Abstract

KAKODKAR, ATISH PANDURANG. Interpolation and Extrapolation in Color Systems. (Under the direction of Prof. Sarah A. Rajala and Prof. H. Joel Trussell.)

With the increased use of color in desktop publishing applications has come a desire for greater control of the color quality. Colorimetric reproduction requires calibrated color output devices. One approach to color calibration is to characterize a color output device with a three-dimensional look-up table. The look-up table maps the color specification values to the control values of the color output device.

This thesis looks at the problem of obtaining the look-up table. The problem can be posed in the following manner. Given a set of control values $\{\mathbf{c}_i\}$ on a regular grid and the corresponding set of color specification values $\{\mathbf{t}_i\}$ obtained from data collection, find the $\{\mathbf{c}_g\}$ for different $\{\mathbf{t}_g\}$ on a grid in the color specification space. This grid should be fine enough so that simple interpolation is adequate to obtain control values for color specification values that are not in the table. The grid is obtained from a relatively sparse data set with an appropriately defined interpolation scheme. This interpolation scheme can be very complex since it is used only once to compute the grid. This regular finer grid can be used in real-time to obtain the control value for any color specification value located inside a given color gamut. Two interpolation schemes are evaluated in this thesis.

One of the problems with constructing the fine grid of values for the LUT is the determination of values that are near the edge of the gamut of the output device.

While the functions which represent the device are usually well behaved and smoothly varying, the truncation of the data can cause a problem with interpolation methods. An approach to solving the truncation problem is to extrapolate the data outside the gamut. The extrapolated points are then used in the interpolation to estimate the values for the fine grid. The use of extrapolated values permits the use of a single interpolation algorithm over the entire gamut of the device, rather than using a modified algorithm in regions near the edge of the gamut. The results of this method are comparable to other interpolation methods but it is simpler to implement. Three different extrapolation schemes are evaluated in this thesis and the best one is selected to obtain the look-up table.

This thesis discusses the calibration of a thermal dye transfer color printer. Calibration is carried out initially for a mathematical model. The behaviour of this model is compared to the behaviour of the printer, both with and without the presence of noise. This can be used to determine whether the model is a good representation of the printer and can be used to predict the performance of the printer.

INTERPOLATION AND EXTRAPOLATION IN COLOR SYSTEMS

by

Atish P. Kakodkar

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Department of Electrical and Computer Engineering

Raleigh
May 1994

Approved By:

H.J. Trussell

G.L. Bilbro

S.A. Rajala
Chairman of Advisory Committee

For my parents and my brother,
whose love, support and encouragement
have made this possible.

Biography

Atish P. Kakodkar, was born in Panaji, India on April 19, 1970. He completed the Indian Certificate for Secondary Education (equivalent to the 10th Grade) in 1986, at Mount St. Mary's School, New Delhi, India. He went on to complete his Higher Secondary School Examination (equivalent to the 12th grade) in 1988 at Dhempe College of Arts & Science, Panaji, India. Following this, he continued his studies at the Indian Institute of Technology, Madras, India, where he was awarded a Bachelor of Technology in the field of Electronics & Electrical Communication Engineering in 1992. He was a Research Assistant with the Center for Communications and Signal Processing at North Carolina State University from August 1992 to June 1994. He is also a recipient of the National Talent Search Examination (NTSE) Scholarship awarded by the Government of India to the most promising students of science in the country.

Acknowledgements

I would like to thank Professor Sarah A. Rajala for acting as my advisor and giving me the opportunity to pursue a graduate degree at NCSU. I would also like to immensely thank Professor H. Joel Trussell for his guidance and help throughout the course of this work. Thanks must also be given to Dr. Griff L. Bilbro for being a member of my committee.

To my colleagues, Dr. Poorvi Vora, Dr. Mike Vrhel and Gaurav Sharma, I owe my gratitude for the exchange of ideas and information which were of great help. A special mention must be made of John Coffie who not only helped me take some of the measurements for my research work, but also proved to be a supportive and helpful friend. I would also like to thank my other friends and colleagues - Valerie La, Ahmet Akyamac, Manish Kulkarni and Christophe Grard.

Lastly, I would like to thank some of my friends who in their own special way made all the work worthwhile - Mary Easo, for being such a wonderful person and an understanding friend, Anadi Srivastava, whose immense knowledge never ceases to amaze me, Jayashri Deshmukh, for giving me a different perspective to life, Paritosh (Lapi) Dikshit, who kept me smiling, and Rajika Bhandari, for teaching me the art of stretching my dollar.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Problem Description

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 is the characterization and calibration of color output devices, for example printers or electronic displays such as a CRT. In order to make accurate judgements about the images they produce, such devices must be calibrated. To do this a relationship needs to be defined between a set of output values and the control values of the output device. There are a number of factors which impact the solution of this problem. They are:

- The model of the output device and its limitations
- The type and location of the measurement data
- The variability in the measurement
- The variability in the output device under normal operating conditions
- The interpolation and extrapolation methods used
- The perceptibility of errors.

A simple method of calibrating an output device would be to characterize the device with a three-dimensional look-up table. The look-up table maps a set of color specification values $\{\mathbf{t}_i\}$ to a set of control values $\{\mathbf{c}_i\}$ of the output device. The look-up table could be built directly, unfortunately this would require an excessive number of measurements. 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. If the grid is fine enough, i.e. we have a look-up table for a large number of points, then we can compute control values for different color specification values not on this grid by simple interpolation schemes such as trilinear interpolation. If the look-up table is generated for a small number of points then we could interpolate for color specification points not on the grid by using complex interpolation schemes. 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 color space using complex interpolation. The finer grid can then be used in real-time using simple interpolation.

The problem can be posed in the following manner. Using vector notation, let $\mathbf{c} = [c_1, c_2, c_3]$ be the three-dimensional vector of control values which map on to the color specification values $\mathbf{t} = [t_1, t_2, t_3]$. The functional form of the output device is then given by:

$$\mathbf{t} = \mathbf{F}(\mathbf{c}) \tag{1.1}$$

where the purpose of calibration is to define an inverse mapping from the color specification values to control values. Although the function $\mathbf{F}(\cdot)$ has no closed form, it can be approximated by a closed form, e.g. polynomial. The forward mapping is then defined by interpolation from a table of values. Likewise, the inverse mapping is defined by tables.

The problem of obtaining the look-up table can be posed in the following manner. For a set of control values $\{\mathbf{c}_i\}$ on a coarse, regular grid of size $N \times N \times N$ in the control value space, we obtain the corresponding set of color specification values $\{\mathbf{t}_i\}$

by measuring the output produced by the control values. Using this measured data, we create a finer, regular look-up table of size $M \times M \times M$ $(M > N)$ in the color space, that is we find the control values, $\{\mathbf{c}_g\}$, for different color specification values, $\{\mathbf{t}_g\}$, on this finer, regular grid. This grid in the color space should be fine enough so that simple interpolation is adequate to obtain control values for color specification values that are not in the table. 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 trilinear interpolation to obtain the control value for any color value located inside the color gamut of the device.

One of the problems with constructing the fine grid of values for the LUT is the determination of values near the edge of the color gamut of the output device. While the functions which represent the device are usually well behaved and smoothly varying, the truncation of the data can cause a problem near the edge of the gamut. Special algorithms may be developed to take care of interpolation problems at the edge of the gamut. However, another approach to solving the truncation problem is to extrapolate the data outside the gamut. These extrapolated points can then be used in the interpolation. The use of extrapolated values permits the use of a single interpolation algorithm over the entire gamut of the device, rather than using a modified algorithm in regions near the edge of the gamut (Figure 1.1).

A background in color science is important to understanding this work. The next section discusses the basic concepts of color matching, a vector space approach to color imagery and the mathematical basis for some subjective color phenomena. The last section discusses uniform color spaces introduced by the C.I.E. and the perceptibility of errors between two colors in these spaces.

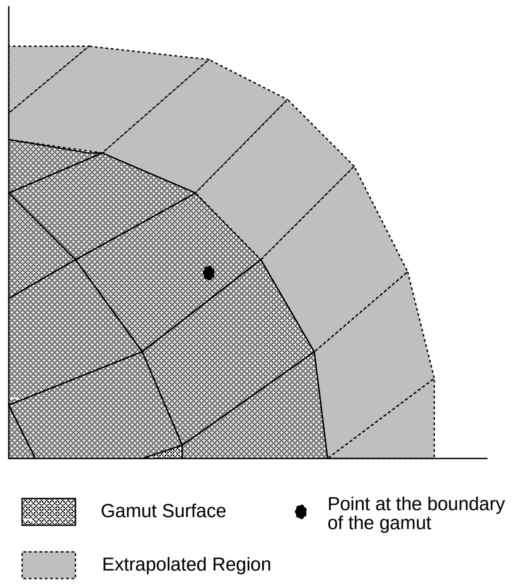

Figure 1.1: Extrapolation at the Boundary of the Gamut

1.2 Color Background

Color perception in humans is a process which has both a physiological and a psychological component. The physiological component involves the sensing of the light signal by sensors (rods and cones) in the eye. The responses of these sensors reach the brain where the psychological component of color perception takes place. Color perception is due to the three types of cones in the human eye which are the color sensitive receptors. The receptor responses are functions of the incident light and can be denoted by β_λ, γ_λ, and ρ_λ. These functions have maxima in the blue, green, and red regions of the spectrum respectively.

1.2.1 Color Matching

All colors can be matched by a combination of three color primaries, but the primaries may have to be changed in order to match certain colors [7]. Primary colors are colors which are independent of each other in the sense that no primary color can be visually equivalent to a mixture of the remaining two primary colors. Matching a color means that the additive combination of particular amounts of a red, a green, and a blue light appears the same to the viewer as a given color. The reason for this is due to the different color sensitive receptors in the human eye.

In a color matching experiment, the subject is shown the color to be matched on one half of the visual field, and a combination of the three primaries on the other. The observer is allowed to adjust the intensities of the primaries so as to obtain a color match with the test stimulus, see Figure 1.2. Usually, the color to be matched has its units normalized to unity in whatever units are being used. It is possible that the units of one of the primaries may be negative which implies that the particular primary with a negative intensity should be added to the test stimulus for a color match with the other two primaries. The color, C, can be a pure spectral color containing only a single wavelength of light, or it can be a combination of many different wavelengths. The colors P_1, P_2 and P_3 are the primaries. The coefficients of these primaries are known as the tristimulus values of the color C relative to the the primaries used. If the tristimulus values are allowed to take on negative values then all colors can be matched. By performing a color match for all monochromatic colors, three color matching curves can be created.

1.2.2 Vector Space Methods for Color Representation

If a continuous spectrum is sampled at a sufficient number of points, a vector space formulation of the color matching problem can be obtained [15]. Let the N-dimensional sampled spectrum be denoted by $\mathbf{f} = [f(1)\ f(2) \ldots f(N)]$. As discussed earlier, the human visual system has three cone sensors. It can be represented by a matrix, $\mathbf{S} = [\mathbf{s}_1\ \mathbf{s}_2\ \mathbf{s}_3]$, containing a set of three N-dimensional vectors. The response of the eye to the spectrum $\mathbf{f}$ is then given by:

$$\mathbf{c} = \mathbf{S}^T\mathbf{f} \tag{1.2}$$

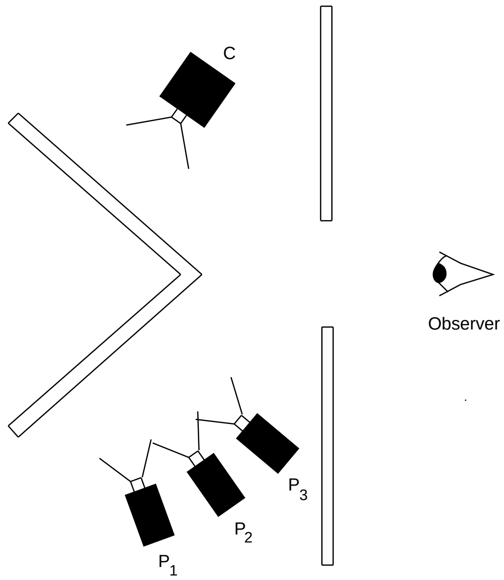

Figure 1.2: Color Matching Experiment

where $\mathbf{c}$ is a three-dimensional vector.

Let $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]$ denote the $N \times 3$ matrix of primaries where $\mathbf{p}_1, \mathbf{p}_2$ and $\mathbf{p}_3$ are three linearly independent, N-dimensional primaries. Let the monochromatic colors be denoted by $\mathbf{e}_i, \quad i = 1, \ldots, N$, where $\mathbf{e}_i$ has a one in the i^{th} component and zeros in all others. We say that the stimulus $\mathbf{e}_i$ is matched by $[m_1(i)\ m_2(i)\ m_3(i)]$ units of primaries $[\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]$ respectively if

$$\mathbf{S}^T\mathbf{e}_i = \mathbf{S}^T(m_1(i)\mathbf{p}_1 + m_2(i)\mathbf{p}_2 + m_3(i)\mathbf{p}_3)$$

$$\Rightarrow \mathbf{S}^T\mathbf{e}_i = \mathbf{S}^T\mathbf{P}\mathbf{m}_i \tag{1.3}$$

where $\mathbf{m}_i = [m_1(i)\ m_2(i)\ m_3(i)]^T$ is the three-dimensional vector of the gains of the primaries.

Matching all the spectral colors gives us the equation

$$\mathbf{S}^T\mathbf{I} = \mathbf{S}^T\mathbf{P}\mathbf{M}^T \tag{1.4}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. The $N \times 3$ matrix $\mathbf{M}$ is known as the color matching matrix. The columns of $\mathbf{S}$ are independent and the columns of $\mathbf{P}$ are also chosen to be independent. Both of them are rank three matrices. Therefore the inverse $(\mathbf{S}^T\mathbf{P})^{-1}$ exists. Equation (1.4) can be solved for $\mathbf{M}$ giving

$$\mathbf{M}^T = (\mathbf{S}^T\mathbf{P})^{-1}\mathbf{S}^T \tag{1.5}$$

Post-multiplying both sides by $\mathbf{P}$ gives us

$$\mathbf{M}^T\mathbf{P} = (\mathbf{S}^T\mathbf{P})^{-1}\mathbf{S}^T\mathbf{P} = \mathbf{I} \tag{1.6}$$

where $\mathbf{I}$ is a 3×3 identity matrix.

Using the color matching matrix, $\mathbf{M}$, one can calculate the tristimulus values, $\mathbf{t}_P$, of an arbitrary spectrum according to

$$\mathbf{t}_P = \mathbf{M}^T\mathbf{f} \tag{1.7}$$

where the subscript $\mathbf{P}$ in $\mathbf{t}_P$ denotes the use of the primaries which form the columns of $\mathbf{P}$. The N-dimensional spectrum, $\mathbf{f}$, is projected onto a three-dimensional subspace that defines a particular color space. Equation (1.5) relates this subspace to the human visual subspace. This reduction in dimension can make two different spectra appear the same to an observer. These spectra are called metamers and the relation is defined mathematically as

$$\mathbf{M}^T\mathbf{f}_1 = \mathbf{M}^T\mathbf{f}_2 \tag{1.8}$$

Mathematically, it is not necessary that the primaries $\mathbf{P}$ be realizable, i.e. all elements of $\mathbf{P}$ are non-negative. We can choose any set of primaries and obtain the corresponding matching functions $\mathbf{M}$. Let $\mathbf{Q} = [\mathbf{q}_1\ \mathbf{q}_2\ \mathbf{q}_3]$ be another set of primaries. If $\mathbf{N}$ is the corresponding set of color matching functions then from Equation (1.4) we have

$$\begin{aligned} \mathbf{S}^T &= \mathbf{S}^T\mathbf{P}\mathbf{M}^T \\ \mathbf{S}^T &= \mathbf{S}^T\mathbf{Q}\mathbf{N}^T \\ \Rightarrow \mathbf{S}^T\mathbf{P}\mathbf{M}^T &= \mathbf{S}^T\mathbf{Q}\mathbf{N}^T \\ \Rightarrow \mathbf{M}^T &= (\mathbf{S}^T\mathbf{P})^{-1}\mathbf{S}^T\mathbf{Q}\mathbf{N}^T \end{aligned} \tag{1.9}$$

Post-multiplying both sides of the equation by $\mathbf{Q}$ and setting $\mathbf{N}^T\mathbf{Q}$ to $\mathbf{I}$ (from equation 1.6), we get

$$(\mathbf{S}^T\mathbf{P})^{-1}\mathbf{S}^T\mathbf{Q} = \mathbf{M}^T\mathbf{Q} \tag{1.10}$$

Substituting back into Equation (1.9), we get

$$\mathbf{M}^T = \mathbf{M}^T\mathbf{Q}\mathbf{N}^T \tag{1.11}$$

or

$$\mathbf{N}^T = (\mathbf{M}^T\mathbf{Q})^{-1}\mathbf{M}^T \tag{1.12}$$

This transformation also gives us the mapping between the tristimulus values in the two different color spaces

$$\mathbf{t}_Q = (\mathbf{M}^T\mathbf{Q})^{-1}\mathbf{t}_P \tag{1.13}$$

1.2.3 C.I.E. XYZ Space

One standard set of primaries is defined by the C.I.E. red, green, and blue primaries consisting of single wavelengths at 700, 546.1 and 435.8 nm. It is seen in Figure 1.3 that the color matching function corresponding to the red primary has negative portions in the blue end of the spectrum. This means that the light at these wavelengths cannot be represented by an additive mixture of these three primaries. It is desirable to have color matching functions which are realizable (i.e. their spectrum contains only positive or zero values) since the tristimulus values can be obtained by optical filters. The primaries however may be non-realizable which means that they contain one or more negative values in their spectrum and are not physically realizable.

