Method for Evaluating the Color Gamut and Quantization Characteristics of Output-Referred Extended-Gamut Color Encodings

Gustav Braun and Kevin Spaulding
Eastman Kodak Company
Rochester, New York

Abstract

The charter of the Commission Internationale de l'Eclairage (CIE) Division 8 Technical Committee 5 is to recommend a "minimal set of techniques that enable unambiguous and efficient communication of the color information in images," including defining "a minimal set of standard color spaces that addresses a wide range of imaging applications." One of the main activities of this committee to date has been the development of a series of objective tests that can be used to compare the characteristics of different color encodings. This paper will describe a number of metrics that have been developed relating to color gamut and color quantization. Example results are presented where these metrics have been applied to a number of recently proposed color encodings: sRGB, e-sRGB, sYCC, e-sYCC, ICC PCS, and ROMM RGB.

Introduction

The need for extended gamut color spaces is becoming increasingly important as the sharing of digital images becomes more common. Standard color encoding spaces such as sRGB have made it easier to share images in an open-systems environment. At the same time, the form of the sRGB color encoding places unacceptable restrictions on color gamut for some applications. As such, digital imaging systems that include sRGB as a central component of their workflow are limited in their ability to accurately produce output colors that are outside the CRT-centric sRGB gamut.

Currently, there are a number of extended-gamut output-referred color encodings that have been proposed in various standards forums. These include sYCC, e-sRGB, e-sYCC, ROMM RGB and ICC PCS LAB. The sYCC, e-sRGB, and e-sYCC color encoding spaces are all extensions of the sRGB color encoding. They have adaptive white point chromaticities equal to CIE Illuminant D65 and are based on the sRGB primaries. While the chromaticities of the sRGB primaries define a gamut limited to that of a typical video display, the encoding processes used in these color encodings extend the range of encodable colors beyond that of sRGB by allowing for negative amounts of the primaries and amounts greater than that used to produce the white point.

In contrast, the ROMM RGB color encoding uses a set of extended-gamut imaginary primaries to achieve a larger color gamut, and PCS LAB is based on the CIELAB color space. Both ROMM RGB and PCS LAB have adaptive white points with the chromaticities of CIE Illuminant D50.

To accurately compare the characteristics of the various color encodings, it was necessary to map them all to a common color space. PCS LAB was selected because well-defined transforms existed to most of the other color encodings, and because it offered convenient access to a number of metrics based on the CIELAB color space (e.g., color difference calculations). Whereas PCS LAB is defined for a D50 white point, the D65-based color encodings (sRGB, sYCC, e-sRGB, and e-sYCC) were chromatically adapted to a D50 white point (using a cone-space von Kries 3 x 3 matrix based on the Hunt-Pointer-Estevez cone responses).

Color Encoding Evaluation Criteria

The members of CIE TC805 have proposed a series of criteria to evaluate the characteristics of various color encodings. These criteria included:

1. Gamut volume characteristics,
2. Color quantization characteristics,
3. Visual uniformity,
4. Complexity of transformation required to and from typical standard spaces (sRGB, ICC PCS, etc.),
5. Compressibility,
6. Compatibility with standard workflows (e.g., Photoshop software),
7. and others.

The focus of this study was to provide evaluation metrics for the first two items in the preceding list (i.e., gamut volume and color quantization characteristics.) The TC will take up the remainder of these criteria in the near future. The following sections provide a description of the objective tests performed and the results of those tests.

Color Gamut Evaluation

A series of color gamut criteria were developed to quantify the gamut sizes of the color encodings and to compare their ability to encode sets of important colors (referred to as target colors). In particular, four objective metrics were developed:

1. Total CIELAB gamut volume of the color encoding for *legal colors*.
2. Percentage of target color gamut contained within color encoding.
3. Percentage of color encoding gamut volume used to encode target color gamut.
4. Percentage of color encoding code values used to encode target color gamut.

