A Little Theory

Color Management

Introduction

If you are not just interested in the practical side of Color Management but also wish to know more about the theories behind color reproduction, then this booklet will provide you with some background information.

It will explain to you in a concise, yet detailed manner the various aspects concerning color reproduction, i.e. the principles of color and what goes on physically, additive color mixing, the different color spaces and color space transformation in its mathematical context.

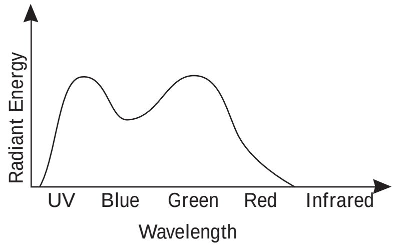
Even if you never advance beyond this introduction in the booklet, you can still work perfectly well with your Color Management system.

However, if you find that the theory which this system is based on is something which attracts to you, you should not just put this booklet aside.

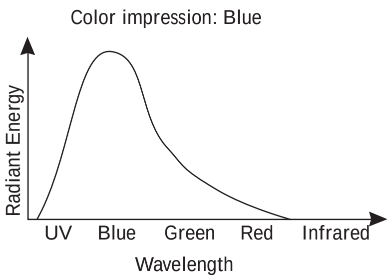
Have fun reading this booklet!

Color as a physical variable

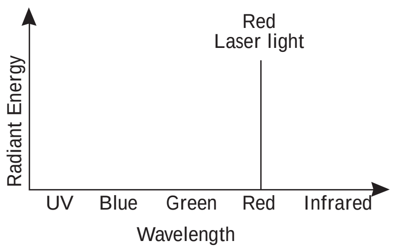
The physical aspect of color is represented by light of different spectral composition. If you record a spectrogram of the color radiation by means of a spectrophotometer you get the color stimulus function as a result of this measurement. This function shows how the radiant energy of the light from a color sample is distributed over the entire visible spectrum varying according to its wavelength.



The highest value of the color stimulus function provides information on the hue. In the following diagram, a blue hue is suggested:



The steeper the rise and fall of the curve in the proximity of the maximum value, the purer (more chromatic) the color. One example of an extremely pure color is laser light. This depicts the color stimulus function of a spectral line:



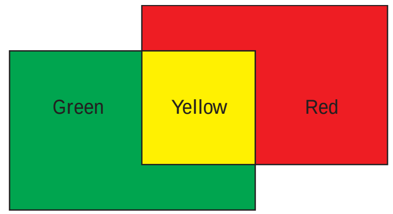
Color as a sensory perception

Man's color vision consists of the analysis, evaluation and encoding of information which the eye receives from such color stimuli. Only after subsequent processing of the stimuli does this information become a color sensation. In this context, the eye can be regarded as a reproduction system. The retina analyzes the light reflected from the objects into three spectral components – namely red, green and blue. The information this provides is, according to the very latest research results, encoded into one element of information defining the brightness and two defining the chroma. This will be examined later in greater detail.

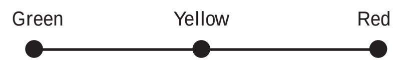
First of all, however, additive color mixing will be described.

Additive color mixing

If we employ additive color mixing to mix red and green light (e.g. by projecting red- and green-filtered light from two slide projectors) we get yellow:

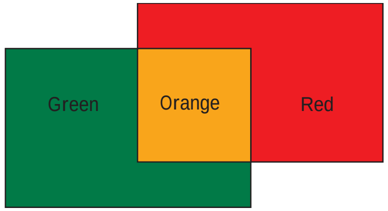


We can represent this mixing process in colorimetric terms by means of a small diagram:



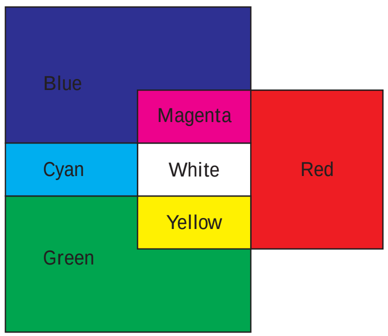
Red and green are two primary colors. Yellow, on the other hand, is a mixed color. In our diagram, the mixed color – consisting of equal components of the primary colors – lies at the exact center of the connecting lines. By stopping down the lens in one of the projectors (and thus changing the intensity of one primary color) it is possible to influence the hue of the mixed color.

