J. J. McCann,

"Color spaces for color mapping",

J. Electronic Imaging, 8, 354-364, 1999.

Journal of Electronic Imaging 8(4), 354–364 (October 1999).

Color spaces for color-gamut mapping

John J. McCann
McCann Imaging
Belmont, Massachusetts 02478
E-mail: mccanns@tiac.net

Abstract. *Before doing extensive color gamut experiments, we wanted to test the uniformity of CIE L*a*b*. This paper shows surprisingly large discrepancies between CIE L*a*b* and isotropic observation-based color spaces, such as Munsell: (1) L*a*b* chroma exaggerate yellows and underestimate blues. (2) The average discrepancy between L*a*b* and ideal is 27%. (3) Chips with identical L*a*b* hue angles are not the same color. L*a*b* introduces errors larger than many gamut mapping corrections. We have isotropic data in the Munsell Book. Computers allow three-dimensional lookup tables to convert instantly any measured L*a*b* to interpolated Munsell Book values. We call this space ML, Ma, and Mb in honor of Munsell. LUTs have been developed for both LabtoMLab and MLabtoLab. With this zero-error, isotropic space we can return our attention to the original problem of color-gamut image processing.* © *1999 SPIE and IS&T.*
[S1017-9909(99)00804-1]

1 Introduction

There is considerable interest in color matching between monitors and printers. The problem is well documented. The color gamut of the print is very different from that of the monitor. The monitor has a wide variety of different, highly saturated, high-value colors near white. Prints, even those using high-extinction coefficient dyes, absorb too much light to generate these colors. Prints, however, have a greater range of colors in the low-lightness browns. The hope is to find automatic calculations of monitor and print colors that make them look more like each other. Before embarking on quantitative evaluations of these ideas, we felt it necessary to study the color spaces used in automatic calculations. The ideal three-dimensional (3D) color space is isotropic. Namely, it has the same properties in all directions, for all parts of the space. We are looking for a practical computational space that has the same properties in all directions, for all parts of the space. Any two samples that are n units apart in a uniform color space will appear equally different, whether the separation is in hue, lightness, chroma, or any combination.

This paper is a revision of a paper presented at the SPIE conference on Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts IV, San Jose, CA, January 1999. The paper appears (unrefereed) in *Proc. SPIE* **3648**.

Paper 008904 received Mar. 25, 1999; revised manuscript received May 21, 1999; accepted May 29, 1999.
1017-9909/99/$10.00 © 1999 SPIE and IS&T.

1.1 *Observations Versus Equations*

There are two different approaches to uniform color spaces. The first is direct observation. Munsell, Ostwald, OSA Uniform Color Space, NCS, and ColorCurve are all examples of extensive research using human observers to select color samples that appear uniformly spaced.

The second approach is to use Colorimetric equations. The cornerstone of Colorimetry is the X,Y,Z 3D space. It is derived from color matching experiments and can predict if two samples will match based on full-spectrum measurements. Independent of X,Y,Z's ability to predict matches, it is not isotropic[1] and it cannot predict the sample's color appearance.[2] CIELAB, CIELUV, CIECAM97 are examples of equations that approximate the observer data.

The *observation* and the *equation* approaches are very different. These two approaches may give the same rendition of color appearance, or they may not. This paper will plot $L^*a^*b^*$'s rendition of observed uniform color spaces. We will concentrate much of our discussion on Munsell data, but will include other observed spaces. Ideally, the differences between the observed data and their calculated approximations are vanishingly small.

The advantage of the Munsell book (Fig. 1) is that one can see the real chips and evaluate for oneself how uniform it appears. I have recently confirmed for myself that each page appears as a single hue and that the colors are isotropically spaced. The 40 pages appear equally spaced in hue.

The advantage of CIELAB is that it has become an almost universal standard in imaging. It is built into many measuring devices for direct measurements and it uses only three equations to calculate L^*,a^*,b^* from X,Y,Z[1]

$$L^* = 116^*(Y/Y_n)^{1/3} - 16,$$

$$a^* = 500^*[(X/X_n)^{1/3} - (Y/Y_n)^{1/3}], \quad (1)$$

$$b^* = 200^*[(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}].$$

2 How Uniform is CIELAB?

Many authors use CIELAB as the space of choice for making decisions on the best color compromise, such as the mapping of extra-gamut points in printed and displayed images. They select CIELAB because hue, lightness, and chroma are approximately uniformly spaced. Before doing an extensive set of these so called gamut mapping experiments, we became curious about the actual size of

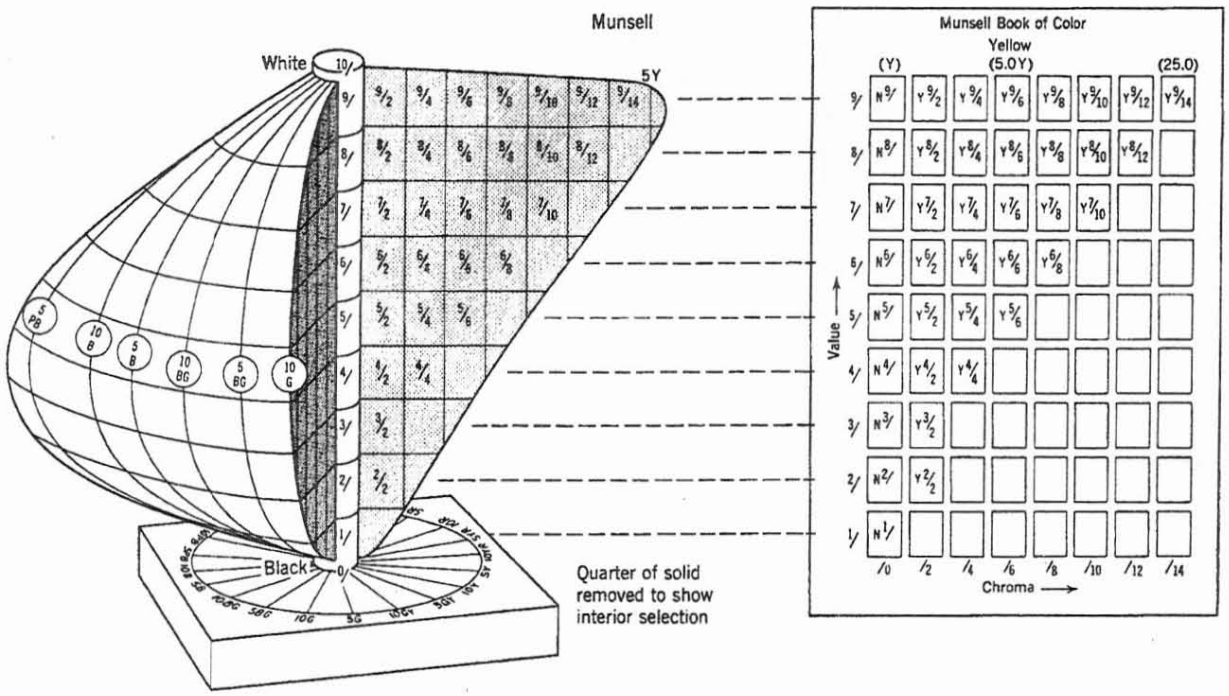

Fig. 1 Nickerson's diagram of the Munsell color space. There are 40 planes of constant hue that share the vertical neutral gray axis at the center. In each plane lightness, value, and chroma are equally spaced.

CIELAB's nonuniformities. There is abundant literature that straddles the line between it "*does a respectable job*" [3] and *it needs improvement*, usually with more clever functions.[4,5] The following comparisons of Munsell observations and CIELAB calculations are an attempt to resolve this issue.