For the purposes of this evaluation, the following target gamuts were defined:[*]

1. Real-world surface colors;[†]
2. "Optimal" surface colors;[‡]
3. "Legal" colors;[§]
4. CRT colors (as defined by sRGB);
5. Photographic print colors;[**]

Metric 1: Calculation of CES Volume

The first color gamut metric (Metric 1) is the total CIELAB gamut volume for the color encoding. For this metric, the color gamut was constrained by the gamut of legal colors[c] to account for the inflated gamut volumes of some of the color encodings that resulted from the fact that certain code value combinations correspond to imaginary colors outside the spectrum locus. Also, any regions of color space that were defined to be illegal by the color encoding specification were also eliminated from the color gamut volume calculation. (A detailed flowchart illustrating the algorithm used to create the gamut and calculate the CIELAB volume is shown in the Appendix Fig. A-1. Note: The color encoding could be completely described by a tetrahedral tessellation defined in a device code value domain. As such, no hulling operations were required.) The gamut volume calculation was performed by defining a tetrahedral tessellation in the device code value domain and then determining the corresponding tetrahedra in PCS LAB. Any tetrahedra containing illegal colors were then truncated accordingly, and the gamut volume was computed by summing the volumes of each tetrahedron. The volume of a tetrahedron was calculated by taking one sixth of the volume of the parallelepiped created by the vector space formed by the four points of the tetrahedron:[f]

$$t_{vol} = \frac{1}{6}\begin{vmatrix} p_1(L^*)-p_0(L^*) & p_1(a^*)-p_0(a^*) & p_1(b^*)-p_0(b^*) \\ p_2(L^*)-p_0(L^*) & p_2(a^*)-p_0(a^*) & p_2(b^*)-p_0(b^*) \\ p_3(L^*)-p_0(L^*) & p_3(a^*)-p_0(a^*) & p_3(b^*)-p_0(b^*) \end{vmatrix} \quad (1)$$

where $\|$ represents the determinant of the matrix representing the vector space formed by the four vertices of the tetrahedron {p0, p1, p2, p3}.

While this metric gives an indication of the total useful color gamut associated with a color encoding, it is difficult to draw conclusions from this metric alone. For many applications, bigger may not always be better if portions of the larger color gamut never get used by any real colors that are likely to be encountered in that particular application. This led to the development of several additional gamut metrics that are better able to answer the question of how well a particular color encoding addresses a set of target colors that are important for a particular application.

Metrics 2: Percentage Target Volume Contained Within Color Encoding

Metric 2 was computed by comparing the volume of the intersection of a particular target color gamut and the color-encoding gamut to the gamut volume of the target gamut:

$$p_2 = \frac{vol(CES \cap T)}{vol(T)}, \quad (2)$$

where *CES* and *T* represent the gamuts of the color encoding and the target color gamut respectively. The *vol()* process is a gamut volume operator that computes total volume by summing the volumes of each individual tetrahedron in the respective gamut volume. This metric is intended to give an indication of how fully the color encoding covers the colors in the target color gamut.

Metric 3: Percentage Color Encoding Gamut Volume used to Encode Target Color Gamut

Similarly, Metric 3 was computed by comparing the volume of the intersection of the target gamut and the color-encoding gamut to the volume of the color-encoding gamut:

$$p_3 = \frac{vol(CES \cap T)}{vol(CES)} \quad (3)$$

This metric is intended to give an indication of how much of the color encoding gamut is used by the colors in the target color gamut. Unfortunately, it was found that the results of this calculation were sometime misleading due to the fact that the color encoding gamut was artificially large

---

* In each case, the tristimulus values were normalized accordingly to be consistent with the PCS LAB color encoding.

† Determined by fitting a convex hull to a set of 1739 measured surface color patches including; a set of Munsell samples (non-fluorescent) from the Color-Cascade set; a set of surface color objects measured by Trussel; a set of DuPont paint samples; and a set of graphic arts spot colors. (Of these 1739 colors, 193 of were on the surface of the convex hull.) The colorimetry was calculated for CIE Illuminant D50.

‡ The *optimal surface color gamut* was calculated for CIE Illuminant D50 using the algorithm developed by MacAdam.[7,8] (The original calculations were performed for CIE Illuminants A and C.[9])

§ The gamut of *legal colors* was defined by points falling inside the spectrum locus having a luminance value between the black and white points of the ICC PCS reference medium.

**Created from dye spectra representing common commercial photographic print papers using a D-max and D-min scaled within the PCS.

f Like a regular cube, a given parallelepiped contains six equally sized tetrahedron.

for some of the color encodings due to the distortions introduced for colors corresponding to negative tristimulus values, etc. This led to the development of Metric 4, which is believed to be a more meaningful metric for addressing this same basic concept. (A detailed flowchart illustrating the algorithm used to calculate the gamut intersections and volume ratios for metrics 2 and 3 is given in Fig. A-2.)