The C.I.E. in 1931 used a 3×3 linear transformation to convert the C.I.E. red, green and blue color matching functions to a realizable set of color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$. The transformation from (R, G, B) to (X, Y, Z) is given by:

$$\begin{aligned} X &= 0.49R + 0.31G + 0.20B \\ Y &= 0.1767R + 0.8124G + 0.0163B \\ Z &= 0.00R + 0.01G + 0.99B \end{aligned} \tag{1.14}$$

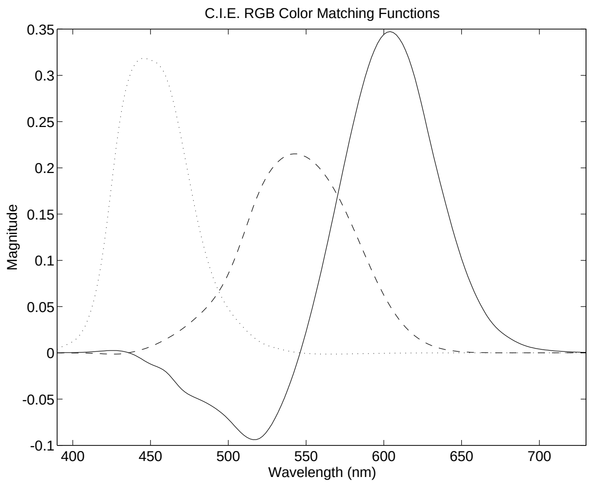

Figure 1.3: C.I.E RGB Color Matching Functions

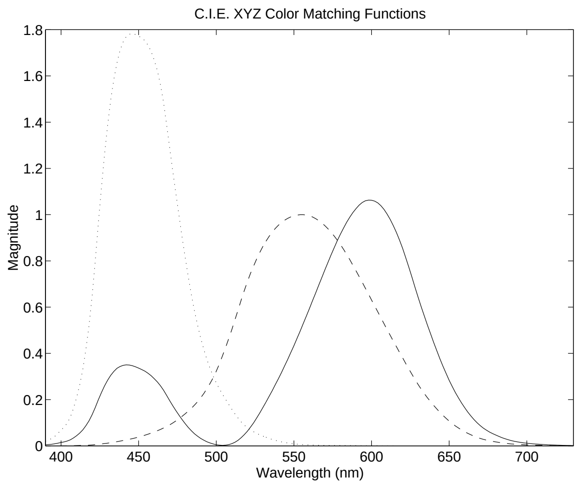

Figure 1.4: C.I.E XYZ Color Matching Functions

This transformation implied a new set of primaries which were not realizable. Figure 1.4 shows the new set of color matching functions which are non-negative for all wavelengths in the visible range. The samples are taken every two nanometers but are connected in the plot. Each of these curves becomes one column of the color matching matrix, denoted by $\mathbf{A}$. If these columns are now denoted by $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ respectively, then the elements of each of the columns are given by the samples at regular intervals in the visible range, i.e. $\mathbf{a}_1 = [a_1(1)\ a_1(2) \ldots a_1(N)]$. Similar sampling is done to obtain $\mathbf{a}_2$ and $\mathbf{a}_3$. The tristimulus values can now be computed in the following manner:

$$\mathbf{t} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{A}^T \mathbf{f} \tag{1.15}$$

where X, Y, Z denote the three-dimensional components of the tristimulus vector in this color space.

1.2.4 Uniform Color Spaces

The CIE colorimetric system includes computational methods designed to aid in the prediction of the magnitude of the perceived color difference between two given object-color stimuli. The determination of a quantity that suitably describes the color difference an observer may perceive between two given color stimuli rests on the ability of the observer to judge the relative magnitude of two color differences possibly perceivable when viewing two pairs of stimuli. The observer's sensitivity varies greatly with the conditions of observation and the kind of stimuli presented. Sizes, shapes, luminances, and relative spectral radiant power distributions of the test stimuli and the stimuli surrounding them are important factors affecting the observer's judgement.

The CIE recommends the use of two approximately uniform color spaces and associated color-difference formulae. The recommendations are all given in terms of the CIE 1931 Standard Colorimetric Observer and Coordinate System [17].

1. **CIE 1976 ($L^*u^*v^*$)-Space and Color-Difference Formulae.** This approximately uniform color space is produced by plotting in rectangular coordinates

the quantities L^*,u^*,v^* defined by:

$$\begin{aligned} L^* &= 116(\frac{Y}{Y_n})^{1/3} - 16 \\ u^* &= 13L^*(u' - u'_n) \\ v^* &= 13L^*(v' - v'_n) \end{aligned} \tag{1.16}$$

with the constraint that $Y/Y_n > 0.01$. If values of Y/Y_n less than 0.01 occur, a somewhat modified procedure is recommended for calculating L^*. For values of Y/Y_n equal to or less than 0.008856, the following L^*_m formula is used:

$$L^*_m = 903.3\frac{Y}{Y_n} \qquad for \ \frac{Y}{Y_n} \leq 0.008856 \tag{1.17}$$

In Equation (1.16), the quantities u',v' and u'_n,v'_n are calculated from:

$$\begin{aligned} u' &= \frac{4X}{X + 15Y + 3Z} \qquad & v' &= \frac{9Y}{X + 15Y + 3Z} \\ u'_n &= \frac{4X_n}{X_n + 15Y_n + 3Z_n} \qquad & v'_n &= \frac{9Y_n}{X_n + 15Y_n + 3Z_n} \end{aligned} \tag{1.18}$$

The tristimulus values X_n, Y_n, Z_n are those of the nominally white object-color stimulus. The total color difference ΔE^*_{uv} between two color stimuli, each given in terms of L^*, u^*, v^* is calculated from

$$\Delta E^*_{uv} = [(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2]^{1/2} \tag{1.19}$$

2. **CIE 1976 ($L^*a^*b^*$)-Space and Color Difference Formula.** The second approximately uniform color space, is produced by plotting in rectangular coordinates the quantities, L^*,a^*,b^* defined by:

$$\begin{aligned} L^* &= 116(\frac{Y}{Y_n})^{1/3} - 16 \\ a^* &= 500[(\frac{X}{X_n})^{1/3} - (\frac{Y}{Y_n})^{1/3}] \\ b^* &= 200[(\frac{Y}{Y_n})^{1/3} - (\frac{Z}{Z_n})^{1/3}] \end{aligned} \tag{1.20}$$

with the constraint that X/X_n, Y/Y_n, $Z/Z_n > 0.01$. in calculating L^*, a^* and b^* values of X/X_n, Y/Y_n, $Z/Z_n < 0.01$ may be included if the normal formulae are replaced by the following modified formulae

$$L^*_m = 903.3\frac{Y}{Y_n} \qquad for\frac{Y}{Y_n} \leq 0.008856$$

and

$$\begin{aligned} a^*_m &= 500[f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})] \\ b^*_m &= 200[f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})] \end{aligned}$$

where

$$\begin{aligned} f(\frac{X}{X_n}) &= (\frac{X}{X_n})^{1/3} && \tfrac{X}{X_n} > 0.008856 \\ f(\frac{X}{X_n}) &= 7.787(\frac{X}{X_n}) + \frac{16}{116} && \tfrac{X}{X_n} \leq 0.008856 \\ f(\frac{Y}{Y_n}) &= (\frac{Y}{Y_n})^{1/3} && \tfrac{Y}{Y_n} > 0.008856 \\ f(\frac{Y}{Y_n}) &= 7.787(\frac{Y}{Y_n}) + \frac{16}{116} && \tfrac{Y}{Y_n} \leq 0.008856 \\ f(\frac{Z}{Z_n}) &= (\frac{Z}{Z_n})^{1/3} && \tfrac{Z}{Z_n} > 0.008856 \\ f(\frac{Z}{Z_n}) &= 7.787(\frac{Z}{Z_n}) + \frac{16}{116} && \tfrac{Z}{Z_n} \leq 0.008856 \end{aligned} \tag{1.21}$$

The tristimulus values X_n, Y_n, Z_n are those of the nominally white object-color stimulus. The total color difference ΔE^*_{ab} between two color stimuli, each given in terms of L^*, a^*, b^* is calculated from

$$\Delta E^*_{ab} = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2} \tag{1.22}$$

The CIE color difference measure ΔE^*_{ab} is often used as a measure of perceptual color difference [4, 17]. The average color tolerance accepted in printing applications has been studied and found to be approximately a ΔE^*_{ab} of six. The standard deviation in the accepted tolerance was 3.63 ΔE^*_{ab} [13].

1.3 Thesis Outline

The thesis is arranged in the following way. The second chapter discusses the mathematical formulation of two different color output systems, a CRT and a color printer. The additive principle is discussed followed by the development of a model for a CRT monitor. This is followed by a discussion of the subtractive principle which is used to develop the model for the color printer. The model limitations for both the CRT and printer models are discussed.

The third chapter discusses the calibration procedure for the simulated printer model developed in the second chapter. The ideal model with all its limitations is used to test the calibration method. The bell function and the cubic B-spline interpolation techniques that are used are discussed in detail. Three different extrapolation techniques, namely, separable linear extrapolation, non-separable linear extrapolation and band-limited extrapolation are discussed. The different interpolation and extrapolation schemes are compared and the best ones are chosen for the actual printer calibration.

The fourth chapter discusses the actual calibration of a KODAK XL7700 thermal dye transfer printer. The entire calibration is carried out using one of each of the interpolation and extrapolation methods discussed in the third chapter. The different sources of error which come up in the calibration procedure are studied and discussed in detail. Errors due to variability of the printer, the measuring device and the interpolation method that is used are all studied in detail. The signal-to-noise ratio

(SNR) of the printer is calculated and a similar noisy situation is simulated for the mathematical model to see how the model compares to the actual printer. The problem of making sure that all points in the printer gamut can be printed using the look-up table is also discussed.

The final chapter contains a summary of the results that were obtained for the printer calibration. The conclusions and the direction of further research are presented at the end.

Chapter 2

Mathematical Models

Calibration of color output devices such as CRTs and color printers involve data interpolation techniques. A mathematical model of these output devices which simulates the devices under certain assumptions helps us to study the behaviour of these output devices. The effect of using different interpolation and extrapolation algorithms in the calibration procedure can be studied by applying them to the calibration of the simulated model. This chapter describes the development of mathematical models for a CRT and a color printer. The basic assumptions made in the course of the model development are listed and their validity explained. The basic principles of additive and subtractive color reproduction are also discussed.

2.1 The Additive Principle

The experimental laws of color matching state that over a wide range of conditions of observation, many color stimuli can be matched completely by additive mixtures of three fixed primary stimuli whose radiant powers have been suitably adjusted. Other color stimuli have to be mixed with one of the primary stimuli before a complete color match with a mixture of the other two primary stimuli can be obtained. For some other sets of primary stimuli, there are certain color stimuli that have to be mixed with two of the three primary stimuli before a color match between this mixture and the third primary stimulus can be obtained. Every color stimulus can be matched in color in one of these ways in terms of three primary stimuli whose radiant power can be adjusted by the observer to suitable levels. This is the additive principle. The choice of the three primary stimuli is not entirely arbitrary. Any set that is such that none of the primary stimuli can be color matched by a mixture of the other two may be used. The term additive mixture as used above represents a color stimulus for which the radiant power in any wavelength interval in any part of the spectrum is equal to the the sum of the powers in the same interval of the constituents of the mixture. The additive primaries are represented by three color stimuli in the red, green and blue regions of the spectrum.

The CRT is an output device which is based on the additive principle. In a CRT the image is produced by what is called a picture tube. A picture tube in a color monitor has three electron guns which produce three electron beams which excite three different phosphors (primaries) on a fluorescent screen in front of the guns. The color of the light output depends on the physical properties of the phosphor. The phosphor screen is capable of emitting red, green and blue light. The intensity of the light output depends on the velocity of the beam, the number of electrons in the beam and the type of phosphor used.

2.2 Mathematical Modelling of a CRT

In a digital graphics system, the video buffer in the monitor drives the three digital-to-analog converters (DACs) that provide the input to the electron guns. The voltages which drive the electron guns are produced by these DACs which in turn convert the code values in the red, green and blue frame buffers to analog voltage levels. To derive a mathematical model for a CRT, it is assumed that the voltage outputs are linearly related to the code values in the respective frame buffers. This has been shown to be a reasonable assumption in [3, 10]. A non-zero voltage is input to the electron guns even at a zero code value, i.e. the black portion of the image has a corresponding non-zero voltage level. For a black and white monitor, which has only a single electron gun the mathematical relationship making use of the above assumption is given by:

$$L = K(n - x_0)^{\gamma} \tag{2.1}$$

where L is the radiance of a black and white phosphor, K is a constant dependent on the units of L, n is the count value in the frame buffer, and x_0 is the cutoff count value for the electron gun.

For a color monitor, the model takes a more complicated form. If we assume that the chromaticity coordinates of the three phosphors remain constant then we get the following equation:

$$\begin{aligned} L_R &= K_{1,R}(R - R_0)^{\gamma_R} + K_{1,G}(G - G_0)^{\gamma_G} + K_{1,B}(B - B_0)^{\gamma_B} \\ L_G &= K_{2,R}(R - R_0)^{\gamma_R} + K_{2,G}(G - G_0)^{\gamma_G} + K_{2,B}(B - B_0)^{\gamma_B} \\ L_B &= K_{3,R}(R - R_0)^{\gamma_R} + K_{3,G}(G - G_0)^{\gamma_G} + K_{3,B}(B - B_0)^{\gamma_B} \end{aligned} \tag{2.2}$$

where L_R, L_G, L_B are the radiance values for red, green and blue phosphors, R, G, B are the red, green and blue count values, R_0, G_0, B_0 are the red, green and blue cutoff count values, and γ_R,γ_G,γ_B are the exponents for thee red, green and blue electron guns. The terms $K_{i,R}, K_{i,G}, K_{i,B},\ i = 1, 2, 3$ are constants. If the condition for gun independence is met, then Equation (2.2) can be further simplified as:

$$\begin{aligned} L_R &= K_{1,R}(R - R_0)^{\gamma_R} \\ L_G &= K_{2,G}(G - G_0)^{\gamma_G} \\ L_B &= K_{3,B}(B - B_0)^{\gamma_B} \end{aligned} \tag{2.3}$$

Equation (2.3) is the fundamental input/output relationship for a color monitor.

2.3 Limitations of the CRT Model

A color monitor has four important properties which are assumed in developing a mathematical model [3, 10]. These assumptions are true to varying degrees. They are:

1. **Temporal Stability:** The monitor maintains its colorimetric parameters over time, so that a calibration is not necessary frequently. In practice, this is true over days but usually monitors should be calibrated about once a month.

2. **Spatial Uniformity:** The monitor has the same colorimetric parameters at different points on its screen, so that a calibration can be done at one place on the screen that will hold for the whole screen. Spatial uniformity is hard to achieve. Sometimes there is as much as a 50 percent variation in the luminance value between the color displayed at the center of the screen and the same color displayed at the corners. This is due to the improper focusing of the electron guns at the corners of the screen and the fact that electrons must travel a longer distance to illuminate the screen corners compared to the center.

3. **Gun Independence:** The colorimetric properties of the monitor are truly additive, so that it is possible to base the calibration scheme on the assumption that the displayed color is the additive mixture of the color produced by each of the three phosphor gun combinations. This property is assumed to hold in most monitors.

4. **Phosphor Constancy:** The chromaticity coordinates of the red, green and blue phosphors remain constant independent of the voltages applied to the electron guns. Mathematically, if the spectral distribution of the phosphor is given by $\phi_1(\lambda)$ at voltage v_1, then the spectral distribution of the phosphor at

voltage v_2 is given by $\phi_2(\lambda) = \alpha\phi_1(\lambda)$ for some constant α. This property is also assumed to hold in most monitors.

2.4 The Subtractive Principle

In additive methods of color reproduction, all colors are produced by the adding together of different proportions of the light from three primary colors, a red, a green, and a blue. Subtractive color reproduction is different in the sense that all the colors are produced by different proportions of three different colorants, cyan, magenta and yellow. It is characterized by the property that color is obtained by removing selected portions of a source spectrum (Figure 2.1).

