A Spatial Color-Gamut Calculation to Optimize Color Appearance

John J. McCann
McCann Imaging
mccanns@tiac.net

Abstract

Colorimetry is limited to image data from a single pixel. Measures of errors between an "original" and a "reproduction" are usually described as the distance between the two values of a pixel in a colorimetric three-dimensional space.

Human color constancy uses spatial comparisons between different parts of the image. The relationships among neighboring pixels are far more important than the absolute differences between the colorimetric values of an original and its gamut-limited reproduction.

If all the pixels in an image have a reproduction error in the same direction (red, green, blue, lightness, hue, chroma), then our color constancy mechanism helps to make large colorimetric errors appear small. However, if all the errors are randomly distributed, then small colorimetric errors appear large.

This paper will describe experiments using constant colorimetric and appearance errors to produce variable quality of reproduction. Further, it describes a technique of calculating the best appearance image using spatial comparisons. This calculation will be applied to a color-gamut problem. This approach minimizes the spatial errors introduced by limited color gamut and employs human color constancy mechanisms, so as to reduce the color appearance differences caused by limited color gamut.

1. Introduction

The Retinex model for estimating apparent lightness was proposed by Land and McCann in 1967.[1] It was applied to color constancy by McCann, McKee and Taylor in 1976.[2] In the early 1980's Frankle and McCann[3] extended the ration-product-reset-average operation to highly efficient multiresolution image processing. Recently, Retinex has been extended for use in calculating the closest color appearance in situations in which the reproduction is made with media having a smaller color gamut than the original[4].

This paper describes the experiments that lead up to the new gamut Retinex calculation and discusses the results of a sample calculation. The central theme here is that the underlying mechanisms that control color constancy can be used to advantage to make images in a small color gamut resemble images in a large color gamut. These color constancy mechanisms are spatial comparisons between pixels. They show almost no dependence on the L, M, S triplet of radiance at a pixel. It suggests that color gamut transformations calculated one pixel at a time produce poor color reproductions, while those done using spatial comparisons can give better results.

Gamut mapping transformations are an essential part of trying to make prints look like monitors despite large differences the gamuts of the display and the print media. Morovic and Luo[5] have reviewed the extensive literature and described three generations of Gamut Mapping Algorithms (GMA). Some of these GMAs are very complex. CARISMA applies different methods of empirical correction for each hue depending on the shapes of the two gamuts. Jan Morovic chairs the CIE Technical Committee 8–03[6] on Guidelines for the Evaluation of Gamut Mapping Algorithms working to establish comparison techniques by unbiased observers. It is beyond the scope of this paper to provide blind comparisons of relative performance of spatial vs. the many different pixel-based procedures. Further experimentation is required to establish the relative merits of the techniques in complete image transformation system (spatial GMA). This paper is limited to describing the ideas supporting spatial mechanisms for gamut mapping.

2. Color Gamut Calculations using a Two-Area Mondrian

We can begin this discussion of Color Gamut calculation with an analysis of calculation strategies. Figure 1 illustrates different choices of strategies used in calculating the best color compromise for a limited gamut reproduction.[7] For simplicity we will study only the combined green record seen by the observer. In other words, we will study the color image as a color densitometer would isolating only the middle-wave, or green light from the combined, final color image. (The same arguments apply to both the long- and short-wave records). In principle the green record is controlled by only the magenta dye. In fact, the green record as seen by observers is controlled by the optical density of the magenta dye and the unwanted optical densities contributed by the yellow and cyan dyes. Monitors, using additive color, have larger gamuts near white because there are no unwanted absorptions. The color gamut of subtractive dye sets is controlled by the amount of unwanted absorptions. Dye sets with larger magenta gamuts have less absorption in the green by the yellow and cyan dyes. In this scheme, green radiance = 100 is the absence of any dye (reflecting all the green light), while green radiance = 0 is maximum dye (reflecting a minimum of green light). The first panel shows the original with radiance A=89 from a light gray patch (left area) and radiance B=95 from a bright blue patch (right area).

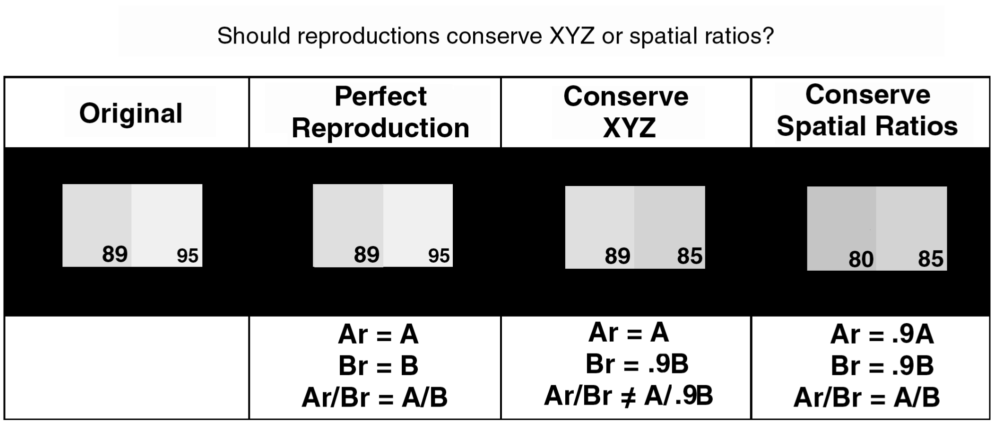


Figure 1 illustrates different approaches to making reproductions. All four panels represent the middle-wave, or green, light coming to the eye from complete color images. The first panel illustrates the original image with two areas with radiances A and B. The second panel illustrates a reproduction with the same color gamut as the original. Here the two reproduction radiances Ar and Br equal those (A and B) of the Original. Both conservation of X, Y, Z and conservation of spatial ratios are successful. The third panel illustrates the approach of Colorimetry. Here each area is treated separately. The reproduction of the left, gray area equals the original (Ar =A). However, the gamut limit found in this example restricts Br to 90% of B. The chroma of the blue area is limited by the unwanted absorptions of the cyan and magenta dyes. By selecting the best fit for the right area and the left area independently we alter the ratio Ar/Br compared to A/B. Now the reproduction reports that the left area is lighter that the right. The fourth panel on the right illustrates the conservation of spatial ratios. Here the limit of 90% on area B is also applied to the area A. The effect of keeping ratios constant, is that we have made both the areas darker. Despite this change in radiance, the relationship of areas A and B are the same. The thesis of this paper is that conserving ratios makes a better reproduction than conserving X, Y, Z.

The second panel shows the case in which the color gamut of the reproduction is as large as the original. In this fairly rare case, the reproduction radiances Ar = A and Br = B. That also means that Ar/Br =A/B.

The third panel illustrates usual color-gamut transformations. Here the dye sets are such that reproduction radiance Ar =89= A is in-gamut. However, reproduction radiance Br=85 is the best possible in-gamut value compared to the original radiance, B=95. This commonly occurs if A is close to white and B is a saturated color with unwanted absorptions from the other dyes In this example we substitute the closest value, namely 85. This approach conserves the colorimetric values X, Y, Z for area A. It leaves all in-gamut pixels unchanged, thus minimizing the cumulative color distance between original and reproduction for all areas. This choice has a highly adverse effect on the ratios. The original ratio is A/B=89/95= 0.94. The reproduction ratio has the value Ar/Br=89/85=1.06. The reproduction now reports that A is lighter than B, while the original reported that B is lighter than A. Such reproductions with distorted edge ratios make poor reproductions.

The fourth panel illustrates the principle of conserving spatial ratios. Here we have the same limitations on gamut. However, the limited gamut of B = 85 causes a shift of A to 80, so as to conserve spatial ratios. Here Ar/Br = A/B is the controlling principle. This mechanism requires an adjustment for area Ar from 89 to 80 even though the 89 is in gamut. This strategy increases the cumulative colorimetric error distance for all areas between the original and the reproduction, yet preserves the information that B>A.

Although Figure 1 is a good illustration of the strategies, it does not convince the reader that one approach is better than the other. For that we need to study complex images.

