Interpolation of Color Data

Sarah A. Rajala and Atish P. Kakodkar
Dept. of Electrical and Computer Engineering
North Carolina State University, Raleigh, North Carolina

Introduction

Recently there has been an increased activity focused on solving the problems associated with the transmission and display of high quality color image/video signals. The objective is to generate an image or video signal whose quality is as high as possible at the output of a given system. Of specific concern in this research are the characterization and calibration of color output devices, for example printers or electronic displays. In order to make accurate judgments about the images produced, such devices must be calibrated. To do this, a relationship needs to be defined between a set of measured CIE $L^*a^*b^*$ values and a set of control values for the output device. Impacting the solution of this problem are: the model of the output device and its limitations, the type and location of the measurement data, the variability in the data, the interpolation and extrapolation methods used, the variability in the output device under normal operating conditions, and the perceptibility of the errors.

One way to characterize a color output device is with a three-dimensional look-up table. The look-up table maps the set of CIE $L^*a^*b^*$ values $\{\mathbf{t}_i\}$ to a set of control values $\{\mathbf{c}_i\}$ of the output device. The control values $\{\mathbf{c}_i\}$ are proportional to the concentration of the inks. The look-up table could be built directly, unfortunately this would require an excessive number of measurements. For a 32×32×32 table 32768 measurements would be required. An alternate approach for creating the look-up table is to make measurements on a coarser grid and then interpolate the values for a finer grid in the CIE $L^*a^*b^*$ space. The problem can be posed in the following manner. Let $\{(c_1, c_2, c_3)_i\}$ be the three-dimensional set of control values which map onto the set of CIE $L^*a^*b^*$ values $\{(t_1,t_2,t_3)_i\}$ such that

$$\mathbf{t} = F\,(\mathrm{c}) \tag{1}$$

where

$$\mathbf{t} = ((t_1, t_2, t_3) \tag{2}$$

and

$$\mathbf{c} = (c_1, c_2, c_3) \tag{3}$$

Given a set of control values {c} on a regular grid and the corresponding set of CIE $L^*a^*b^*$ values $\{\mathbf{t}\}$, find the $\{\mathrm{c}_g\}$ for different $\{\mathrm{t}_g\}$ on a finer grid in the CIE $L^*a^*b^*$ space. With an appropriately defined simpler interpolation scheme such as trilinear interpolation, the regular finer grid can be used in real-time to obtain the control value for any CIE $L^*a^*b^*$ value located inside this regular finer grid.

One of the applications of this work is in the calibration of nonlinear printing devices. For this research, we use the proposed color data interpolation scheme to calibrate a Kodak XL7700 color thermal dye transfer printer. To find the appropriate color interpolation scheme, a nonlinear forward model for the printer is defined ant an interpolation scheme selected and evaluated. Then, the limits of the printer model and its impact on the process, the impact of the variability of the printer under normal operating conditions, the choice of interpolation function and its impact on the results, the sensitivity of the calibration process to variations in the measurement data, and an evaluation of the results are presented.

The Forward Model

A printing process is subtractive in nature. Each of the cyan, magenta and yellow inks removes an amount of its complimentary color, red, green and blue respectively. The amount of a particular color removed is related to the concentration of the ink. Each ink is characterized by its density $\mathbf{D}(\lambda)$ at any wavelength λ. The density is proportional to the concentration c of the ink. Therefore,[1]

$$D_i\,(\lambda) = D_{i,max}(\lambda)\; c_i \;\; i = 1, 2, 3 \tag{4}$$

where $D_{i,max}$ is the density at unit concentration for ink i and c_i is the concentration of ink i. Transmission of a particular ink is related logarithmically to its density as

$$T_i(\lambda) = 10^{-D_{i,\max(\lambda)c_i}} \;\; i = 1, 2, 3 \tag{5}$$

The observed spectrum at a particular wavelength λ is given as [2]

$$\mathrm{g}(\lambda) = l(\lambda)\; T_1(\lambda)\; T_2(\lambda)\; T_3(\lambda) \tag{6}$$

where $T_i(\lambda)$ is the transmission of the i^{th} ink at wavelength λ and $l(\lambda)$ the intensity of the illuminant at wavelength λ. Using equation (5), equation (6) may be rewritten as:

$$g(\lambda) = l(\lambda)\; 10^{-\sum_{i=1}^{3} D_{i,max(\lambda)c_i}} \tag{7}$$

Thus, the forward model can be written mathematically as [2]

$$\mathrm{g} = \mathbf{L}\left[10^{-\mathrm{D}_{max}\mathrm{C}}\right] \tag{8}$$

where $\mathbf{L}$ is a NxN diagonal matrix representing the illumi-nant spectrum, $\mathbf{c}$ is the 3-vector representing the concentration of the inks and $\mathbf{g}$ is the N-vector representing the reflectance spectrum. The concentration values must be between zero and unity, and the $Nx3$ matrix of density spectra D_{max} represents the densities at the maximum concentration. This model ignores nonlinear interactions between the colorant layers. The tristimulus values *xyz* are obtained by the equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^T g \tag{9}$$

where **A** is a *3xN* matrix of the CIE color matching functions. The CIE *L*a*b** values are obtained by the following transformation [3]:

$$\begin{bmatrix} L^* \\ a^* \\ b^* \end{bmatrix} = \Upsilon \begin{bmatrix} (\frac{x}{x_n})^{1/3} \\ (\frac{y}{y_n})^{1/3} \\ (\frac{z}{z_n})^{1/3} \end{bmatrix} - \begin{bmatrix} 16 \\ 0 \\ 0 \end{bmatrix} \tag{10}$$

where

$$\Upsilon = \begin{bmatrix} 0 & 116 & 0 \\ 500 & -500 & 0 \\ 0 & 200 & -200 \end{bmatrix}$$

and x_n, y_n and z_n are normalization factors, i.e. the tristimulus values of a nominal 'white point'. The 'white point' represents the tristimulus values of a standard illuminant or a reference background point. The above formula is used

when $\frac{x}{x_n}, \frac{y}{y_n}, \frac{z}{z_n} \geq 0.01.$. When, $\frac{x}{x_n}, \frac{y}{y_n}, \frac{z}{z_n} < 0.01,$

$$\begin{bmatrix} L^* \\ a^* \\ b^* \end{bmatrix} = \mho \begin{bmatrix} f(\frac{x}{x_n}) \\ f(\frac{y}{y_n}) \\ f(\frac{z}{z_n}) \end{bmatrix} + \begin{bmatrix} 903.3\frac{y}{y_n} \\ 0 \\ 0 \end{bmatrix} \tag{11}$$

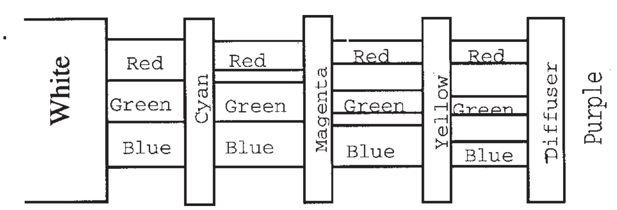

Figure 1. Subtractive Model

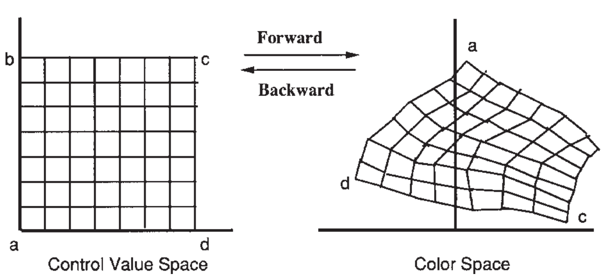

Figure 2. The Forward Model

where

$$\mho = \begin{bmatrix} 0 & 0 & 0 \\ 500 & -500 & 0 \\ 0 & 200 & -200 \end{bmatrix}$$

and

$$f(t) = \begin{cases} (t)^{\frac{1}{3}} & t > 0.008856 \\ 7,7787(t) + \frac{16}{116} & t \leq 0.008856 \end{cases}$$

The perceptual error between two points is defined as the Euclidean distance between two points in CIE *L*a*b** space and is called the ΔE value of the difference. A rule of thumb is that a equation ΔE error of 3 is perceptible, but smaller errors are not.

Developing the Look-up Table

Using the mathematical model, a coarse 8×8×8 look-up table mapping the control values to the output CIE *L*a*b** values is generated with the forward model of equation (8). We have

$$\mathbf{t} = \boldsymbol{F}(\mathbf{c}) \tag{12}$$

where **t** is a 3-vector of the CIE *L*a*b** values and **c** is a 3-vector of the control values.

