Parametric Decomposition for Evaluating Metamerism

Zhaojian Li
Parametric Decomposition for Evaluating Metamerism

Zhaojian Li

B. S. Central University for Nationality, Beijing, China (2001)

M. S. Beijing Institute of Technology, Beijing, China (2004)

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

June 2006

Zhaojian Li

Signature of the Author

Roy S. Berns

Accepted by Academic Coordinator, Roy S. Berns
M.S. Degree Program
The M.S. Degree Thesis of Zhaojian Li has been examined and approved by two members of the color science faculty as satisfactory for the thesis requirements for the Master of Science Degree.

Roy S. Berns
Dr. Roy S. Berns, Thesis Advisor

Mark D. Fairchild
Dr. Mark D. Fairchild
Thesis Release Permission Form

Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology
Rochester, New York

Title of Thesis: Parameric Decomposition for Evaluating Metamerism

I, Zhaojian Li, hereby grant permission to the Wallace Memorial Library of Rochester Institute of Technology to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or for profit.

Zhaojian Li
Signature of the Author

Date
Parameric Decomposition for Evaluating Metamerism

Zhaojian Li

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

In those industries in which materials are colored to close specifications, a means of evaluating the degree of metamerism of colored objects is of considerable importance. Based on Wyszecki’s hypothesis and its application to quantifying metamerism as described by Fairman, parameric decomposition is a technique to adjust one spectrum of a parameric match in order to achieve a perfect (metameric) match under a specific illumination and observer condition. This method can be viewed as batch correction using three “colorants” where the color-mixing model is linear in reflectance.

The research in this thesis presented these methods using the basis functions from the CIE color-matching functions (CMFs) as well as alternative basis functions derived from dimensionality reduction techniques such as principal component analysis (PCA) and independent component analysis (ICA) for a pre-defined DuPont spectral dataset and Munsell dataset. 1,152 parameric pairs surrounding 24 color centers were synthesized using an automotive finish paint system and two-constant Kubelka-Munk turbid-media theory. Each parameric pair was corrected to a metameric pair using these various methods. The corrected spectra were compared with the formulated spectra using Kubelka-Munk theory to evaluate the parameric decomposition accuracy in terms of special and general metameric indices. The results showed that the estimated metameric indices from the CMFs-based process primaries presented relatively poor correlation to those from Kubelka-Munk theory. The process primaries from ICA for the Munsell
dataset showed almost identical performance in estimation of metameric indices to the process primaries from the PCA for Munsell dataset as well as those from ICA for the DuPont dataset. These three sets of process primaries showed slightly better performance in estimation of metameric indices than the process primaries from PCA for the DuPont dataset.
Acknowledgements

I would like to thank Dr. Roy S. Berns, my advisor, for his tireless guidance and inspiring advice – thank you Roy.

I would like to thank my parents for their patience and support. I could not come to this point without their unselfish support materially and spiritually.

I would like to thank all the people at the Munsell Color Science laboratory. Thanks to you all for the good times and suggestions. A special note of thanks to Val, Ken, Rod, and Lawrence.

This research was supported by the DuPont Graduate Fellowship in Color Science. I would like to express my appreciation to Dr. David H. Alman from DuPont for his advice, wisdom, and motivation. It has been a very valuable experience.

Thanks to you all – I could not have done this without a single one of you.
# TABLE OF CONTENTS

1. Introduction .................................................................................................................. 1

2. Background .................................................................................................................... 4
   2.1 Metamerism ................................................................................................................ 4
   2.2 Indices of Metamerism ............................................................................................. 6
      2.2.1 Special Index of metamerism, change in illuminant ........................................ 7
      2.2.2 General index of metamerism ............................................................................ 8

2.3 Paramerism ................................................................................................................... 10
   2.3.1 Linear Mixing Model .......................................................................................... 11
   2.3.2 Wyszecki’s hypothesis ....................................................................................... 12

2.4 Parameric Decomposition ............................................................................................. 13
   2.4.1 Matrix R Method ............................................................................................... 17
   2.4.2 PCA and ICA Method ....................................................................................... 18

2.5 Principal Component Analysis ..................................................................................... 18

2.6 Independent Component Analysis ............................................................................... 20
   2.6.1 ICA Using the Mean-Subtracted Data ............................................................... 21
   2.6.2 ICA Using the Original Data ............................................................................. 23

