

The Use of Extrapolation in Computing Color Look-Up Tables

S. A. Rajala
A. P. Kakodkar

Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University

TR-94/2
March 1994

The use of extrapolation in computing color look-up tables

Sarah A. Rajala, H. Joel Trussell, Atish P. Kakodkar
Dept. of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7911

ABSTRACT

Colorimetric reproduction requires calibrated color output devices. One way to characterize a color output device is with a three-dimensional look-up table which maps the tristimulus values, \mathbf{t} , to the control values, \mathbf{c} of the output device. The functional form of the output device can be written in vector notation as $\mathbf{t} = F(\mathbf{c})$. The purpose of calibration is to define an inverse mapping from tristimulus values to control values. Since the function $F(\cdot)$ has no closed form, it is defined by interpolation from a table of values.

Given a set of control values $\{\mathbf{c}_i\}$ on a regular grid and the corresponding set of tristimulus values $\{\mathbf{t}_i\}$ obtained from data collection, we wish to find the $\{\mathbf{c}_g\}$ for different $\{\mathbf{t}_g\}$ on a grid in the tristimulus space. The grid is obtained from a relatively sparse data set with an appropriately defined interpolation scheme. This interpolation scheme can be complex since it is used only once to compute the grid. The regular finer grid can be used in real-time with simple interpolation.

While the functions which represent the device are usually well behaved and smoothly varying, the truncation of the data can cause a problem with interpolations near the edge of the gamut. An approach to solving this problem is to extrapolate the data outside the gamut using bandlimited or linear extrapolation methods. The extrapolated points along with the measured data are used in a single interpolation algorithm over the entire gamut of the device. The results of this method are comparable to other interpolation methods but it is simpler to implement. It has the additional advantage of allowing physical constraints, such as bandlimits, to be easily imposed.

1 INTRODUCTION

One way to characterize a color output device is with a three-dimensional look-up table. The look-up table maps the set of color specification values $\{\mathbf{t}_i\}$ to a set of control values $\{\mathbf{c}_i\}$ of the output device. For this work color is specified by CIE $L^*a^*b^*$ values. The control values $\{\mathbf{c}_i\}$ determine the concentration of the colorants. For a $P \times P \times P$ table P^3 measurements would be required. It is clear that P must be small relative to the number of distinct control values. 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. Given a set of control values $\{\mathbf{c}\}$ on a regular grid and the corresponding set of CIE $L^*a^*b^*$ values $\{\mathbf{t}\}$, find the $\{\mathbf{c}_g\}$ for different $\{\mathbf{t}_g\}$ 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 the calibration of nonlinear printing devices. For this research, we use the proposed color data interpolation scheme to calibrate a Kodak XL7700 thermal dye transfer printer. To determine the appropriate color interpolation scheme, a forward model for the printer is defined and interpolation schemes are evaluated using simulated data. In particular, we realize that while the functions which represent the device are well behaved and smoothly varying, the truncation of the data can cause a problem with interpolations near the edge of the gamut. An approach to solving this problem is to extrapolate the data outside the gamut and use these points in a single interpolation algorithm over the entire gamut of the device. The results using these extrapolated points are discussed in this paper.

2 THE FORWARD MODEL

A printing process is subtractive in nature. Each of the cyan, magenta and yellow dyes 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 dye. Each dye is characterized by its density $D(\lambda)$ at any wavelength λ . The density is proportional to the concentration c of the dye. Therefore¹,

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

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

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

The observed spectrum at a particular wavelength λ is given as²

$$g(\lambda) = I(\lambda)T_1(\lambda)T_2(\lambda)T_3(\lambda) \tag{3}$$

where $T_i(\lambda)$ is the transmission of the i^{th} dye at wavelength λ and $I(\lambda)$ is the intensity of the illuminant at wavelength λ . The effect of paper or the transparency has been neglected by assuming it to be uniform. Using equation (2), equation (3) may be rewritten as:

$$g(\lambda) = I(\lambda)10^{-\sum_{i=1}^3 D_{i,max}(\lambda)c_i} \tag{4}$$

Thus, the forward model can be written mathematically as²

$$g = L[10^{-D_{max}c}] \tag{5}$$

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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^T g \tag{6}$$

where A is a $N \times 3$ matrix of the CIE color matching functions. The (x,y,z) values are then converted to the CIE $L^*a^*b^*$ values. 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 ΔE error of 3 is perceptible, but smaller errors are not.

