Color Theory and Color Imaging Systems: Past, Present and Future

John J. McCann*

Consultant, 161 Claflin Street, Belmont, Massachusetts 02178

James Clerk Maxwell demonstrated the first color photograph in a lecture to the Royal Society of Great Britain in 1861. He used the demonstration to illustrate Thomas Young's idea that human vision uses three kinds of light sensors. This demonstration led to a great variety of color photographic systems using both additive and subtractive color. Today, we have image-capture devices that are photographic, video, still, and scanning. We have hardcopy printers that are electrophotographic, ink jet, thermal and holographic, as well as displays that use cathode ray tubes, liquid-crystal and other light emission color devices. The major effort today is to get control of all these technologies so that the user can, without effort, move a color digital image from one technology to another without changing the appearance of the image. The strategy of choice is to use colorimetry to calibrate each device. If all prints and displays sent the same colorimetric values from every pixel, then the images, regardless of the display, would appear identical. The problem with matching prints and displays is that they have very different color gamuts. A more satisfactory solution is needed. In my view, the future emphasis of color research will be in models of human vision. The purpose of these models will shift from calculating color matches to calculating color sensations. All the technologies listed above work one pixel at a time. The response at every pixel is dependent on the input at that pixel, regardless of whether the imaging system is chemical, photonic, or electrical. Humans are different. The color they see at a pixel is controlled by that pixel and all the other pixels in the field of view. Human color vision uses a spatial calculation involving the whole image. Except for human vision, all other color systems have the same output from a single input. In other words, if an input pixel has a value of 128, and the image processing changes that value to 155, then all pixels with 128 in will have 155 out. Human vision is unique among color imaging systems because a single input value (128) will generate a range of output values (0, or 55, or 128, or 255), depending on the values of other pixels in the image. Despite the remarkable progress in our ability to control the placement of dyes and pigments on paper, we must now return to the study of Maxwell's interest—color theory—for the next advancements in color systems. In the future, we will see more models that compute the color appearance from spatial information and write color sensations on media, rather than attempting to write the quanta catch of visual receptors.

Journal of Imaging Science and Technology 42: 70–78 (1998)

The Past: The Conflict Between Physics and Psychology

Human interest in color images goes back to the Lascaux caves (about 25,000 B.C.) in northern Spain and southern France.[1] Here the cave artists used oxides of iron for red, oxides of manganese for blue–black and dark brown, and iron carbonate for yellow. Painted images have accompanied human activities throughout history.

The centerpiece of the intellectual rebirth in the Renaissance was the explosion of imagery, both secular and religious.[2] Perspective was rapidly developed and quickly accepted as a standard tool for imaging. Trompe l'oeil was reinvented by painters' interest in light and shadows. Color was used with a boldness that often is overlooked today.[3] The most notable example is the restoration of the Last Judgment and the Sistine Chapel ceiling.[4] Many people were astonished when the Vatican Museums painstakingly removed 500 years of candle soot and horse glue applied by previous centuries' conservators. The removal of this veil of light absorbers revealed the bright, brilliant images of the Renaissance.

The Pixel Tradition from Physics. Monochrome photography burst forth in a very dramatic way between 1816 and 1851. In that short time Niepce successfully fixed camera images, Dagauerre invented the positive metallic images that bear his name, Fox Talbot developed the negative-to-positive system so common today, and Archer invented the wet collodion plate that took photography away from the dark room.[5] The first "color photographs" were made by the application of paints and dyes on what we now call black and white photographs. Such images were common by 1850.[6]

In 1861, James Clerk Maxwell made the first color imaging system that recorded different spectral records and generated a multispectral chemical color image. The image of the colored ribbon (Fig. 1) was a demonstration in a discourse to the Royal Institution, London. Maxwell illustrated Thomas Young's idea that the human eye had three different types of light receptors: one sensitive to long-wave (red), one sensitive to middle-wave (green), and one sensitive to short-wave (blue) light.[7,8,*]

Maxwell also wrote the first color matching equations. He experimented with circles of paper, slit along their radii. Two or more of these papers can slide over each other,

Orginail manuscript received September 3, 1997.

* IS&T Member

* The interesting observation made by R. Evans was that although the demonstration worked, it did not work exactly as Maxwell described. Silver halide is inherently sensitive to blue and ultraviolet light, and insensitive to red and green light. The red, green, and blue filters modulated the blue and ultraviolet light to generate three different separations.

sharing the same perimeter and allowing variable percentages of each paper to occupy the front surface. To make a middle gray he arranged black and white papers to each show 50%. He used a top carrying a circular plate whose circumference was divided into 100 equal parts. Two sets of papers rested on the plate: a large diameter set on the bottom and a smaller diameter on top. When he spun the papers at high speed, the black and white fused to gray.

In one such experiment presented to the Royal Society of Edinburgh in 1855, Maxwell used papers prepared by Mr. T. Purdie with the unmixed pigments used in the arts. He used 37 parts vermilion, 27 parts ultramarine, and 36 parts emerald green to match 28 parts snow white and 72 parts black. This led to the first color matching equation.[†,9]

$$37\ V + 27\ U + 36\ EG = 28\ SW + 72\ Bk$$

The twentieth century experiments by David Wright[10] and others, led to international standards of colorimetry.[11] These equations can predict with great accuracy whether two stimuli viewed next to each other will match. These colorimetry equations have their predictive ability because they mimic the spectral response of the retinal receptors.[12] Unfortunately, many people mistakenly use colorimetry to predict color appearance. Gunter Wyszecki, co-author of the most extensive compilation of 20th-century colorimetry experimentation,[13] would warn his students about hand-painted chromaticity diagrams. Wyszecki wrote, "...tristimulus values and thus the chromaticity of a color stimulus do not offer any direct clues as to what color will be predicted. A multitude of other parameters concerning the viewing conditions must be considered before perceived colors can be predicted."[14]

In 1890, F. E. Ives patented the idea that photographic films should use colorimetric primaries.[15] The curves Ives used in the patent were the color mixture curves of Maxwell.[16] Mees and Pledge did an extensive study on film spectral sensitivity and did not to use film sensitivities that mimicked human sensitivities as suggested by Ives. Instead, they decided to emphasize color analysis incorporating non-overlapping spectral sensitivities.[17] So far, no industrial product has used colorimetric primaries as sensors to record the light from real life scenes. The overlap in spectral sensitivities of human cone pigments is so great that it allows almost no color isolation. By skillfully selecting film sensitivities, color photography successfully records virtually all colors.