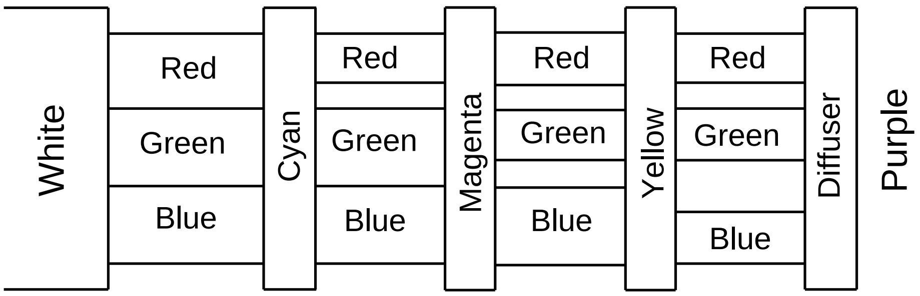

Figure 2.1: Subtractive Model

A printing process is subtractive in nature. Each of the cyan, magenta and yellow colorants removes an amount of its complimentary color. The function of the cyan colorant is to absorb red light, that of the magenta colorant to absorb green light, and that of the yellow colorant to absorb blue light. The transmission curves of each of the three colorants vary over the entire spectrum as shown in Figure 2.2. The amount of a particular color removed is related to the concentration of the colorant. For the yellow colorant it is seen that, for all concentrations, the transmission in the reddish part of the spectrum is high. In the greenish part, the variation of

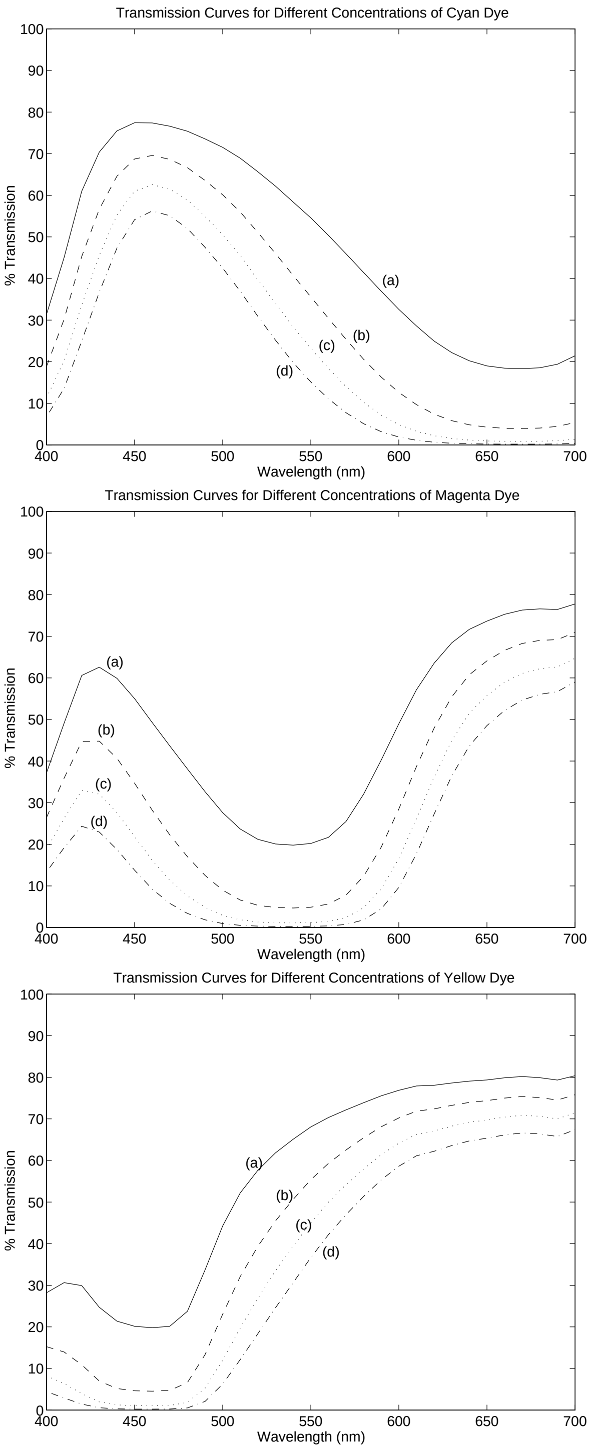

Figure 2.2: Spectral Transmission Curves of Subtractive Colorants for Different Concentrations (a)25% (b)50% (c)75% (d)100%

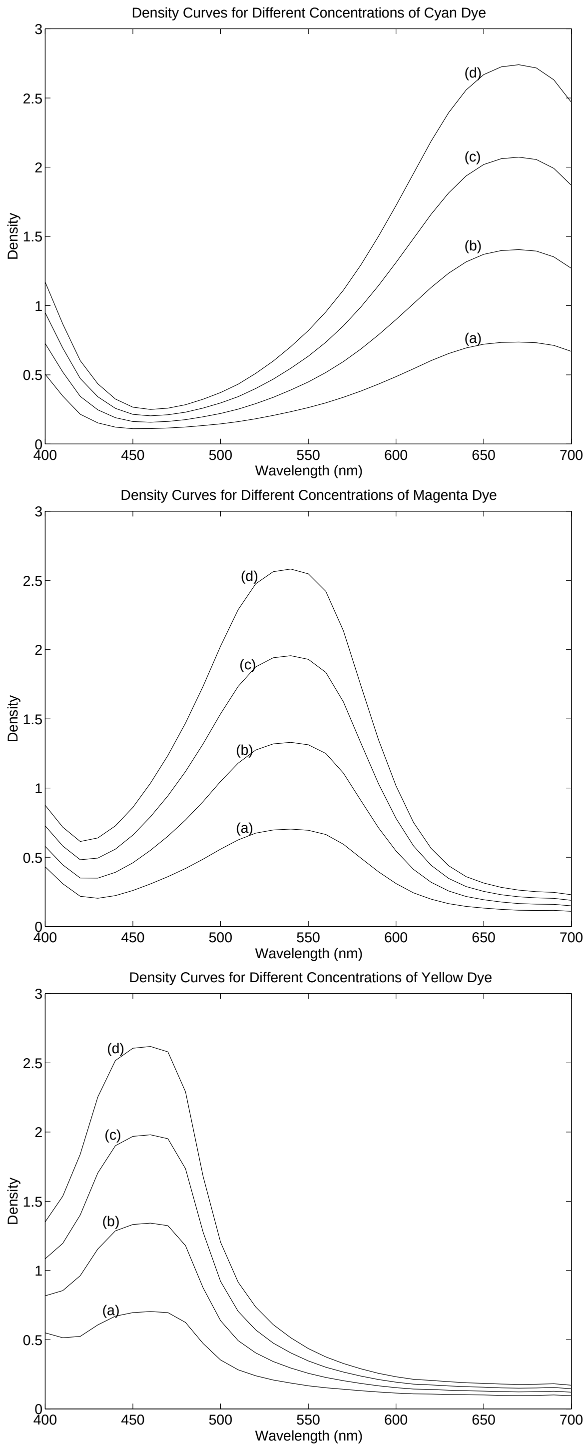

Figure 2.3: Density Curves of Subtractive Colorants for Different Concentrations (a)25% (b)50% (c)75% (d)100%

transmission with concentration is not large but, in the bluish part of the spectrum, the transmission depends very markedly on the concentration of the colorant. This shows that the amount of bluish light reflected from a piece of white paper, viewed in daylight, would be altered by the concentration of a yellow colorant on its surface. Similarly, the main effect of altering the magenta colorant is to vary the transmission in the greenish part of the spectrum, while altering the bluish and reddish parts to a smaller extent. Different concentrations of the cyan colorant alter the transmission in the reddish portion of the spectrum and to a lesser extent the transmissions in the greenish and reddish parts of the spectrum.

2.5 The Forward Model of the Printer

Each of the colorants can be characterized by their optical density spectra, the $N \times 3$ matrix $\mathbf{D}$. The optical density of a colorant is defined as $-\log_{10}(T)$, where T represents the transmittance or reflectance of the colorant. The density of a colorant is proportional to its concentration c. Thus we can write [7]

$$D_i(\lambda) = D_{i,max}(\lambda)\ c_i \qquad i = 1, 2, 3 \tag{2.4}$$

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

$$T_i(\lambda) \ = \ 10^{-D_{i,max}(\lambda)c_i} \qquad i = 1, 2, 3 \tag{2.5}$$

If we look at the density spectra for the different colorants, we see that the curves should be linear in density, i.e. $D_i(c_{i,1} + c_{i,2}) = D_i(c_{i,1}) + D_i(c_{i,2})$ for different concentrations $c_{i,1}$ and $c_{i,2}$ of a particular colorant i. This can be seen in Figure 2.3. The observed spectrum at a particular wavelength λ is given as [16]

$$g(\lambda) \ = \ l(\lambda)T_1(\lambda)T_2(\lambda)T_3(\lambda) \tag{2.6}$$

where $T_i(\lambda)$ is the transmission of the i^{th} colorant at wavelength λ and $l(\lambda)$ is the intensity of the illuminant at wavelength λ. This is the case for the colorants being placed upon film. In the case where paper is used, the paper reflects the light and the light passes through each layer of the colorant twice. Substituting Equation (2.5) into Equation (2.6) we get:

$$g(\lambda) = l(\lambda)10^{-\sum_{i=1}^{3} D_{i,max}(\lambda)c_i} \tag{2.7}$$

Thus, the forward model can be written algebraically as [16]

$$\mathbf{g} = \mathbf{L}[10^{-\mathbf{D}_{max}\mathbf{c}}] \tag{2.8}$$

where $\mathbf{c}$ is the 3-vector representing the concentration of the colorants, $\mathbf{D}_{max}$ is an $N \times 3$ matrix of the densities at the maximum concentration, $\mathbf{L}$ is an $N \times N$ diagonal matrix of the illuminant spectrum and $\mathbf{g}$ is the N-vector representing the radiant spectrum. The concentration values must be between zero and unity. The exponential term is computed componentwise i.e.

$$10^{\mathbf{r}} = [10^{r_1}10^{r_2}.....10^{r_N}]^T \tag{2.9}$$

This model ignores non-linear interactions between colorant layers.

The tristimulus values XYZ are obtained by the equation

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{A}^T\mathbf{g} \tag{2.10}$$

where $\mathbf{A}$ is a $N \times 3$ matrix of the CIE color matching functions.

2.6 Limitations of the Printer Model

The mathematical model for the printer described in the previous section represents an ideal printer which makes a number of basic assumptions. The important properties which this model possesses are as follows:

1. **Temporal Stability:** The amount of colorant that is printed onto the paper (for a particular control vector) remains the same each time it is printed. Therefore the colorimetric parameters do not change over time so that calibration needs to be done only once.

2. **Spatial Uniformity:** The printer behaves identically over different regions of the printer sheet. The amount of colorant laid down on the sheet does not depend on its spatial location on the sheet. In the case with printers which have different printer heads, this implies that each of these heads are identical in nature and there is no variation in the amount of colorant put down due to variation of these heads. This would imply that the calibration can be done for all the heads simultaneously.

3. **Ideal Transmissions and Reflections:** This is one of the most important assumptions of the mathematical model. We assume that the paper reflects all the light which is incident onto its surface without any absorptions. This is followed by ideal absorptions and transmissions by each of the colorants. We assume that there are no first-surface reflections at a dye surface when light is incident on it, see Figure 2.5. We also assume that there are no multiple internal reflections and we model the path that is indicated in Figure 2.5.

4. **No Interaction Between Colorant Layers:** A dye laid down on top of another dye does not adhere as well as it would to paper. This is called dye inhibition which we assume is absent in the model. In a thermal dye transfer printer, the heating for a subsequent dye can melt the dye that has been previously laid down. This is called back-transfer which we assume is absent in this model.

5. **Spectral Characteristics of Dye are not Changed by Interaction:** We assume that placing the dyes on top of one another does does change the

spectral characteristics of each of the individual dyes.

The above properties represent an ideal printer. However, in reality, a printer may behave quite differently. The assumption of having identical printing heads may not be a reasonable assumption as was the case with the printer that was calibrated in the course of this work. Each of the printer heads show a certain amount of variability. This would imply different calibrations i.e. look-up tables, for each of the heads.

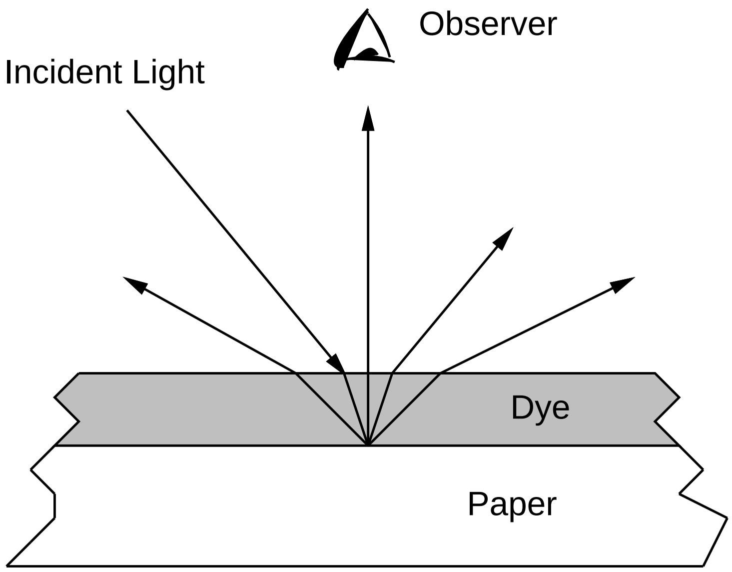

Figure 2.4: Scattering of Light

Different optical effects may take place in a layer of dye placed on paper [18]. The major path of light is shown in Figure 2.4. The incident light passes through the colorant layer, is reflected in all directions by the paper, passes through the colorant layer again and emerges from the surface of the colorant. The observer sees that portion of it which is travelling in the direction of his eye. There are several other factors to be taken into account such as first-surface reflections, multiple internal reflections, and absorption and transmission by the paper. First-surface reflections and multiple internal reflections are shown in Figure 2.5. First-surface reflection is the

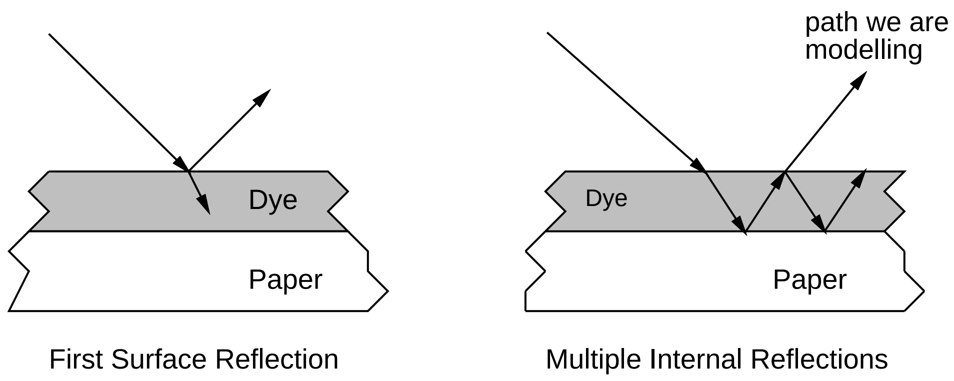

Figure 2.5: Other Paths of Light

reflection of the incident light at the surface of the dye. Multiple internal reflections occur when some of the light reflected by the paper is reflected back toward the paper by the upper surface of the colorant, passing through the colorant to be reflected again by the paper. Eventually it either emerges or is absorbed.

Dye inhibition and back transfer are also problems which occur in printers. These problems give rise to image dependent variations over different regions of the paper onto which an image is printed. These and the other factors described above are the main causes of the variability of any thermal dye-transfer printer.

Chapter 3

Mathematical Formulation

This chapter deals with the detailed description of the calibration procedure performed for a simulated printer. The mathematical model developed in the second chapter is used in the simulation. The calibration of this simulated printer will help us understand and study some of the problems encountered in the actual calibration of a printer which is described in the next chapter. This chapter describes in detail the different interpolation and extrapolation techniques that are used in the calibration. These schemes are then implemented with the mathematical model and the results compared so that the best schemes can be chosen for the actual printer calibration which is discussed in the next chapter.

3.1 Interpolation Technique

Using the mathematical model developed in Chapter 2, a coarse $8 \times 8 \times 8$ look-up table mapping the control values to the output CIE $L^*a^*b^*$ values is generated with the forward model of Equation (2.8). This forward model is used for testing only and to help us predict how well each of the the interpolation functions that we use, will perform for the actual printer calibration. We have from the forward model

$$\mathbf{t} = \mathbf{F}(\mathbf{c}) \tag{3.1}$$

where $\mathbf{c}$ is a 3-vector of control values and $\mathbf{t}$ is a 3-vector of the CIE $L^*a^*b^*$ values. The function $\mathbf{F}$ here refers to the interpolation function that is used to define this mapping. The forward model is shown in Figure 3.1.

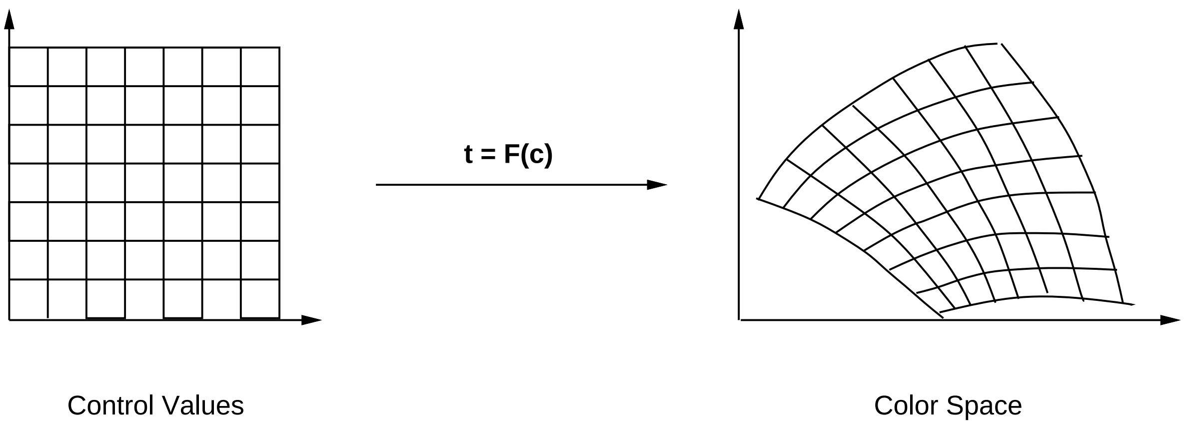

Figure 3.1: The Forward Model

The mapping from the control value space to the CIE $L^*a^*b^*$ space is highly non linear. Our problem now consists of creating a finer, regular grid in the CIE $L^*a^*b^*$ space, such that each point on this fine grid in the color space has a known control vector associated with it, i.e. 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. Since the forward mapping is not defined for the actual printer, we make the same assumption with the mathematical model and define the interpolation function as an approximation to the function $\mathbf{F}$. A fixed set of data points $(8 \times 8 \times 8)$ are used taking into account the fact that it is convenient to take only a small number of measurements from the printer.