Gamut Intersection Calculation

The gamuts of the intersections, needed for the computation of Metrics 2 and 3, were calculated by identifying vertices shared between the CES and target gamuts. This subset of points was fit using a 3D α-shapes process similar to the one presented by Cholewo and Love. The α-shapes process was used in place of a convex-hull operation because many of the color encoding gamuts have concave features. Fitting the intersection of a convex set and a concave set would result in an overestimation of the resulting solid in the regions dominated by the concave surface of the CES. The α-shapes process results in a concave set of tetrahedra that represents a complete tessellation of the intersection vertices. Thus, the summed tetrahedral volume calculations could be used to accurately compute the gamut volume.

Metric 4: Percentage Color Encoding Code Values used to Encode Target Color Gamut

Metrics 3 was intended to give an indication of how efficiently a particular color encoding is used to encode a target gamut by comparing the gamut volume used by colors in the target color gamut to the total gamut of the color encoding. However, as was mentioned earlier, this metric turned out to be susceptible to misleading results due to distortions in the gamut volume calculations for regions of the color encoding code value space corresponding to imaginary colors. A more meaningful metric was found to be the fraction of the code values used to encode colors in the target color gamut, rather than the fraction of the gamut volume. The procedure for computing this metric is as follows. A set of random code value vertices were converted into PCS LAB values and compared to the target gamut. Of these random vertices, the sub set of in-gamut points was identified by determining if a given color value was inside any of the tetrahedra comprising the target gamut. Metric 4 was calculated by determining the ratio of the in-gamut vertices to the total number of vertices in the random set. (A detailed flowchart describing the algorithm for this test is shown in Fig. A-3 of the Appendix.)

Color Gamut Evaluation Results

Using the techniques and the target gamuts outlined in the previous section, gamut volume comparisons were generated for the six example color encodings. Table 1 shows the results for Metric 1. Calculations were also performed for Metrics 2, 3 and 4 using the various target color gamuts described earlier. However, these results are too extensive to be given in their entirety in these proceedings. Therefore, an illustrative set of results is given in Tables 2 and 3 for the real-world surface color gamut and the optimal surface color gamut, respectively. (Note: The results from the photographic paper gamut, real-world surface color gamut, and the CRT color gamut were the same in terms of the overall rankings.)

Table 1. Metric 1: Total color encoding gamut volume (constrained by legal color gamut).

Color Encoding	Total Legal Gamut Volume (ΔE_{ab}^3)
sRGB	821,000
sYCC	2,725,000
e-sRGB	3,750,000
e-sYCC	4,600,000
ROMM RGB	2,520,000
PCS LAB	4,110,000

Table 2. Gamut intersection results for target color gamut of real-world surface colors.

Color Encoding	Metric 2	Metric 3	Metric 4
sRGB	61%	85%	75%
sYCC	99%	7%	28%
e-sRGB	100%	3%	16%
e-sYCC	100%	2%	7%
ROMM RGB	100%	40%	29%
PCS LAB	100%	18%	17%

Table 3. Gamut intersection results for target color gamut of optimal surface colors.

Color Encoding	Metric 2	Metric 3	Metric 4
sRGB	29%	100%	100%
sYCC	90%	16%	54%
e-sRGB	100%	7%	35%
e-sYCC	100%	5%	16%
ROMM RGB	83%	82%	74%
PCS LAB	95%	41%	41%

The results of total legal gamut volume show that all the extended-gamut color encodings have significantly larger color gamuts than sRGB. As expected the YCC based color gamuts (sYCC and e-sYCC) are larger than their RGB based counterparts (sRGB and e-sRGB). (Not surprisingly, as shown in the next sections, these color encodings pay a quantization price for their extra gamut.)

For metrics 2, 3, and 4 larger numbers tend to indicate higher performance. The misleading results that can be obtained with Metric 3 can be illustrated by looking at the e-sRGB line in Table 2. While Metric 3 indicates that only 3% of the gamut volume is used for colors in the real world surface color gamut, Metric 4 shows that a more respectable 16% of the code values are used for this purpose.