A reduction in green while maintaining the intensity of red, for example, will produce orange:

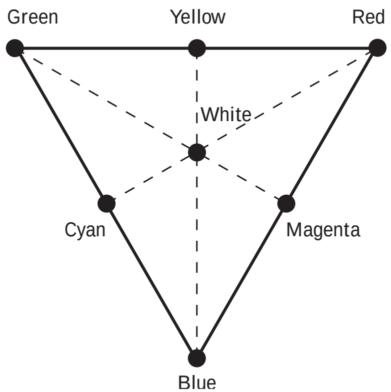


The color locus for the mixed color now moves along the straight connecting line towards the primary color red. The original hue of the yellow has thus changed in the direction of orange:

In order to mix a color like cyan, a third primary color (blue) is needed since mixing the primary colors red and green cannot produce cyan. Three primary colors are all that is needed to mix most colors:



The diagram now assumes the form of a color triangle:



This color triangle shows mixed colors cyan, magenta and yellow in the middle of the straight connecting lines between the primary colors red, green and blue. According to the law of additive color mixing, changes in hue – seen in colorimetric terms – thus mean the movement of color loci along the sides of the color triangle.

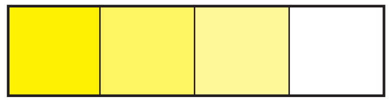
Color spaces

When observing a color, there are three characteristics important to us:

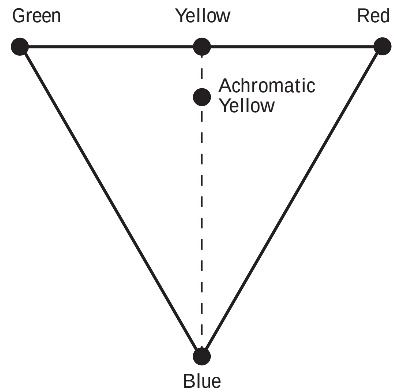
- the hue
- the chroma
- the brightness (or luminance)

The term "hue" refers to the basic color of an object such as green or red. When we look at something, it is the first criterium we use to discriminate colors.

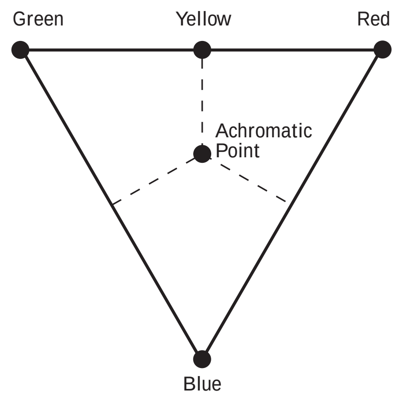
The term "chroma" describes the purity of colors. If blue is gradually added to a yellow which has been mixed from red and green this gives rise to yellow steps of decreasing purity. These are then less chromatic:



Colors of different chroma retain their original hue since the relationship between the color values of red and green has not changed. In the color triangle, they move along the straight connecting line for example from the yellow color locus towards blue:

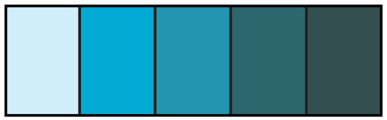


Increasing the amount of the third primary color until all three primary colors are present in equal components will give white. The chroma level is then equal to zero. The achromatic point lies in the middle of the color triangle:

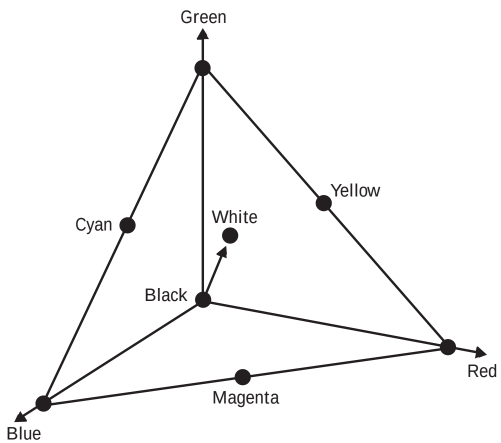


All other colors which can be produced by additive mixing of the three primary colors red, green and blue lie in the area enclosed by this color triangle. The further away they are from the center of the triangle the higher their chroma. A mixed color has a high chroma level if it has only a small amount of its third component or indeed none at all. A maximum of chroma is found in all colors mixed from only two primary colors.

