A New Approach to the Representation of Color Gamuts

Patrick G. Herzog and Bernhard Hill
Technical Electronics Institute, Aachen University of Technology, Germany

Abstract

A new method is presented that analytically describes the surfaces of color gamuts in a closed form. Extending the method to the whole volume of the gamut, it can be used to analytically represent the relationship between a color space and the space of color control signals of a reproduction device.

Introduction

A typical color reproduction process is controlled by three color control signals at the input. These may be RGB signals for the case of a CRT monitor or CMY colorant concentrations in case of a print process. Controlled by the three color signals at the input, colors are reproduced at the output that can be described in any of the well-known color spaces, e. g. CIELAB, CIELUV or CIEXYZ. The entire range of colors that a reproduction device is able to produce defines the device's color gamut. The color gamuts of reproduction devices differ from one another in most cases. Generally, the gamut of a color image to be rendered does not match the gamut of the supposed reproduction device. The image might contain colors that cannot be generated by the output device, and thus image colors somehow have to be adapted to the color gamut of the device, a procedure called *gamut mapping.*[1,2,3,4,5]

For that, a basic requirement is the knowledge of the color gamuts of both image and device, and that these gamuts have to be represented in any color space. As the user may be confronted with several output devices like offset print, several CRT monitors, photo-realistic printers (thermal dye diffusion printers), and low budget printers (ink printers), a compact gamut representation is highly necessary. In this paper, a new method is proposed that analytically represents the gamut of a color reproduction device in a compact manner using any of the well-known color spaces (e. g. CIELAB, CIELUV or CIEXYZ) as basis.

Principles

Before getting into the method's explanations, some fundamentals and definitions must be given. Each of the three color control signals of an output device can be modified independently between a minimum and a maximum, i. e. between zero and one in normalized form. Thus all color control signals the device is capable of processing are contained in a cube. In view of the important role of this cube for the proposed method, it is called the *kernel gamut* of the color control signals and, consequently, the respective space is called the *kernel space* in the following. The eight corners of the cube control the full- and zero-tone colors and all the integer mixtures of them.

The color gamut normally has a strongly distorted surface since the relationship between the color space and the kernel space is strongly non-linear. However, some significant characteristics of the cubic kernel gamut remain inherent. The color gamut, being a non-linear transform of the kernel gamut, has eight corners, twelve edges and six planes, though edges and planes are somehow distorted. Figure 1 shows the color gamut of a thermal dye diffusion printer.

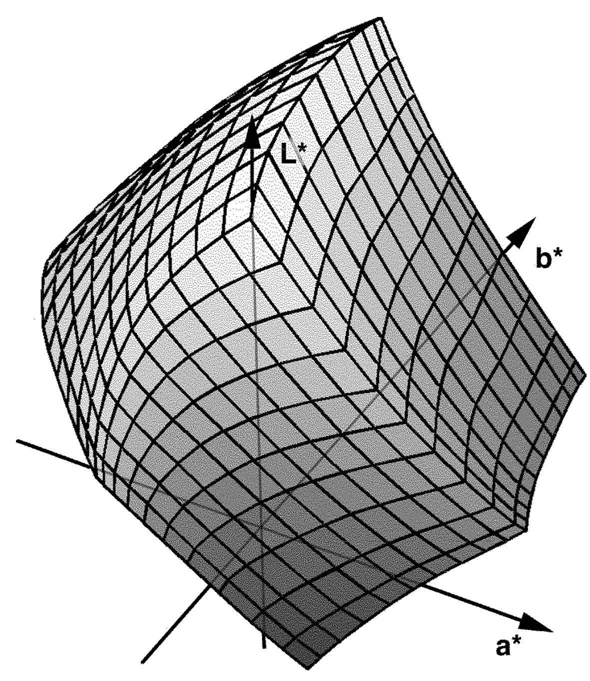

Figure 1. The color gamut of a thermal dye diffusion printer

The method is based on the characteristics of the gamuts that appear in both spaces. The surface of the kernel gamut is mapped onto the color gamut surface in a manner that the eight corners of the kernel gamut come to lie at the corners of the color gamut. Similarly the twelve edges and six planes of the kernel gamut are mapped onto the corresponding edges and planes of the color gamut. This is done by analytically distorting the kernel gamut surface.