Other than sRGB, all of the color encodings were, essentially, able to encode the real-world surface color gamut in its entirety. The sYCC and the ROMM RGB CES used the largest percentages of quantization states to encode this gamut while the e-sYCC CES used the least number of quantization states.

The results of the optimal surface color gamut intersections indicate the percentages of idealized surfaces colors that are encodable within a color encoding. Both e-sYCC and e-sRGB were able to encode the entire set of optimal surface colors. The ROMM RGB, PCS Lab, and sYCC CES were able to encode most colors (i.e. more that 80 percent). While the sRGB CES was only able to encode about 30 percent of the optimal surface colors. (It should be remembered that the optimal surface color gamut corresponds to idealized block dye spectra, and is significantly larger than is likely to be encountered in any real surface colors.)

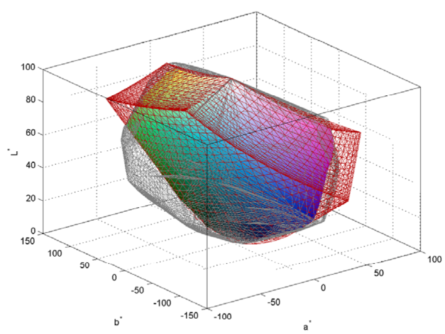

Figure 1. sRGB/real-world surface color gamut intersection. “Red” wire-frame gamut represents sRGB gamut. “Gray” wire-frame gamut represents real world surface color gamut. “Color-shaded” solid gamut is intersection of the two wire-frame gamuts

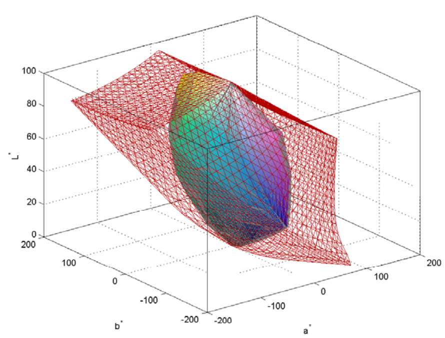

Figure 2. ROMM RGB/real-world surface color gamut intersection. “Red” wire-frame gamut represents ROMM RGB gamut. “Gray” wire-frame gamut represents real-world surface color gamut. “Color-shaded” solid gamut is the intersection of the two wire-frame gamuts.

An example gamut intersection is shown in Fig. 1. The target gamut in this example is the real-world surface color gamut described above, and the color encoding is sRGB. (Notice the concave components of the sRGB gamut that predicate the use of the α-shapes hulling algorithm.) From this illustration, it is clear that the sRGB CES is not capable of encoding all the target colors. Additionally, not the entire sRGB gamut is used by colors in the target color gamut. In comparison, the ROMM RGB color encoding is capable of encoding all of the colors within the target gamut as is illustrated in Fig. 2.

Color Quantization Evaluation

The color gamut evaluations described in the previous sections indicate which color encodings have the gamut necessary to encode important colors. This analysis does not, however, give a direct indication of how well the colors within the gamut are sampled. For example, a given color encoding may have an extremely large color gamut, and therefore the color difference between consecutive code values may also be large. Additionally, the quantization states within that gamut may be non-uniformly spaced. As a result, the quantization errors in different regions of the color space may vary significantly. This may present problems depending on the size of the quantization errors within the particular set target colors that are important for a particular application.

In order to quantify the quantization errors associated with various color encodings, the color difference for a single code value change in the color encoding was determined for a set of random color values within the target color gamut. For each color value in the set, the closest code value in the color encoding was determined, and the code value was incremented by one in the direction of each color component separately. The color differences (ΔE^* and ΔE^*_{94}) between the original encoded colors and the incremented colors were computed. Average and maximum color differences were then calculated for the set of test colors. A detailed flowchart describing the algorithm used in this test is shown in Fig. A-4 of the Appendix. Because the quantization efficiency of the color encoding is directly linked to the bit depth used in the encoding, the quantization statistics were calculated for encoding bit depths of 8, 10, 12, and 16 bits (realizing that not all of the color encodings were explicitly defined for each of these bit depths).

In order for the comparisons of different color encodings to be meaningful, it is important that the quantization errors be evaluated using a common set of test colors. The only way to ensure that this is possible is to limit the test colors to fall within the intersection of all of the color encoding gamuts being tested. In this case, the intersection of all of the color encoding gamuts turns out to be the sRGB gamut.