Let us define a new space called MLab.[6] *M* is in honor of Munsell. The idea is to mimic the format and formulas of CIELAB format so we can easily compare the two. MLab just replaces the equations in $L^*a^*b^*$ with isotropic data. CIELAB lightness varies from 100 to 0. Munsell notation varies from 9/ to 1/. By multiplying Munsell value by 10, we get 90 to 10, more closely approximating L^*. In Munsell space lightness units /1 are equal to chroma units/2. By multiplying Munsell chroma by 5 we make MLab's chroma 10 equals MLab's lightness 10.

A simple test is to look at the area of biggest concern in gamut mismatch, namely saturated yellows. A yellow Munsell 5.0 *Y* 8/14 has an expected Munsell Chroma (MC) value of 70, $(14\times5=70)$. However, when we measure the chip we get $C^*=103$. $L^*a^*b^*$ has introduced an error of $\Delta E=33$. CIELAB overestimates 5.0 *Y* 8/14's chroma (MC) by $(33/70=47\%)$. This is a serious problem. In yellow, $L^*a^*b^*$ introduces an error of 47% compared with observer data.

Another example is Munsell 10 *B* 6/8. Observers place the chroma at 40 $(5\times8=40)$. CIELAB places chroma at 31.9. $L^*a^*b^*$ has introduced an error of $\Delta E=8$. CIELAB underestimates Munsell chroma by 20%. Both these discrepancies between observation and calculation are much too large. If the entire CIELAB space has similar problems, it should not be used in gamut mapping applications because the space will introduce very large errors in uniformity.

2.1 *Real Chips in the Munsell Book*

The following is a systematic study of all colors. It compares CIELAB values with the MLab values derived from Munsell notation. The data for CIELAB values come directly from Newhall, Nickerson, and Judd's 1943 paper.[7] This familiar table[8] gives *Y*, *x*, *y* for 2742 chips. This set contains both 1317 real chip data found in the Munsell book and the remainder are defined by extrapolation to the spectrum locus.[9] In this analysis we used only the real chips supplied by Munsell. We used Newhall, Nickerson, and Judd (NNJ's) illuminant *C* as the standard illuminant (see Fig. 2).

If we return for a moment to Fig. 1, we can review the design of Munsell's uniform space.

- There are eight horizontal lightness planes perpendicular to the gray axis.
- There are concentric circles of chroma, with gray at the center.
- There are 40 hues that define vertical planes that are parallel to the gray axis down the center.
- Each plane is made of colors with constant hue.
- All hue planes are uniformly spaced $(360/40=9^\circ$ apart).

Figure 3 plots the same data as Fig. 2 in the a/b plane. Here all lightnesses are compressed into a single plane. CIELAB is irregularly spaced. Some hues, such as yellow,

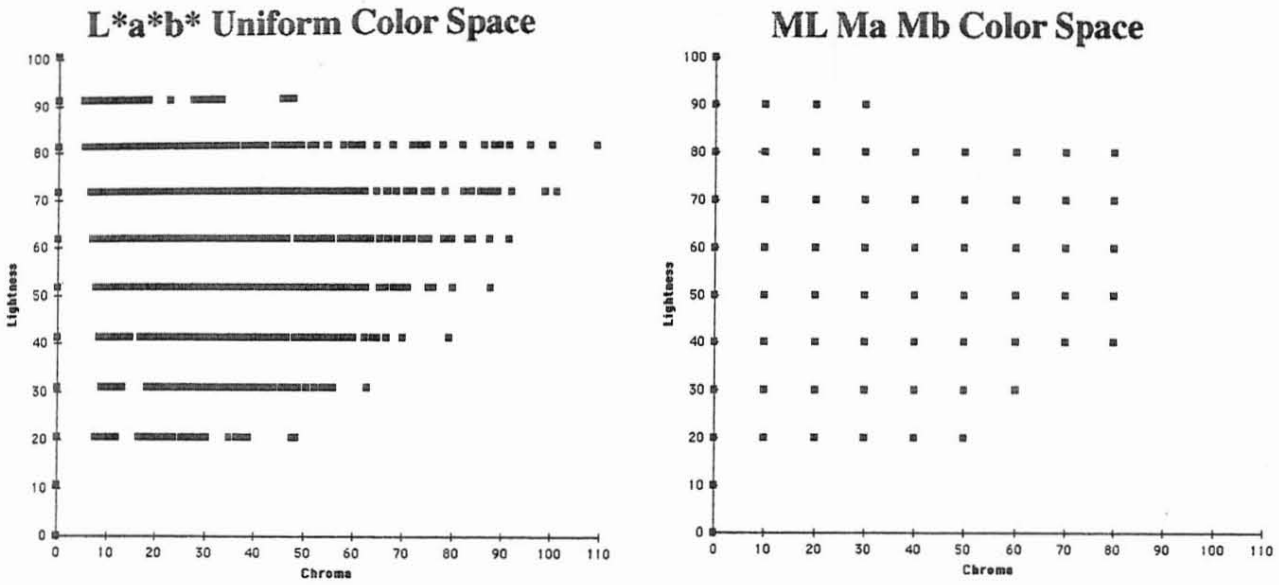

Fig. 2 Comparison of $L^*a^*b^*$ uniform color space with the proposed MLab. In both graphs chroma are plotted on the horizontal axis and lightness is plotted on the vertical axis. All hue planes are compressed into the L/C plane. $L^*a^*b^*$ is not uniformly spaced in chroma.

expand the chroma, others such as blues, compress the chroma. The distinct radial lines created by Munsell have been blurred. The MLab plot on the right shows a regularly spaced array.

2.2 *Planes of Constant Hue*

Observers picked the chips on a single page of the Munsell as a set of samples with the same hue. Each of these pages is uniformly spaced around the hue circle. The values of a^* and b^* are derived from X, Y, and Z. Nevertheless, constant hue does not correlate with constant dominant wavelength.[11] The discrepancy between hue angle, derived from a^*, b^*, and observed color can be seen clearly in Fig. 4. Here we look at all the chips found on the 2.5 R, 5 R, and 7.5 R pages. We plot the hue angle in CIELAB space on the vertical axis and give the name of the chip on the horizontal

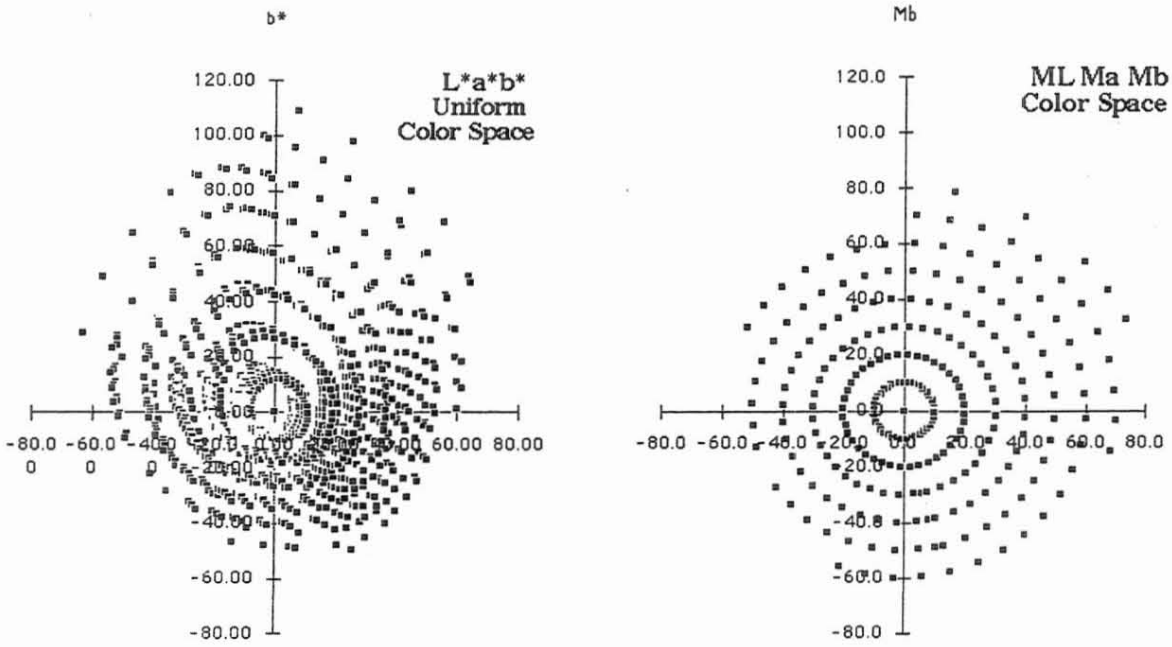

Fig. 3 Comparison of $L^*a^*b^*$ uniform color space with the proposed MLab. In both graphs a^* is plotted on the horizontal axis and b^* is plotted on the vertical axis. All lightness planes are compressed in the a^*/b^* plane. Again, $L^*a^*b^*$ is irregularly spaced.