Iterative techniques with interpolation are used to estimate the control vector corresponding to every vector $\mathbf{t}_g$ on the finer grid in the CIE $L^*a^*b^*$ space. The mapping from the control value space to the CIE $L^*a^*b^*$ space is not given by an analytic function. This non linear mapping is also a one-to-one mapping which helps us use a fixed point iteration. 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. A special form of Newton's method in three dimensions is used to obtain a new estimate of the control value. This method is called Broyden's method [5] and is equivalent to finding a root of the equation

$$\mathbf{F}(\mathbf{c}) \ = \ \mathbf{t}_g \quad or \quad \mathbf{F}(\mathbf{c}) - \mathbf{t}_g \ = \ 0 \tag{3.2}$$

where the function $\mathbf{F}$ refers to the interpolation function that is used. This iteration is continued until Equation (3.2) is solved to the selected degree of accuracy. Broyden's method is described in Appendix A.

Two kinds of interpolation functions were used, the bell function and the cubic B-spline function. They are described below:

1. **Bell function**: This function is obtained by the convolution of a triangle function with a square function and is defined as [12]

$$S_b(x) \ = \ \begin{cases} \frac{1}{2}(x \ + \ \frac{3}{2})^2 & \frac{-3}{2} \leq x \leq \frac{-1}{2} \\ \frac{3}{4} \ - \ x^2 & \frac{-1}{2} \leq x \leq \frac{1}{2} \\ \frac{1}{2}(x \ - \ \frac{3}{2})^2 & \frac{1}{2} \leq x \leq \frac{3}{2} \end{cases} \tag{3.3}$$

   The sample spacing is assumed to be unity.

2. **Cubic B-spline function**: This function is obtained by the convolution of the bell function with a square function. It is defined as [12]

$$S_c(x) = \begin{cases} \frac{2}{3} \ + \ \frac{1}{2}|x|^2 \ - \ x^3 & 0 \leq |x| \leq 1 \\ \frac{1}{6}(2 \ - \ |x|^3) & 1 \leq |x| \leq 2 \end{cases} \tag{3.4}$$

   The sample spacing is assumed to be unity. The bell and cubic B-spline waveforms are both shown in Figure 3.2.

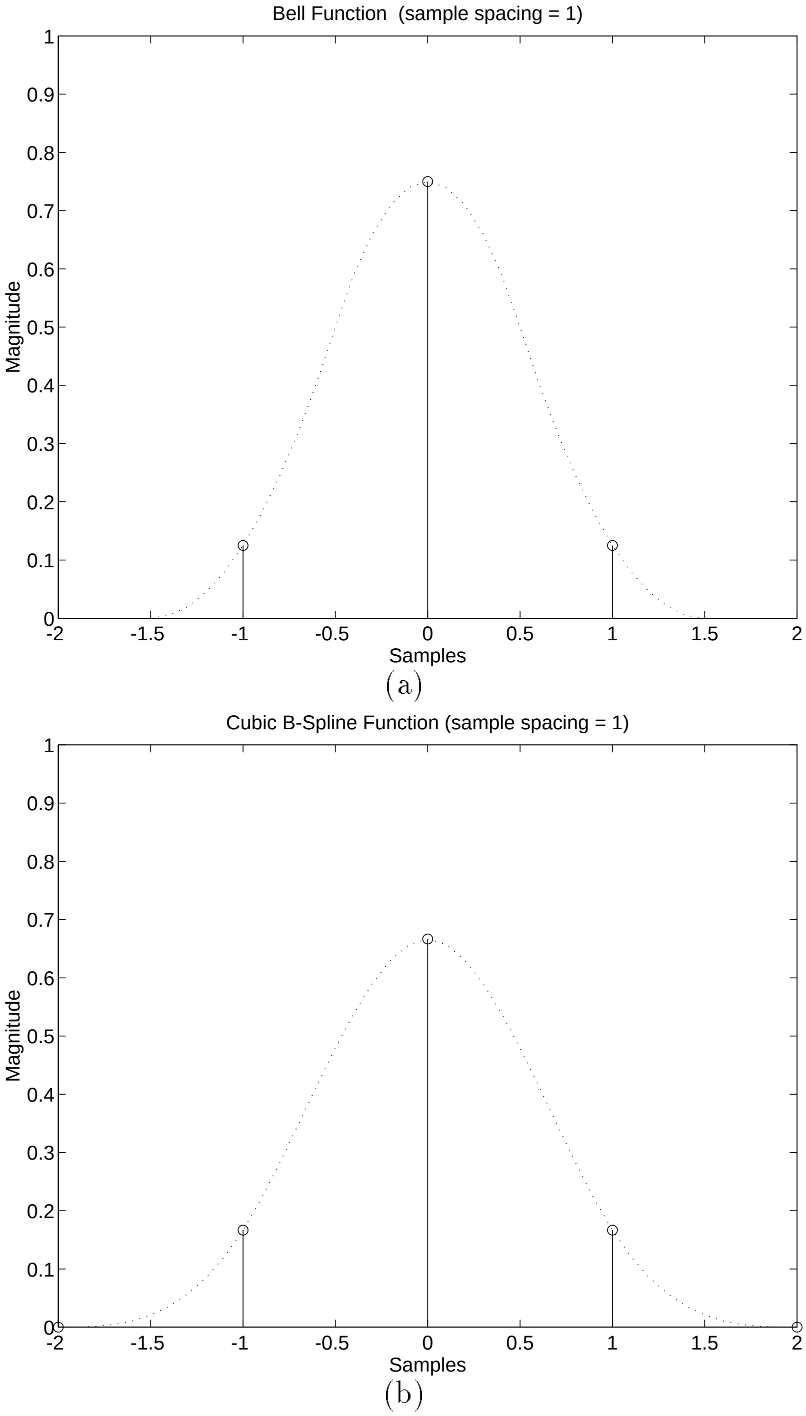

Figure 3.2: Interpolation Waveforms (a)Bell Function, (b)Cubic B-Spline Function

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

$$\begin{aligned} t_{g,l}(c_{1,p}, c_{2,q}, c_{3,r}) \ = \\ \sum_k \sum_j \sum_i \ w_l(i,j,k) S(c_{1,p} - i) S(c_{2,q} - j) S(c_{3,r} - k) \\ p, q, r = 1, \ldots, 8 \quad l \ = 1, 2, 3 \end{aligned} \tag{3.5}$$

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 weights $w_l(i,j,k)$ are obtained by solving three sets of linear equations, one for each l, as shown above. The values $t_{g,l}(c_{1,i}, c_{2,j}, c_{3,k})$ are assumed to 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 $l = 1, 2, 3$.

3.2 Extrapolation

These interpolation schemes assume the presence of a sufficient number of data points close to the interpolated point. This condition is satisfied by points lying well within the printer gamut. Truncation of the data can cause a problem with interpolations near the boundary of the printer gamut. One way of solving this problem is to develop special algorithms to take care of points close to the boundary. We can make these methods as complex as possible since they will be used only once in the development of the fine, regular LUT. However, in this research we wish to study the effect of extrapolation in color systems. Therefore, an approach to solving the truncation problem is to extrapolate the data set for points outside the gamut, see Figure 3.3. These extrapolated points can then be used along with the known data set in an interpolation algorithm which does not require treating the boundary as a special case. Three types of extrapolation schemes were proposed (i) separable linear extrapolation, (ii) non-separable linear extrapolation and (iii) band-limited extrapolation. The last scheme was investigated since the mapping from the CMY colorants to the CIE $L^*a^*b^*$ space for a printer is usually well behaved and smoothly varying.

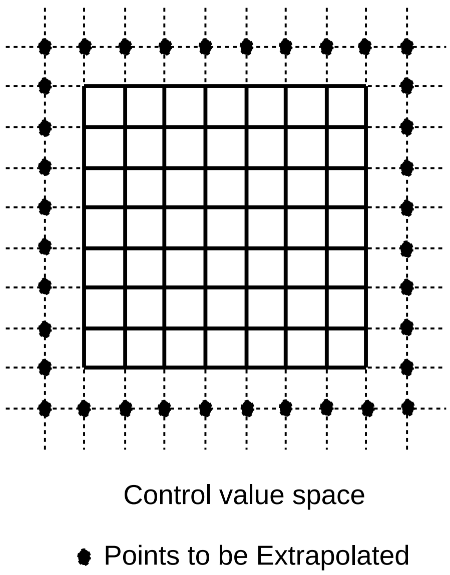

Figure 3.3: Extrapolation in the Control Value Space

3.3 Extrapolation Methods

The CIE $L^*a^*b^*$ components are each assumed to be separable functions of the control values of cyan, magenta and yellow, i.e. we can write

$$L^* = f(c_1, c_2, c_3) = f_1(c_1)\ f_2(c_2)\ f_3(c_3) \tag{3.6}$$

$$a^* = g(c_1, c_2, c_3) = g_1(c_1)\ g_2(c_2)\ g_3(c_3) \tag{3.7}$$

$$b^* = h(c_1, c_2, c_3) = h_1(c_1)\ h_2(c_2)\ h_3(c_3) \tag{3.8}$$

Since we assume separability, the separable 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 corresponding to the L^* values, were created by keeping two of the colorants constant and varying the third. The true L^* values corresponding to these control values are then compared with the original L^* values to calculate the error. Two different separable extrapolation schemes were used, linear extrapolation and band-limited extrapolation. These methods are discussed in detail in the following subsections. The third method of non-separable linear extrapolation is also discussed.

3.3.1 Separable Linear Extrapolation

Linear extrapolation is the simpler extrapolation of the two and is carried out one dimension at a time. The extrapolations were carried out along each of the colorants, one at a time. The order, cyan, magenta and yellow was selected. Let $t_i,\ \ i = 1, ...8$ be the set of eight data points representing the L^* values which correspond to different concentrations of cyan in the range (0,1) and zero concentrations of magenta and yellow. The extrapolated points, t_0 and t_9, corresponding to concentrations of -1/7 and 8/7 of cyan respectively, for the same concentrations of magenta and yellow are obtained by simple linear extrapolation as follows:

$$t_0 = t_1 + (t_1 - t_2) \ \ or \ \ t_0 = 2t_1 - t_2 \tag{3.9}$$

Similarly,

$$t_9 = t_8 + (t_8 - t_7) \ \ or \ \ t_9 = 2t_8 - t_7 \tag{3.10}$$

In this manner, all the extrapolations are done along the cyan direction for different concentrations of magenta and yellow. These extrapolated points are now included in the data set to perform extrapolations in the directions of each of the other colorants. This then involves taking a set of eight data points for which the concentrations of cyan and yellow are kept constant and magenta is allowed to vary from (0,1). The extrapolation is carried out in the same way as before. These points are now included in the data set and extrapolations are carried out in the yellow direction by selecting eight data points with constant cyan and magenta concentrations but different concentrations of yellow in the range (0,1).

The same procedure as above is followed for the data points corresponding to the a^* and b^* values. This gives us the final set of extrapolated data. In this process, the $8 \times 8 \times 8$ grid is now augmented to a $10 \times 10 \times 10$ grid. All these data points can now be used in the interpolation algorithm.

3.3.2 Linear Extrapolation using Taylor Series Expansion

This method of linear extrapolation performs a non-separable vector extrapolation. If $\mathbf{c}_0$ is a measured control value point on the boundary of the control value space of the printer then let the corresponding point in the CIE $L^*a^*b^*$ space, $\mathbf{t}_0$, be denoted by $\mathbf{F}(\mathbf{c}_0)$. Using the Taylor series expansion we can make a linear approximation for any point $\mathbf{c}_i$ not in the control value space in the following manner:

$$\mathbf{F}(\mathbf{c}_i) \approx \mathbf{F}(\mathbf{c}_0) + \mathbf{A}(\mathbf{c}_i - \mathbf{c}_0) \tag{3.11}$$

where $\mathbf{A}$ is a 3×3 matrix and has the following form:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial L}{\partial c_{0,1}} & \frac{\partial L}{\partial c_{0,2}} & \frac{\partial L}{\partial c_{0,3}} \\ \frac{\partial a}{\partial c_{0,1}} & \frac{\partial a}{\partial c_{0,2}} & \frac{\partial a}{\partial c_{0,3}} \\ \frac{\partial b}{\partial c_{0,1}} & \frac{\partial b}{\partial c_{0,2}} & \frac{\partial b}{\partial c_{0,3}} \end{bmatrix} \tag{3.12}$$

$\mathbf{A}$ can be approximated by using least-squares. We first estimate $\mathbf{A}$ by using the points neighbouring the boundary point $\mathbf{c}_0$ of concern. Each of these points forms a different $\mathbf{c}_i$ and the corresponding measured value forms the $\mathbf{F}(\mathbf{c}_i)$. In the case of a point which is at the corner of the cube we have seven neighbouring points which means that we have to solve three sets of equations with seven equations each. Each set of equations gives us the least-squares approximation of each row of the matrix $\mathbf{A}$. For a point on the edge of the gamut, the boundary point of concern, $\mathbf{c}_0$ has eleven neighbours. This results in three sets of simultaneous equations with eleven equations each. A least-squares approach is used again to solve for $\mathbf{A}$. Finally, a point lying on one of the faces of the gamut, is surrounded by seventeen neighbours. This involves solving three sets of simultaneous equations with seventeen equations each. Again, each set of equations solves for each row of the matrix $\mathbf{A}$. Once $\mathbf{A}$ is known, we can easily approximate $\mathbf{F}(\mathbf{c}_i)$ for any $\mathbf{c}_i$ just outside the printer range using Equation (3.11). This gives us a set of extrapolated data. Using this extrapolation method our $8 \times 8 \times 8$ data set is augmented to a $10 \times 10 \times 10$ data set. All these points can then be used in the interpolation algorithm to build the finer, regular grid in the CIE $L^*a^*b^*$ space.

3.3.3 Band-Limited Extrapolation

A well-known technique for solving the extrapolation problem is an iterative method known as the Papoulis-Gerchberg algorithm. Before we discuss this algorithm, let us state the band-limited extrapolation problem for a discrete periodic sequence. For a one-dimensional problem, let $g(m), \quad m \in \mathcal{Z}$ be a discrete periodic sequence with a period $N = 2M + 1$ (without any loss of generality) such that

$$g(m) = g(m+N) \quad m \in \mathcal{Z} \tag{3.13}$$

Assume also that g is band-limited to $[-k_0, k_0]$, so that

$$\sum_{m=0}^{N-1} g(m) e^{2\pi i m k/N} = 0 \quad if \quad M \geq |k| > k_0 \tag{3.14}$$

If we are given a piece of g: $g(m), \quad m \in [-L, L]$, our goal is to recover $g(m), \quad m \notin [-L, L]$. We can now state the discrete-discrete band-limited extrapolation problem as follows:

$$\text{given } g(m), \quad m \in [-L, L],$$

$$\text{find } f(m), \quad -M \leq m \leq M,\ M > L, \text{ such that}$$
$$f(m) = g(m), \quad m \in [-L, L];$$
$$f \text{ is band-limited to } [-k_0, k_0].$$

An iterative extrapolation technique is discussed by Papoulis in [11]. However, this technique has to be modified when only sampled data is available [6]. A possible technique of implementing the iterative technique is by means of the DFT. The first iteration is given with the DFT

$$G(m) = \sum_{j=-M}^{M} g(j) e^{-2\pi i j m/N} \tag{3.15}$$

The n^{th} iteration step proceeds as follows. We form the function

$$F_n(m) = G_{n-1}(m) p(m) \tag{3.16}$$

where

$$p(m) = \begin{cases} 1 & |m| \leq k_0 \\ 0 & |m| > k_0 \end{cases} \tag{3.17}$$

This is equivalent to band-limiting G_{n-1} to the band-limit of f. We then compute the inverse DFT of F_n as below:

$$f_n(j) = \frac{1}{N} \sum_{m=-M}^{M} F_n(m) e^{2\pi i m j/N} \tag{3.18}$$

We next form the function

$$g_n(j) = \begin{cases} g(j) & |j| \leq L \\ f_n(j) & |j| > L \end{cases} \tag{3.19}$$

obtained by replacing the segment of $f_n(j)$ in the interval $[-L, L]$ by the known segment $g(j)$ of $f(j)$. The n^{th} step ends by computing the transform

$$G_n(m) = \sum_{j=-M}^{M} g_n(j) e^{-2\pi i j m/N} \tag{3.20}$$

of the function $g_n(j)$ so formed. The iteration is carried out until the selected degree of accuracy is achieved i.e. until $g_n \approx g_{n+1}$.

The method described above is used to extrapolate the data that we have. One-dimensional band-limited extrapolation is carried out, one dimension at a time. As in the case of separable linear extrapolation, the extrapolation are done first in the cyan direction, i.e. the concentrations of magenta and yellow are kept constant while those of cyan are allowed to vary. This is followed by extrapolations in the magenta and yellow directions.

To compare the separable linear and band-limited extrapolation schemes, a series of lines of the luminance (L^*) values are chosen by keeping two of the dyes constant and varying the third. Each of these lines consists of eight points (by varying one of the dyes from 0 to 1 in steps of 1/7, and keeping the concentrations of the other two dyes constant). Extrapolations are carried out for one point on either side to give us a set of ten points. These ten points are then used in an interpolation routine to estimate the control value for different values of the luminance which lie within the range of the original eight data points. Figure 3.4 shows a typical comparison of interpolation errors using points generated by both these extrapolation schemes.