A finer regular grid is constructed in the CIE *L*a*b** space. For a certain CIE *L*a*b** vector $\mathbf{t}_g$ on this finer regular grid the corresponding control vector $\mathbf{c}_g$ is obtained. Iterative techniques with interpolation are used to calculate the control vector corresponding to every vector $\mathbf{t}_g$, on the finer grid. For a given CIE *L*a*b** vector $\mathbf{t}_g$, an initial estimate $\mathbf{c}_o$ of the control vector is made and the corresponding CIE *L*a*b** vector is obtained by interpolating over the three dimensional regular grid of control values. The choice of $\mathbf{c}_o$ is arbitrary and can be selected anywhere in the control value space. In this work, the control vector $[0.5\ 0.5\ 0.5]^T$ which represents the center of the control value space was used as the initial estimate. Newton's method in three dimensions[4] is used to obtain a new estimate of the control value. This procedure is equivalent to finding a root of the equation

$$\mathbf{f}(\mathbf{c}) = \mathbf{t}_g \text{ or } \mathbf{f}(\mathbf{c}) - \mathbf{t}_g = 0 \tag{13}$$

where the function **f** refers to the interpolating function that is used. This iteration is carried out until equation (13) is solved to the selected degree of accuracy.

Two kinds of interpolating functions were used, the bell function and the cubic B-spline function. A bell interpolating function is obtained by the convolution of a triangle function with a square function and is defined as:[5]

$$S_b(x) = \begin{cases} \frac{1}{2}(x+\frac{3}{2})^2 & \frac{-3}{2} \leq x \leq \frac{-1}{2} \\ \frac{3}{4} - x^2 & \frac{-1}{2} \leq x \leq \frac{1}{2} \\ \frac{1}{2}(x-\frac{3}{2})^2 & \frac{1}{2} \leq x \leq \frac{3}{2} \end{cases} \tag{14}$$

A cubic B-spline is defined as:[5]

$$S_c(x) = = \begin{cases} \frac{2}{3} + \frac{1}{2}|x|^2 - x^3 & 0 \le |x| \le 1 \\ \frac{1}{6}(2 - |x|^3) & 1 \le |x| \le 2 \end{cases} \quad (15)$$

Three-dimensional separable interpolation is used to interpolate for the CIE $L^*a^*b^*$ values $t_g = (t_{g,1}, t_{g,2}, t_{g,3})$ on the finer grid

$$t_{g,1}(c_1, c_2, c_3) = \sum_k \sum_j \sum_i w_l(i,j,k) S(c_1 - i) S(c_2 - j) S(c_3 - k)$$
$$l = 1,2,3 \quad (16)$$

where $w_l(i, j, k)$ are the weights used to interpolate, S denotes the interpolation function used and (c_1,c_2,c_3) denotes the control vector. The values $t_{g,l}(c_{1,i},c_{2,j},c_{3,k})$ should be equal to the value of the ijk^{th} scalar entry in the given set of data points $t_{g,l}(c_1, i, c_2, j, c_{3,k})$, for $i, j, k = 1, 2, ... 8$ and $l = 1, 2, 3$. This condition provides a set of linear equations which are solved to provide suitable values for the weights $w_l(i, j, k)$. The work presented here is restricted to the determination of control values corresponding to CIE $L^*a^*b^*$ on the finer grid. For CIE $L^*a^*b^*$ values not lying on this finer regular grid, simple trilinear interpolation may be used.

Interpolation problems occur for points which are close to the boundary of the color gamut. To overcome this problem, a linear extrapolation is used to estimate the CIE $L^*a^*b^*$ values for points outside the given gamut. These points are then used along with the coarse 8×8×8 grid in the interpolation process. The experimental results of this simulation are presented in the next section.

Application to the Kodak XL7700 Printer

The Kodak XL7700, a color thermal dye transfer printer, was used to generate printouts of color patches by varying the concentrations of the cyan, magenta and yellow dyes in a uniform manner. The control value space was divided into eight equi-spaced samples in each dimension to generate 512 color patches. A GRETAG SPM50 colorimeter was used to measure the CIE $L^*a^*b^*$ values of each of these color patches. This set of CIE $L^*a^*b^*$ values forms known 8×8×8 look-up table.