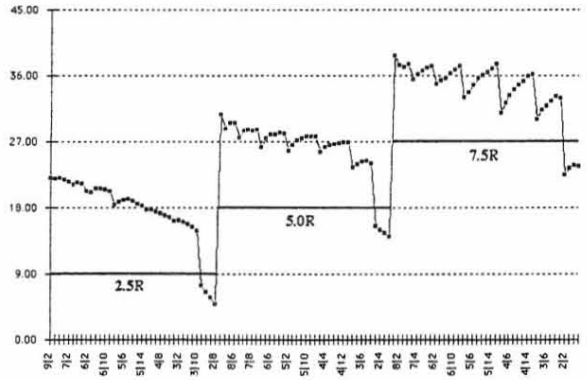

Fig. 4 The plot of hue angle on the vertical axis vs all the chips on three pages of the Munsell book. Ideally all the $L^*a^*b^*$ points should fall on the solid lines. A hue angle of 22° could be a light 2.5*R* or a dark 5.0 or 7.5 *R* chip. CIELAB does not present the same hue in a plane parallel to the gray scale. Rather constant hue is a slanted, rippled surface.

axis. Munsell notation (MLab) places all 2.5 *R* hues on the solid line at 9°, all 5 *R* chips on the 18° line, and all 7.5*R* chips on the 27° line. CIELAB portrays a constant hue angle as drifting across three pages of the Munsell book. The drift in Fig. 3 is caused by the fact that CIELAB does not position chips with constant hue in a single vertical plane. It creates a slanted plane with lightness that undulates with chroma.

We have defined the expected hue angle in MLab space by setting 10 RP to 0°. Then, all 2.5*R* chips are at 9°, all 5.0 *R* chips are at 18°,... and all 7.5 RP chips are at 351°. We can measure the hue error by subtracting MLab hue angle from CIELAB hue angle. Figure 5 shows this error plotted on the vertical axis versus all the real chips in the Munsell book. Ideally, all the errors would fall close to the 0 line. The data, however, show that there is considerable variability for each page in the book, as well as drift around the hue circle.

In order to summarize the results so far, we can review how CIELAB portrays the Munsell book.

- There are horizontal lightness planes perpendicular to the gray axis. This is in excellent agreement with Munsell.
- There are irregular shapes instead of concentric circles of increasing chroma.
- The 40 hues are not vertical planes parallel to the gray axis.
- Constant hues form corrugated, slanted planes.
- Hue planes are nonuniformly spaced.

Except for lightness, CIELAB distorts the fundamental design principles of the Munsell book. Since virtually all the errors are in the a/b planes, we can evaluate the magnitude of the errors by measuring the distance between CIELAB position and the MLab position in the a/b plane.

Figure 6 defines the $L^*a^*b^*$ error term ΔEr. The proceedings,[10] with an earlier version of this paper, includes a list of ΔE_r and $\Delta E_r/C$ for all the real chips in the Munsell book. The rows report data for a particular hue and lightness. The columns report values for different chromas. Averages are presented at the end of the table. The Munsell chip 2.5*R*9/2 has a ΔE_r of 5.0. $\Delta E_r/C$ is 50%. 2.5*R*8/2, 4, 6 have ΔE_rs of 4.4, 6.4, 8.8 and $\Delta E_r/C$s are 44%, 32%, 29%, averaging 35%. The average for all the chips on 2.5*R* page is 23%. The average of the eight *R* pages is 25%. The *Y*, *G*, *B*, *P* page averages are 36%, 19%, 14%, and 29%.

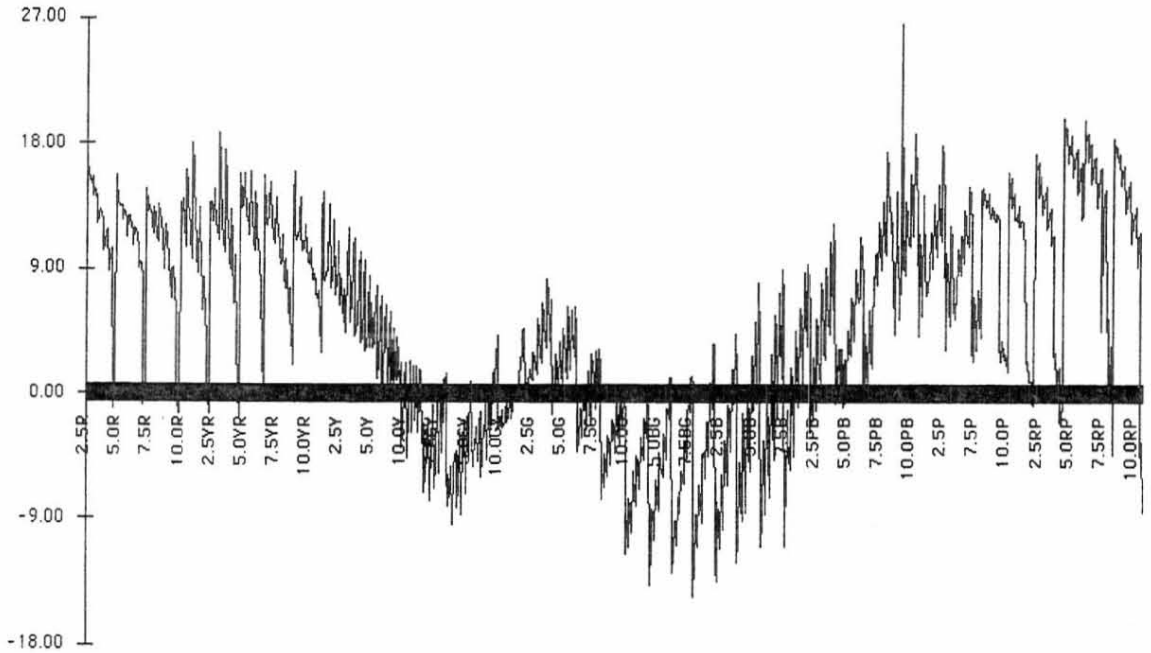

Fig. 5 The error in hue angle introduced by CIELAB. The vertical axis is the difference [CIELAB hue angle−MLab hue angle]. The horizontal axis comprises all the real chips in the Munsell book. In $L^*a^*b^*$ space Munsell red samples vary from 0° to +18° discrepancy from the ideal angle. Greens fall between +9° and −14°. Blues and purples fall between 0° and +20°.

