Evaluating the 1931 CIE Color Matching Functions

Mark Q. Shaw

B.Sc. Graphic Media Studies,

The University of Hertfordshire, England (1997)

Evaluating the 1931 CIE Color Matching Functions

Mark Q. Shaw

B.Sc. Graphic Media Studies,
The University of Hertfordshire, England (1997)

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Color Science
in the Center of Imaging Science
Rochester Institute of Technology

June 1999

_____

Signature of the Author

_____

Accepted by

Coordinator, M.S. Degree Program

CENTER FOR IMAGING SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Mark Q. Shaw
has been examined and approved
by two members of the color science faculty
as satisfactory for the thesis requirement for the
Master of Science degree

Dr. Mark Fairchild, Thesis Advisor

Dr. Ethan Montag

THESIS RELEASE PERMISSION FORM

Rochester Institute of Technology

Center For Imaging Science

Title of Thesis

Evaluating the 1931 CIE Color Matching Functions

I, Mark Q. Shaw, hereby grant permission to the Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or part. Any reproduction will not be for commercial use or profit.

_____

Signature of Author

_____

Date

Evaluating the 1931 CIE Color Matching Functions

Mark Q. Shaw

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Color Science in the
Center of Imaging Science, Rochester Institute of Technology

Abstract

The use of colorimetry within industry has grown extensively in the last few decades. Central to many of today's instruments is the work of the CIE system, established in 1931. Many have questioned the validity of the assumptions made by Wright (1928-29) and Guild (1931), some suggesting that the 1931 color matching functions are not the best representation of the human visual systems' cone responses.

A computational analysis was performed to evaluate the 1931 color matching functions against other responsivity functions using metameric data. The underlying principle is that an optimal set of responsivity functions will yield minimal color difference errors between pairs of visually matched metamers. The difference of average color differences found in the six chosen sets of responsivity functions were small. The CIE 1931 2° color matching functions, on average, provided the largest color difference, 4.56 ΔE^*_{ab}. With the best performance coming from the CIE 1964 10° color matching functions, yielding an average color difference of 4.02 ΔE^*_{ab}.

An optimization was then performed on the CIE 1931 color matching functions. The concept was that color differences between metamers can be used to improve predictions of color matching functions. If one is to take all pairs, and perform an optimization that globally minimizes the average color difference, then one can hope to obtain an optimal set of responsivity functions. The optimum solution was to use a weighted combination of each of the different sets of responsivity functions. The optimized set, the 'Shaw and Fairchild' responsivity functions, were able to reduce the average color difference down to 3.92 ΔE^*_{ab}.

The final part of the work was to build a computer based simulation of the color differences between the different sets of responsivity functions. This simulation allows a user to load a spectral radiance, or reflectance, data file and display the tristimulus match predicted by each of the seven sets of responsivity functions.

Acknowledgements

I am sincerely grateful to the following people for their time and support in completion of this thesis:

Dr. Mark Fairchild for his wisdom, guidance, and motivation. It has been a very valuable experience.

Dr. Noboru Ohta for his guidance in performing the computational analysis, and general encouragement.

Dr. Roy Berns, Dr. Ethan Montag, Dave Wyble, Lisa Reniff, and Colleen Desimone for their encouragement, guidance and relentless homework's. It has been a thoroughly enjoyable, unforgettable experience.

Sally, first my girlfriend, then fiancée, and now my wife. For the love, support, encouragement, and friendship she has shown me throughout this time.

My friends and family for their support and encouragement.

Thank you all!

TABLE OF CONTENTS

TABLE OF CONTENTS I

TABLE OF TABLES IV

TABLE OF FIGURES V

1 INTRODUCTION 1

1.1 COLORIMETRY 3

1.1.1 CIE 1931 Color Matching Functions 4

1.1.2 Stiles and Burch Color Matching Functions 6

1.1.3 Thornton Color Matching Functions 8

1.2 CONE SPECTRAL SENSITIVITIES 9

1.2.1 Vos and Walraven 10

1.2.2 Smith and Pokorny 12

1.2.3 Stockman, MacLeod and Johnson 14

1.2.4 Stockman and Sharpe 15

1.3 CURRENT RESEARCH 17

2 RESPONSIVITY FUNCTION SELECTION 18

2.1 SELECTED SETS 18

2.1.1 CIE 1931 Color Matching Functions 18

2.1.2 CIE 1964 Color Matching Functions 19

2.1.3 Stiles and Burch Color Matching Functions 20

2.1.4 Demarco, Smith and Pokorny Cone Fundamentals 20

2.1.5 Stockman, MacLeod and Johnson Cone Fundamentals 21

2.1.6 Vos and Walraven Cone Fundamentals 22

2.2 Rejected Responsivity Functions 22

3 Experimental Data Sources 24

3.1 Experiment 1 – Alfvin and Fairchild (1995) 24

3.2 Experiment 2 – Shaw and Montag (1998) 26

3.3 Experiment 3 – Shaw and Montag (1999) 29

3.4 Experimental Discussion 30

4 Comparing the Accuracy of Various Data Sets 31

4.1 Rotation to near CIE space 32

4.2 Tristimulus Calculation 36

4.3 CIELAB Representation 37

4.4 Observer Variability 39

4.5 Statistical Analysis 39

4.5.1 Students t-test 39

4.5.2 Multivariate 95% Confidence Ellipse 40

5 Computational Analysis Results and Discussion 43

5.1 Alfvin Data Set 43

5.1.1 Combined Alfvin Data Set 43

5.1.2 Segmented Alfvin Data Set 46

5.2 Lab '98 and '99 Data Sets 47

5.3 All Three Data Sets Combined 48

5.4 Thornton Metameric Pair 50

5.5 Derivation of Optimized Responsivity Functions 51

6 Deriving an Optimized Set of Responsivity Functions 52

6.1 Techniques Used 52

6.1.1 Unconstrained Non-linear optimization, delta XYZ function 54

6.1.2 Constrained Non-linear optimization, delta XYZ function 55

6.1.3 Constrained Non-linear optimization, using 18 cubic spline functions, delta XYZ function 57

6.1.4 Constrained Non-linear optimization, baseline 0, using 18 cubic spline functions, delta XYZ function..... 59

6.1.5 Constrained Non-linear optimization, using 18 cubic spline functions, baseline 0, direct XYZ functions..... 61

6.1.6 Constrained Non-linear optimization, using 4 cubic spline functions, delta XYZ function 62

6.1.7 Linear Regression, using 18 cubic spline functions, delta XYZ function..... 64

6.1.8 Linear Regression, using PCA basis vectors, delta XYZ function 66

6.1.9 Linear Regression – I, using Sine and Cosine basis vectors, delta XYZ function... 69

6.1.10 Linear Regression – II, using Sine and Cosine basis vectors, delta XYZ function 71

6.1.11 Constrained Non-linear optimization, sum eq. 1, using multiple color matching functions, direct XYZ functions 73

6.1.12 Non-linear optimization, using CIE Standard Deviate Observer functions, delta XYZ functions..... 75

6.1.13 Monte Carlo Method, using Sine and Cosine basis vectors, delta XYZ function ... 77

6.2 OPTIMIZATION SUMMARY 78

7 COMPUTER SIMULATION 81

7.1 SPECTRAL CHARACTERIZATION OF CRT 82

7.2 SOFTWARE SIMULATION PACKAGE..... 84

7.3 USER EXPERIENCES 88

REFERENCES 92

TABLE OF TABLES

TABLE 5.1 : AVERAGE RESULTS OF COMPUTATIONAL ANALYSIS FOR ALL COLOR CENTERS AND OBSERVERS IN THE ALFVIN DATA SET. 44

TABLE 5.2 : MEAN COLOR DIFFERENCE (ΔE^*_{AB}) RESULTS OF COMPUTATIONAL ANALYSIS USING THE SEGMENTED ALFVIN DATA SET. MEAN OF MEANS IS THE MEAN COLOR DIFFERENCE OF THE AVERAGED RESULTS FOR EACH COLOR CENTER, GIVING AN INDICATION TO THE OVERALL PERFORMANCE OF EACH SET OF RESPONSIVITY FUNCTIONS. 46

TABLE 5.3 : AVERAGE RESULTS OF COMPUTATIONAL ANALYSIS FOR THE LAB '98 DATA SET. 47

TABLE 5.4 : AVERAGE RESULTS OF COMPUTATIONAL ANALYSIS FOR THE LAB '99 DATA SET. 47

TABLE 5.5 : AVERAGE RESULTS FOR THE COMPUTATIONAL ANALYSIS OF ALL THREE DATA SETS COMBINED 48

TABLE 5.6 : RESULTS FOR THE COMPUTATIONAL ANALYSIS USING THE THORNTON METAMER PAIR 50

TABLE 6.1 : RESULTS OF LINEAR REGRESSION WITH PCA BASIS VECTORS. IT IS EVIDENT THAT THE OPTIMIZATION MINIMIZED THE TRISTIMULUS DIFFERENCE VERY WELL, BUT THAT DOES NOT REPRESENT AN ADEQUATE MINIMIZATION OF COLOR DIFFERENCE. 67

TABLE 6.2 : RESULTS OF LINEAR REGRESSION WITH SINE AND COSINE BASIS VECTORS. IT IS EVIDENT THAT THE OPTIMIZATION MINIMIZED THE TRISTIMULUS DIFFERENCE VERY WELL, BUT THAT DOES NOT REPRESENT AN ADEQUATE MINIMIZATION OF COLOR DIFFERENCE. 71

TABLE 6.3 : RESULTS OF LINEAR REGRESSION WITH SINE AND COSINE BASIS VECTORS. IT IS EVIDENT THAT THE OPTIMIZATION MINIMIZED THE TRISTIMULUS DIFFERENCE VERY WELL, BUT THAT DOES NOT REPRESENT AN ADEQUATE MINIMIZATION OF COLOR DIFFERENCE. 72

TABLE 6.4 : SUMMARY TABLE OF THE RESULTS OF EACH DIFFERENT SET OF OPTIMIZED WEIGHTING FUNCTIONS. 79

TABLE OF FIGURES

FIGURE 1.1 : SCHEMATIC REPRESENTATION OF THE ZONE MODEL OF COLOR VISION 11

FIGURE 2.1 : SPECTRAL RESPONSIVITY OF THE CIE 1931 COLOR MATCHING FUNCTIONS 19

FIGURE 2.2 : SPECTRAL RESPONSIVITY OF THE CIE 1964 COLOR MATCHING FUNCTIONS 19

FIGURE 2.3 : SPECTRAL RESPONSIVITY OF THE STILES AND BURCH COLOR MATCHING FUNCTIONS 20

FIGURE 2.4 : SPECTRAL RESPONSIVITY OF THE DEMARCO, SMITH AND POKORNY CONE FUNDAMENTALS 21

FIGURE 2.5 : SPECTRAL RESPONSIVITY OF THE STOCKMAN, MACLEOD AND JOHNSON CONE FUNDAMENTALS 21

FIGURE 2.6 : SPECTRAL RESPONSIVITY OF VOS AND WALRAVEN CONE FUNDAMENTALS 22

FIGURE 2.7 : THE FOUNDATION OF VARIOUS SETS OF CONE FUNDAMENTALS 23

FIGURE 3.1 : SCHEMATIC OVERHEAD VIEW OF ALFVIN EXPERIMENTAL SETUP USED FOR COLOR MATCHING 25

FIGURE 3.2 : THE SPECTRAL POWER DISTRIBUTIONS OF A GRAY TRANSPARENCY AND A CRT DISPLAY ADJUSTED TO YIELD A METAMERIC MATCH FOR AN OBSERVER. 26

FIGURE 3.3 : ACS VCS 10 ADDITIVE MIXTURE DEVICE 27

FIGURE 3.4 : ACS VCS 10 IN USE. COLORED PRIMARY DISC CONTRIBUTIONS WERE CONTROLLED BY THE USER CONTROL PANEL. METAMERIC MATCH SPECTRA ARE TAKEN USING PHOTORESEARCH PR650 MOUNTED BY OBSERVERS RIGHT SHOULDER. 28

FIGURE 3.5 : THE SPECTRAL POWER DISTRIBUTION OF THE MUNSELL N5 PATCH AND ACS VCS ADJUSTED TO YIELD A METAMERIC MATCH UNDER DAYLIGHT FLUORESCENT ILLUMINATION. 29

FIGURE 3.6 : THE SPECTRAL POWER DISTRIBUTION OF THE FUJIX PATCH AND ACS VCS ADJUSTED TO YIELD A METAMERIC MATCH UNDER DAYLIGHT FLUORESCENT ILLUMINATION. 30

FIGURE 4.1 : CONCEPTUALLY IDEAL SET OF COLOR MATCHING FUNCTIONS RESULTING IN EQUAL INTEGRAL RESPONSE OF THE CONES 31

FIGURE 4.2 : ROTATION OF THE STILES AND BURCH COLOR MATCHING FUNCTIONS INTO A NEAR CIE REPRESENTATION 33

FIGURE 4.3 : ROTATED CIE 1964 COLOR MATCHING FUNCTIONS (SOLID LINE) PLOTTED OVER CIE 1931 COLOR MATCHING FUNCTIONS (MARKED LINE) 34

FIGURE 4.4 : ROTATED DEMARCO, SMITH, AND POKORNY CONE FUNDAMENTALS (SOLID LINE) PLOTTED OVER CIE COLOR MATCHING FUNCTIONS (MARKED LINE) 35

FIGURE 4.5 : ROTATED STOCKMAN, MACLEOD, AND JOHNSON CONE FUNDAMENTALS (SOLID LINE) PLOTTED OVER CIE COLOR MATCHING FUNCTIONS (MARKED LINE) 35

FIGURE 4.6 : ROTATED VOS AND WALRAVEN CONE FUNDAMENTALS (SOLID LINE) PLOTTED OVER CIE COLOR MATCHING FUNCTIONS (MARKED LINE) 36

FIGURE 4.7 : CIELAB A* B* PLOT OF ALL THREE COMBINED DATA SETS 37

FIGURE 4.8 : CIELAB L* A* PLOT OF ALL THREE COMBINED DATA SETS 38

FIGURE 4.9 : CIELAB L* B* PLOT OF ALL THREE COMBINED DATA SETS 38

FIGURE 4.10 : 95% AND MEAN CONFIDENCE ELLIPSES FOR COMBINED DATA SET 41

FIGURE 5.1 : CONFIDENCE ELLIPSE PLOTS OF THE SHAW AND FAIRCHILD RESPONSIVITY FUNCTIONS, USING THE WHOLE ALFVIN DATA SET. 95% CONFIDENCE REGION - OUTER ELLIPSE, MEAN – INNER ELLIPSE 45

FIGURE 5.2 CONFIDENCE ELLIPSE PLOTS OF THE LAB 98 DATA SET, USING THE CIE 1931 STANDARD OBSERVER. THE TWO 'CLOUDS' OF DATA POINTS ARE THOUGHT TO REPRESENT THE COLOR MATCHES ATTAINED BY THE DIFFERENT PRIMARY SETS USED. 48

FIGURE 5.3 : 95% AND MEAN CONFIDENCE ELLIPSES OF SHAW AND FAIRCHILD RESPONSIVITY FUNCTIONS, USING ALL THREE COMBINED DATA SETS 49

FIGURE 6.1 : OPTIMIZED WEIGHTING FUNCTIONS USING THE POWELL METHOD WITHOUT CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE 1931 WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 55

FIGURE 6.2 : OPTIMIZED WEIGHTING FUNCTIONS USING THE POWELL METHOD WITH CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 56

FIGURE 6.3 : OPTIMIZED WEIGHTING FUNCTIONS USING THE LEVENBERG-MARQUARDT METHOD WITH CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 58

FIGURE 6.4 : OPTIMIZED WEIGHTING FUNCTIONS USING THE LEVENBERG-MARQUARDT METHOD WITH CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 60

FIGURE 6.5 : OPTIMIZED WEIGHTING FUNCTIONS USING THE LEVENBERG-MARQUARDT METHOD WITH CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. GENERATED USING 18 CONSTRAINED CUBIC SPLINES. 62

FIGURE 6.6 : 4 CUBIC SPLINE FUNCTIONS 63

FIGURE 6.7 : OPTIMIZED WEIGHTING FUNCTIONS USING THE LEVENBERG-MARQUARDT METHOD. GENERATED USING 4 CONSTRAINED CUBIC SPLINES, WITH CONSTRAINT. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 64

FIGURE 6.8 : DELTA FUNCTION DERIVED USING LINEAR REGRESSION WITH WEIGHTED CUBIC SPLINES. 65

FIGURE 6.9 : (TOP) VARIANCE EXPLAINED BY EACH OF THE EIGENVECTORS IN THE DELTA SPECTRA. IT CAN BE SEEN THAT THE DATA IS HIGHLY DEPENDENT, AND THE MAJORITY OF VARIANCE CAN BE EXPLAINED BY THE LAST 4 VECTORS. (BOTTOM) PLOTS OF THE 4 MOST SIGNIFICANT EIGENVECTORS. 66

FIGURE 6.10 : OPTIMIZED WEIGHTING FUNCTIONS USING LINEAR REGRESSION WITH PCA BASIS VECTORS. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 68

FIGURE 6.11 : SINE AND COSINE WAVES TO BE USED AS BASIS VECTORS 69

FIGURE 6.12 : OPTIMIZED WEIGHTING FUNCTIONS USING LINEAR REGRESSION WITH SINE AND COSINE BASIS VECTORS. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 70

FIGURE 6.13 : OPTIMIZED WEIGHTING FUNCTIONS USING LINEAR REGRESSION WITH SINE AND COSINE BASIS VECTORS. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 73

FIGURE 6.14 : OPTIMIZED WEIGHTING FUNCTIONS USING MULTIPLE SETS OF COLOR MATCHING FUNCTIONS. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 74

FIGURE 6.15 : OPTIMIZED WEIGHTING FUNCTIONS USING THE CIE STANDARD DEVIATE OBSERVER. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION 76

FIGURE 6.16 : OPTIMIZED WEIGHTING FUNCTIONS USING SINE AND COSINE SPLINES AND THE MONTE CARLO METHOD. IN EACH CASE THE BLACK LINE REPRESENTS THE CIE STANDARD WEIGHTING FUNCTION, THE COLORED LINE THE OPTIMIZED FUNCTION. 78

FIGURE 6.17 : FINAL SELECTION OF OPTIMIZED COLOR WEIGHTING FUNCTIONS 80

FIGURE 7.1 : CONCEPTUAL VISUAL EXPERIMENT. OBSERVERS COMPARISON OF VARIOUS SETS OF RESPONSIVITY FUNCTIONS..... 81

FIGURE 7.2 : CHARACTERIZATION OUTLINE OF CRT DISPLAY USING MULTIPLE SETS OF RESPONSIVITY FUNCTIONS..... 82

FIGURE 7.3 : COMPUTATIONAL PREDICTION OF AN INPUT SPECTRA USING THE 7 RESPONSIVITIES 83

FIGURE 7.4 : SIMULATION WELCOME SCREEN 84

FIGURE 7.5 : MAIN CONTROL WINDOW - USER SELECTION OF SPECTRAL DATA FILE. USER SELECTS EITHER RADIANCE OR REFLECTANCE SPECTRAL DATA, AND CAN THEN LOAD A DATA FILE FROM DISC..... 85

FIGURE 7.6 : MAIN CONTROL WINDOW - USER SELECTION OF SPECTRAL DATA FILE. WHEN REFLECTANCE IS SELECTED THE USER MUST ALSO SPECIFY THE ILLUMINANT UNDER WHICH THE SAMPLE IS VIEWED. 85

FIGURE 7.7 : PREDICTION WINDOW - SIMULATION PREDICTION OF EACH OF THE 7 RESPONSIVITY FUNCTIONS. IT CAN BE SEEN THAT SUBTLE DIFFERENCES ARE NOTICEABLE. 86

FIGURE 7.8 : PREDICTION WINDOW - WHEN A TRISTIMULUS VALUE TO BE DISPLAYED IS OUT OF GAMUT OF THE CRT, A GAMUT WARNING IS PRESENTED TO THE USER INSTEAD. 87

1 INTRODUCTION

The use of colorimetry within industry has grown extensively in last few decades. The cost of hardware is constantly decreasing, and the ‘quality drive’ pushing companies to optimize productivity have been strong contributory factors. Although the technology is only really still in its infancy, many take for granted that instrumental tolerances will suffice and human interaction can be minimized.

Central to many of today’s instruments is the work of the CIE system (Commission Internationale de l'Éclairage) established in 1931. The CIE system allows the specification of color matches for a standard observer using color matching functions. These color matching functions of normal human observers are the fundamental basis of colorimetry. Studies performed by Guild (1931) and Wright (1928-29) are central to the work of the CIE, providing the foundation for the derivation of the CIE 1931 standard observer. Since the original work of the CIE, the standard has withstood an onslaught of technical pressures and remained a useful international standard for many years (Fairchild, 1998).

Stockman and Sharpe (1998), Stockman, MacLeod and Johnson (1993), and others, have questioned the validity of the assumptions made by Wright (1928-29) and Guild (1931). Stockman and Sharpe (1998) suggest that the 1931 color matching functions are not the best representation of the human visual systems’ cone responses. They were constructed from the relative color matching data of Wright (1928-29) and Guild (1931), with the assumption that the color matching functions must be a linear combination of the 1924

CIE V(λ) luminous efficiency function. The implications of the questions are not known, one could consider that the concern be purely one of theoretical rigor, bearing no significance in the application of the current technology. Or, should the concern be of valid cause, one that may affect the whole of colorimetry as it stands to date.

If one is to assume that the CIE 1931 color matching functions are sufficient when they are not, then much of the work being done to build device models, minimizing colorimetric error is attempting to attain an impossible goal. If the CIE 1931 color matching functions are not the ideal solution, would a new set lead to better colorimetry? Or is the error within the bounds of statistical insignificance?

This question is of great importance to the color science community. Applications that assume that the CIE 1931 color matching functions are sufficient, if they are not, will inherently lead to a systematic error throughout all applications of colorimetry.

Existing works clearly document some of the issues that are cause for concern when using the CIE color matching functions and their impending effects on the color community. But, although many are quick to comment on their opinions and suggest new cone fundamentals or color matching functions, little documented work has been done to actually compare the benefits of using the modified functions and sensitivities.

It is the aim of this work to evaluate the accuracy of the CIE 1931 color matching functions against more recent estimations and modifications. To determine whether the new cone fundamentals / color matching function derivations are truly better, or merely within the bounds of statistical insignificance.

The objective of the work is to undergo a computational analysis with different sets of color matching functions and cone fundamentals using existing measurements of

metameric pairs. The computation will determine which set best approximates the observers' visual perception.

The foundation upon which the theory of color matching functions are based, and the different sets of cone fundamentals / color matching functions are summarized to provide the reader a background to the work.

1.1 Colorimetry

At its birth, it was clear that CIE colorimetry was seen primarily as a means of identifying and specifying colors, and for defining color standards. Few envisaged colorimetry and the CIE system as a working tool in the color industries (Wright, 1981). Following the establishment of the CIE 1924 luminous efficiency function ($V\lambda$), attention was turned to the development of a system of colorimetry. At that time cone fundamentals were not available, therefore a system of colorimetry was developed based on the principles of trichromacy and Grassmann's laws of additive color mixture (Fairchild, 1998, Wright, 1981).

The concept of the system is that any color can be matched by an additive mixture of three primary colors. The primaries are defined as a set of three lights, either real or imaginary that cannot be matched by any additive combination of the other two.

$$C \equiv R(\mathcal{R}) + G(\mathcal{G}) + B(\mathcal{B}) \qquad (1.1)$$

Any color, C, can therefore be represented by a trichromatic equation (Eqn 1.1), where $\mathcal{R}$, $\mathcal{G}$, $\mathcal{B}$ are the primaries and R, G, B are the amounts of each primary needed to match a given color. These amounts are called *tristimulus values*. The tristimulus values are extended such that they can be obtained for any given stimulus, defined by a spectral

power distribution. This extension takes advantage of Grassmann's laws of additivity and proportionality to sum tristimulus values for each spectral component of a stimuli (Fairchild, 1998).

Wyszecki surmises that basic colorimetry is governed by the experimental laws of color matching valid for all human observers with normal color vision. From these experimental laws three equations evolve which express the conditions for a color match between two given color stimuli, shown in Equation 1.2

$$\begin{aligned} \int_{\lambda} S_1(\lambda)\bar{x}(\lambda)d\lambda &= \int_{\lambda} S_2(\lambda)\bar{x}(\lambda)d\lambda \\ \int_{\lambda} S_1(\lambda)\bar{y}(\lambda)d\lambda &= \int_{\lambda} S_2(\lambda)\bar{y}(\lambda)d\lambda \\ \int_{\lambda} S_1(\lambda)\bar{z}(\lambda)d\lambda &= \int_{\lambda} S_2(\lambda)\bar{z}(\lambda)d\lambda \end{aligned} \qquad (1.2)$$

Where S_1 and S_2 are the spectral power distributions of the two color stimuli compared by the observer, and $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ define the color matching properties of the observer (Wyszecki, 1973).

1.1.1 CIE 1931 Color Matching Functions

The CIE 1931 standard observer is based on two independent experiments performed by Guild (1931) and Wright (1928-29) that measured the chromaticity coordinates for a total of 17 observers. Wright (1928-29) used monochromatic primaries, whereas Guild (1931) used broadband primaries. Since the primaries from one experiment can be specified in terms of tristimulus values the other experiment, a linear transformation (3x3 matrix transform) is possible to change the results from the one set of primaries into the other. Thus a transformation was derived to convert both Wright's and Guild's data into a set of common $\mathcal{RGB}$ primaries (Fairchild, 1998).

In order to eliminate the negative values in the color matching functions, the CIE transformed to a set of imaginary primaries, $\mathfrak{XYZ}$. The goal of the transformation to eliminate the negative portions of the color matching functions and to force one of the color matching functions to equal the CIE 1924 photopic luminous efficiency function, $V(\lambda)$. Forcing one of the color matching functions to equal the $V(\lambda)$ function serves the purpose to incorporate the CIE system of photometry into the CIE system of colorimetry.

The validity of such decisions are still in much debate to date. Wright (1981) suggests that no reference can be found to additivity experiments in the papers of Ives (1923) and Guild (1931) on the transformation of color mixture data, in spite of the fact that the assumption is fundamental to the legitimacy of such transformations. He writes that Guild, in his 1926 Survey, states: “Newton’s law of mixtures follows directly from the geometry of the color triangle. The law is that if any two colours are mixed, the colour produced lies on the line joining the two constituents on the colour chart and dividing it inversely as the quantities in the mixture.” Guild was either relying on Newton’s experiments to justify the principle, or claiming that the geometry of the colour triangle proved the law (Wright, 1981). Wright then goes on to add that the additivity principle has been studied since 1931, and found to hold reasonably well for 2° color matching. Although it has been shown, however, that if light adapted to level above about 10,000 to 15,000 trolands, then color matches break down, implying additivity failure at this level.

In 1951, Judd proposed a revision to the 1931 CIE standard observer, (Judd, 1951) concluding that the CIE had given weight to early measurements that led to low average values at short wavelengths. Vos (1978) has further refined Judd’s modification by (i) making use of more precise computational procedures than that available to Judd in 1951, (ii) extending the data further into the far red, (iii) reducing by 0.2 log units the values of Judd’s modified $V(\lambda)$ function for wavelengths below 410nm, making them follow more

closely to the values obtained by Stiles (1955), (iv) slightly smoothing the CIE color matching functions from 380 to 400nm and (v) truncating the color matching functions at 380nm, because the CIE 1931 data below that wavelength are extrapolated (Vos, 1978).