Additive Color. In 1891 Lippman described a color photography system that recorded color by the interference of light in a single layer, very fine grain emulsion. These images were photographed using a mirror of mercury on the back of the emulsion. In 1893 Joly produced the first additive color screens (Fig. 2). Here a panchromatic, positive emulsion was coated on a lattice of individual red, green, and blue filters. White was represented by no density behind all the color filters; Red by no density behind the red filter and full density behind the green and blue filters.

In the period from 1903 to 1907, the Lumiere Autochrome became available. Here the color mechanism used a random array of stained starch grains containing light-sensitive silver salts. In 1925 Zworkin described the additive, shadow-mask color television tube. In 1928 the first Kodacolor movie film was introduced using red, green, and blue filter stripes in the lens with a black and white lenticular film. Each lenticule imaged a triplet of different color separation on the emulsion behind it. The projector used a matching lens to project a full-color image. In 1931 Dufay introduced a 300 triplet per inch additive screen with parallel green and blue stripes and perpendicular red stripes.[18] In 1971 Land introduced Polavision instant movies with 1500 triplets per inch. This manufacturing process embossed the film base with lenticules, then used them to form the parallel red, green, and blue filter stripes. The next step removed the lenticules and then coated a positive AgX layer over the colored filters. Figure 2 compares the different structures used in additive color from 1883 to today.

Subtractive Color Systems. In 1869 du Huron and C. Clos made the first color prints. They used three-color separation photographs to record the color information. They used three different transfers of yellow, magenta, and cyan dyes to paper. In 1871 the process was improved greatly by Vogel's invention of spectral sensitizers for silver halides. Until then all AgX films used only the inherent sensitivity, which lacked response to red light. Emulsions with balanced sensitivity over the entire visual range improved color performance.

For the next 40 years color photography used a wide variety of single-shot three-color cameras. These unwieldy devices had one lens, a three-way beam-splitter, three color filters, and holders for three separate films. In Friedman's *History of Color Photography*, he devotes three chapters to five types of color cameras using 28 ray-trace diagrams and citing 126 patents. In 1935 three-color cameras became obsolete with Mannes and Godowski's single-sheet color film based on layered emulsions with adjacent dye couplers. It was first sold as a movie film, and then in 1941 it was made into Kodacolor print film.[19]

Land and Rogers invented instant color photography based on dye developers. It was first sold in 1963. Versions of color electrophotography were in the laboratory as early as 1955, but did not make a substantial market impact until the 1980s. Dye sublimation thermal printers and color liquid crystal displays were sold in the 1980s. In 1987 Endo and Vaught invented the thermal ink-jet processes that created the desktop printer revolution.[‡]

Spatial Tradition from Psychology. There is a second tradition in color science that goes back to the Renaissance. DaVinci observed that the color appearance of a pigment mixture changed with the choice of pigments placed around it. The German poet Goethe riled against the physical formalisms of Newton's ideas of human color vision, citing after-image experiments. Count Rumford reported color shadows.[20] Chevreul, France's most famous chemist, was placed in charge of the Gobelin tapestry factory to make new fabric dyes that were as black as the German tapestry dyes. His research led to his famous book on color contrast, in which he explained that better blacks are achieved by white surrounds, not by improving the chemistry of the dyes.[21] Hering introduced the idea of opponent colors.[22] Gestalt psychologists studied a whole family of image constancies, including color constancy. Their experiments showed that the color of objects in real, complex images stayed the same, despite significant changes in illumination. Land and McCann's color Mondrian experiments[23] show that nearly any color can be generated by the same stimulus on the retina. They measured with

† The data were recorded on 6 March 1855 in daylight without sun.

‡ See article by Le in this issue of the *Journal of Imaging Science and Technology*.

a telephotometer the light coming to the eye from a gray paper in a complex image. They moved the telephotometer to a red paper and adjusted the long-, middle-, and short-wave illuminants until it sent the same stimulus as the gray paper. Despite this large change in illumination, the red paper looks red and the gray paper looks gray.

A real life example of an interesting phenomenon caused by illumination is the image shown in Fig. 3. It is an image of a boy holding a white card under a tree. In the foreground, in the sun is a Munsell color chart. Light meter readings of the illumination in the scene showed that the shade was 5 stops, or a factor of 32, darker than similar readings in the sun. What makes this value interesting is that the ratio of reflectances (white to black) on the Munsell chart is also 32 to 1. If the reflected light from the white in the Colorchecker is 100, then the light from the black in the Colorchecker is 3. Because the shade is 32 times darker, the white card in the boy's hand is 3. There is a second white to black scale in the shade ending with the black visor on the boy's cap, light reading 0.1.

Measurements showed that the white in the shade is sending to the camera the same radiance as the black in the sun. The camera will record the same quanta catch from white and black. The rods and cones in the retina will also record the same quanta catch on the retina from the white area and the black area. The photographic negative has a very wide dynamic range of light sensitivity. It can meaningfully record both the 32-to-1 range of radiances in the sun and the 32-to-1 range in the shade—a combined range of 1000 to 1.

Both the photographic print and the neurons in the optic nerve impose limitations on the dynamic range of an image signal. The high end of the print's reflectance range is the white paper [90-95%]. The low end of the print's reflectance range is limited by the surface properties of the print; typically the surface reflects around 3%. The photographic print and the printed page cannot recreate the 1000 to 1 range of the scene. Photographic prints use a nonlinear response function that compresses the whites, expands the grays, and compresses the blacks. This response curve has been optimized to be the best compromise for many different scenes.

