Is CIE L*a*b* Good Enough for Desktop Publishing?

E. M. Granger
Light Source, Inc., Larkspur, California

Abstract

The CIE L*a*b* system of color coordinates is not optimal for use in desktop publishing (DTP) systems, because it is non-uniform, not well matched to human visual dynamics, and computationally inconvenient.

CIE L*a*b* uses cube roots of differences of color matching functions. In complex scenes, perceived lightness varies quadratically, not cubically, with intensity. Chroma varies in a more complex way but is also not well represented by cubic polynomials. As a result, CIE L*a*b* exaggerates highlights and compresses shadows. It distorts the concentric circles and radii of the Munsell chart into ellipses and curves. This makes it difficult to achieve appearance equivalence during gamut compression. Gamut compression is an essential element in working with the variety of inexpensive DTP devices.

Users expect DTP operations to proceed rapidly on inexpensive equipment. This dictates heavy reliance on integer arithmetic and lookup tables. These techniques do not mesh well with the computational complexities of CIE L*a*b*.

Introduction

At least as far back as Newton's color wheel in the seventeenth century, scientists have tried to devise schemes for ordering colors. These schemes aim not only at order, but also at providing an idea of distances. A Euclidean metric, however, remains an elusive goal.

The foundation for such work in this century was laid by the artist and teacher A.H. Munsell, who devised a three-dimensional system based on hue, saturation, and lightness. Munsell's systems, first published in 1905, has flaws—notably its use of a fifth primary, purple. Today's opponent theories of color, such as Guth's ATD, suggest that there are four unique primaries: red and green, yellow and blue, with white and black forming the remaining set of unique vision sensations. Indow's survey[2] shows that the Munsell intervals correlate strongly with reported perceptual intervals. For this reason the Munsell system provides a useful norm against which to test proposed theoretical color models.

Since the goal of DTP color is a perceived image, it makes sense to judge the suitability of CIE L*a*b* for DTP by looking at how well its colorimetric scales correspond to the Munsell intervals. This gives an idea of how well these scales can help us to predict and control color reproduction errors. It also provides an indication of whether these scales provide a helpful basis for the gamut-compression computations that are essential to accurate appearance modeling on inexpensive DTP devices.

Another requirement for DTP programs is that they must use computational and storage resources efficiently. Most computer systems use 8-bit bytes as the smallest efficiently manipulated data unit, so any system probably needs to use coordinates that are an integral number of bytes in size. Three one-byte coordinates can represent over 16 million colors, which is clearly an adequate number if used effectively. Loss of precision during computations is a major consideration in dealing with 8-bit color coordinates. We can judge the suitability of CIE L*a*b* for DTP by looking at how well it lends itself to the use of 8-bit coordinates and what kinds of computations it requires us to use.

The body of this paper details the shortcomings of CIE L*a*b* in all of the above areas.

What's Wrong With CIE L*A*B*

In 1978, CIE established and recommended use of two approximately uniform color spaces to serve as standards until something better came along. These standards are the CIE 1976 (L* a* b*)-space and the CIE 1976 (L* u* v*)-space.[3] For purposes of this paper, the first thing to note about both of these spaces is that they assign luminance values by the formula

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16$$

One problem with the cube root is that it forces loss of precision when working with 8-bit quantities. A typical DTP application entails acquisition of an RGB-encoded scanner image, transformation into the color space for manipulation, then transformation into a CMYK form for printing. There is no efficient way to use the full 8-bit ranges for the RGB, color space, and CMYK values if the sets are related by cubic functions. This virtually forces the computations to be done in floating point, which is impractical for DTP.

The placement of a cube root function between a scanner and the color space, then its approximate inverse between the color space and a printer is highly inefficient. It is like two native English speakers who speak Japanese poorly, communicating through a messenger who speaks only Japanese. It should be avoided if possible.