It can be seen from the figure that separable linear extrapolation gives better results than band-limited extrapolation in all simulations. Linear extrapolation is more attractive than band-limited extrapolation since it is easier to implement and takes less time. Band-limited extrapolation has the advantage of allowing physical constraints to be imposed. However, since an iterative algorithm is used, it has the disadvantage of taking a long time and not giving significantly different results obtained by separable linear extrapolation.

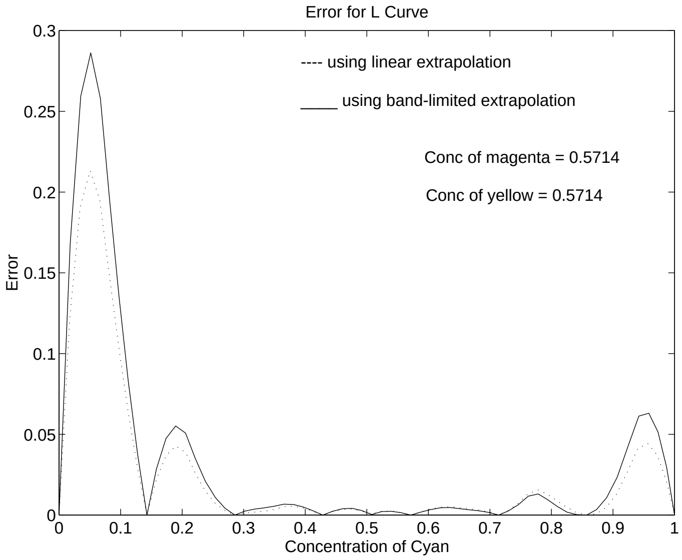

Figure 3.4: Errors due to Different Extrapolation Methods

3.4 Interpolation Methods

The points obtained by linear extrapolation are then used along with the data points from the mathematical model to create a finer, regular LUT in the CIE $L^*a^*b^*$ space. The bell and spline interpolations functions described in section 3.2 are used in the interpolation routine to create the grid. 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

- Divide the range in each of the L^*,a^* and b^* directions (obtained from the $8 \times 8 \times 8$ LUT) by the minimum grid size i.e. form the spacings $(L^*_{max} - L^*_{min})/31$, $(a^*_{max} - a^*_{min})/31$ and $(b^*_{max} - b^*_{min})/31$.

- Choose the minimum of the spacings formed above as the spacing between the two points on the finer grid in the CIE $L^*a^*b^*$ space.

This algorithm gives us a regularly spaced grid such that the spacing in each of the L^*, a^* and b^* directions is the same.

Each of the CIE $L^*a^*b^*$ points on this regular grid are fed through the interpolation routine to obtain a corresponding control vector. The ΔE errors for each of these points can be easily calculated since the mathematical model gives us the 'true' CIE $L^*a^*b^*$ value corresponding to the control vector which is obtained by the interpolation routine. In the case of the actual printer, this would correspond to the actual color that the printer prints on the paper as opposed to the color it is supposed to print.

The bell and spline interpolation functions perform almost identically for the mathematical model. Figure 3.5 and 3.6 show the histograms of the ΔE errors for each of the two cases. The histogram includes all the points which have converged to control vectors which lie within the range of the printer i.e. each of the control values lie in the range (0,1). It can be seen from the two figures that the bell function shows a smaller average ΔE error. As a result, the bell function was the interpolation function chosen for the actual printer calibration which is discussed in the next chapter.

Another comparison was performed using the bell interpolation function. Both the linear extrapolation schemes discussed in the previous sections were used with the mathematical model and the results compared. Histograms of ΔE errors are plotted for both cases in Figures 3.7 and 3.8. Both these histograms now also include the points outside the gamut which converge in the iterative interpolation routine. The mean ΔE errors have increased because of the larger errors associated with points which lie outside the gamut and converge in the interpolation routine. However, the use of vector extrapolation shows slightly better results than the case where separable linear extrapolation is used, see Figures 3.7 and 3.8.

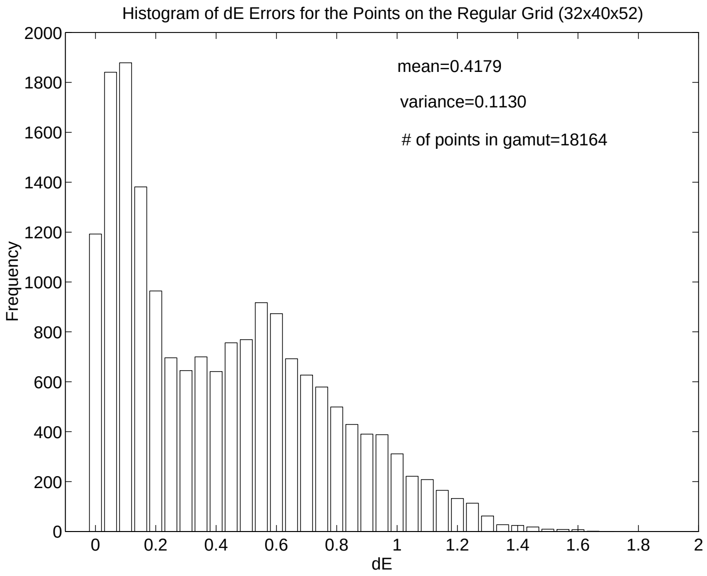

Figure 3.5: ΔE Errors Using Bell Interpolation

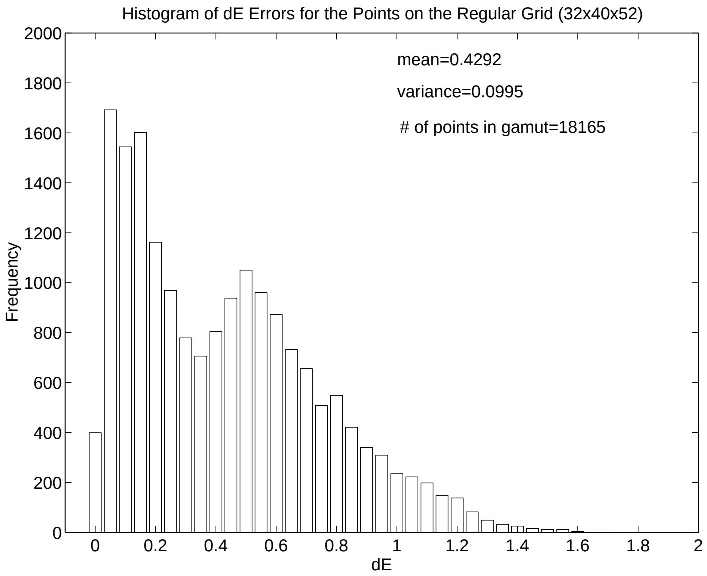

Figure 3.6: ΔE Errors Using Spline Interpolation

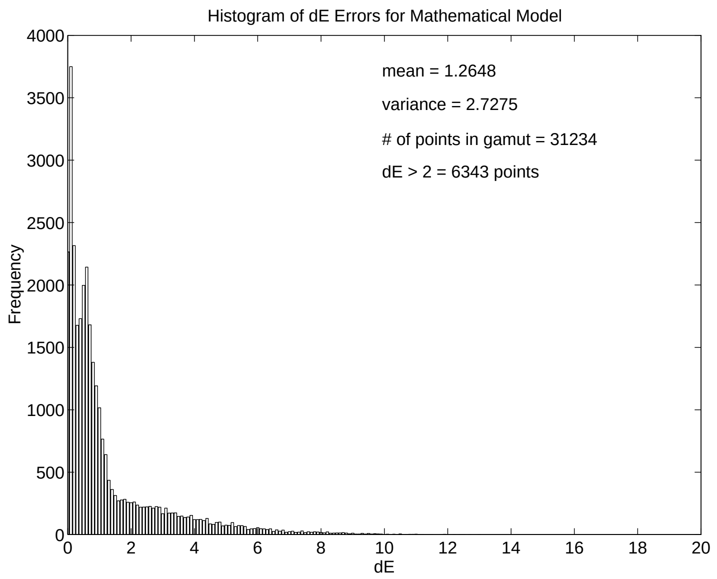

Figure 3.7: ΔE Errors Using Bell Interpolation & Separable Linear Extrapolation (Includes Out-of-Gamut Points)

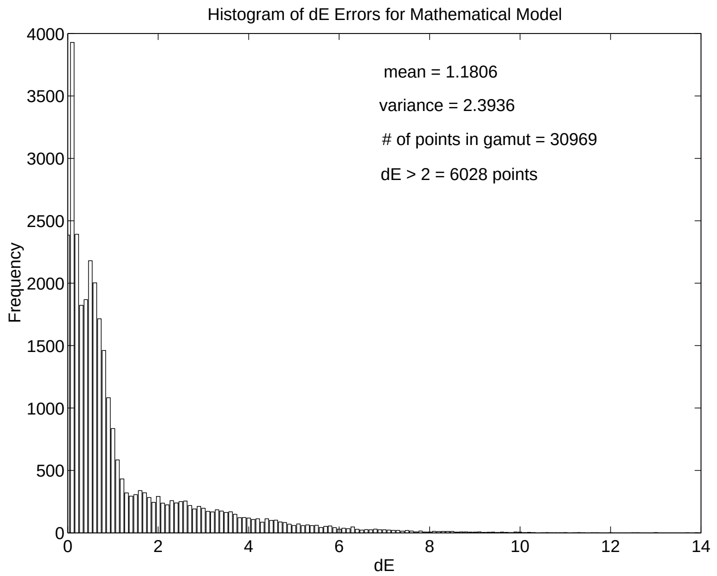

Figure 3.8: ΔE Errors Using Bell Interpolation & Non-Separable Linear Extrapolation (Includes Out-of-Gamut Points)

3.5 Signal-to-Noise Ratio Description

The results obtained in the previous section describe the performance of the interpolation routine under the assumption that no noise is present in the data collection. However, in reality, a color printer will not be noise free and the noise will manifest itself as a positive or negative variation about the true value of the data that is collected. The primary causes are the variability of the printer, the interactions between the colorant layers and the variability in the measuring device.

To simulate this environment, noise is added to the data from the mathematical model and its effect on the change in ΔE errors is observed. The noise is added in the XYZ tristimulus space. The XYZ values of the data points are generated from the mathematical model as described in Chapter 2. If $x_i, \quad i = 1, ..., 512$ is the set of data points without the noise then the signal power is calculated in the following manner:

$$\sigma^2_{signal} = \sum_{i=1}^{512} x_i^2 \tag{3.21}$$

The noise power depends on the SNR (in db) that is desired and is calculated in the following manner:

$$\sigma^2_{noise} = \sigma^2_{signal} 10^{\frac{-SNR}{10}} \tag{3.22}$$

The signal and noise power are calculated separately for each set of X values, each set of Y values and each set of Z values. The noise power thus calculated is used to calculate the standard deviation of the noise sequence which is to be added to the original signal. The noise sequence is chosen to originate from a Gaussian distribution.

One of the problems with using Gaussian noise is that there is always a finite probability of obtaining large values for the noise which may cause very large ΔE errors while creating the LUT. This may not be very realistic. One way to reduce the noise variance is to use correlated noise. This process is described in the next chapter.

Another problem with using the Gaussian distribution is that care must be taken to see that the XYZ values do not fall below zero when the noise component is added to the original signal, since the XYZ values are non-negative. It makes sense to test the mathematical model with the addition of noise only after we have an estimate of the SNR of the printer that we want to calibrate. This will enable us to compare the performance of the printer with the mathematical model and to observe whether the mathematical model is a sufficiently accurate representation of the printer.

Chapter 4

The Printer Calibration

The previous chapters have discussed the calibration of a printer simulated by the forward model developed in Chapter 2. The interpolation functions suggested were used in the forward model with very good results. Low average ΔE errors were obtained with both the bell and the cubic B-spline interpolation functions. These results were obtained under the assumption that noise is absent. However, in reality a printer will have an appreciable noise component which will affect the calibration process. It is an aim of this research to calibrate the printer and to estimate it's SNR. This SNR can then be used to simulate a noisy mathematical model and compare the relative ΔE errors between the two. This will help us distinguish between the functional problems and noise problems of the printer.

4.1 Description of the Printer

The printer that was calibrated as part of this research was the XL7700 printer manufactured by KODAK. The XL7700 printer is a digital continuous tone printer that converts digital data from a host computer into color or black-and-white pictures of a quality approaching that of traditional photographs. The printer does this without light forming optics, wet processing chemicals or a separate fusing process. The XL7700 produces pictures through a process called thermal dye transfer [9].

The basic thermal dye transfer process consists of transferring the dye which is on a carrier ribbon to another substrate, paper or transparency [14]. The ribbon is made with successive strips of yellow, magenta and cyan dyes. The yellow image is laid down first; the paper is repositioned; the magenta image is then laid down; the paper is repositioned and the final cyan image is laid down. The paper and ribbon are moved simultaneously over a thermal head which melts or vaporizes the dye and thus transfers it to the paper (Figure 4.1). The printing head consists of one individually controlled heating element for each pixel on a line. This allows the thermal printers to produce high resolution images. The XL7700 head contains 2048 individual heating elements. Thermal control of each element drives the appropriate amount of dye from the ribbon onto the paper or transparency material, thus forming a continuous tone picture.

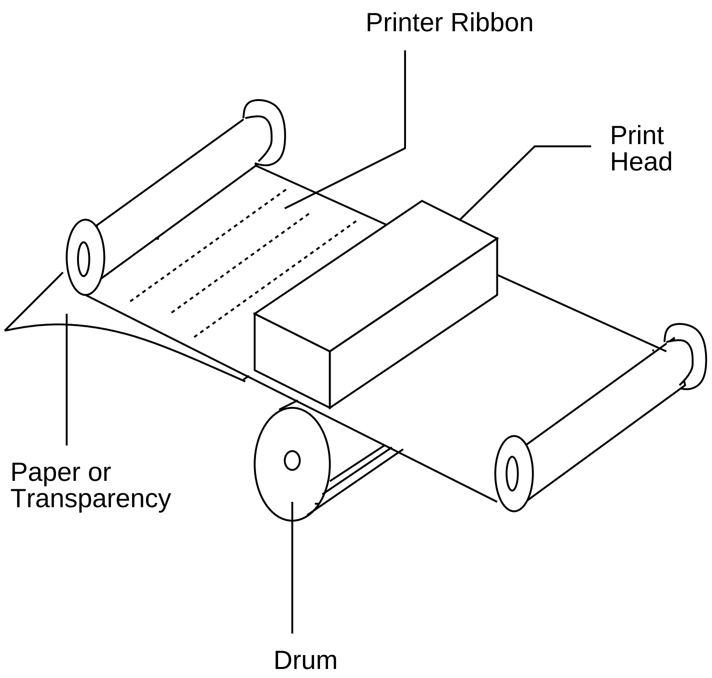

Figure 4.1: The Thermal Dye Transfer Process

A number of problems can arise with the thermal dye transfer methods. These include:

- variability of the heating elements
- variable warm-up time and behaviour of the heads
- variation caused by ambient conditions
- hysteresis of the heating elements
- dye inhibition, a dye laid down on top of another dye does not adhere as well as it would to paper
- back transfer - the heating for a subsequent dye can melt the dye that has been previously laid down

In addition to these errors, errors are also introduced due to the calibration method that is used. These errors will mainly be errors due to measurements and interpolation schemes that are used. All these errors will be discussed in greater detail with reference to the calibration of the XL7700.

4.2 Data Collection and Observations

A calibration chart was generated from the the Kodak XL7700 by varying the concentrations of the cyan, magenta and yellow dyes in a uniform manner. The control value space was divided into eight equispaced samples in each direction (from 0 to 1 in steps of 1/7) 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 $8 \times 8 \times 8$ data set.

As seen in Figure 4.2, there are a number of steps in the calibration process which coupled with the variability of the printer, contribute to the total error. Before proceeding further with the calibration process, it was important to determine if there were any variations in the printouts produced by the XL7700. Different factors contribute to the error during the calibration process of the printer. These errors appear at every stage of the process. One of the important contributions to this error is the inherent variability of the printer. Four kinds of inconsistencies were observed due to the printer variation. These variations of the printer were used to give us an estimate of the noise for the SNR calculation of the printer.