Analogous to the limited range of prints, the signals sent to the brain along the optic nerve are encoded in frequency modulation, namely, spikes per second. The ratio of maximum spikes per second to spikes at resting level is much closer to 32 to 1 than 1000 to 1. In summary, the range of intensities is much greater than the range of visual response in neurons and the range of reflectances in prints. The eye, and the ideal image processor that mimics the eye, must transform light intensities into light sensations that preserve visual detail, but use only a small range of data.

Although well known, color contrast ideas are difficult to apply to industrial color imaging systems. The reason is that all imaging systems begin with independent sampling devices such as rods and cones in the eye, silver halide grains in film, and pixels in electronic sensors. All these light receptors count quanta at a point in the image. At this stage there is a unique response of the sensors for a particular quanta catch. Because each receptor is independent of all other receptors, any pair of receptors sensing the same quanta would give the identical responses. Silver halide grain and CCD pixels respond to the number of photons at a point, regardless of the number of photons at other points. It follows that regardless of where a particular stimulus (photon count) falls in the image, the sets of responses remain identical.

Figure 3 shows three different pixel-based image strategies. Each optimizes one aspect of the scene: the sunny area, the shade, and the total range. Each distorts the image an observer would see if viewing the original scene. Each of these pixel-based strategies creates a bad picture.

Human vision is different. After the receptor level, it compares the information at one pixel with all the others in the field of view. This comparison is not absolute. It depends on the distance between areas, the size of the areas, and how much one area wraps around another. The same input, sent to different parts of the same image, does not generate identical responses. For example, the Mondrian experiments and the Boy at Yosemite showed a white and a black appearance from identical quanta catches in the same image. Human vision is unique. A particular quanta catch can generate any color sensation, depending on what else is in the field of view and where it is in that field. (Fig. 4)

Computer programs, and electronic devices have been made that calculate sensations. Early computer programs modeled human color matching data in Mondrians[24] and real life scenes. Later work processed real images to write *onto film* (calculated color sensations).[23]

The interesting paradox of the past is the conflict between physics and psychology. The physics of colorimetry has produced a robust quantitative model of quanta catch in the retina. The psychology of color sensation has pointed out that human biological processes use spatial comparisons. Pixel-based colorimetry cannot predict color sensations because human vision is a spatially dependent imaging system. Vision does not depend on the responses at a single pixel. It responds to the relative responses all across the field of view. This distinction between pixel- and field-based perception is important to the future of imaging.

The Present: Development of Color Management Systems

In the 135 years since Maxwell's lecture to the Royal Institution, the world of color imaging systems has exploded. In the early days of color photography both additive and subtractive color systems were developed. Color printing became the international standard for imaging quality. Subtractive photographic prints became universally available at low cost. In the past 25 years we have added video, still, and scanning image-capture devices. We have also added electrophotographic, ink-jet, thermal, and holographic hardcopy systems, as well as, color cathode ray tubes, liquid-crystal displays and other light-emission color devices.

Ten years ago color imaging systems were supplied as a complete system calibrated by the system provider. The personal computer revolution has changed all that. With personal computers the hardware manufacturers began to make components that interface to other manufacturers' components. A common interface is required to pass meaningful information from one imaging component to another.

The major effort today is to get control of all these technologies so that the user can, without effort, move his color digital image from one technology to another without changing the appearance of the image. A number of companies have collaborated to form the International Color Consortium. The founding members include: Adobe Systems Inc., Agfa-Gevaert N.V., Apple Computer, Inc. Eastman Kodak Company, FOGRA (Honorary), Microsoft Corporation, Silicon Graphics, Inc., Sun Microsystems, Inc., and Taligent, Inc. Other companies that commit to support the consortium specifications will be invited to join as soon as the consortium charter is finalized. The details of the specification are found on the World Wide Web at www.inforamp.net/~poynton/ICC_3.0a/icc_0.html.

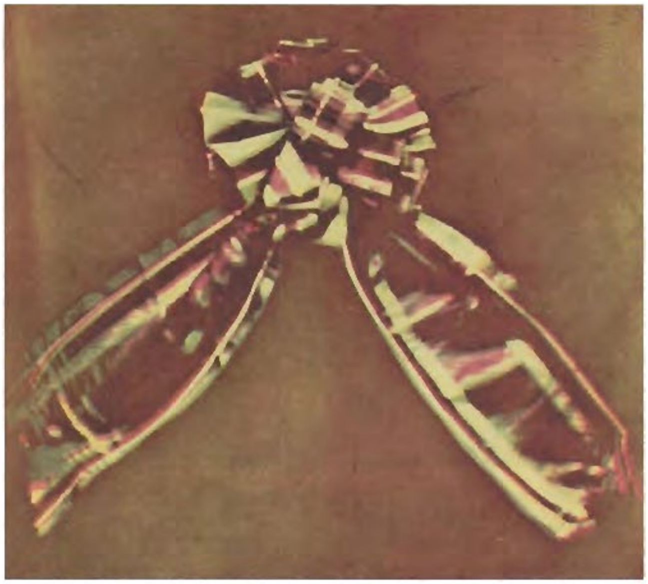

Figure 1. The first color photograph was created by projecting in superposition three black and white photographs, each taken and projected with a different color filter. [1]

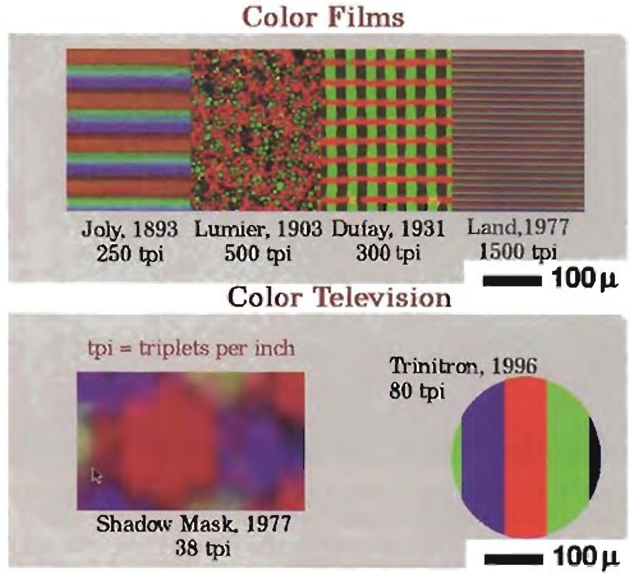

Figure 2. Color films are shown on the top in the range of 250 to 1500 triplets per inch. Color TV is shown below in the range of 38 to 80 triplets per inch.