Since the representation of a cube (i. e. the kernel gamut) is well-known, the color gamut of the device is represented by these analytical distortion functions.

In the method, a cylindric coordinate system is used. Therefore the coordinates can be interpreted as lightness, chroma and hue if the method is applied to an adequate color space. The method is applicable to all processes based on three primary colors or colorants, including CRT monitors and print processes like ink printing, offset printing etc. Furthermore, four-colorant print processes can also be included if well-defined separation algorithms are used (e. g. GCR, UCR; see for example References 6 and 7). In this case, these processes can be handled like three-colorant processes.

The determination of color gamuts of electronic *images* to be reproduced is still a problem today as non trivial surface-fitting algorithms must be used. Since the gamut representation of the image is given numerically and therefore memory-consuming, storage or transmission of image gamuts in addition to the image data is nowadays out of question. But if the process that generated the original image is known, then it may be helpful to transmit the gamut of the process together with the image, and for that, the analytical method proposed here is well suited because the additional amount of data is negligible. This is possible as well when color images are commercially distributed by e. g. Photo CD from different sources like film transparencies, computer generated images etc.

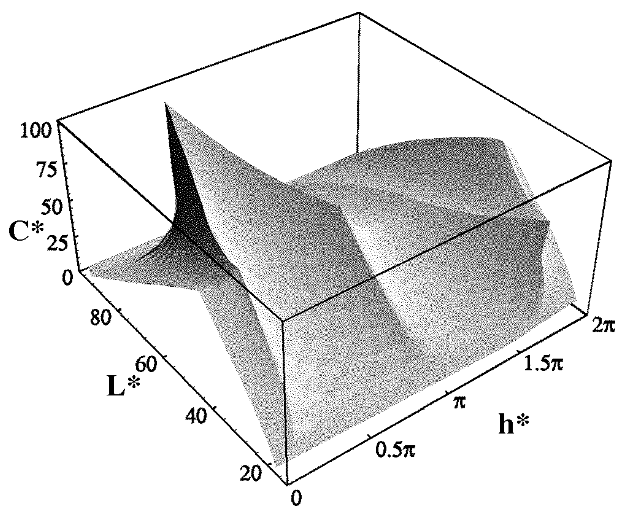

Figure 2. The color gamut of a thermal dye diffusion printer in the two dimensional presentation

The Analytical Representation

In this paper, the color gamut surface is represented by the maximum chroma as a function of lightness and hue:

$$C^* = f(L^*, h^*) \tag{1}$$

Thus, the gamut surface can be shown as chroma mountains over the lightness-hue plane (see Figure 2). Because chroma is a function of two parameters, this presentation is called the two dimensional presentation as opposed to the three dimensional presentation of the gamut in the color space. If the gamut's white point is lying on the L*-axis, then in the two dimensional presentation it is depicted by the straight line $L^* = L^*_{white}$ and, by analogy, the black point is depicted by $L^* = L^*_{black}$. The remaining six vertices possess lightness values in between, and the six edges connecting them define the *middle kink-line*.

The kernel gamut (a regular cube) can be presented similarly (Figure 3).

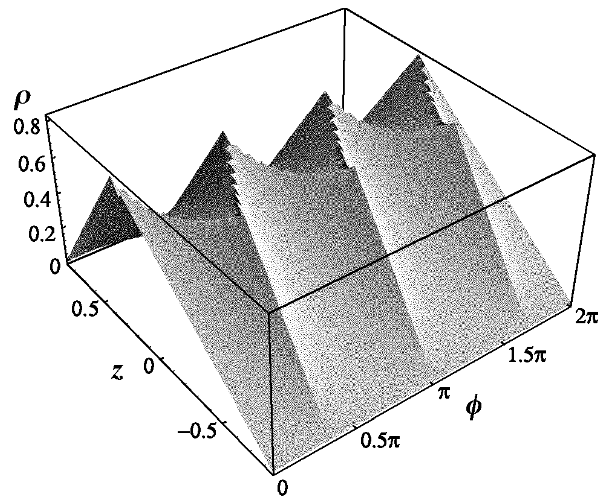

Figure 3. The kernel gamut in the two dimensional presentation

The kernel gamut has now to be distorted in order to match the color gamut.