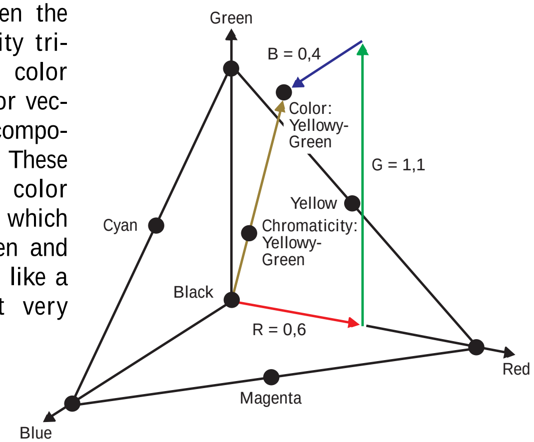
If all three color components of a combination consisting of three primary colors are reduced simultaneously while retaining their mixing ratio, the hue remains unchanged. The color decreases in brightness, however. If the components of all three primary colors are reduced to zero, the resulting color will be black. Like white, black has a chroma level of zero.



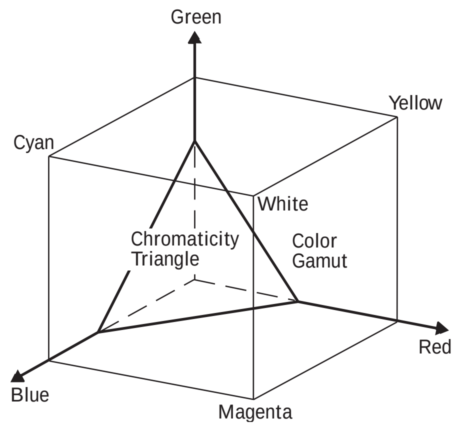
In the color triangle we were able to define both hue and chroma. The colors are thus expressed in terms of their chromaticity and form a chromaticity triangle. All colors in the chromaticity triangle are defined by their hue and chroma only and not by their brightness. There can be any amount of brightness in the chromaticity triangle. In order to integrate brightness into our diagram we need to turn the two-dimensional chromaticity triangle into a spatial body known as a color space. The color space is a three-dimensional coordinate system with coordinates for red, green and blue:



The triangle drawn between the coordinates is the chromaticity triangle. A color locus in the color space is defined by three color vectors which represent the components of the primary colors. These components are known as color values. A color, for example, which consists of 0.6 red, 1.1 green and 0.4 blue, would be something like a yellowy-green which is not very chromatic:



The point of intersection of the resulting vector represents the chromaticity "yellowy-green", the end point of the vector being the color yellowy-green with the inclusion of its brightness value.



The further the color loci of the primary colors from the origin, the greater the volume of the cuboid color gamut which is formed and thus the higher the quality of any color reproduction system which is based on it.

All colors lying inside the color gamut can be reproduced by a reproduction system which is based on these primary colors (for example a color monitor). Colors outside the color space cannot be reproduced. The primary colors of a color space are determined essentially by the equipment which generates them.

The CIE color system

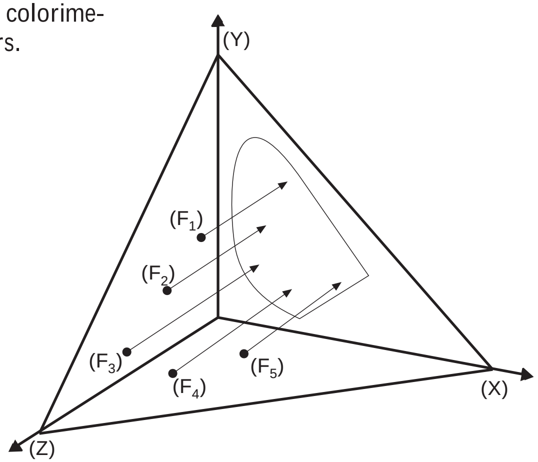
The creation of a global color standard which embraces the most important color gamuts and allows colors to be communicated is of great importance for the development of a color reproduction system.