Before proceeding further with the calibration process, it was important to determine if there were any variations in the printouts produced by the XL7700. Two types of inconsistencies were found:

1. The CIE $L^*a^*b^*$ values produced by identical control values varied depending on the position of the color patch on the paper on which it was printed. Preliminary data collected on several patches printed at different positions on a sheet typically gave ΔE errors greater than 3.5. Since the XL7700 has four heads printing over different areas of the paper a natural explanation for the error was that the amount of ink printed on to the paper was a function of the particular head that printed it. This has lead us to the conclusion that the calibration should be carried out for each individual head.

2. A certain amount of difference in the CIE $L^*a^*b^*$ corresponding to identical control values are observed when the printouts are obtained on different sheets. This indicated the need to measure the 512 color patches on different sheets to study the average ΔE error due to this observed anomaly. It was concluded that the 8×8×8 lookup table should be generated on the basis of an average over several printouts.

The above observations were taken into account to create the known 8×8×8 look-up table. The details are explained below.

Table 1: ΔE Errors Due to Interpolation

Statistics	Bell Function	Spline function
Max ΔE	0.6211	0.6252
Average ΔE	0.1178	0.1554
Variance(σ^2)	0.0105	0.0127

Experimental Results and Future Work

The two interpolating functions, the bell function and the cubic B-spline function worked well for the mathematical model of equation (8). For each of the CIE $L^*a^*b^*$ values on a test 16×16×16 finer grid in the $L^*a^*b^*$ space, the control value was obtained by Newton's iteration along with bell interpolation. Table 1 lists out the ΔE errors obtained when each of the interpolating functions described earlier were used.

These ΔE errors correspond to the case where the extrapolated points are used along with the known 8×8×8 grid. This error corresponds to only the interpolation error. Another error factor will be added after the finer regular grid is created and trilinear interpolation is used to determine the control value of any arbitrary point in CIE $L^*a^*b^*$ space which lies in the color gamut. If the grid is fine enough the error due to trilinear interpolation is expected to be small. The work presented in this paper indicates that it would be worthwhile to study the effects of grid size on trilinear interpolation.

In the actual calibration of the XL7700, the inconsistencies discussed in the previous section were studied.

1. Due to the first inconsistency, a representative head of the XL7700 was calibrated. Creating the known 8×8×8 grid required taking printouts of all the 512 color patches under the particular head.

2. Due to the second inconsistency, color patches were printed out on several different sheets. This enabled the study of the error caused by printing on different sheets using identical control values. Preliminary work to compare the ΔE errors among all pairwise combinations of sheets was carried out. The ΔE error was averaged out over all the 512 color patches and over all the sheets that were used. The average ΔE

computed in this fashion was 1.527. The variance was 0.5380 and the maximum ΔE (averaged over the different pairs) was 5.0323. On an average the total number of color patches which showed a $\Delta E > 3$ was 19 out of the 512 different color patches, which is about 3.71%. The average of the CIE $L^*a^*b^*$ values over different sheets of paper of the 512 color patches forms the coarse 8×8×8 look-up table.

Future work consists of testing out the calibration at different points in the color space by taking printouts of several color patches corresponding to points in the control value space which are obtained by statistical sampling. Statistical sampling seems to be the best choice to test the entire space in the most efficient manner. Calibration will be carried out for all the four heads. Error components will be introduced in the practical situation when creating the finer grid using interpolation, and when trilinear interpolation is used once the finer regular grid is created. The ΔE errors thus produced will be analyzed and studied in detail.

References

1. Hunt, R. W. G., *The Reproduction of Colour in Photography, Printing & Television*, Fountain Press, England, 1987.
2. Trussell, H. J., and Sullivan, J. R., A Vector Space Approach to Color Imaging Systems, *Proceedings*, *SPIE Conference on Image Processing Algorithms and Techniques*, Santa Clara, CA, 12-14 Feb. 1990.
3. Wyssecki, G. and Stiles, W. S., *Color Science: Concepts and Methods, Quantitative Data and Formulae*, John Wiley and Sons, 1982.
4. Dennis, J. E. and Schnabel, R. B., *Numerical Methods for Unconstrained Optimisation and Non-Linear Equations*, Englewood Cliffs, N.J., Prentice Hall, 1983.
5. Pratt, W. K., *Digital Image Processing*, John Wiley and Sons, 1991.