Color Spaces for Image Representation and Image Processing

Lucien Xu and Richard Holub

AGFA Division, Miles, Inc.
Wilmington, Massachusetts 01887

Properties which make a color space well-suited for the representation of color images are a subset of those which make it powerful for the editing and processing of images. We will elaborate this point while examining a list of eight desirable attributes of a color space for digital imaging applications. Several color spaces will be compared with reference to many of the attributes. The presentation will emphasize efficiency. Compressibility will be considered with respect to the efficiency of utilization of bits and to perceptual uniformity. Data on the efficiency of utilization of a regular quantization scheme for storing images will be discussed in relation to the gamuts of common Graphic Arts media.

0.0 INTRODUCTION

Product announcements and introductions of the last year or two have created *de facto* standards in the area of color representation for electronic publishing. Nevertheless, it is yet of more than academic interest to consider the relative strengths and weaknesses of some common spaces with respect to image representation and image processing.

Publishing applications generally involve cross-rendering in which the medium of the reproduction is different from that of the original. Since the reproduction is a four-color reflection image, the problems of handling unprintable colors and a black printer are especially daunting. It is for this reason that the automation of cross-rendering has been slow. And it is partly for this reason that image retouching applications which can compensate for deficiencies of automation are necessary. Of course, they are also necessary for the introduction of elective edits into images and are increasingly useful for the composition and layout of complex graphics. Therefore, it will be a thesis of this paper that attributes which make a color space advantageous for image processing rather than for mere image representation and transmission should receive added weight.

Following is a reasonably complete list of desirable properties of a standard color space. Since they are not all mutually consistent with one another, trade-offs and weightings are necessary in deciding which color space is better for a stated purpose. The weighting which is implied by our ordering of the list is controversial. In keeping with the comments of the preceding paragraph, we have placed atop the list two properties which we think are important when the application involves cross rendering and lots of image processing.

1. It should have approximately independent and perceptually uniform numerical scales for the dimensions of hue, saturation and lightness. Especially, it should separate the gray scale information of an image from the chromatic.
2. It should incorporate a model of white point adaptation so that the whites of different media are independent of the neutral point in the image representation. The translation of one white to another can be accounted for within calibrated transformations into and out of the standard space.
3. It should be device independent, with coordinates which are measurable and related to the Standard Observer. This is really implied by one and two.
4. It should be possible to transform into and out of the space by easy computations on inexpensive hardware. Counterbalancing this requirement, however, is a need to be able to convert into and out of device coordinate systems accurately.
5. It should utilize bits efficiently, consistent with minimizing the width of a data path. It should have competitive image compression ratios.
6. It should make efficient use of a regular quantization scheme. This attribute is not necessarily consistent with number five. Both five and six presuppose a digital image representation.
7. It should have a well-understood and accepted notion of color difference or error.
8. It should facilitate additive color mixture.

We will discuss all eight factors. However, some will receive greater scrutiny than others because, for example, a color space either models white point adaptation or it doesn't, whereas the relative compressibility of images stored in different color spaces is a matter for quantitative study.

The color spaces to be considered in detail are "calibrated RGB," CIEXYZ, Kodak PhotoYCC, CIELUV, CIELAB and CIEYu'v'. It is not possible to compare all spaces with respect to all aspects of the study. We omit the spaces GLHS (Generalized Lightness, Hue and Saturation, cf. Levkowitz and Gabor, 1988,) MOTR (mutually opposed trichromatic response model, cf. Huntsman, 1988,) Munsell (Newhall, Nickerson and Judd, 1943) and TekHVC (Tektronix Hue Value Chroma, cf. Taylor, McManus and Murch, 1988) because they do not (yet) have broad application in the Graphic Arts and for limitations of time and space.

"Calibrated RGB" is generally taken to refer to an "emancipated" device coordinate system which is an apparently simple means of storing quantized images for display on a CRT. The monitor's coordinates readily admit of a device independent interpretation because the device observes (apart from gamma) approximately linear color mixture laws. Given calibration data on the phosphor emissions, it is possible to calculate a 3X3 matrix useable in converting device coordinates to CIEXYZ. The CIE's XYZ coordinates, of course, are the TriStimulus Values of the 1931 (two degree) Standard Observer.

Different "Calibrated RGBs" can be distinguished on the basis of their (phosphor-based) primaries and, potentially, on the gamma function imputed to the device. If a gamma, or companding, function is part of the definition of the color space, then a non-linear encoding of data precedes the quantization for storage. Gamma, of course, refers to the slope of the log Intensity vs. log Voltage function of the device. The motivation for encoding data with the inverse of gamma is to precondition the entries of a frame buffer for direct display on a CRT. In other words, the popularity of these spaces is related to the attribute "easy computations on inexpensive hardware." Another rationale which has been given for "calibrated RGB" is that the space efficiently utilizes the cuboidal RGB addressing scheme; i.e., all 255 levels in R, G and B dimensions can be occupied by valid image data. We will revisit both of these points in connection with attributes four and six.

The CIE spaces are explained in Publication 15.2 (1978) of the Commission. A still more complete treatment of that subject may be found in Bartleson (1980.) Kodak's PhotoYCC is a special case of "Calibrated RGB" specified in detail in connection with Kodak Photo CD (1991.) In addition to offerring an encoding of RGB, it defines a rotation of axes to yield "Luma" and two "chroma" channels in a manner resemblant of television techniques.

The balance of this paper is organized into sections, treating the eight attributes in turn. The section on hue, saturation and lightness presents results of a small, perceptual study of the independence of lightness and chromatic variables and of the relative uniformity of several spaces. Sections two and three are brief elaborations of comments in this introduction. In section four, we discuss the issues surrounding computational complexity and why it is difficult to draw simple conclusions, rather than to attempt to rank color spaces in this regard. We present a metric with which to study the question of the relative simplicity and accuracy of several spaces with respect to color transformations of practical interest in the Graphic Arts.