3. Search for a Color Metric

There is a dual history of color; one based on art and psychology of images and the other based on the physics of a pixel[8] Leonardo, Goethe, von Guerica, Rumford, Helmholtz, Chevreul and the Gestalt school all pointed out the roll of the image. In the second half of the 20th century, Campbell and Robson paper[9] on the visibility of sine-wave gratings lead to a the entire field of spatial vision. In a review paper Braddick, Campbell and Atkinson[10] describe the wide range of experiments using gratings as a visual stimuli. These experiments demonstrate the dependence of visual response on spatial properties of the stimulus. The paper begins with a description of receptive fields and channels, continues with periodic stimuli, and human responses to spatial frequency, orientation and temporal stimuli. This great wealth of data led to numerous spatial models.[11,12]

The physics of a pixel story of color was established by Newton, Young and Maxwell. They described the eye's response at a point of light on the retina. In the 20th century international standards were adopted for Maxwell's color matching functions.[8,13] Today, as a practical matter, color differences are almost always described by ΔE in L*a*b* color space. This space, defined by the 1976 CIE report[14], is calculated using the Tristimulus values X, Y, Z, defined in the CIE 1931 report[15]. Further, a complex image is often evaluated by averaging the individual ΔEs to calculate a Color Metric for the color difference between two images.

Colorimetry provides easy-to-use equations for wavelength matches, derived from the properties of the retinal light receptors. However, these equations do not predict appearance representing the sensation image after spatial interactions in the human visual system[16]. Physical models of the image at the retina do not predict appearance in the brain's visual cortex. The question we are asking is whether the commonly used average ΔE metric is appropriate for predicting the quality of reproductions or for evaluating the quality of gamut mapping. In other words, can color differences between copy and reproduction be used to quantify the quality of the reproduction?

Although researchers are usually concerned with matches across media, such as display to print, these experiments were restricted to single media to eliminate prop-

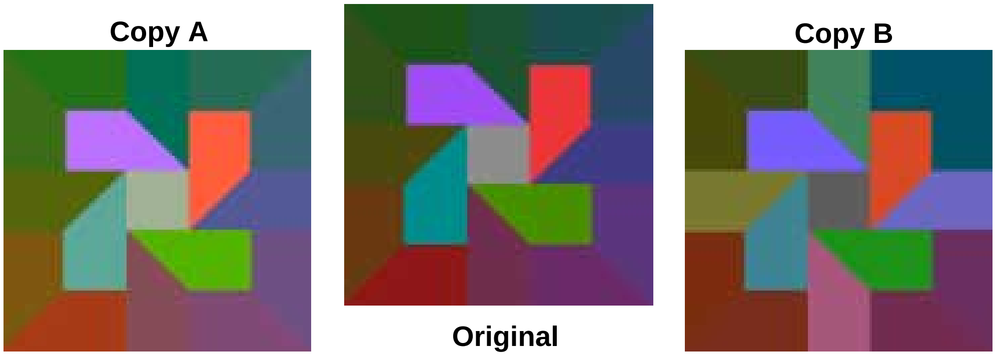


Figure 2. Copy A was made so that each area in Copy A was lighter ($\Delta L=+10$) and less red ($\Delta a=-10$) and yellower ($\Delta b=+10$) than the Original. The first important observation is that Copy A is a fairly good reproduction, considering that it has a $\Delta E=17$ for each area. Copy B was made so that the $\Delta E=17$ changes were in many different directions. In this case the ΔE s were chosen to change the appearance of the display. The outer corner patches moved closer in color to each other. The other, mid-side patches moved closer in color to the inner areas. The net effect is that Copy B does not look like the Original. It looks like a different display. Nevertheless, if judged by the ΔE Color Metric, Copy B is exactly as good a reproduction as Copy A.

Area	Original			Copy A				Copy B			
	L*	a*	b*	L*	a*	b*	DE*	L*	a*	b*	DE*
1	54.4	0.6	-8.1	65.0	-7.3	-0.8	15.0	38.5	-1.1	-6.5	16.1
2	43.8	57.2	26.5	56.9	44.5	36.0	20.5	48.6	43.2	44.3	23.2
3	43.0	-31.5	48.4	50.6	-35.0	58.8	13.3	37.4	-49.1	33.1	24.0
4	35.4	-36.4	-23.7	54.1	-22.8	-14.2	25.0	43.9	-20.8	-22.4	17.8
5	42.5	41.6	-39.8	55.2	33.2	-34.6	16.1	44.0	23.2	-43.4	18.8
6	23.3	-25.0	1.5	29.3	-36.7	-8.7	16.6	42.6	-28.4	-0.9	19.8
7	23.9	-19.8	-12.9	33.0	-28.4	-8.4	13.4	24.8	-21.7	-22.0	9.3
8	24.4	-5.6	-24.6	35.2	-14.1	-19.4	14.6	24.8	-22.0	-21.6	16.7
9	24.4	13.6	-35.1	37.0	5.8	-29.9	15.7	45.2	12.3	-35.1	20.8
10	27.0	25.1	-29.3	37.8	16.4	-22.1	15.7	27.8	27.6	-16.7	12.9
11	26.1	32.5	-20.9	37.8	24.4	-16.1	15.1	27.2	29.8	-9.8	11.5
12	26.4	29.3	-9.1	37.9	21.4	-1.4	15.9	47.0	28.0	-8.9	20.6
13	25.0	41.3	25.9	36.7	36.4	39.5	18.6	27.6	28.3	25.2	13.3
14	27.0	16.9	26.8	38.1	9.8	41.3	19.6	28.0	29.2	30.9	1.3
15	26.8	-4.4	25.7	36.6	-12.2	39.7	18.8	46.8	-5.2	30.1	20.5
16	24.6	-16.9	19.4	33.4	-25.1	30.4	16.4	28.2	-6.6	29.2	14.7
17	23.3	-26.4	17.6	31.5	-35.5	28.0	16.0	27.2	-14.8	24.4	14.0
						Average	16.8			Average	16.9

Table 1 lists the measured L*, a*, b* values for the Original, CopyA and Copy B. Beside the L*a*b* data is the ΔE* value measuring the color difference between the Copy and the Original. The average distance is 16.8 for A and 16.9 for B. The printer measurements have more variability than those from the display, but on average the color differences are close to the desired value of 17.

erties of materials and calibration variables. These experiments on identical media were aimed at observer preference of Color Metric without confounding media issues. Initially, all experiments were performed on computer displays. Later, these experiments were repeated using print materials with the same results.

Select the Better Reproduction

The following series of demonstrations creates triplets of targets: Original, Copy A and Copy B (Figure 2)[16]. The colors in the Copy A were selected to be significantly different from those in the Original. Each area in Copy A is 10 units lighter, 10 units less red and 10 units yellower in Lab space. The combined distance is $\Delta E=17$. For each area the color difference was $\Delta E=17$ between the Original and Copy A. Copy B was made with each area $\Delta E=17$ compared to the Original, but it was designed have the color shifts go in many different directions. Copy B significantly changed the local relationships, while Copy A preserved them.

As described above the original demonstration was created on a computer CRT. Following that the image was printed on a Fugix printer. Table 1 lists the measured L*a*b* values of the printed Original, Copy A and copy B.

If average ΔE could be used as Color Metric for quality of reproduction, half the observers would select Copy A as the best reproduction and half the observers would pick Copy B. If spatial parameters, namely the relationship between different

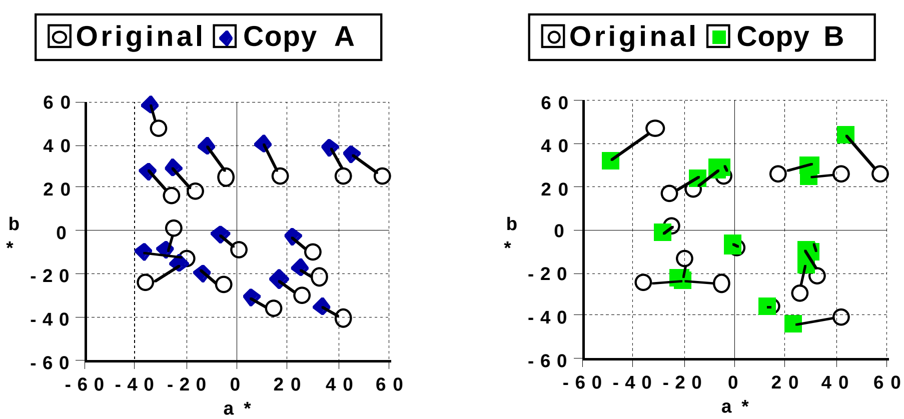


Figure 3 (Left) plots the measured print a* vs. b* value from the Original [open circles] and Copy A [diamonds]. The solid lines illustrate the shift in a*/b* plane caused by Copy A. With two exceptions (introduced by printer) all the vectors go in the same general direction and have similar lengths. By comparison (Right) plots the measured print a* vs. b* for Original [open circles] and Copy B [squares]. Here the solid lines illustrate the shift in a*/b* plane caused by Copy B. All the vectors go in the different directions and with variable lengths.

areas within the test target is important, then observers will select Copy A, preserving the spatial relationships. Observers were asked to pick the preferred reproduction.