More important than the computational inefficiency is the fact that the cube root law actually models perceived brightness poorly in complex scenes. Bartelson & Breneman[4] showed subjects a variety of complex black and white pictures and asked them to judge the brightness of identified areas of known luminance. Figure 1 shows three plots. The first is a log-log graph of the mean responses for each of the

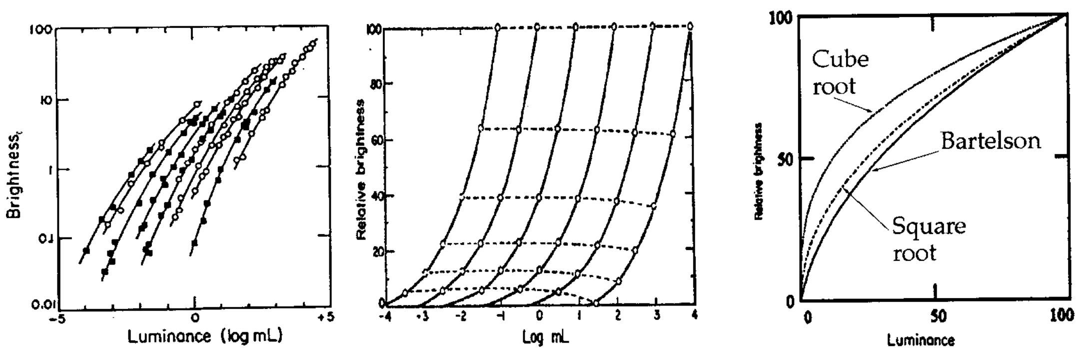


Figure 1. Bartelson & Breneman Data—raw and composite

ten images viewed at different luminance levels. The second plot shows the relative response at several luminance levels. It plots brightness relative to a reference white vs the log of luminance for six levels of luminance. In the third plot we have combined these curves into a single curve and compared it with the curves that would result if the relationship followed a cube root law or a square root law. The Bartelson & Breneman curve resembles the square root curve but differs significantly from the cube root curve.

Since most DTP images qualify as complex scenes, CIE L*a*b* fails to represent lightness correctly. Instead, it wastes much of its limited numerical range expanding the shadows, leaving too few numerical steps in the highlights, where they are needed. This suggests that we need to reexamine the appropriateness of L* as a model of brightness.

The incorrect representation of lightness is common to CIE L*a*b* and CIE L*u*v*. However, CIE L*a*b* has additional distortions in the area of chroma. Kuehni[5] studied this problem by comparing observed color differences with those predicted using CIE L*a*b* coordinates. He based part of his analysis on how CIE L*a*b* distorts the MacAdam ellipses.[6] At any given point in a chromaticity space, observers' sensitivity to changes varies with the direction of the change. MacAdam's ellipses plot this differential sensitivity. A colorimetry system that correlates well with human vision should map these ellipses accurately. Figure 2 shows two plots of the MacAdam ellipses. The first[7] shows the ellipses in CIE (x,y)-chromaticity diagram. The second represents Kuehni's data, which shows how CIE L*a*b* distorts the MacAdam ellipses, particularly in the yellow and red regions.

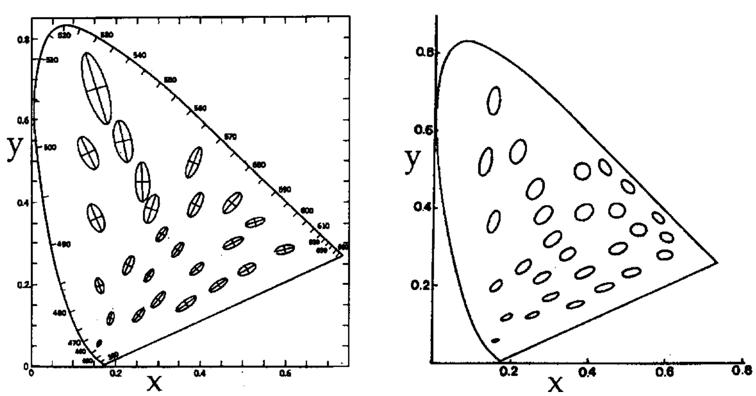


Figure 2. MacAdam Ellipses in CIE (x,y) Chromaticity Diagram and as transformed by CIE Lab