The surface of a unit cube standing on its corner with its center located at the origin of the mathematical (x,y,z)-space is given by the following equations:

$$\rho(z,\phi) = \rho_k(\phi)\frac{\frac{3}{4} - z\,z_k(\phi) - \frac{\sqrt{3}}{2}\left|z - z_k(\phi)\right|}{\frac{3}{4} - z_k^2(\phi)} \tag{2}$$

with

$$\rho_k(\phi) = \frac{\sqrt{2}}{\left|\sin\phi\right| + \left|\sin\left(\phi - \frac{2\pi}{3}\right)\right| + \left|\sin\left(\phi - \frac{4\pi}{3}\right)\right|} \tag{3}$$

and

$$z_k(\phi) = \frac{\left|\cos\phi - \frac{1}{2}\right| - \left|\cos\phi + \frac{1}{2}\right|}{2\left(\left|\sin\phi\right| + \left|\sin\left(\phi - \frac{2\pi}{3}\right)\right| + \left|\sin\left(\phi - \frac{4\pi}{3}\right)\right|\right)}\cdots$$
$$\cdots\frac{+\left|\cos\left(\phi + \frac{2\pi}{3}\right) - \frac{1}{2}\right| - \left|\cos\left(\phi + \frac{2\pi}{3}\right) + \frac{1}{2}\right|}{\cdots}\cdots \tag{4}$$
$$\cdots\frac{+\left|\cos\left(\phi + \frac{4\pi}{3}\right) - \frac{1}{2}\right| - \left|\cos\left(\phi + \frac{4\pi}{3}\right) + \frac{1}{2}\right|}{\cdots}$$

$\rho_k(\phi)$ is the cylindric radius and $z_k(\phi)$ the z value of the middle kink-line; $\phi = \arg(x+iy)$, $i=\sqrt{-1}$.

The final chroma function $\hat{C}(L^*,h^*)$ is derived from the kernel function $\rho(z,\phi)$ in two steps. In the first step, the kernel function is scaled by multiplication with the scaling function $s(z,\phi)$. This operation transforms into the correct amplitudes of $\hat{C}$, but at coordinates z and ϕ. In the second step therefore, the two distortion functions $z_d(L^*,h^*)$ and $\phi_d(L^*,h^*)$ are introduced to move the transformed amplitude values to the right positions in the two-dimensional (L^*,h^*) plane. To improve the performance of the scaling, a further function $s_a(L^*,h^*)$ is added.

$$\rho_s(z,\phi) = \rho(z,\phi)\, s(z,\phi) \tag{5}$$

$$\hat{C}(L^*,h^*) = \rho_s\big(z_d(L^*,h^*),\phi_d(L^*,h^*)\big) + s_a(L^*,h^*) \tag{6}$$

The color gamut is represented by the distortion functions $z_d(L^*,h^*)$ and $\phi_d(L^*,h^*)$ and the scaling functions $s(z,\phi)$ and $s_a(L^*,h^*)$.

Limitations

Until now it was assumed that the white point and the black point are lying on the L* axis. This is true for the white point if colors are referred to the paper white of the print device. This is the normal procedure if gamut mapping is to be applied.[3,8] But in many cases, the black point is located off the grey axis. Then, all colors having lightnesses lower than the darkest neutral grey cannot be considered by the method at issue. In practice, this loss is not very meaningful. Considering that one would like to avoid inversions when mapping colors which run from dark neutral grey to lighter, colorful colors, one would anyway dispense with utilizing these color regions.

Another limitation is caused by the definition of CIELAB when bright and saturated colors are expressed in CIELAB coordinates. Since colors of the same hue are not located on half-planes of constant hue-*angle* in CIELAB, the relation between chroma and hue-angle is ambiguous in this area and this cannot be compensated by the used method.

Results

For the practical application of the method, the required distortion and scaling functions describing the gamut surface are approximated by limited polynomials or limited series expansions to keep the number of parameters as low as possible.