The experiment consisted of asking people to choice the better reproduction. In a poster session at Fourth IS&T/SID Color Imaging Conference all non-colorblind observers said A was better than B.

Figure 3 plots the shifts in the a* vs. b* plane measured from the prints. In the Original vs. CopyA graph (left) the vectors go in the same direction and have similar lengths. In the Original vs. CopyB graph (right) the direction and magnitude of the shifts are different. If we compare plots of L* vs. a* and L* vs. b* we would see the same pattern.

Observers selected Copy A, the image in which all areas had the same color shift, as the better reproduction. Observers chose the image that maintained the ratios across edges. The fact that the displays had constant ΔE values was not apparent to the observers. Using this experimental design we looked at a number of variations. In each case the observer preferred the systematic, constant ratio reproductions. These experiments support the other experiments pointing to the observation that color appearance is a spatial calculation in humans. Obviously, the response of the different spectral sensitivity receptors is very important, but falls far short of explaining the entire color vision calculation. ΔE is a function of quanta catch of retinal receptors, but cannot be used to evaluate color reproductions evaluated later in the visual system. Minimizing ΔE searches for best matches independent of spatial information; humans use spatial information to calculate contrast.

Spatial Color Metric

Color reproduction needs a different kind of Color Metric. It has to be based on the fundamental difference between imaging systems and human vision. All chemical and electronic imaging systems operate one pixel at a time. The exposure on a silver halide grain controls the concentration of metallic silver deposited at that pixel. The calories delivered to a thermal ribbon, the number of pulses delivered to the inkjet cavity, electrons delivered to the TFT cell control the response of that pixel. The signal sent to other pixels far away from the pixel in question have no effect on it. Human vision at the receptor level behaves the same way. The rod and cone receptors respond to the quanta caught by the chromophores in the cell. This is modeled well by colorimetry. But appearance is the relative response of all the receptors across the field of view. The metric for color appearance needs to build on colorimrtry and accumulate relationships all across the image.

When printer and display technologies introduce different color gamuts, the problem becomes more interesting. In fact most displays have much larger gamuts than prints in high lightness at full saturation. Most printers have larger gamuts than displays in low lightnesses at full saturation. Typically one finds that the common vol-

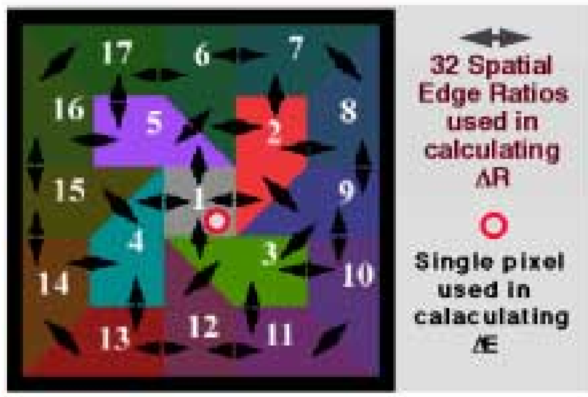


Figure 4 illustrates of the 32 adjacent edges found in "Original" image. The display has 17 different color patches. The double-headed, black arrows indicate edges between different patches in the image. The black circle represents the information at one pixel that is used to calculate Tristimulus Values of one area. The Retinex Color Metric for color appearance compares the Tristimulus Values across each of these 32 edges to find the best reproduction of the entire image.

ume of a three-dimensional color-space is half the combined display and printer volumes. A test image that represents all parts of display plus printer color space will reproduce accurately only half the pixels on a printer or on a display. (See papers by Marcu[17] and Morovic and Luo[5] for specific data.)

An approach to solving the problem is to use the information learned from the comparison of Copy A with Copy B. The human eye cares more about the relationships of the parts of the image than it does about the absolute value of the match. Figure 4 shows the Original image with identification numbers. Area 1 is the gray square in the center, the numbers are assigned in a clockwise spiral up to Area 17 on the top left. The round black circle in the bottom right corner of Area 1 represents the input information used in Colorimetry to calculate the Tristimulus values of Area 1. Relative colorimetry compares the information from a single pixel to the ratio of media white to illuminant; that is information that cannot be derived from the image itself. The media white and the illuminant have to be independently measured.

The double-headed black arrows in Figure 4 show the 32 comparisons possible between adjacent areas in the image. This is the information the human eye uses to calculate color appearance. We can propose a Color Metric more like human vision by comparing X from one area with X' of an adjacent area in the same image ($X_{\text{Copy A}}/X'_{\text{Copy A}}$). See Figure 5.

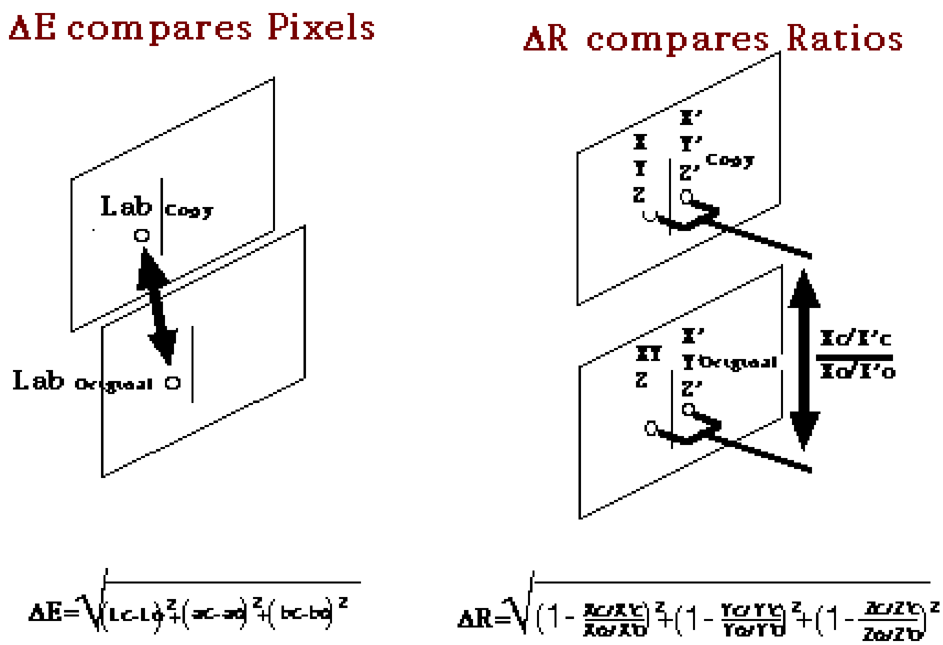


Figure 5. An illustration comparing ΔE, which uses corresponding pixels with ΔR, which compares corresponding ratios.

We can calculate a corresponding spatial comparison for the Original image ($X_{Original}/X'_{Original}$). We can compare the two with a ratio.

$$[\ (X_{CopyA}/X'_{CopyA})\ /\ (X_{Original}/X'_{Original})\ ]$$

The results calculated in X, Y, and Z for the 32 edges are plotted in Figure 6. Here we see that shifting all the colors the same amount (CopyA, Top Right) has little effect in altering the 32 ratios across edges. The ratio of ratios is very close to 1.0 for all 32 edges in X, Y, Z. In fact, the departure from ratios of 1.0 are experimental errors introduced by color media limitations. Individual shifts that mimic separate pixel-by-pixel color gamut mapping show significant shifts in the 32 spatial comparisons (CopyB, Bottom Right).