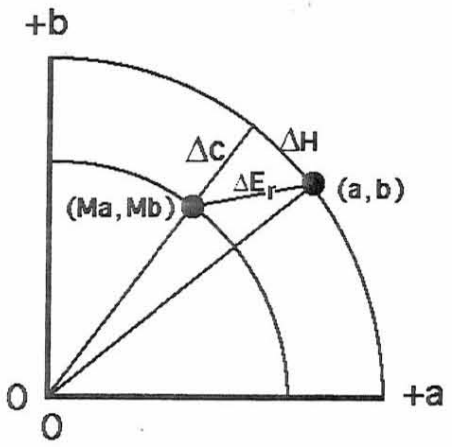

Fig. 6 The (*Ma,Mb*) coordinates show the location in MLab space. The (a^*,b^*) coordinates show the location in CIELAB space. The line *C* is the distance from the origin to (*Ma,Mb*). The line ΔE_r is the magnitude of the error introduced by CIELAB. It is the distance between (*Ma,Mb*) and (a^*,b^*). We evaluated each chip in the Munsell book as a percentage error by the ratio of ΔE_r to *C*. Further, we resolved ΔE_r into its components ΔC and ΔH to evaluate the relative sizes of chroma and hue errors.

The average of all 9/ chips is 42%, 8/ chips is 38%, 7/ chips is 32%, 6/ chips is 29%, 5/ chips is 25%, 4/ chips is 22%, 3/ chips is 21%, 2/ chips is 20%. The average of all /2 chips is 30%, /4 chips is 26%, /6 chips is 24%, /8 chips is 24%, /10 chips is 26%, /12 chips is 31%, /14 chips is 31%, and /16 chips is 33%. Studying these data suggests that the pattern of these discrepancies is quite complex.

Figure 7 shows the histogram of $\Delta E_r/C$ for all chips in the Munsell book. On average CIELAB distorts Munsell data by 27% of chroma. Figure 8 plots the direction of hue errors.

Figure 9 shows the breakdown of the Munsell book distances into component vectors, ΔH and ΔC (see Fig. 6). The ΔH graph on the left shows that hue discrepancies are significant. The middle ΔC graph shows that chroma discrepancies are somewhat larger. The resultant ΔEr error is shown in the right graph.[12] Here we see that both ΔH and ΔC are each significant contributors to ΔE_r. In the diagram on the left we see 0 ΔH at the point of normalization (90°). In the reds the errors range from -5 to $+20$. In this

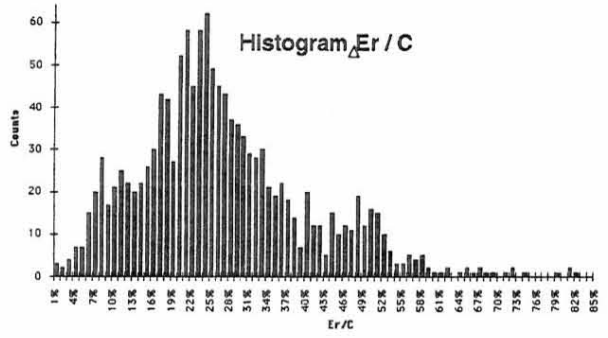

Fig. 7 The histogram of $\Delta E_r/C$ for all 1731 real chips in the Munsell book. The maximum error is 82%; the minimum is 0.01%. The mean $\Delta E_r/C$ is 27%, the standard deviation is 13%, and the median is 24%.

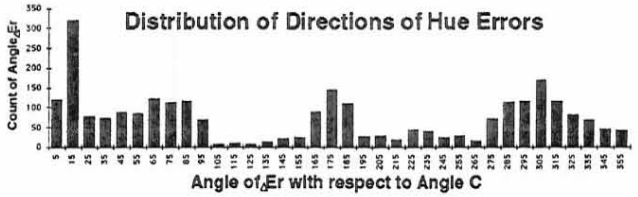

Fig. 8 Histogram of the angle of ΔE_r with respect to the angle of *C*. If the problem in $L^*a^*b^*$ could be fixed by an overall decrease in saturation, then all the errors would fall at 180°. If the problem in $L^*a^*b^*$ could be fixed by overall increase in saturation, then all the errors would fall at 0°. These data demonstrate that there is no apparent simple relationship between observed hue shift data (MLab) and calculations ($L^*a^*b^*$). The error in hue angle is distributed in virtually all directions from the point (*Ma,Mb*).

color region the ΔC increases from -5 to $+5$. Combined, the ΔE_r varies from 0 to 30 $L^*a^*b^*$ units. In the yellow region the ΔH errors range from $+20$ to -25. Here ΔC varies from $+5$ to $+30$ to $+5$. Combined the ΔE_r varies from 0 to 30 $L^*a^*b^*$ units. In the green region the ΔH errors range from 0 to -10. In this color region the ΔC is the smallest and varies from -5 to 0. Combined, the ΔE_r varies from -5 to $+5$ $L^*a^*b^*$ units. In the blue region the ΔH errors range from 0 to -10. Here the ΔC varies from 0 to -10. Combined, the ΔE_r varies from 0 to 15. In the purple region the ΔH errors range from 0 to 5. Here ΔC varies from 0 to -10. Combined, the ΔE_r varies from 0 to 15. The green to blue section of color space is best represented by $L^*a^*b^*$. The reds suffer from hue angle distortion; the yellows suffer from chroma exaggeration; and the blue–purples suffer hue angle distortion in the opposite direction.

Putting the whole story together, we see a very complex pattern of local variations between $L^*a^*b^*$ and observer appearance. Figure 7 shows a histogram mean at 27% $\Delta E_r/C$. Figure 8 shows that these errors are in all directions. Figure 9 showed that ΔH is a slightly smaller contributor than ΔC. Nevertheless ΔH is significant. On average the $\Delta E_r/C$ values are greater for higher lightnesses than for lower ones. Figure 9 shows that the smallest errors fall in the green–blue region.

3 $L^*a^*b^*$ Portrayal of Other Uniform Color Spaces

So far we have seen that Munsell book observer data and $L^*a^*b^*$ calculations do not agree. The $L^*a^*b^*$ portrayal of Munsell colors agrees in lightness, but is markedly different in both chroma and hue. If these distortions are common to all uniform color spaces, then we can conclude that the equation set fails to model color appearance. If the above observed overestimation of chroma in yellow and its underestimation in blue are found in many color spaces, then it is a $L^*a^*b^*$ problem. If, however, these properties are unique to Munsell space, then we cannot find fault with $L^*a^*b^*$.

We can compare Munsell with other color spaces such as Ostwald, NCS, and OSA Uniform Color Space and ColorCurve. First, we need to identify sources of colorimetric data for each space. OSA data are documented in Mac-

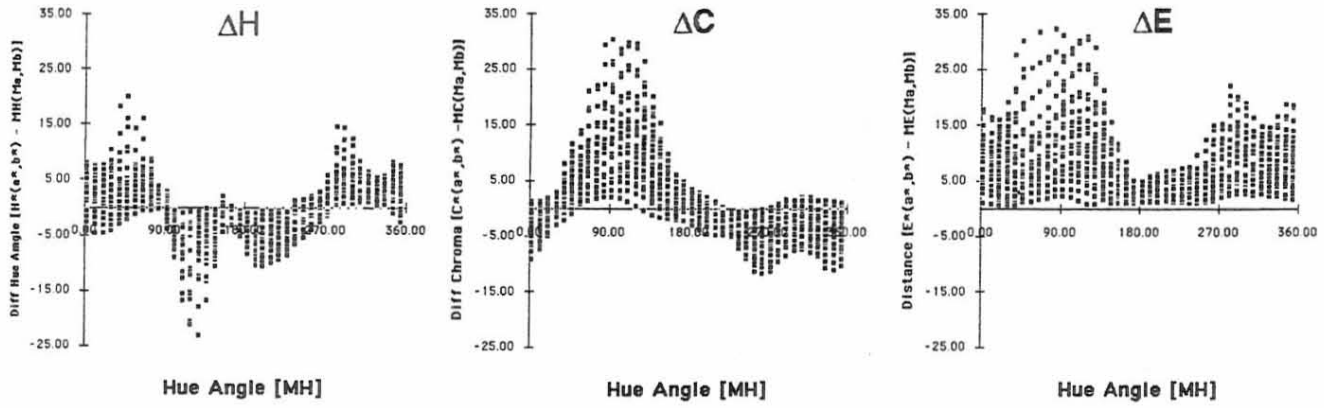

Fig. 9 Comparison of the differences between $L^*a^*b^*$ spatial rendition of Munsell and the ideal rendition MLab. The three graphs break the distances into ΔH, ΔC, and ΔE. The left graph shows the plot of difference of hue angle $[H(a^*,b^*)\text{-}MH(Ma,Mb)]$ vs ideal hue angle. The middle graph shows the plot of difference of chroma $[C(a^*,b^*)\text{-}MC(Ma,Mb)]$ vs ideal hue angle. The right graph shows the plot of distance ΔE between (a^*,b^*) and (Ma,Mb). Both ΔH and ΔC make substantial contributions to the distance between (a^*,b^*) and (Ma,Mb).