Figure 3. These three images show three different pixel-based image processing alternatives. On the left, the reproduction of the Color Checker chart in the sun is optimized. Details in the shade are lost. In the middle the shade is correctly printed, but the entire color chart is reproduced as white. The image on the right lowers the contrast of the entire image. Here we have increased discrimination but the color contrast everywhere is reduced. It looks like a foggy day.

Figure 4.This is a computer-processed image of the scene as Yosemite. The range of the input is 1000 to 1. The range of the output is 32-to-1. The computer algorithm maintained the ratios across edges and created a new 32 - to - 1 image that appears nearly the same as the original 1000 to 1 image. The process and the computer program are detailed in Frankle and McCann.[2,3] Today's technology makes it possible to calculate the color sensation for each pixel in the field of view.

Figure 5. Copy A was made so that each area in Copy A was lighter ($\Delta L=+10$) and greener ($\Delta a=-10$) 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 $\Delta E=18$ for each area. In Copy B, the color shifts were in different directions for each area. In this case the ΔE's were chosen to change the appearance of the display. The outer corner patches are closer in color to each other. The mid-side patches are 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. Judged by the ΔE Color Metric, Copy B is exactly as good a reproduction as Copy A.

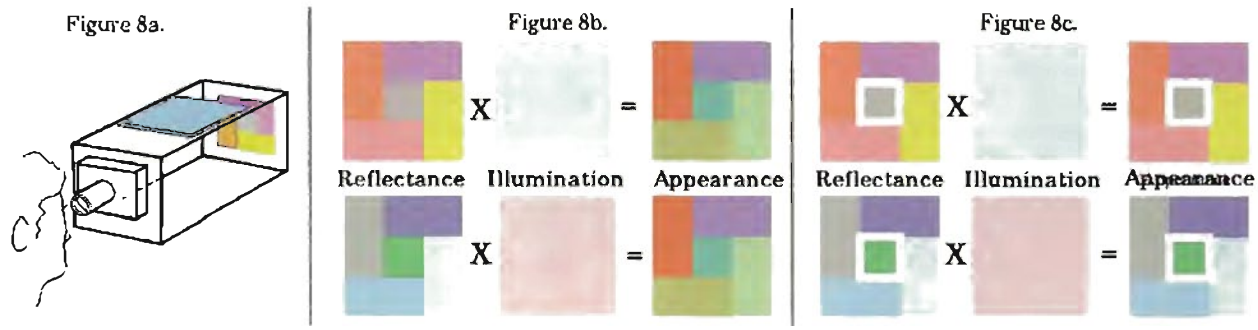

Figure 8a. Maximov's shoebox experiment. The observer sees a small Mondrian through a black tube that restricts the field of view. The filters placed over the hole on the top of the box control the illumination.
Figure 8b. The reflectance of papers is illustrated on the left, the illumination filters place over the hole in the top of the box is in the center, and the appearance of the combination is illustrated on the right. The papers in the Mondrians were chosen to complement the two illuminants. When combined they sent the same stimuli to the eye everywhere in the image. They look alike, in violation of color constancy.
Figure 8c. New maxima, such as this white, changes the normalization values for these two images, thus restoring color constancy, the color areas no longer look identical. Also, observers report that the two white squares are not identical. One is pink: the other light cyan.

A color management module (CMM) is the central element of the information transfer. Each input, display, and output device has an individual spectral response quantized to a digital format called the native device color space. To communicate between all the different native color spaces, the consortium approach is to convert all spaces to a single, device-independent space. Device profile is the name of the collection of information used to covert between native device and device-independent spaces.

The profile connection space (PSC)[§] is based on relative colorimetry; that is, the long-wave tristimulus values from the sample are multiplied by the ratio of media white to illuminant. The specification provides a choice of 8 bit/component CIELAB, 16 bit/component CIELAB, and 16 bit/component CIEXYZ encoding to accommodate conflicting accuracy and storage space requirements. To allow individual processing, the specification has three levels of tags: required, optional, and private data. Tags are used to contain information about viewing environments, D-min of materials, white point, black point, etc.

This specification uses relative colorimetry; that is, the tristimulus values of the area of interest, scaled by the ratio of the material white to the illuminant. Most important, it scales the long-wave information (X) independent of the middle-wave (Y) and short-wave (Z) information. This is a major advance over absolute colorimetry that just counts the absolute quanta catch. It works well because it normalizes to the maxima in the image. This step alone gives the calculation the properties of color constancy.

Where the ICC color profiles approach falls short is that it compares all the pixels in the image to one triplet of values. The new CIE color appearance model (CIECAM) is the first CIE standard to include surround parameters in its calculations.[25] Both ICC and CIECAM calculations

§ Based on ANSI CGATS.5 1993 STANDARD "Graphic technology—Spectral measurement and colorimetric computation for graphic arts images" using D50, 1931 CIE standard observer, 0/45 or 45/0 reflectance measurement geometry and ANSI PH2.30-1989 viewing conditions. See www.inforamp.net/~poynton/ICC3.0a/icc0.html, section 2.6.

do not, however, pass the test of acting like the human visual system:

1. The data used to normalize the entire image is derived independent of the image. ICC and CIECAM¶ use information measured from a white media sample that may or may not be present in the image. Further, they use measured information about the illumination not derived from the scene. Humans do not have independent media and illumination measurements. Humans calculate color from the scene itself.
2. If two pixels have the same input quanta catch; they will have the same output value. They cannot generate multiple outputs, such as white and black, from the same input values in different parts of the image. ICC and CIECAM do not act like the human eye. They can calculate what we see successfully only in special cases. Consider the Boy at Yosemite. The successful reproduction maps two areas with the same inputs to two very different outputs—white and black.