We use these ratio of ratios to formulate a spatial metric ΔR analogous to ΔE for pixels. We can calculate ΔR for each ratio in the image.

$$\Delta R=\sqrt{\left(1-\frac{Xcopy/X'copy}{Xorig/X'orig}\right)^2+\left(1-\frac{Ycopy/Y'copy}{Yorig/Y'orig}\right)^2+\left(1-\frac{Zcopy/Z'copy}{Zorig/Z'orig}\right)^2}$$

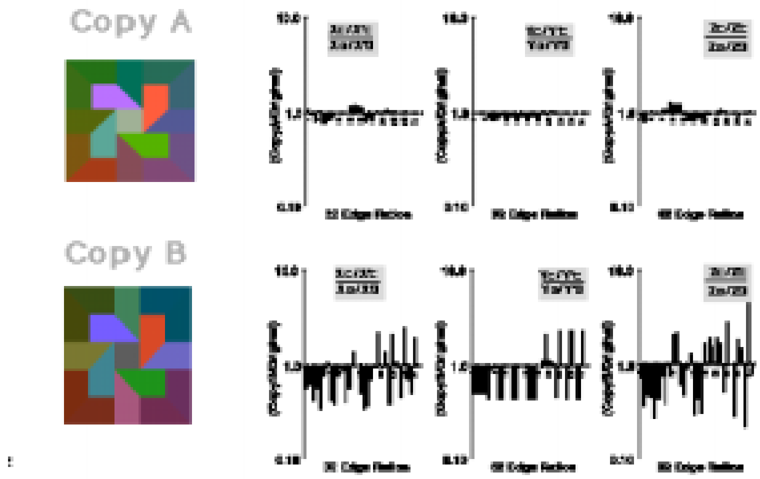


Figure 6. Result for the 32 edge ratios. A Color Metric imitating human vision compares X from one area with X' of an adjacent area in the same display image ($X_{CopyA}/X'_{Copy\ A}$). The metric compares the corresponding ratio data for the Original image ($X_{Original}/X'_{Original}$). We can compare the two with a ratio $[(X_{CopyA}/X'_{Copy\ A}) / (X_{Original}/X'_{Original})]$. The results X, Y, and Z for the 32 edges are plotted in here. Results for Copy A (top row) show that all the comparisons of X/X'(left column), Y/Y' (middle column) and Z/Z' (right column) fall very close to 1.0. The same analysis of Copy B shows that the comparisons of areas are very different for Copy B and the Original. The spatial metric can discriminate between CopyA and Copy B, while colorimetry cannot.

We can average ΔR to accumulate a net quantity that describes the entire image. Average ΔR for Copy A is 0.17±0.11 and 1.05±0.83. for Copy B. The spatial metric ΔR predicts that Copy A is six times better than Copy B. The displays were made to have equal ΔE's. The spatial metric analysis of Copy A and Copy B corresponds with observer reports, namely that Copy A is significantly better than Copy B.

This particular spatial metric is very primitive. X,Y, Z are proportional to cone response, but are not isotropic in color appearance. Future psychophysical experiments are needed to optimize the shape of an optimal spatial psychometric function. The point here is that spatial comparisons can discriminate between Copy A and CopyB while pixel comparisons cannot.

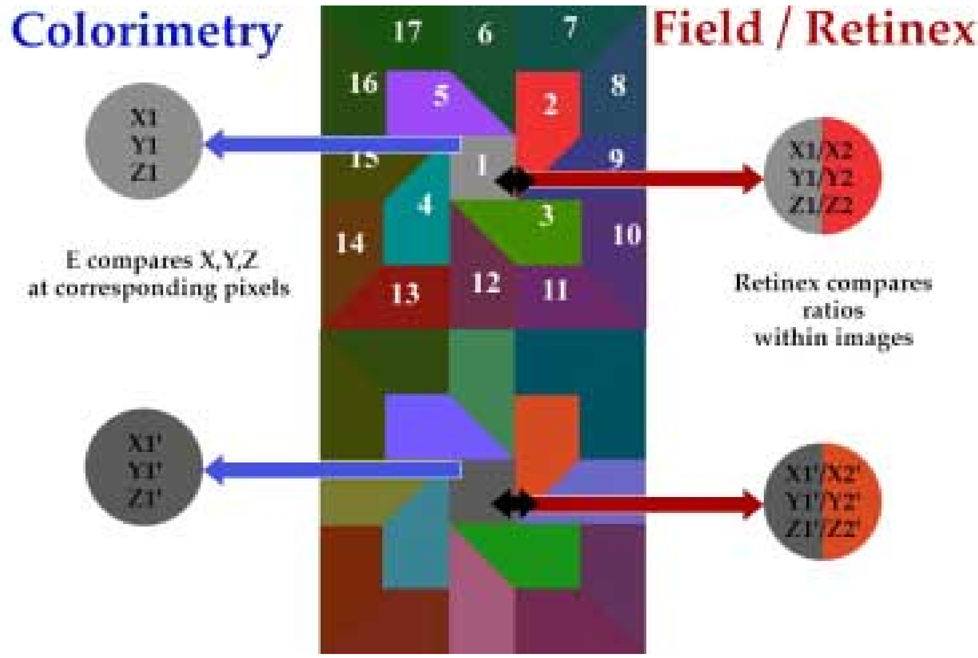


Figure 7 illustrates the different strategies used to calculate color appearance. The top square image is the Original. The bottom image is the reproduction called CopyB. Colorimetry restricts input data to one pixel. Field calculations, such as retinex, uses the ratios of radiances from different pixels. These images were created to have identical colorimetric differences, but different edge ratios. The observer evaluations are consistent with human vision using spatial comparisons rather than colorimetric information.

Figure 7 is a summary of the strategic ideas established so far. The central theme is the decision between evaluating color one pixel at a time, as in colorimetry, as opposed to spatial comparisons. The Original (Top) is compared with the Copy B reproduction (Bottom) one pixel at a time. The data from the rest of the image is irrelevant. Retinex and field calculation advocate relative spatial information as the basis for color appearance. Absolute intensity at a pixel is less important than spatial comparisons.

The color demonstration shown in Figure 2 was designed to have constant ΔE, and different edge ratios. When the edge ratios were nearly constant the Copy A makes a fairly good reproduction despite the $\Delta E=17$. When the edge ratios were distorted in Copy B, the appearance of the display changed. Figure 6 shows that the ratios =1.0 characterize a good copy. Ratios that depart from 1.0 characterize the bad copy. Local ratios provide an important tool in finding a meaningful Color Metric.

The comparison of Copy A and Original shows an additional important point. Although Copy A has the same edge ratios as the Original, it does not match the Original. Edge ratios are an important tool for successful metric, but not a complete metric by themselves. Human vision normalization depends on both spatial and absolute quanta catch information.

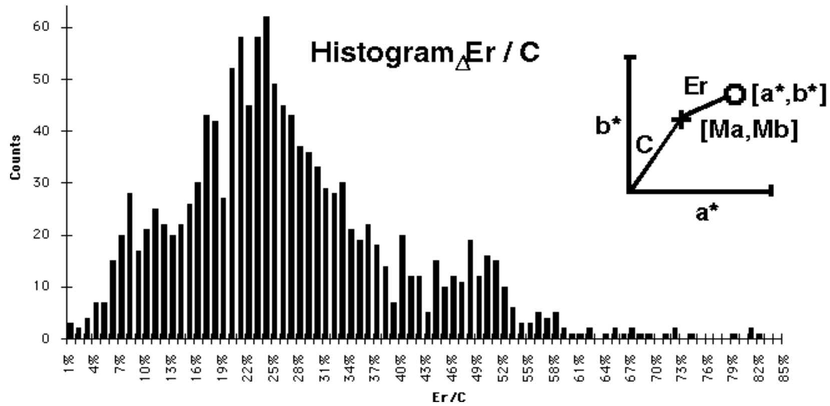


Figure 8. The coordinates in the upper right corner describe the definition of vectors C and Er. The point +[Ma, MB] is the location in color space selected by Nickerson's experiments. C is the distance from the origin to [Ma, Mb]. The point O is at the coordinates assigned to the same X, Y, Z by CIEL*a*b*. The vector Er is the error introduced by L*a*b*. The histogram of the ratio [Er / C] for all 1731 real chips in the Munsell book is shown on the left. The maximum error is 82%, the minimum is 0.01%. The mean is 27%; the standard deviation is 13% and the median is 24%.