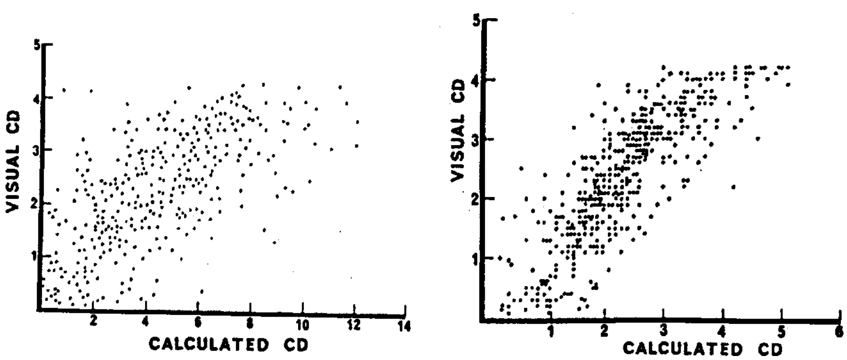


Figure 3. Observed vs predicted. CIE Lab and Best fit MacAdam Ellipses

Kuehni then devised his own replacement for the CIE L*a*b* coordinates. His system preserved the shapes and orientations of the MacAdam ellipses. Figure 3 shows correlation plots of observed vs predicted color differences. The first plot uses CIE L*a*b* for the predictions. The second uses the MacAdam ellipse data for the computed color differences. The substantially better correlation in the second plot shows that CIE L*a*b* distorts and misrepresents chroma. This shows that CIE L*a*b* is not a good space to use for predicting and controlling color errors or for carrying out gamut compression computations. These are both important aspects of DTP.

Figure 4 shows a plot of Munsell data points, at lightness level 5, in the CIE L*a*b* and CIE L*u*v* spaces. In the CIE L*a*b* plot, the points do not lie on concentric circles and the points with equal hue do not lie on straight lines. In this plot the red to yellow to green region has large steps. Indow[8] shows that they should be uniform. This distortion places emphasis on color increments in yellow, not in blue. In fact, the opposite emphasis is required for accurate color reproduction in DTP. Again, CIE L*a*b* is seen to do a poor job of measuring color differences.

While CIE L*a*b* and CIE L*u*v* use the same L* function, CIE L*u*v* actually comes closer to representing the Munsell solid accurately than CIE L*a*b* does, as the plots in Figure 5 show.[9]

The plots show the biggest problem with using CIE L*a*b* in DTP. The first plot shows the structure of the Munsell space. Note that the gamut of usable colors comes to a point at the bottom. There are few differences below value 2 (lightness 0.20). On the other hand, CIE L*a*b* has a broad flat floor. In the range of lightness values of .10 to .20, corresponding to densities 2.0 to 1.4, i.e., full black to 75% dot in graphic arts terms, CIE L*a*b* has a 50 to 75 unit radius. In fact, there are few observable steps in that area. Using the CIE L*a*b* metric to predict and control color reproduction errors in that region would direct valuable scarce computing and storage resources to an area where they are not needed. The CIE L* a*b* metric predicts errors 50 times greater than they are actually perceived.

Since CIE L*a*b* seems so badly matched to the needs of DTP, the question of what to use instead naturally arises. As we noted above, CIE L*u*v* approximates the shape of the Munsell solid much better than CIE L*a*b* does. With a modified definition of L*, it would definitely be superior to CIE L*a*b* in DTP applications.

As companion papers to this one show, there is an even better choice, namely, Guth's ATD space. In our own companion paper we describe a color space based on Guth's space.

Conclusion

CIE L*a*b* is poorly suited to DTP applications in three main areas:

- It correlates poorly with observed color differences.
- The cube root definition of L* does not model the