These two properties are necessary conditions for models of human vision.

A very important axiom with regard to successful image reproduction is if one uses colorimetry and successfully matches every pixel in the original to every pixel in the reproduction, the result is a perfect copy. Regardless of how the human eye processes information, if the reproduction quanta catch of every pixel is identical to the original quanta catch, the two images have to match. If every pixel in the field view is identical, the reproduction is perfect.[26] The problem becomes more interesting when the color gamuts of the original and the reproduction do not match. When significant fractions of the gamut of possible color matches are beyond the range of the reproduction media, for example prints and displays, then matching all pixels is impossible. The underlying axiom becomes invalid. The problem shifts from color match to minimizing the apparent color error. The eye's image processing principles now must be considered. The best color reproduction will be one that maintains the relationships within the image.

The Future: Color Field Calculations

The ideal solution to the color gamut problem is to write a computer program that minimizes 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 metric that quantifies the color mismatch. Once the color metric is shown to be accurate, then a computer can reliably find the best compromise.

Color Metric. Color differences are almost always described by ΔE in L*a*b* color space. This space, defined by the 1976 CIE report,[27] is calculated using the tristimulus values X, Y, Z, defined in the CIE 1931 report. Further, a complex image is often evaluated by averaging the individual ΔE's 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 what we see, namely, the sensation image after spatial interactions in the human visual system. Physical models of the image at the retina do not predict appearance in the brain's visual cortex. The question we are asking in this experiment is whether the average ΔE metric, which we all use for color differences at the first stage in the retina, is appropriate for predicting color appearances.

The following series of experiments creates triplets of targets: original, copy A, and copy B. Although researchers are usually concerned with matches across media, such as display to print, these experiments were restricted to single media to eliminate properties 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. The colors in copy A were selected to be significantly different from those in the original. Each area in copy A is 10 units lighter, 10 units greener, and 10 units yellower in Lab space. The combined distance in L*a*b* space is $\Delta E = 18$. For each area the color difference was $\Delta E = 18$ between the original and copy A. Copy B was made with each area $\Delta E = 18$ compared to the original, but it was designed to have the color shifts go in many different directions. Copy B changed the local relationships, while copy A preserved them. If ΔE is an isotropic color metric for color appearance, half the observers will select copy A as the best reproduction and half the observers will pick copy B. If spatial parameters, namely, the relationship between different areas within the test target is important, then observers will select copy A, which preserves the spatial relationship. Observers were shown a trio of similar displays in the same field of view (Fig. 5).

Observers selected copy A as the better reproduction, the image in which all areas had the same color shift. 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. The colors we see are 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. The value ΔE correlates with the quanta catch of retinal receptors, but cannot be used to evaluate color appearances later in the visual system. Minimizing ΔE searches for best matches independent of spatial information. However, humans use spatial information to calculate color.

Color appearance needs a spatial 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 ink-jet cavity, and electrons delivered to the TFT cell control the response of that pixel. The signals 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 quantum catch 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 accumulate relationships all across the image.

¶ CIECAM is a new standard adopted by the CIE. Color appearance models are based on the radiance of a single pixel with four imagewise constants to represent the rest of the image. This set of 64 equations uses input from one pixel. The chromaticites of the white in the scene, the background, the white reference and four surround parameters and the luminance of test adapting field are assumed as given.

** Most people use the 1978 Color difference formula even though there is a new CIE 1994 formula. This new formula reduces the color differences between color of high chromaticities. Using this formula a ΔE (1978) = 18 becomes ΔE(1994) = 10. Although a new target is necessary to make Copy B exactly equivalent to Copy A in the 1994 standard, the relative changes to each area are small.

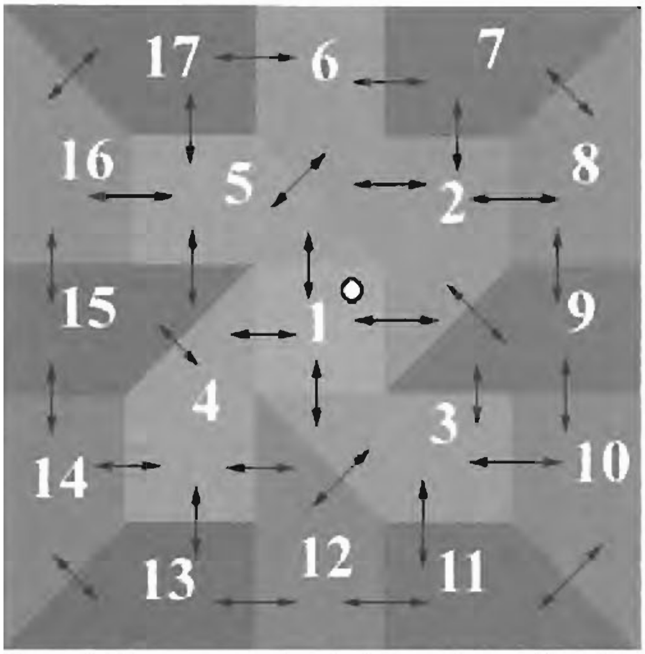

Figure 6. Diagram comparing the information used in calculating color appearance. Colorimetry uses information from a single pixel in the scene (black open circle). A ratio metric uses all 32 comparisons between each of 17 different areas (double headed arrows).

After different technologies, such as printers and displays introduce different color gamuts, the problem becomes more interesting. In fact most printers have larger gamuts in low lightnesses at full saturation; displays have much larger gamuts in high lightnesses at full saturation. Usually, one finds that the common volume of a three-dimensional color space is half of 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.