4. Working in an Isotropic Color Space

An important tool in the color gamut calculation is the minimization of the discrepancy between the original and the copy. Minimization programs are frequently used to find optimal compromises. The hard part is to provide the Color Space that quantifies the color mismatch. First, this space should isotropic, that is equally spaced in appearance in all three dimensions. That is, a forced compromise between hue, chroma and lightness must be made in a space that has been experimentally verified as isotropic in these dimensions. Second, it should be based on a large body of data across many observers and viewing surrounds. Third, it should include data from reflective papers and extrapolations to the spectrum locus. Displays using phosphors and LEDs emit colors on, or close to, the spectrum locus in x,y. Once the Color Space has been shown to be appropriate, then a computer can reliably find the best compromise.

Except for Lightness, CIEL*a*b* distorts the design principles of the Munsell book.[18-21] Newhall et. al.'s observers selected hue chroma and lightness to be equally spaced.[22] To facilitate comparison between Munsell and L*a*b*, we can convert the original Munsell Notation by scaling Chroma and Hue proportional to Lightness. The Munsell Book uses chips that are spaced 1/ units apart in Lightness. The Munsell real papers cover the range 9.5/ to 1.75/. L* covers the range of 100 to 0. It is convenient to set ML = 10*Munsell Value. The chips on a given page of the book have the notation /2. To make them isotropic with value MC = 5* Munsell Chroma.

A Munsell designation of 5 Y8/10 becomes ML = 80, Ma = 0, Mb = 50 (MLab notation)[21,17]. In other words, we will use the position in color space selected by

Munsell, but describe it the same 3-space notation as L*a*b*. If we assume that the experimentally based Munsell space is our starting point, then departures from it introduced by L*a*b* can be described as errors introduced by the equations. Virtually all the errors are in the a/b plane. We can evaluate the magnitude of the errors by measuring the distance between [a, b] the CIE L*a*b* position and [Ma, Mb] the MLab position in the a/b plane. In the upper right of Figure 8 we see a plot in the a*, b* plane. The vector C from the origin to [Ma, Mb] shows the position chosen by observer data. The vector Er shows the magnitude of the error introduced by L*a*b*. The ratio Er/C is plotted in Figure 8. The chroma error introduced by L*a*b* averages 27%[18].

5. Experiment in Munsell Space

It order to remove any distortions in the demonstration shown in Figure 2, we can make new displays using Munsell Notation, instead of L*a*b*. To illustrate the point that colorimetric distance is a poor predictor of best color appearance we can study Figure 9.[7] Here we have made a set of nine different "Umbrella" displays. We call the one in the middle [E] the Original. All the others are reproductions. Which of the 8 different reproductions are acceptable and which are not?

Design of Umbrella Targets

The specifications for all areas are in Munsell notation. This avoids the problems introduced by the inconsistencies found in L*a*b* spacing.[18] The "Original" [E] has 10 triangular areas. The top area has the same L*a*b* values as Munsell paper 5R 7/6. The remaining 9 areas are all 7/6, each of them are 4 pages apart [5YR, 5Y, 5GY, 5G, 5BG, 5B, 5PB, 5P, 5RP]. Thus we have placed the 10 areas in a circle of constant lightness and chroma in color space.

All 10 colors in all the remaining 8 targets differ from the original by only one chip in the Munsell Book. The [DEF] series varies in lightness. Each corresponding area in [D] is 8/, and is 6/ in [F]. This change in stimulus is a global shift. It produces noticeable, but acceptable changes in appearance. The changes for all 10 areas are about the same magnitude.

The [AEI] series varies in chroma. Here all 10 papers have /4 in [A] and /8 in [I]. The chances are noticeable, but regarded as acceptable reproductions.

The [BEH] series varies in hue. Here 5R is replaced with 2.5R in [B] and 7.5R in [H]. The same shift was applied to all 10 areas. The chroma /6 was chosen so that one hue page in the Munsell book equals one chip in lightness value and one chip in chroma. Here the change in color appearance is small, but noticeable and acceptable as good reproductions.

So far all the changes have been global shifts. All 10 papers have moved in the same direction in color space. Such uniform movement is reminiscent of the changes found in color constancy experiments in which the paper display is constant, while

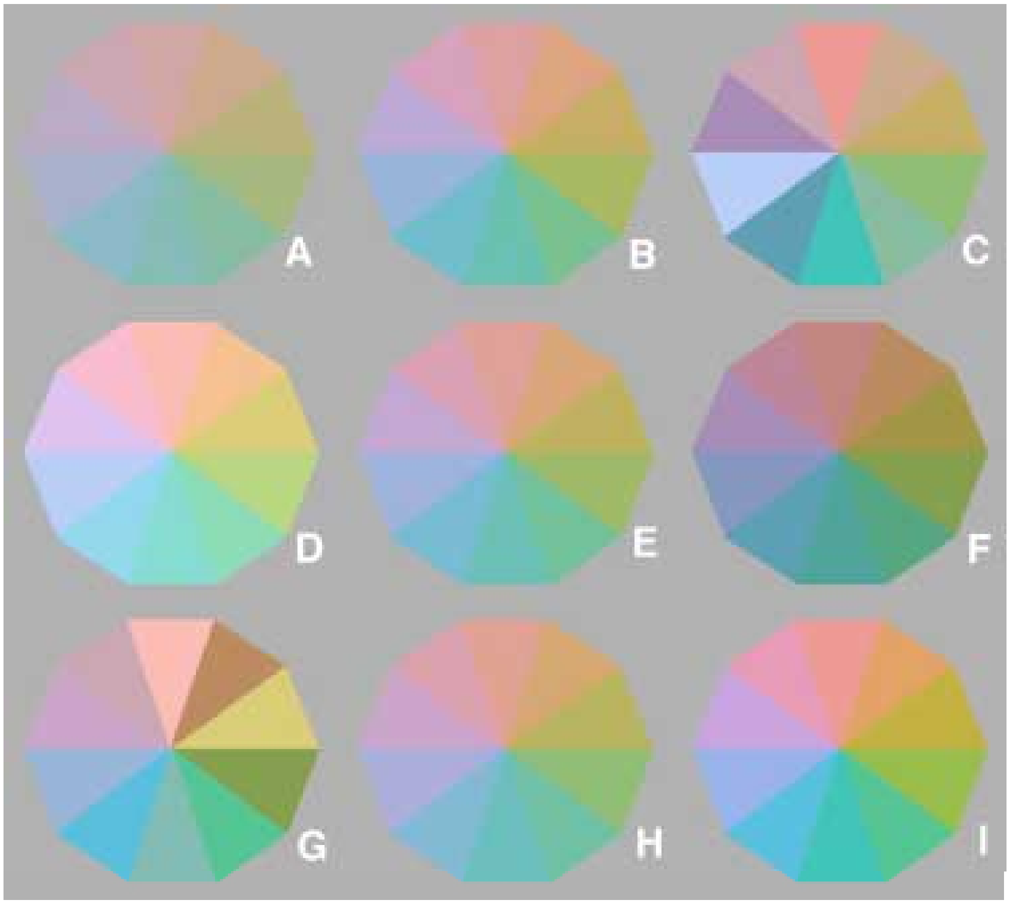


Figure 9. An “Original” display [E] surrounded by 8 reproductions. In all cases there are 10 pie-shaped color patches. For all 8 reproductions each individual patch differs from the original by one chip in the Munsell Book of Color. That means that each individual chip is a constant distance from the original in the Munsell Uniform Color Space. Thus, changes in lightness, chroma and hue are equal. Along the left-to-right axis [D-E-F] the Original and reproductions vary only in lightness. All 10 patches in D are one Munsell chip lighter than the original. All 10 patches in F are one Munsell chip darker than the original. Along the top-to-bottom axis [B-E-H] the Original and reproductions vary only in hue. All 10 patches in B are shifted counterclockwise one Munsell chip from the original. All 10 patches in H are shifted clockwise one Munsell chip from the original. Along the upper-left to bottom-right axis [A-E-I] the Original and reproductions vary only in chroma. All 10 patches in A are one Munsell chip less saturated than the original. All 10 patches in I are one Munsell chip more saturated than the original. All of the above reproductions [ABDFHI] are still reasonable reproductions despite the one chip color shifts. The remaining two reproductions C and G are examples of individual color shifts in hue, lightness and saturation. Unlike systematic shifts, individual shifts create unacceptable distortions of the original. An error of 1 Munsell chip is acceptable if it is global, but not if it is local.

the intensity and color of the illumination changes. In color constancy experiments we are familiar with this kind of result, namely that colors change very little with large global shifts in illumination. As well, it is common practice in photographic systems to find small apparent changes with large global density shifts, while local changes of the same magnitude are clearly visible.