Another target gamut that is important for many applications is the real-world surface color gamut. Quantization error statistics were also computed for colors within this target gamut. Because the sRGB color encoding does not encode the entire real-world surface color gamut, sRGB could not be included in the evaluation performed using this set of test colors.

Color Quantization Evaluation Results

The results of the quantization analysis are shown in Tables 4-7 and Fig. 3. These results were calculated using the CIELAB ΔE^*_{94} color difference equation with the default parametric factors (i.e., $k_l = k_c = k_h = 1$). Results are shown for two target gamuts: the sRGB gamut and the real-world surface color gamut.

Table 4. Average Quantization Error (Within sRGB Gamut).

Color Encoding Space	8 bit ΔE^*_{94}	10 bit ΔE^*_{94}	12 bit ΔE^*_{94}	16 bit ΔE^*_{94}
sRGB	0.29	0.07	0.018	0.0011
sYCC	0.45	0.11	0.028	0.0017
e-sRGB	0.57	0.14	0.036	0.0022
e-sYCC	0.77	0.19	0.048	0.0030
ROMM RGB	0.48	0.12	0.030	0.0019
PCS LAB	0.47	0.12	0.029	0.0018

Table 5. Maximum Quantization Error (Constrained Within sRGB Gamut).

Color Encoding Space	8 bit ΔE^*_{94}	10 bit ΔE^*_{94}	12 bit ΔE^*_{94}	16 bit ΔE^*_{94}
sRGB	0.92	0.23	0.057	0.0036
sYCC	1.49	0.37	0.093	0.0058
e-sRGB	1.83	0.46	0.115	0.0072
e-sYCC	2.94	0.74	0.186	0.0116
ROMM RGB	2.18	0.53	0.133	0.0083
PCS LAB	1.00	0.25	0.062	0.0039

Table 6. Average Quantization Error (Within Real-World Surface-Color Gamut).

Color Encoding Space	8 bit ΔE^*_{94}	10 bit ΔE^*_{94}	12 bit ΔE^*_{94}	16 bit ΔE^*_{94}
sRGB	N/A	N/A	N/A	N/A
sYCC	0.46	0.11	0.029	0.0018
e-sRGB	0.58	0.15	0.036	0.0023
e-sYCC	0.78	0.20	0.049	0.0030
ROMM RGB	0.48	0.12	0.029	0.0019
PCS LAB	0.45	0.11	0.028	0.0017

Table 7. Maximum Quantization Error (Within Real-World Surface-Color Gamut).

Color Encoding Space	8 bit ΔE^*_{94}	10 bit ΔE^*_{94}	12 bit ΔE^*_{94}	16 bit ΔE^*_{94}
sRGB	N/A	N/A	N/A	N/A
sYCC	1.43	0.37	0.09	0.0058
e-sRGB	1.80	0.46	0.12	0.0072
e-sYCC	2.94	0.72	0.18	0.0115
ROMM RGB	2.16	0.54	0.14	0.0084
PCS LAB	1.00	0.25	0.062	0.0039

As expected, the quantization results for the YCC color encodings had higher average and maximum quantization errors compared to their RGB-based counterparts. As the number of bits used in the encoding increased, the corresponding color difference values dropped by a factor d given by:

$$d = \frac{2^{N_1}}{2^{N_2}}, \tag{4}$$

where N_1 and N_2 are the number of bits used in different encodings. That is, to convert the mean color error for 8 bits to the mean color error for 12 bits $N_1 = 8$ and $N_2 = 12$. Intuitively, this property makes sense and servers as a good check to the algorithms used in this analysis.

Quantization Efficiency

A useful extension of this analysis was to determine the required bit depth to achieve a threshold level of quantization. For example, consider the case where it is desired to encode colors within the real-world surface-color gamut with a mean color quantization error of less than 0.5 ΔE^*_{94} units, and maximum color quantization of less than 1.0 ΔE^*_{94} units. The corresponding bit depths needed to achieve these limits can be determined directly from the curves shown in Fig. 3 by setting the appropriate threshold. The resulting quantization efficiency values for this example are shown in Table 8.