An approach to solving the problem is to use the information learned from the comparison of copy A with copy B, above. 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 6 is the original image in Fig. 5 with each area given a number for identification. Area 1 is the gray square in the center. Numbers increase in a clockwise spiral 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 measured independently. The double-headed black arrows 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 $(X_{Copy\,A}/X'_{Copy\,A})$. We can calculate this result with the corresponding data for the original image $(X_{Original}/X'_{Original})$. We can compare the two with a ratio, $[(X_{Copy\,A}/X'_{Copy\,A})/(X_{Original}/X'_{Original})]$. The results X, Y, and Z for the 32 edges are plotted in Figs. 7(a) and 7(b).

Copy A is a good reproduction because it has the same edge ratios as the original. Copy B is a bad reproduction because the edge ratios are different from those in the original.

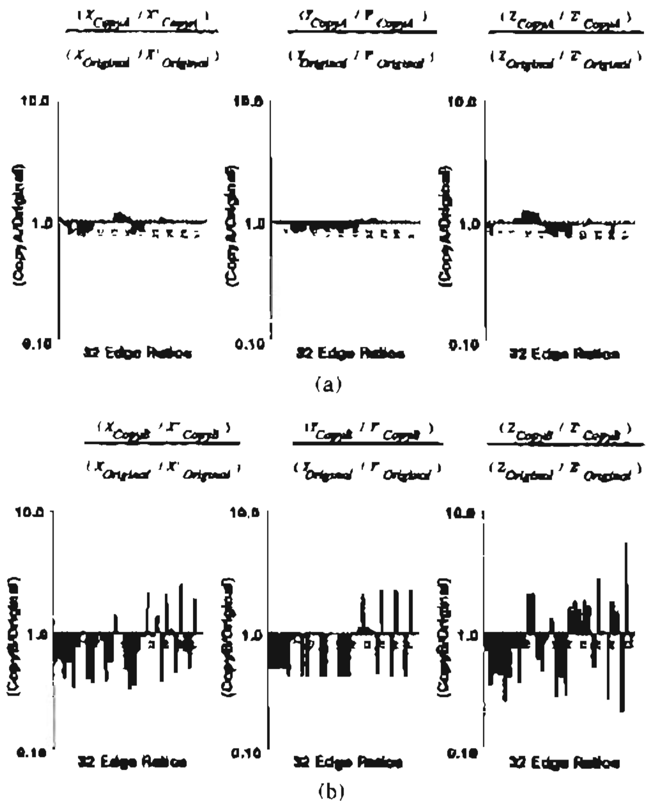

Figure 7. (a) The three graphs demonstrate a color metric for color appearance. Results for all the comparisons of X/X' are shown in the left graph, Y/Y' in the middle graph, and Z/Z' in the right graph. The edge ratios from the comparison of copy A/original are very close to 1.0, indicating that the two displays are relatively the same. This similarity shows a good reproduction; (b) Results for all the comparisons of X/X' are shown in the left graph, Y/Y' in the middle graph, and Z/Z' in the right graph. The edge ratios from the comparison of copy B/original are very far from 1.0, indicating that the two displays are relatively very different. Copy B is a bad reproduction because its edge ratios differ from those in the original.

Calculating Color Sensations

As described above, there are two key properties of human color image processing. The first is that the eye calculates sensations, or appearance, from the scene itself. It is not given a priori information about the illumination and "whites." The second is that a single input is transformed to many different outputs by the spatial information in the rest of the image.

The common denominator of all color models that address color constancy is that the quanta catch of the cones is normalized by a value relevant to the scene. There the similarity ends. The CIECAM normalizes each pixel to an illuminant value measured directly from the scene. Other, gray world approaches normalize with an average of all the pixels in the scene.

Experiments with small Mondrians show that human color vision normalizes to the maxima in each waveband,[28] corresponding to long-, middle-, and short-wave cones(L,M,S). The experiment used two small, five area Mondrians viewed in a Maximov shoebox, with different filters to control the color intensity of the illuminant. (See Fig. 8a).

One Mondrian was viewed with the left eye, the other with the right eye. The reflectances of each area were carefully generated on a Canon CLC500 printer so that each Mondrian shared all the same ratios of radiance across edges. The pink Mondrian reflected two times more long-wave or "red" light everywhere than the cyan

Mondrian. The cyan Mondrian reflected two times more "green-blue" light everywhere than the pink Mondrian.

When the cyan Mondrian is viewed in daylight with a pale red filter that doubles the "red" light, and when the pink Mondrian is viewed with a pale cyan filter that doubles the "green–blue" light, the two Mondrians appear the same [Fig. 8b]. They must! They generate exactly the same quanta catch everywhere on the retina. The reflectances were generated to have exactly the same ratios at edges. They differed only in the amount of long-wave and short-wave light they reflected. The two different illuminants were chosen to compensate for the overall shift in reflectances.

The fact that two identical stimuli look the same is not very interesting. Nevertheless, the fact that this combination of reflectances and illuminants destroys color constancy is very interesting. Our common experience is that moderate changes in illumination such as a Wratten 50CC red and 50CC cyan filter will not change the color appearance significantly for almost any scene. Yet in combination with these carefully composed Mondrians they destroyed "color constancy." With such mild shifts in illumination we ordinarily would expect color constancy to work.

This combination of illuminant and reflectances is the ideal test for identifying the mechanism of the human normalization process. We can add a wide variety of different papers to the matching Mondrians. If a white paper added to both destroys the match, that indicates that the human normalization process normalizes to white. If a black paper added to both destroys the match, that indicates that the human normalization process normalizes to black. If a new reflectance, higher than the maxima in the Mondrians in any of the long-, middle-, and short-wave bands all destroy the match, then the human mechanism normalizes to the maxima in all three wavebands independently [Fig. 8c].[29] These experiments show that this is the case. A red, green, blue, yellow, magenta, or white pair of papers will destroy the match. A black or dark-hued pair of papers will not. Grays will not destroy the match, but black and white that average to the same gray value will. Humans normalize to the maxima for each cone type.