Given a pair of values (L^*, h^*), the task of the representation formula is to furnish an optimal approximation of the gamut's maximum chroma. Therefore, a visual error is defined solely as a difference of the original and the approximated chroma:

$$\Delta E_{ab} = C^*_{orig} - \hat{C} \tag{7}$$

Given an errouness approximation $\hat{C}$ of the gamut hull for a given pair (L^*, h^*), the nearest point on the actual gamut hull is generally not located in the radial direction. It follows that the distance of the approximated chroma from the gamut hull is actually less than given by (7).

Extensive studies with several output devices (from CRT monitor to offset proof process) have shown that the amount of parameters to represent the gamuts is only about one hundred if mean visual errors are kept below 2.2 ΔE_{ab} units. Thus, a number of about 100 to 200 Bytes is sufficient to describe the surface of a color gamut. These visual errors are in the range of the perceptibility threshold of pictorial images[9,10] and below the acceptability threshold of normal images[10,11]

Since the method is based on the corners and edges of the gamut, these can be representated very accurately. This is important as they contain the most saturated colors the device can produce.

The maximum errors are below 10 ΔE_{ab} units and are located in the regions of the light yellow edge ([CMY] = [001]) and the dark blue edge ([CMY] = [110]). As mentioned before, the visual error is only an error of chroma which always means an error in the radial direction. If the McAdam and the Brown-McAdam ellipses are plotted in the a*b* plane,[12,13] respectively, it can be observed that right in these yellow and blue regions the ellipses are strongly non-uniform and that the main axes extend exactly in the direction of the error, i. e. the radial direction. The same is true for the Wyszecky-Fielder ellipses[13] and can also be seen in the CIE 1931 chromaticity diagram.[14] Also Luo and Rigg[15] show in a compound investigation that ellipses for the yellow and blue centers tend to point along lines of constant dominant wavelength. In other words: Certainly is the numerical value of the maximum error relatively high, but the error is expected to be visually much less noticeable.

Figures 4 and 5 show examples of the distortion functions $z_d(L^*,h^*)$ and $\phi_d(L^*,h^*)$, respectively. These functions turn out to be relatively smooth in practical applications.

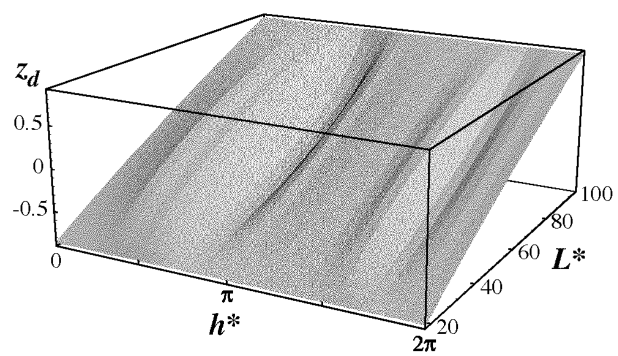

Figure 4. Example of the distortion function $z_d(L^*,h^*)$ for a thermal dye diffusion printer

Extension to the Transformation of Color Spaces

The analytical distortion functions mentioned above can be considered as a manner to analytically represent the transformation between the kernel space and the color space. In fact they describe only the colors contained on the surface of the gamut. The extension to the whole space is carried out by defining subsurfaces e. g. parallel to the outer surface and finding relationships between the subsurfaces described by smooth transitions between the parameters of the analytical representations.

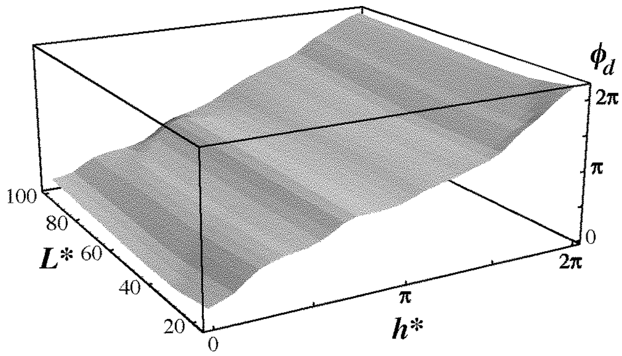

Figure 5. Example of the distortion function $\phi_d(L^*,h^*)$ for a thermal dye diffusion printer

Thus, an analytical formula of the transformation between a color space and a kernel space of a device can be defined to immediately specify the color control signals that lead to the demanded reproduced color. Thereby, calibration of color output devices is greatly simplified and the time-consuming process of inverting color maps of output devices becomes obsolete.