In section five, we return to the issue of uniformity as it affects efficiency of digital storage. We also study the relative compression ratios attainable with several spaces in the reproduction of images with a criterion quality. In section six, we turn to the gamuts of several media of interest in the Graphic Arts and study the efficiency with which said gamuts can be represented by regular quantizations of the spaces. In other words, we find a bounding cuboid for each gamut in the coordinates of one of the color spaces and compute the percent occupation of the cuboid by colors within the gamut in question. The resulting numbers are compared as a measure of the efficiency of utilization of a regular quantization scheme by various color spaces. Sections seven and eight are also brief elaborations of the points in the introduction followed by summary conclusions. The authors acknowledge that some of our conclusions are opinions, albeit scientifically informed ones, and should not be construed as an official position of AGFA.

1.0 PERCEPTUAL INDEPENDENCE OF COORDINATES AND UNIFORMITY

Of the color spaces we are considering, CIE Yu'v', CIELAB, CIELUV and Kodak PhotoYCC purport to separate gray scale from chromatic variables. CIELAB and CIELUV have well defined and supposedly independent hue, saturation and lightness variables. The latter have been found to be subjectively preferred perceptual primitives for many color editing circumstances, especially those involving untrained operators. Although the PhotoYCC specification (Kodak, 1991) does not define hue, saturation and lightness variables, it may be possible to define them by analogy to the CIE spaces; this certainly can be done for Yu'v'.

We asked nine subjects of various ages and backgrounds to compare color spaces, in pairs, with respect to 1) which more accurately modelled lightness constancy along a trajectory of saturation at constant hue, 2) which produced more uniform steps of saturation change at constant hue and lightness, 3) which better preserved hue in the course of a saturation change at constant lightness, 4) which better preserved hue in the course of a lightness change at constant saturation and 5) which of two notions of saturation was more *perceptually* independent of lightness. As we present results, we will clarify how the pairings were decided. All stimuli were presented on a calibrated monitor (19 inch Sony Trinitron) against a full-field white background. Controls for spatial and temporal variations of the device were built into our protocol. Likewise, our design included means of checking for subjective inconsistencies. Our stimuli consisted of the six hues of nominal SWOP primaries and secondaries (Cyan, Magenta, Yellow, Red, Green and Blue.)

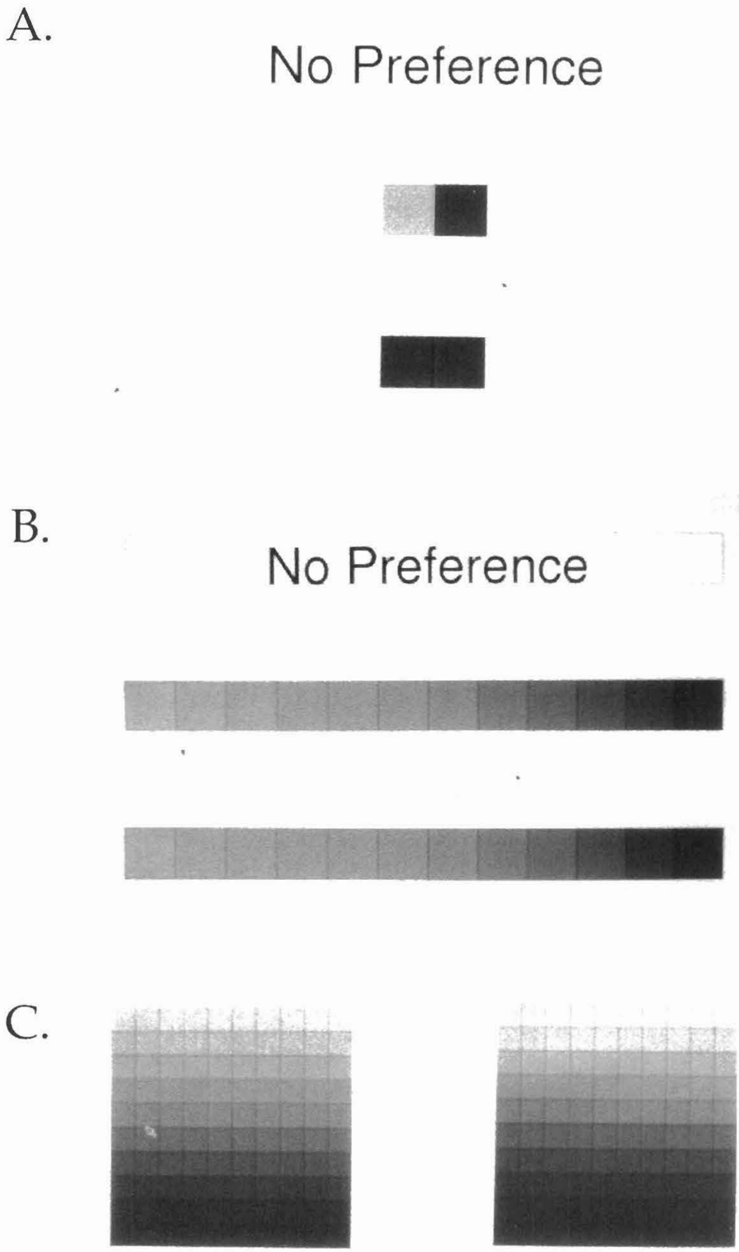

Figure One

CIELAB and CIELUV model aspects of perception, such as the uniformity of small color differences throughout the perceivable gamut and the separation of gray scale from chromatic dimensions of experience. The designers of Kodak's PhotoYCC may not have set the foregoing properties as goals, but it remains legitimate to ask how YCC compares to CIE spaces in realizing them when applied to applications in which they are important.

1.1 Lightness Constancy

Figure 1A is a semblance, in black and white, of the display employed in the first experiment. (Fig. 1 is derived from 35 mm slides of the display, which was converted to a monochrome image and halftoned.) Subjects clicked a mouse to choose one of two pairs of adjacent boxes or a "no preference" panel and their votes were recorded automatically by the interactive program. One of the pairs of boxes was a saturated version of one of the six test hues and the other was a neutral of the same L* in one pair and a neutral of equal Luma in the other. L*, of course, is the CIE psychometric lightness variable. Luma is the nominally independent luminance or lightness variable in Kodak PhotoYCC.