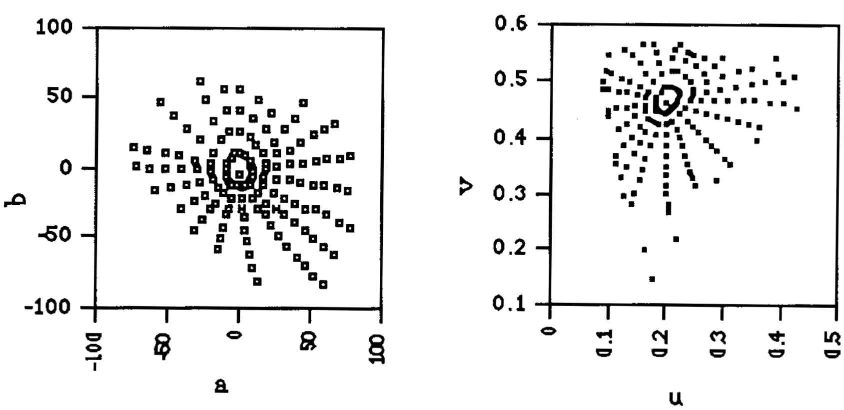
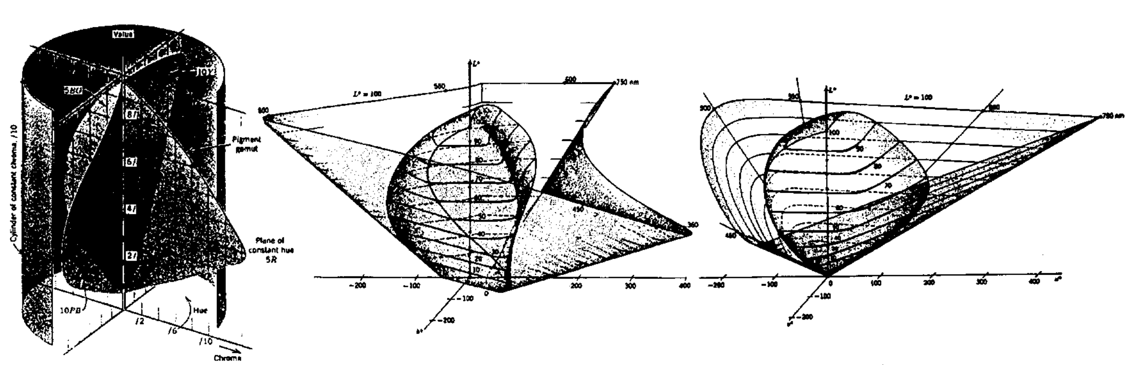


*Figure 4. Munsell data at level 5 transformed by CIE L*a*b* and CIE L*u*v*.*



Munsell color solid | CIE 1976 Lab | CIE 1976 Luv

Figure 5. Comparison of color spaces

perception of brightness in complex scenes; a square root law models human vision better.
- In the dark region (below 50% dot), CIE L*a*b* is a poor predictor of color errors in DTP.

CIE L*a*b*, rather than giving us uniform color space, has opened color tolerancing up to question. It reports large errors where none exist, and it is not uniform in lightness or chroma. CIE L*u*v*, with a modified definition of L*, might provide a better basis for DTP. However, companion papers to this one show that Guth's ATD space, based on a full opponent model of human vision, is an even better choice.

References

1. S.M. Newhall, D. Nickerson, & D.B. Judd, "Final Report of the OSA Subcommittee on the Spacing of the Munsell Colors," *J. Opt. Soc. Am.*, Vol. **33**, pp. 385-418, 1943.
2. T. Indow, "Multidimensional Studies of the Munsell Color Solid," *Psych. Rev.*, Vol **95**, No. 4, pp. 456-70, 1988
3. G. Wyszecki & W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., pp. 164-169, Wiley, NY, 1982
4. J.C. Bartelson & E.J. Breneman, "Brightness Perception in Complex Fields," *J. Opt. Soc. Am.*, Vol. **57**, pp. 953-57, 1967.
5. R.G. Kuehni, "Color-tolerance Data and the Tentative CIE 1976 L*a*b* Formula," *J. Opt. Soc. Am.*, Vol. **66**, No. 5, pp. 497-500, May 1976.
6. D.L. MacAdam, "Visual Sensitivities to Color Differences in Daylight," *J. Opt. Soc. Am.*, Vol. **32**, p. 247, 1942.
7. G. Wyszecki & W.S. Stiles, *op. cit*., p. 308.
8. T. Indow, *op. cit.*
9. G. Wyszecki & W.S. Stiles, *op. cit*., pp. 166-67, 510.

published previously in SPIE, Vol. 2170, page 144