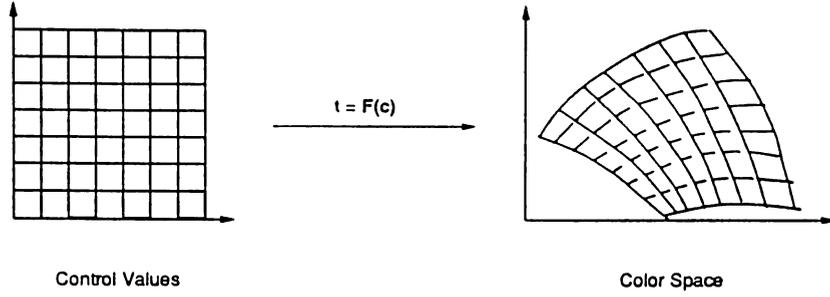


Figure 1: The Forward Model

3 DEVELOPING THE LOOK-UP TABLE

Using the mathematical model, a coarse 8x8x8 look-up table mapping the control values to the output CIE $L^*a^*b^*$ values is generated with the forward model of equation (5). This forward model is used for testing only and to help us predict how well each of the interpolation functions will perform for the ideal model. We have

$$t = F(c) \quad (7)$$

where t is a 3-vector of the CIE $L^*a^*b^*$ values and c is a 3-vector of the control values. The forward model is shown in Figure 1.

A finer regular grid is constructed in the CIE $L^*a^*b^*$ space. For a certain CIE $L^*a^*b^*$ vector t_g on this finer regular grid the corresponding control vector c_g is obtained. Iterative techniques with interpolation are used to estimate the control vector corresponding to every vector t_g on the finer grid. For a given CIE $L^*a^*b^*$ vector t_g , an initial estimate 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. Newton's method⁴ in three dimensions is used to obtain a new estimate of the control value. This procedure is equivalent to finding a root of the equation

$$f(c) = t_g \text{ or } f(c) - t_g = 0 \quad (8)$$

where the function f refers to the interpolating function that is used. This iteration is carried out until equation (8) 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 as⁵:

$$S_b(x) = \begin{cases} \frac{1}{2}(x + \frac{3}{2})^2 & -\frac{3}{2} \leq x \leq -\frac{1}{2} \\ \frac{3}{4} - x^2 & -\frac{1}{2} \leq x \leq \frac{1}{2} \\ \frac{1}{2}(x - \frac{3}{2})^2 & \frac{1}{2} \leq x \leq \frac{3}{2} \end{cases} \quad (9)$$

A cubic B-spline is defined as⁵:

$$S_c(x) = \begin{cases} \frac{2}{3} + \frac{1}{2}|x|^2 - x^3 & 0 \leq |x| \leq 1 \\ \frac{1}{6}(2 - |x|^3) & 1 \leq |x| \leq 2 \end{cases} \quad (10)$$

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

$$t_{g,l}(c_1, c_2, c_3) = \sum_k \sum_j \sum_i w_l(i, j, k) S(c_1 - i) S(c_2 - j) S(c_3 - k) \quad l = 1, 2, 3 \quad (11)$$

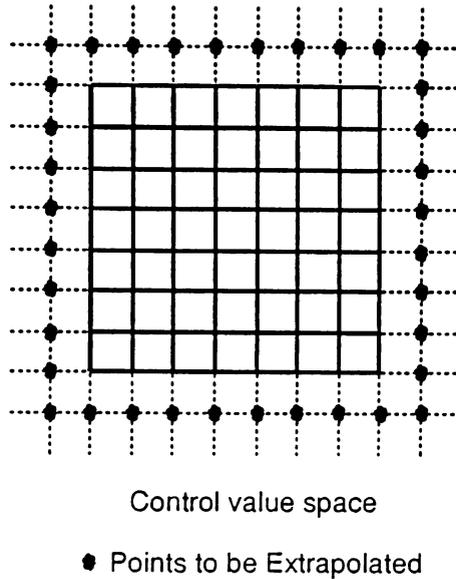


Figure 2: Extrapolation in the Control Value Space

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. In order to determine the weights, the interpolated values are set to the measured values on the coarse grid.

These interpolation schemes assume the presence of a sufficient number of data points close to the interpolated point. However, the truncation of the data can cause a problem with interpolations near the boundary of the printer gamut. In order to solve this problem, we extrapolate the data set for points outside the gamut (Figure 2). These extrapolated points are then used in an interpolation algorithm which does not require treating the boundary as a special case. Two types of extrapolation schemes were proposed (i) simple linear extrapolation and (ii) band-limited extrapolation. The latter scheme was proposed since the mapping from the CMY colorants to the CIE $L^*a^*b^*$ space is usually well-behaved and smoothly varying.