The final [CEG] series shows the effect of pseudorandom color shifts. As above, each color patch is one Munsell chip different from the Original [E]. Nevertheless, independent changes produce both large and small visual changes in appearance. These local area changes make for poor reproductions. They distort the appearance of the original in way that make them unacceptable reproductions.

The analogy to color constancy is very compelling. Changing the colors independent of the neighbors disrupts the spatial ratios (Figure 1). Changing the local ratios in the context of a color constancy experiment is the same as changing the reflectances of the areas. Changing reflectances of individual papers (local shifts) cause big changes in appearances, while changes in illumination (global shifts) cause small changes. It should be noted that changing the spectral character of the illumination is not the direct analog of changing lightness, hue and chroma, as we did in the above umbrella experiment. The direct analog of color constancy experiments would be to uniformly lower or raise the long-wave reflectances of all the papers, and make analogous changes to the middle-, and short-wave reflectances. However, we began this experiment with the design of substituting paper of known differences of color appearance. That has been provided by the Munsell Book data. Shifting long- middle- and short-wave reflectances may be a better experiment, but equal color difference data is not available.

Evans describes a "consistency principle" for color reproductions. He says, "As long as the chromatic relations within a given scene are consistently like those of the original scene, the colors will actually seem to be the same in both, unless of course, direct point-for-point comparisons are made. The actual discrepancies can be enormous, providing there is no inconsistency."[25] Evans attributes this obsevation to poor color memory. By introducing spatial comparisons we can add computational properties to Evan's principle. The suggestion from these experiments is that color gamut calculations, using spatial comparisons can lead to better in-gamut reproductions. Colorimetrically they will have larger color-difference errors, but they should look better. The same color constancy mechanism, that reduces large physical shifts in illumination to small appearance changes, can be employed to make gamut-limited reproductions look better.

6. Retinex - A Spatial Calculation

There has been little change in the fundamental operation of Retinex model (Figure 10) since first proposed in 1967 at Land's Ives Medal Address to the Optical Society of America.[1] The original proposal used the Ratio, Product, Reset and Average. Figure 10(Right) documents the details of the a Retinex model. There are only four operations: Ratio, Product, Reset and Average. Ratio makes spatial comparisons. The product propagates local comparisons in nearly global ones. Reset performs a normalization process relative to the maxima in the image. Average combines the different reports of spatial information from different locations. In implementing these calculations we convert the input to log radiance. The consequence is that ratio and product operations are simplified to subtraction and addition. Reset is a simple logical operator, that normalizes the image.

The model, that evolved from the study of Mondrians, can as well calculate appearances in color constancy experiments, high-dynamic-range real life scenes and Gestalt phenomena. Examples of these images can be found in a recent summary publication.[16] The details of the spatial calculations, along with sample Matlab code can be found in Funt, Ciurea and McCann.[23]

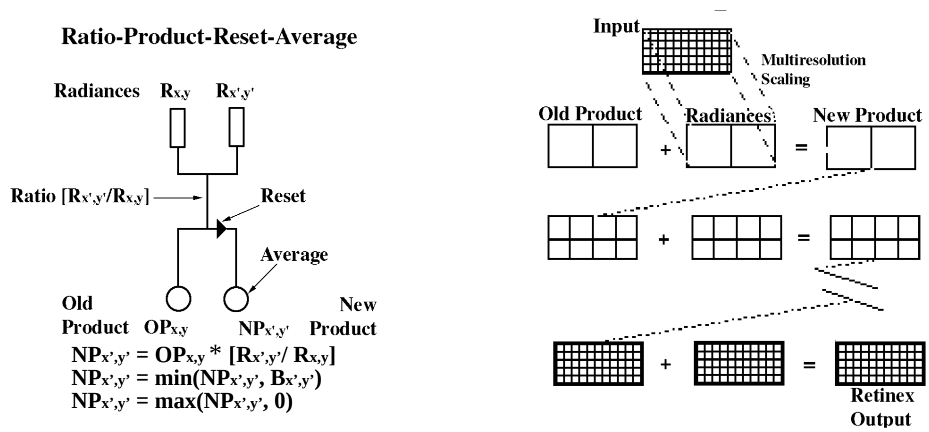


Figure 10 (Left). The explanation of Ratio-Product-Reset-Average operation. Here we calculate the New Product (NP) for the output pixel x',y'. We begin at the starting pixel x,y using the Old Product (OP). All OP's are initialized with the value in the Best image [Bx,y] for that waveband. The product of the radiance Ratios times the Old Product is reset if greater than the Bx,y and averaged with the previous New Products. Figure 10(Right). An illustration of the Multiresolution aspect of the Retinex calculation[23]. The calculation uses three data planes. The Old Product is initialized to the Best image. The original full-resolution image is illustrated as Input at the top. The input is averaged down to make a series of multiresolution planes ending with two pixels. This average Radiance image is the second data plane. The third data plane is for the output of each iteration and is called the New Product. Starting with two pixels we multiply the Old Product at the starting pixel and multiply it by the ratio of Radiances for the starting and output pixels. That product is reset to the Bestin Old Product value, if it exceedes it. It is then averaged with previous Old Product at the output pixel. To get to the next level, the New Product is interpolated to twice the size and placed in the Old Product data plane. The Radiance data plane uses the next larger (8 by 2) average of the Input. The Ratio-Product Reset-Average calculation illustrated in Figure 9 (Left) are repeated. The Process continues until New Product at full-resolution is complete and is used as one Retinex separation. The process is repeated for each of the three R,G,B separations.

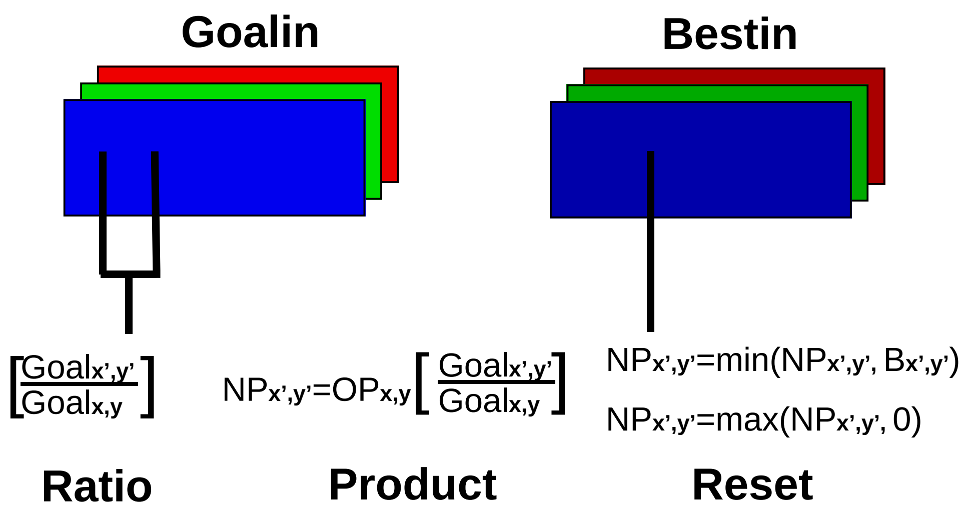


Figure 11. Schematic diagram of the Color Gamut Retinex Calculation. This calculation creates long-, middle-, and short-wave separations for both Goalin and Bestin images. It uses the ratios from The Goalin image and reset from Bestin image to put spatial comparisons in the search for best reproduction using a limited gamut.