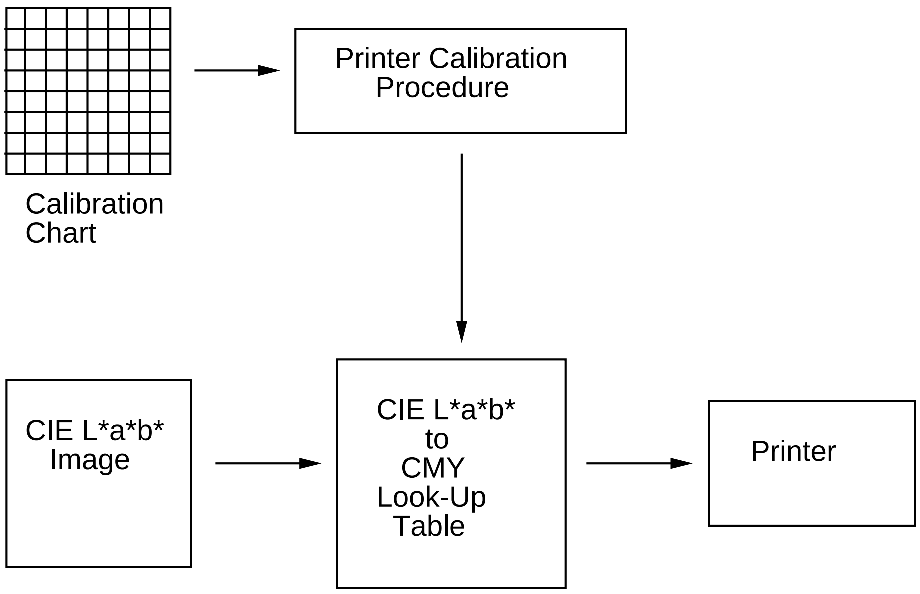

Figure 4.2: Processing the Image for Printing

1. **Initial Warm-Up Time:** The XL7700 has an initial warm-up time. Several identical images of the calibration chart were printed out in succession as soon as the printer was ready for printing. However, after measurements of the first two charts were made, it was noticed that there was an appreciable color difference between the CIE $L^*a^*b^*$ values of the color patches on the first chart as compared to those on the second. ΔE errors of the order of 10 or

more were observed between the first print and the second print. However, when successive prints were compared with these two initial prints they were much closer to the the second print than the first. ΔE errors of the order of 10 were still observed between each of these prints and the first. It was concluded that the first print produced by the XL7700 after it is switched on is an inaccurate representation of the true image. An initial print was the warm-up time assumed for the printer for all future measurements i.e. only the prints after the first were used in the calibration process.

2. **Measurements on Different Heads:** 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 (corresponding to the same control value) printed at different positions on the 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 dye printed on to the paper was also a function of the particular head that printed it. This led to the conclusion that the calibration should be carried out for each individual head. In the calibration procedure used, the calibration chart was printed only under the second head.

3. **Measurements on Different Sheets of Paper:** The XL7700 also displayed a certain amount of variability when the same color patch was printed at the same location on different sheets after the first sheet. The standard deviation over different sheets ranged from 1.06ΔE to 1.93ΔE. This provides a lower bound on the accuracy of the calibration. This indicated the need to measure the calibration chart over several printouts and to study the average ΔE errors that are observed due to this anomaly. Therefore, the

measurements of the calibration chart were measured over several printouts and the average values over all the sheets were used as the points of the known $8 \times 8 \times 8$ coarse data set. Another problem arose at this stage since the entire calibration chart (which consisted of 512 color patches) could not be printed on one sheet under the second head only. Each color patch had to be large enough for it to lie completely under the aperture of the GRETAG SPM-50 spectrophotometer. This meant that the calibration chart had to be printed on two separate sheets, each sheet containing half the total number of color patches on the chart. Since all the color patches were generated sequentially by changing the concentration of the cyan, magenta and yellow dyes successively in that order, it was obvious that the measured errors would incur a block change when half of the chart was printed on a new sheet. Therefore, larger errors would occur in the regions where adjacent color patches of the calibration chart were printed on different sheets. To keep this error random and not just in one region of the CIE $L^*a^*b^*$ space, the calibration chart was printed out in a random manner. These measurements were then averaged over several printouts to obtain average CIE $L^*a^*b^*$ values for all the 512 points of the coarse LUT. Figure 4.3 shows the calibration chart which was printed out in an ordered fashion under the second head.

4. **Error due to the Measuring Device:** Another source of error is due to the measuring device itself. The GRETAG SPM-50 spectrophotometer was used to measure the CIE $L^*a^*b^*$ values of the color patches on the calibration chart. The GRETAG has a certain amount of error when measuring the same color patch. This is known as the reproducibility of the instrument. The specifications from the GRETAG manual mention a reproducibility of ΔE ≤ 0.2. To average out this error and to take care of variations under the same head, each color patch was measured five times at different regions in the same

Figure 4.3: Both (a) and (b) form the calibration chart and are printed on separate sheets under the second head

patch and the average of these was taken as the CIE $L^*a^*b^*$ value for that particular patch on a particular sheet under the second head.

4.3 Interpolation Error (Creating The Look-Up Table)

The next step of the calibration process involves creating the finer, regular LUT from the coarse set of 512 measured values. The errors discussed in the previous section were errors introduced in the calibration process due to the variability of the XL7700 and the measuring instrument used. However, in trying to create a finer, regular grid in the CIE $L^*a^*b^*$ space interpolation is used over the known coarse $8 \times 8 \times 8$ set of data that was measured. This introduces an additional error component.

As discussed in Chapter 3, interpolation problems occur for points close to the boundary. To overcome this problem, linear extrapolation was used to estimate the CIE $L^*a^*b^*$ values for points outside the printer gamut. The printer gamut was thus augmented to a $10 \times 10 \times 10$ grid. These extrapolated points are used along with the measured 512 data points in the interpolation process to create the finer LUT.

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 of 0.1) was used. The testing procedure is described schematically in Figure 4.4. 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 the CIE $L^*a^*b^*$ values of the test chart. The bell interpolation function was used in the interpolation process since it gave a smaller ΔE error for the

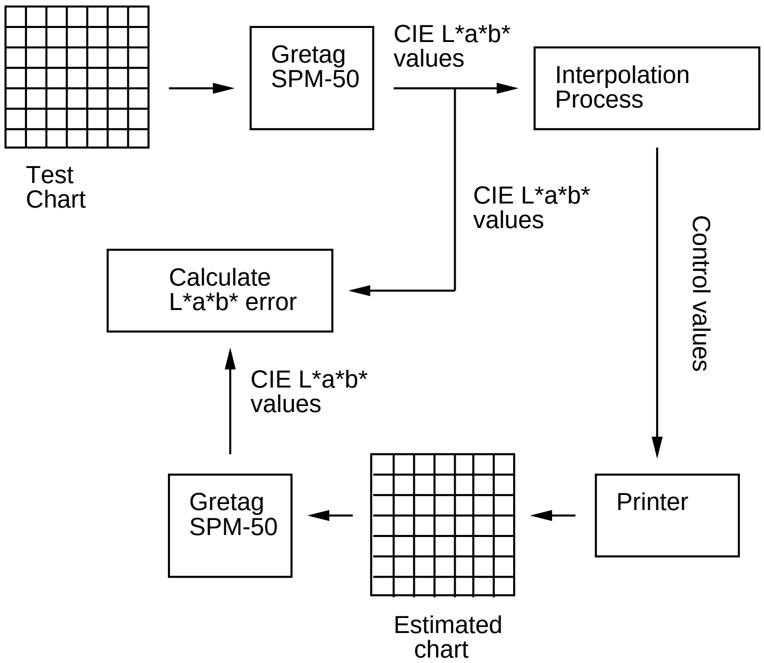

Figure 4.4: Testing the Interpolation Scheme

mathematical simulation.

The same algorithm that was used in the mathematical model 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 same manner as before i.e.

- divide the range in each of the L^*, a^* and b^* directions (obtained from the $8 \times 8 \times 8$ LUT) by the minimum grid size i.e. form the spacings $(L^*_{max} - L^*_{min})/31$, $(a^*_{max} - a^*_{min})/31$ and $(b^*_{max} - b^*_{min})/31$.
- 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.

Using this algorithm a grid size of $32 \times 48 \times 59$ was created in the CIE $L^*a^*b^*$ space. Only a fraction of these points were within the boundary of the the printer gamut.

The points which fell outside the gamut and did not converge in the iteration were flagged. The ΔE errors which arise due to the interpolation algorithm are tabulated in Table 4.1. These errors here indicate a lower bound on the errors obtained by the calibrated printer since the errors for arbitrary points will include additional interpolation error caused by a less accurate interpolation scheme such as trilinear interpolation. Figure 4.5 shows a histogram of these ΔE errors.

Table 4.1: ΔE Errors Due to Interpolation

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

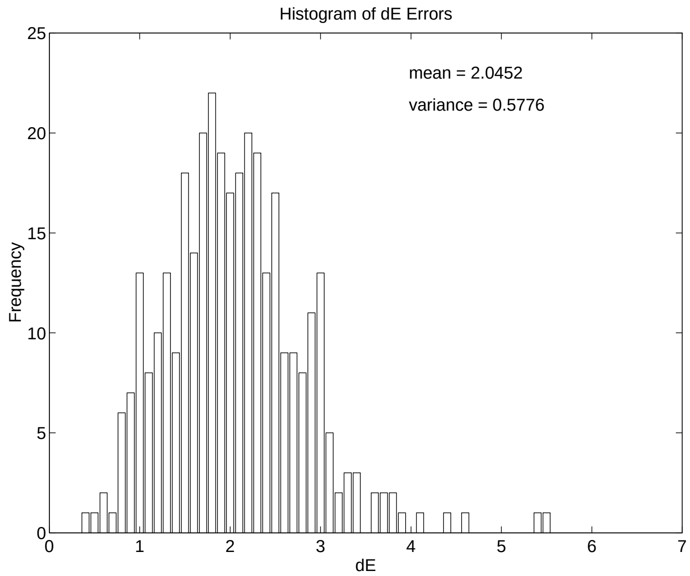

Figure 4.5: ΔE Errors Due to Interpolation

4.4 Trilinear Interpolation

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. The process of trilinear interpolation is described in Appendix B.

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. To do this, the CIE $L^*a^*b^*$ value of each test point from the image is used and the corresponding cube containing this point is located in the fine LUT that has been created. A simple trilinear interpolation as described in Appendix B is performed using the eight points on the corners of the cube as data points. This gives us a control value for each of these 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 4.4. The interpolation procedure now, refers to trilinear interpolation.

We can predict that the error that is added by trilinear interpolation will not be significantly large if our assumptions that the forward mapping is well behaved and smoothly varying and that the grid size is fine enough to make linear approximations are valid. The relevant statistics of the ΔE errors are shown in Table 4.2. It is observed that the average errors for the test data set that was used did not change appreciably after trilinear interpolation. However, the variance of the ΔE errors increased almost two-fold. The histogram of ΔE errors is plotted in Figure 4.6 and should be compared to Figure 4.5.

Table 4.2: ΔE Errors Due to Trilinear Interpolation

Statistics	
Max ΔE	6.9392
Average ΔE	2.1939
Variance ΔE	0.9849

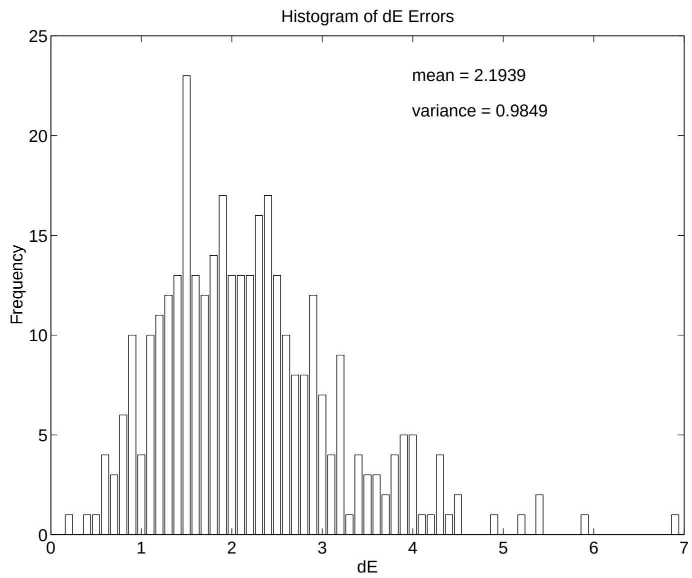

Figure 4.6: ΔE Errors Due to Trilinear Interpolation

Trilinear interpolation may not be feasible for points close to the boundary if all the eight corner points of the cube in which the test point lies, do not have corresponding control values (Figure 4.7). This would imply that in the process of creating the grid these points did not converge in our iteration. This is a serious problem since we would like to make sure that we can print at least all the points lying in the printer gamut. An easy initial solution to this problem was to augment our data set further from a $10 \times 10 \times 10$ set to a $12 \times 12 \times 12$ set. This would involve another set of linear extrapolations at the boundary of our already augmented $(10 \times 10 \times 10)$ data set. This gives us a wider range for the test point to converge to some control vector, each element of which can now lie between (-2/7,9/7) which is greater than the range (-1/7,8/7) which we had for our smaller augmented grid $(10 \times 10 \times 10)$.

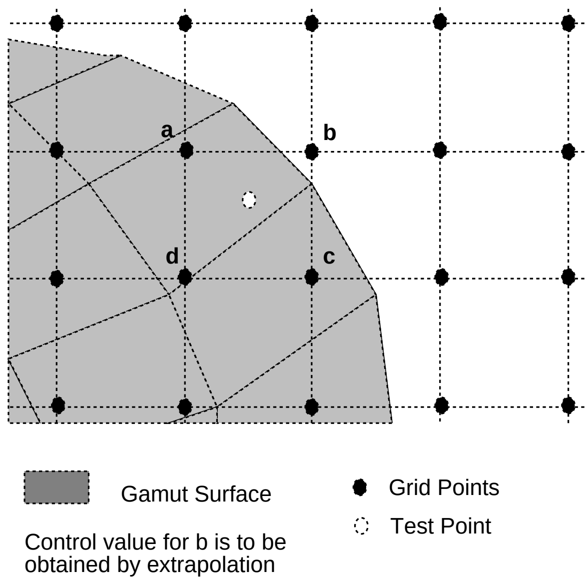

Figure 4.7: Extrapolation to Cover Entire Gamut

The smaller augmented grid of $10 \times 10 \times 10$ gave convergence problems for points on the grid which were in the darker (low $L*$ values) regions of the gamut. A test pattern of 54 points consisting of color patches close to the boundaries was printed on the XL7700. These points corresponded to all combinations of control values ranging from 0.03 to 0.09 in steps of 0.03, and another combination of control values ranging from 0.91 to 0.97 in steps of 0.03. The former set gives us colors of high luminance while the latter set consists of the darker colors. The smaller ($10 \times 10 \times 10$) augmented grid gave convergence problems for points on the grid which were in the darker regions. Most of the test points in this region did not lie in cubes for which all the eight corner points had converged. However, by augmenting the size of the data set to $12 \times 12 \times 12$ by another linear extrapolation at the boundary, the corner points containing the same test points had converged to their corresponding control values. This occurred for all but one point of one cube containing one of the test points. This would have been a problem in the lighter regions of the gamut since a small change in control values gives a corresponding large change in ΔE between the two colors. However, in the darker regions small changes in control values would imply a small ΔE error between the two colors. This means that it is reasonable to do simple linear extrapolations for such corner points of the cube knowing that the ΔE error in doing so would not be large.

The histogram corresponding to this test set is shown in Figure 4.8. The errors below three correspond to errors in the darker regions of the gamut (high control values), whereas the errors larger than three correspond to the lighter points in the gamut (low control values).

From the mathematical model, in Chapter 3 (Figures 3.7 and 3.8), it was seen that the average ΔE errors were smaller using non-separable linear extrapolation. To test whether the same result applies to the actual data, the measured data set was augmented to a $10 \times 10 \times 10$ grid using linear extrapolation by Taylor series expansion. The same set of color patches which were used in the previous case were tested for the LUT created by using this new set of extrapolated points. It was seen that all the 54 test points near the boundary were located in cubes for which each of the corner points had converged in the process of creating the LUT. This is a definite improvement over the previous LUT that was created by using separable linear extrapolation. Furthermore, extra extrapolations were not required, unlike the previous case. It was concluded that non-separable linear extrapolation was more accurate than separable linear extrapolation. The histogram of ΔE errors that were associated with these color patches after trilinear interpolation is similar to Figure 4.8. The LUT that was created using the extrapolated points generated by this method was the final LUT.

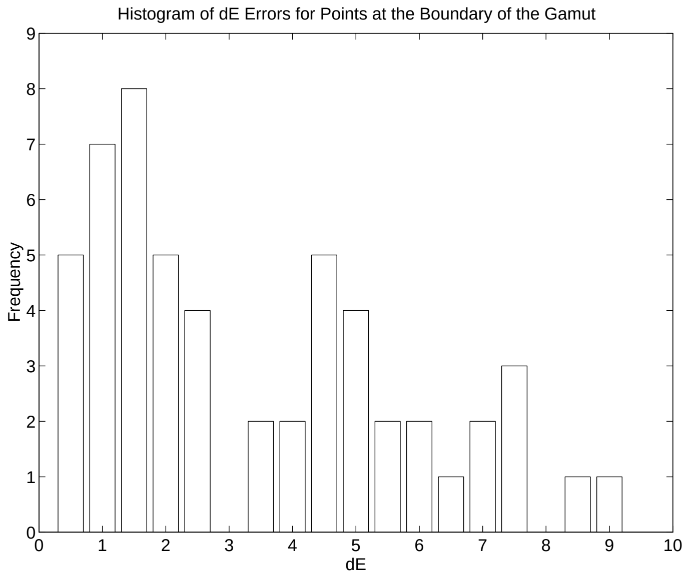

Figure 4.8: ΔE Errors for Points Close to the Boundary of the Gamut Using Separable Linear Extrapolation

4.5 Signal-to-Noise Ratio of the XL7700

As described earlier, the coarse $8 \times 8 \times 8$ data set was measured over seven different sheets and an average value over these sheets was taken as the measured value. An additional averaging was done over each sheet. Due to the variation of the printer on the same sheet, an average of five different measurements was used for each of the 512 color patches that were printed out. This was done for all the seven sets of 512 color patches which were measured. Finally an average over all these seven sets was then used as the measured value. Measurements were made in the CIE $L^*a^*b^*$ space by using the GRETAG SPM-50 spectrophotometer. The SNR calculation is done in the XYZ tristimulus space since the GRETAG makes measurements in this space and then converts it to the CIE $L^*a^*b^*$ space. The kind of averaging that was used gives rise to the fact that two different kinds of printer variations give rise to two different SNR calculations

(i) SNR calculation due to inter-sheet variations

(ii) SNR calculation due to variation on the same sheet.

