reflects well the type of the TRC curve. Plots of d(logYn)/d(logRGBn) vs. RGB and d(logYn)/d(logRGBn) vs. logRGBn are shown in **Fig. 22**. L-star curve has extensive linear parts in the logRGBn plot and continuously decreasing "gamma" in the RGB plot. On the other hand, the plot of "classical" gamma curve has a typical S-shape in the logRGBn plot while the RGB plot shows more constant gamma in highlights. However, as indicated earlier, more data is needed to formulate any general characteristics from these plots. Due to the larger size of the images, the corresponding analysis is available in [this document (v. 1.0)](http://www.marcelpatek.com/download/L-star_gamma.pdf) (check here for the latest version).

| **Figure 22: 1st derivatives of logYn vs. RGB or logRGBn for L-star and gamma TRCs** |
| --- |
| g_eval_deriv.m |

**Method 3:**

This method provides only a qualitative comparison between two or more TRC curves. All are subtracted from a reference TRC (in our case the ideal camera sRGB TRC curve). When display TRC curve is also sRGB, a straight line from 0 to 1 will

| **Figure 23: L-curves subtracted from camera sRGB γ** |
| --- |
| **(place mouse over the image to toggle)** |
| [gamma_L](http://www.marcelpatek.com/gamma.html) |

result from the subtraction. This is the case shown in **Fig. 23** (black diagonal line). Other three lines are TRCs of the L-star curves relative to sRGB curve. Blue line shows an L-star curve coming from the L-star icc profile, the red line is an L-star curve coming from the icc profile of calibrated/profiled monitor, and the orange line is the experimentally measured TRC based on the same icc profile. Measurement was done directly off the screen using 5-step gray ramp. In general, the L-star curves make midtones slightly darker than the typical gamma based TRCs. We have already seen the same trend in **Fig. 21**. For more detailed inspection of these curves, place mouse over the image in **Fig. 23**. On the toggled image, differences from the sRGB TRC are exaggerated four times. As one can see, the major differences from "classical" gamma curves are in midtones, specifically in R'G'B' ranges of 25-100 (8-bit encoding). It is mostly the curve steepness (contrast) that differentiates the curves from each other.

**Method 4:**

The relationship between measured tristimulus Y-values (normalized) and RGB normalized values for a 5-step gray ramp is plotted in **Fig. 24** as Yn vs. RGBn. Left panel shows data obtained from L-star calibrated Eizo CG19 monitor, the right panel shows data obtained from similar gamma 2.2 calibration both using ColorNavigator and the X-rite DTP-94 colorimeter. In both cases, the Y-values were also transformed into L\* values of the CIELAB color space.

Here is the formula for Y to L\* transformation: **L\*** = 116 . (Y/100)1/3 - 16 for Y/100 > 0.008856, Y ∈ <0-100> and **L\*** = 903.3 . Y/100 for Y/100 < 0.008856 Least squares fitting method was used to fit a line to the L\* vs. RGBn function. Fitted straight line shown in the right panel (in black) is nearly identical to the original L\* vs. RGBn function. On the other hand, the linear regression performed on the "classical' gamma curve (left panel) shows hints of deviation from the straight line. The corresponding sums of squares due to error (SSEs) are 0.0003 (L-star) and 0.0010 (gamma 2.2.). Analysis of residuals reveals further details of the fit. While residuals for the L-star curve have generally monotonic concave character, residuals of the gamma curve have a characteristic convex shape typical for other measured or theoretical gamma curves. Due to the uncertainty of [curve behavior in the dark areas](http://www.marcelpatek.com/gamma.html#rectangle), evaluate only parts of the plot starting from about R'G'B'=50 (RGBn=0.2) and up. [Here](http://www.marcelpatek.com/download/gamma_vs_L.xls) is the Excel worksheet that performs all the calculations.

| **Figure 24: Linear fit to L\* vs. RGB for gamma and L-star TRCs with plotted residuals** |
| --- |
| g_eval_deriv_nolog.m |

### [↑](http://www.marcelpatek.com/gamma.html#top)Links and References:

**General references:**

* H.R. Kang, Computational Color Technology, SPIE Press, Bellingham, Washington USA (2006)
* B.A. Wandell, L.D. Silverstein, [Digital color reproduction](http://white.stanford.edu/~brian/papers/pdc/OsaHandBookChapter.pdf), The Science of Color, pp.281-316 (2003)
* G. Wyszecki, W.S. Stiles, Color Science, Concepts and Methods, Quantitative Data and Formulae (2nd Edition), (1982)
* M. A. Karim, Ed., Electro-Optical Displays. New York: Marcel Dekker, 1992.

**:**

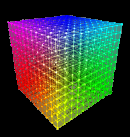
* [A comparison of four multimedia RGB spaces - pdf (Danny Pascale)](http://www.babelcolor.com/download/A%20comparison%20of%20four%20multimedia%20RGB%20spaces.pdf)
* [Monitor calibration and gamma (Norman Koren)](http://www.normankoren.com/makingfineprints1A.html)
* [Gernot Hoffmann](http://www.fho-emden.de/~hoffmann/howww41a.html) (computer vision section - [gamcurve26082001.pdf](http://www.fho-emden.de/~hoffmann/gamcurve26082001.pdf), [gamquest18102001.pdf](http://www.fho-emden.de/~hoffmann/gamquest18102001.pdf), [measgamma10022004.pdf,](http://www.fho-emden.de/~hoffmann/measgamma10022004.pdf) [optigray06102001.pdf](http://www.fho-emden.de/~hoffmann/optigray06102001.pdf))
* [Gamma correction (Wikipedia)](http://en.wikipedia.org/wiki/Gamma_correction)
* [Calibrating monitor gamma (Tom Niemann)](http://www.epaperpress.com/monitorcal/index.html)
* [Computer Graphics Systems Development (CGSD) on Gamma](http://www.cgsd.com/papers/gamma.html)
* [Gamma Correction and Precision Color at libpng.org](http://www.libpng.org/pub/png/book/chapter10.html)
* [L\* gamma](http://homepage.mac.com/hanspeterharpf/LStar-RGB/FileSharing36.html)
* [Links browser Gamma Calibration page](http://links.twibright.com/calibration.html)
* [Monitor Calibration Wizard](http://www.hex2bit.com/products/product_mcw.asp)
* [Advanced Gamma Corrector](http://www.fsc-soft.com/agc.htm)
* [Epaperpress.com](http://epaperpress.com/monitorcal/index.html) (monitor calibration, gamma check)

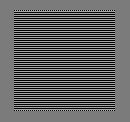
**General:**

* [Emission spectra of elements (I.N. Galidakis)](http://ioannis.virtualcomposer2000.com/spectroscope/elements.html)
* [Monitor Calibration Help - Click Here!](http://www.momentskept.com/MonitorCalibration.htm)
* [Dry Creek Photo - Monitor Black Point Check](http://www.drycreekphoto.com/Learn/Calibration/monitor_black.htm)
* [Akvis - Calibration Theory](http://akvis.com/en/articles/monitor-calibration/calibration-theory.php)

**Links ICC:**

* [Andrew Shepherd's ICC Profile Toolkit - use to change the name of icc profile](http://www.tlbtlb.com/links/)
* [Introduction to digital colour management](http://www.999inks.co.uk/introduction-to-digital-colour-management.html) (collection of links to different topics)
* [iccView (Tobias Huneke)](http://www.iccview.de/index_eng.htm)



[](http://www.tsi.enst.fr/~brettel/TESTS/Gamma/Gamma.html)