Color Gamut Calculation

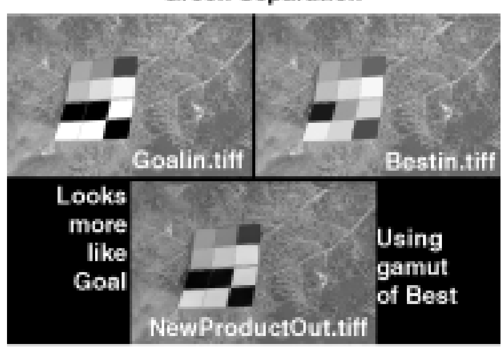
The hypothesis connecting the experiments in this paper is that humans calculate color using spatial comparisons. The above experiments show that the sum of errors (distances in color space) is a very poor predictor of the quality of a reproduction. In fact, good reproductions make all their errors in similar directions.[7]

If that premise is true, then spatial comparisons could be helpful in finding a set of in-gamut colors that look like the out-of-gamut original. Figure 11 illustrates the color gamut Retinex calculation. For gamut mapping we begin with two input images, instead of one. We have the Goal image that has the large, original gamut. Second, we have the Best image that represents the limited gamut of the reproduction media. If the shape of the limited gamut is complicated, we may substitute a three-dimensional LUT for the Best image. Again, we begin by averaging down each of the R, G, B separations to a small number of pixels for both the Goal and the Best image (Figure 10 Right). We take the Old Product initialized to maximum and multiply by the Goal ratio. This New Product is reset to the Best image or the Best data LUT. This process is repeated and the New Product values from this resolution are interpolated up to the next resolution. The process is repeated for R, G, B.

The gamut retinex process takes the spatial comparisons from the Goal and limits the product by the Best image (Figure 11)[4,24,28]. The iterative process keeps reinforcing the ratios found in the Goal while the reset forces the New Product to migrate toward an image with all the same ratios, regardless of the absolute input values of the

Blue Separation

Green Separation



Red Separation

NewProductOut images. Figure 12a (Above) shows the 3 color images. Here we see that the NewProductOut image looks much closer to the Goalin image than the Bestin. Both Bestin and NewProductOut are within the limited gamut of the Toyo inks (uncoated). Figures 12b, 12c & 12d (Right) show the blue, green and red color separation images. By comparing the NewProductOut image with the Goalin image we can see that the gamut retinex product preserved the spatial relationships in each separation. The gamut Toyo profle that modified the Goalin image to make the Bestin image, adjusted each pixel independent of all the others. The NewProductOut image used spatial comparisons and gamut information to calculate the value at each pixel. The improvement in matching appearances can be seen by studying the three sets of color separations. In each image the 12 patch Jobo target shows us the rendition of 6 gray areas in the top two rows and the colors blue, green and red above yellow, magenta and cyan in the bottom two rows.

Goal image. The resulting image NewProductOut shows a big improvement in appearance compared to the Best (Figure 11). The data shows that this new image is in-gamut.

A demonstration of the process can be found in Figure 12. The Goalin.tiff image is a scanned photograph of a Jobo photographic test chart. In Photoshop the six color patches have been replaced with the most saturated colors available. Red digits are [255,0,0]; green digits are [0, 255,0]; blue digits are [0, 0, 255]; yellow digits are [255, 255,0]; magenta digits are [255, 0, 255]; cyan digits are [0, 255, 255].

The Besin.tiff image was created in Photoshop from the modified Goalin.tiff image using the Toyo (uncoated) profile. The Toyo (uncoated) gamut is much smaller than the display. The NewProductOut image is the output of the calculation described in Figures 10 and 11.

In understanding these results it is helpful to study both the color rendition and the R, G, B separation images. The Goalin, Bestin and NewProductOut images are shown in color in Figure 12 (Top Left). Figures 12b, 12c, and 12d show the R,G,B separations.

In Figure 12b, the blue separation, the Goalin colors of [blue, green, red] are [white, black, black]; while [yellow, magenta, cyan] are [black, white, white]. The Bestin image is much lower in contrast. All the blue separation values are nearly the same. The spatial gamut retinex process [NewProductOut] finds in-gamut combinations that maintain, as well as possible, the lightness differences for each color. In this blue record the Goalin whites have been limited to NewProductOut light grays. However, the blacks associated with red, green and yellow have remained black. The comparison with the Bestin image is very interesting. The Toyo ink profile found blue separation values much lower in contrast, in fact all the lightness are close to middle gray.

In Figure 12c the green separation the Goalin colors of [blue, green, red] are [black, white, black]; while [yellow, magenta, cyan] are [white, black, white]. Again the spatial gamut retinex process finds in-gamut combinations that maintain, as well as possible, the lightness differences in each color separation. In this record the Goalin whites have been limited to NewProductOut light grays for green and cyan. However, the separation values associated with red, blue and magenta have remained very close to black. The comparison with the Bestin image shows that the non-spatial, Toyo profile found generated a green separation with much lower in contrast.

In Figure 12d, the red separation, the Goalin whites have been limited to NewProductOut light grays for yellow and light-middle gray for red and magenta. The rendition of the other colors is close to black. The comparison with the Bestin image shows that the Toyo profile generated green-patch values much lighter than desired.

Fig. 13 plots the Goalin and NewProductOut values of the six colored areas in ML, Ma, Mb space. The Goalin values are plotted as large solid red squares; the NewProductOut are plotted as solid blue circles. The solid lines are drawn between Goalin and the NewProductOut values. These lines are beyond the NewProductOut values. Fig 13 (Left) plots the data in the Ma, Mb plane. Here many of the vectors pass near the central point (Ma = Mb = 0), but do not intersect there. Fig 13b plots the data in the ML, Ma plane. Here the vectors pass near the central point (ML =50, Ma = 0), but do not intersect there. Fig 13c plots the data in the ML, Mb plane. Here the

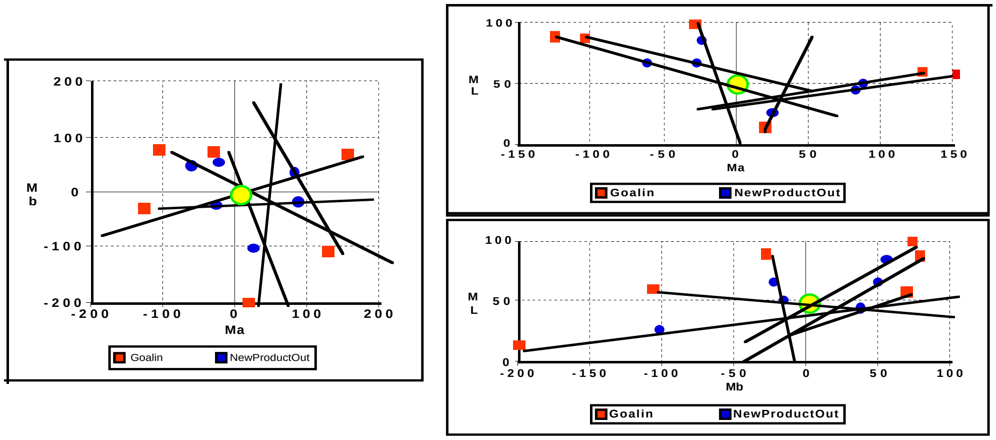


Figure 13. (Left) The plot of the 6 Goalin to NewProductOut vectors in Ma vs. Mb plane. The plot of 6 Goalin to NewProductOut vectors in ML vs. Ma plane (Top Right). The plot of 6 Goalin to NewProductOut vectors in ML vs. Mb plane (Bottom Right). The yellow circular spot plots the coordinates ML=50, Ma=0, Mb=0. The square grid represents equal distances in MLab space.

vectors pass below the central point (ML = 50, Mb = 0). These graphs show the improved color seen in NewProductOut image are not caused by simple color space projections as: reduction in chroma (Fig 13 Left), projecting toward middle gray (Figs 13 Top Right & 13 Bottom Right). The NewProductOut image was created by optimizing spatial comparisons. Such a process is somewhat similar in that the results project in the vicinity of middle gray (ML = 50, Ma = 0, Mb = 0). However, their projections sweep out a substantial, non-symmetrical volume. This is an argument that the Gamut Retinex process is fundamentally different from a single, 3D color space transformation.

The lack of convergence to a single point in color space is similar to projections generated by the CARISMA and Lee[26] algorithms. CHARIMA used empirical non-global transformations. Further, it is based on how experienced printers would treat images on small-gamut media. Printers are trained to interpret images in terms of the quality of their separartions. As we saw in Figure 12, the gamut retinex calculation retains the appearance of objects in their separations and hence does well in color.