Both the extrapolation schemes were implemented in a one-dimension problem to compare the relative errors that occur in the interpolation of points close to the boundary. A series of lines were created by keeping two colorants constant and varying the third. The true L^* values corresponding to these control values are compared with the original L^* values to calculate the error. Figure 3 compares typical errors due to both the schemes. It is observed that the linear extrapolation gives better results for all simulations.

Linear extrapolation is more attractive than the band-limited extrapolation scheme since it is much easier to implement and takes less time. Band-limited extrapolation has the advantage of allowing physical constraints to be imposed. However, since an iterative algorithm⁶ is used, it has the disadvantage of taking a long time and not giving significantly different results obtained by linear extrapolation.

The points obtained by linear extrapolation are then used in the mathematical model to create a finer, regular LUT in the CIE $L^*a^*b^*$ space. The bell and the spline interpolation functions perform almost identically. As a result, the bell function, which showed a smaller average ΔE error, was used in the actual printer calibration (Figure 4 and Figure 5).

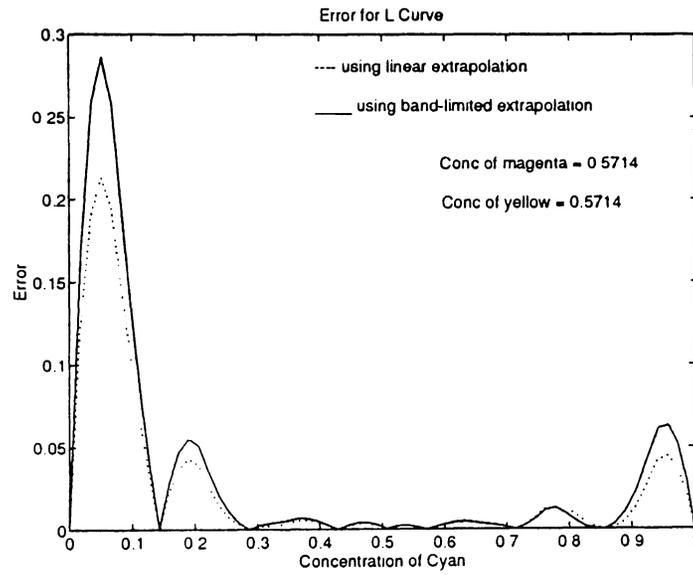


Figure 3: Errors due to Extrapolation

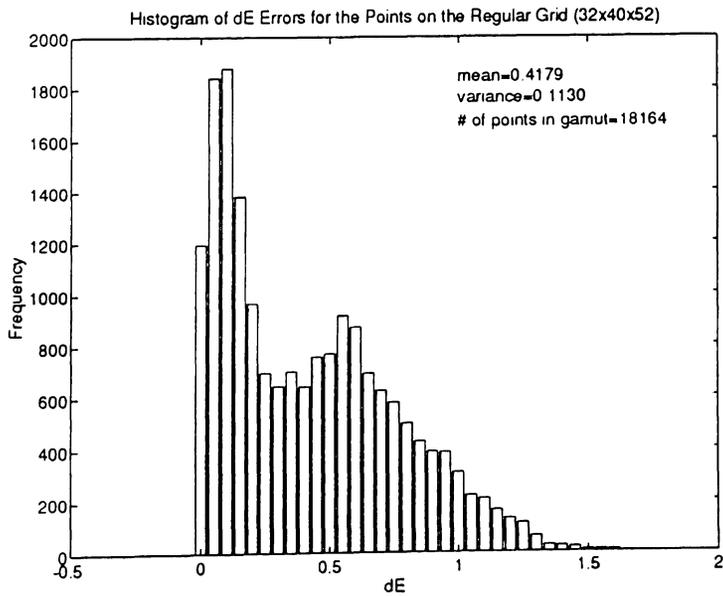


Figure 4: ΔE Errors Using Bell Interpolation

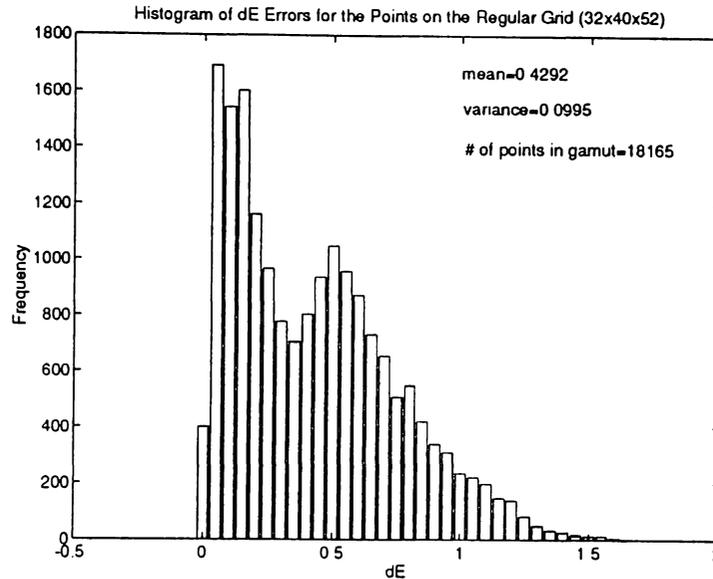


Figure 5: ΔE Errors Using Spline Interpolation

4 THE PRINTER CALIBRATION

A calibration chart was generated from the KODAK XL7700 by dividing the control value space into eight equi-spaced samples in each dimension to generate 512 color patches. A GRETAG SPM-50 spectrophotometer was used to measure the CIE $L^*a^*b^*$ values of these color patches. This set of CIE $L^*a^*b^*$ values forms the known coarse data set.