In each case the SNR is calculated for the single head under which measurements are made.

4.5.1 SNR calculation due to inter-sheet variations

The measured value of each of the 512 color patches is the average over seven different sheets, all measured under the same head. These average values are defined as the signal value associated with each color patch. Let :

$$X_k(i) \quad i = 1, \ldots, 7 \quad k = 1, \ldots, 512 \tag{4.1}$$

represent the X tristimulus value of the k^{th} color patch on the i^{th} sheet. Then the average X value of the k^{th} color patch is given by:

$$\bar{X}_k = \frac{1}{7} \sum_{i=1}^{7} X_k(i) \tag{4.2}$$

The signal power corresponding to the X value only is given by:

$$\sigma_X^2 = \sum_{k=1}^{512} \bar{X}_k^2 \tag{4.3}$$

The noise variance for the k^{th} color patch is given by:

$$\sigma_{X_k}^2 = \frac{1}{7} \sum_{i=1}^{7} (X_k(i) - \bar{X}_k)^2 \tag{4.4}$$

The average noise power is calculated by taking the average over all the 512 color patches as follows:

$$\sigma_N^2 = \frac{1}{512} \sum_{k=1}^{512} \sigma_{X_k}^2 \tag{4.5}$$

The signal to noise ratio is then given by:

$$SNR = 10 \log_{10} \frac{\sigma_X^2}{\sigma_N^2} \tag{4.6}$$

Similar SNR calculations were done using the average Y and Z values as the signal components. Corresponding noise powers were calculated as above and a SNR was obtained.

Using the data that was measured, the SNR was calculated for each of the three components of the tristimulus values i.e. X, Y and Z. The SNR values that were obtained are shown in Table 4.3.

Table 4.3: SNR of the XL7700 due to Inter-Sheet Variations

	SNR
X	59.4382
Y	58.7063
Z	56.9435

4.5.2 SNR calculation due to variation on the same sheet

To calculate the SNR of the printer due to the variation of the printer over one particular sheet (under a single head), we need to make multiple measurements for each color patch on a sheet. In making the measurements for the 512 data points, an average of five measurements for each color patch was used as the measured value. However, each of these measurements was not recorded since the averaging was done by the GRETAG while the measurements were made. Therefore a subset of the data set corresponding to 14 different colors were printed again. Each of these color patches were measured at eight different regions in the patch and these values were recorded. As before, CIE $L^*a^*b^*$ values were measured and converted to XYZ tristimulus values. The SNR calculation was done as before except that the averaging is now done over the values measured on the same sheet. As before let

$$X_k(i) \quad i = 1, \ldots, 8 \quad k = 1, \ldots, 512 \tag{4.7}$$

represent the i^{th} measured value of the X tristimulus value of the k^{th} color patch on the same sheet. The average signal power and noise power is calculated in the same way as before except that the averaging is done over the eight different measurements that are made.

The SNR calculations are again done separately for the X,Y and Z values. Table 4.4 shows us the SNR that was obtained for each of the three components.

Table 4.4: SNR calculation due to variations on the same sheet

	SNR
X	53.4497
Y	52.5841
Z	50.5775

4.5.3 SNR calculation taking into account both kinds of variations

Two more sheets with the same 14 color patches were printed out. These color patches were measured again in the same way as the first sheet that was printed. As before eight measurements were made at different locations of the same color patch on each sheet. Each of these measurements also showed variations about the mean. These variations are plotted in Figure 4.9 for each of the three sheets. The y-axis denotes the luminance(L^*) value for each of the 14 color patches. The x-axis has an arbitrary numbering sequence.

An attempt was made to calculate a SNR for the XL7700 using all the data that we now have of the 14 different color patches, each set of which is printed on three different sheets. Each color patch on each sheet has an average luminance value associated with it. The mean of these average values over all the three sheets is used as the luminance value for each color patch. Variations of luminance for each patch are compared to this mean level and an SNR calculation is done using this data set. The SNR is again calculated in the XYZ tristimulus space since the GRETAG makes measurements in this space and then performs the conversions.

Let $X_k(i,j) \quad i = 1,\ldots,8 \quad j = 1,2,3 \quad k = 1,\ldots,14$ be the X value of the i^{th} measurement of the k^{th} color patch on the j^{th} sheet. The average signal value for the k^{th} color patch is given by:

$$\bar{X}_x^2 = \frac{1}{8 \times 3} \sum_i \sum_j X_k(i,j) \tag{4.8}$$

The signal power is given by:

$$\sigma_X^2 = \sum_{k=1}^{14} \bar{X}_k^2 \tag{4.9}$$

The noise power for each color patch is calculated as follows:

$$\sigma_{X_k}^2 = \frac{1}{8 \times 3} \sum_i \sum_j (X_k(i,j) - \bar{X}_k)^2 \tag{4.10}$$

The average noise power is now given by:

$$\sigma_N^2 = \frac{1}{14} \sum_{k=1}^{14} \sigma_{X_k}^2 \tag{4.11}$$

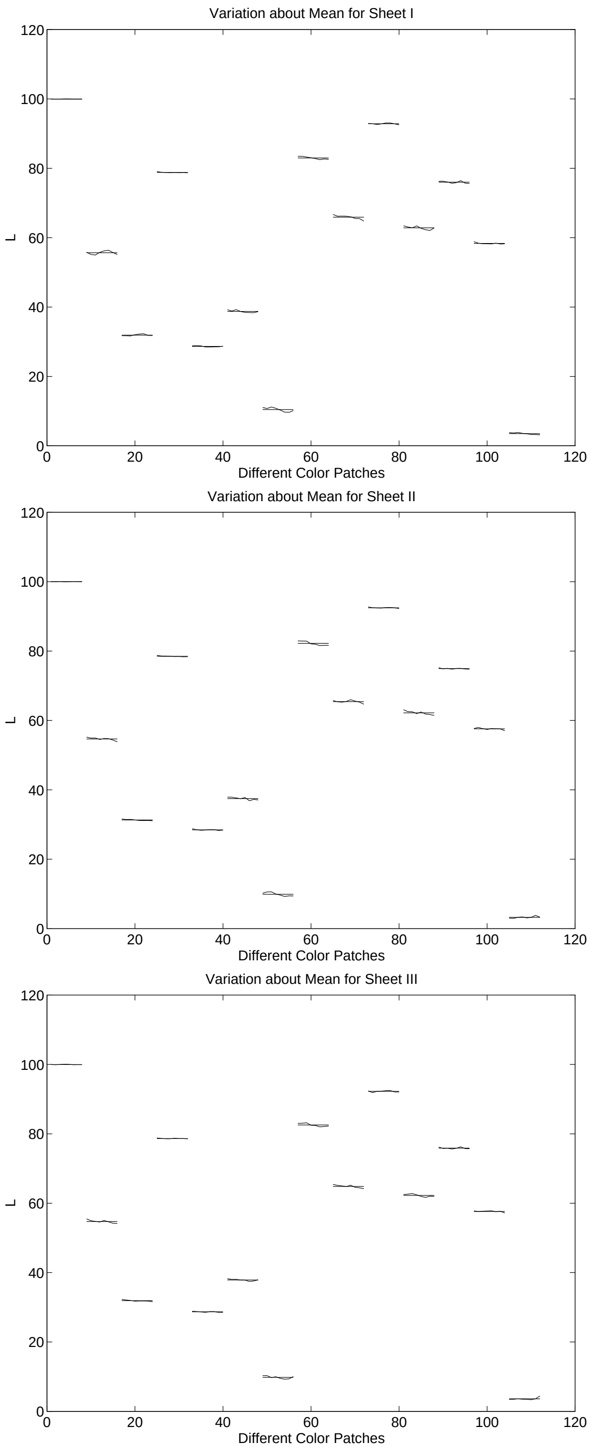

Figure 4.9: Variation of L^* values on Different Sheets

The SNR is given as before:

$$SNR = 10 \log_{10} \frac{\sigma_X^2}{\sigma_N^2} \tag{4.12}$$

Figure 4.10 shows a plot of the mean luminance values for each color patch on the three different sheets. Variation is observed between the means on different sheets. However, if we assume that this variation is small compared to the variation on each sheet about its own mean then the above calculation seems to be a reasonable calculation for the SNR. We do, however, expect that the SNR calculated as above will be smaller than an individual SNR calculation for each sheet. This is because by taking the mean over the three sheets as the signal level, and all variations about it due to the noise, we are assuming a larger noise variance. This would give us a smaller SNR. The SNR that was calculated in this case is given in Table 4.5. As expected the SNR showed a slight decrease.

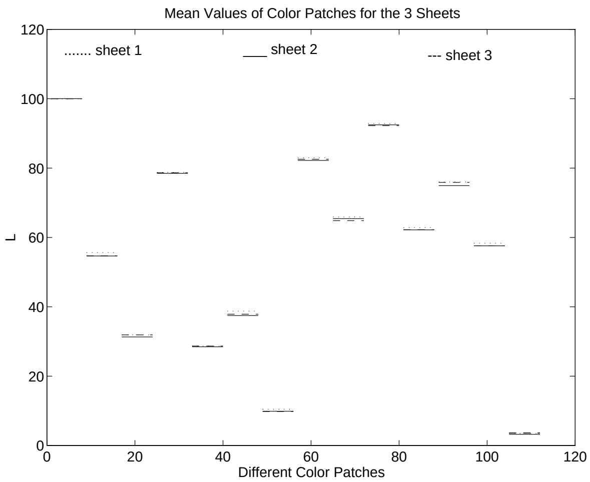

Figure 4.10: Mean L^* Values for Each Sheet

An attempt was made to relate the noise variance to the luminance level of the color signal. For the 343 points which we had measured earlier to test the interpolation

Table 4.5: SNR calculation with both variations

	SNR
X	49.9488
Y	49.8337
Z	49.6174

scheme, the ΔE errors corresponding to each of the color patches were plotted against the luminance level of the color signal. Figure 4.11 shows the corresponding plot, which shows a positive correlation between the two. The correlation coefficient was calculated between the ΔE values and the corresponding L^* values. This gave us a value of about 0.44 which shows a distinct positive correlation i.e. an increase in the luminance level of the signal would most likely increase the ΔE error after interpolation.

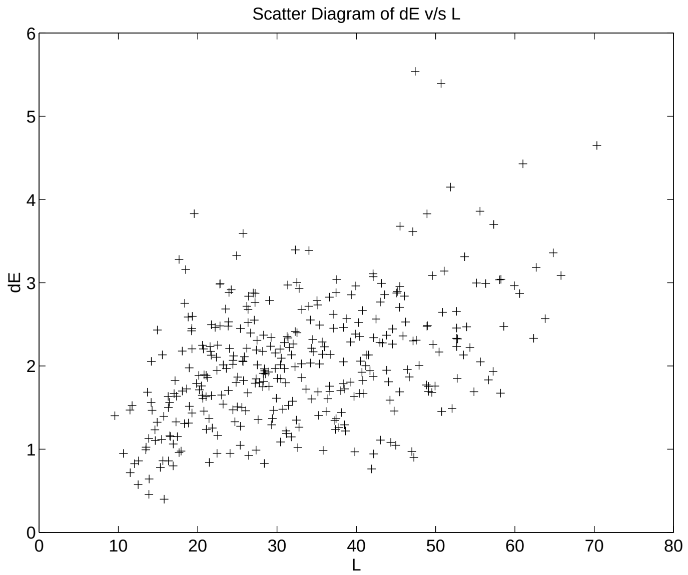

Figure 4.11: Luminance v/s ΔE Errors

4.6 SNR Simulation with Mathematical Model

Now that we have an estimate of the SNR of the XL7700 color printer, we can use the mathematical model developed in chapter 2, to simulate a printer with the same SNR. This is done by adding noise to the set of XYZ tristimulus values as described in chapter 3. If $X_i \quad i = 1, \ldots, 512$ is the set of X values of the simulated data points, then the signal power is given by:

$$\sigma_X^2 = \sum_{i=1}^{512} X_i^2 \tag{4.13}$$

The noise power is calculated using the knowledge of the SNR that is desired,

$$\sigma_N^2 = \sigma_X^2 10^{\frac{-SNR}{10}} \tag{4.14}$$

The noise power is used to calculate the standard deviation of the noise sequence that is to be added to the original signal. This corrupted signal is now used as the set of X values of the data set. Similar manipulations are done for the Y and Z values of the simulated data. The noise sequence is chosen to originate from a Gaussian distribution.

The SNR of the XL7700 was estimated to be in between 55 to 60dB. The mathematical model is simulated with a noise power of 60dB that is added according to the method described above. The LUT is constructed and the ΔE errors are calculated. The distribution of ΔE errors for the grid points which lie in the color gamut of the simulated printer are shown in the histogram in Figure 4.12. As can be seen from the figure, the variance of the errors is quite large. This is because the noise that was added to the XYZ data was uncorrelated Gaussian noise and results in a larger ΔE error that would be obtained if the noise on each channel were correlated as it actually is.

To generate correlated noise for the simulation, for each XYZ triple generate a 31-vector (since the spectrum is sampled at steps of 10nm between 400nm and 700nm) of white Gaussian noise of standard deviation σ_n. Let this vector be denoted by $\mathbf{n}$. The noise for each channel is obtained from the dot product of the corresponding color matching function with the noise spectrum. The simulated measured value is then given by:

$$X_m = X + \bar{\mathbf{x}}^T \mathbf{n}, \tag{4.15}$$

where X_m is the simulated measured value, X is the actual value and $\bar{\mathbf{x}}$ is the color matching function corresponding to the X value. The same procedure is carried out with the actual Y and Z values to get Y_m and Z_m. The value of σ_n has to be adjusted to get the appropriate SNR.

Correlated noise was added to the XYZ data such that an SNR of 61dB was used. This data that was generated by the addition of correlated noise was run through the interpolation routine to generate the LUT. Control values were obtained for these points and the ΔE errors were calculated. The histogram of these ΔE errors is shown in Figure 4.13. It can be seen that the variance of the ΔE errors reduced appreciably with the use of correlated noise.

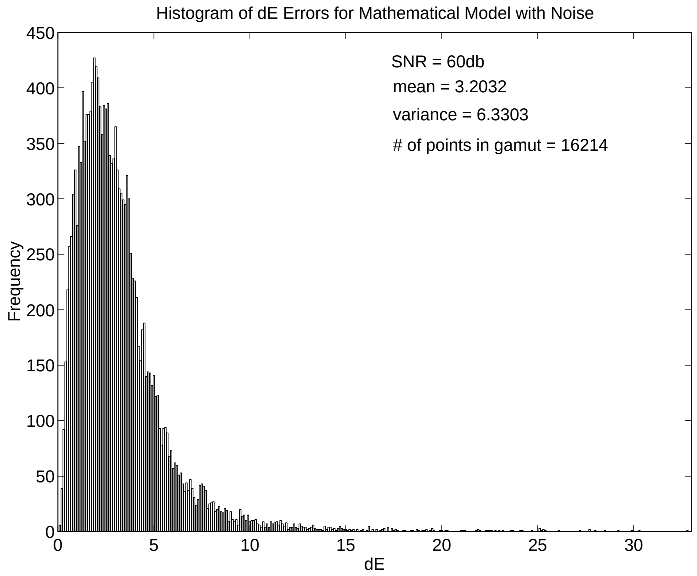

Figure 4.12: ΔE Errors for Model with 60dB Uncorrelated Noise

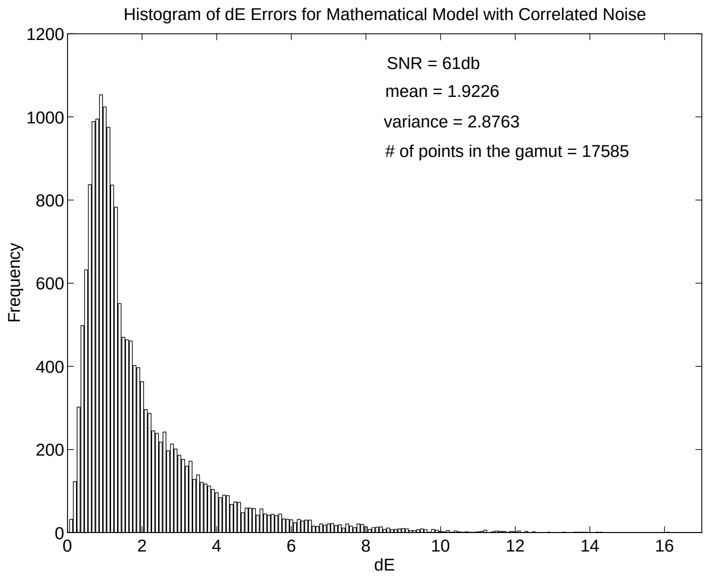

Figure 4.13: ΔE Errors for Model with 61dB Correlated Noise

Chapter 5

Summary and Conclusions

5.1 Summary

Chapter 1 discussed the research problem and the motivation for looking into it. The fundamentals of color science were necessary to understand the work. The basics of color matching, a vector space approach to color imagery and the mathematical basis for some subjective color phenomena were discussed. The CIE L^*u^*v* and the CIE $L^*a^*b^*$ uniform color spaces were discussed. The CIE $L^*a^*b^*$ space was used to calculate ΔE errors between colors in all future work.

Mathematical models for a CRT and a printer were developed in Chapter 2. The additive and subtractive principles on which the CRT and the printer are based respectively are discussed in detail. The properties of an ideal CRT and a printer are mentioned along with the limitations of each of the mathematical models. The mathematical models help us to simulate these output devices and understand their behaviour under the model limitations.