2.7 Additive Correction and Multiplicative Correction ...................................................... 24
   2.7.1 Additive Correction ........................................................................................... 24
   2.7.2 Multiplicative Correction ................................................................................... 24

2.8 Kubelka-Munk theory ................................................................................................. 25

3. Spectral Database and Dimensionality Reduction ......................................................... 28
   3.1 DuPont Spectral Database Definition .................................................................. 28
   3.2 Principal Component Analysis (PCA) ................................................................. 34
   3.3 Independent Component Analysis (ICA) ........................................................... 40
   3.4 Conclusions ............................................................................................................ 44

4. Metameric and Parameric Pairs Preparation ................................................................. 45
   4.1 Selection of the Neutral Samples ......................................................................... 45
   4.2 Generation of the Metameric Pairs ...................................................................... 47
   4.3 Generation of the Parameric Pairs ....................................................................... 53
   4.4 Conclusions ............................................................................................................ 57

5. Parameric Decomposition .............................................................................................. 58
   5.2 Fundamental Stimuli .............................................................................................. 63
   5.3 Metameric Correction ............................................................................................. 66
   5.4 The Estimation of Metameric Indices ................................................................... 72
      5.4.1 Special Index of Metamerism: Change in Illuminant .................................... 73
      5.4.2 Index for Metamerism Potential ..................................................................... 82
   5.5 Conclusions ............................................................................................................ 83

6. Summary and Conclusions ........................................................................................... 86
   6.1 Summary .................................................................................................................. 86
LIST OF FIGURES

Figure 2-1. Example of metamerism ......................................................... 5
Figure 2-2. Spectral power distribution of some typical illuminants .................. 8
Figure 3-1. The chroma distribution, $C_{ab}$, of the defined DuPont 44,100 samples .... 31
Figure 3-2. The hue distribution, $h_{ab}$, of the defined DuPont 44,100 samples .................. 31
Figure 3-3. The lightness distribution, $L^*$, of the defined DuPont 44,100 samples ...... 32
Figure 3-4. Pigment color gamut of the defined DuPont database in CIELAB $a^*-b^*$ plane. .................................................................................................................. 32
Figure 3-5. Pigment color gamut of the defined DuPont database in CIELAB $L^*-a^*$ plane .................................................................................................................. 33
Figure 3-6. Pigment color gamut of the defined DuPont database in CIELAB $L^*-b^*$ plane .................................................................................................................. 33
Figure 3-7. The first six eigenvectors and the mean derived from the reflectance spectra of the defined DuPont spectral data set ........................................................................... 35
Figure 3-8. An example of the spectral fits using the three approaches of PCA .......... 38
Figure 3-9. The three basis functions of the mean-subtracted DuPont dataset using ICA. .................................................................................................................. 40
Figure 3-10. The three basis functions of the original DuPont dataset using ICA ....... 40
Figure 3-11. An example of the spectral fits using the two approaches of ICA ........... 43
Figure 4-1. The chroma distribution of the selected neutral samples ..................... 46
Figure 4-2. The hue distribution of the selected neutral samples ......................... 46
Figure 4-3. The lightness distribution of the selected neutral samples .................. 47
Figure 4-4. Colorimetric plots of the selected 24 color centers ............................. 49
Figure 4-5. The spectral reflectance of the twenty-four metameric pairs. The green lines represent the standard and the red lines represent the metameric batch. Note that each plot has a different range of the reflectance factor ......................................... 52
Figure 4-6. A set of parameters in CIELAB space. The red point represents the standard and the green points represent the 48 parametric batches .................................................................... 54
Figure 4-7. The reflectance curves of a set of parameters and the metameric pair ...... 54
Figure 5-1. Seven sets of process primaries to perform parametric decomposition .... 55
Figure 5-2. Seven generated projection operators from seven sets of process primaries 61
Figure 5-3. The diagonal of Matrix $S$ ................................................................... 62
Figure 5-4. The fundamental stimuli from seven sets of process primaries and the metamers of the 24 color centers ...................................................................... 64
Figure 5-5. The fundamental stimuli and the metameric pair of the 9th color center .... 65
Figure 5-6. An example of the corrected spectra using five sets of process primaries, the metameric spectrum using Kubelka-Munk theory, the parametric batch, and the standard .................................................................................................................. 67
Figure 5-7. Difference plot of corrected spectra using matrix R method from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters

Figure 5-8. Difference plot of corrected spectra using DuPont PCA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters

Figure 5-9. Difference plot of corrected spectra using DuPont ICA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters

Figure 5-10. Difference plot of corrected spectra using Munsell PCA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters

Figure 5-11. Difference plot of corrected spectra using Munsell ICA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters

Figure 5-12. The mean spectral RMS error between the corrected spectra from five sets of process primaries and the formulated metameric spectra from Kubelka-Munk theory for the 24 color centers

Figure 5-13. Testing MI residual data against a normal distribution

Figure 5-14. Correlation of MI (D65->A) between Kubelka-Munk theory and seven metameric correction methods

Figure 5-16. Correlation of MI (D65->F11) between Kubelka-Munk theory and seven metameric correction methods

Figure A. 1. Scatter plot of a two-dimensional sample set

Figure A. 2. The procedure of PCA to a two-dimensional sample set

Figure A. 3. The procedure of ICA to a two-dimensional sample set

Figure A. 4. Color gamut of the Munsell database in CIELAB a*-b* plane

Figure A. 5. Color gamut of the Munsell database in CIELAB L*-a* plane

Figure A. 6. Color gamut of the Munsell database in CIELAB L*-b* plane

Figure A. 7. The three basis functions of the Munsell spectral dataset from PCA and the mean spectrum of the Munsell spectral dataset

Figure A. 8. The three basis functions of the Munsell spectral dataset from ICA using the mean-subtracted data

Figure A. 9. The three basis functions of the Munsell spectral dataset from ICA using the original data

Figure A. 10. An example of the spectral reconstruction using three basis functions from PCA and ICA for Munsell dataset
LIST OF TABLES

Table 3-1. DuPont pigments used in pigment mixing .................................................. 29
Table 4-1. The recipes, and the color difference of the selected 24 metameric pairs ...... 50
Table 4-2 The eight parameric batches along eight directions with component difference of 1.5 from the standard under illuminant D65 ................................................... 53
Table 4-3. The recipes and color attributes of the paramers and the metamer for the 9th color center ............................................................... 55
Table 5-1. The spectral RMS error between the corrected spectrum using five sets of process primaries and that formulated using Kubelka-Munk theory for the pooled data of all color centers ........................................... 72
Table 5-2 The recipes and the color difference of a representative parameric pair and the corresponding formulated metameric pair using Kubelka-Munk theory ........... 73
Table 5-3. The estimated special index of metamerism of a representative corrected parameric pair using seven correction techniques and the formulated metameric pair using Kubelka-Munk theory for three test illuminants ........................................ 74
Table 5-4. Significance of the difference between the MI residuals from various metameric correction methods for test illuminant A ......................................................... 76
Table 5-5. Significance of the difference between the MI residuals from various metameric correction methods for test illuminant F2 ................................................. 77
Table 5-6. Significance of the difference between the MI residuals from various metameric correction methods for test illuminant F11 ................................................. 77
Table 5-7. Significance of the difference between the WRMS residuals using the five sets of process primaries .............................................................. 83
Table A. 1. The average colorimetric and spectral accuracy of the spectral reconstruction for the Munsell dataset with different number of basis functions from PCA and ICA. .................................................................................................................. 99
Table B. 1. The derived first three eigenvectors and three independent components for Munsell data ........................................................................................................ 102
Table B. 2. The derived first three eigenvectors and three independent components for DuPont data ........................................................................................................ 103
1. Introduction

Metamerism occurs when two stimuli match under one viewing condition but mismatch under other conditions. It is a problem for all color reproduction processes. So a means of describing the magnitude of metamerism for a given pair of stimuli is of considerable practical importance, particularly to those industries where materials are colored to close specifications, such as the textile and paint industries (Wyszecki 1982). For this, indices of metamerism are used. The CIE has recommended two ways to calculate indices of metamerism: special index of metamerism and general index of metamerism. According to the CIE, in order to calculate an index of metamerism for a sample pair, they (metameric pair) must have an exact tristimulus match in a reference viewing condition (CIE 1986). However, for most real samples, it is rare to achieve an exact tristimulus match. That is, there is a residual color difference for the pair (parameric pair) under the reference condition. In order to calculate the meaningful index of metamerism, it is imperative to eliminate this residual color difference. In 1987, Fairman reported a method to decompose the spectra of the parameric pair into the spectral fundamental and metameric black stimuli based on Cohen and Kappauf’s method (Cohen 1982) of spectral decomposition. This method was called parameric decomposition (Fairman 1987). In this method, color-matching functions are used as the process primaries to adjust a spectrum to achieve a specific set of tristimulus values. However, they are not the colorant primaries with which the real samples can be moved in color space (Fairman 1991). That
is, they are not the intrinsic primaries of a given real sample. So this kind of metameric correction can only approximate the correction that would be made by real colorants (Fairman 1991). The primary objective of this research was to find alternative process primaries to better approximate the correction that would be made using real colorants.