The paired colors are the same for CIELAB, CIELUV and Yu'v', so these spaces were not treated independently in this experiment. Neither CIEXYZ nor "calibrated RGB" were considered in these experiments because gray scale and chromatic information are not separable in them. Each comparison was presented four times (six hues times four presentations for 24 judgments.) Saturated color and neutral were alternated left to right in two presentations and CIE and YCC were alternated top to bottom. Position dependencies and subjective inconsistencies were thus detected and controlled. An inconsistency was logged as a "no preference." In general, throughout the five experiments, we did not detect position dependencies.

TABLE ONE

Constancy of Lightness

cyan	magenta	yellow	red	green	blue
L* 6 Ycc 2	L* 3 Ycc 1	Ycc 5 L* 2	L* 4 Ycc 3	L* 6 Ycc 3	L* 4 Ycc 2

Table 1 summarizes the results. For each of the test hues, the number of votes for L* and Luma are shown. Where the sum of votes was not 9, the remainder were "no preference." For some of the hues, especially cyan and green, the difference in lightness between the saturated hue and the neutral of equal Luma was striking. Post test questioning of subjects who voted YCC in these cases revealed that they were responding to greatest perceived contrast, not equal perceived lightness. The non-constancy of lightness encoded by YCC Luma is not surprising on mathematical grounds, because Luma is a weighted sum of non-linear RGB values rather than a projective transformation of the CIE luminosity variable. The results indicate that PhotoYCC is not a good space for users who want to edit hue or saturation without changing lightness.

1.2 Uniformity of Changes in Saturation

The pronounced (and non-uniform) changes in apparent lightness from one end of a saturation trajectory at constant hue to the other also made it difficult to judge the uniformity of PhotoYCC with respect to saturation. Therefore, we compared only CIELAB and CIELUV in this regard. Figure 1B attempts a gray scale facsimile of the stimulus panel. Two step wedges, each consisting of twelve patches were displayed along with a "no preference" window. Equal steps of CIELUV and CIELAB Chroma, respectively, were displayed from the greatest Chroma of the test hue to neutral in the two wedges. Increasing Chroma is indicated by increasing density in the gray scale of 1B. Controls were as in 1.1. Subjects were asked to judge in which of the two wedges the gradation was more uniform. Table two reveals that it was a draw, except for hue angles near blue. This finding is consistent with the considerable body of evidence cited by Pointer (1981.)

TABLE TWO

Uniformity

cyan	magenta	yellow	red	green	blue
NP 9	NP 8 LUV 1	LUV 4 LAB 3 NP 2	LAB 4 LUV 3 NP 2	NP 8 LAB 1	LUV 8 LAB 1

1.3 Hue Constancy with Saturation Changes

The panels were identical in this experiment to those of 1.2. In this case, the subjects were asked which of CIELAB or CIELUV trajectories better preserved the hue of the saturation extreme. Table three shows results. Subjects voted strongly for CIELUV in the case of reds and blues, but couldn't decide for other hues. They felt that the discriminations were difficult in the tasks of sections 1.2, 1.3 and 1.4. The underlying question in this experiment is whether CIELAB or CIELUV has a better model of hue constancy or better independence of the perceptual dimensions of hue, saturation and lightness. If one were a clear winner, then it would be a better space in which to perform gamut scalings or to perform edits in which saturation alone is to be changed.

TABLE THREE

Constancy of Hue with Chroma

cyan	magenta	yellow	red	green	blue
NP 9	NP 9	NP 4 LAB 3 LUV 2	LUV 7 LAB 1 NP 1	NP 7 LAB 1 LUV 1	LUV 9

1.4 Hue Constancy with Lightness Changes

This test was motivated particularly by Pointer's (1981) reference to potential anomalies of CIELAB's model of hue-constancy. The stimulus panel of this experiment had the same general organization as in the previous two, except that the gradations were in psychometric lightness (L*) at constant CIELAB and CIELUV hue angle and Chroma. Subjects were asked to judge in which space *perceived* hue was more constant through the 12 steps in lightness. At each of the six hue angles, a Chroma was chosen as far from neutral as possible, consistent with a sufficient range of realizable, discriminable lightnesses. Table four shows that the two spaces were largely indistinguishable in this test, at least in the range of luminances presented.

1.5 Saturation vs. Chroma

CIELUV coordinates are related to the u',v' chromaticities of the 1976 Uniform Chromaticity Scale (CIE.) In other words, CIELUV has an associated chromaticity diagram and many of a scientific bent consider this to be an important advantage over CIELAB. The hue

TABLE FOUR

Constancy of Hue with Lightness

cyan	magenta	yellow	red	green	blue
NP 6 LUV 3	LUV 5 NP 4	LUV 7 NP 2	LAB 4 NP 3 LUV 2	LAB 4 NP 3 LUV 2	LAB 4 NP 4 LUV 1

angle calculable from u',v' is the same as that yielded by u*,v*, but the length of the color vector is independent of luminance or lightness in the case of u',v', but proportional to lightness in the case of u*,v*. The length of the chromaticity vector (u',v') is called saturation while the length of the chromaticness vector (u*,v*) is called Chroma (C*.) This experiment sought evidence regarding which model of the relationship between lightness and colorfulness (relative mixture of the greatest purity of hue with gray) better approximates perception. As implied, CIELUV (with its chromaticity diagram) is the only one of the spaces we are considering to which this question applies.