Wright believes that the integrity of the CIE 1924 luminous efficiency function V(λ) being incorporated into the color matching functions has not proved to be so error free (Wright, 1981). Stockman and Sharpe (1998) re-iterate this concern, "The CIE functions were constructed from the relative color matching data of Wright and Guild with the assumption that the CMFs must be a linear combination of the V(λ) function. Not only is the validity of the V(λ) curve questionable, even after the corrections of Judd and Vos have been applied, but so too is the assumption that V(λ) must be a linear combination of the CMFs."

1.1.2 Stiles and Burch Color Matching Functions

Twenty years passed between the standardization of the 1931 standard observer and a new investigation to re-determine the color matching functions, performed by Stiles in the 1950's. The investigation was encouraged by the CIE Colorimetry Committee, having expressed concern for the correctness of the 1931 standard observer data. Specific concern surrounded (i) the use of the 1924 standard luminous efficiency function V(λ) at the ends of the spectrum, (ii) discrepancies between perceived and measured color differences of certain titanium pigments observed and computed from the CIE 1931 data, and (iii) interest in applicability to color matching in large visual fields used for industrial production control (Stiles *et al.,* 1955, Wyszecki and Stiles, 1982). Their concerns are very valid, especially that of (ii), indicating that the standard observer does not work for the titanium pigments under evaluation.

Stiles' investigation began with a pilot study involving a group of 10 observers that made color matches and heterochromatic brightness matching in 2° and 10° visual bipartite fields on a specially designed colorimeter, known as the Stiles (or NPL) trichromator. (Stiles *et al.,* 1955) The report showed that there were experimentally significant differences between the 2° pilot data and the standard observer, but were not large enough to warrant a change in the standard data for practical colorimetry. (Wyszecki and Stiles, 1982).

Although no further work was done by Stiles on 2° color matching, his 2° pilot data provide a valuable set of color matching functions. Others differ in the opinion of the CIE, Stockman and Sharpe (1998) prefer either the 2° or 10° data because they were directly measured and are uncontaminated by changes introduced by standards committees. (Stockman *et al.,* 1998)

Estevez (1979) performed a study of the data from each of the investigations for the 2° observer because he felt that the CIE standard observer imposed additional processing to the experimental data that may have added error to the experimentation performed by Wright and Guild. He also documents correcting for a calibration error in the original study, and concludes

> "The conclusion one draws from all of these comparisons is straightforward: In *no* case can a photometric function be a linear combination of CMFs ... Most important, the CIE V(λ) is not itself a linear combination of CMFs, and therefore the synthesized CIE CMF must be in error, especially in the spectral region around 465nm." (Estevez, 1982)

But, Estevez does also conclude on the basis of his observations that the three studies (Wright, Stiles and Burch) agree with each other and, therefore represent the same average normal trichromatic observer.

1.1.3 Thornton Color Matching Functions

In a series of articles covering over 20 years of research, Thornton discusses his work 'toward a more accurate and extensible colorimetry.' (1992abc, 1997, 1998ab) Early work focussed on determining an observer's color matching functions using several disparate sets of primary lights for each observer. The visual colorimeter-spectro-radiometer utilized bipartite 10° visual fields, natural pupil, dim surround, binocular viewing, and adaptation to a steady match-point. Both maximum saturation and Maxwell method visual matches were used, resulting in large tristimulus error when computed using the 1964 CIE standard observer (Thornton *et al.*, 1998a). Thornton's tests included an evaluation of Grassmann's Additivity Law, resulting in small visual differences between the summed colors being often seen, but with low consistency among the 3 observers. He concludes that Grassmann's additivity law holds approximately, since only small visual mismatches were seen. However, the 1964 Standard Observer showed large systematic errors, and again showed itself to be non-representative of normal human observers (Thornton, 1992a). In a later article (Thornton *et al.,* 1998a), he summarizes his earlier findings, stating that the CIE standard Observers represent human vision no better at high brightness than at low, and no better with small visual fields than with large.

In his quest to find a more accurate and extensible colorimetry, Thornton published a new method of extraction of color matching functions (Thornton, 1998b). Thornton's justification for the new method was his finding that color matching functions by either the Maxwell method or by the maximum saturation method lead to large errors (discrepancies) in computed chromaticities of pairs of visually matching lights. The direct extraction method utilizes 10 pairs of visually matching lights and 8 observers. Each observer matches a constant reference white in the bottom field of view by adjusting the spectral power distribution of a white light mixture in the top field of view.

Using an iterative approach, a computer program uses a set of arbitrarily chosen color matching functions to predict the tristimulus values of the reference and visual match. The total tristimulus error (TTE_o) is then calculated between the two sets, for each of the 10 visually matching pairs of lights. The color matching functions are then weighted, renormalized, and the TTE again computed. The ratio of TTE/TTE_o is then added to successive wavelengths of the weighting function. After each iteration, the correction function is modified, suggesting an appropriate change to the working color matching functions in order to reduce the total tristimulus error.

The result is a new function already accredited with almost completely eliminating the initial large tristimulus error, committed by the beginning CIE function in computing tristimulus values of highly metameric starting material. Having used the CIE 2° functions as an initial set, new functions have been extracted, with reductions of total tristimulus error to 1-2%.

1.2 Cone Spectral Sensitivities

The properties of human color matching are defined by the spectral responsivities of the three cone types. Thus, if the spectral responsivities of three cone types are known, two stimuli, denoted by their spectral power distributions $\Phi_1(\lambda)$ and $\Phi_2(\lambda)$, will match in color if the product of their spectral power distributions and each of the three cone responsivities, $L(\lambda)$, $M(\lambda)$, and $S(\lambda)$, integrated over wavelength, are equal. (Fairchild, 1998) This is shown in Equation 1.3.

$$\begin{aligned} \int_{\lambda} \Phi_1(\lambda) L(\lambda) d\lambda &= \int_{\lambda} \Phi_2(\lambda) L(\lambda) d\lambda \\ \int_{\lambda} \Phi_1(\lambda) M(\lambda) d\lambda &= \int_{\lambda} \Phi_2(\lambda) M(\lambda) d\lambda \\ \int_{\lambda} \Phi_1(\lambda) S(\lambda) d\lambda &= \int_{\lambda} \Phi_2(\lambda) S(\lambda) d\lambda \end{aligned} \qquad (1.3)$$

It follows therefore from Equations 1.2 and 1.3 that the relationship between colorimetry and cone spectral sensitivities is very close indeed. Because of the difficulty in measuring the spectral sensitivity of cones, the need for a colorimetry defined by color matching predated vision research by several decades. Thus the CIE, in establishing the 1931 standard observer previously mentioned, needed to take a less direct, psychophysical approach. Since then, scientific techniques have progressed a long way. Cone spectral sensitivities (a.k.a. cone fundamentals) have been proposed by Vos and Walraven (1971), Smith and Pokorny (1975), Stockman, MacLeod and Johnson (1993), and Stockman and Sharpe (1998). Each set of cone fundamentals are briefly discussed.

1.2.1 Vos and Walraven

Vos and Walraven set out to try and determine the spectral response curves of the foveal receptor system, their work using the underlying assumption that dichromatic vision can be conceived as a reduced form of normal trichromatic vision. This assumption was not universally agreed upon, Helmholtz hypothesized that deuteranopia is attributed to a lack of G receptors, whereas Fick believes that it is due to a fusion of the R and G systems. (Vos and Walraven, 1970)

Their foundation was that the spectral sensitivities of the foveal receptors (R, G, B) are a linear transform of the CIE system primaries (X, Y, Z), by an unknown transform, shown in Equation 1.4.

$$\begin{pmatrix} R_\lambda \\ G_\lambda \\ B_\lambda \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \cdot \begin{pmatrix} X_\lambda \\ Y_\lambda \\ Z_\lambda \end{pmatrix} \qquad (1.4)$$

Because of the well known limitations of the 1931 CIE standard observer below 460nm, Judd (1951) proposed a revision, upon which their work was based. The characteristics upon which the vision model is built, are based on a Helmholtz-type three receptor cone, converted to a Hering-type antagonistic red-green and yellow-blue color signal.

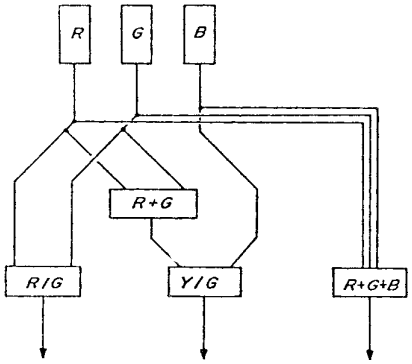


Figure 1.1 – Schematic representation of the zone model of color vision

A schematic diagram is shown in Figure 1.1. The scheme was developed to account for a wide variety of color vision phenomena such as color discrimination, the Bezold-Brucke effect and the Stiles-Crawford hue shift. (Vos and Walraven, 1970)

It was assumed that the luminance contribution Y_λ was the sum of the receptor contributions, thus can be written as in Equation 1.5.

$$R_\lambda + G_\lambda + B_\lambda = Y_\lambda \qquad (1.5)$$

$$r_\lambda + g_\lambda + b_\lambda = 1 \qquad (1.6)$$

Equation 1.6 is in terms of relative sensitivity, thus three of the 9 unknowns in Equation 1.4 are known ($A_{12,}$ $A_{22,}$ A_{32}). The other six were derived from dichromatic color confusion data.

Dichromatic color confusion is a simplified form of trichromacy, building upon the assumption that the confusion of colors in a color deficient observer occur along straight convergent lines in a chromaticity diagram. Using tritanopic, deuteranopic and protanopic color observers, the remaining unknowns in Equation 1.4 can be calculated.

The data on the tritanopic confusion center was re-evaluated by Walraven (1974), his study leading to a new center. On the basis of his work, Vos (1978) redefined the cone system fundamentals.

$$\begin{pmatrix} R_\lambda \\ G_\lambda \\ B_\lambda \end{pmatrix} = \begin{pmatrix} 0.1551646 & 0.5430763 & -0.0370161 \\ -0.1551646 & 0.4569237 & 0.0296946 \\ 0 & 0 & 0.0073215 \end{pmatrix} \cdot \begin{pmatrix} X_\lambda \\ Y_\lambda \\ Z_\lambda \end{pmatrix} \qquad (1.7)$$

The final matrix transformation proposed by Vos and Walraven (1971), updated by Walraven (1974), based on the Judd (1951), Vos (1978) modified X, Y, Z, functions is shown in Equation 1.7.

1.2.2 Smith and Pokorny

Smith and Pokorny (1971), showed that the spectral sensitivities of protanopes and deuteranopes are similar in shape on their long wavelength slope to that of isolated visual pigment absorption spectra. Their extended research analyzed the spectral sensitivity of protanopes and deuteranopes in the wavelength range 400-520nm. This specific wavelength range posing a number of problems. Firstly, the 400-520nm region is one where two photopigments are active, and secondly, the accuracy of the CIE color

matching functions within that range is questionable. To combat the problems encountered with the color matching functions, Smith and Pokorny utilized additional tools. Thirdly, Smith and Pokorny believed that when adjusting the luminosity curves and their CIE transforms for comparison with pigment absorption curves, the corrections for the inert ocular pigments become crucial. Fourth, the problems in the short wavelength region possibly arise from fluorescence within either the optic media, or the receptors themselves. Such fluorescence could give rise to heightened sensitivity curves, which would deviate from predicted absorption spectra of the visual pigments. However, they believe that the effect is confined to wavelength 400-410nm and is probably negligible. (Smith and Pokorny, 1975)

In light of the problems just mentioned, Smith and Pokorny decided to calculate a set of receptor characteristics based on a set color confusion (copunctal) points, whilst restricting the model such that the sum of the R and G functions represent the Y function. Smith and Pokorny chose the same protanopic and deuteranopic copunctal points as Vos and Walraven (1971), with a different tritanopic copunctal point.

$$\begin{pmatrix} R_\lambda \\ G_\lambda \\ B_\lambda \end{pmatrix} = \begin{pmatrix} 0.15514 & 0.54312 & -0.03286 \\ -0.15514 & 0.45684 & 0.03286 \\ 0 & 0 & 0.00801 \end{pmatrix} \cdot \begin{pmatrix} X_\lambda \\ Y_\lambda \\ Z_\lambda \end{pmatrix} \qquad (1.8)$$

The final matrix transformation proposed by Smith and Pokorny, based on the Judd (1951), Vos (1978) modified X, Y, Z, functions is shown in Equation 1.8.

Their results suggest that the dichromatic data used and their transformation of the CIE color matching functions in the wavelength range 430-700nm approximate the visual pigment absorption coefficient predicted from the iodopsin absorption spectra. Thus

concluding that the iodopsin spectra can provide a reasonable description of the visual pigment absorption spectra of the middle and long wavelength sensitive foveal cones.

Later, DeMarco, Smith and Pokorny (1992) published new cone fundamentals derived from X-chromosome linked anomalous trichromats for the wavelength range of 400-700nm.

1.2.3 Stockman, MacLeod and Johnson

Using a new technique, transient chromatic adaptation, Stockman *et al.* measured the spectral sensitivity of middle-wave sensitive (M-) cones in 11 normal and 2 protanopic observers, and the long-wave sensitive (L-) cones in 12 normal and 4 deuteranopic observers.

Smith *et al.* (1983) evaluated the differences between the CIE 2° CMF's, the Judd revision and the Stiles and Burch 2° color matching functions, and concluded that the different color matching functions are equivalent except for variations in pre-receptoral absorption's. Stockman *et al.,* believe that to the extent that Smith *et al.* analysis is true, the differences between the Stiles and Burch cone fundamentals and the CIE 2° CMF's will disappear if differences in lens and macular pigmentation are allowed for. Therefore, Stockman *et al.,* based their cone fundamentals on the revised Stiles and Burch (1955) cone fundamentals due to their consistency with deuteranopic and protanopic spectral sensitivities, with tritanopic color matching and with their own spectral sensitivity data measured in color normal observers.

The proposed Stockman *et al.,* cone fundamentals are based on the Stiles and Burch (1955) $\bar{r}$, $\bar{g}$ and $\bar{b}$ 2° color matching functions, calculated using Equation 1.9.

$$\begin{pmatrix} L_\lambda \\ M_\lambda \\ S_\lambda \end{pmatrix} = \begin{pmatrix} 0.214808 & 0.751035 & 0.045156 \\ 0.022882 & 0.940534 & 0.076827 \\ 0 & 0.016500 & 0.999989 \end{pmatrix} \cdot \begin{pmatrix} R_\lambda \\ G_\lambda \\ B_\lambda \end{pmatrix} \qquad (1.9)$$

Where R, G and B, are the Stiles and Burch 2° color matching functions, L, M and S, are the foveal receptors. The S-cone sensitivities have been further extended at wavelengths greater than 525nm using Equation 1.10.

$$\log_{10} S_\lambda = \frac{10{,}402.1}{\lambda - 21.7185} \qquad (1.10)$$

Because the Stiles and Burch (1955) cone fundamentals are not a linear transform of Judd's CIE V(λ) function, an approximation can be made using Equation 1.11.

$$V_M(\lambda) \cong 0.68273 L_\lambda + 0.35235 M_\lambda \qquad (1.11)$$

1.2.4 Stockman and Sharpe

In extended research, Stockman and Sharpe measured the spectral sensitivities of the short (S-), middle (M-) and long (L-) wavelength cones in normal trichromats, in dichromats, and monochromats of known genotype. (Stockman *et al.*, 1998)

As with Vos and Walraven (1971), Smith and Pokorny (1975), and Stockman, MacLeod and Johnson (1993), the use of monochromatic and dichromatic observers to define normal cone spectral sensitivities requires that their vision be a reduced form of normal color vision.

An important aspect of the research is that the M- and L- cone photopigment genes of the dichromat observers and those of the blue-cone monochromat observers were genetically

sequenced. The dichromat observers, with the exception of three protanopes, had only a single M- or L- cone photopigment gene. According to Stockman *et al.*, using single gene dichromats simplified the interpretation of the spectral sensitivity data. Of the nine protanopes, three had a single L1M2 gene, three had a single L2M3 gene, one had both an L1M2 and an M gene, and two had both an L2M3 and an M gene. The spectral sensitivities of protanopes with a single L1M2 and protanopes with a single L2M3 photopigment gene were practically indistinguishable.

Stockman and Sharpe compared the normal, dichromat and monochromat spectral sensitivities and tritanopic color matching data with the Smith and Pokorny (1975) fundamentals and with the Stockman *et al.* (1993) fundamentals. Their results favor the Stiles and Burch 2° color matching functions and the CIE 1964 10° color matching function based L-, M-, and S- cone fundamentals of Stockman, MacLeod and Johnson. Both sets are consistent with each other and with their dichromat data, but inconsistent with the blue cone monochromat S-cone spectral sensitivities at long wavelengths.

Using the data derived from their experiments, Stockman *et al.* (1999) proposed a new transformation from the Stiles and Burch 10° color matching functions into cone spectral sensitivities. The matrix transformation is shown in Equation 1.12.

$$\begin{pmatrix} \bar{l}_\lambda \\ \bar{m}_\lambda \\ \bar{s}_\lambda \end{pmatrix} = \begin{pmatrix} 2.802296 & 10.973728 & 1.000000 \\ 0.167289 & 8.235906 & 1.000000 \\ 0 & 0.010600 & 1.000000 \end{pmatrix} \cdot \begin{pmatrix} \bar{r}_\lambda \\ \bar{g}_\lambda \\ \bar{b}_\lambda \end{pmatrix} \qquad (1.12)$$

where $\bar{r}, \bar{g}$ and $\bar{b}$ are the Stiles and Burch (1955) 10° color matching functions, and $\bar{l}$, $\bar{m}$ and $\bar{s}$ are the relative foveal receptor sensitivities.

1.3 Current Research

The aim and objectives of this research were outlined in the introduction. The computational analysis builds upon the work of Alfvin (1995), Shaw and Montag (1998), and Shaw and Montag (1999). Alfvin's work involved a visual experiment designed to explore observer metamerism by quantifying the precision and accuracy of three sets of color matching functions, and the magnitude of observer variability found in hard-copy to CRT matches. The works by Shaw and Montag involved visual experiments in which observers created metameric matches to a reference stimuli using an additive color mixture device.

Using the three metameric data sets the accuracy of each of the sets of color matching functions / cone fundamentals can be evaluated and confidence bounds calculated. Due to the fact that Alfvin's data was taken from cross-media matches, the matches are inherently metameric. The work relies on the assumption that any color match will have equal cone responses, thus for each set of color matching functions / cone fundamentals, the optimum match would be $0\Delta L$, $0\Delta M$, $0\Delta S$ cone response.

One can also consider that when dealing with metameric matching response data, one can optimize the existing sensitivity functions to find a set that will optimally predict the metameric response, given the bounds of observer variance.

2 RESPONSIVITY FUNCTION SELECTION

To save confusion between the differences of color matching functions and cone sensitivity functions, in this work both are called responsivity functions. Each of the different sets of responsivity functions claim to have their benefits over the standard CIE set. Some are merely a linear transformation of that derived by the CIE into cone space to better predict the cone responses, whilst others include modifications derived by the authors. This chapter documents the sets of color matching functions and cone responsivities chosen for this work, and provides reasoning for the rejected sets.

2.1 Selected Sets

A total of six sets were chosen to evaluate the accuracy of the CIE 1931 color matching functions, including the CIE 1931; CIE 1964; Stiles and Burch; Demarco, Smith and Pokorny; Stockman, MacLeod and Johnson; and Vos and Walraven. A tabulated set of the chosen responsivity functions can be found in Appendix 7, the most up to date sets can be found on the internet at the Color and Vision Research Laboratory home page : http://www-cvrl.ucsd.edu/ .

2.1.1 CIE 1931 Color Matching Functions

This set has been included by default, the responsivity curves are shown in Figure 2.1.

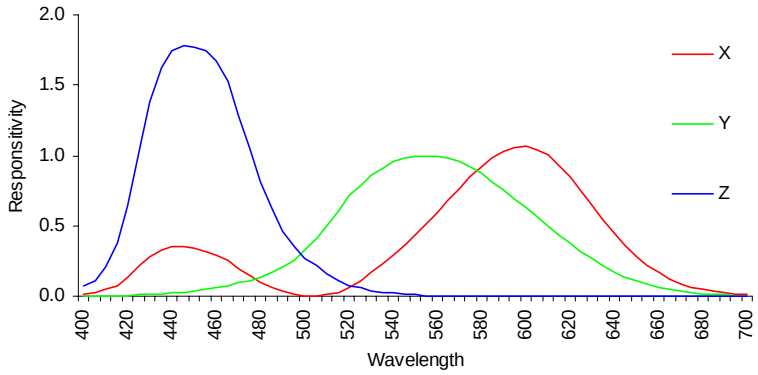


Figure 2.1 : Spectral responsivity of the CIE 1931 color matching functions

2.1.2 CIE 1964 Color Matching Functions

The CIE 1964 color matching functions have been included in order to evaluate the use of the 10° functions instead of the standard 2° functions. Technically the 1964 functions should only be used when viewing a stimuli that subtends a visual angle of greater than 4° on the retina. By including the 10° functions in the analysis, with the experimental setup being for a 2° observation, one can determine if the 1964 functions provide an overall improvement over the CIE 1931 set, although the application is not considered correct. The responsivity functions are shown in Figure 2.2.

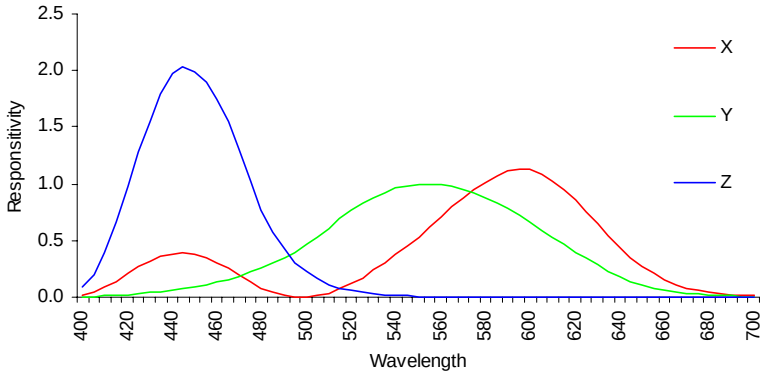


Figure 2.2 : Spectral responsivity of the CIE 1964 color matching functions

2.1.3 Stiles and Burch Color Matching Functions

Some have considered the ‘pilot data’ proposed by Stiles and Burch (1955), to be the most accurate set of color matching functions available to date. The study involved a group of 10 observers making color matches and hetero-chromatic brightness matches in 2° and 10° bipartite fields. The final decision of the CIE committee was that the differences between the pilot data and the CIE standard observer were not large enough to warrant a change. With this in mind, the inclusion of this set provides an interesting benchmark for the accuracy of other sets of responsivity function. The Stiles and Burch responsivity functions are shown in Figure 2.3.

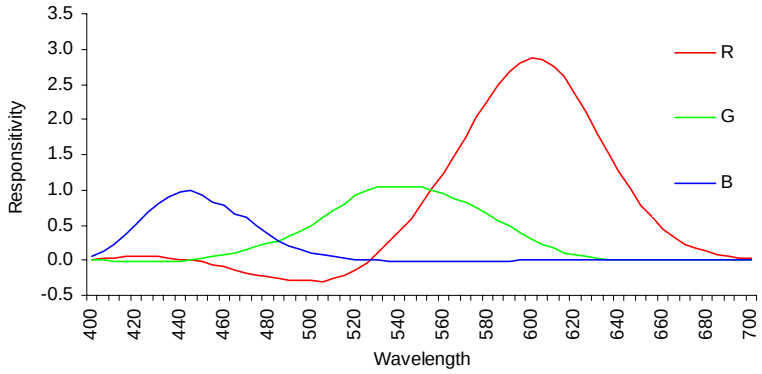


Figure 2.3 : Spectral responsivity of the Stiles and Burch color matching functions

2.1.4 Demarco, Smith and Pokorny Cone Fundamentals

The work by Demarco, Smith and Pokorny (1992), proposes a set of cone fundamentals for X-chromasome-linked anomalous trichromats. The work used the cone fundamentals proposed by Smith and Pokorny (1975) to provide an initial estimate of the normal L and M cone responsivities for an observer whose color matches are those of the Judd observer. The spectral responsivity functions are shown in figure 2.4

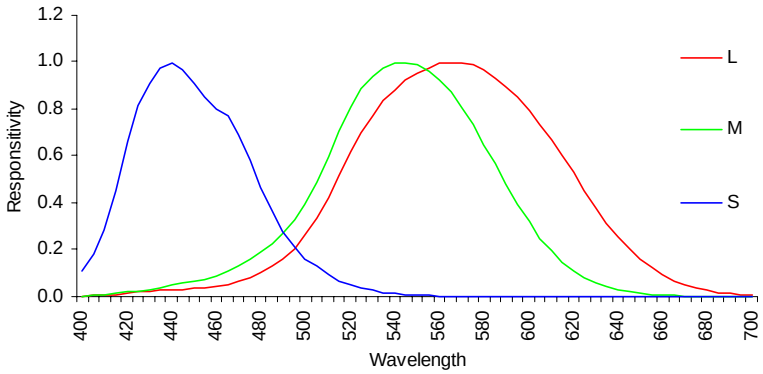


Figure 2.4 : Spectral responsivity of the Demarco, Smith and Pokorny cone fundamentals

2.1.5 Stockman, MacLeod and Johnson Cone Fundamentals

The responsivity functions proposed by Stockman, MacLeod and Johnson (1993) used a new technique to measure the spectral responsivity of middle-wave sensitive, and long-wave sensitive cones in observers. The cone absorption spectra are based on the Stiles and Burch data, but the S-cone responsivities are modified in the long wavelength region. The spectral responsivity functions are shown in Figure 2.5.

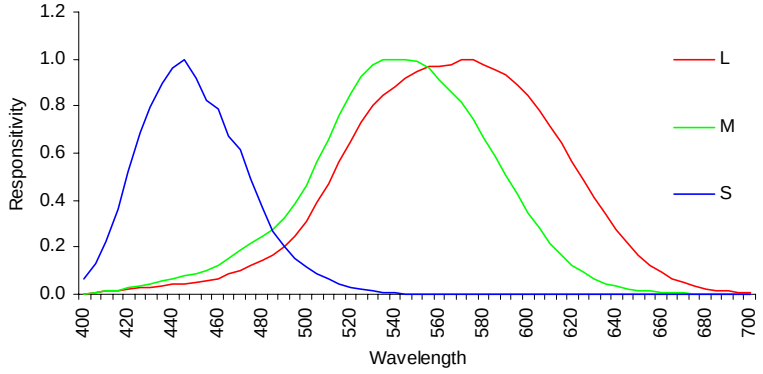


Figure 2.5 : Spectral responsivity of the Stockman, MacLeod and Johnson cone fundamentals

2.1.6 Vos and Walraven Cone Fundamentals

The work of Vos and Walraven (1970) used the assumption that dichromatic color vision can be conceived as a reduced form of trichromatic vision. Their foundation being that the spectral responsitivities of the foveal receptors are a linear transform of the CIE primaries. The original work by was re-evaluated again by Walraven (1974) and Vos (1978) based on the Judd (1951), Vos (1978) modified color matching functions. They are not therefore a linear transformation of the CIE 1931 color matching functions. The spectral responsivity functions are shown in Figure 2.6.