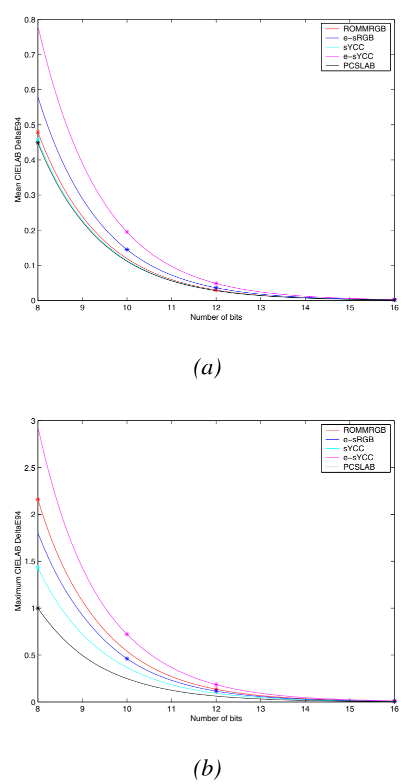

Figure 3. Mean (a) and maximum (b) quantization errors as a function of encoding bit depth.

Table 8. Quantization Efficiency for Various Threshold Quantization Limits (Within Surface Color Gamut).

Color Encoding Space	**Average $\Delta E^*_{94} = 0.5$**	**Max $\Delta E^*_{94} = 1.0$**
sRGB	N/A	N/A
sYCC	7.9	8.5
e-sRGB	8.2	8.9
e-sYCC	8.6	9.5
ROMM RGB	7.9	9.1
PCS LAB	7.8	8.0

Conclusions

This paper has described a number of color gamut and quantization metrics that can be used to evaluate various color encodings. The results of this analysis prove to be useful in gauging the efficiency of proposed color encodings to encode important sets of color data. In general, it can be seen that there is a trade off between the color encoding gamut volume and corresponding quantization efficiency. However, as the bit depth of the color encoding increases beyond 8 bits, the quantization errors quickly drop below acceptable thresholds. These metrics are part of a larger set of criteria being developed by the CIE TC 805 working group for purposes of objectively comparing different color encodings. The relative importance placed on any one criterion will be a function of the particular application for which the color encoding will be used.

Appendix

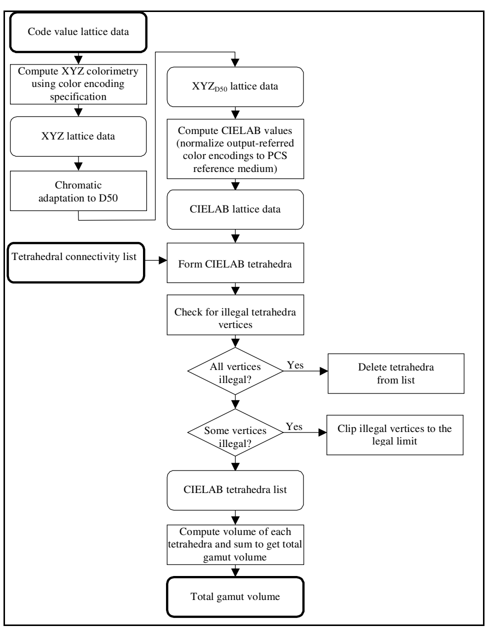

Figure A-1. Algorithm used to calculate the total gamut volumes of the CES (Metric 1).

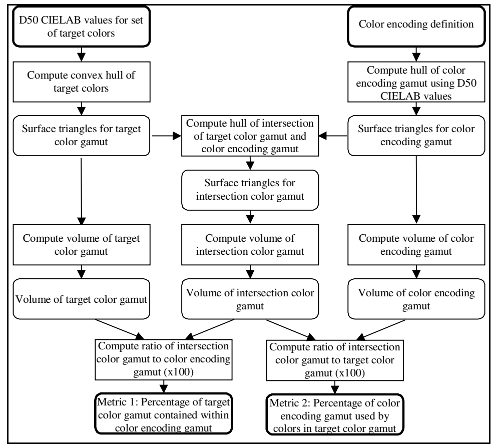

Figure A-2. Algorithm used to calculate the gamut volume intersections (Metrics 2 and 3).