TABLE FIVE

Chroma and Saturation

cyan	magenta	yellow	red	green	blue
C* 5 sat 4	sat 7 C* 2	C* 5 sat 4	sat 7 C* 2	sat 5 C* 4	C* 5 sat 4

In this experiment, subjects were presented with two ten by ten arrays of color patches as suggested, in grayscale, by Fig. 1C. The left column of each array was a series of ten equal steps in lightness along the neutral axis. The highest lightness occupied the top row. The rightmost column in one of the panels had a lightness series at constant C* while the other had a lightness series at constant saturation. Subjects were not offerred a "no preference" option, but were forced to vote for one of the panels. Position was controlled for by presenting each pairing of panels twice and alternating which of C* and saturation were uppermost. Table Five shows that neither model was consistently preferred. Rather, subjects felt that something intermediate to the two models was best. The trajectory of constant perceived colorfulness appears to depend on hue angle.

1.6 Summary of Perceptual Studies

Results of the five experiments have been presented in such a way as to clarify the reasons for choosing the spaces compared. Spaces such as YIQ or HSL and HSV (Smith, 1978) are easy to compute from RGB, but their coordinates have marked perceptual interdependencies. The luminance and chrominance coordinates of Kodak PhotoYCC are also computed directly from RGB by a non-projective transformation. Therefore, experiments one and two compared YCC to the CIE spaces with respect to *perceptual* separability of luminance and chromatic variables and *perceptual* uniformity. Although we did not show tabular results on the uniformity of YCC, the results of pilot observations during the definition of experiment two indicated that YCC was not uniform along hue trajectories, possibly due to confounding changes in perceived lightness. Given this, we simplified the experiment to a comparison of two spaces.

Experiments three and four afforded a closer look at the independence of perceptual dimensions in the CIE spaces. Although the subjects felt that the discriminations in experiments two, three and four were difficult, they also felt that CIELUV and CIELAB did a reasonable job of modelling what they were designed to model. In experiment five, however, subjects felt that the best model of the interaction between lightness and "colorfulness" was somewhere between saturation and chroma. The question of how to treat saturation during a tonal remapping affects not only the design of image editing application, but also the design of automatic gamut scalings among media of differing dynamic ranges.

2.0 WHITE POINT ADAPTATION

In a publishing application, it is common to scan a transmissive original with a tungsten lamp and to view the reflection copy under a fluorescent D5000 simulator. What the operator chooses as highlight in the original has a color close to that of film base + fog as lit by the scanner lamp and optics. That color has an "absolute" chromaticity which is likely to be very different from the chromaticity of the printing stock under a fluorescent. The operator or the system must map the chromaticity of "white" in the original to that of the reproduction. All other colors must be handled "consistently." There is no way to translate all colors exactly, except by processing each color in the spectral domain, since

$$\text{CIE TSV}_{X,Y\text{or}Z} = \sum_{\lambda} \text{sample}(\lambda) * \text{illum}(\lambda) * \text{cmf}(\lambda)\,\Delta\lambda$$

where TSV refers to TriStimulus Values (of the CIE Standard Observer) and cmf to the color matching functions or imputed spectral sensitivities of said Observer. Equation 1 shows that exact translation requires substitution of one illuminant spectrum for another, which is not a linear problem.

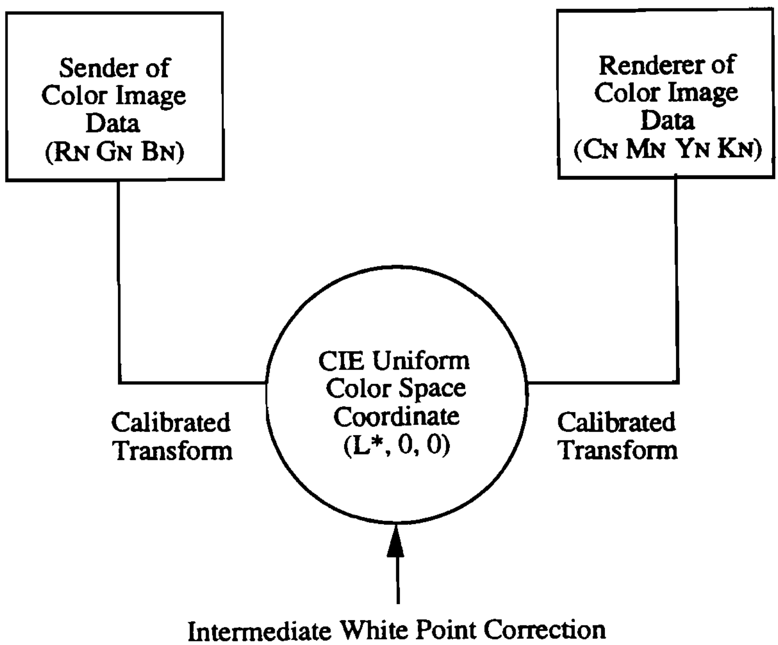

Figure Two: Mapping of a neutral pixel denoted by Rn, Gn, Bn to a different neutral on a different medium through a medium-independent neutral.

Both CIELAB and CIELUV incorporate a model of "white point adaptation" which maps colors consistently, if not exactly (neutrals are handled exactly.) In CIELAB, a*=b*=0.0 is a *relative* neutral, as is u*=v*=0.0 in CIELUV. The calculation of the starred coordinates uses the absolute TriStimulus Values or chromaticity, respectively, of the white point in a particular medium. However, this calculation can be incorporated into transformations between device coordinates and CIELUV or CIELAB so that the white point is handled transparently as colors are passed through the system (Fig. 2.) Of the spaces considered in this paper, only Lab and Luv have this property.

3.0 DEVICE INDEPENDENCE

All of the spaces considered in this paper meet this criterion. There are instruments capable of measuring the CIE XYZ values which would be selected by a Standard Observer in a color matching experiment. These coordinates are readily convertible to those of other spaces by exact equations. Parenthetically, the last condition is not satisfied by the Munsell Color Order System, although it is the earliest device-independent notation and the most uniform, perceptually.

4.0 EASY COMPUTATIONS ON INEXPENSIVE HARDWARE

This section has three parts. In the first, the "systems" and architectural issues affecting cost of computations are considered. This is *en lieu* of a detailed analysis of how many adds, multiplies and table look ups are needed to transform from one color representation into another or of relative performance based on various assumptions about hardware realization. The second section elaborates a point raised in the first regarding the relative accuracy of color transformations carried out in various spaces. The third reviews in a general way how the factors considered in 4.1 and 4.2 influence the implementation.