For the purpose of this thesis, we chose to calibrate the KODAK XL7700 thermal dye-transfer printer. Before the calibration of the actual printer was carried out, the calibration of the mathematical model was performed in Chapter 3. In this chapter, the different interpolation and the extrapolation techniques that were used were discussed and implemented in the calibration of the mathematical model of the printer. Two types of interpolation functions were used, the bell function and the cubic B-spline function. Extrapolation of the data set was carried out to reduce interpolation errors for points at the boundary of the color gamut. Three types of extrapolation were looked at, separable linear extrapolation, non-separable linear extrapolation and band-limited extrapolation. It was shown that non-separable linear extrapolation gave us the best results. The LUT for the mathematical model was developed without taking into account the presence of noise. The bell and the cubic B-spline interpolation functions performed almost identically. The bell function gave a lower average ΔE error of 0.4179 in the creation of the LUT and was therefore selected as the function to be used in the calibration of the actual printer.

The actual calibration of the KODAK XL7700 was discussed in Chapter 4. The calibration procedure involves a number of factors which contribute to the error in the calibration procedure. The different errors that were studied were those due to the initial warm-up time of the printer, measurements being made on different heads and on different sheets of paper and finally due to the measuring device that was used. It was determined that the printer had an initial warm-up time of one print. Since there was considerable variation between different heads, it was decided that the calibration be carried out separately for each head. The calibration procedure was carried out for the second head only. Since there was a variation between different sheets, an average of seven different sheets was used in the data collection. To average out errors due to the measuring device and to take care of variations under the same head, each color patch was averaged over five measurements on each sheet.

The data that was collected in this manner was then used to create the LUT. The bell function, which gave better results than the cubic B-spline was used in the interpolation routine. The interpolation error in creating the LUT was estimated for 343 points in the color gamut. An average ΔE error of about 2.05 was obtained as the interpolation error which is lower than the threshold value of 3, above which color differences can be detected. This value of 2.05 provides us with a lower bound on the errors obtained by the calibrated printer since additional interpolation error is introduced when trilinear interpolation is used to estimate the control values for points not on the grid. Once the LUT is created, we use trilinear interpolation to estimate the control values for points not on the grid. Using the same 343 points, a total average error of 2.19 is obtained which is again less than 3.

The data that was available was then used to estimate the SNR of the printer. The SNR calculation was done in different ways and these values are tabulated in Tables 4.3, 4.4 and 4.5. This estimate of the SNR value is then used in the mathematical simulation to generate Gaussian noise with the required power. It was seen that a large variance was observed in the interpolation errors using uncorrelated noise indicating that it is an inappropriate model. To reduce the variance, correlated noise was used which reduced the noise variance appreciably.

5.2 Conclusions

The XL7700 printer was calibrated for a single print head. The procedure is the same for the other three heads. 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 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.

Simulations using the mathematical model showed close resemblance to the performance of the printer. However, since white Gaussian noise was used to add noise in the model, the variance of the noise was quite large. This was reduced by using correlated noise. This showed that the mathematical model could be used to predict the performance of the printer.

5.3 Future Work

The work that has been carried out can be extended by looking at certain problems that come up in the course of the research. We have assumed that the printer that was calibrated possessed the property of temporal stability. However, in reality the colorimetric parameters may change over time thus requiring the printer to be calibrated every once in a while. To facilitate this process we should consider adaptive updating methods for the LUT.

Another interesting problem is to determine the correlation of quantitative error with subjective ratings of errors. We should try and extend the determination of error for a pixel of concern by using the colorimetric properties of the pixels that surround it. The method of calibration developed in this research could be used along with the improvements suggested to calibrate other color output devices.

Bibliography

[1] Andrews, H.C. and Patterson III, C.L., "Digital Interpolation of Images", *IEEE Transactions on Computers*, Feb 1976.

[2] Berns, R.S., Motta, R.J., and Gorzynski, M.E., "CRT Colorimetry Part I: Theory and Practice", *Color Research and Application*, Vol 18, Number 5, Oct 1993.

[3] Berns, R.S., Gorzynski, M.E., and Motta, R.J., "CRT Colorimetry Part II: Metrology", *Color Research and Application*, Vol 18, Number 5, pp 321-325, Oct 1993.

[4] CIE, "Recommendations on Uniform Color Spaces, Color-Difference Equations, Psychometric Color Terms", *Supplement No. 2 of CIE Publ. No. 15*(E-1.3.1) 1971, Bureau Central de la CIE, Paris, 1978.

[5] Dennis, J.E., and Schnabel, R.B., *Numerical Methods for Unconstrained Optimisation and Non-Linear Equations*, Englewood Cliffs, N.J., Prentice Hall, 1983.

[6] Huang, T.S., *Advances in Computer Vision and Image Processing*, Vol 1, Jai Press Inc, 1984.

[7] Hunt, R.W.G., *The Reproduction of Colour in Photography, Printing and Television*, Fountain Press, England, 1987.

[8] Kasson, J., Sigfredo, N., Plouffe, W., "Tetrahedral Interpolation Technique for Color Space Conversion", *Research Report IBM Research Division, Almaden Research Center*, pp 4-5, Jan 1993.

[9] Kodak XL7700 Digital Continuous Tone Printer, User's Manual.

[10] Mani, V., "Calibration of Color Monitors", Master's Thesis, Department of Electrical and Computer Engineering, North Carolina State University, 1991.

[11] Papoulis, A., "A New Algorithm in Spectral Analysis and Band-Limited Extrapolation", *IEEE Transactions on Circuits & Systems*, Vol. CAS-22, September 1975.

[12] Pratt, W.K., *Digital Image Processing*, John Wiley and Sons, 1991.

[13] Stamm, S., "An Investigation of Color Tolerance", *TAGA Proceedings*, 1981, pp156-173.

[14] Trussell, H.J., ECE 591I Class Notes, North Carolina State University, 1993.

[15] Trussell, H.J., "A Digital Signal Processing Approach to Color Systems", May 1992.

[16] Trussell, H.J., 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.

[17] Wyszecki, G. and Stiles, W.S., *Color Science: Concepts and Methods, Quantitative Data and Formulae*, John Wiley and Sons, 1982.

[18] Yule, J.A.C., *Principles of Color Reproduction*, John Wiley and Sons Inc, 1967.

Appendix A

Newton's Method when Derivatives are Unavailable

In one-dimension, our basic problem of solving a non-linear equation involves the following:

$$\text{given } f : \mathcal{R} \longrightarrow \mathcal{R}, \text{ find } x_* \in \mathcal{R} \text{ such that } f(x_*) \;=\; 0$$

In many practical applications, $f(x)$ is not given by a formula. Since $f'(x)$, the derivative of $f(x)$ is not available then, Newton's method must be modified to require only values of $f(x)$. In the case where the derivative is available, $f'(x)$ is used in modelling f near the current estimate x_c by the line tangent to f at x_c. When $f'(x)$ is unavailable, we replace the model by the secant line that goes through f at x_c, and some nearby point $x_c + h_c$. The slope of this line is:

$$a_c \;=\; \frac{f(x_c + h_c) - f(x_c)}{h_c} \tag{A.1}$$

So the model we obtain is the line [5]:

$$\hat{M}_c(x) \;=\; f(x_c) + a_c(x - x_c) \tag{A.2}$$

This is equivalent to replacing the derivative in the original model $M_c(x) = f(x_c) + f'(x_c)(x - x_c)$ by the approximation a_c.

In multi-dimensions, the problem reduces to the solution of a system of non-linear equations:

$$\text{given } \mathbf{F} : \mathcal{R}^n \longrightarrow \mathcal{R}^n, \text{ find } \mathbf{x}_* \in \mathcal{R}^n \text{ such that } \mathbf{F}(\mathbf{x}_*) \ = \ 0$$

In multiple dimensions, the analogous affine model is [5]:

$$\mathbf{M}_+(x) = \mathbf{F}(\mathbf{x}_+) + \mathbf{A}_+(\mathbf{x} - \mathbf{x}_c) \tag{A.3}$$

which satisfies $\mathbf{M}_+(\mathbf{x}_+) = \mathbf{F}(\mathbf{x}_+)$ for any $\mathbf{A}_+ \in \mathcal{R}^{nxn}$. In Newton's Method (where the derivative is available) $\mathbf{A}_+ = \mathbf{J}(\mathbf{x}_+)$, where $\mathbf{J}$ refers to the Jacobian. If $\mathbf{J}(\mathbf{x}_+)$ is not available, the extension of the one-dimensional secant method is $\mathbf{M}_+(\mathbf{x}_c) = \mathbf{F}(\mathbf{x}_c)$, that is,

$$\mathbf{F}(\mathbf{x}_c) = \mathbf{F}(\mathbf{x}_+) + \mathbf{A}_+(\mathbf{x}_c - \mathbf{x}_+) \tag{A.4}$$

or

$$\mathbf{A}_+(\mathbf{x}_+ - \mathbf{x}_c) = \mathbf{F}(\mathbf{x}_+) - \mathbf{F}(\mathbf{x}_c) \tag{A.5}$$

This is referred to as the secant equation. Let $\mathbf{s}_c = \mathbf{x}_+ - \mathbf{x}_c$ denote the current step and $\mathbf{y}_c = \mathbf{F}(\mathbf{x}_+) - \mathbf{F}(\mathbf{x}_c)$ denote the yield of the current step, so that the secant equation is written as:

$$\mathbf{A}_+\mathbf{s}_c = \mathbf{y}_c \tag{A.6}$$

We choose $\mathbf{A}_+$ by trying to minimize the change in the affine model, subject to satisfying equation A.6. The difference between the new and old affine models at any $\mathbf{x} \in \mathcal{R}^n$ is

$$\begin{aligned} \mathbf{M}_+(\mathbf{x}) - \mathbf{M}_c(\mathbf{x}) &= \mathbf{F}(\mathbf{x}_+) + \mathbf{A}_+(\mathbf{x} - \mathbf{x}_+) - \mathbf{F}(\mathbf{x}_c) - \mathbf{A}_c(\mathbf{x} - \mathbf{x}_c) \\ &= \mathbf{F}(\mathbf{x}_+) - \mathbf{F}(\mathbf{x}_c) - \mathbf{A}_+(\mathbf{x}_+ - \mathbf{x}_c) + (\mathbf{A}_+ - \mathbf{A}_c)(\mathbf{x} - \mathbf{x}_c) \end{aligned} \tag{A.7}$$

using equation A.5, we get

$$\mathbf{M}_+(\mathbf{x}) - \mathbf{M}_c(\mathbf{x}) = (\mathbf{A}_+ - \mathbf{A}_c)(\mathbf{x} - \mathbf{x}_c) \tag{A.8}$$

For any $\mathbf{x} \in \mathcal{R}^n$, let

$$\mathbf{x} - \mathbf{x}_c = \alpha\mathbf{s}_c + \mathbf{t} \tag{A.9}$$

where $\mathbf{t}^T\mathbf{s} = 0$. Then the term we wish to minimize becomes

$$\mathbf{M}_+(\mathbf{x}) - \mathbf{M}_c(\mathbf{x}) = \alpha(\mathbf{A}_+ - \mathbf{A}_c)\mathbf{s}_c + (\mathbf{A}_+ - \mathbf{A}_c)\mathbf{t} \tag{A.10}$$

We have no control over the first term on the right side, since the secant equation implies $(\mathbf{A}_+ - \mathbf{A}_c)\mathbf{s}_c = \mathbf{y}_c - \mathbf{A}_c\mathbf{s}_c$ However, we can make the second term zero for all $\mathbf{x} \in \mathcal{R}^n$ by choosing $\mathbf{A}_+$ such that $(\mathbf{A}_+ - \mathbf{A}_c)\mathbf{t} = 0$ for all $\mathbf{t}$ orthogonal to $\mathbf{s}_c$. This gives us

$$\mathbf{A}_+ = \mathbf{A}_c + \frac{(\mathbf{y}_c - \mathbf{A}_c\mathbf{s}_c)\mathbf{s}_c^T}{\mathbf{s}_c^T\mathbf{s}_c} \tag{A.11}$$

as the least change in the affine model consistent with the secant equation. This update was proposed by C.Broyden and is known as the *Broyden's update* or the *secant update*. The resultant algorithm for Broyden's method is as follows:

Given $\mathbf{F} : \mathcal{R}^n \longrightarrow \mathcal{R}^n, \mathbf{x}_0 \in \mathcal{R}^n, \mathbf{A}_0 \in \mathcal{R}^{nxn}$

where $\mathbf{x}_0$ is the initial approximation to the solution

and $\mathbf{A}_0$ is the initial approximation to the Jacobian

Do for $k = 0, 1, ...$

Solve $\mathbf{A}_k\mathbf{s}_k = -\mathbf{F}(\mathbf{x}_k)$ for $\mathbf{s}_k$

$\mathbf{x}_{k+1} := \mathbf{x}_k + \mathbf{s}_k$

$\mathbf{y}_k := \mathbf{F}(\mathbf{x}_{k+1}) - \mathbf{F}(\mathbf{x}_k)$

$\mathbf{A}_{k+1} := \mathbf{A}_k + \frac{(\mathbf{y}_k - \mathbf{A}_k\mathbf{s}_k)\mathbf{s}_k^T}{\mathbf{s}_k^T\mathbf{s}_k}$

We use finite differences to get the initial approximation $\mathbf{A}_0$ to the Jacobian $\mathbf{J}(\mathbf{x}_0)$.

Appendix B

Trilinear Interpolation

In trilinear interpolation, the sample function values may be arranged into a three dimensional table which are indexed by the independent variables [8], e.g. x, y and z. The range of each input variable is usually evenly sampled. Let x be in the range $x_0 \ldots x_a$, y be in the range $y_0 \ldots y_b$, and z be in the range $z_0 \ldots z_c$. One possible sampling would be (x_i, y_j, z_k), where $0 \le i \le a$, $0 \le j \le b$ and $0 \le k \le c$ and

$$x_i = x_0 + \frac{x_a - x_0}{a} i \tag{B.1}$$

$$y_j = y_0 + \frac{y_b - y_0}{b} j \tag{B.2}$$

$$z_k = z_0 + \frac{z_c - z_0}{c} k \tag{B.3}$$

Then the function F, may be approximated for a target point (r, s, t) as follows:

$$\begin{aligned} F(r,s,t) \;=\; & (1-d_r)\left\{ \begin{array}{l} (1-d_s)[(1-d_t)F(x_i,y_j,z_k) + d_t F(x_i,y_j,z_{k+1})] \\ +d_s[(1-d_t)F(x_i,y_{j+1},z_k) + d_t F(x_i,y_{j+1},z_{k+1})] \end{array} \right\} \\ & +d_r \left\{ \begin{array}{l} (1-d_s)[(1-d_t)F(x_{i+1},y_j,z_k) + d_t F(x_{i+1},y_j,z_{k+1})] \\ +d_s[(1-d_t)F(x_{i+1},y_{j+1},z_k) + d_t F(x_{i+1},y_{j+1},z_{k+1})] \end{array} \right\} \end{aligned} \tag{B.4}$$

where

$$x_i \leq r < x_{i+1} \tag{B.5}$$

$$y_j \leq s < y_{j+1} \tag{B.6}$$

$$z_k \leq t < z_{k+1} \tag{B.7}$$

and

$$d_r = \frac{r - x_i}{x_{i+1} - x_i} \tag{B.8}$$

$$d_s = \frac{s - y_j}{y_{j+1} - y_j} \tag{B.9}$$

$$d_t = \frac{t - z_k}{z_{k+1} - z_k} \tag{B.10}$$

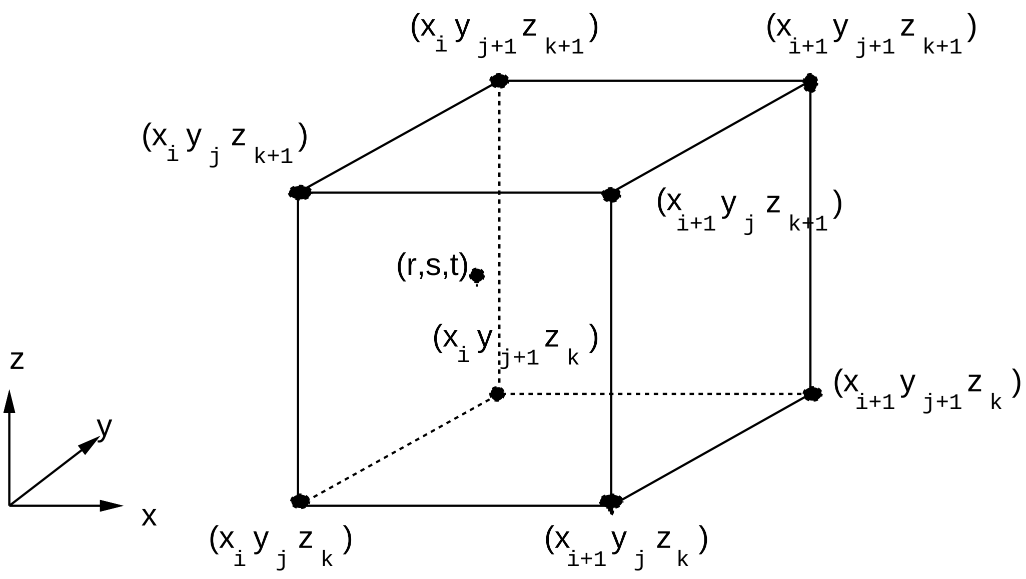

Figure B.1: Trilinear Interpolation Geometry