The CIE color system is such a standard.

This CIE color standard is based on the imaginary primary colors XYZ which cannot be realized physically. They have been generated on a purely theoretical basis and are thus independent of device-dependent color gamuts such as RGB or CMYK. These virtual primary colors have, however, been selected so that all colors which can be perceived by the human eye lie within this color space.

The XYZ system is based on the response curves of the eye's three color receptors. Since these differ slightly from person to person, the CIE has defined a "standard observer" whose spectral response corresponds more or less to the average response of the population. This approach objectifies the colorimetric determination of colors.

The three primary colors of the CIE XYZ reference system initially call for a spatial model with coordinates (X), (Y) and (Z). Here, too, we can also draw a chromaticity triangle. To arrive at a two-dimensional diagram (the shoesole or horse shoe), this chromaticity triangle is projected into the red-green plane:



This is only meaningful, however, if appropriate standardization is performed at the same time which allows the lost value (Z) to be read from the new two-dimensional model. This standardization is achieved by introducing the chromaticity coordinates x, y and z. The following are defined:

$x = X / (X + Y + Z)$
$y = Y / (X + Y + Z)$
$z = Z / (X + Y + Z)$

where:

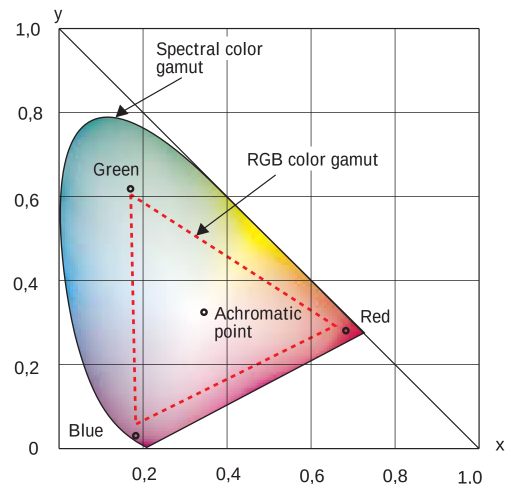
$x + y + z = 1$

The value z of any desired color can be obtained by subtracting the chromaticity coordinates x and y from 1:

$1 - x - y = z$

A color is not defined fully by its chromaticity (x and y). A brightness coefficient also needs to be specified. The eye response curve for green is standardized in the XYZ system so that it simultaneously reflects the sensation of brightness. It is thus identical to the V (l) curve. A color is described in full if it contains the values x and y plus the brightness coefficient Y.

In the standard color triangle, the right-angled chromaticity triangle drawn between zero, x = 1 and y =1 represents the boundaries of this reference system. Chromaticities cannot lie outside the triangle. The closed curve section represents the position of the spectral colors:



While it is possible to define colors between the triangle and spectral color gamut, they cannot be realized but on virtual basis, i.e. not physically. The primary colors RGB of a reproduction device, e.g. a color monitor, form a triangle within the spectral color gamut. Such a triangle then represents a relatively smaller color gamut with the achromatic point more or less in the center.

The introduction of the CIE color system has made it possible to transform color determination from a quality-describing process (bright red) into a process which can be expressed in quantitative and numerical terms.

In addition to the quantitative judgement it allows, the CIE color system also permits the results of additive color mixing to be presented in simple form. The results always lie on straight lines between the colors being mixed. The CIE standard also allows any desired color transformations from one color gamut to another. For example, the transformation of a given color from the RGB color gamut of a monitor to the CMYK gamut of a printing process is facilitated by this standard.

However, the CIE chromaticity diagram not only has advantages, it also has some drawbacks:

- brightness is difficult to include in this diagram
- there is a discrepancy between what we perceive as differences in color and the actual spacing of color in the system.

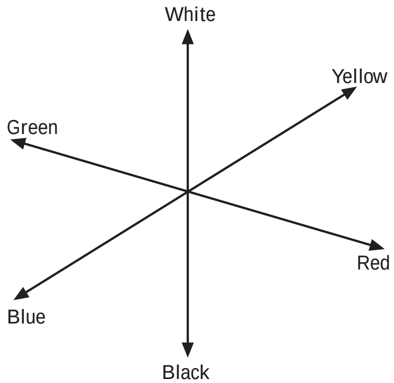
This brings us to the Lab color space.