Adam *et al.*[13] and the data are reprinted in Wyszecki and Stiles.[8] Ostwald data were measured from the most colorful samples of the book.[14] NCS data came from Derefeldt and Sahlin.[15] ColorCurve provides $L^*a^*b^*$ data on its color chips in the book.[16]

Figure 10 plots the OSA Uniform Color Space in the a^*,b^* plane. On the left we again see that $L^*a^*b^*$ significantly expands the saturation of the yellows compared to MLab on the right. Similarly we find the same behavior in ColorCurve data plotted in Fig. 11.

We see now that Munsell, OSA, and ColorCurve have the same characteristics. All three spaces agree with $L^*a^*b^*$ in lightness. $L^*a^*b^*$ overestimates the chroma of yellows in all three spaces. $L^*a^*b^*$ fails to portray the observer selected papers on an equally spaced grid. The Ostwald and NCS system papers have been chosen with a different procedure. Each page of constant hue is an equilateral triangle, with neutral gray along the vertical side. The apex of the triangle is the most saturated color, regardless of lightness. In these two spaces we cannot make the same compression of the lightness axis shown in Figs. 2, 10, and 11. Lightness and chroma have different observer definitions.

Nevertheless, the Ostwald and NCS spaces are extremely valuable because they have hue plane definitions similar to Munsell. If we want to compare how the spaces handle the hue circle, we have to exclude OSA and ColorCurve spaces because each defines the color dimension by formula, rather than by observation. This is where Ostwald and NCS are very important. Here the spacing of the hue plane is observer based, and we can see if $L^*a^*b^*$ introduces the same hue plane distortions as it does for Munsell.

The next step is to convert the data to a common colorimetric space ($L^*a^*b^*$). From this we calculated hue

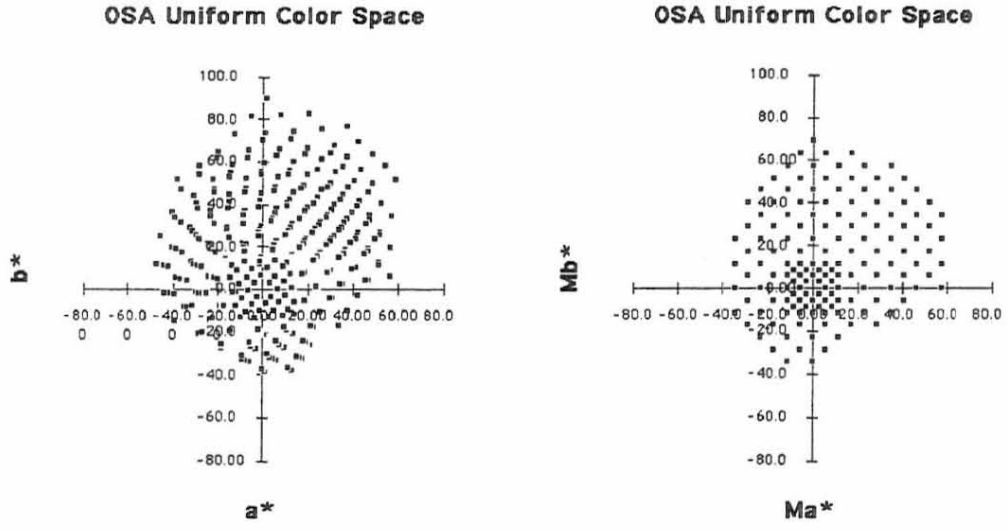

Fig. 10 Comparison of the $L^*a^*b^*$ representation of OSA Uniform Color Space with the proposed MLab representation. In both graphs a^* is plotted on the horizontal axis and b^* is plotted on the vertical axis. All lightness planes are compressed in the a^*/b^* plane. We see the same exaggeration of yellow chroma found in Fig. 2's plot of Munsell space.

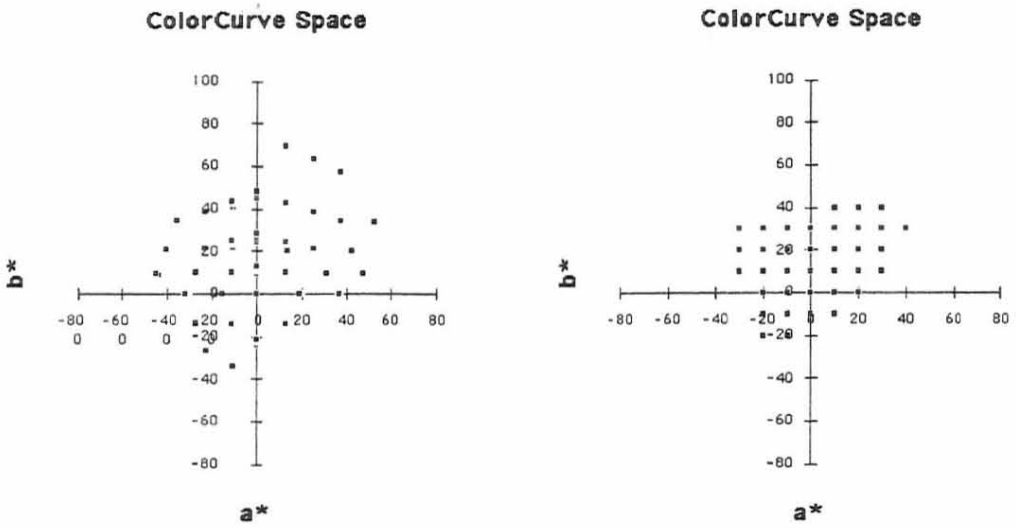

Fig. 11 Comparison of $L^*a^*b^*$ representation of ColorCurve Uniform Color Space with the proposed MLab representation. In both graphs a^* is plotted on the horizontal axis and b^* is plotted on the vertical axis. All lightness planes are compressed in the a^*/b^* plane. We see the same exaggeration of yellow chroma found in Figs. 2 and 10. In all three plots $L^*a^*b^*$ fails to portray equally spaced colors on an equally spaced grid.

angle (H) and chroma (C). Next we need to rotate the different hue circles so that they are equal at one point in the circle. We decided to assign 90° to $a^*=0$, for maximum $+b^*$ value. We took the (a^*,b^*) data for the most saturated yellow papers and calculated $H(a,b)$. We interpolated between papers to find the hue angle of the paper nearest 90°. For each color space we can now assign an ideal hue angle. For example, when hue plane $5.0Y$ is placed at 93.2°, the $a^*=0$ is at 90°. Since the Munsell space has 40 hue planes, they should be 9° apart around the circle if they are uniform. $2.5Y$ falls at 84.2° and $7.5Y$ falls at 102.2. Similarly, NCS has 40 hue planes and the Y hue page falls at 85.6° when $a^*=0°$. In this case, $Y10R$ falls at 76.6° and $G90Y$ falls at 95.6°. Ostwald has 24 planes, each 15° apart. Plane 2 falls at 90.2°, plane 1 falls at 75.2°, and plane 3 falls at 105.2°.