Experiments measuring color appearance demonstrate the human normalization process is different from the mathematical normalization used in image processing. The mathematical normalization scales each pixel to the highest value found in the scene. All values with a particular value g will be divided by the same maxima m to generate a fixed output of g/m for all pixels with input g.

Here again, human vision does not work that way. Humans have a normalization system that differs from an image processing normalization by two important properties:

1. overall intensity
2. spatial data in the image.

McCann, McKee, and Taylor[30] measured the shift in appearance of whites with overall intensity. It is close to the image processing normalization, but there is a measurable dependence on the intensity of the illuminant. In the shoebox experiments the two new whites that destroy the match and restore color constancy do not look the same.

Measurements of the appearance of grays in the vicinity of white have shown that apparent lightness is influenced by the spatial relation of the gray and white. Figure 9 shows six different lightnesses from the same amount of white with different spatial placements. The gray patch with nothing except the white in the field of view will appear darkest [L = 1.5] when the white surrounds the gray. It appears lightest [L = 3.9] when the same total area of white is placed

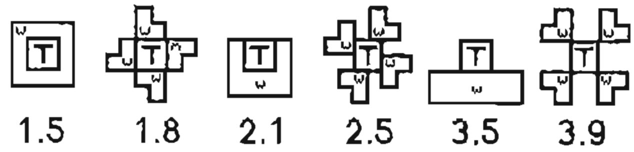

Figure 9. A diagram of six targets; the numbers refer to lightness matches to a standard display (white L = 9.0, black = 1.0). Here the experiments measure the changes in lightness as a constant area of White is spatially repositioned. The test field (T) is darkest when surrounded by white (W) on all four sides (L = 1.5). It is lightest when white is placed on the four corners (L = 3.9). The four intermediate targets that enclose T less (left to right) have higher lightnesses (L= 1.8, 2.1, 2.5, and 3.5).

all on the corners of the gray test area T. Five other intermediate placements of the white give five intermediate gray values. How the white surrounds the gray is important in the normalizing process in human vision.[31]

A similar experiment measures the influence of white placed on one, two, three, and four sides of a gray patch. Further, it measured the influence of maxima both in terms of separation between the white and the gray, and enclosure by maxima. The data are shown in the 3-D graph in Fig. 10.

The values found in these measurements are not as important as the strong message that human normalization is a spatial process. A mathematical image processing normalization would generate a flat plane parallel to the horizontal axes. All that matters is the value of the white. Humans normalize spatially so that the amount of enclosure and the distance from the white have a significant influence on the appearance. Data from the entire field of view are necessary to predict human normalization responses.

Summary

Although the entire history of color imaging systems is 135 years old, it is growing today at a phenomenal rate. Color imaging systems are everywhere and they are doing everything.

We are on the threshold of standards between many different companies worldwide. These companies are joining color groups that will provide systems for automatic translation of one manufacturer's signal to all other manufacturers' hardware.

However, research in the field of color is far from being finished and put on the shelf. The color experiments shown in Fig. 5 were designed to have constant ΔE and different edge ratios. When the edge ratios were nearly constant the copy A made a fairly good reproduction despite that ΔE = 18. When the edge ratios were distorted in copy B, the appearance of the display changed. Figure 7 shows that when ratios in the reproduction match the ratios in the original we have a good copy. Reproduction ratios that depart from those of the original characterize a bad copy. Local ratios provide an important tool in finding a meaningful color metric. The comparison of copy A and the 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 a successful metric, but not a complete metric by themselves. Human vision normalization depends on both spatial and absolute quanta catch information. The best color metric that will empower the color information exchange process is yet to be described.

The use of the human normalization processes allows us to process images to produce written images having the same sensations as looking at the scene. As shown in Figs.

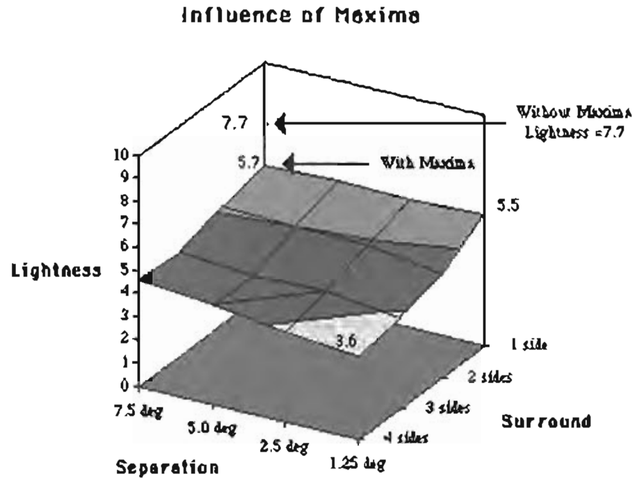

Figure 10. The vertical axis is matching lightness chosen by observers for the same gray patch. The two axes in the horizontal plane plot are "Separation" between the gray and the white, and the "Sides of White" surrounding the gray. The furthest point in the background shows the match for gray surrounded by white on only one side at 7.5°, and the furthest distance (L = 5.7). Above that point is the control match (L = 7.7) for the same patch without any white in the field of view. The closest point has the darkest appearance for the set of measurements. The gray area with four sides of white at the closest distance was matched to a lightness of 3.6. All other combinations showed an intermediate value of matching lightness.

3 and 4, successful reproductions have to destroy the unique relationship between input and output pixel values. In human vision, any particular input value will generate a range of output values that will be scene dependent. In a scene with perfectly uniform illumination and using exclusively one spectral composition of light source with white surrounding all areas in the scene, it is possible to conceive that for each input pixel value to the eye there would be a singular output value. In all other cases, single input values will generate a range of different output values. In scenes such as Yosemite, the same input value at the retina will generate a range of outputs from white to black.