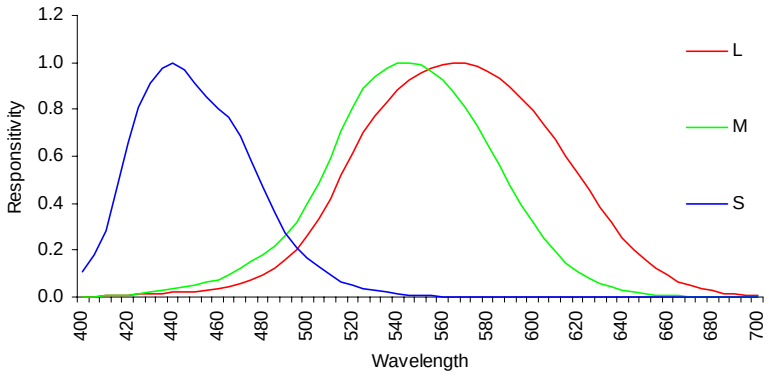


Figure 2.6 : Spectral responsivity of Vos and Walraven cone fundamentals

2.2 Rejected Responsivity Functions

Not all of the responsivity functions were used in the analysis. A selection of different sets of color matching functions were analyzed, those being a linear transformation of another set were automatically discounted.

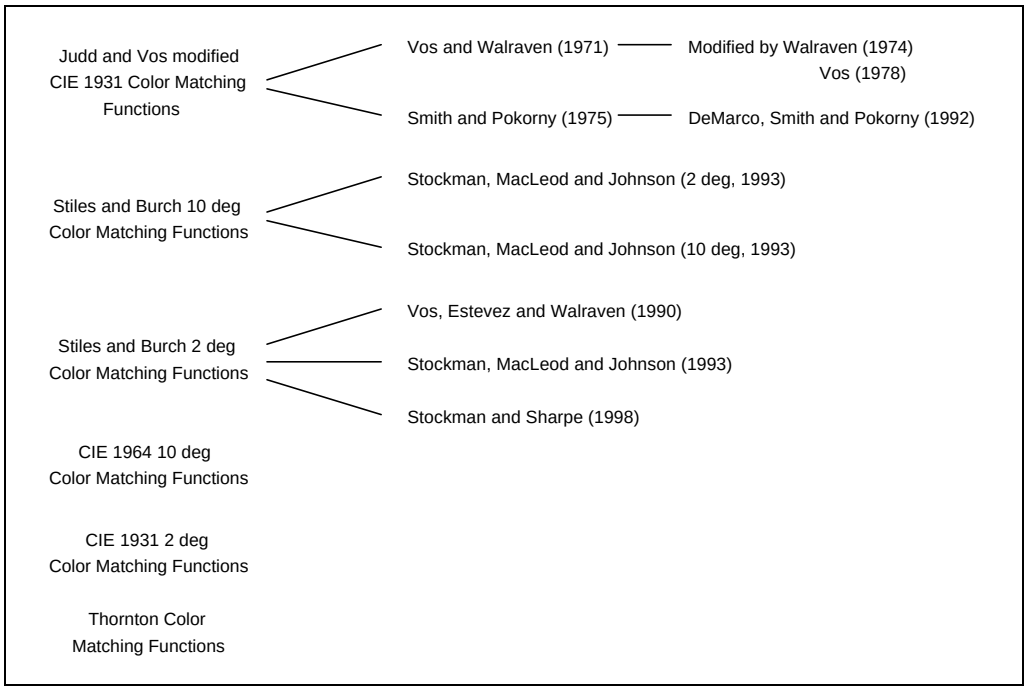


Figure 2.7 : The foundation of various sets of cone fundamentals

Figure 2.7 shows the different sets of color matching functions considered for the computational analysis. The Thornton (1998) set was not included in the analysis since it is still under revision by the author.

The Stockman and Sharpe (1998), and Vos, Estevez and Walraven (1990) sets were not considered because they are a linear transformation of the Stiles and Burch (1955) set. The Stockman, MacLeod and Johnson (1993) set was included due to the s-cone modifications made to the Stiles and Burch (1955) data at wavelengths greater than 525nm.

The Stockman, MacLeod and Johnson (1993) modifications were not included because they are a linear transformation of the CIE 1964 10° color matching functions.

3 EXPERIMENTAL DATA SOURCES

This work utilized experimental data from previous visual experiments by Alfvin (1995), Shaw and Montag (1998), and Shaw and Montag (1999). This chapter discusses the experimental setup and data collection procedures of the three experiments.

3.1 Experiment 1 – Alfvin and Fairchild (1995)

Alfvin and Fairchild (1995) designed a visual experiment to permit observers to make critical color matches between prints or transparencies and a CRT display. Seven color prints, and seven color transparencies were prepared as fixed matching stimuli. The seven colors included red, green, blue, gray, cyan, magenta, and yellow. The color print samples were designed in Adobe PhotoShop (v2.5) and imaged with a Fujix Pictrography 3000 color printer. The color transparencies were imaged using an MGI Solitaire 8_{xp} film recorder using 4x5 Ektachrome 100 Plus Professional film. The chromaticities of the fixed hard-copy samples illuminated with a fluorescent D50 simulator were designed to effectively sample the color gamut of the Sony Trinitron CRT display, model PVM-1942Q used to generate the soft-copy color matches (Alfvin, 1995).

With the use of a simple optical setup, consisting of an equilateral glass prism mounted on an optical bench, the observers were able to simultaneously view both the soft and hard-copy matching stimuli. The experimental setup is shown in Figure 3.1.

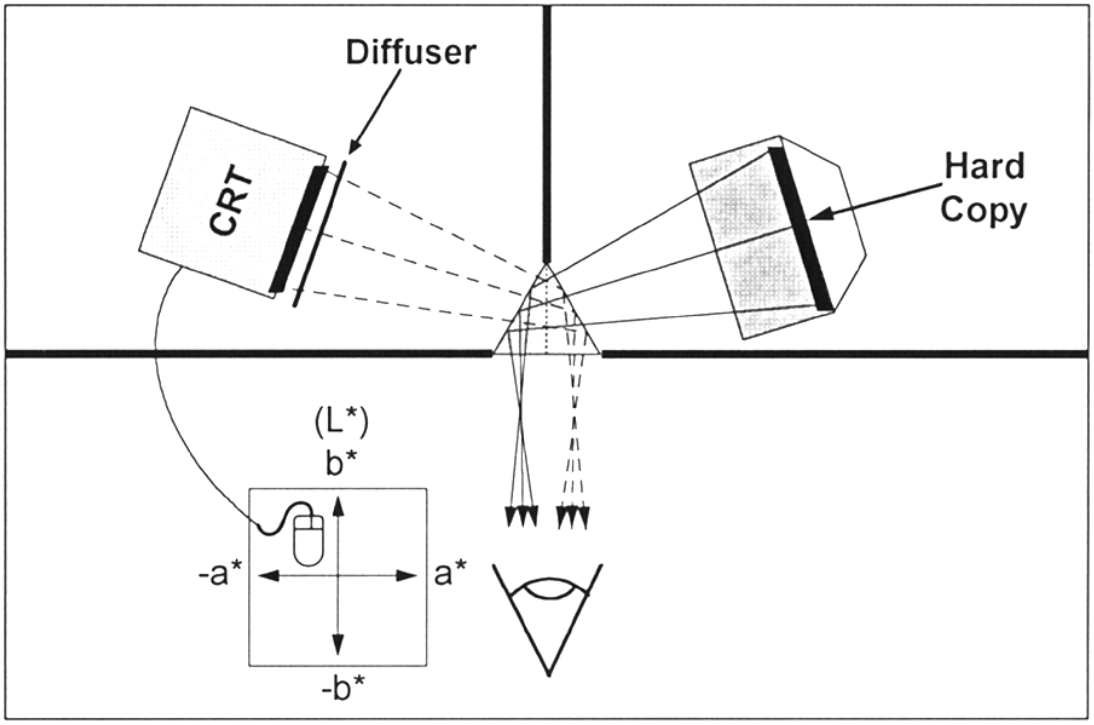


Figure 3.1 : Schematic overhead view of Alfvin experimental setup used for color matching

Both the CRT Display and combination light-booth / light box were aligned with the optical prism and shielded from the observer. The fixed hard copy stimulus and the adjustable soft-copy stimulus were presented in a vertical symmetric bipartite field. The color matching stimuli were presented as solid colors appearing self-luminous in a darkened room. A neutral translucent diffusion material placed in front of the CRT display eliminated the appearance of any visual texture in the soft copy stimulus, rendering the soft-copy stimulus identical to the hard-copy stimuli in terms of spatial characteristics (Alfvin, 1995).

Observers were seated at a distance of approximately 1 meter from the matching stimuli. The 5x5 cm matching field subtending a visual angle of 2.9°. The observers were asked to adjust the color appearance of the soft-copy stimulus to create an exact match for each of the fourteen different hardcopy stimuli (Alfvin, 1995).

In order to minimize any potential effects associated with observer fatigue resulting from extended duration of the experimental tasks, the experiment was divided into two separate sessions. Each session lasting approximately 45-60 minutes (Alfvin, 1995).

After a match was attained, a PhotoResearch PR-650 telespectroradiometer incorporating a half-height triangular bandpass of 4nm, recorded the spectral radiance at 4nm wavelength intervals across the visible spectrum between 380 and 780 nm.

Figure 3.2 shows an example of the different spectral power distributions required to produce a metameric match between the transparency sample and the CRT display.

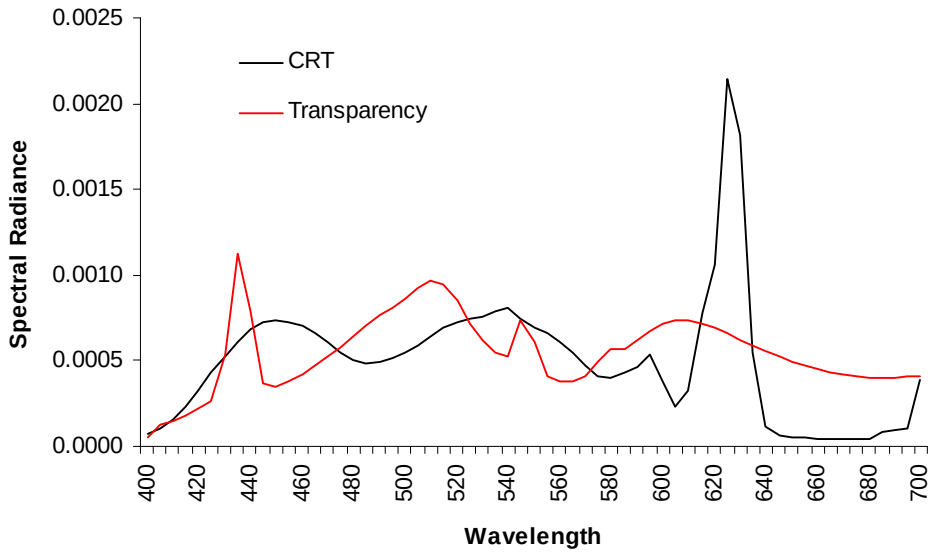


Figure 3.2 : The spectral power distributions of a gray transparency and a CRT display adjusted to yield a metameric match for an observer.

3.2 Experiments 2 – Shaw and Montag (1998)

Shaw and Montag (1998) designed a visual experiment to allow observers to perform color match between a gray card of Munsell N5 and an ACS VCS 10 additive mixing device.

The ACS VCS 10 consisted of 7 colored discs, all rotating at high speed to simulate an additive integral visual response. The proportions of each colored disc could then be adjusted by an observer using the controls to simulate a color. The viewing booth had fluorescent daylight and incandescent illumination to view the colors. Figure 3.3 shows the ACS VCS 10 with the front fascia panel removed displaying all 7 colored discs.

Figure 3.3 : ACS VCS 10 additive mixture device

The seven discs in the ACS VCS 10 were White, Red, Green, Blue, Yellow, purple and Black. Independent control was allowed of any 3 primaries at any one time by the user control panel. The goal being to generate a metameric match between the Munsell N5 paper and the three primaries.

Of the 5 colored primary discs, two sets of 3 Primaries were chosen – Red, Green, Blue (RGB) and Blue, Yellow, Purple (BYP). Figure 3.4 shows the ACS VCS 10 in use.

Observers participated in the color matching experiment to assess inter- and intra-observer variability. All observers claimed to be color normal, but this was not tested. Each observer performed the color matching experiment 10 times in succession for each primary triplet. The matching field was 8cm by 9cm, subtending a visual angle of 7°.

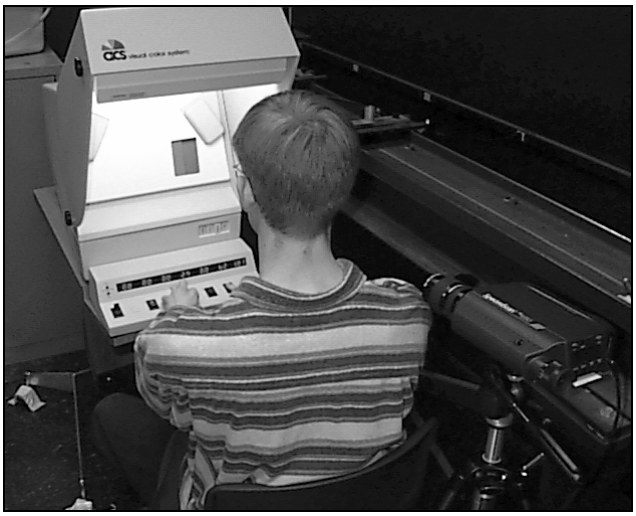


Figure 3.4 : ACS VCS 10 in use. Colored primary disc contributions were controlled by the user control panel. Metameric match spectra are taken using PhotoResearch PR650 mounted by observers right shoulder.

Observers were seated 30 inches from the stimuli and asked to make an exact color match to the Munsell N5 gray card using only the three primaries specified. When a color match had been achieved, the PhotoResearch PR650 was used to measure the spectral radiance of the metamer from the observers angle of view. This was considered very important due to the angular properties of the colored discs, whereby a color match was perceived a different color when viewed from a different viewing angle. Each observer was asked to repeat the experiment 10 times for each primary set.

The precision and accuracy of the PhotoResearch was not tested. A previous evaluation performed by Alfvin (1995) indicated that the systematic and random error associated with the instrument were minimal and acceptable for the purpose of their research.

Figure 3.5 shows an example of the different spectral power distributions required to produce a metameric match between the Munsell N5 sample and the ACS VCS 10 additive color mixture device.

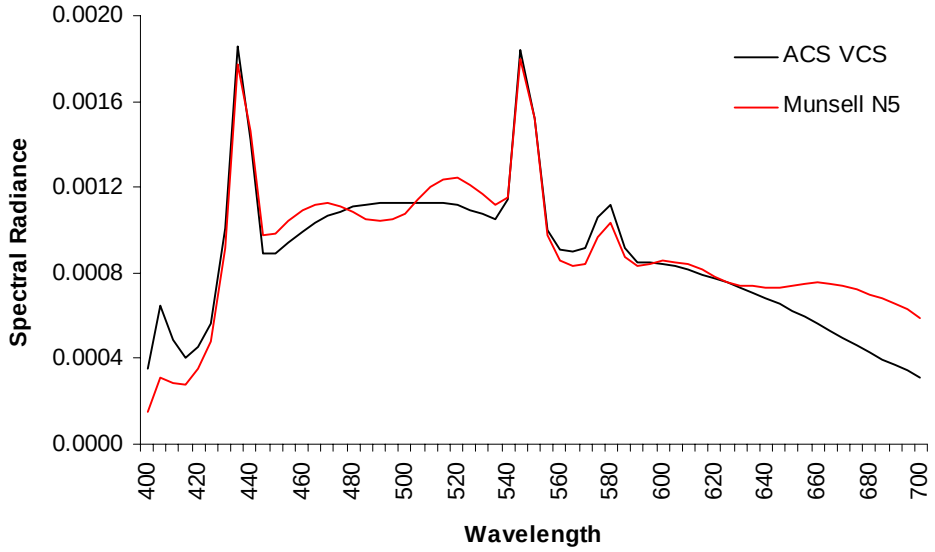


Figure 3.5 : The spectral power distribution of the Munsell N5 patch and ACS VCS adjusted to yield a metameric match under daylight fluorescent illumination.

3.3 Experiment 3 – Shaw and Montag (1999)

Using an identical experimental setup to that of Shaw *et al.* (1998), Shaw *et al.* (1999) repeated the visual experiment using a different stimuli, and different primary sets.

The stimuli was a neutral gray of $L^*=50$, created with a Fujix Pictrography 3000 color printer. The Fujix printer is a hybrid photographic / thermal-transfer continuous-tone digital printer.

The primary sets chosen were Red, Green, Blue (RGB) and Green, Yellow, Purple (GYP).

Figure 3.6 shows an example of the different spectral power distributions required to produce a metameric match between the Fujix sample and the ACS VCS 10 additive color mixture device.

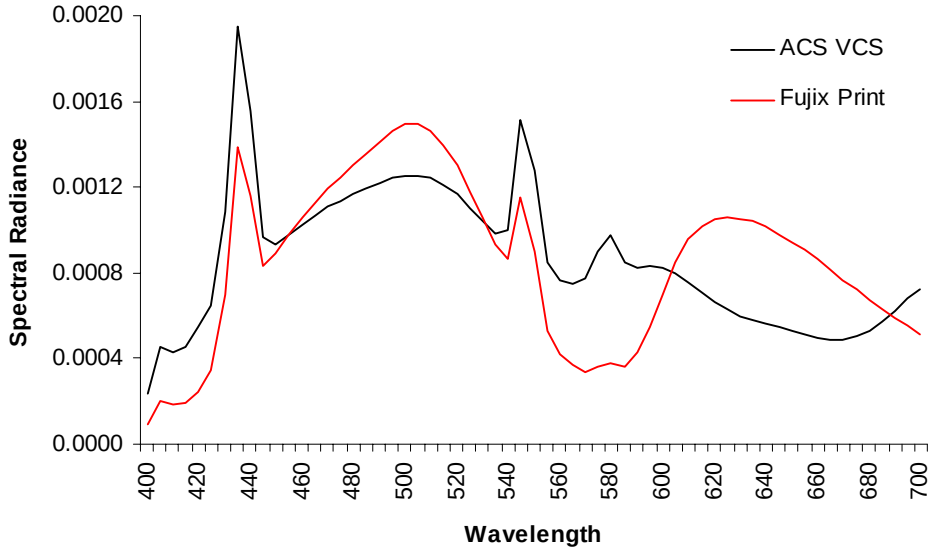


Figure 3.6 : The spectral power distribution of the Fujix patch and ACS VCS adjusted to yield a metameric match under daylight fluorescent illumination.

3.4 Experimental Discussion

The three experimental data sets combined comprise of a total of 468 metameric pairs, matched by observers. All the data was collected at 4 nm wavelength intervals using a PhotoResearch PR-650 telespectroradiometer.

Observers in the Alfvin experiment made a match with a 2.9° visual field, whereas the observers for Shaw and Montag (1998, 1999) made a match with a 7° visual field.

4 COMPARING THE ACCURACY OF VARIOUS DATA SETS

One of the key underlying assumptions of this work is that an optimum set of color matching functions / cone fundamentals, will predict that the integrated cone responses of a metameric pair are equal. This is illustrated in Figure 4.1, thus yielding a minimum color difference in a color space like CIELAB.

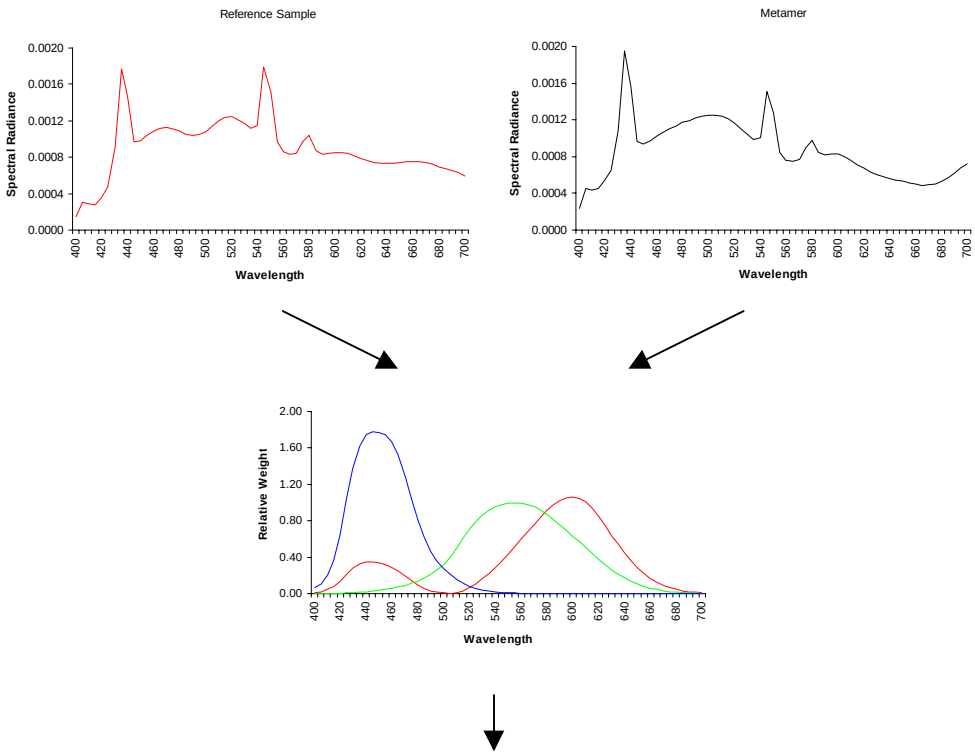


Figure 4.1 : Conceptually ideal set of color matching functions resulting in equal integral response of the cones

One can therefore evaluate the performance of a set of color matching functions using metameric data, by calculating the color difference over all the metameric pairs in that color space.

A problem arises when one wishes to compare the performance of different sets of color matching functions / cone fundamentals. It is not sufficient to assume that one can use the CIELAB color space as the comparative space by just transforming the tristimulus values of each set of color matching functions into CIELAB co-ordinates using the standard equations. One must find a common color space in which each set can be compared.

4.1 Rotation to near CIE space

The problem of disparate cone functions was overcome by assuming the common color space to be CIELAB, using the CIE 1931 standard observer. In order to evaluate the other sets of functions accurately, a linear transformation was made to rotate each set into an approximate CIE representation, shown in Equation 4.1.

$$\begin{bmatrix} x_{400} & y_{400} & z_{400} \\ \vdots & & \vdots \\ x_{700} & y_{700} & z_{700} \end{bmatrix}_{pseudo\ CIE} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}_{\substack{rotation \\ transform}} \cdot \begin{bmatrix} r_{400} & g_{400} & b_{400} \\ \vdots & & \vdots \\ r_{700} & g_{700} & b_{700} \end{bmatrix}_{Stiles} \qquad (4.1)$$

A 3x3 rotation matrix was calculated for each set of color matching functions using least squares. The rotated color matching functions / cone fundamentals, were then used to calculate the ‘pseudo CIE’ tristimulus values for each set of functions.

By rotating each set into its CIE representation, it enables one to use the standard quality metrics, including CIE ΔE_{ab}, and CIE ΔE_{94} in a color space that is commonly used for performance evaluation.

It is important to note that the transformation is a linear rotation of the cone fundamentals. Differences will still exist between the various sets of functions and the CIE 1931 set after rotation, it must be understood that a direct linear transformation between any of the sets and the CIE 1931 fundamentals is not wanted. Linear transforms of the CIE set were discarded in chapter 2.

The rotation of the Stiles set of color matching functions is demonstrated in Figure 4.2. The upper diagrams showing the original, un-rotated functions, the 3x3 rotation matrix, and the rotated Stiles and Burch functions overlaid on to of the CIE 1931 color matching functions.

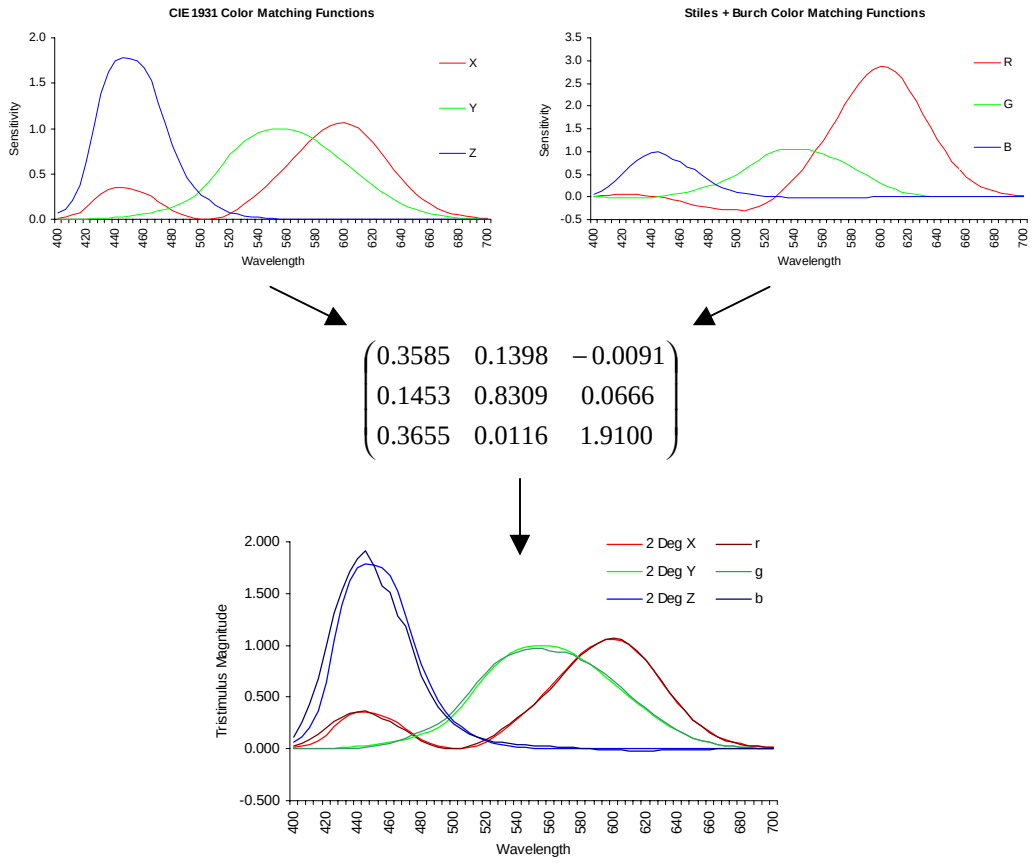


Figure 4.2 : Rotation of the Stiles and Burch Color Matching Functions into a near CIE representation

$$\begin{pmatrix} 0.9968 & 0.0282 & -0.0399 \\ -0.0863 & 0.9476 & 0.0495 \\ -0.0061 & -0.0428 & 0.9154 \end{pmatrix}$$

(4.2): CIE 1964 Transform

$$\begin{pmatrix} 0.3585 & 0.1398 & -0.0091 \\ 0.1453 & 0.8309 & 0.0666 \\ 0.3655 & 0.0116 & 1.9100 \end{pmatrix}$$

(4.3): Stiles and Burch Transform

$$\begin{pmatrix} 1.8910 & 0.6356 & -0.0092 \\ -1.3893 & 0.3949 & 0.0099 \\ 0.3589 & -0.0096 & 1.7697 \end{pmatrix}$$

(4.4): Demarco, Smith and Pokorny Transform

$$\begin{pmatrix} 1.8067 & 0.6081 & -0.0496 \\ -1.2948 & 0.3987 & 0.0774 \\ 0.3834 & -0.0464 & 1.9063 \end{pmatrix}$$

(4.5): Stockman, MacLeod and Johnson Transform

$$\begin{pmatrix} 1.8907 & 0.6355 & -0.0095 \\ -1.3876 & 0.3943 & 0.0099 \\ 0.3609 & 0.0023 & 1.7713 \end{pmatrix}$$

(4.6): Vos and Walraven Transform

The rotation matrices for the five sets of weighting functions are shown in Equations 4.2-4.6. These rotations are then applied to the weighting functions and then plotted against the CIE 1931 standard observer, shown in figures 4.3-4.6.