Now each uniform color space has a common *ideal* hue angle assigned to it. We can compare the hue angle estimated by $L^*a^*b^*$ with the ideal hue angle. We can also compare these hue plane positions between different color spaces. If all the color spaces behave identically, then Munsell hue plane positioning is the same as the others. If the different spaces all behave differently, then there are inherent errors in some, or all of the color spaces.

Figure 12 plots the difference in hue angle $[H^*(a^*b^*) - \text{ideal}\, H(Ma,Mb]$ versus ideal hue angle. The hue angle $[H]$ is calculated from a^*, b^*. The ideal hue angle $[MH]$ is calculated from Ma, Mb. These are the coordinates of the 3D LUT space. The values are calculated from the chip's Munsell notation. It represents what it should be, rather than a colorimetric calculation from the reflectance spectrum. If Munsell notation for a chip is 8/12, then lightness is 80 (8×10) and chroma is 60 (12×5).

First, we plot all the real chips in the Munsell book. Next we plot all the chips in NCS, and finally we plot the most saturated chips in the Ostwald book. The results in Fig. 12 show a general similarity between these spaces. All curves have been normalized to 0 difference at 90°. Between 90° and 270° $L^*a^*b^*$ underestimates hue angle compared to ideal angles for all three color spaces (except for three pages in Ostwald). Between 90° and 180° Munsell and NCS are in close agreement. The average errors are between 10° and 15°. Between 270° and 0° to 90° $L^*a^*b^*$ overestimates the hue angle compared to ideal angles for all three color spaces. Munsell space has the smallest discrepancies, Ostwald next, and NCS has the largest differences.

We are left with the conclusion that Munsell position of hue planes is consistent with other color systems, but not exactly the same. The spaces were defined in different illuminants and under different viewing conditions. $L^*a^*b^*$ introduces similar distortions to the placement of hue planes in Munsell, Ostwald, and NCS.

4 Solutions

Over the years since 1976 when $L^*a^*b^*$ became a CIE standard, there have been many papers discussing the problems and recommending improvements. There are far too many to review in detail. Needless to say, none of these improvements has enjoyed such common use as CIE $L^*a^*b^*$.

4.1 *Find the "Magic Bullet" Function*

The most frequently found approach has been to find mathematical functions that can fit the observer data better.[17–21] RLAB replaces the coefficients 500 and 200 with 430 and 170.[22] This makes the average error for all chips approach zero, but does little to improve the individual errors. The advantages of using equations is that they are easy to compute by hand, slide rule, or with a calculator. The disadvantage is that all of these functional approximations lack the precision necessary to render Munsell with vanishingly small errors. Munsell's entire color space is much more complex than the equations used to approximate it.

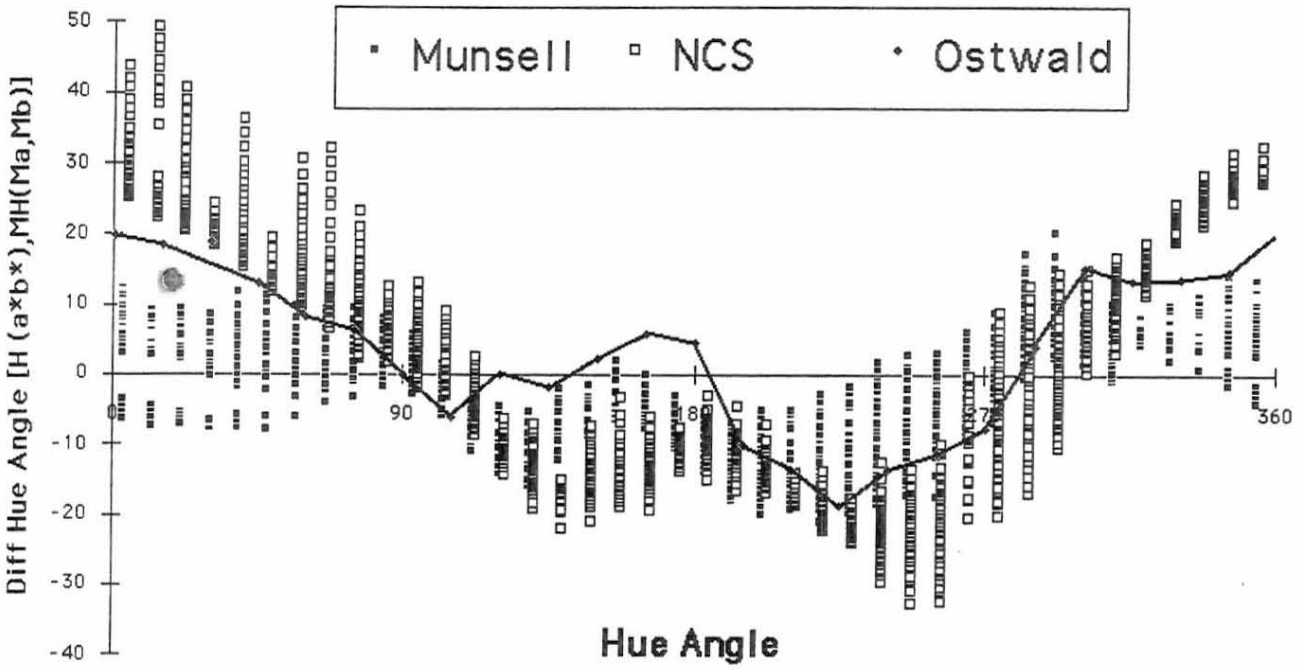

Fig. 12 The plot of difference of hue angle $[H(a^*,b^*) - MH(Ma,Mb)]$ vs ideal hue angle. All the real chips in the Munsell book are plotted as solid squares. All of the chips in the NCS are plotted as open squares. The most saturated chips in the Ostwald book are plotted as diamonds, connected by a solid line. All data rotated so that $a^* = 0$ at 90°. The graph shows that $L^*a^*b^*$ overestimates the hue angle between 90° and 0°–270° and underestimates it between 90° and 270°. The comparison of different spaces is far from perfect agreement. The differences between curves are due to the inherent difference in Munsell, NCS and Ostwald color spaces. Nevertheless, the trends are the same for these three spaces. $L^*a^*b^*$ distorts all three color spaces in the same way.

4.2 *3D LUTs*

In 1960 Rheinboldt and Menard,[23] working at the National Bureau of Standards (NBS), Washington, reported the first mechanized conversion of colorimetric data to Munsell renotations. They described a program that read Y, x, y and automatically calculated Munsell renotation. They avoided using an inverse-equation solution because it would lack experimental verification. Instead, they electronically mimiced the hand calculations using tables described in Newhall, Nickerson, and Judd.[3] They describe that in their first version of code the 4996 data points had to be stored on magnetic tape, because core memory was limited. In the second version, using the NBS IBM704 computer, with 8000 words of memory, it was possible to store the data in core.

In the 1980s the Polaroid Vision Research Laboratory began a series of experiments attempting to characterize the behavior of Polacolor instant film in devices that exposed the film using digital control of three narrow-band illuminants. Instant film has highly nonlinear properties. The three sets of dyes that form the final image are in a nine layer structure that also contains the three light-sensitive emulsions. First, the cyan dye has to migrate past the red-sensitive emulsion that controls its concentration in the final image. Then, it must pass through the magenta dye layer, the green sensitive emulsion, the yellow dye layer, the blue-sensitive emulsion, and the viscous developer to reach the mordant layer. Obviously, the amount of cyan dye in the final image depended primarily on the red exposure, but it was significantly influenced by the green exposure and the blue exposure.