In order to derive the possible colorant primaries, the first mission of this research was to define a spectral database containing the spectral properties of an automotive paint system supplied by DuPont. Two spectral reflectance dimensionality reduction techniques, principal component analysis (PCA) and independent component analysis (ICA), were performed to get the characteristic spectra of this database as well as the Munsell dataset. The process primaries could be these characteristic spectra obtained from linear modeling of the reflectance spectra, thereby generating a set of basis spectra.

Based on this database, a number of metameric and parameric pairs surrounding several selected color centers were synthesized using the paint system and two-constant Kubelka-Munk turbid-media theory. The generated parameric pairs were corrected to metameric pairs using the parameric decomposition method with the derived process primaries from PCA and ICA as well as the process primaries based on color-matching functions. The question of how does the selection of the process primaries for fundamental stimulus approximation influence the estimation of parameric decomposition was explored. The formulated metameric spectra using Kubelka-Munk turbid-media theory were used as the reference to evaluate the corrected spectra from various process primaries.
The final aim of this research was to evaluate the metamerism of the corrected parameric pair from various process primaries. The influence estimation of special index of metamerism, change in illuminant, and general index of metamerism were explored in this research.
2. Background

In this chapter, the background of evaluating metamerism of a parameric pair will be reviewed. It involves metamerism, paramerism, indices of metamerism, parameric decomposition, and Kubelka-Munk turbid media theory. Two dimensionality reduction techniques, PCA and ICA, will also be described in this chapter.

2.1 Metamerism

If two samples have the same reflectance curve, they will match under all conditions of illumination and observation. The opposite case, where two samples with different reflectance curves look different under all conditions, is not valid. That is, two samples with different reflectance curves can match under specified illumination and observation. This phenomenon can occur because our color vision is the result of integration of light source, the object, and the observer, resulting in what is called trichromacy (Berns 2000). These two samples are called metamers (a chemical term), metameric pair, or metameric objects, and the underlying concept is referred to as metamerism (Berns 2000, Wyszecki 1982). There are many definitions of metamerism. According to ASTM E 284, a metameric pair are two specimens that match under a specified illumination and to a specified observer and whose spectral reflectance or transmittances differ in visible wavelengths (ASTM E 284). Mathematically (Berns 2000):
\[
\begin{align*}
\Phi_{\lambda,1} & = k \int \Phi_{\lambda,1} \, d\lambda \\
\Phi_{\lambda,2} & = k \int \Phi_{\lambda,2} \, d\lambda \\
\Phi_{\lambda,2} & = k \int \Phi_{\lambda,2} \, d\lambda \\
\Phi_{\lambda,1} & = \Phi_{\lambda,2}
\end{align*}
\]
\[\text{(2.1)}\]

where \(\Phi_{\lambda}\) can be a light source, \(L_{\lambda}\), or the product of an illuminant's relative spectral power distribution, \(S_{\lambda}\), and a sample's reflectance factor, \(R_{\lambda}\).

An example of a metameric pair is seen in Figure 2-1 below.

![Graph](image)

**Figure 2-1. Example of metamerism.**

In Figure 2-1, the standard and the trial which have the different spectral reflectance match for illuminant D65 and the CIE 1931 standard observer. In this case, \(\Phi_{\lambda,1} \neq \Phi_{\lambda,2}\) in Equation (2.1) is replaced with \(R_{\lambda,1} \neq R_{\lambda,2}\).
For a metameric pair, their match is a conditional match. This match may fail if the light source or the observer changes because the fact that the color match depends on the interactions among the light source, the object, and the observer (Berns 2000). So there are two types of memamerism, illuminant metamerism and observer metamerism, due to the change in light source and the observer, respectively. In this research, only illuminant metamerism will be explored.