It was noticed that the ΔE errors between two color patches corresponding to identical control values was significant when they were printed in different regions of the printer paper. This indicated that the printer needed to be calibrated for each of its heads. Another observation was that there was a significant difference between the ΔE errors between two color patches corresponding to identical control values when they were printed on different sheets. This meant that the coarse LUT had to be created out of an average over several printouts. The standard deviation over different sheets ranged from $1.06\Delta E$ to $1.93\Delta E$. This provides a lower bound on the accuracy of the calibration.

Making use of the above observations, the printer was calibrated for only one of its heads. The $8 \times 8 \times 8$ LUT was measured and averaged over several printouts. As mentioned earlier, extrapolation of points at the boundary of the printer gamut was used in the interpolation routine to create the finer, regular LUT in the CIE $L^*a^*b^*$ space.

Using the bell interpolation function, a finer regular grid was created in the CIE $L^*a^*b^*$ space. Trilinear interpolation is used to find control values not on the finer grid. To measure the average interpolation error that was obtained, a test pattern of 343 points (corresponding to all combinations of the three dyes by varying their concentrations from 0.2 to 0.8 in steps on 0.1) were used. The testing procedure is described schematically in Figure 6. The test pattern whose control values are known is printed and its CIE $L^*a^*b^*$ values are measured by the spectrophotometer. These CIE $L^*a^*b^*$ values are then used in the interpolation routine to obtain the corresponding estimate of the control values. These control values are then used to generate a new print. The CIE $L^*a^*b^*$ values of the color patches on this sheet are then compared with CIE $L^*a^*b^*$ values of the test chart. The ΔE errors are discussed in the next section. The errors here indicate a lower bound on the errors

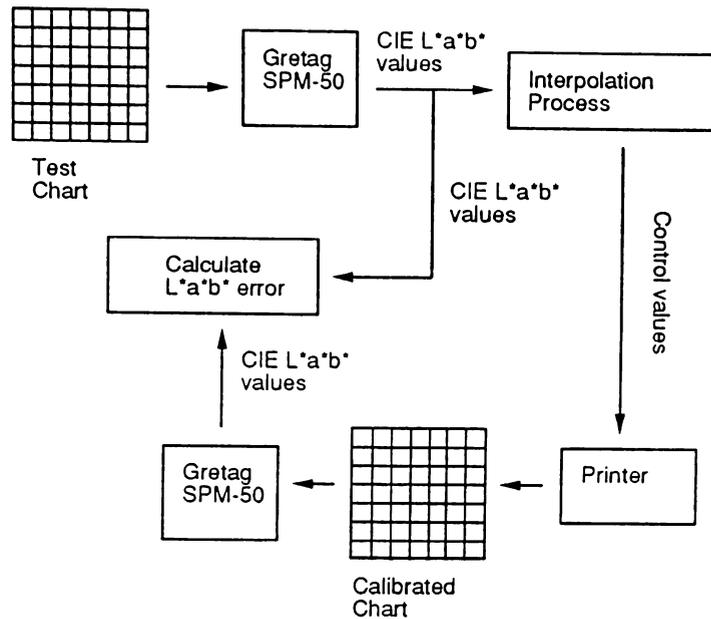


Figure 6: Testing the Interpolation Scheme

obtained by the calibrated printer.