The technology of image making has progressed at an astonishing pace. Our ability to control dyes and pigments began with factory film coaters, progressed to high-quality ink on paper, then to very high-quality electrophotographic machines and now to small inexpensive thermal ink-jet printers. The price for this hardware has gone from tens of millions of dollars to hundreds of dollars in 60 years. The image making machinery has moved from the factory to the desktop.

Maxwell's invention of color photography was not for commercial purposes. He just wanted to demonstrate Thomas Young's ideas about human color vision. We have come full circle. We have the ability to print and display images that Maxwell never imagined. Today, our limitations return to our understanding of color theory. If we could calculate color sensations better, then we could make better displays and prints. To do this we need to incorporate in our equations the lessons of psychology, as well as those of physics, by using field calculations. Edwin Land often quipped, "What if Maxwell had done his color experiments (1853 to 1861) after he did his electromagnetic field experiments (1864 to 1877). Maybe the whole history of color vision would have been different."

Acknowledgments. The author wishes to thank Norbert Herzer for his work on creating and evaluating the targets used in the color metric experiments. The author wishes to thank Jim Burkhardt, Karen Houston, and Brian Hallstrom for their work on the Yosemite experiments.

References

1. Arts of Stone Age European Peoples, in *Encyclopedia Britannica*, 15th Ed., 17, p. 704 (1974).
2. G. Vasari, Le vita de' piu eccellenti architetti, pittori e scultori italiani, G. Milanesi, Ed., Milan, 1878-85, reprinted 1906.
3. F. Mancinelli, J. McCann, and V. Ruzdic, *A Masterpiece Close-up: The Transfiguration by Raphael*, Vatican Museums and Galleries, Vatican City, 1980.
4. F. Mancinelli, *Michelangelo at Work: The Painting of the Ceiling in The Sistene Chapel: The Art, the History and the Restoration*, M. Giacometti, Ed., Harmony Books, New York, 1986.
5. H. Gernsheim, *A Concise History of Photography*, 3rd ed., Dover, New York, 1986, pp. 8–18, 1986.
6. B. Newhall, *The History of Photography*, Little, Brown and Company, Boston, 1982, p. 270.
7. J. C. Maxwell, On color vision, in *The Scientific Papers of James Clerk Maxwell*, vol. 2, W. D. Niven, Ed., Dover, New York 1965, pp. 267–279.
8. R. M. Evans, *Sci. Amer.* **205**, 118 (Nov. 1961).
9. J. C. Maxwell, in *Scientific Papers of James Clerk Maxwell*, W. D. Niven, Ed., Dover Publications, Inc., New York, 1965, vol. 1, p. 128.
10. W. D. Wright, *The Measurement of Colour*, 3rd ed., Hilger & Watts, London, 1964.
11. *CIE Proceedings 1931*, Cambridge University Press, Cambridge, 1932, p. 19.
12. V. C. Smith and J. Pokorney, Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm, *Vision Res.* **15**, 161 (1975).
13. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed, John Wiley & Sons, New York, 1982.
14. G. Wyszecki, Colorimetry, in *Color Theory and Imaging Systems*, R. Eynard, Ed., (Society of Photographic Scientists and Engineers, Washington, D.C., 1973, p. 24.
15. F. E. Ives, U.S. Patent 432,530 (1890).
16. J. S. Friedman, *History of Color Photography*, The American Photographic Publishing Company, Boston, 1945, p. 13.
17. C. E. K. Mees, *Brit. J. Photog.*, **55**, 41 (1908); C. E. K. Mees and Pledge, *Phot. J.*, **50**, 197 (1910).
18. J. S. Friedman, *History of Color Photography*, The American Photographic Publishing Company, Boston, 1945.
19. J. S. Friedman, pp. 34-63.
20. J. McCann, Human color perception in *Color Theory and Imaging Systems*, edited by R. Eynard, Society of Photographic Scientists and Engineers, Washington, D.C. (1973) p. 1.
21. M. E. Chevreul, *The Principles of Harmony and Contrast of Colors and Their Application to the Arts*, 1854, Republished, Reinhold, New York, 1967.
22. E. Hering, *Outline of A Theory of the Light Sense*, trans. by L. M. Hurvich and D. H. Jameson, Harvard University Press, Cambridge, MA, 1964.
23. J. Frankle and J.J. McCann, Method and apparatus of lightness imaging, U.S. Patent 4,384,336. J. J. McCann and K. L. Houston, Calculating color sensations from arrays of physical stimuli, *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC 13**, pp. 1000-1007 (1983); J. J. McCann and K. L. Houston, Color Sensation, Color Perception and Mathematical Models of Color Vision, in *Colour Vision*, J. D. Mollon, and L. T. Sharpe, Eds., Academic Press, London, 1983, pp. 535-544.
24. E. H. Land and J. J. McCann, Lightness and retinex theory, *J. Opt. Soc. Amer.* **61**, 1 (1971).
25. R. W. G. Hunt, *AIC Proceedings*, Kyoto, in press, 1997.
26. J. J. McCann, Digital color transforms applied to fine art reproduction, in *Proc. of 2nd International Conference on Imaging Science and Hardcopy, Reprographic Scientists and Engineers Society*, China Instrumentation Society, Guilin, 1995, p. 221
27. 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.
28. J. J. McCann, Rules for colour constancy, *Ophth. Physiol Opt.* **12**, 175 (1992).
29. J. J. McCann, Color constancy: small overall and large local changes, in *Human Vision, Visual Processing, and Digital Display III*, B. Rogowitz, *Ed., Proc SPIE*, **1666**, pp. 310–321 (1992).
30. J. J. McCann, S. McKee, and T. Taylor, Quantitative studies in retinex theory, a comparison between theoretical predictions and observer responses to Color Mondrian experiments, *Vision Res.* **16**, 445 (1976).
31. J. J. McCann and R. L. Savoy, Measurement of lightness, dependence on the position of a white in the field of view, *Proc. SPIE* **1453**, 402–411 (1991).