Kang, Cho, J. Morovic and Luo[27] recently described experimental matches to find the observers' choice of best compromise for each color independently. Their data lead to a model with three ranges. In the upper level (high lightnesses) there is a high point of convergence, in the middle level (middle lightnesses) there is no point of convergence, in the lower level (low lightnesses) there is a lower point of convergence. MacDonald, Morovic, and Xiao[28] described their recent addition to GMA models. Here they described three zones. The first is the region between small gamut and large gamut. The third is the region in color space close to the gray axis (vertical axis in L*a*b* space) that is unchanged by the GMA. The second is the region between the two that is shaped by the GMA. MacDonald et. al. have described a new method that introduces a different point of convergence for different lightnesses. Green and Luo[29] extended CARISMA to the conversion of transparency to newsprint workflows. The included comparisons of CHARISMA with algorithms by Lee and Kang.

A common property between CARISMA, Lee's , Kang's and MacDonald's obsevations and models is that there is no single projection to a single convergence point in color space. The "Out-of-Gamut to Gamut" transforms got their speed and efficiency from the fact that there was only one rule. (Project all the out-of-gamut points to a common point, such as L*=50, a*=0, B*=0). As already mentioned, observers prefer multiple end points and complex transforms. The more complex the transform, the greater the colorimetric error between the large-gamut original and the optimal small-gamut reproduction.

Best Color Gamut Compromise

A familiar gamut mapping process is to evaluating the absolute colorimetry of a pixel, see if it is in-gamut, and then replacing it with the nearest in-gamut color. This process distorts color appearance. Take two areas next to each other. Let us assume that one area is in-gamut and the other is not. If we leave the in-gamut pixel value unchanged, while changing the out-of-gamut pixel, we have replaced the ratio of these two areas with a new ratio and a new color relationship[30]. This paper argues that it is better to change both pixel values, so as to leave the spatial comparisons constant. The best reproduction is the one that reproduces the most spatial comparisons.

7. Conclusions

Retinex calculations extended to the problem of gamut limited reproductions show promise. The argument states that global shifts in color similar to those found in color constancy produce much smaller changes in appearance than local, individual color shifts. Further, this paper argues that color gamut transformations using spatial comparisons can generate in-gamut reproductions that look more like the original, because it employs the benefits of human color-constancy processing. These reproductions have a greater colorimetric difference between original and reproduction, but look better. Color is a spatial calculation in humans.

Acknowledgments

I want to thank Norbert Herzer, Mike McGuire, Irwin Sobel, Paul Hubel, and Mary McCann for their thoughtful discussions and suggestions.

References

[1] E. H. Land and J. J. McCann, "Lightness and Retinex Theory", *J. opt. Soc. Am.* **61,** p. 1, (1971).

[2] J. J. McCann, S. P. McKee and T. Taylor, "Quantitative Studies in Retinex theory, A comparison between theoretical predictions and observer responses to Color Mondrian experiments", *Vision Res.,* **16**, pp. 445-458, (1976).

[3] J. Frankle and J.J. McCann, "Method and apparatus of lightness imaging", U.S. Patent 4,384,336, (1983)

[4] J.J. McCann, "Lessons Learned from Mondrians Applied to Real Images and Color Gamuts", Proc. IS&T/SID Seventh Color Imaging Conference, pp. 1-8, (1999).

[5] J. Morovic and R. Luo, "Developing Algotithms for Universal Colour Gamut Mapping", in Colour Imaging Vision and Technology, Ed. L. W. MacDonald and M. R. Luo, Wiley & Son Ltd., Chichester, pp. 253-282 1999.

[6] Members of CIE Technical Committee 8–03. Ján Morovic – chairman (Colour & Imaging Institute, UK), Peter Bodrogi (Veszprém University, Hungary), Gus Braun (Eastman Kodak, USA), Fritz Ebner (Xerox, USA). Mark D. Fairchild (Rochester Institute of Technology, USA), Phil Green (London College of Printing, UK), Patrick Herzog (RWTH, Germany), Tony Johnson (UK), Byoung–Ho Kang (ETRI, Korea), Naoya Katoh (Sony, Japan), M. Ronnier Luo (Colour & Imaging Institute, UK), Marc Mahy (Agfa, Belgium), Gabriel Marcu (Apple, USA), John McCann (USA), Ethan Montag (Rochester Institute of Technology, USA), Hideto Motomura (Matsushita Research Institute Tokyo, Japan), Todd Newman (Canon Information Systems, USA), Raimondo Schettini (Italian National Research Council, Italy), J. A. Stephen Viggiano (RIT Research Corporation, USA), Geoff J Woolfe (Eastman Kodak, USA). See website [www.colour.org/tc8-03].

[7] J.J. McCann, "Making Reproductions of Masterpieces: Managing all the Colors within the Gamut", Proc. International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives, Chiba, (1999).

[8] J. McCann, "Human Color Perception" in Color Theory and Imaging Systems, Society of Photographic Scientists and Engineers, R. Eynard, ed., Washington, p. 1, 1973.

[9] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings", J. Phyiol. (Lond.) 197, 551-566, (1968).

[10] O. Braddick, F. W. Campbell and J. Atkinson, "Channels in Vision: Basic Aspects", in Handbook of Sensory Physiology VII, R. Held, H. W. Leibowitz and H. Teuber, pp. 3-38, (1978).

[11] J.J. McCann, "Visibility of gradients and low-spatial frequency sinusoids: Evidence for a distance constancy mechanism", *J. Photogr. Sci. Eng.,* **22,** 64-68, (1978).

[12] H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision", Vision. Res., 19, 19-32, (1979).

[13] R. Luo, "Colour Science: Past, Present and Future", in Colour Imaging Vision and Technology, Ed. L.W. MacDonald and M. R. Luo, Wiley & Son Ltd., Chichester, pp. 381-404,(1999).

[14] CIE "Recommendations on Uniform Color Spaces, Color Difference Equations, Psychometric Color Terms", Supplement No. 2 of CIE Publ. No 15 (E-1.3.1) 1971, Bureau Central de la CIE, Paris, 1978.

[15] CIE Proceedings 1931, p. 19, Cambridge University Press, Cambridge, (1932).

[16] J.J. McCann, "Color Theory and Color Imaging Systems: Past, Present and Future", *J. Imaging. Sci. and Technol., 42, 70, (1998).*

[17] G. Marcu, "Gamut Mapping in Munsell Constant Hue Sections", in Proc. 6th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, p. 159, (1998)

[18] J.J. McCann, "Color spaces for color mapping", *J. Electronic Imaging, **8**, pp. 354-364, (1999);*

[19] 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, Proc. SPIE, **3648,** pp. 68-82, (1999).

[20] J.J. McCann, Uniform Color Spaces: 3D LUTs vs, Algorithms, Proc. Second IS&T PICS Conference,

2, pp. 204-208, (1999);
[21] [J.] J. McCann and M. Stokes, “Color Spaces and Image Quality”, Proc. IS&T PICS, **1**, *pp.* 140-144, (1998)..
[22] 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).
[23] B. Funt, F. Ciurea and J. J. McCann, “ Retinex in Matlab” in Proc. 8th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, 112-121, (2000).
[24.] J.J. McCann, “Using Color Constancy to Advantage in Color Gamut Calculations, Proc. Third IS&T PICS Conference, **3,** pp. 169-176, (2000).
[25.] R. M. Evans, “The Perception of Color”, John Wiley & Sons, p 205, (1974).[26.]B. Kang, M. Cho, J.Morovic and R. Luo, “ Gamut Compression Algorithm Development on the Basis of Observer Experimental Data” in Proc. 8th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, in 311-317, (2000).
[26.] C. S Lee, K. Sim, E. J.Lee &Y. A. Ha, “Gamut mapping using variable anchor points”, in Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts IV, Proc. SPIE, **3648,** pp.126-130, (1999).
[27.] B. Kang, M. Cho, J.Morovic and R. Luo, “Gamut Compression Algorithm Development on the Basis of Observer Experimental Data” in Proc. 8th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, 311-317, (2000).
[28.]L. MacDonald, J. Morovic, and K. Xiao, “A Topographic Gamut Mapping Algorithm Based on Experimental Observed Data” in Proc. 8th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, in 311-317, (2000).
[29.] P. Green and R. Luo, “Color Gamut Measurements and Mapping: The Role of Color Spaces, in Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VI, Proc. SPIE, **4300,** pp. 139-149, (2001).
[30.] J. J. McCann, “Color Gamut Mapping Using Spatial Comparisons”, in Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VI, Proc. SPIE, **4300,** pp.126-130, (2001).