The CIE Lab color space

Color vision is more complex than merely combining color values in the eye. While the retina does at first register three color stimuli which relate essentially to red, green and blue light rays, it is not until a further processing stage that three sensations are generated:

- a red-green sensation
- a yellow-blue sensation
- a brightness sensation

This can be used to develop a system which is known as the complementary color system. It is based on the differences of three elementary color pairs: red-green, yellow-blue and black-white.

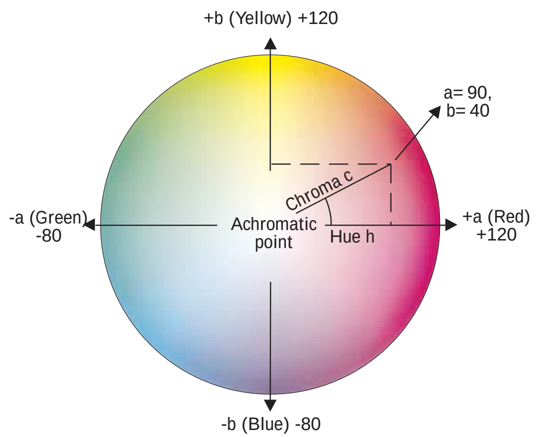


We know from our own experience that red can never contain green components, blue cannot contain yellow components and white can never contain black. When asked about the primaries, people with no knowledge of printing or the color monitor industry will not name three colors such as red, green and blue, but rather four – red, green, blue and yellow.

If one examines colors such as black, gray or white, these are acknowledged as colors only with some reluctance (if at all) since they appear to be of an entirely different quality in the sensations they produce. The absence of chrominance in a black-and-white movie on a TV screen, for example, is something we accept fully after a brief process of mental adjustment. It follows from this that in a reference system which has been designed correctly in sensational terms the achromatic brightness information and the color information should be clearly separated not only quantitatively but also qualitatively.

This is precisely what the Lab color space (which was developed by the CIE in 1976) does. On the one hand, this system is based on the XYZ primary colors, but yet at the same time it also includes the complementary color model described above.

Hue and chroma are defined by the coordinates a and b which can have both positive and negative values:



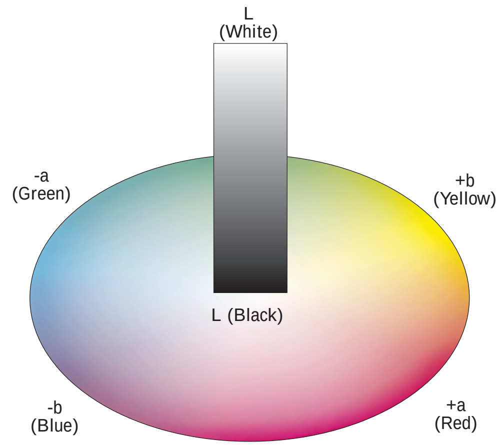
As with the standard color triangle this color system represents all conceivable colors.

Numerical values for chroma and hue can be derived from a and b:

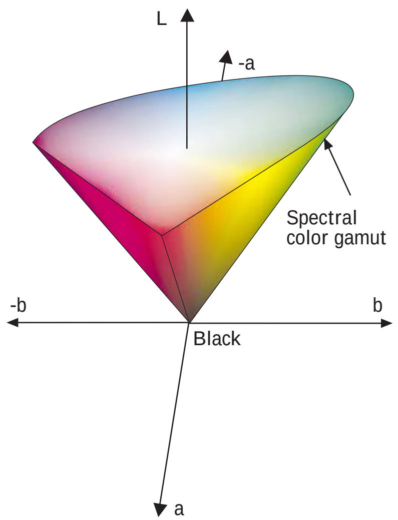
Hue: h = arctan (b/a) (This corresponds to the angle between the color vector and the +a axis)

Chroma: $c = (a^2 + b^2)^{1/2}$ – (This corresponds to the distance between the color locus and the mid-point)

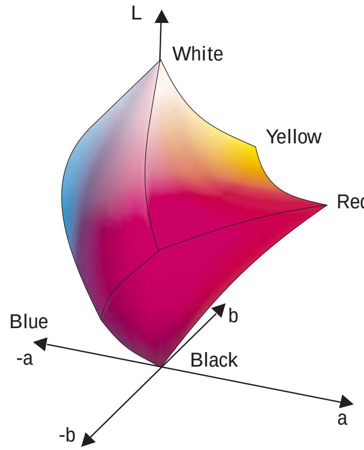
The third characteristic, brightness, is represented vertically by means of a brightness scale designated L with scale values ranging from 0 (black) to 100 (white).