4.1 A Systems View of Computational Expense

Ease of computation might well appear at the head of many lists of desirable properties, especially those coming from developers of desktop publishing applications. From their perspective, the primary device is the color monitor and it is tempting to think that RGB signals from a scanner can be preconditioned in simple ways and loaded directly into a frame buffer for display. Considerations such as these explain the strong lobby for "calibrated RGB" among suppliers to the low end of the market. As long as neither the color fidelity of the reproduction nor the ease of performing color edits are issues, RGB offers the lowest cost of computations.

As expectations of quality rise in the desktop market, three of the limitations of the foregoing scenario will become apparent. First, few, if any, scanners currently on the market produce RGB signals which themselves bear an exact interpretation in terms of the Standard Observer. Because scanner channel sensitivities are generally *not* linearly related to human color matching functions (cf. Holub, Kearsley and Pearson, 1988) the scanner codes can be related to TriStimulus Values only by inexact and non-linear functions. Nonetheless, such functions are very useful in preparing images which more closely resemble the original copy than if no transform were used at all. They speed the processing of reproductions (system throughput) and improve their quality. The functions are sufficiently complicated that the computational cost of transformation does not depend on the color space into which device signals are transformed; however, the accuracy of conversion does (see section 4.2.)

Second, the price to be paid in converting into a space satisfying conditions 1 and 2 of the Introduction may be more than compensated by the savings in editorial processing time. For example, it is almost always sufficient to restrict spatial filtering to the lightness channel in CIELUV or CIELAB, rather than perform roughly three times the work to filter in RGB. Likewise, the ability to modify tone reproduction with a control independent of hue and saturation might improve throughput and reduce make-over.

Third, the highly efficient utilization of available quantization levels in the RGB cube which is cited for "calibrated RGB" is at the expense of gamut. When the encoding of RGB is modified to accomodate gamuts of practical interest in Graphic Arts, two things happen: 1) the efficiency of utilization decreases (see data of section 6) and 2) the feeding of RGBs to the frame buffer becomes more complicated (computational expense increases.)

4.2 Comparative Modelling

This section considers the accuracy of fitted models of the relationship between ink and color expressed as dependent variable in several color spaces. The model is generic, yielding color as a polynomial function of ink, after Pobboravsky (1962) and Vachon (1988). Spectral measurements of samples from an off-press proofing system and a sheet fed press were taken in 45/90 geometry and converted into the various color spaces of comparison by incorporating a D50 illuminant. Our metric consisted of the average color difference (CIELUV delta E*) between the measured values of color samples and the values predicted for the samples by the model. In other words, we compare color spaces on the basis of the average color error of their fits. We chose to model 4-color reflection devices because these are most non-linear and challenging.

We anticipate the criticism that one or more of the spaces might excel in concert with a different or physical model to which it is appropriate. Variations on Neugebauer's equations might be cited as physical models based directly on the TriStimulus Values, CIE XYZ. In our view, Neugebauer's model is not a physical one but a perceptual one. In it, the TSVs of ink solids and their overprints are treated as primaries analogous to the primaries used in the color matching experiments which defined the Standard Observer.

The original equations used Demichel's polynomial to predict the colors of ink mixtures. Close inspection reveals that Demichel's polynomial has precisely the form (for three inks) of the trilinear interpolatory function (each ink raised at most to the first power with the maximum sum of powers in a term equal to three.) The simplest Neugebauer model predicts the TSVs of inkings by trilinear interpolation in the ink cube. Some of the modifications to Neugebauer's equations, introduced to improve their predictive power, have physical rationales. Nevertheless, to our knowledge, they all depend on optimizing parameters, usually by a least squares technique. Thus, the class of model we utilize here can be viewed as a generalization of the most popular family of models of halftone color reproduction.

The results are assembled in Table Six. CIELUV and CIELAB perform very well. The linear spaces do poorly. Companded RGB and its derivative Kodak PhotoYCC do well, which reflects the fact that their companding functions resemble the transformation from XYZ to CIELAB. The larger average errors associated with the press have been shown to be related to process variation intrinsic to the device (Holub and Kearsley, 1989.)

TABLE SIX

Comparative Model Error of Fit

Color Space	SWOP Proof		Sheet Fed	
	Mean ΔE*	Max ΔE*	Mean ΔE*	Max ΔE*
CIE XYZ	2.83	28.70	3.64	19.70
CIE xyY	2.21	14.90	3.22	14.50
CIELUV	0.77	2.90	2.84	14.50
CIE Yu'v'	2.22	14.90	3.22	14.50
CIELAB	0.88	3.30	3.24	22.90
Linear RGB	2.97	30.29	3.60	16.90
Companded RGB	1.13	5.50	3.08	14.30
Kodak YCC	0.88	4.50	2.83	11.00

4.3 Implementation of Color Transformations

In section 4.1, it was argued that accurate conversions of color from scanners requires nonlinear, multivariate transformations. On the output side, the functions of device signals which yield color must be inverted in order to render imagery. Currently, real time inversion of output device functions is even less practical than real time evaluation of accurate input transformations. Dating back at least as far as Korman's (1971) work, it has been common practice to evaluate output transformations while rendering by interpolating in sparsely sampled tables of inverse function values. It is sensible to extend this technique to input transformations where color fidelity is required. In this case, the cost of computations is independent of color coordinate system. In the foregoing, we have not considered conversions for a CRT display. For high accuracy or the kinds of nonlinear transformations required for soft proofing, the considerations remain the same.

It should be acknowledged, however, that it is possible to achieve higher fidelity on a CRT without device calibration than is possible for scanners and hard copy devices. One may define "calibrated RGB" in terms of the average TSVs associated with the phosphor set of the monitor shipped with "the system" and compand the RGB signals so as to "precompensate" the gamma of the average device. In this case, the accuracy of color on the display will depend on the deviations of particular devices from the average; it may be possible to keep the deviations to a low level. In this scenario, "calibrated RGB" has a cost-of-computations advantage over other spaces. However, it is wedded to the particular class of display device and it may, as an image representation, sacrifice significant amounts of the gamut of some input media.