A simple algorithm was used to divide the CIE $L^*a^*b^*$ space into an equi-spaced grid. A minimum grid size of 32 was used in each direction. The spacing in the grid was calculated in the following manner:

1. divide the range in each of the L^* , a^* and b^* directions (obtained from the 8x8x8 LUT) by the minimum grid size i.e. form the spacings $(L_{max}^* - L_{min}^*)/31$, $(a_{max}^* - a_{min}^*)/31$ and $(b_{max}^* - b_{min}^*)/31$.
2. choose the minimum of the spacings formed above as the spacing between two points on the finer grid in the CIE $L^*a^*b^*$ space.

Once this LUT (mapping points on a finer, regular grid in the CIE $L^*a^*b^*$ space to the control value space) is created, we can use it with simple trilinear interpolation to estimate the control values for any CIE $L^*a^*b^*$ point in the printer gamut. If the mapping from the control value space to the CIE $L^*a^*b^*$ space is assumed to be smoothly varying, and if the size of the uniform grid in CIE $L^*a^*b^*$ is sufficiently fine then we can assume that the contribution to the error due to the trilinear interpolation will not be large. Since both these conditions are satisfied in our case, we hope to observe that trilinear interpolation does not increase the ΔE errors by a large factor.

To test out our hypothesis, we use the same test print consisting of 343 equi-spaced samples and measure their CIE $L^*a^*b^*$ values. Using trilinear interpolation over the fine, regular grid that we have created we estimate the control values for each of the test points. These control values are then used to print a new chart whose CIE $L^*a^*b^*$ values are measured again. The CIE $L^*a^*b^*$ values of the color patches on this sheet are compared with the CIE $L^*a^*b^*$ values of the color patches of the test sheet and the ΔE errors are calculated. These results are discussed in the next section. The procedure is described in Figure 6. The interpolation procedure now, refers to trilinear interpolation.

While carrying out the trilinear interpolation, points that fell in a cube in which at least one of the corner points on the cube was not in the gamut of the printer, were neglected in the error computation since they could

here →

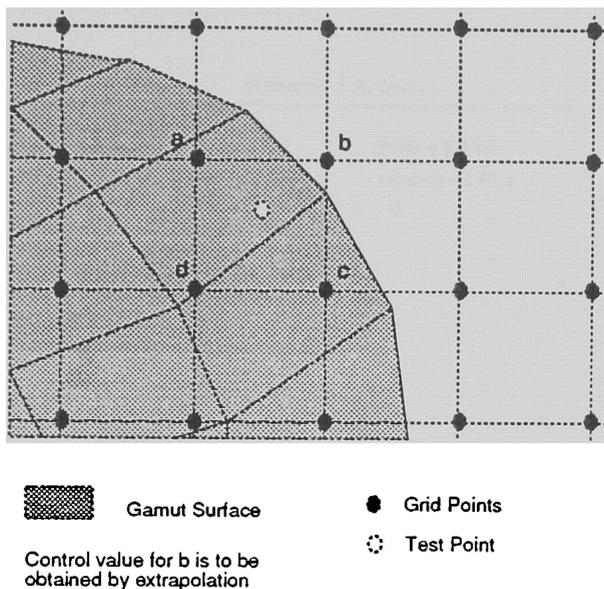


Figure 7: Extrapolation to Cover Entire Gamut

not be interpolated with the grid that we have (Figure 7). However, this is a serious problem since we would like to make sure that we can print all points lying in the printer gamut. To overcome this problem, points on such cubes which are outside the printer gamut should be assigned control values (outside the printer range, obviously) by extrapolation. This problem will be looked at in the near future and will be an additional source of error for points close to the boundary of the printer gamut.

5 EXPERIMENTAL RESULTS

The ΔE errors due to the interpolation routine used to create the fine, regular grid in the CIE $L^*a^*b^*$ space are calculated and the relevant statistics are tabulated in Table 1. Figure 8 shows a histogram of these ΔE errors.

Table 1: ΔE Errors Due to Interpolation

Statistics	Bell Function
Max ΔE	5.5391
Average ΔE	2.0452
Variance ΔE	0.5776

Using the same test pattern as the one used to measure the error in creating the grid, we measure the average error due to trilinear interpolation. We can predict that the error that is added by the trilinear interpolation

5 EXPERIMENTAL RESULTS

The ΔE errors due to the interpolation routine used to create the fine, regular grid in the CIE $L^*a^*b^*$ space are calculated and the relevant statistics are tabulated in Table 1. Figure 8 shows a histogram of these ΔE errors.