A color gamut in the Lab reference system could appear as follows in an idealized form:



A color gamut based on real colors could then have the following form for example:



For the sake of clarity, we have chosen not to show the different brightnesses of the spectral color curve in their entirety. The model is delimited at the top by a horizontal section. On the outer surface of this ideal color gamut lie all colors of maximum chroma. It can be seen quite clearly that as colors become darker they also lose chroma. This seems logical if one considers that, when the minimum brightness value is reached, every color becomes black and the chroma value is thus zero.

We can see two things here:

- As the brightness of the colors increases and decreases, the chroma decreases down as far as zero when white or black is reached.
- In contrast to the CIE color triangle, the connecting lines between the primary colors here are not straight. The reason for this lies in the visual equispacing of colors in the Lab system. This has been achieved through a non-linear transformation of the XYZ values into Lab values.

The formulas for the transformation of XYZ to Lab are based on (for X, Y, Z the values must not be too small):

$$L = 116\, Y^{1/3} - 16$$

$$a = 500\, (X^{1/3} - Y^{1/3})$$

$$b = 200\, (Y^{1/3} - Z^{1/3}),$$

with X, Y and Z standardized to 1.

Advantages of the Lab color space in color reproduction

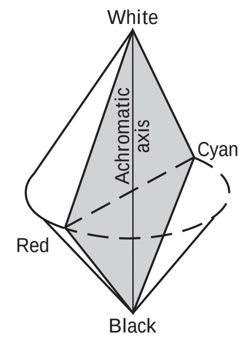
The Lab color space has many advantages. We would like to go into more detail about how the color system does not depend on any particular device and how it is possible to set colors as we perceive them when operating a repro system.

Device independence

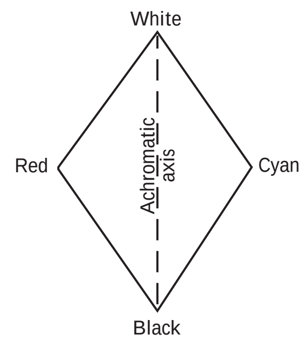
The Lab color space – like the XYZ color space – is able to represent all real color gamuts as subsets.

Let us assume a reproduction system which is based on the RGB color space. The RGB color values need to be converted to CMYK color values for the printing process. The two color spaces coincide in neither size nor location. Due to the fact that the system has RGB as its reference system, it follows that colors of the CMYK color space which cannot be represented in RGB cannot be printed in CMYK either, even though the CMYK color space does not forbid this. RGB acts as a restriction to CMYK. This applies, for example, to a dark and chromatic cyan which cannot be represented on the RGB monitor and, under such circumstances, becomes a non-reproducible color.

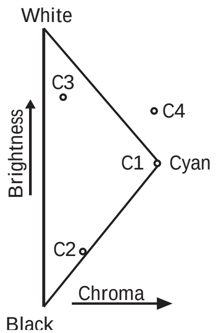
To illustrate this in graphic form, let us take a cross-section through a stylized color gamut:



We can now see the cyan-red plane, for example:



The problem can be shown in simpler form if only one of the two planes is observed – in this case the cyan plane:

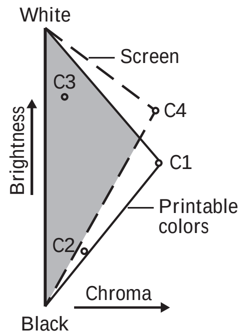


The colors depicted show:

- C1 the cyan with maximum chroma
- is a cyan which has the highest possible chroma for this brightness value
- C3 is a pale achromatic cyan
- C4 is a cyan which lies outside the color space

In the diagram, all the colors have the same hue, namely cyan. It is also possible to reproduce them all except C4 which lies outside the color space.