We conclude this section with one other practical aspect of implementation. In an interpolation table or addressing scheme based on the RGB cube, neutrals are arrayed along the diagonal through the cube. This can be shown (cf., Franklin, 1982) to result in oscillations in Gray Component Replacement along the neutral axis when standard, trilinear interpolation techniques are employed. The oscillations manifest themselves as variations in a gray scale in individual separations. Franklin described countermeasures derived from analytical work by Gallagher (1975.) Addressing schemes defined in terms of coordinate systems in which lightness (or neutral) forms one of the principal axes do not have this problem.

5.0 UNIFORMITY, BIT EFFICIENCY AND COMPRESSIBILITY

Poe and Gordon (1987) have treated the relationship between encodings of color and the efficiency of utilization of bits. They proved that the more perceptually uniform an encoding the fewer the bits needed to represent all discriminable colors. In this section we develop a metric for evaluating this relationship based on image compression. We compressed images using methods outlined in the draft ANSI standard JPEG-8-R5 (1990).

The proposed standard calls for Discrete Cosine Transformation of image data, followed by stages of quantization and Huffman encoding. Decompression reverses the stages to the extent possible, since information losses occur with quantization. We selected three images for study with varying degrees of complexity of color and spatial structure. The first transparency is included in the Kodak Q60 Color Reproduction Guides. It has a 35mm format on Ektachrome 64 Professional film and depicts an artist with a palette standing in front of a randomly painted background. The second (also a "Q60" transparency on 120mm Ektachrome 100) is an architectural scene including a dome, colonnades and a pond with fountain. The third is a 4x5 Ektachrome dupe of a young woman's head. The image is rich in grayscale information as well as fleshtones, but lacks color complexity.

All images were scanned and processed through a calibrated transformation to produce 24-bit, companded RGB images in TIF Format. By virtue of the calibration, each pixel has a colorimetric (device independent) interpretation and conversion among spaces is meaningful. The primaries and implied white point for this calibrated RGB were based on P22 phosphors and D5000, respectively. The companding applied to R, G and B signals was identical to the CIE's function for transforming cap Y to L*. We chose this as the reference space because something very much like it may be adopted by ANSI. Fig. Three presents a flow diagram of the processing thenceforth. The goal was to use a common treatment of the three images represented in each color space in order to derive a compression ratio, least squares error and average color error. These are the quantities presented in Tables Seven (one for each image.)

Compression was via the "baseline sequential codec" spelled out by JPEG (revision 5 of the draft standard, 1990; also see Wallace, 1991, for a readable review of the committee's progress.) The 24-bit stored TIFF image was converted to a floating representation in the target color space. Each channel was rescaled to range between 0 and 255, but was retained as a floating point quantity. Then, the Forward Discrete Cosine Transform, FDCT, was applied by simulating

Compression Data Flow

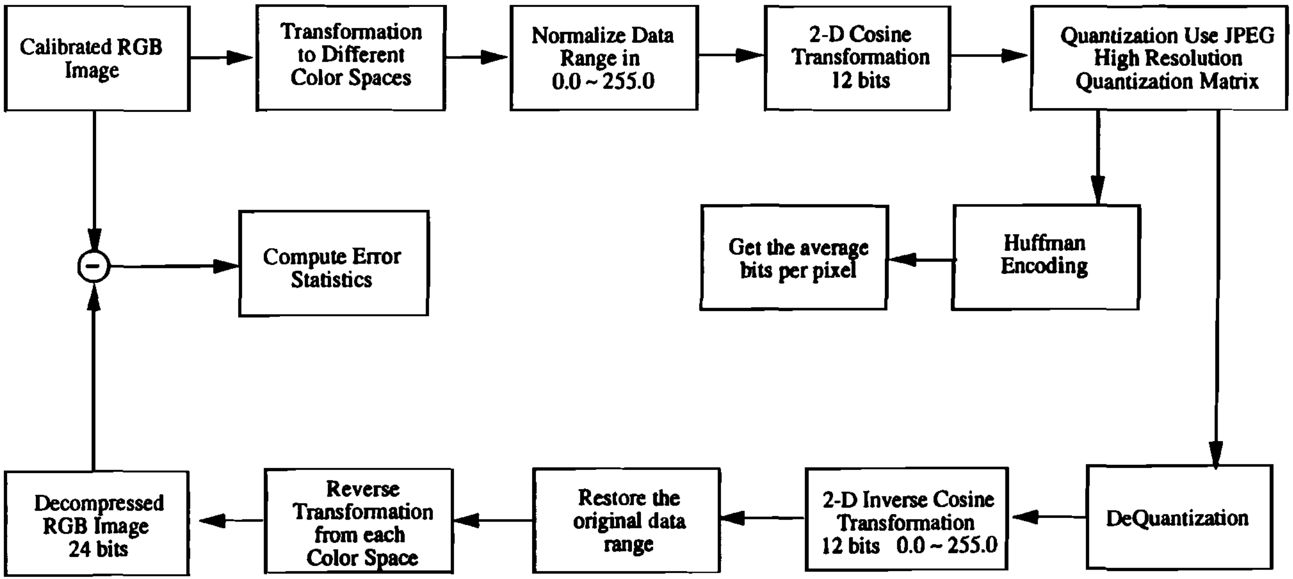

Figure Three

12-bit integer computations. Thereafter, each channel was processed through a suggested luminance quantization matrix (JPEG-8-R5, table 5.3.3.2.1.)

To derive the average bits per pixel from which compression ratio is computed, Huffman Encoding (via table 7.3.5.1.2.1) was performed as suggested by one branch of the flow diagram. Although we verified that we could decode images successfully to be sure our compression ratios are trustworthy, we did not ordinarily store them in coded form. Rather, we went to the second branch of the flow diagram and reversed the steps leading to the quantized representation of the image. The Least Square Errors quoted in Table Seven are averages of the Pythagorean distances of decompressed image pixels (in 24-bit companded RGB coordinates) from the original values. We also computed the average CIELUV color difference in order to state the changes due to compression/decompression in more meaningful, visual terms.