Table 1: ΔE Errors Due to Interpolation

Statistics	Bell Function
Max ΔE	5.5391

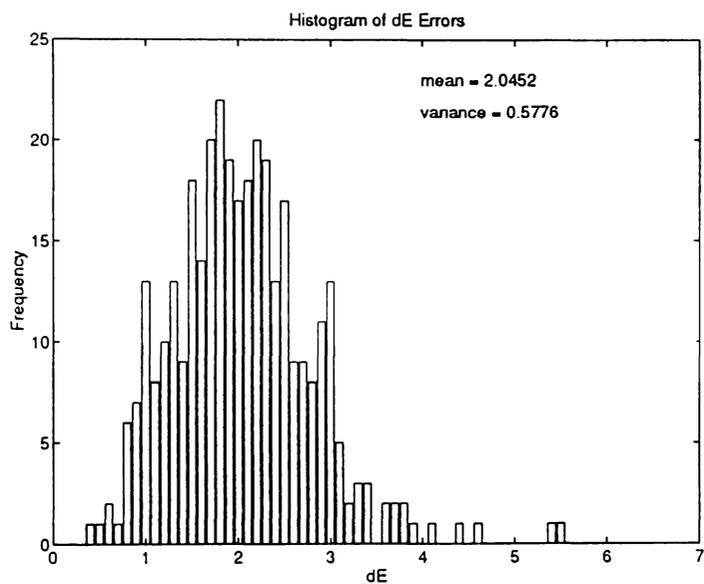


Figure 8: ΔE Errors Due to Interpolation

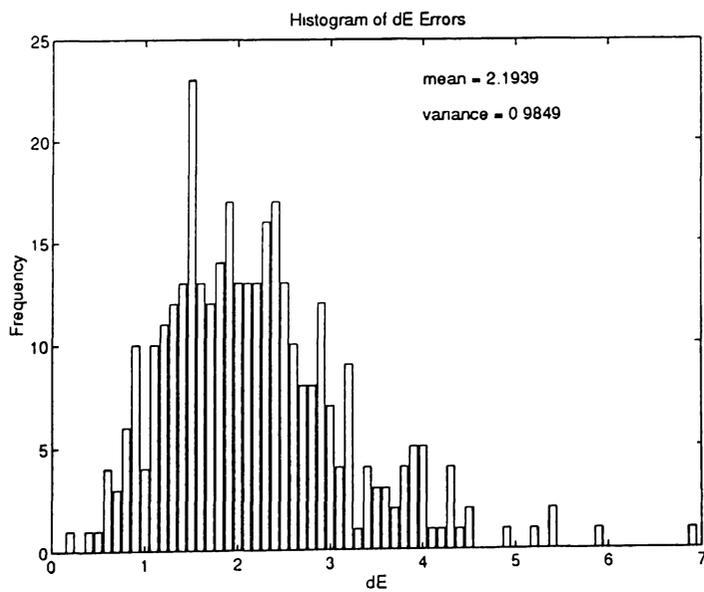


Figure 9: ΔE Errors Due to Trilinear Interpolation

Table 2: ΔE Errors Due to Trilinear Interpolation

Statistics	Bell Function
Max ΔE	6.9392
Average ΔE	2.1939
Variance ΔE	0.9849

the ΔE errors increased almost two-fold. The histogram of the ΔE errors is plotted in Figure 9. The standard deviation of the printer over different sheets was about $1.64\Delta E$ which is of the same order as the average ΔE of 2.1939 that was obtained after trilinear interpolation.

6 CONCLUSION

The printer was calibrated for a single print head. The extrapolated points were used along with the measured values to reduce the interpolation errors for points close to the boundary of the printer gamut. The look-up table method with trilinear interpolation worked well and showed that little could be gained by using more precise interpolation methods for all points. The average ΔE errors due to the interpolation methods are of the same order as the variability of the printer. Therefore, no significant improvement in the average ΔE errors may be expected by changing the interpolation scheme. The interpolation functions that were used seem to be good choices for the calibration of the KODAK XL7700 thermal dye transfer printer.

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, 12-14Feb. 1990.
- [3] Wyszecki, 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.
- [6] Huang, T.S., *Advances in Computer Vision and Image Processing*, Vol 1, Jai Press Inc, 1984.