If the printable colors are included in this figure, you can see that the two color spaces are not identical.



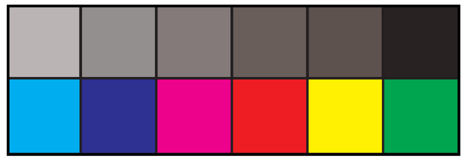
The fact that the two color gamuts overlap each other means that only the colors in a common subset (colored area) are reproduced in identical form both on the monitor and in print.

In a device-dependent reference system like RGB or CMYK, colors lying outside this reference system cannot be reproduced even if they are present in the target color gamut. This is where we can see the advantage of global reference systems such as the XYZ or the Lab color space which are unrestricted in this way.

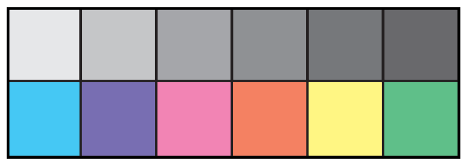
With the help of gamut mapping, color gamuts can be adapted to each other so that the entire range of the target color gamut can be utilized.

Operation based on our perception of color

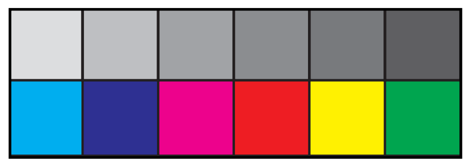
Let us imagine a color transparency which contains very dark achromatic tonal value areas alongside bright colors. We want to brighten the dark image parts while retaining the chromatic colors. In order to illustrate what happens in such a case, let us first examine a combination consisting of a number of very chromatic colors and a grayscale. This is intended to represent an original:



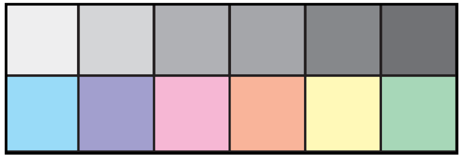
The tonal values of the gray scale are so dark that it seems they will need to be brightened. In a conventional CMYK or RGB repro system, a graduation function would be employed for this purpose. This would influence all the color channels to the same degree.



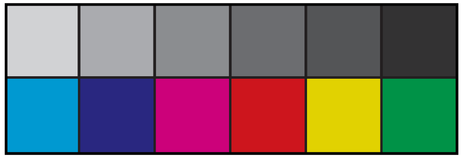
The corrected graduation makes it easier to distinguish the tonal values of the gray scale but the chromatic colors are pale and less chromatic. In order to retain the chroma of the original colors, it would be necessary to perform extensive subsequent correction of each individual color. This disadvantage can be avoided by using a Lab editor which processes the achromatic colors independently of the chromatic ones. Subsequent correction thus becomes superfluous. Increased reliability and savings in time during editing are the result:



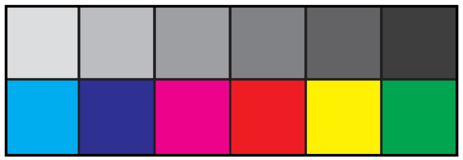
A similar advantage is obtained with over-exposed originals. The colors of such originals are generally too pale and their chroma needs to be improved. Changing the gradation alone will not do the job since the weak colors will merely become darker and thereby become even less chromatic. Using conventional techniques, each individual color would need to be corrected – brightness and chroma would have to be adapted accordingly, a process which is time-consuming and not the best method from the print of view of quality. The fact that luminance and chrominance are separated in the Lab color space means that the chroma of all colors can be enhanced in a single processing stage. A Lab editor is provided with a chroma editing function for precisely this purpose.



Over-exposed original
(Poor chroma)



Conventional reproduction
(Dirty colors, tonal values too dark)



Chroma characteristic corrected
in a Lab editor

Even these brief explanations show the advantages obtained when working with the Lab color space. Other advantages can be seen within the framework of image processing. The separation of brightness and chroma greatly improve the quality of the focus filters. More precise color corrections can be performed as a result which in turn leads to an increase in the quality of the image data.

Make use of these advantages!!!

This booklet is part of the book
"The Creative World of Digital
Data – Color Management"
published by
Heidelberger Druckmaschinen AG
Siemenswall
D-24107 Kiel