In considering the results, several comments are relevant. First, 24 bits are more than enough to represent images in some color spaces but insufficient in the linear (uncompanded) RGB and CIEXYZ spaces, and possibly in others. (For example, we have found that many images look acceptable when represented by uncompressed CIELUV using 8 bits of L* and 6 each of u* and v*.) This means that the compression ratios quoted for linear, "calibrated" RGB and CIEXYZ are probably too generous.

TABLE 7-1

Compression Results for Image One

Color Space	Compression Ratio	Least Square Error	CIELUV ΔE*
CIE XYZ	5.15	3.98	2.91
CIE xyY	5.04	3.17	1.93
CIELUV	5.1	2.92	1.85
CIE Yu'v'	5.31	3.48	2.09
CIELAB	5.69	3.23	2.27
Linear RGB	4.36	2.66	1.73
Companded RGB	3.86	1.83	1.38
Kodak YCC	3.86	1.84	1.34

Second, nothing was done to optimize the compression for any color space(s). (For example, it is tempting to downsample, or otherwise to treat differentially, the chrominance channels of spaces such as Yu'v' or Kodak PhotoYCC.) Instead, we sought a level playing field by using generic procedures. Although we did not give

TABLE 7-2

Compression Results for Image Two

Color Space	Compression Ratio	Least Square Error	CIELUV ΔE*
CIE XYZ	5.91	3.77	2.23
CIE xyY	5.89	3.07	1.5
CIELUV	5.89	2.56	1.32
CIE Yu'v'	6.19	3.31	1.54
CIELAB	6.49	2.73	1.55
Linear RGB	5.13	2.66	1.41
Companded RGB	4.36	1.6	1.05
Kodak YCC	4.43	1.68	1.03

TABLE 7-3

Compression Results for Image Three

Color Space	Compression Ratio	Least Square Error	CIELUV ΔE*
CIE XYZ	9.2	3.89	2.37
CIE xyY	10.23	3.01	1.3
CIELUV	10.03	2.3	1.01
CIE Yu'v'	11.13	3.04	1.32
CIELAB	11.16	2.27	1.03
Linear RGB	8.92	3.02	1.51
Companded RGB	5.16	1.99	1.06
Kodak YCC	5.64	2.14	1.11

chrominance channels (in spaces which have them) special treatment (doing so would have increased the compression ratios) these spaces are the most compressible. CIELAB consistently has the highest compression ratios, but CIELUV, overall, has the best compressibility at a given color error. This was evaluated in terms of the quotient of compression ratio to color error. Kodak PhotoYCC also did well in this regard. It should be acknowledged that CIELAB might have shown better quotients if color error were expressed in its own color difference units rather than CIELUV's.

Third, the average errors for companded RGB represent a kind of baseline due to the compression simulation itself, since it was the reference space. Incremental errors for other spaces include a modest quantization error due to conversion from 24-bit companded RGB. We estimate this error to be about 0.5 JND or delta E* unit. The errors are this small because of the precision we carried through the calculations.

Fourth, errors for the linear spaces are probably as good as they are due to the use of 12-bit integer simulations. The use of the baseline codec with 12-bit simulations explains why the compression ratios, overall, are somewhat low.

Fifth, the relative advantage of CIELUV and CIELAB is most evident in the compression ratios for image three. This is consistent with the true perceptual separation of gray scale information in these spaces (image three does not have great color variety.)

6.0 MAXIMUM OCCUPATION OF ENTRIES IN A REGULAR QUANTIZATION SCHEME

Section five examined one aspect of digital efficiency. This section is a treatment of another aspect which may not be entirely compatible with the former. On the one hand, a color space may be able to represent the largest number of discriminable colors with the fewest bits by virtue of its perceptual uniformity. On the other, its representations of practical gamuts may have shapes which leave large open areas in a regular digital addressing scheme.

We studied this with reference to three gamuts: 1) that of a positive reversal film, 2) Pointer's (1980) gamut of real surface colors and 3) that of an off-press proofing system. We converted the sets of TriStimulus Values representing these gamuts into the various spaces of comparison. We then defined the coordinates of the minimum bounding cuboid, i.e., the box whose faces impinged the gamut on each side along each coordinate axis. We then divided the boxes up into equal numbers of quantization cells and computed the numbers of cells that were in gamut as a percentage of the total. We did not consider the minimum bounding cuboid for the union of the three gamuts although it would be interesting to do so. The results are assembled in Table Eight.

Some of our labels deserve explanation. CIE Yu'v' used the chromaticity coordinates of the 1976 Uniform Chromaticity Scale and the cap Y luminance variable. Its utilization would probably have been somewhat better had we used psychometric lightness, L*, instead. YCC consists of the Luma and C1, C2 chroma components of Kodak PhotoYCC. Linear and companded RGBs are based on the CCIR 709 recommended primaries (for high definition TV) with an implied D65 neutral. These are the primaries underlying PhotoYCC. CIEXYZ stands for the linear form of the TriStimulus Values themselves.

TABLE EIGHT

Utilization of Quantization Scheme

	POINTER	FILM	SWOP
CIELAB	23.6%	30.9%	20.8%
CIELUV	23.8%	32.0%	21.0%
CIEYu'v'	14.4%	12.9%	16.6%
Kodak YCC	21.6%	23.5%	13.8%
CIEXYZ	11.3%	13.6%	7.0%
Linear RGB	37.4%	39.9%	28.7%
Companded RGB	35.9%	36.1%	27.6%

The poorest utilization is that of CIEXYZ. The best is that of linear RGB. Nonetheless, it is worth noting how much of the "perfect" utilization of the RGB cube is lost when non-physical, for a high definition device, values of the coordinates are used in order to accomodate the gamuts of other devices. Note that in going from linear RGB to companded RGB, there is a substantial gain in uniformity without much loss in utilization and that the same can be said of going from companded RGB to CIELUV.