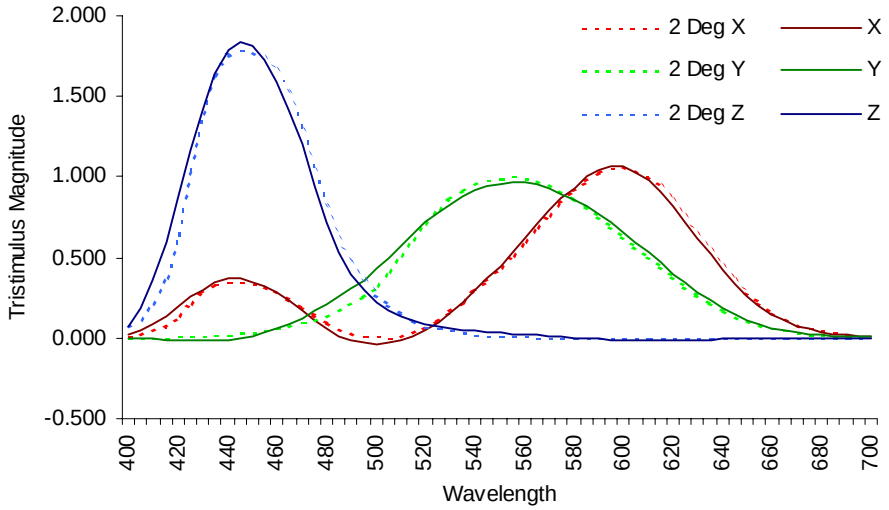


Figure 4.3 : Rotated CIE 1964 Color Matching Functions (solid line) plotted over CIE 1931 Color Matching Functions (marked line)

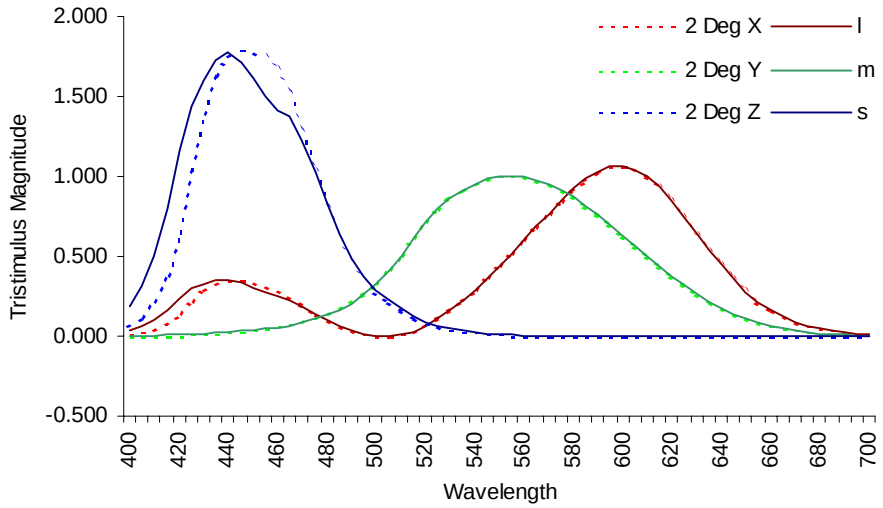


Figure 4.4 : Rotated Demarco, Smith, and Pokorny cone Fundamentals (solid line) plotted over CIE Color Matching Functions (marked line)

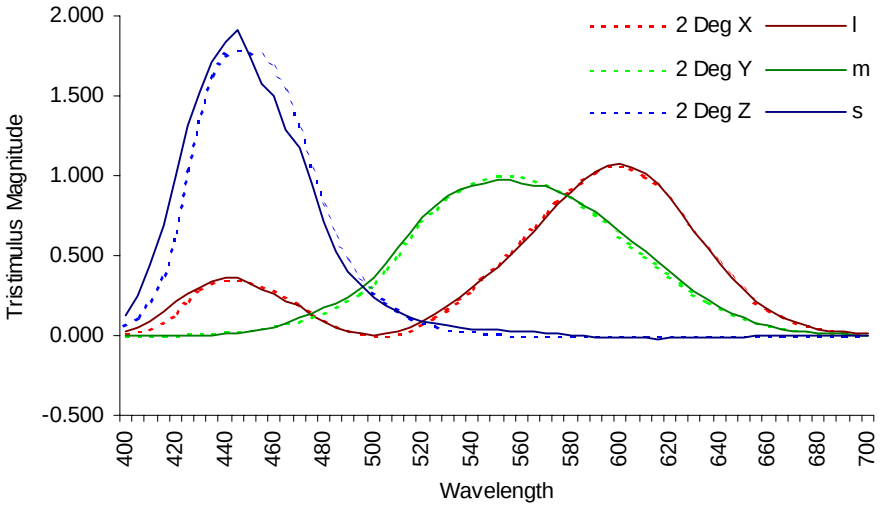


Figure 4.5 : Rotated Stockman, MacLeod, and Johnson cone Fundamentals (solid line) plotted over CIE Color Matching Functions (marked line)

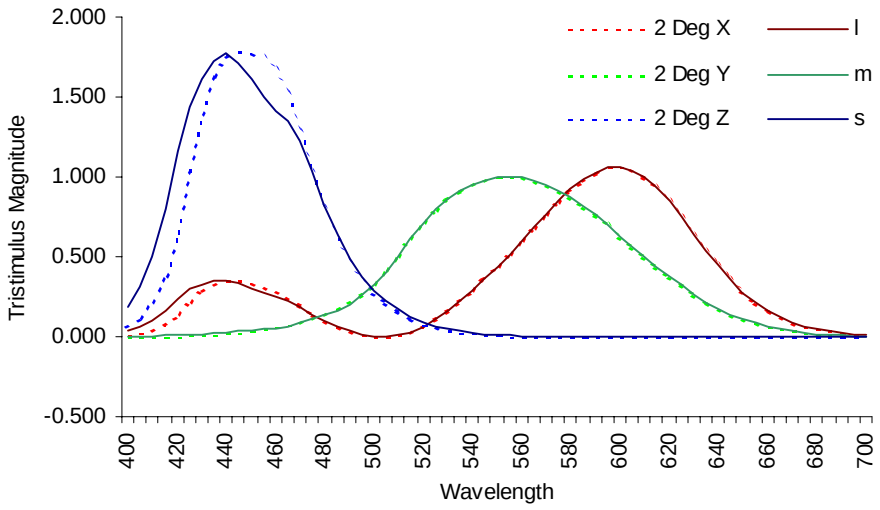


Figure 4.6 : Rotated Vos and Walraven cone Fundamentals (solid line) plotted over CIE Color Matching Functions (marked line)

It can be seen from Figures 4.2-4.6 that the different cone fundamentals differ mainly in shape below 500 nm, the main affect applying to the $\bar{z}'$ function. The rotated functions nearly all show a shift in peak sensitivity in the blue region, and also a shift in the low wavelength region of the $\bar{x}'$ function. In most cases the $\bar{y}'$ function has remained unchanged.

This indication is in line with the thoughts of the vision community, that the CIE $\bar{z}$ function inadequately describes the visual systems sensitivity to the blue end of the spectrum.

4.2 Tristimulus Calculation

Absolute tristimulus integration was used to calculate the tristimulus values for each of the metamer pairs from the spectral data. The 4 nm wavelength increment spectral data was linearly interpolated up to a 5 nm wavelength increment, and the wavelength range

cropped to 400-700 nm. The wavelength increment was chosen to correspond with that of the various sets of color matching functions / cone fundamentals, the wavelength range determined by the source of cone function data.

4.3 CIELAB Representation

Reference white tristimulus values for the measured radiance data were unknown. Therefore, since the gray patch in each data set had an L^* of approximately 50, the reference white can be approximated using the tristimulus values of the reflectance spectra scaled by 5. CIELAB coordinates were calculated according to standard CIE methods (1986), using each of the rotated sets of color matching functions. CIE ΔL^*, Δa^*, Δb^*, ΔC^*, ΔH^* values were calculated for each metameric pair. Color difference values were calculated using both CIE ΔE_{ab} and CIE ΔE_{94}.

Figures 4.7-4.9 show the positions of the metameric pairs in CIELAB space.

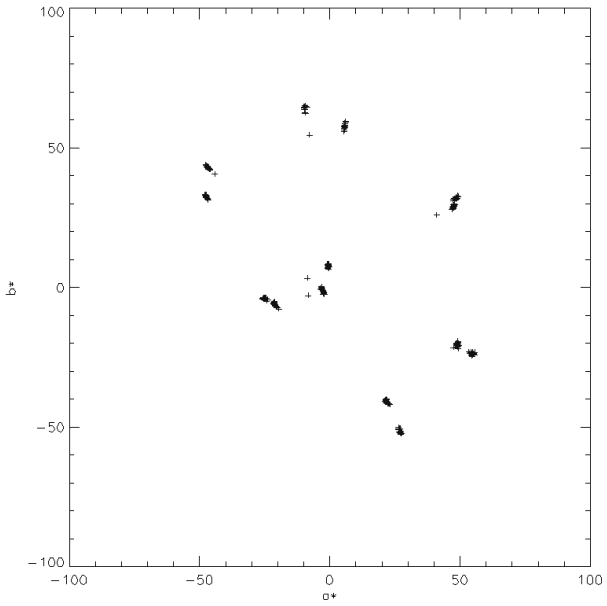


Figure 4.7 : CIELAB a^* b^* plot of all three combined data sets

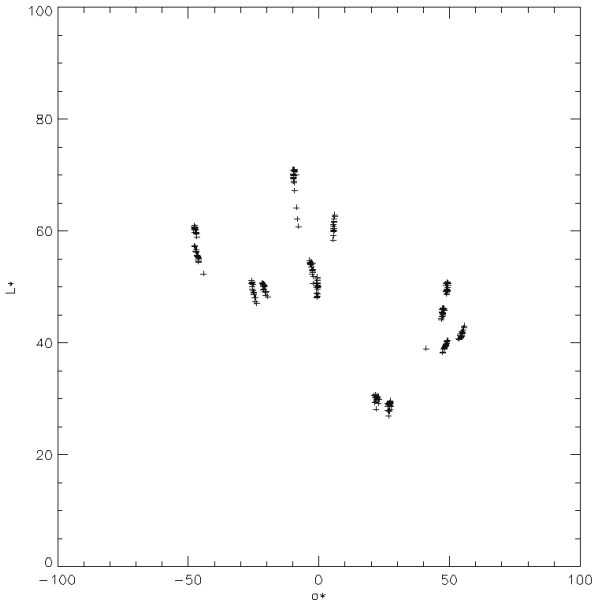


Figure 4.8 : CIELAB L^* a^* plot of all three combined data sets

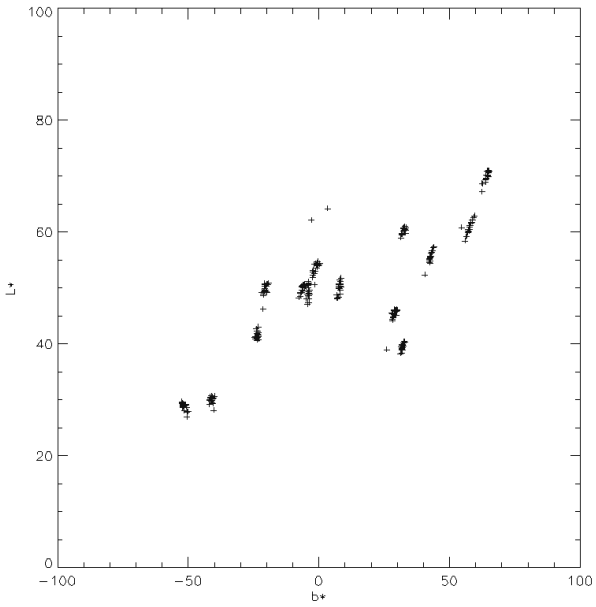


Figure 4.9 : CIELAB L^* b^* plot of all three combined data sets

4.4 Observer Variability

When considering metameric data from multiple observers, one must also consider the implications of observer variability. Inter observer variations occur through differences in observer macular pigment, lens absorptions, and other pre-receptoral features. Readers interested in understanding further the effects of inter- and intra-observer variability are advised to look to the work of North (1991), and Alfvin (1995).

For the purposes of this work it is assumed that inter- observer variability is inherent within the data. The variance can be considered a gain parameter that is present within the color difference results, the best set of weighting functions minimizing the color difference as possible. It is therefore important that one understands that the expectation of this work is *not* that an optimal set of color matching functions will yield $0\Delta E_{ab}$ over all samples, but that it will yield the lowest mean color difference.

4.5 Statistical Analysis

4.5.1 Students t-test

In order to performance each set of weighting functions, two tailed t-tests were used. The t-test was used to compare the mean color difference vector (ΔL^*, Δa^*, or Δb^*) against a mean of zero. An ideal set of color matching functions would yield a mean ΔL^*, Δa^*, Δb^* of zero, plus an offset for observer variance. The tests were run given the following hypothesis;

$$H_0 : \mu = 0.0 \qquad \alpha = 0.05$$

$$H_1 : \mu \neq 0.0$$

First, the color difference between each metamer pair, ΔL^*, Δa^*, Δb^* were calculated for each set of weighting functions. The test statistic was then calculated for each difference vector, using Eqn 4.7.

$$T_2 = \frac{\overline{X} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} \qquad (4.7)$$

where T_2 is the test statistic, $\overline{X}$ is the mean vector difference, μ_0 is equal to zero (test to see if mean is significantly different from zero), s is the sample standard deviation, and n is the number of samples.

If a set of weighting functions introduce systematic error to the CIELAB values, then the results of the t-test should indicate that fact. A random distribution of error is desirable.

4.5.2 Multivariate 95% Confidence Ellipse

A sample covariance matrix, S_{Lab}, defined by the sample covariances and variances of ΔL^*, Δa^*, and Δb^* values, as shown in equation 4.8, was calculated for each set of weighting functions.

$$S_{Lab} = \begin{bmatrix} S_{\Delta L}^2 & S_{\Delta L} S_{\Delta a} & S_{\Delta L} S_{\Delta b} \\ S_{\Delta a} S_{\Delta L} & S_{\Delta a}^2 & S_{\Delta a} S_{\Delta b} \\ S_{\Delta b} S_{\Delta L} & S_{\Delta b} S_{\Delta a} & S_{\Delta b}^2 \end{bmatrix} \qquad (4.8)$$

Assuming a multivariate normal distribution, the inverse of the sample covariance matrix S_{Lab} can be used to construct a 95% confidence region for the sample distribution of the CIE ΔL^*, Δa^*, and Δb^* multivariate data set. (Alfvin, 1995)

The *p*-dimensional confidence region for a sample distribution with probability (1-α) is represented by the quadratic distance function, shown in matrix form in Equations 4.9-4.11.

$$DS^{-1}D' = Q \tag{4.9}$$

$$D_{Lab} = \left[(\Delta L^*)\ \ (\Delta a^*)\ \ (\Delta b^*)\right] \tag{4.10}$$

$$Q_{\text{Sample Distribution}} = \frac{p(n-1)}{(n-p)} F_{(p,n-p,\alpha)} \tag{4.11}$$

where *p* is the number of variables, *n* is the number of observations, α is the Type I alpha risk associated with a 100(1-α)% confidence region, *F* is the value of the F distribution for *p*, and *n-p* degrees of freedom with an alpha risk of α. An example of a $\Delta a^* \Delta b^*$ bivariate ellipse bound by a 95% confidence region for the sample distribution defined by Equation 4.9 is shown in Figure 4.10, relative to the CIE 2° Standard Observer.

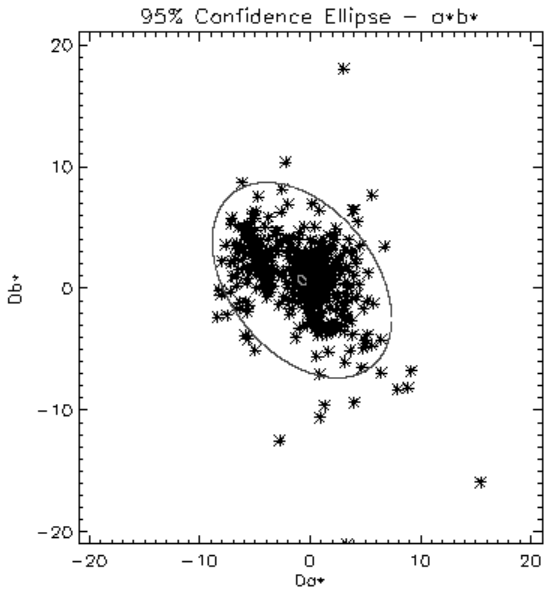


Figure 4.10 : 95% and mean confidence ellipses for combined data set

A *p*-dimensional confidence region for the mean of a sample distribution with probability (1-α) can also be described with Equation 4.9, by modifying the Q term :

$$Q_{\text{Sample Mean}} = \left(\frac{1}{n}\right)\frac{p(n-1)}{(n-p)}F_{(p,n-p,\alpha)} \qquad (4.12)$$

As shown in Equation 4.12, the Q term for the sample mean is a function of the reciprocal of sample size; as *n* increases, the magnitude of the Q term decreases. Thus, the uncertainty with which the mean is known decreases with increasing numbers of observations. (Alfvin, 1995) The mean ellipse can also be seen in Figure 4.10.

In addition to the Δa^*-Δb^* relationship shown in figure 4.10, one must also consider the ΔL^*-Δa^*, and ΔL^*-Δb^* planes because the experimental color matches involved adjustments to each of the three independent variables defining the CIELAB color space. If any one of the three confidence regions defining the sample means of the Δa^*-Δb^*, ΔL^*-Δa^*, or ΔL^*-Δb planes do not contain the theoretical mean match for a given standard observer, the mean color matches are considered to be statistically significantly different. (Alfvin, 1995)

The results of the statistical tests documented in this chapter are discussed in Chapter 5, ‘Computational Analysis Results and Discussion’.

5 COMPUTATIONAL ANALYSIS RESULTS AND DISCUSSION

The computations discussed in Chapter 4 were performed on each of the data sets individually, as well as all three data sets combined. The results of each data set are discussed in this chapter. As well as the final six sets of responsivity functions chosen, discussed in Chapter 2, the Shaw and Fairchild set derived in Chapter 6 have been included in the results for ease of comparison. The numerical results and confidence ellipse plots for each data set can be found in Appendices 1-5.

5.1 Alfvin Data Set

As discussed in Chapter 3, the Alfvin experiment was designed to permit observers to make critical color matches between color prints or transparencies and a CRT display. Seven color prints and seven color transparencies were chosen by Alfvin as fixed matching stimuli, the seven colors, red, green, blue, gray, cyan, magenta, and yellow were chosen. In order to evaluate how the region of color space also varies with the different set of responsivity functions, the computation was performed on a color center basis also.

5.1.1 Combined Alfvin Data Set

The total Alfvin data set, containing all 7 color centers is comprised of 268 metamer pairs.

The combined evaluation utilizes all metamer pairs to determine the performance of each set of responsivity functions. The complete tables of results can be found in Appendix 2, *'Prediction of Metamer Pair – Alfvin Data'*.

The results show that little difference can be found between the different sets of responsivity functions for the Alfvin Data. The CIE 1931 functions yield an average color difference of 4.39 ΔE^*_{ab}, 2.72 ΔE_{94}, with the best set of responsivity functions being that of Demarco, Smith and Pokorny, yielding an average color difference of 4.33 ΔE^*_{ab}, 2.67 ΔE_{94}. A summary of the results in Appendix 2 can be found in table 5.1.

Color Difference	CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
ΔE^*_{ab}	4.39	4.88	4.37	4.33	4.38	4.33	4.56
ΔE^*_{94}	2.72	3.01	2.72	2.67	2.70	2.67	2.81

Table 5.1 : Average results of computational analysis for all color centers and observers in the Alfvin data set.

The Shaw and Fairchild responsivity functions derived by the optimization in Chapter 6 yield an interesting result, the average color difference is the second highest. This indicates that the contribution of the Alfvin data set to the optimization may have been less significant than the other data sets, with the optimization minimum being found even though an increase in the Alfvin data set occurred.

The statistical tests performed on the data show some interesting features and can be found in Appendix 2. A two tailed t-test for means was applied on each of the ΔL^*, Δa^*, Δb^* values. The null hypothesis being that the means are equal to zero. One can see that the CIE 1931 functions fail the test in both the Δa^* and Δb^* planes, indicating that there are systematic deviations from a mean of zero. This implies that the functions are not optimum. If one then looks at the Shaw and Fairchild results, it is found that the null

hypothesis is *'failed to be rejected'* in all three dimensions, indicating that the mean cannot be shown to deviate from zero. The Shaw and Fairchild confidence ellipse plots in Figure 5.1, and the whole set of plots in Appendix 2, show the 95% confidence ellipse and mean confidence ellipse for each of the responsivity functions. It is clearly evident that the mean ellipse is far smaller than the variation of the data, in all cases.

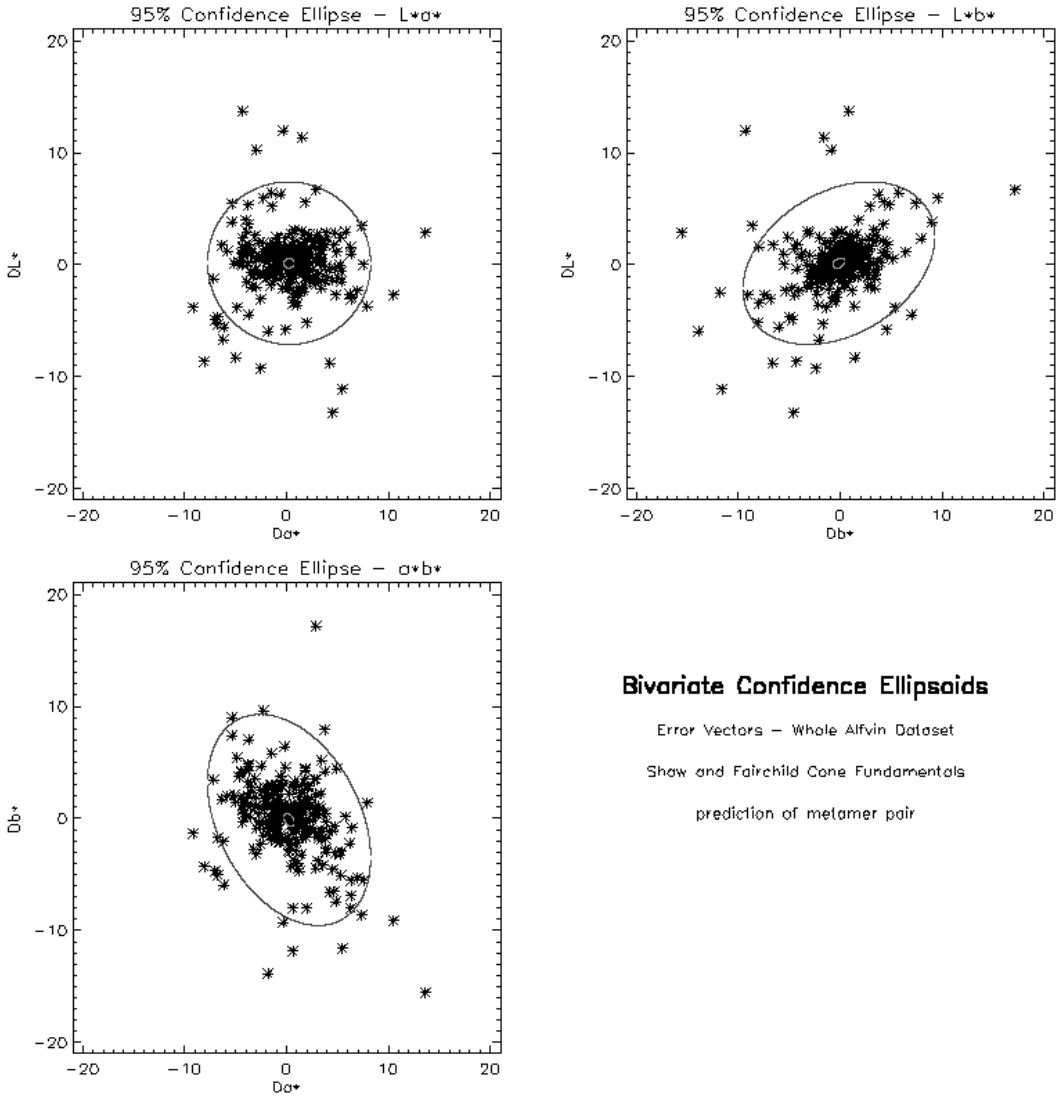


Figure 5.1 : Confidence ellipse plots of the Shaw and Fairchild responsivity functions, using the whole alfvin data set.
95% Confidence region - Outer ellipse, Mean – Inner ellipse

The shapes of the ellipses can be seen to vary from plot to plot, but the change is usually only slight and of little significance.

5.1.2 Segmented Alfvin Data Set

The segmented analysis was done both by color center and by media type, yielding 14 separate analyses. Table 5.2 summarizes the average color difference results (ΔE^*_{ab}) for each separate analysis.

Color Center		CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
Gray	Print	3.69	3.65	3.28	3.38	3.28	3.38	3.31
	Trans.	3.08	4.20	3.24	3.03	3.25	3.03	3.49
Cyan	Print	3.62	4.13	3.54	3.34	3.54	3.34	3.44
	Trans.	3.32	5.95	3.88	3.33	3.88	3.32	4.65
Magenta	Print	4.06	4.43	3.92	4.04	3.91	4.05	4.54
	Trans.	4.42	4.42	4.32	4.43	4.32	4.43	4.31
Yellow	Print	5.78	7.18	5.69	5.75	5.69	5.76	6.34
	Trans.	4.61	4.77	4.44	4.61	4.44	4.61	4.41
Red	Print	3.63	3.54	3.52	3.57	3.52	3.57	3.51
	Trans.	4.61	4.87	4.75	4.60	4.76	4.60	4.78
Green	Print	4.43	4.39	4.36	4.34	4.36	4.34	4.38
	Trans.	4.43	5.13	4.60	4.20	4.60	4.20	4.62
Blue	Print	4.92	4.76	4.63	4.65	4.63	4.65	4.89
	Trans.	6.14	6.08	6.09	6.37	6.09	6.38	6.17
Mean of Means		4.34	4.82	4.30	4.26	4.31	4.26	4.49

Table 5.2 : Mean color difference (ΔE^*_{ab}) results of computational analysis using the segmented Alfvin data set. Mean of means is the mean color difference of the averaged results for each color center, giving an indication to the overall performance of each set of responsivity functions.