Summary

An analytical method for the representation of color gamuts is presented. It is shown that the method can be extended to analytically represent the relationship between a color space and a device's space of color control signals with a low number of parameters.

The analytical method has the following advantages:

- Representation of edges and corners is especially accurate in opposition to traditional methods like Fast Fourier Transform or Karhunen Loéve Transform.
- The method contains an inherent interpolation since it is analytical.
- No numerical methods like interpolation or inversion are necessary: output devices can be calibrated analytically.
- Any of the known color spaces (e. g. CIELAB) can be used for the representation of color gamuts.
- As the representation is very compact it is excellently suited for storage and transmission of gamuts together with color images.
- No mathematical case discrimination between adjacent planes is necessary as opposed to other analytical methods.
- Gamut mapping can also be carried out analytically in the future.

It becomes clear that the CIELAB space is not the best suited color space for gamut mapping and device calibrations, and the existence of a better color order system is worthwhile.

Acknowledgement

This research was supported by the DFG (Deutsche Forschungsgemeinschaft), Bonn (Germany).

References

1. Maureen C. Stone, William B. Cowan, John C. Beatty, "Color Gamut Mapping and the Printing of Digital Color Images", *ACM Transactions on Graphics*, Vol. **7**, No. 4, October 1988, pp. 406–422.
2. R. S. Gentile, E. Walowit, and J. P. Allebach, "Comparison of Techniques for Color Gamut Mismatch Compensation", *Journal of Imaging Technology*, Vol. **16**, No. 5, October 1990, pp. 176–181.
3. J. A. Viggiano and C. Jeffrey Wang, "A Comparison of Algorithms for Mapping Color Between Media of Differing Luminance Ranges", *TAGA Proceedings*, 1992, pp. 959–974.
4. Bruce J. Lindbloom, "Accurate Color Reproduction for Computer Graphics Applications", *Computer Graphics*, Vol. **23**, No. 3, July 1989, pp. 117–126.
5. Giordano Beretta, "Is Color Appearance Matching necessary?", *SPIE Proceedings*, Vol. **2171**, 1994, pp. 220-227.
6. Henry R. Kang, "Gray Component replacement using Color Mixing Models", *SPIE Proceedings*, vol. **2171**, 1994, pp. 287–296.
7. William F. Schreiber, "A Color Prepress System using Appearance Variables", *Journal of Imaging Technology*, vol. **12**, 1986, pp. 200–211.
8. James Gordon, Richard Holub, and Robert Poe, "On the Rendition of Unprintable Colors", *TAGA Proceedings*, 1987, pp. 186–195.
9. Mike Stokes, Mark D. Fairchild, and Roy S. Berns, "Precision Requirements for Digital Color Reproduction", *ACM Transactions on Graphics*, Vol. **11**, No. 4, October 1992, pp. 406–422.
10. Mike Stokes, Mark D. Fairchild, and Roy S. Berns, "Colorimetrically Quantified Visual Tolerances for Pictorial Images", *TAGA Proceedings*, 1992, pp. 757–777.
11. Scott Stamm, "An Investigation of Color Tolerance", *TAGA Proceedings*, 1981, pp. 156–173.
12. Alan R. Robertson, "The CIE 1976 Color-Difference Formulae", *Color Research and Application*, 1977, vol. **2**, pp. 7–11.
13. M. Mahy, L. Van Eycken, A. Oosterlinck, "Evaluation of Uniform Color Spaces Developed After the Adoption of CIELAB and CIELUV", *Color Research and Application*, 1994, vol. **19**, pp. 105–121.
14. J. Romero, E. Hita, and L. Jiménez del Barco, "Chromaticity Differential Thresholds and Successive Stimuli Presentation", *Die Farbe*, vol. **32/33**, 1985/1986, pp. 264–274.
15. M. R. Luo and B. Rigg, "Chromaticity-Discrimination Ellipses for Surface Colours", *Color Research and Application*, vol. **11**, 1986, pp. 25–42.