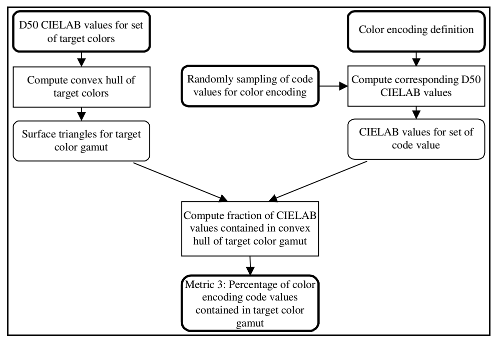

Figure A-3. Algorithm used to calculate the percentage of CES nodes contained used to encode the target color gamut (Metric 4).

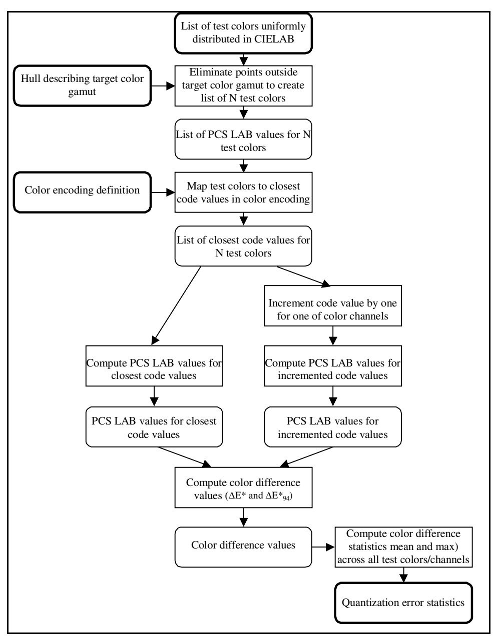

Figure A-4. Method used to calculate the quantization error statistics for the CES

References

1. "Multimedia Systems and Equipment – Colour Measurement and Management – Part 2-1: Colour Management – Default RGB Colour Space – sRGB," IEC 61966-2-1 (1999).
2. Committee Draft for Vote of "Amendment 1 to Multimedia Systems and Equipment - Colour Measurement and Management – Part 2-1: Colour Management – Default RGB Colour Space – sRGB," IEC 61966-2-1/A1/ Ed.1, 2001.
3. "Photography – Electronic still picture imaging – Extended sRGB color encoding – e-sRGB," PIMA 7667:2001.
4. "Photography – Electronic still picture imaging – Reference Output Medium Metric RGB Color encoding: ROMM-RGB," PIMA 7666:2001.
5. International Color Consortium, File Format for Color Profiles, ICC.1:2001
6. M. D. Fairchild, Color Appearance Models, Addison Wesely, Reading MA, pp. 250, 1998.
7. D. L. MacAdam, The theory of maximum visual efficiency of colored materials, *J. Opt. Soc. Am.*, **25**, 249, 1935.
8. D. L. MacAdam, Maximum visual efficiency of colored materials, *J. Opt. Soc. Am.*, **25**, 361, 1935.
9. G. Wyszecki and W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, John Wiley & Sons, New York, New York, Second Edition, pp. 179-183, 1982.
10. T. J. Cholewo and S. Love, Gamut boundary determination using α-shapes, *IS&T Seventh Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Az, pp. 200-204, 1999.
11. Commission International de l'Eclairage, Technical Report 116, Industrial Color Difference Evaluation, 1995.

Biography

Gustav Braun is a Principal Scientist in the Research Laboratories at the Eastman Kodak Company. He received his Ph.D. in Imaging Science from the Center for Imaging Science at the Rochester Institute of Technology in 1999. His current research interests include digital color reproduction, color gamut mapping, digital printing, and vision modeling. He is active in two CIE standards groups - Division 8 technical committees for Communication of Colour Information (TC805) and Gamut Mapping (TC803).

Kevin Spaulding received a BS in Imaging Science from Rochester Institute of Technology in 1983, and MS and PhD degrees in Optical Engineering from the University of Rochester in 1988 and 1992, respectively. He has been with Eastman Kodak Company since 1983 where he is currently a Senior Principal Scientist in the Imaging Science Division. He serves on several international standards committees that are working on the unambiguous communication of color images. He is project leader for the ISO 22028-1 standard, and is technical secretary for the CIE TC8-05 committee. His research interests include color reproduction, digital halftoning, image processing algorithms for digital camera and printers, and image quality metrics.