It can be seen that some color centers yielded higher average color difference results than others, the blue transparency and yellow print being good examples.

That said, the majority of mean color differences tended to vary around 3-4 ΔE^*_{ab}. This indicates that observer variance exists, irrespective of color center. The overall performance (Mean of means) of each set of responsivity functions are shown in Table

5.2, it is evident that the CIE 1931 color difference is lower than the Shaw and Fairchild color difference for this data set. This again confirms the suspicion that the optimization minimization was not necessarily the Alfvin minimum.

5.2 Lab ’98 and ’99 Data Sets

The results and confidence ellipse plots of the Lab ’98 data can be found in Appendix 3, that of the Lab ’99 data in Appendix 4. Both sets of results exhibit interesting trends that can be seen in Tables 5.3 and 5.4.

Color Difference	CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
ΔE^*_{ab}	3.80	2.03	2.90	3.39	2.90	3.38	2.08
ΔE^*_{94}	3.49	1.85	2.57	3.11	2.57	3.11	1.86

Table 5.3 : Average results of computational analysis for the Lab ’98 data set.

Color Difference	CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
ΔE^*_{ab}	4.53	3.07	3.86	4.36	3.86	4.36	3.42
ΔE^*_{94}	4.16	2.79	3.50	4.01	3.50	4.01	3.12

Table 5.4 : Average results of computational analysis for the Lab ’99 data set.

The average color difference for both data sets is lower when calculated using the CIE 1964 10° functions, and the Shaw and Fairchild functions. This is interesting to note, since the general trend in the Alfvin data set was an increased color difference for both the CIE 1964 and Shaw functions. Also, the performance improvement is quite substantial, in both cases.

Although the CIE 1964 functions provide the lowest average color difference, the tabulated results in Appendices 3 and 4 show systematic deviations from a mean of zero in at least two of the three dimensions. This indicates that although the average color

difference is low, it is offset from a mean of zero. This is not surprising when one looks at the confidence ellipse plots of some of the responsivity functions, shown in Figure 5.2, it is clear that two distinct clouds exist. One possible reason can be linked to the two primary sets used to match the reference color. The Shaw and Fairchild set, exhibit systematic deviations in two of the three dimensions. The same possible explanation exists.

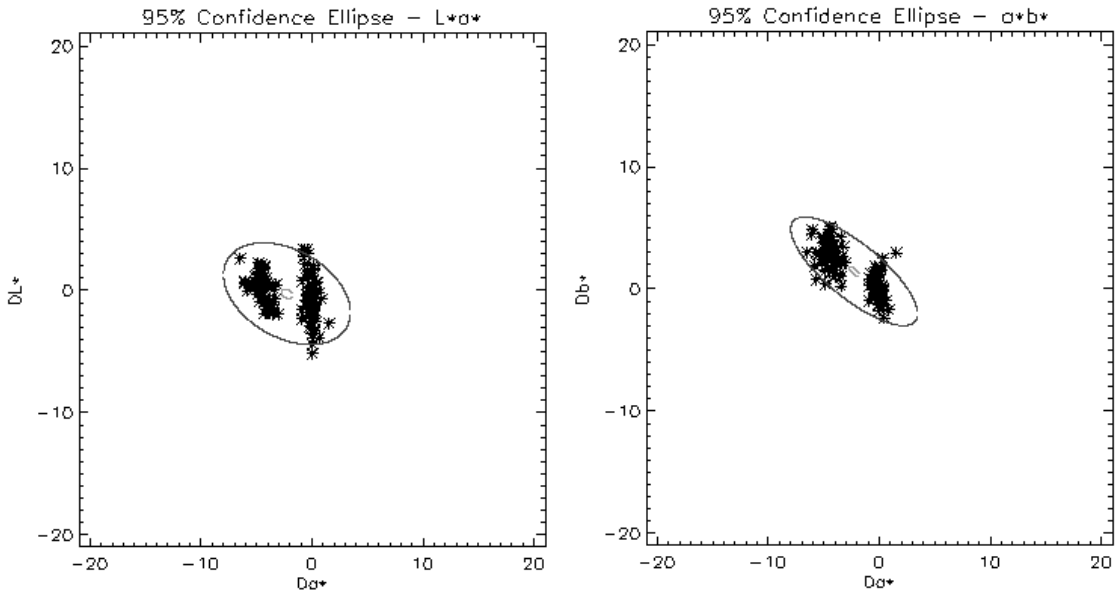


Figure 5.2 Confidence ellipse plots of the Lab 98 data set, using the CIE 1931 standard observer. The two 'clouds' of data points are thought to represent the color matches attained by the different primary sets used.

5.3 All Three Data Sets Combined

The results and confidence ellipse plots of the three data sets combined can be found in Appendix 1. The results have been summarized below in Table 5.5.

Color Difference	CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
ΔE^*_{ab}	4.56	4.02	4.14	4.36	4.14	4.37	3.92
ΔE^*_{94}	3.41	2.82	3.01	3.24	3.00	3.24	2.78

Table 5.5 : Average results for the computational analysis of all three data sets combined

It can be seen from Table 5.5 that the Shaw and Fairchild responsivity functions yield the lowest color difference, averaged over all of the samples in the combined data set. But, it is clear that the difference between each of the sets is only slight, ranging from 4.56 ΔE^*_{ab} at worst, to 3.92 ΔE^*_{ab} at best.

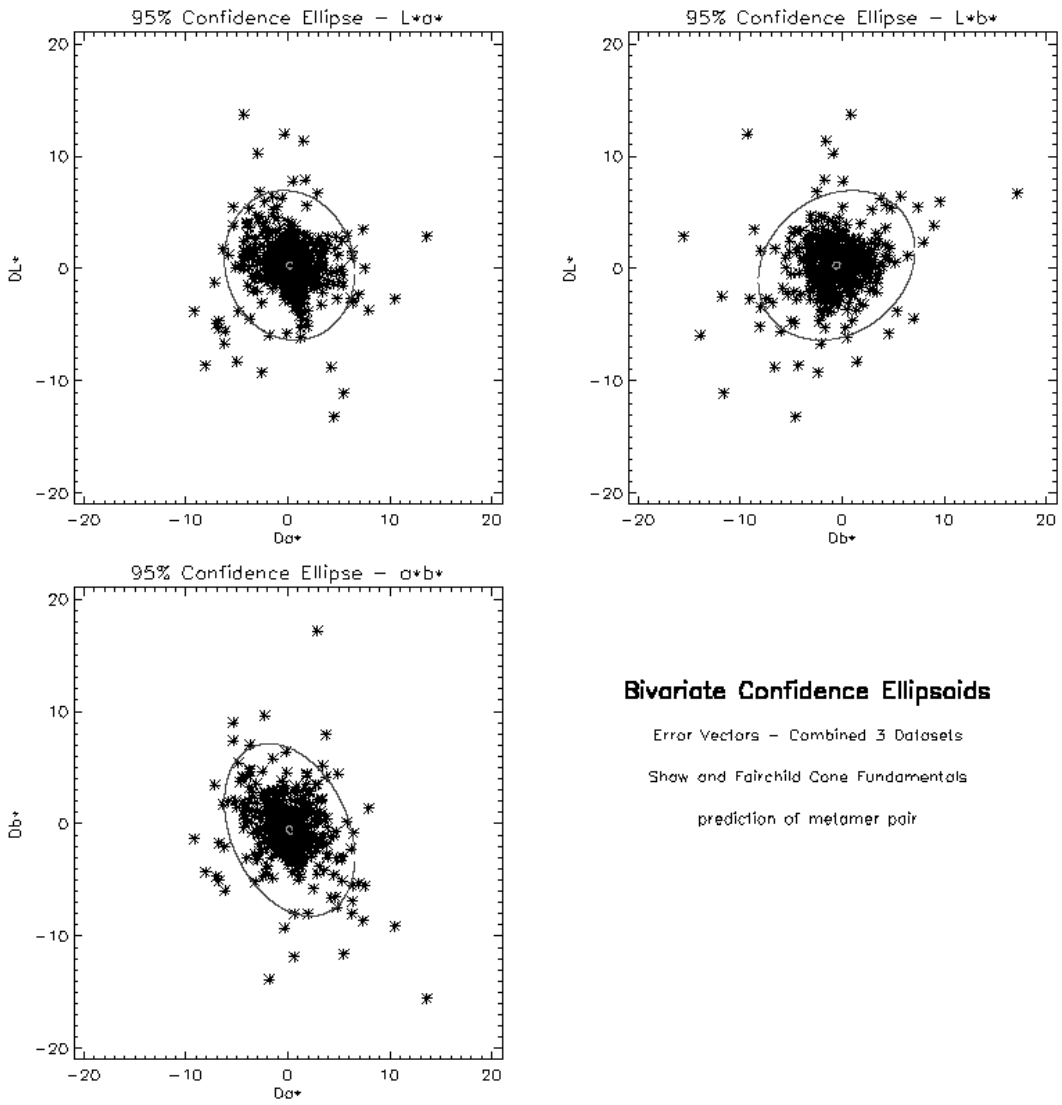


Figure 5.3 : 95% and mean confidence ellipses of Shaw and Fairchild responsivity functions, using all three combined data sets

When looking at the statistical results, it is clear that no one set of responsivity functions performs optimally in all three dimensions. There is always at least one dimension in which it is possible to reject the null hypothesis.

Although the statistics tend toward the decision that the means are significantly different from zero, it is important also to consider that the ΔL^*, Δa^*, Δb^* means are not far off zero, in all cases. The discrepancy is due to the fact that the data set consists of 468 metamer pairs, creating a very small mean ellipse.

The Shaw and Fairchild confidence ellipse plot for the combined data sets are shown in Figure 5.3. The mean ellipse can be seen in the confidence ellipse plots of Figure 5.3 and Appendix 1, the mean confidence ellipse being very small in comparison with the outer ellipse.

5.4 Thornton Metameric Pair

As discussed in chapter 1, the recent works of Thornton (1992abc, 1997, 1998ab) have focussed on determining an observer's color matching functions using several sets of primary lights for each observer. Using one of Thornton's data points, shown in Figure 5.4, a pair of spectral power distributions that stress the deficiency of the standard observer, the performance of each set of weighting functions was ascertained. The two stimuli visually matched to eight color normal observers.

Color Difference	CIE 1931	CIE 1964	Stiles	Demarco	Stockman	Vos	Shaw
ΔE^*_{ab}	5.35	9.75	6.75	4.84	6.74	4.83	7.96
ΔE^*_{94}	3.50	6.23	4.32	3.03	4.31	3.02	4.94

Table 5.6 : Results for the computational analysis using the Thornton metamer pair

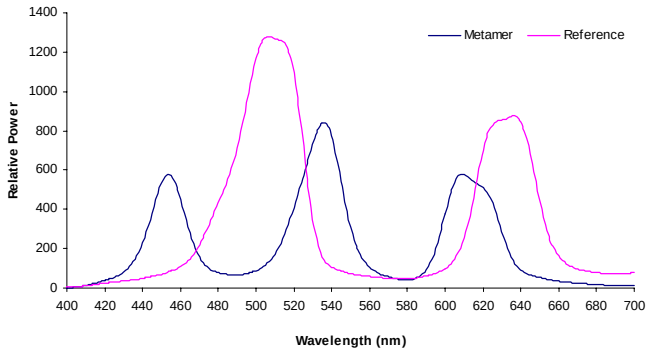


Figure 5.4 : Thornton's metameric pair

Table 5.6 shows the results of the computational analysis, it can be seen that the CIE 1964, and Shaw and Fairchild responsivity functions performed worst. Surprisingly the results point towards the Demarco, *et al.* and Vos and Walraven responsivity functions as being the best. It is important to note though that the Thornton data set only consisted of one metameric pair that was considered equivalent by eight observers. A much more rigorous data set is needed to make any conclusive decisions. The Thornton data does however agree with the Alfvin and Fairchild range of observer variability.

5.5 Derivation of Optimized Responsivity Functions

The derivation of an optimized set of responsivity functions is discussed later in Chapter 6. The results discussed above have shown that the optimized weighting function performs very well on the Lab '98 and Lab '99 data sets, but not so well on the Alfvin data set. This is surprising, since the Alfvin data set comprises over half the samples in the combined data set on which the optimization was performed. It is also evident that the Shaw and Fairchild responsivity functions exhibit similarities to the CIE 1964 10° color matching functions, producing similar results in all three data sets. One could conceivably improve the performance of the Shaw and Fairchild responsivity functions on the Alfvin data by weighting the importance of the Alfvin data set.

6 DERIVING AN OPTIMIZED SET OF RESPONSIVITY FUNCTIONS

It is possible to consider that an optimal set of responsivity functions can be derived from visual color matching data by a modification of an existing set of responsivity functions using the visual data to weight the adjustments. Using the three combined metameric data sets, one can derive an optimized set by weighting the minimization function with the color difference of the metameric pairs for multiple observers. Thus optimizing the relative weights by minimizing the average color difference over all observations.

6.1 Techniques Used

A variety of approaches have been taken to finding the optimal set of responsivity functions that best describes the data. The results of each of the techniques listed below are summarized in further detail, with possible reasoning as to why each technique obtained the results they did.

.i. Unconstrained Non-linear optimization, delta XYZ function

.ii. Constrained Non-linear optimization, delta XYZ function

.iii. Constrained Non-linear optimization, using 18 cubic spline functions, delta XYZ function

.iv. Constrained Non-linear optimization, baseline 0, using 18 cubic spline functions, delta XYZ functions

.v. Constrained Non-linear optimization, baseline 0, using 18 cubic spline functions, direct XYZ functions

.vi. Constrained Non-linear optimization, using 4 cubic spline functions, delta XYZ function

.vii. Linear Regression, using 18 cubic spline functions, delta XYZ function

.viii. Linear Regression, using Principal Component Analysis basis vectors, delta XYZ function

.ix. Linear Regression – I, using Sine and Cosine basis vectors, delta XYZ function

.x. Linear Regression – II, using Sine and Cosine basis vectors, delta XYZ function

.xi. Constrained Non-linear optimization, using multiple color matching functions, direct XYZ functions

.xii. Constrained Non-linear optimization, using CIE Standard Deviate Observer functions, delta XYZ functions

.xiii. Monte Carlo Method, using Sine and Cosine basis vectors, delta XYZ function

All analyses were performed using all 468 metamer pairs, combining 3 data sets – Alfvin (1995), Shaw and Montag (1998), and Shaw and Montag (1999).

6.1.1 Unconstrained Non-linear optimization, delta XYZ function

This optimization approach uses an in-built IDL™ non-linear minimization routine called Powell, using the Powell method. The optimization routine does not allow for constraints on any of the parameters. The cost function is shown in Equation 6.1.

$$\bar{x}_{i\,optimum} = \bar{x}_i + \Delta\bar{x}_i \qquad \longrightarrow \qquad \Delta\bar{x}_i = \text{Arg min}\sqrt{\frac{\left(X_{i,1} - X_{i,2}\right)^2}{n}} \qquad (6.1)$$

For the purposes of the optimization, each of the three individual responsivity functions shall be referred to as weighing functions. Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$). $X_{i,1}$ is the tristimulus value of the reference for the i^{th} weighting function, $X_{i,2}$ is the tristimulus value of the visual match (sample) for the i^{th} weighing function, and *n* is the number of observations. *Arg min* is the argument of $\Delta\bar{x}_i$ that minimizes the RMS error between the reference and sample tristimulus values.

The routine iterates to find the best fit delta function ($\Delta\bar{x}_i$) for $\bar{x}$, $\bar{y}$ and $\bar{z}$ to minimize RMS($\bar{x}$), RMS($\bar{y}$) and RMS($\bar{z}$) respectively. Freedom was given to all 61 wavelengths, thus allowing each wavelength to be individually optimized for each weighing function.

The optimized, average color difference obtained by this minimization method is $2.468\Delta E^*_{ab}$. It can be seen from Figure 6.1 that although the average color difference predicted was very low, the weighting functions are highly jagged and unsuitable as a set of weighting functions.

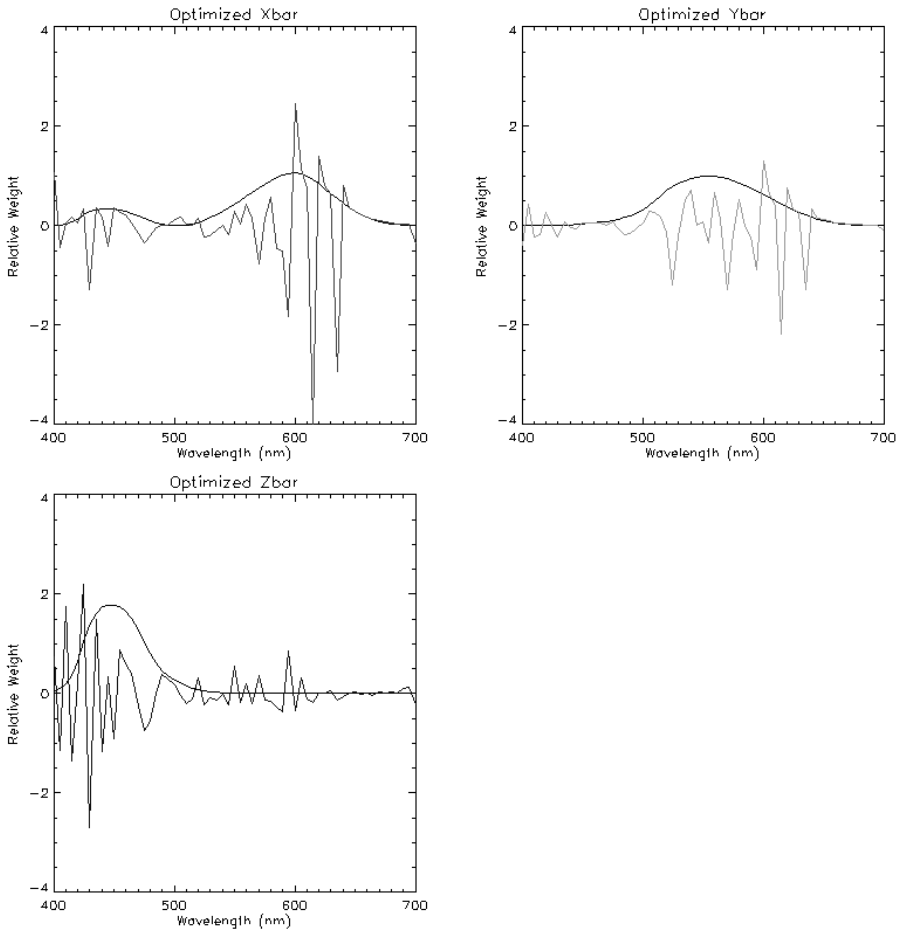


Figure 6.1 : Optimized weighting functions using the Powell method without constraint. In each case the black line represents the CIE 1931 weighting function, the colored line the optimized function.

6.1.2 Constrained Non-linear optimization, delta XYZ function

This technique uses the same optimization method as discussed in section 6.1.1, but this time an artificial constraint is built into the cost function that limits the magnitude of the delta function. The delta function is constrained to ± 0.4, shown in Equation 6.2.

$$\bar{x}_{i\,optimum} = \bar{x}_i + \Delta\bar{x}_i \qquad \longrightarrow \qquad \Delta\bar{x}_i = \text{Arg min}\sqrt{\frac{\left(X_{i,1} - X_{i,2}\right)^2}{n}} \qquad (6.2)$$

$$where \;\; -0.4 \geq \Delta\bar{x}_i \leq 0.4$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$). $X_{i,1}$ is the tristimulus value of the reference for the i^{th} weighting function, $X_{i,2}$ is the tristimulus value of the visual match (sample) for the i^{th} weighing function, and *n* is the number of observations. *Arg min* is the argument of $\Delta\bar{x}_i$ that minimizes the RMS error between the reference and sample tristimulus values.

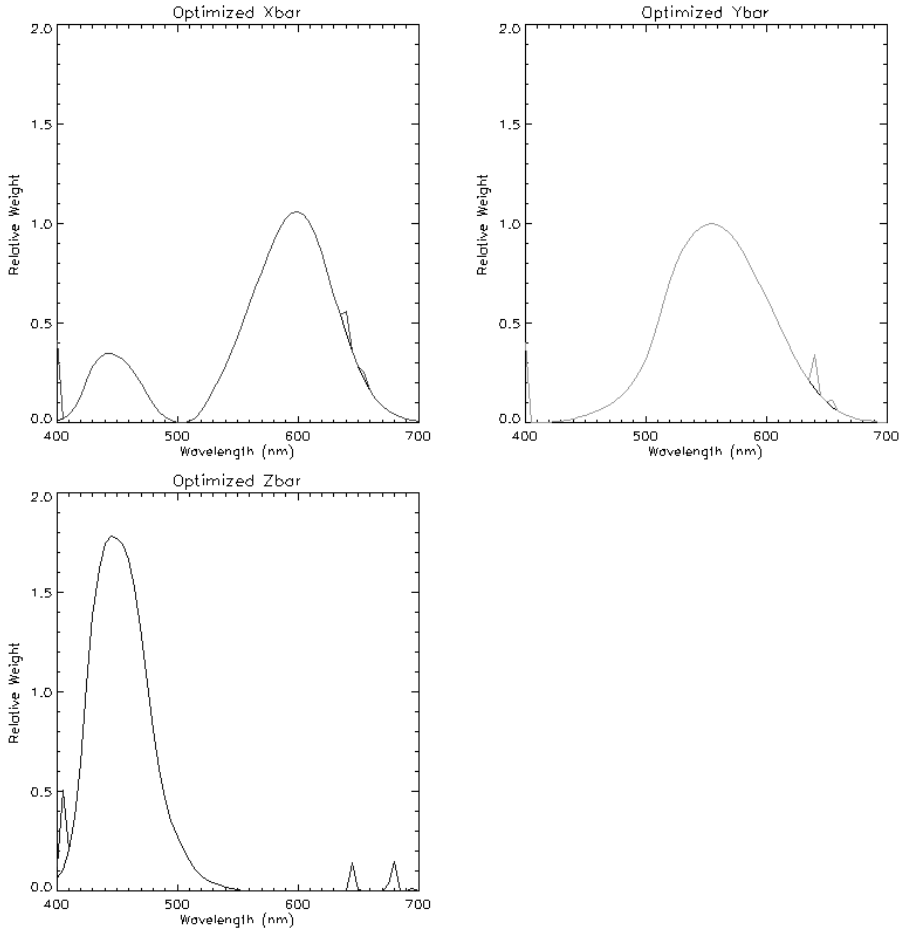


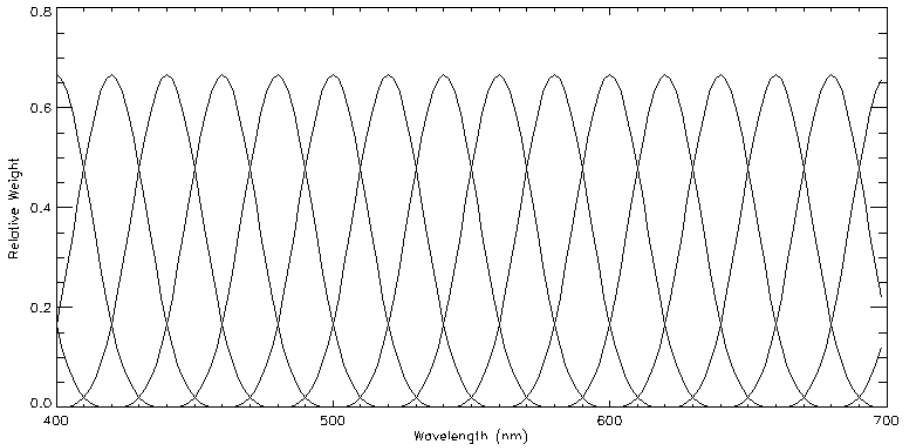
Figure 6.2 : Optimized weighting functions using the Powell method with constraint. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

The routine iterates to find the best fit delta function ($\Delta\bar{x}_i$) for $\bar{x}$, $\bar{y}$ and $\bar{z}$ to minimize RMS($\bar{x}$), RMS($\bar{y}$) and RMS($\bar{z}$) respectively. Freedom was given to all 61 wavelengths,

thus allowing each wavelength to be individually optimized for each weighing function. The optimized, average color difference obtained by this minimization method is $4.454\Delta E^*_{ab}$. It can be seen from Figure 6.2 that although the average color difference predicted varied little from that of the un-weighted functions, also, the weighting functions are jagged and unsuitable as a set of weighting functions.

6.1.3 Constrained Non-linear optimization, using 18 cubic spline functions, delta XYZ function

This optimization function uses the Levenberg-Marquardt technique, allowing constraints to be applied to different parameters in the optimization process. The results of techniques 6.1.1 and 6.1.2 have shown that allowing the optimization to vary all 61 wavelengths yields very jagged curves. To alleviate this, the use of cubic spline functions are employed. Splines give the benefit of a set of smooth functions being applied to the original spectra, the splines are shown below:



18 cubic spline functions

The use of splines also reduces the number of degrees of freedom (from 61 to 18) required in the data to adequately model the delta spectra. The weighted color matching

functions are then rotated in the cost function to represent the CIE xyz functions and the color difference calculated.

$$\bar{x}_{i,\,optimum} = \bar{x}_i + \sum_{j}^{18} w_{i,j} C_j \quad \longrightarrow \quad w_j = \text{Arg min}\, \Delta E^*_{\text{ab rotated}} \qquad (6.3)$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{j,i}$ is the j^{th} spline weight for the i^{th} weighting function, and C_j is the j^{th} cubic spline weighing function. *Arg min* is the argument of w_j that minimizes the color difference between the reference and sample tristimulus values.

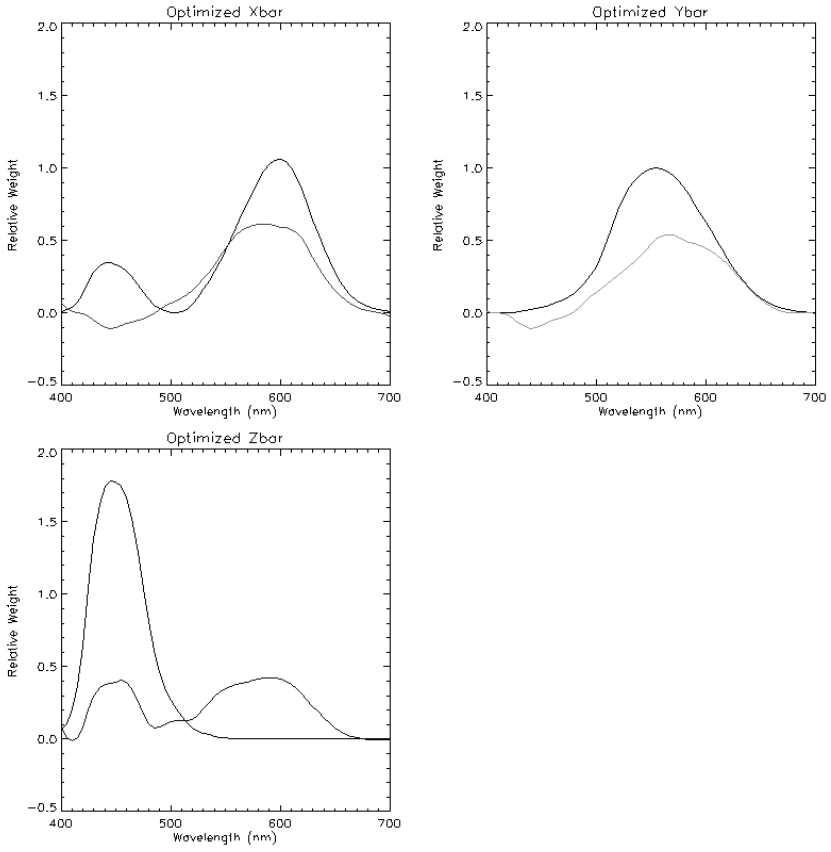


Figure 6.3 : Optimized weighting functions using the Levenberg-Marquardt method with constraint. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

In order to stop the optimized color matching functions averaging to zero, (computationally where ΔE^*_{ab} is minimum) a constraint is applied that restricts each spline weights to ± 1.0 about the CIE functions.