We studied the problem from two perspectives: mathematical models and 3D LUTs. Although we were able to model the system with high order polynomials we found that process too slow and painful to apply to practical working environments. By using the inherent power of scanners and computers, we found it faster, simpler, and easier to measure the response of a film to all possible combinations of the three exposures, than to accurately model its complete behavior. With a known digital image test target we exposed all combinations of eight levels of R, G, and B. Scanners read that image and programs selected the pixels associated with each test patch. This made sure that any error in alignment of the print was corrected.

The film's response in all parts of the color space was used to create any desired color. This system was used in a color transform board designed by the Vision Research Lab that resided in the MacDonald Detweiller FIRE 300 and 1000 film recorders. This device was an early digital prepress proof system. Starting with digits used to make the press plate, along with press calibrations, we calculated the desired color for each pixel. Using the calibrated response of the film we can calculate the three exposures to write, pixel by pixel, a highly accurate prepress proof. All image calculations were done by the 3D hardware lookup table. The calibration of the press and of the film were combined so that there was only one operation performed in the hardware color transformation board. This board resides in the

printer. Digits used to make printing plates were sent to the film recorder. The color transform board stored map coefficients derived from the press and film calibrations. The input digits were sent to the lookup table. All combinations of the most significant five digits were prestored in the board. Using prestored information from all nearest neighbors we interpolated the best output in the least significant bits. The hardware board calculated the LUT output value much faster than the film recorder could read image data from the magnetic disk. A more complete description can be found in a U.S. Patent entitled *Method and Apparatus for Transforming Color Image Data on the Basis of an Isotropic and Uniform Colorimetric Space*.[24]

Kotera and his colleagues[25–29] have used 3D LUTs in a wide variety of applications, including high speed video applications. The 3D LUTs techniques are used to make accurate color reproductions of paintings.[30] Here the LUTs are used to remove the characteristic sensitivity functions found in films.[31] These reproductions have a slope 1.0 relationship with the painting throughout the entire color space. Many commercial products use 3D LUTs as the most efficient way to calculate accurate information in real-time applications. For more information on 3D LUTs, see Ref. 32.

If we think of the uniform color space problem as a candidate for 3D LUTs we have:
Advantages

- All 1317 Munsell chips have zero error (they are table lookups).
- All intermediate values are computed from all their nearest neighbors.

Requirements

- Requires a computer program and LUT.
- Size of program=80 KBytes.
- Size of LUT=806 KBytes.
- Run time=1 s for all 1317 chips in the Munsell book.

Disadvantages

- Difficult to do with a slide rule.
- There are no disadvantages for anyone using a computer.

The lookup table function has been possible for many years, first by using the Newhall, Nickerson, and Judd, table as originally designed. The real-time conversion from Munsell notation to $L^*a^*b^*$ and the inverse interpolation can be done with a Munsell conversion program sold by Munsell of Greytag Macbeth. It can be found on the web at http://munsell.com/Download.htm

Munsell notation is far from ideal in format. $L^*a^*b^*$'s format was chosen so that the one can calculate distances and angles. The ideal case is to use Munsell information to calculate a MLab space and use it as we would CIE $L^*a^*b^*$. To be effective we need 3D LUTs to convert $L^*a^*b^*$ to MLab and MLab to $L^*a^*b^*$.

5 LabtoMLab And MLabtoLab Working 3D LUT

The best source of data for a 3D LUT is the data from Newhall, Nickerson, and Judd (NNJ).[7] Here they list Munsell notation colors that are uniformly spaced in appearance. The foundation is a very large study of real papers selected by extensive observer experiments. They interpolate colors out to the spectrum locus. There is another advantage to the Munsell data. In situations in which the range of colors exceeds the set of reflectance papers, we can use the rest of the NNJ table to reach the total range of all possible colors. This simply means that the 3D LUT has 2742 input chips, instead of 1317.

They provide an atlas of all possible colors that are uniformly spaced. First we need to make a 3D table of equivalents. We used the NNJ Y, x, y data to calculate $L^*a^*b^*$ in illuminant C for each chip. The MLab designation was calculated from Munsell notation as follows:

- Hue angle for 10 RP was set to 0°.
- A table of the 40 hue pages was spaced 9°. $2.5R=9°$, $5.0R=18°$, ...$7.5RP=351°$.
- H=hue angle from the above table.
- ML=10* value notation.
- MC=5* chroma notation.
- Ma=MC* cos H.
- Mb=MC* sin H.

With Gary Dispoto and Michael McGuire of HP Labs we created a pair of 3D LUT LabtoMLab and MLabtoLab and two programs to use them. The input and output channels of the LUTs are scaled to 8 bits. The tables are $65\times65\times65\times3$ planes, 824 Kbytes which means the six most significant bits are looked up and the two least significant bits are interpolated. The programs do the transform in either direction, depending on the table used. One program makes single calculations, with input and output in normal units. The other program is for transforming TIFF images and uses the standard representation of $L^*a^*b^*$. The programs are console applications for Win NT or WIN95/98.

All programs ran on an HP Kayak PC with a 300 MHz processor. The LabtoMLab table is 806 KBytes. The program accepts either a triplet of $L^*a^*b^*$ values, or an array of them, and calculates the MLab values for all inputs. The program's size is under 80 KBytes and it takes less than 1 second to convert a 512×512 array of $L^*a^*b^*$ to MLab. The second set of LUTs makes the inverse estimation. It reads in MLab data and puts out $L^*a^*b^*$. The MLabtoLab calculation has roughly the same sizes and speeds. This means that anyone concerned with working in a truly uniform color space can do so by taking advantage of 3D LUT technology. It takes only a second to get there.

Gabriel Marcu presented a recent paper using the same approach.[33] He used a 3D LUT to convert $L^*a^*b^*$ to MLab. He optimized portions of an out-of-gamut image in MLab space. He got improved pictures by interpolating in MLab space.

We are very fortunate these days to have such a wide variety of reliable colorimetric equipment. Many of these devices provide direct $L^*a^*b^*$ measurements of high ac-

curacy and reliability. The problem is not the devices, but the $L^*a^*b^*$ space, as seen above. In the future we can speculate that these devices will include MLab LUTs. Until then, pairs of efficient 3D LUT computer programs can easily transform our data into and out of the *uniform* Munsell color space.

6 When Do We Care?

Any time that we need to compare two colors and we want to express that information as a distance we imply an isotropic color space. If the colors vary only in lightness, then L^* is perfectly appropriate. If the colors vary in hue angle and chroma, then $L^*a^*b^*$ will introduce on average a 27% error at each end of the color difference. A numerical color space must conform to color appearance experiments in color appearance.

7 Discussion

$L^*a^*b^*$ calculations designed to create a uniform color space do not correspond to observer data found in Munsell, Ostwald, NCS, OSA, and ColorCurve spaces. The extent and direction of the discrepancies between observer and equation are complex and vary considerably in local regions of the space. Although these observer based spaces are all different because of their design, they all show similar discrepancies with $L^*a^*b^*$'s portrayal of uniformity. The convenience of using $L^*a^*b^*$ made us all use it beyond its experimental definition. No one I know has ever said that $L^*a^*b^*$ is proven to be sufficiently accurate for the delicate rearrangement of complex scenes required in color-gamut mapping. Nevertheless, it is almost universally used in the evaluation of images and in instrumentation. Whenever we report a separation between two colors as a distance (ΔE), we imply an isotropic space. Whenever we use $L^*a^*b^*$, each point has an average isotropic discrepancy of 27%.

All the problems of complexity disappear with the use of 3D LUTs. The user of the LabtoMLab tables does not notice the increased computational load between the LUT and a formula. They both report their answers in an instant. However, the LUT takes into account all the local variations found in the observer data. The LUT has zero error for all chips in the Munsell book.