7.0 NOTION OF COLOR TOLERANCE

Only CIELAB and CIELUV purport to be three-dimensionally uniform and have a defined metric of color difference. Of the two, CIELAB's Just Noticeable Difference has the wider currency and a long history of application in industry and science.

8.0 FACILITY WITH ADDITIVE COLOR MIXTURE

All the spaces with acceptable bit efficiency use non linear encodings of the TriStimulus Values and must therefore be decoded for additive mixture. For CIELUV, this involves backing out of L*u*v* to Yu'v'. For CIELAB, it involves converting from L*a*b* to XYZ. YCC must be converted back to companded RGB and then uncompanded. The computations can probably be organized such that there is no marked penalty associated with any of the spaces.

9.0 SUMMARY AND CONCLUSIONS

1. CIELUV, CIELAB and CIE L*u'v' provide useable models of hue, saturation and lightness as independent perceptual primitives. This is a significant advantage in image editing applications, especially those involving unskilled operators. In a number of respects, a majority of subjects judged CIELUV to be a better model than CIELAB, but for most practical purposes, the two spaces were indistinguishable in our experiments.

2. CIELAB and CIELUV incorporate a useable model of white point adaptation which can significantly lessen the computational burden in cross-rendering applications. None of the other spaces considered offer this advantage.

3. The question of computational expense should be viewed from the architectural perspective of what the reprographic system is intended to do. Once it is required that the colors at various stages in the processing have a known colorimetric relationship to one another, none of the spaces has a clear computational advantage because of the nature of practical means of implementing the color conversions.

4. Consistent with their perceptual uniformity, CIELAB and CIELUV yielded the best compression ratios and the best ratios in relation to the average color error in decompressed images. Kodak PhotoYCC also performed well.

5. Linear RGB made the most efficient use of a regular quantization scheme and companded RGB was second. CIELUV and CIELAB performed almost as well.

Although the CIE spaces were not specifically designed for digital image processing, they are the fruit of decades of evolutionary work by many eminent scientists. The design goals incorporated into CIELUV and CIELAB by those scientists are very relevant to imaging applications and so the findings that the spaces excel in a digital milieu should come as no surprise. Although we think that CIELUV is a better space, scientifically, we cannot fault CIELAB for practical purposes.

LITERATURE CITED

ANSI (American National Standards Institute - X3L2.8)
1990 "JPEG draft technical specification (revision 5)" X3 Information Systems Secretariat: Computer and Business Equipment Manufacturers Association, Wash. D. C. pp. 125.

Bartleson, C.
1980 "Colorimetry" Ch. 3 in Grum, F. and Bartleson, C., Eds., *Color Measurement,* (Academic Press, New York), pp. 33-148.

C I E (Commission Internationale de l'Eclairage)
1978 "Recommendations on uniform color spaces, color-difference equations and psychometric color terms" Supplement No. 2 to *Colorimetry,* Publication No. 15 (Bureau Central de la CIE, Paris)

Franklin, P.
1982 United States Patent 4,334,240 (June 8, 1982)

Gallagher, R.
1975 Finite element analysis: Fundamentals (Prentice-Hall, Englewood Cliffs, N. J.), 420 pp. Esp. chapter 8, pp. 211-248.

Holub, R., Kearsley, W. and Pearson, C.
1988 "Color systems calibration for Graphic Arts. I. Input devices" *J. Imag. Technol.* **14:** 47-52.

Holub, R. and Kearsley, W.
1989 "Color to colorant conversions in a colorimetric separation system" *Neugebauer Memorial Seminar on Color Reproduction* SPIE Vol. **1184:** 24-35.

Huntsman, J.
1988 "A planar vector model of color vision and the application of its color space to color reproduction" *Proc. Tech. Assoc. Graphic Arts* **40:** 29-61.

Kodak (Eastman Kodak Company)
1991 "Photo CD: A planning guide for developers" Kodak Information Center, Rochester, N.Y. 14650, p. 8ff.

Korman, N.
1971 United States Patent 3,612,753 (October 12, 1971)

Levkowitz, H. and Herman, G.
1988 "Towards a uniform lightness, hue, and saturation color model" *Image Processing, Analysis, Measurement, and Quality* SPIE Vol. 901: 215-222.

Newhall, S., Nickerson, D. and Judd, D.
1943 "Final report of the O. S. A. subcommittee on the spacing of the Munsell colors" *J. Opt. Soc. Amer.* 33: 385-418.

Pobboravsky, I.
1962 "A proposed engineering approach to color reproduction" *Proc. Tech. Assoc. Graphic Arts* 14: 127-165.

Poe. R. and Gordon, J.
1987 "Quantization effects in digital imaging systems" *Proc. Tech. Assoc. Graphic Arts* 39: 230-255.

Pointer, M.
1980 "The gamut of real surface colours" *Color Research and Application* 5: 145-155.

Pointer, M.
1981 "A comparison of the CIE 1976 colour spaces" *Color Research and Application* 6: 109-118.

Smith, A.
1978 "Color gamut transform pairs" *Proceedings, Fifth Annual Conference on Computer Graphics and Interactive Techniques,* Assoc. Comp. Mach., pp. 12-19.

Taylor, J., Murch, G. and McManus, P.
1988 "TekHVC: A uniform perceptual color system for display users" Technical Report No. UIRL-901-001, Tektronix Laboratories, Beaverton, OR 97077.

Vachon, G.
1988 "Modeling the mixing behavior of inks with polynomials" *Color Res. Appl.* 13: 46-49.

Wallace, G.
1991 "The JPEG still picture compression standard" *Communications of the ACM,* **34:** 31-44.

ACKNOWLEDGMENTS

Photo CD and PhotoYCC are trademarks of the Eastman Kodak Company. The authors thank Ms. S. Warren for help in preparing the manuscript and Messrs. R. Ewbank and J. Gordon for helpful discussions.