The optimized, average color difference obtained by this minimization method is $2.160\Delta E^*_{ab}$. It can be seen from Figure 6.3 that although the average color difference predicted was very low, the weighting functions are 'unrealistic', and unsuitable as a set of weighting functions. The tendency of the three functions to approach zero is indicative of the lower color difference.

6.1.4 Constrained Non-linear optimization, baseline 0, using 18 cubic spline functions, delta XYZ function

Using the 18 cubic spline functions defined in section 6.1.3, a solution for the optimum set of color matching function weights was sought that would constrain the difference functions to all positive values.

$$\bar{x}_{i\,optimum} = \bar{x}_i + \sum_{j}^{18} w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = \text{Arg min}\, \Delta E^*_{ab\,rotated} \qquad (6.4)$$

$$where \;\; w_{i,j} \geq 0$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{j,i}$ is the j^{th} spline weight for the i^{th} weighting function, and C_j is the j^{th} cubic spline weighing function. *Arg min* is the argument of w_j that minimizes the color difference between the reference and sample tristimulus values.

In order to stop the optimized difference functions going below zero, each spline weighting was constrained to ≥ 0. The optimum set of color matching functions are shown in figure 6.4

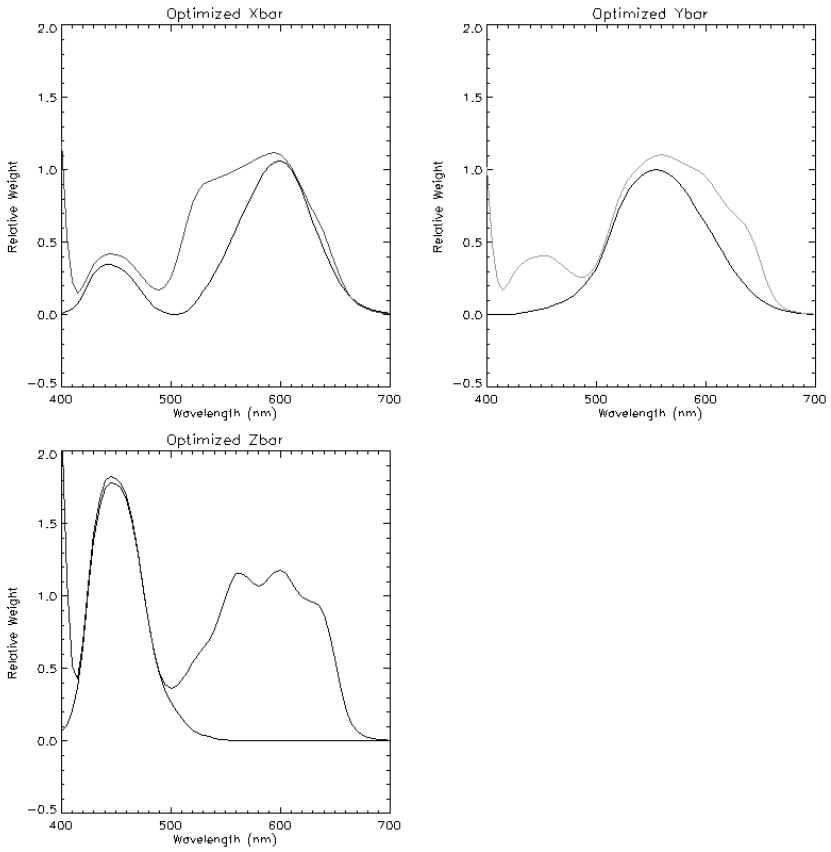


Figure 6.4 : Optimized weighting functions using the Levenberg-Marquardt method with constraint. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

The optimized, average color difference obtained by this minimization method is $2.512\Delta E^*_{ab}$. It can be seen from Figure 6.4 that although the average color difference predicted was very low, the weighting functions are 'unrealistic', and unsuitable as a set of weighting functions. The tendency of the blue function to inflate the $\bar{z}$ responsivity in the green region is indicative of an unrealistic set of functions.

6.1.5 Constrained Non-linear optimization, using 18 cubic spline functions, baseline 0, direct XYZ functions

Using the 18 cubic spline functions defined in section 6.1.3, a direct solution for the optimum set of color matching functions was sought that would restrict the functions to all positive values.

$$\bar{x}_{i,\,optimum} = \sum_{j}^{18} w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = \text{Arg min } \Delta E^*_{\text{ab rotated}} \qquad (6.5)$$

$$\textit{where } w_{i,j} \geq 0$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{j,i}$ is the j^{th} spline weight for the i^{th} weighting function, and C_j is the j^{th} cubic spline weighing function. *Arg min* is the argument of w_j that minimizes the color difference between the reference and sample tristimulus values.

As previously discussed, the weighted spline functions are then rotated in the cost function to represent the CIE $\bar{x}\ \bar{y}\ \bar{z}$ functions and the color difference calculated. In order to stop the optimized color matching functions going below zero, a constraint is applied that restricts the minimum spline weight to 0. The optimum set of color matching functions are shown in figure 6.5

The optimized, average color difference obtained by this minimization method is $1.076\Delta E^*_{ab}$. It can be seen from Figure 6.5 that although the average color difference predicted was very low, the weighting functions are ‘unrealistic’, and unsuitable as a set of weighting functions. The tendency of all of the functions to approach zero is a characteristic of the low color difference, but by no means a desirable property.

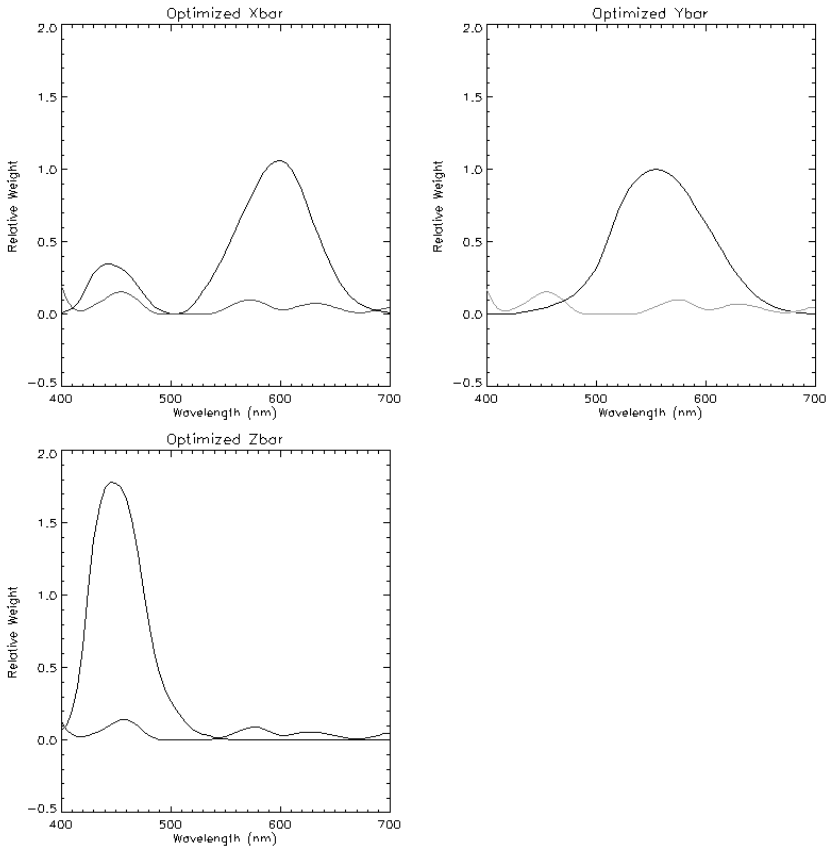


Figure 6.5 : Optimized weighting functions using the Levenberg-Marquardt method with constraint. In each case the black line represents the CIE standard weighting function, the colored line the optimized function. Generated using 18 constrained cubic splines.

6.1.6 Constrained Non-linear optimization, using 4 cubic spline functions, delta XYZ function

Using the same technique as in 6.1.3, but with only 4 cubic splines instead of 18, the optimization was again performed.

Whilst evaluating the data it was noticed that the metameric spectra were highly dependent. This is not desirable since dependent data will have inherent trends build into the results. In order to alter the color matching functions more accurately, one must ensure that the number of degrees of freedom in the data is more that the number of

spline functions, thus only 4 splines are used in this evaluation. The 4 cubic splines are shown in Figure 6.6 below.

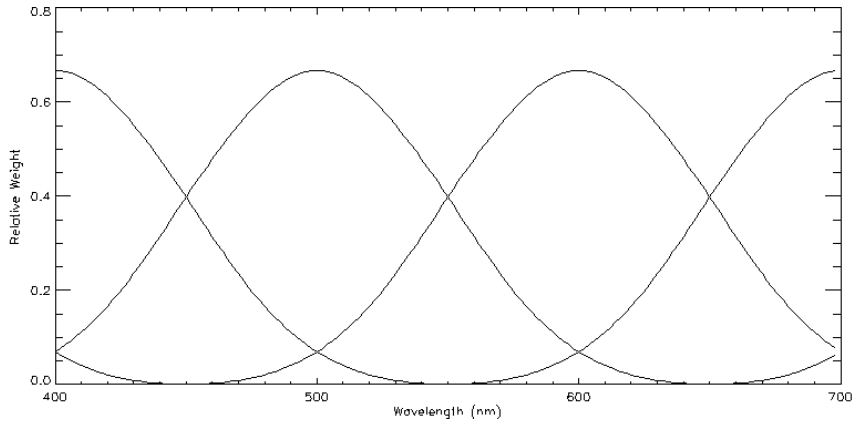


Figure 6.6 : 4 cubic spline functions

The weighted color matching functions are then rotated in the cost function to represent the CIE $\bar{x}\ \bar{y}\ \bar{z}$ functions and the color difference calculated.

$$\bar{x}_{i,optimum} = \bar{x}_i + \sum_j^4 w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = \text{Arg min}\, \Delta E^*_{\text{ab rotated}} \qquad (6.6)$$

where $-1 \le w_{i,j} \le 1$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{j,i}$ is the j^{th} spline weight for the i^{th} weighting function, and C_j is the j^{th} cubic spline weighing function. *Arg min* is the argument of w_j that minimizes the color difference between the reference and sample tristimulus values.

Again, a constraint is applied that restricts the spline weight $(w_{j,i})$ to ± 1.0 about the CIE functions. The optimized color matching functions are shown in Figure 6.7.

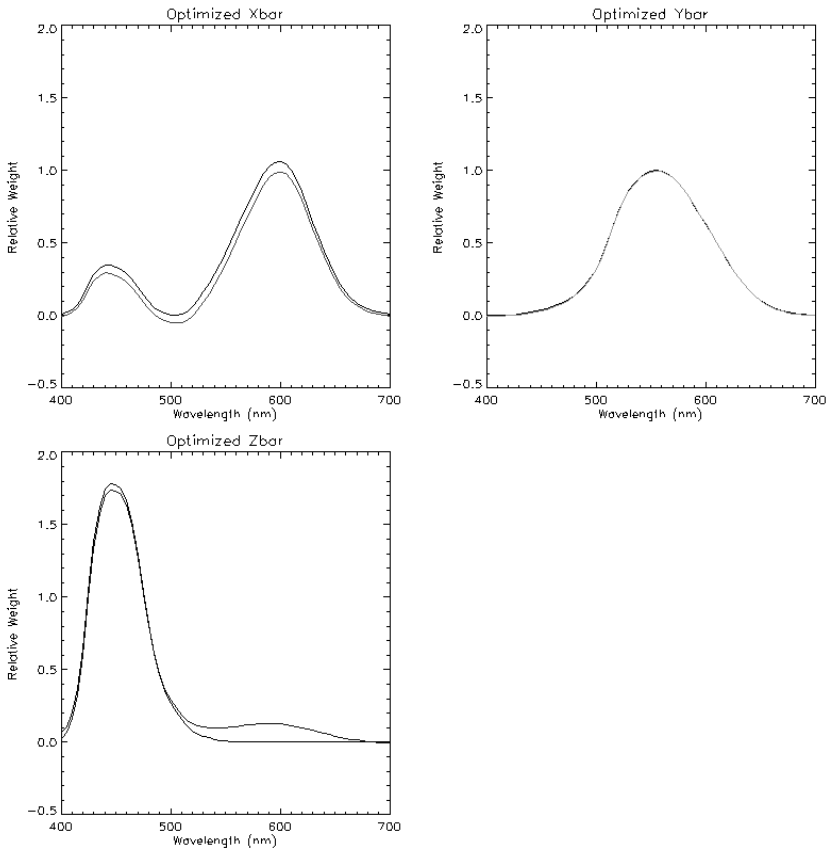


Figure 6.7 : Optimized weighting functions using the Levenberg-Marquardt method. Generated using 4 constrained cubic splines, with constraint. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

The optimized, average color difference obtained by this minimization method is $4.345\Delta E^*_{ab}$. This value of color difference suggests that although the weighting functions are smooth, little improvement can be made over the existing set of weighing functions using only 4 cubic splines.

6.1.7 Linear Regression, using 18 cubic spline functions, delta XYZ function

If one is to consider that the best fit color matching functions would result when ΔX, ΔY, and ΔZ (the tristimulus difference between the metameric sample and reference) are at a minimum, then one can derive a set of linear equations to calculate the change in the

color matching functions with respect to tristimulus error. The change in each weighting function (a.k.a. delta function) is constrained by the use of cubic spline functions, as documented in section 6.1.3.

As before, the optimized color matching functions are based on an existing set, with the addition of a delta function :

$$\bar{x}_{i,optimum} = \bar{x}_i + \sum_{j}^{18} w_{i,j} C_j \qquad \longrightarrow \qquad \frac{\partial}{\partial w_{i,j}} \sum_{k}^{n} \left(X_{i,1}^{k} - X_{i,2}^{k} \right)^2 = 0 \qquad (6.7)$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{j,i}$ is the j^{th} spline weight for the i^{th} weighting function, and C_j is the j^{th} cubic spline weighing function. $\frac{\partial}{\partial w_{i,j}}$ is the partial derivative with respect to $w_{i,j}$ that leads to the minimum tristimulus difference between the reference and sample tristimulus values.

Taking the derivative with respect to $w_{i,j}$, the spline weightings, and using least squares regression over all samples should yield the weights. The weights calculated for the $\bar{x}$ function are shown in Figure 6.8.

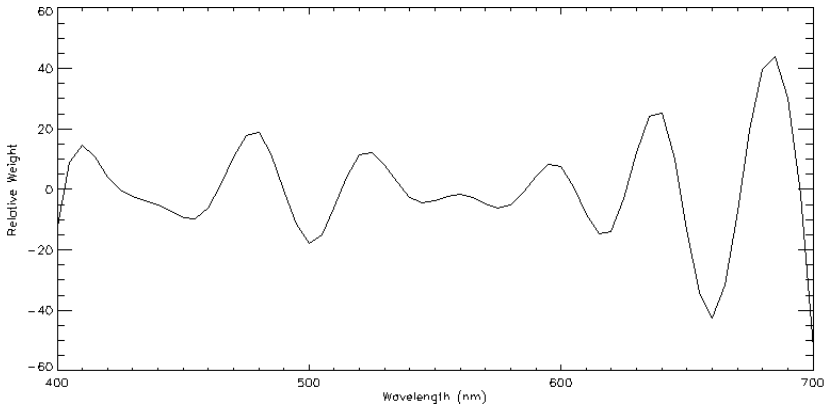


Figure 6.8 : Delta $\bar{x}$ function derived using linear regression with weighted cubic splines.

The delta $\bar{x}$ function is clearly over predicting the error within the matches. An in depth analysis led to the conclusion that errors result from a high dependency of the data. Linear regression assumes that all variables are independent, not so in this case. Therefore the use of this technique is not based on a solid foundation.

6.1.8 Linear Regression, using PCA basis vectors, delta XYZ function

This analysis looked at the number of eigen-vectors in the difference spectra (spectral difference between the metameric pair), and found that 4 vectors were sufficient to describe a high proportion of the variance within the data.

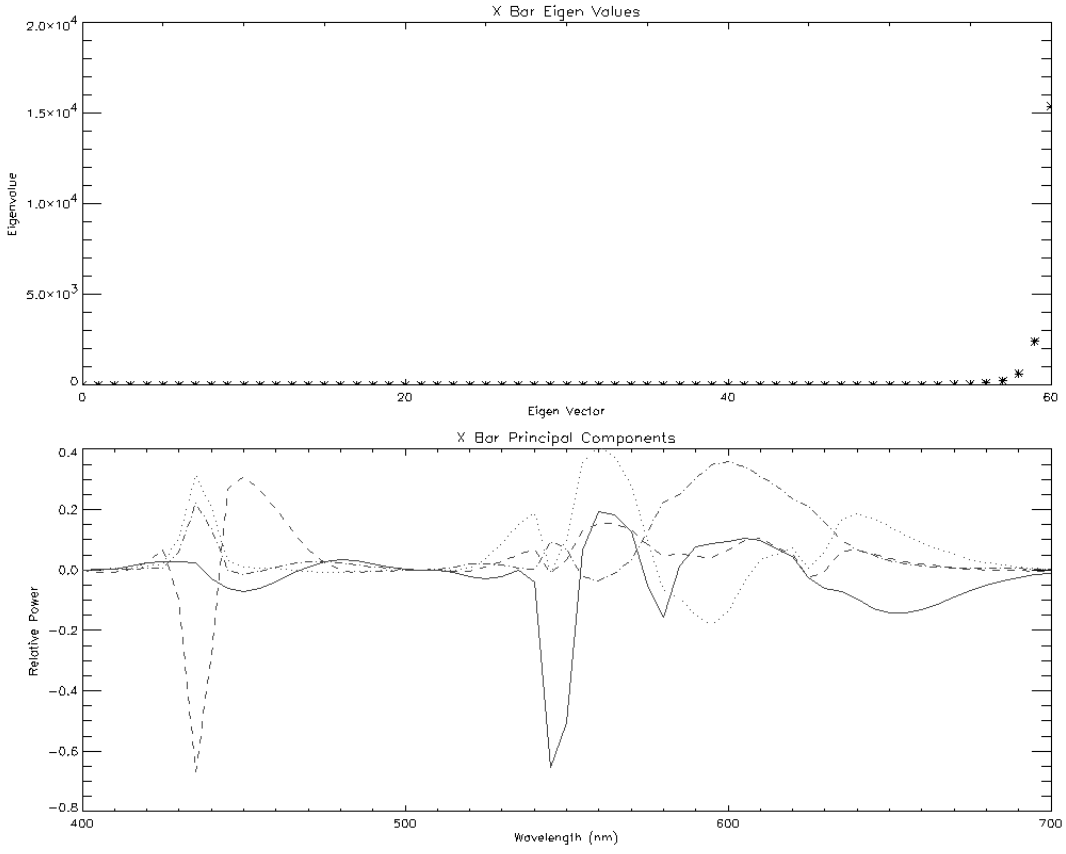


Figure 6.9 : (Top) Variance explained by each of the eigenvectors in the delta $\bar{x}$ spectra. It can be seen that the data is highly dependent, and the majority of variance can be explained by the last 4 vectors. (Bottom) Plots of the 4 most significant eigenvectors.

Figure 6.9 shows the variance explained by each eigenvector, and the four vectors that explain the highest proportion of variance within the delta $\bar{x}$ data. Using the assumption that an optimized set of color matching functions can be derived from a linear combination of the four principal eigenvectors plus the weighting functions, the same regression as in section 6.1.7 was performed using the first 4 principal components as the splines.

$$\bar{x}_{i,\,optimum} = \bar{x}_i + \sum_{j}^{4} w_{i,j} C_{i,j} \qquad \longrightarrow \qquad \frac{\partial}{\partial w_{i,j}} \sum_{\mathrm{k}}^{\mathrm{n}} \left(X_{i,1}^{k} - X_{i,2}^{k} \right)^2 = 0 \qquad (6.8)$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} vector weight for the i^{th} weighting function, and $C_{i,j}$ is the j^{th} eigenvector for the i^{th} weighting function. $\frac{\partial}{\partial w_{i,j}}$ is the partial derivative with respect to $w_{i,j}$ that leads to the minimum tristimulus difference between the reference and sample tristimulus values.

Taking the derivative with respect to $w_{i,j}$, the eigenvectors, and using least squares regression over all samples yields the weights. The weights calculated for the $\bar{x}$, $\bar{y}$, and $\bar{z}$ functions are shown in Figure 6.10. The regression was performed without any constraints.

	Optimized	*Original*
$\Delta\bar{x}$	38.67	14664.73
$\Delta\bar{y}$	81.87	17350.03
$\Delta\bar{z}$	0.42	25057.55
ΔE^{*}_{ab}	10.29	4.56

Table 6.1 : Results of Linear Regression with PCA basis vectors. It is evident that the optimization minimized the tristimulus difference very well, but that does not represent an adequate minimization of color difference.

Table 6.1 shows the optimized average tristimulus difference, and color difference calculated by the regression. It can be seen that although the average tristimulus difference was minimized very well indeed, no constraints were made that the resulting functions had to be ‘realistic’ weighing functions. This can be seen both by looking at the optimized average color difference in Table 6.1, and by looking at the resulting curves in Figure 6.10.

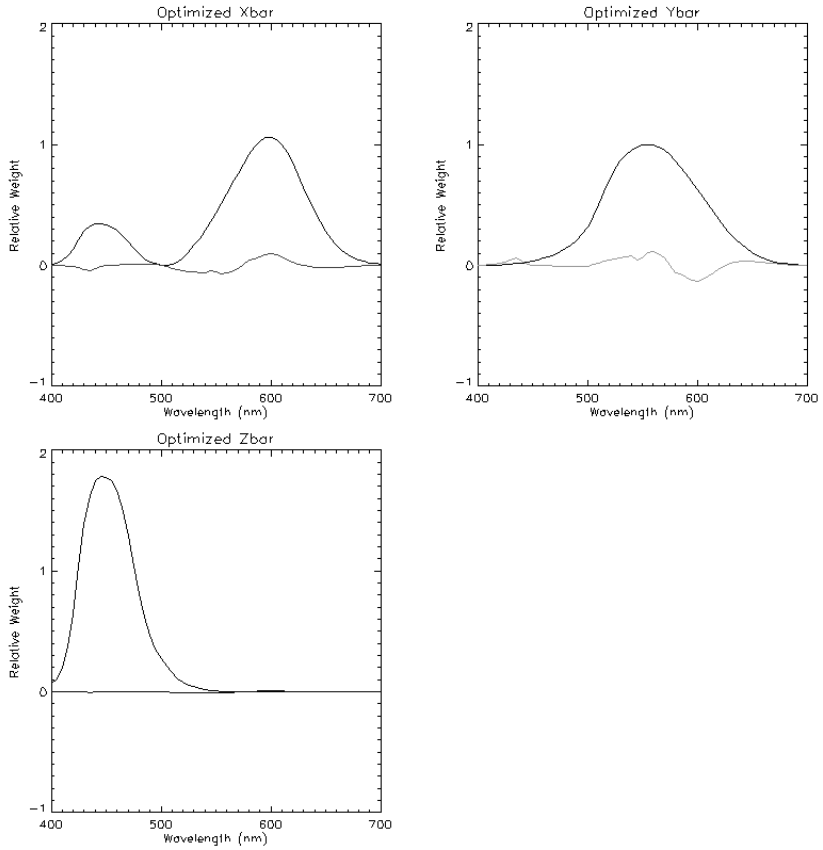


Figure 6.10 : Optimized weighting functions using linear regression with PCA basis vectors. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

It is also interesting to note in Table 6.1 that although the tristimulus difference decreases, the color difference increases from 4.56 to 10.29 ΔE^*_{ab}. This can be explained by the fact that the optimized weighting functions approach zero.

6.1.9 Linear Regression – I, using Sine and Cosine basis vectors, delta XYZ function

This analysis builds upon the assumption of fourier analysis that states that any wave-pattern can be decomposed into a collection of sine and cosine waves. Therefore the regression can be performed using sine and cosine waves as the basis vectors. These basis vectors are shown in Figure 6.11.

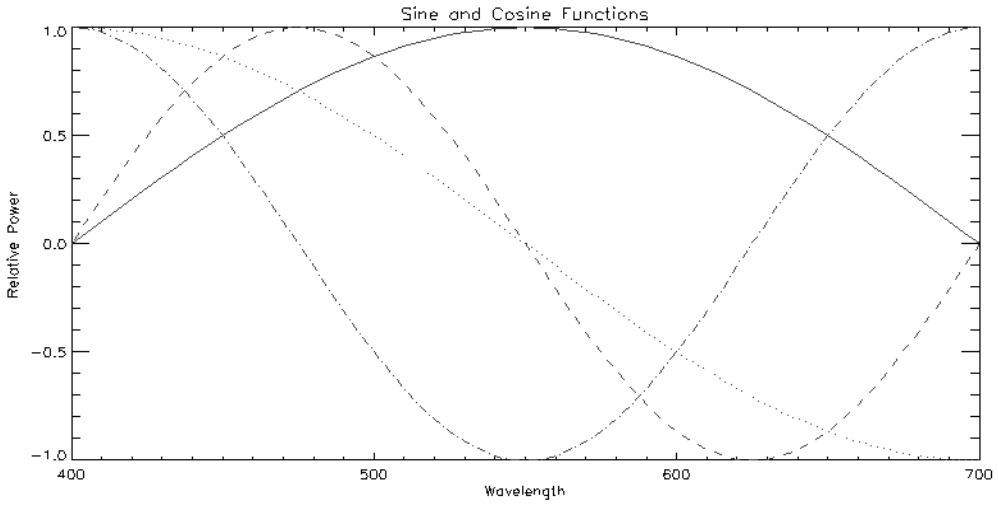


Figure 6.11 : Sine and Cosine waves to be used as basis vectors

Using the assumption that an optimized set of color matching functions can be derived from a linear combination of the four basis vectors plus the basis functions, the same regression as in section 6.1.7 was performed using the basis vectors as the splines.

$$\bar{x}_{i,\,optimum} = \bar{x}_i + \sum_j^4 w_{i,j} C_{i,j} \qquad \longrightarrow \qquad \frac{\partial}{\partial w_{i,j}} \sum_{\mathrm{k}}^{\mathrm{n}} \left(X_{i,1}^{k} - X_{i,2}^{k} \right)^2 = 0 \qquad (6.9)$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} vector weight for the i^{th} weighting function, and $C_{i,j}$ is the j^{th} eigenvector for the i^{th} weighting function. $\frac{\partial}{\partial w_{i,j}}$ is the partial derivative with respect to $w_{i,j}$ that leads to the minimum tristimulus difference between the reference and sample tristimulus values.