We reviewed the selection of Munsell space as the data for a 3D LUT. The Munsell space shares the same goals as Ostwald, NCS, OSA, and ColorCurve spaces. They all attempt to provide a set of colors that are uniformly spaced in hue, chroma, and lightness. The experiments that defined these spaces were different and therefore the data are somewhat different. Nevertheless, Munsell, NCS, and Ostwald show similar properties when compared in $L^*a^*b^*$ space. In OSA space, j and g have a specific relationship defined by the equations[34] and are described by CIE 1964 Tristimulus values. ColorCurve used a^* and b^*. These hue plane placements are different from those of Munsell, NSC, and Ostwald.[13]

The decomposition of the distance between (a^*,b^*) and (Ma,Mb) into ΔH and ΔC shows that hue distortions introduced by $L^*a^*b^*$ are substantial, but somewhat smaller than those introduced by chroma. Lightness discrepancies are very small.

Munsell space remains the preferred space for a uniform color 3D LUT. It is unique in that it provides data out to the spectrum locus. It is the compilation of 3 million observations that is most highly relevant to the problem. There could possibly be some errors in the hue plane placement, but it is not obvious how to improve the data collecting process.

8 Conclusions

This paper describes the substantial discrepancies between observed uniform color spaces and calculated $L^*a^*b^*$. The discrepancies are complex and apparently unique in each subsection of color space.

We recommend the use of 3D LabtoMLab LUT to convert colorimetric measurement (X,Y,Z or $L^*a^*b^*$) to isotropic MLab. Distances in color space can be accurately and conveniently calculated in MLab. When required inverse MLabtoLab LUTs can accurately return values to CIE notation.

Calculations that involve finding minimal error in hue, chroma, and lightness, such as color-gamut mapping, will benefit from analysis in an isotropic space with zero errors.

Acknowledgments

The author wishes to thank Gary Dispoto and Mike McGuire for their work on the very fast 3D LUT running in Windows. He also wants to thank Gunilla Derefeldt for providing colorimetric data on NCS and Peter Engledrum and Max Satzman for leads to NBS work in the 1960s, and to John Meyer, Gabriel Marcu, and Mary McCann for their thoughtful discussions and comments.

References

1. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., pp. 306–330, 506–513, Wiley, New York (1982).
2. E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.* **61**, 1 (1971).
3. M. D. Fairchild, *Color Appearance Models*, p. 221, Addison Wesley, Reading, MA (1998).
4. M. D. Fairchild, "Refinement of the RLAB color space" *Color Res. Appl.* **21**(5), 338–346 (1996).
5. F. Ebner and M. D. Fairchild, "Gamut mapping: Evaluation of chroma clipping techniques for three destination gamuts," *Proc. 6th IS&T/SID Color Imaging Conf.*, p. 57, Scottsdale, Arizona (1998).
6. J. J. McCann and M. Stokes, "Color spaces and image quality," *Proc. IS&T PICS*, pp. 140–144 (1998).
7. S. M. Newhall, D. Nickerson, and D. B. Judd, "Final report of the O.S.A. subcommittee on spacing of the Munsell colors," *J. Opt. Soc. Am.* **33**, 385 (1943).
8. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., pp. 840–852, Wiley, New York (1982).
9. D. L. MacAdam, "Maximum visual efficiency of colored materials," *J. Opt. Soc. Am.* **25**, 361 (1935).
10. J. J. McCann, "Color gamut measurements and mapping: The role of color spaces," in *Color Imaging: Device independent Color, Color Hardcopy, and Graphic Arts IV*, G. Beretta and R. Eschbach, Eds., *Proc. SPIE* **3648**, 68–82 (1999).
11. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., pp. 420–424, Wiley, New York (1982).
12. J. J. McCann, "Uniform color spaces: 3D LUTS vs algorithms," *Proc. IS&T PICS* (in press).
13. D. L. MacAdam, "Uniform color scales," *J. Opt. Soc. Am.* **64**, 1691–1702 (1974).
14. *Color Harmony Manual*, 4th ed., Container Corporation of America, Chicago (1958).
15. G. Derefeldt and C. Sahlin, "Colour order systems for computer graphics I. Transformation of NCS data into CIELAB colour space," Report C53013-H2, *FOA Rapport* (1984).
16. ColorCurve, Minneapolis (1988).

17. P. Hung and R. S. Berns, "Determination of constant hue loci for a CRT gamut and their predictions using color appearance space," *Color Res. Appl.* **20**(5), 285–295 (1995).
18. CIE, Technical Report CIE No. 116 (1995).
19. P. Glatz and E. Luebbe, "Colour gamut transformation using visually assessed colour difference formula," *Proc AIC*, Vol. 1, p. 586, Kyoto (1997).
20. F. Ebner and M. D. Fairchild, "Finding constant hue surfaces in color space," *Proc. SPIE* **3300**, 107–117 (1998).
21. Braun and Fairchild, "Color gamut mapping in a hue-linearized CIELAB color space," in *Proc. 6th IS&T/SID Color Imaging Conf.*, p. 163, Scottsdale, Arizona (1998).
22. M. D. Fairchild, "Refinement of the RLAB color space," *Color Res. Appl.* **21**(5), 338–346 (1996).
23. W. C. Rheinboldt and J. P. Menard, "Mechanized conversion of coloimetric data to Munsell renotaions," *J. Opt. Soc. Am.* **50**, 802–807 (1960).
24. M. Abdulwahab, J. L. Burkhardt, and J. J. McCann, "Method and apparatus for transforming color image data on the basis of an isotropic anrm colorimetric space," U.S. Patent, Application No. 4,839,721 (filed Jun. 13, 1989).
25. K. Kanamori, H. Kawakami, and H. Kotera, "A novel color transformation algorithm and its applications," *Proc. SPIE* **1244**, 272–281 (1990).
26. K. Kanamori and H. Kotera, "Color correction technique for hard copies by 4-neighbors interpolation method," *J. Imaging Sci. Technol.* **36**(1), 73–80 (1992).
27. K. Kanamori, H. Kotera, O. Yamada, H. Motomura, R. Iikawa, and T. Fumoto: "Fast color processor with programmable interpolation by small memory (PRISM)," *J. Electron. Imaging* **2**(3), 213–224 (1993).
28. H. Kotera, K. Kanamori, T. Fumoto, O. Yamada, H. Motomura, and M. Inoue: "A single chip color processor for device independent color reproduction," in *Proc. First CIC*, pp. 133–137 (1993).
29. T. Fumoto, K. Kanamori, O. Yamada, H. Motomura, and H. Kotera: "SLANT/PRISM Convertible structured color processor MN5515," *Proc. 3rd CIC*, pp. 101–105 (1995).
30. J. J. McCann, "High resolution color photographic reproductions," *Proc. SPIE* **3025**, 7 (1997).
31. J. J. McCann, "Color imaging systems and color theory: Past, present and future," *Proc. SPIE* **3299**, 38–46 (1998).
32. H. Kang, "3D lookup table with interpolation," in *Color Technology of Electronic Imaging Devices*, Vol. 28, SPIE Press, Bellingham, WA (1997).
33. G. Marcu, "Gamut mapping in munsell constant hue sections," in *Proc. 6th IS&T/SID Color Imaging Conf.*, p. 159, Scottsdale, Arizona (1998).
34. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., pp. 506–513, Wiley, New York (1982).

John J. McCann has studied color and calculated sensations since college. As a freshman at Harvard he started to work part-time for Edwin Land and Nigel Daw at Polaroid. He did an undergraduate thesis under John Dowling on human adaptation mechanisms. He is co-author of Retinex theory and has written more than 50 papers on color in complex images, color from rods and long-wave cone interactions, spatial-frequency response of the eye, visibility of gradients, and lightness. He has written a dozen patents on imaging. At Polaroid he managed the Vision Research Laboratory, Large Format Photography, and the Polaroid Replica Collection, a digital reproduction service for fine art, until he retired in 1996. He now consults and continues his research on color, lightness, and calculated color sensations. John was elected fellow of IS&T in 1984 and is past president of IS&T.