Taking the derivative with respect to $w_{i,j}$, the eigenvectors, and using least squares regression over all samples yields the weights. The weights calculated for the $\bar{x}$, $\bar{y}$, and $\bar{z}$ functions are shown in Figure 6.12. The regression was performed without any constraints.

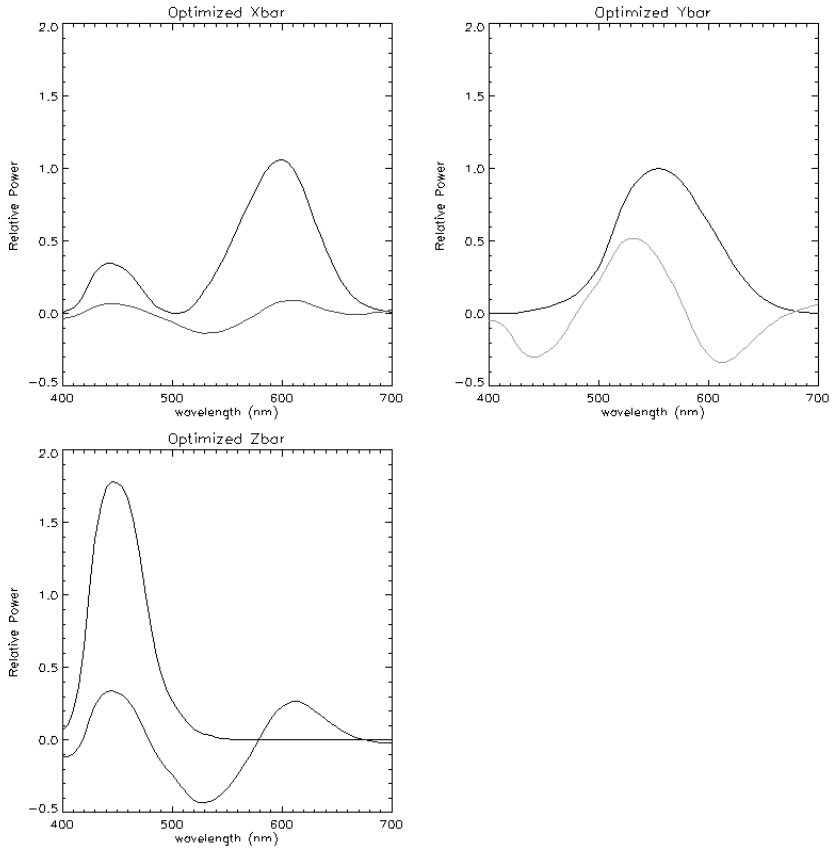


Figure 6.12 : Optimized weighting functions using linear regression with Sine and Cosine basis vectors. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

Table 6.2 shows the results of the minimization. As with the previous part, the average color difference is very high due to the minimization being optimized for tristimulus error.

	Optimized	*Original*
$\Delta\bar{x}$	77.87	14664.73
$\Delta\bar{y}$	1443.75	17350.03
$\Delta\bar{z}$	970.04	25057.55
ΔE^*_{ab}	22.32	4.56

Table 6.2 : Results of Linear Regression with Sine and Cosine basis vectors. It is evident that the optimization minimized the tristimulus difference very well, but that does not represent an adequate minimization of color difference.

One can see that the optimized vectors in Figure 6.12 are by no means 'realistic', the resulting vectors approaching an integrated area of zero.

6.1.10 Linear Regression – II, using Sine and Cosine basis vectors, delta XYZ function

This approach uses a similar technique to that in section 6.1.9 but differs in that the regression minimizes the distance between the pair of tristimulus values and a set of tristimulus values resulting from a linear combination of the basis vectors, rather than the tristimulus difference.

$$\bar{x}_{i,\,optimum} = \bar{x}_i + \sum_j^4 w_{i,j} C_{i,j} \qquad \longrightarrow \qquad \frac{\partial}{\partial w_{i,j}} \sum_{\mathrm{k}}^{\mathrm{n}} \left(\left(\frac{X_{i,1}^k - X_{i,2}^k}{2} \right) \cdot \left(X_{i,3}^k - X_{i,4}^k \right) \right) = 0 \qquad (6.10)$$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} vector weight for the i^{th} weighting function, and $C_{i,j}$ is the j^{th} eigenvector for the i^{th} weighting function. $\frac{\partial}{\partial w_{i,j}}$ is the partial derivative with respect to $w_{i,j}$ that leads to the minimum

tristimulus difference between the reference and sample tristimulus values. $X_{i,1}$ is the tristimulus value of the reference using the i^{th} weighting function, $X_{i,2}$ is the tristimulus value of the match using the i^{th} weighting function. $X_{i,3}$ is a vector of tristimulus weights of the reference integrated against each of the basis vectors, $X_{i,4}$ is a vector of tristimulus weights of the match integrated against each of the basis vectors.

Taking the derivative with respect to $w_{i,j}$, the eigenvectors, and using least squares regression over all samples yields the final weights. The weights calculated for the $\bar{x}$, $\bar{y}$, and $\bar{z}$ functions are shown in Figure 6.13. The regression was performed without any constraints.

Table 6.3 shows the results of the minimization. It can be seen that although the optimized tristimulus differences are larger, the color difference is much more realistic.

	Optimized	*Original*
$\Delta\bar{x}$	17740.62	14664.73
$\Delta\bar{y}$	18373.20	17350.03
$\Delta\bar{z}$	30146.77	25057.55
ΔE^*_{ab}	6.04	4.56

Table 6.3 : Results of Linear Regression with Sine and Cosine basis vectors. It is evident that the optimization minimized the tristimulus difference very well, but that does not represent an adequate minimization of color difference.

One can see that the optimized vectors in Figure 6.13 are much more realistic than the sets of color matching functions generated in section 6.1.9. Unfortunately the color difference is larger than the original set of weighting functions.

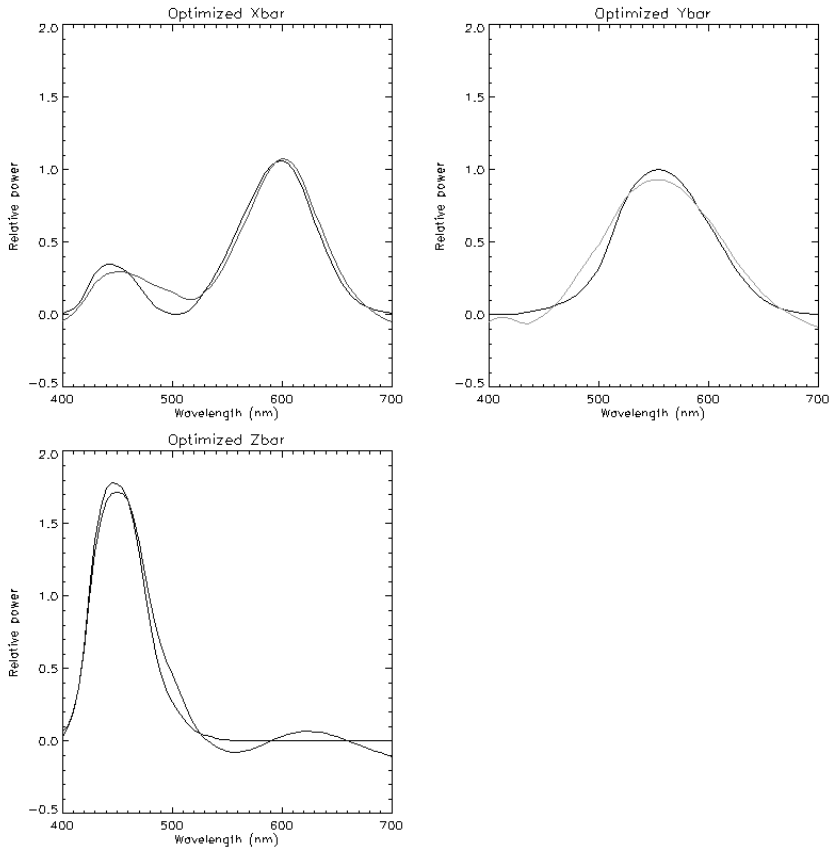


Figure 6.13 : Optimized weighting functions using linear regression with Sine and Cosine basis vectors. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

6.1.11 Constrained Non-linear optimization, sum eq. 1, using multiple color matching functions, direct XYZ functions

Using a similar approach as that defined in section 6.1.3 a solution for the optimum set of color matching function weights was sought that would utilize existing responsivity functions. The sum of the weights being constrained to 1.

$$\bar{x}_{i,optimum} = \sum_{j}^{6} w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = \text{Arg min}\, \Delta E^*_{\text{ab rotated}} \qquad (6.11)$$

where $\quad w_{i,1} = 1 - \sum_{k=2}^{6} w_{i,k}$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} weight for the i^{th} weighting function, and C_j is the j^{th} responsivity function. *Arg min* is the argument of $w_{i,j}$ that minimizes the color difference between the reference and sample tristimulus values.

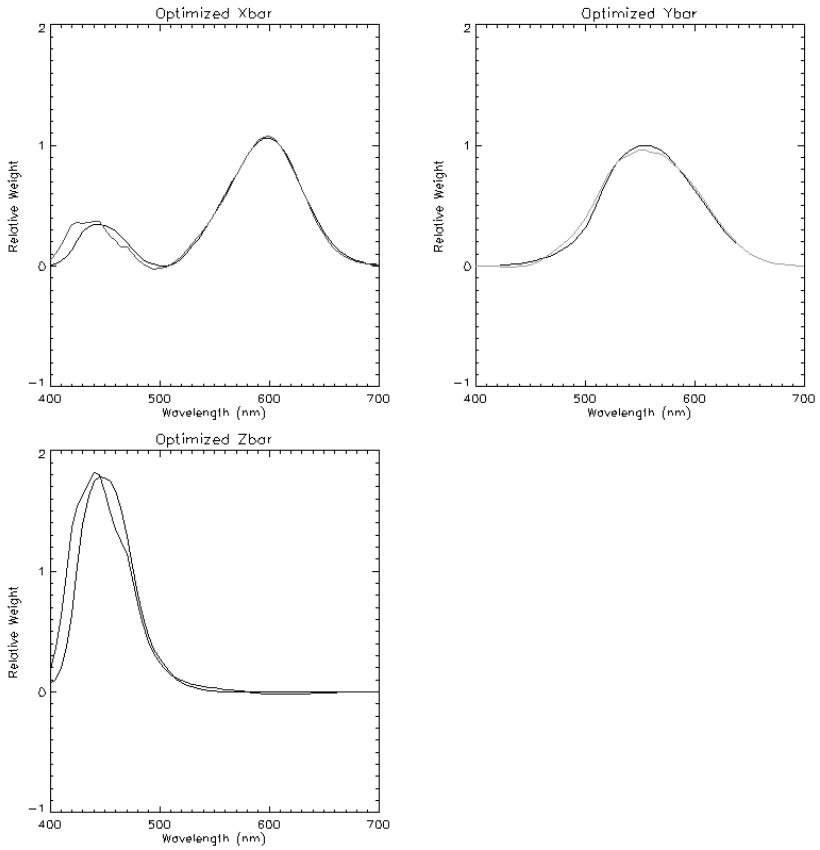


Figure 6.14 : Optimized weighting functions using multiple sets of color matching functions. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

The weighted responsivity functions are rotated in the cost function to represent the CIE $\bar{x}\ \bar{y}\ \bar{z}$ functions and the color difference calculated. In order to stop the optimized color matching functions going below zero, a constraint is applied that restricts the sum of the weights to 1. The optimum set of color matching functions are shown in figure 6.14

The optimized, average color difference obtained by this minimization method is $3.921\Delta E^*_{ab}$. It can be seen from Figure 6.14 that the weighting functions are very realistic, and potentially suitable as a set of weighting functions.

It is clear to see that there are distinct differences between the CIE 1931 color matching functions and the optimized responsivity functions. Both the $\bar{x}$ and $\bar{z}$ functions exhibit a sensitivity shift in the short wavelength region, the $\bar{y}$ function has also changed slightly.

6.1.12 Non-linear optimization, using CIE Standard Deviate Observer functions, delta XYZ functions

Using a similar approach as that in section 6.1.11, a solution for the optimum set of color matching function weights was sought that would utilize existing standard deviate observer functions. In order to maintain the shape of the color matching functions, the optimization was constrained to ± 5 deviation functions about the CIE functions.

$$\bar{x}_{i,optimum} = \bar{x}_i + \sum_j^4 w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = \text{Arg min}\ \Delta E^*_{\text{ab rotated}} \qquad (6.12)$$

where $\quad -5 \le w_{i,j} \le 5$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} weight for the i^{th} weighting function, and C_j is the j^{th} responsivity function. *Arg min* is the argument of $w_{i,j}$ that minimizes the color difference between the reference and sample tristimulus values.

The weighted responsivity functions are rotated in the cost function to represent the CIE $\bar{x}\ \bar{y}\ \bar{z}$ functions and the color difference calculated. In order to stop the optimized color matching functions going below zero, a constraint is applied that restricts each weight to ±6. The optimum set of color matching functions are shown in Figure 6.15

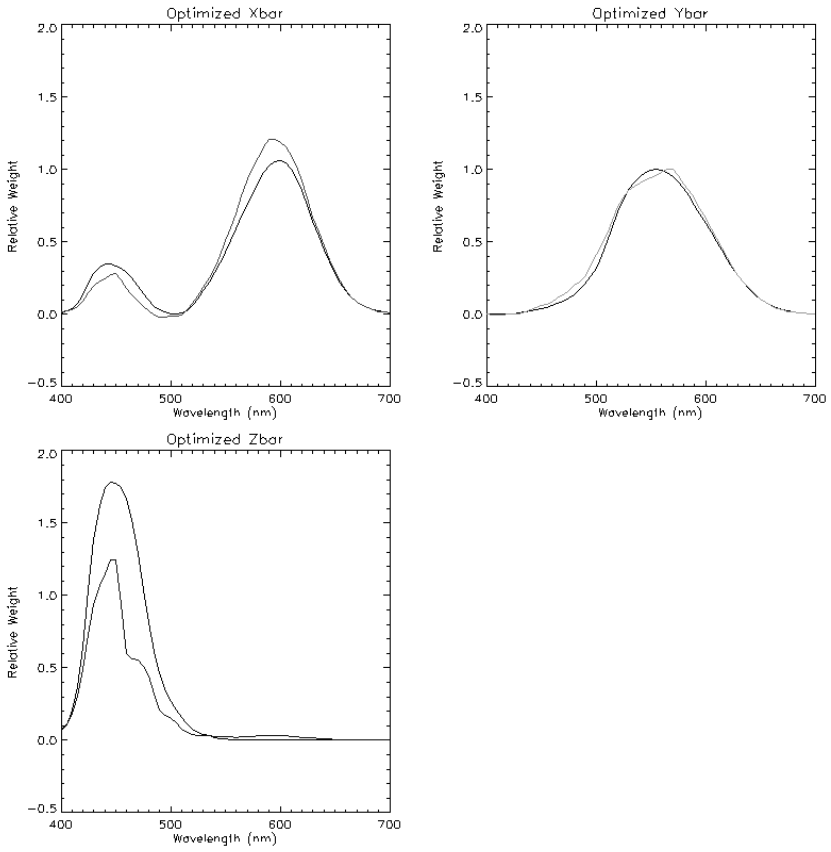


Figure 6.15 : Optimized weighting functions using the CIE Standard Deviate observer. In each case the black line represents the CIE standard weighting function, the colored line the optimized function

The optimized, average color difference obtained by this minimization method is $3.71\Delta E^*_{ab}$. It can be seen from Figure 6.15 that the weighting functions are somewhat realistic, and potentially suitable as a set of weighting functions. The main concern being the sensitivity of the $\bar{z}$ function is very low compared to the original CIE function.

6.1.13 Monte Carlo Method, using Sine and Cosine basis vectors, delta XYZ function

Using the assumption that an optimized set of color matching functions can be derived from a linear combination of four basis vectors plus basis functions, the monte carlo method was employed using randomly generated weights for each of the basis vectors. Sine and cosine basis vectors were used to generate smooth adjustments to the existing weighting functions. The computation was performed constraining the weights to ± 0.1 about the CIE functions.

$$\bar{x}_{i,optimum} = \bar{x}_i + \sum_{j}^{4} w_{i,j} C_j \qquad \longrightarrow \qquad w_{i,j} = Arg\ Min\ \Delta E^*_{ab\ rotated} \qquad (6.13)$$

where $-0.1 \le w_{i,j} \le 0.1$

Where $\bar{x}_i$ is the i^{th} weighting function under evaluation, ($\bar{x}$, $\bar{y}$, or $\bar{z}$), $w_{i,j}$ is the j^{th} weight for the i^{th} weighting function, and C_j is the j^{th} responsivity function. *Arg min* is the argument of $w_{i,j}$ that minimizes the color difference between the reference and sample tristimulus values.

Values of $w_{i,j}$ were chosen using a random number generator, assuming a uniform distribution with CPU time indexed seed. The weighted responsivity functions are rotated in the cost function to represent the CIE $\bar{x}\ \bar{y}\ \bar{z}$ functions and the color difference calculated. The optimum set of color matching functions are shown in Figure 6.16

The optimized, average color difference obtained by this minimization method is $4.15\Delta E^*_{ab}$. It can be seen from Figure 6.16 that the weighting functions are realistic, and potentially suitable as a set of weighting functions. It is interesting to note that the peak sensitivity of each of the three responsivity functions has increased or decreased to some degree.

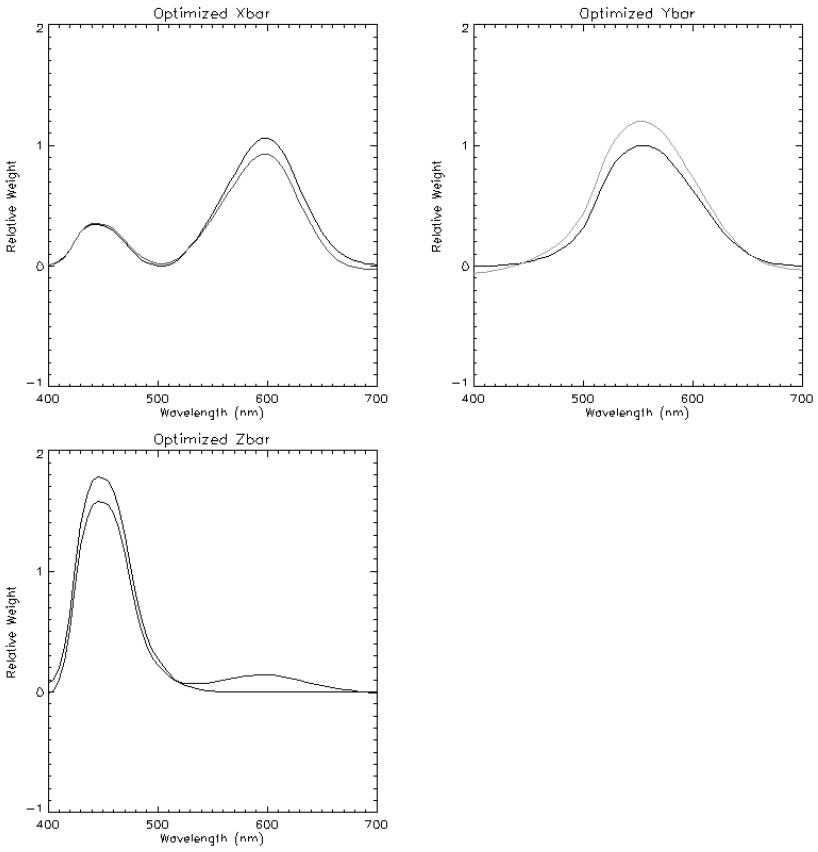


Figure 6.16 : Optimized weighting functions using Sine and Cosine splines and the Monte Carlo method. In each case the black line represents the CIE standard weighting function, the colored line the optimized function.

6.2 Optimization Summary

Of the thirteen techniques mentioned it is immediately clear that techniques i, ii, iii, iv, v, vii, viii, and ix are unsuitable. Those using the linear regression, minimizing tristimulus error, could potentially be improved by weighting the radiance spectra with Neugebauers CQF weighting (Neugebauer 1956, Engeldrum 1993), but one would not expect an exceptional improvement. The problem is partly related to the fact that an optimal color space minimizes color difference in a color space, such as CIELAB. The conversion of

tristimulus values into CIELAB coordinates includes the cube-root non-linear function, that is not accounted for in the tristimulus minimization.

Other reasons can be attributed to the dependency of the data itself. Ideally, when using a linear regression technique one assumes that each variable is independent. When using spectral data an optimization of all wavelengths would require independence of all wavelengths, this is clearly not the case in the color matches. In Alfvin's experiment the color matches were created with a CRT, thus only giving three dimensions in which that data can vary. The same being true for the Shaw '98 and Shaw '99 experiments, matches being made with only three primaries.

It is evident therefore that an optimization based on minimizing spectral RMS error will be underdetermined, and produce erroneous results.

Technique	*Av.* ΔE^*_{ab}	*Decision*
.i.	2.468	Rejected due to unrealistic weighting functions
.ii.	4.454	Rejected due to unrealistic weighting functions
.iii.	2.160	Rejected due to unrealistic weighting functions
.iv.	2.512	Rejected due to unrealistic weighting functions
.v.	1.076	Rejected due to unrealistic weighting functions
.vi.	4.345	Rejected due to little improvement
.vii.	N/A	Rejected due to unrealistic weighting functions
.viii.	10.290	Rejected due to unrealistic weighting functions
.ix.	22.320	Rejected due to unrealistic weighting functions
.x.	5.040	Rejected due to increase in average error
.xi.	3.921	Best, most realistic set chosen.
.xii.	3.710	Rejected due to $\bar{z}$ suspect sensitivity below 500nm.
.xiii.	4.150	Acceptable result, slight improvement over CIE

Table 6.4 : Summary table of the results of each different set of optimized weighting functions.

Of all the approaches shown in Table 6.4, technique xi performed the best. It did not yield the lowest color difference of all the techniques, but it did yield the most reasonable color difference and most realistic set of weighting functions. The final set of functions can be seen in Figure 6.17, resulting in an average color difference of $3.921\Delta E^*_{ab}$ over all samples.

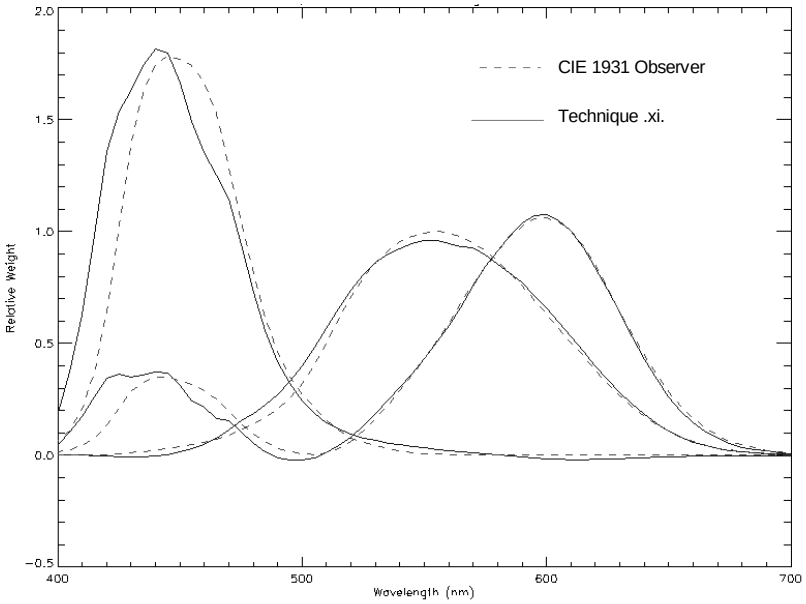


Figure 6.17 : Final selection of optimized color weighting functions

7 COMPUTER SIMULATION

In order to simulate the differences found in the different sets of responsivity functions, a computer simulation was designed using a characterized CRT. This chapter documents the computer simulation user interface and the characterization approach taken. The explicit techniques used to characterize the CRT are not documented, interested readers are directed to the work of Berns (1996), CIE (1996), and Berns *et al* (1993).

If one is to conceptualize a visual experiment in which an observer is seated in front of a viewing booth and CRT, see Figure 7.1. A spectro-radiometer is positioned by the observers left shoulder, measuring the radiance of samples positioned in the booth.

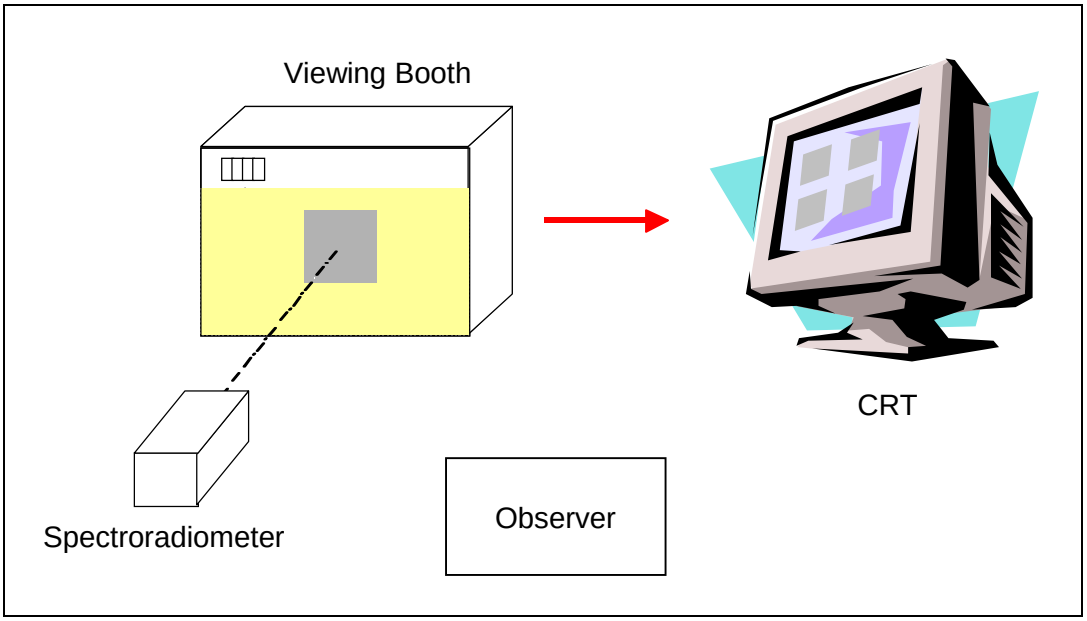


Figure 7.1 : Conceptual visual experiment. Observers comparison of various sets of responsivity functions

The observer places a sample in the viewing booth, under the controlled viewing conditions and measures the spectral radiance of the sample. The sample spectra is then passed to the computer and the simulation program displays the predicted CRT metameric match for that spectral radiance, using a spectral CRT model and each set of rotated responsivity functions.

The observer then looks at the predictions of each set of responsivity functions and can distinguish the differences between predictions.

7.1 Spectral Characterization of CRT

In order to model the CRT, enabling one to display the results of different sets of responsivity functions it was important to measure the CRT spectrally. Measuring the CRT primaries using a tristimulus colorimeter is insufficient since a characterization of the primaries must be done relative to each set of responsivity functions. Therefore, Figure 7.2 outlines the process involved in characterizing the display for each set of weighting functions.

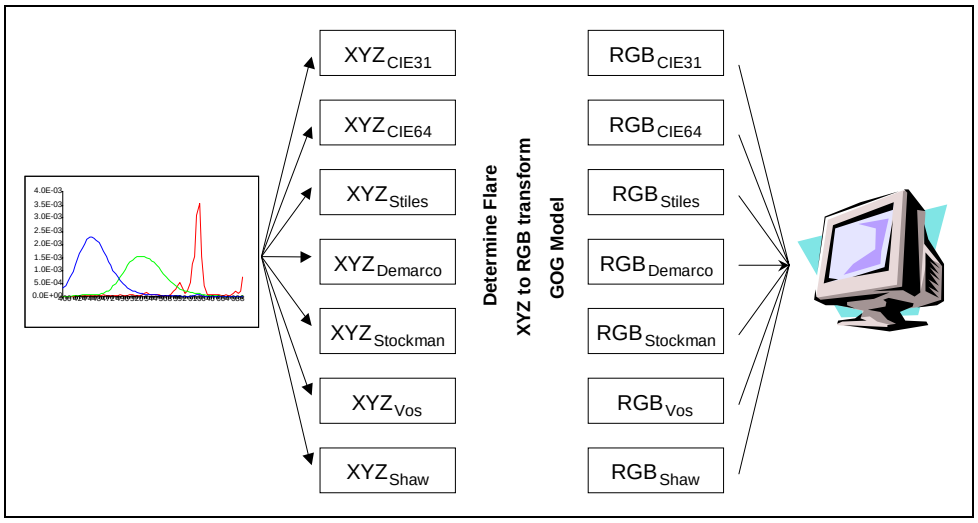


Figure 7.2 : Characterization outline of CRT display using multiple sets of responsivity functions

The spectral radiance of each primary, flare characteristics, and channel crosstalk properties were measured using a Photo-Research PR650 spectro-radiometer incorporating a half-height triangular bandpass of 4nm, recorded the spectral radiance at 4nm wavelength intervals across the visible spectrum between 380 and 780 nm.

Using each set of responsivity functions rotated into CIE representation, the tristimulus values were calculated. Then, for each set of tristimulus values, the CRT model is performed. First subtracting the flare, building the flare free XYZ to RGB transform matrix, then determining the GOG model parameters.

When the software application is used, the spectral radiance of the sample is treated in a similar fashion, shown in Figure 7.3.

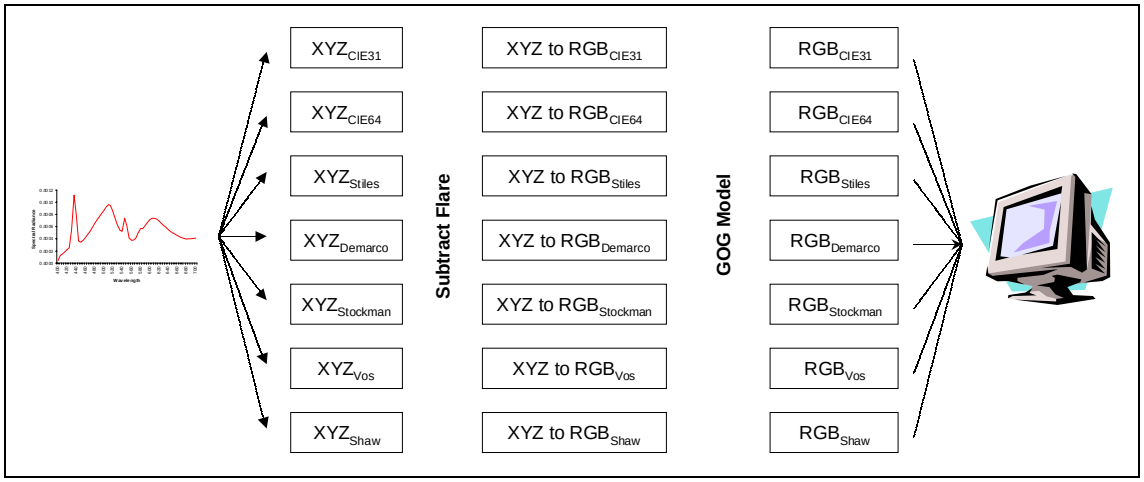


Figure 7.3 : Computational prediction of an input spectra using the 7 responsivities

The spectral radiance data is first used to calculate tristimulus values for each set of rotated responsivity functions. Each set of tristimulus values then has the flare subtracted by its respective flare measurement. The different sets of flare free tristimulus values are then transformed into RGB scalars using the respective transform. The GOG model then converts each set of RGB scalars into RGB dac's to be displayed by the application.

7.2 Software Simulation Package

The simulation was written in IDL[Ψ], the code can be found in Appendix 7. When launched, the application displays a start up screen welcoming the user to the application, shown in Figure 7.4.

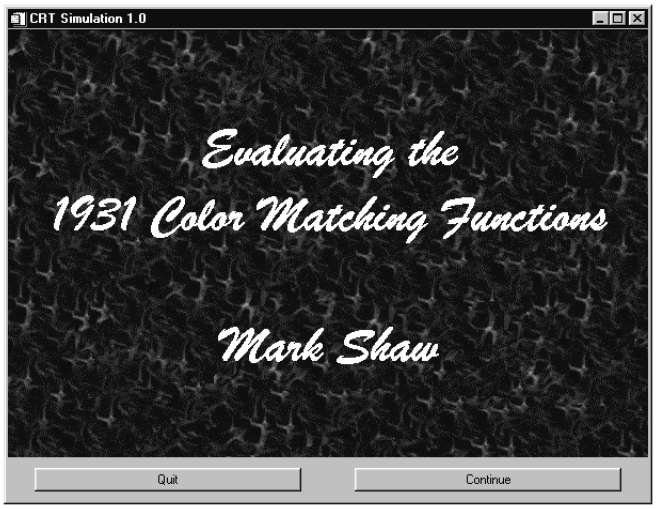


Figure 7.4 : Simulation welcome screen

When ready, the user clicks *continue* and the main control window is displayed, shown in Figure 7.5. The user must first select the type of spectral data being used, Radiance or Reflectance.

When *Radiance* is selected, clicking on the load spectral data button opens up a file selector window in which the user can load the specific spectral data file desired. The files are automatically sorted by extension and *'.rad'* files are considered to be spectral radiance data files.

[Ψ] IDL is a registered trademark of Research Systems, Inc. IDL, the Interactive Data Language, is software for data analysis, visualization, and cross-platform application development. For information go to the IDL home page : http://www.rsinc.com/

When *Reflectance* is selected, the *Reference Illuminant* selector becomes active, shown in Figure 7.7. The user must both select the illuminant, and load the spectral reflectance data file to be displayed. The files are automatically sorted by extension and *'.ref'* files are considered to be spectral reflectance data files.

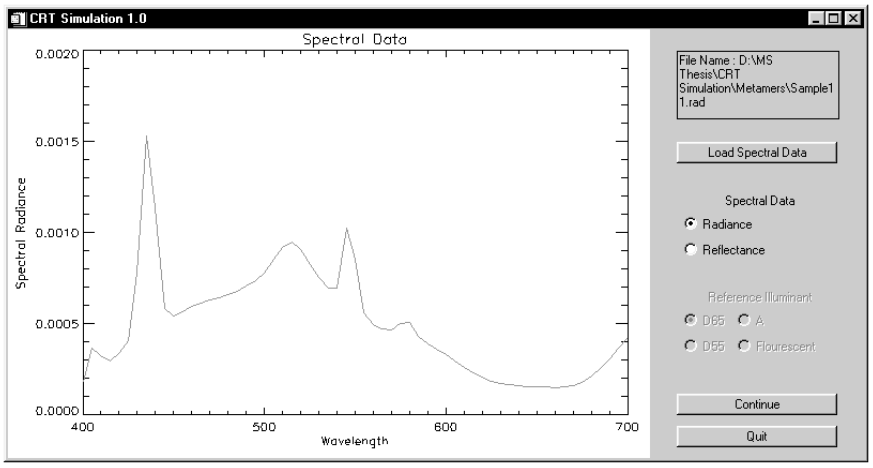


Figure 7.5 : Main Control Window - User selection of spectral data file. User selects either radiance or Reflectance spectral data, and can then load a data file from disc.

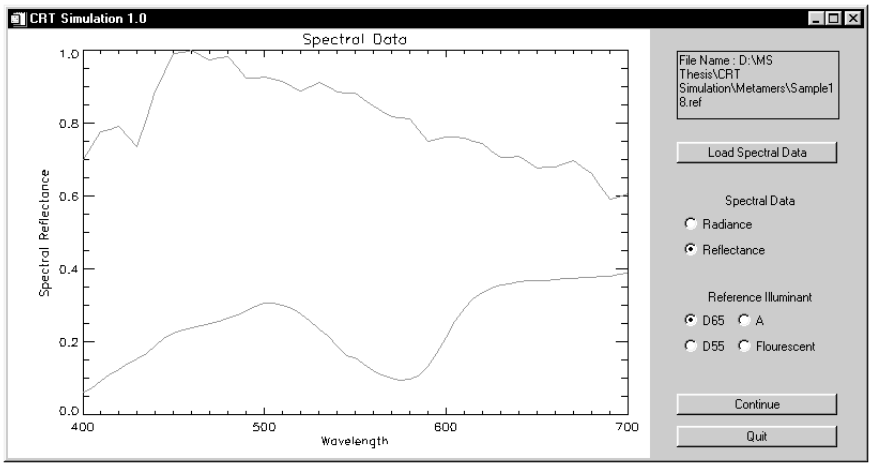


Figure 7.6 : Main Control Window - User selection of spectral data file. When Reflectance is selected the user must also specify the illuminant under which the sample is viewed.

Due to reflectance factor data being relative measurements, the absolute tristimulus values are not known. Therefore the illuminant white point is a assumed to be Y=100, and the transformation matrices constrained respectively.

When the spectral data file has been selected, the user the clicks the *continue* button. The main control window disappears and the prediction window opens, encompassing the whole screen. Shown in Figure 7.7.

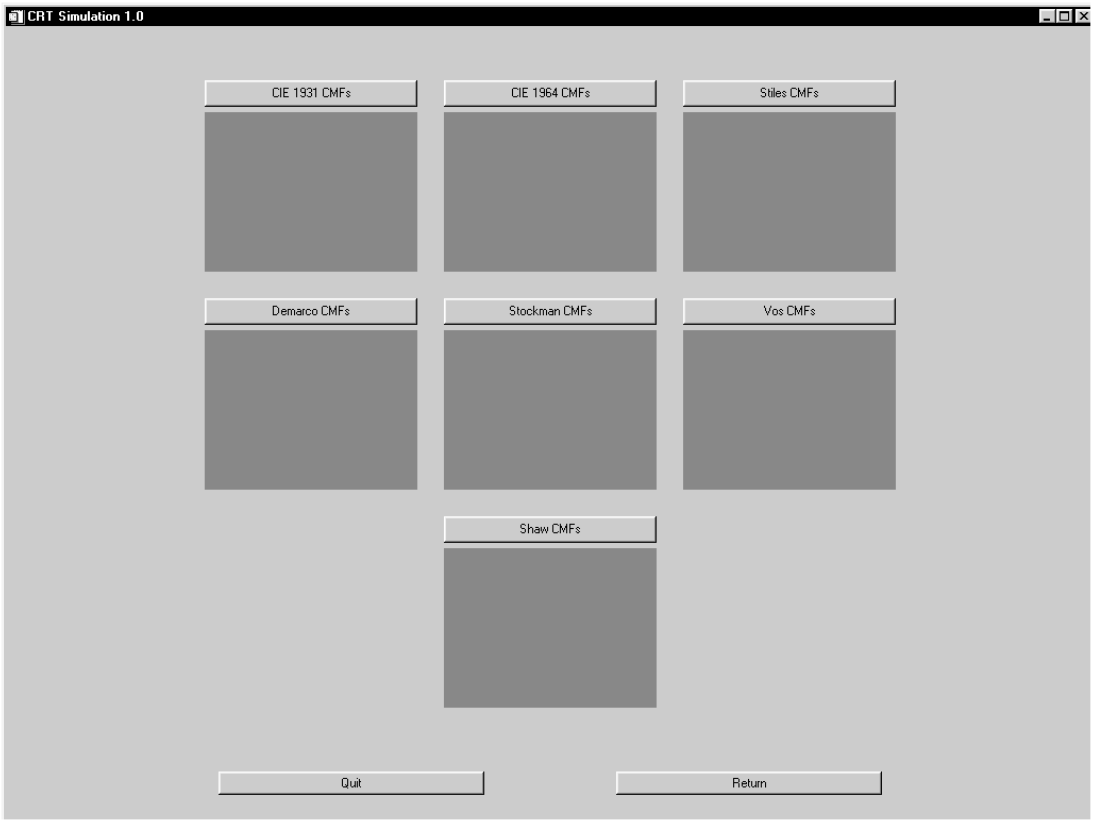


Figure 7.7 : Prediction Window - Simulation prediction of each of the 7 responsivity functions. It can be seen that subtle differences are noticeable.

When the user clicks on the *continue* button the spectral data file is used to calculate tristimulus values under each set of responsivity functions. The CRT model discussed earlier is then used to calculate the RGB digital counts (dac's) that will yield a

colorimetric match to the measured spectral data. The RGB dac's for each patch are then displayed in their respective position.

If for some reason the measured spectral reflectance or radiance is out of the gamut of the display, then a gamut warning is displayed in place of the color, shown in Figure 7.8.

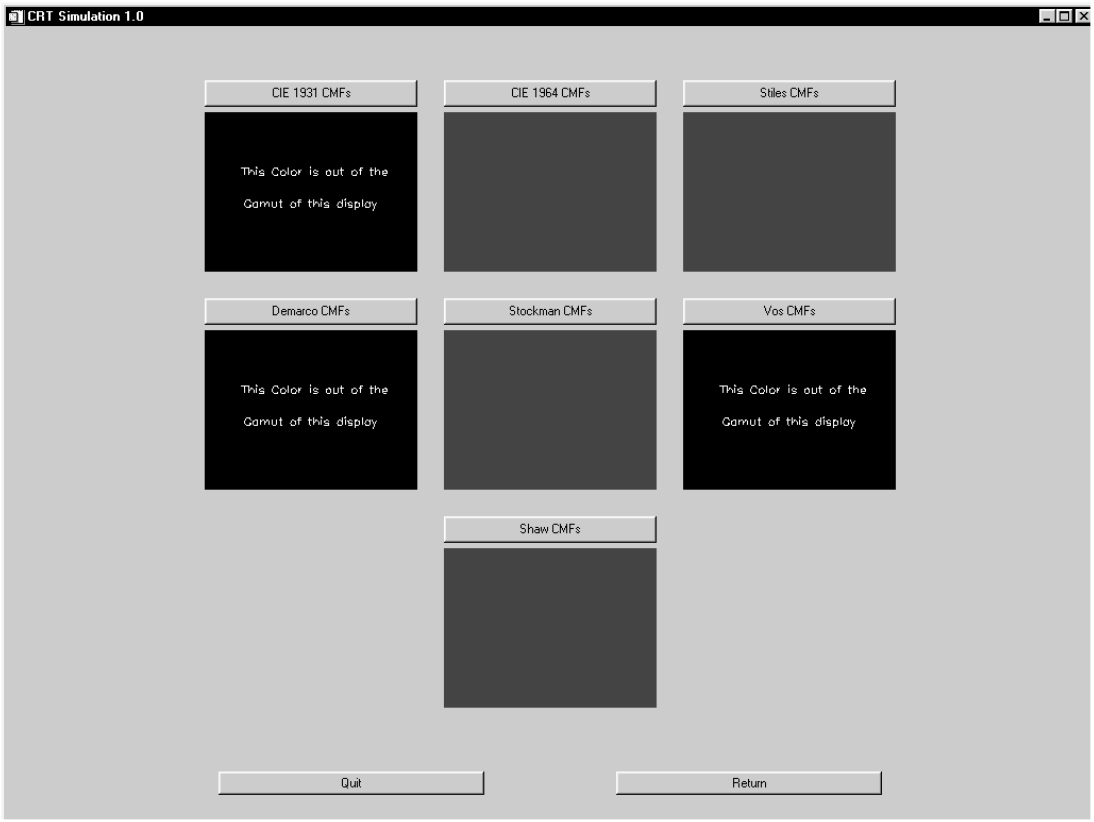


Figure 7.8 : Prediction Window - When a tristimulus value to be displayed is out of gamut of the CRT, a gamut warning is presented to the user instead.

The user can then either click *Return* to return to the main control window and select another spectral data file, or *Quit* to exit the application.

7.3 User Experiences

Although it is impossible to show the simulation in this report, one can comment on the phenomenology of the program. Sharing observations based upon users own color vision as to which set of color matching functions performs best.

Running the simulation was quite straightforward. Users picked spectral data files, both radiance and reflectance, from a previously determined set of measurements, and then displayed the calibrated color from each set of weighting functions. It is important to note that different spectra performed differently, even if they appeared similar under controlled viewing conditions, as was expected. For example, two grays of similar color appearance had a very different spectral profile. One was a Munsell Neutral gray card (N5), having a very flat spectrally non-selective profile, the second gray was imaged using a Fujix Pictrography. Both samples resulted in an $L^*=50$, and $a^*=b^*=0$ when viewed under D65 controlled viewing conditions. When displayed on the simulation the perceived relative differences were greater for one of the samples than for the other. The spectral profile being the major contributory factor.

In general the color predictions showed similar characteristics, although of different magnitudes for each stimuli. The results of the CIE 1964, and Shaw and Fairchild color matching functions tended to be similar, and the predictions of the others again were similar. For example, when looking at a gray sample, the CIE 1964, and the Shaw and Fairchild set looked slightly greenish, whereas the other predictions were all slightly reddish. Although, when looking at a pale green sample, only the CIE 1964 functions were significantly different, looking slightly darker.

8 CONCLUSIONS

A computational analysis was performed to evaluate the 1931 color matching functions against other responsivity functions using metameric data. The underlying principle is that an optimal set of responsivity functions will yield minimal tristimulus error between a pair of visually matched metamers.

In order to be able to compare the performance of the different responsivity functions a common color space was used in which the results could be compared. The common color space was CIELAB, based upon the CIE 1931 2° standard observer functions. The five other responsivity functions were transformed using a linear rotation matrix into near CIE approximations. The rotated functions were then used to calculate the pseudo tristimulus values for each set and the CIELAB coordinates.

Color differences were calculated between the pair of CIELAB coordinates using the standard ΔE^*_{ab}, and ΔE_{94} color difference formulae.

The difference of average color differences found in the six chosen sets of responsivity functions were small. The CIE 1931 2° color matching functions, on average, provided the largest color difference, 4.56 ΔE^*_{ab}. With the best performance coming from the CIE 1964 10° color matching functions, yielding an average color difference of 4.02 ΔE^*_{ab}.

An optimization was then performed on the CIE 1931 color matching functions. The concept was that color differences between metamers can be used to improve predictions of color matching functions. If one is to take all pairs, and perform an optimization that

globally minimizes the average color difference, then one can hope to obtain an optimal set of responsivity functions.

A total of thirteen optimization techniques were tested, but only two were found that were capable of both maintaining the integrity of the color matching functions, and reducing the average color difference. The optimum solution was to use a weighted combination of each of the different sets of responsivity functions chosen in chapter 2. The optimized set, the 'Shaw and Fairchild' responsivity functions, were able to reduce the average color difference down to 3.92 ΔE^*_{ab}.

An important lesson was learnt during performing the optimizations. It is possible for one to describe what one would like the optimization to do, but one must also know how to program the computer to perform this optimization and the constraints that are to be applied. For example, maintaining the integrity of the color matching functions may include constraining the area under each function, not allowing a function to go negative, minimizing noisy adjustment to the functions by ensuring smooth modifications are made. These are all concepts that are easy to grasp and contemplate, but formulating these concepts in mathematical terms is not so straightforward.

The final part of the work was to build a computer based simulation of the color differences between the different sets of responsivity functions. This simulation allows a user to load a spectral radiance, or reflectance, data file and display the tristimulus match predicted by each of the seven sets of responsivity functions.

This work provides insight into many areas of color science. One can conclude from the work that the work of the CIE has stood its ground for the last sixty years, and is still, without doubt, a standard that performs well when compared with more recent research.

It is clear that the standard is not perfect, having its areas of weakness that some are very ready to criticize.

It was found that the magnitude of observer variability was nearly eight times that of the variability found between the responsivity functions. One should be more concerned with the problems introduced by observer metamerism than the accuracy of the CIE 1931 functions themselves.

REFERENCES

Alfvin R, (1995) *'A computational analysis of observer metamerism in cross-medi color matching,'* MSc Thesis, Munsell Color Science Laboratory, Rochester Institute of Technology

Brill M.H, (1998) *'Report on Improved Colorimetry to ISCC Board of Directors – October 1998,'* ISCC News, No 376, November/December 1998

Berns R.S, Motta R.J and Gorzynski M.E. (1993) *'CRT Colorimetry. Part I: Theory and Practice,'* Col. Res. Appl., **18**, No. 5, pp. 299-325

Berns R.S, (1996) *'Methods for characterizing CRT displays,'* Displays, **16**, No. 4, pp. 173-182

CIE, (1986) *'Colorimetry,'* 2nd Ed., Central Bureau of the CIE, Vienna, CIE Publication No. 15.2

CIE, (1996) *'The relationship between digital and colorimetric data for computer-controlled CRT displays,'* Central Bureau of the CIE, Vienna, CIE Publication No. 122

Engeldrum P. G, (1993) Color scanner colorimetric design requirements. *SPIE.* **1909**, pp.75-83

Estevez O, (1979) *'On the Fundamental Data-Base of Normal and Dichromatic Color Vision,'* Thesis, University of Amsterdam, Krips Repro, Meppel

Estevez O, (1982) *'A better Colorimetric Standard Observer for Color Vision Studies: The Stiles and Burch 2° Color Matching Functions,'* Col. Res. Appl., **7**, No. 2, pp. 131-134

Fairchild M.D, (1998) *'Color Appearance Models,'* Addison Wesley, Reading, Massachusetts, USA

Guild J, (1931) *'The Colorimetric Properties of the Spectrum,'* Phil. Trans. Roy. Soc. (London), **A 230**, p. 149-187

Ives H.E, (1923) *'The transformation of color-mixture equations from one system to another,'* J. Franklin Inst., vol **180**, pp. 673-701 (1915) and vol **195**, pp. 23-44 (1923)

Judd D.B, (1951) *'Colorimetry and Artificial daylight,'* Proc. 12th session CIE, Stockholm, **I**, Tech. Committee No 7, p. 11

Kuo W.G. and Luo M.R, (1996a) *'Methods for quantifying metamerism. Part 1 – Visual assessment,'* JSDC, **112**, pp. 312-320

Kuo W.G. and Luo M.R, (1996b) *'Methods for quantifying metamerism. Part 1 – Instrumental Methods,'* JSDC, **112**, pp. 354-360

North A, (1991) *'Investigation of observer variability using a new method for determining color matching functions,'* MSc Thesis, Munsell Color Science Laboratory, Rochester Institute of Technology

Neugebauer H. E. J, (1956) Quality factor for filters whose spectral transmittances are different from color mixture curves and its application to color photography, *Jour. Opt. Soc. Amer.* **46,** No. 821, pp. 821-824

Ohta N, (1983) *'Optimization of Spectral Sensitivities,'* Photographic Science and Engineering, **27**, No. 5, pp. 193-201

Shaw M.Q and Montag E.D, (1998) *'Analyzing Observer Metamerism,'* Color Measurement Laboratory Project, Munsell Color Science Laboratory, Rochester Institute of Technology

Shaw M.Q and Montag E.D, (1999) *'Analyzing Observer Metamerism,'* Color Measurement Laboratory Project, Munsell Color Science Laboratory, Rochester Institute of Technology

Smith V.C and Pokorny J, (1975) *'Spectral sensitivity of the foveal cone photopigments between 400 and 500nm,'* Vision Research, **15**, pp. 161-171

Smith V.C, Pokorny J and Q. Zaidi, (1983) *'How do sets of color matching functions differ?'* in Colour Vision: Physiology and Psychophysics, J.D. Mollon and L.T. Sharpe, eds. Academic Press, London

Stiles W.S. and Burch J.M, (1955) *'Interim report on the Commission Internationale de l'Eclairage, Zurich, on the National Physical Laboratory's Investigation of Colour Matching,'* Optica Acta, **2**, pp. 168

Stiles W.S. and Burch J.M, (1959) *'NPL Colour Matching investigation: Final Report,'* Optica Acta, **6**, pp. 1-24

Stockman A, Macleod D.I.A and Johnson N.E, (1993) *'Spectral sensitivities of the human cones,'* Optical Society of America, **10**, No 12, pp. 2491-2516

Stockman A, Sharpe L.T, (1998) *'Human cone spectral sensitivities: a progress report'*, Vision Research, 38, 19, pp. 3193-3206

Stockman A, Sharpe L.T, (1999) *'Cone spectral sensitivities and color matching,'* Yet to be published

Thornton W.A, (1992a) *'Toward a more accurate and extensible colorimetry. Part 1 – Introduction. The visual Colorimeter-Spectroradiometer. Experimental results,'* CR&A, **17**, No.2, pp. 79-121

Thornton W.A, (1992b) *'Toward a more accurate and extensible colorimetry. Part II – Discussion,'* CR&A, **17**, No.3, pp. 162-186

Thornton W.A, (1992c) *'Toward a more accurate and extensible colorimetry. Part III - Discussion,'* CR&A, **17**, No.4, pp. 240-262

Thornton W.A, (1997) *'Toward a more accurate and extensible colorimetry. Part IV – Visual experiments with bright fields and both 10° and 1.3° field sizes,'* CR&A, **22**, No.3, pp. 189-198

Thornton W.A, Fairman H.S, (1998a) *'Toward a more accurate and extensible colorimetry. Part V – Testing visually matching pairs of lights for possible rod participation on the Angular-Stiles model,'* CR&A, **23**, No.2, pp. 92-103

Thornton W.A, (1998b) *'Toward a more accurate and extensible colorimetry. Part VI – Improved weighting functions. Preliminary results,'* CR&A, **23**, No.4, pp. 226-233

Vos J.J, (1978) *'Colorimetric and Photometric Properties of a 2° Fundamental Observer,'* Col. Res. and Appl., **3**, No. 3, pp. 125-128

Vos J.J and Walraven P.L, (1971) *'On the derivation of the foveal receptor primaries,'* Vision Research, **11**, pp. 799-818

Vos J.J, Estevez O, and Walraven P.L, (1990) *'Improved color fundamentals offer a new view on photometric additivity,'* Vision Research, **30**, No. 6, pp. 937-943

Wright W.D, (1928-29) *'A re-etermination of the trichromatic coefficients of the spectral colours,'* Trans. Opt. Soc. **30**, pp. 141-161

Wright W.D, (1981) *"50 Years of the 1931 CIE Standard Observer for Colorimetry,"* AIC Colour 81, Paper A3

Wyszecki G, (1973) *"Current developments in colorimetry,"* AIC Colour 73, pp. 21-51

Wyszecki G, and Stiles W.S, (1982) *'Color Science: Concepts and Methods, Quantative Data and formulae,'* John Wiley & Sons, New York, USA, 2nd Edn