Illuminant metamerism occurs when two objects match under one illuminant, but do not match under one or more other illuminants (Fairchild 2005). This happens when the spectral reflectance functions of the two objects differ, but those differences can be compensated by the integration with the spectral power distribution of the first illuminant. However this compensation fails when the illuminant is changed. Illuminant metamerism is often a serious problem in industries that produce colored materials. A customer expects all the parts of a car that are the same color to match whether it is a sunny or cloudy day; or in the morning or mid-day.

2.2 Indices of Metamerism

Since metamerism is a common and important phenomenon, a means of describing the magnitude of metamerism for a given pair of samples is highly desirable. For this, *indices of metamerism* are used. That is, an index of metamerism is a single-number index, which indicates how well two objects that match under one illuminant-observer condition will match under another illuminant-observer condition. The indices of metamerism are
classified into two categories: special index of metamerism and general index of metamerism. For the special index of metamerism, there are two types of index, change in illuminant and change in observer. In this research, the special index of metamerism, change in illuminant, will be explored as well as the general index of metamerism.

### 2.2.1 Special Index of metamerism, change in illuminant

CIE 15.2 (CIE 1986) recommends a technique to calculate an index of metamerism for the change in illuminant with the same observer. For two specimens whose tristimulus values are identical with respect to a reference illuminant and reference observer, the index of metamerism, $MI$, is set equal to the color difference between the two specimens under the test illuminant $t$. One of the currently recommended color-difference formulae should be used for the computation when appropriate. The recently developed color difference equation CIEDE2000 is recommended to use. However, in this research, CIE94 was used since it is current DuPont practice. The preferred reference illuminant is CIE standard illuminant D65 called the primary illuminant. The preferred test illuminant is CIE standard illuminant A or one of the fluorescent illuminants. The most appropriate choice of test illuminant depends on application, and in some instances, it may be useful to determine the metamerism index with respect to several test illuminants (Berns 2000, Wyszecki 1982). In this research, illuminants A, F2, and F11, representing incandescent, cool-white fluorescent, and narrow-band fluorescent, respectively, were used as test illuminants. These three illuminants are called the secondary, tertiary, quaternary
illuminant, respectively (Berns 2000). The special index of metamerism for a specific change in illuminant can be denoted as \( MI(r \rightarrow t) \), for example, \( MI(D65 \rightarrow A) \) or \( MI(A \rightarrow F2) \). The spectral curves of these illuminants are plotted in Figure 2-2.

![Spectral power distribution of some typical illuminants](image)

**Figure 2-2.** Spectral power distribution of some typical illuminants.

### 2.2.2 General index of metamerism

In order to specify the degree of metamerism that is illuminant independent, a general index of metamerism has been recommended. It is based on spectral difference between the metamic pair (CIE 1986). The simplest form of this index was proposed by Bridgeman (Bridgeman 1969). Essentially, it is the root mean square (RMS) error between the reflectance factors of the metamic pair. But this calculation does not consider human cones sensitivities. So Nimeroff and Yurow (Nimeroff 1965) proposed
the weighted RMS error between the metameric pair using color-matching functions. However, because color-matching functions are not physiological, this weighted index does not correlate with visual evaluations (Berns 2000). Viggiano (Viggiano 2001) proposed the perception-reference method comparing radiance ratio spectra. It is a refinement of a spectral-based metameric index based on a weighted sum of the absolute differences between two spectra proposed by Nimeroff and Yurow. Another weighting of the RMS error is the diagonal of Cohen’s matrix R (Imai 2002). The matrix R method will be described in detail later in this chapter.

However, the term general index of metamerism is a misnomer since it provides no information at all about what happens under any conditions in which the sample pair does not match (Fairman 1991). Fairman proposed calling such an index, an index for metameric potential, which is solely derived from spectral information. In this research, the index for metamerism potential is expressed by the weighted spectral RMS error using the diagonal of matrix R:

\[
WRMS = \sqrt{\frac{\sum_{\lambda=1}^{n}(w_{\lambda} \Delta \rho_{\lambda})^2}{n}}
\]  

(2.2)

where \( n \) is the number of the wavelengths, \( \Delta \rho_{\lambda} \) is the spectral difference between the metameric pair, and \( w_{\lambda} \) is the weighting function from the diagonal of matrix R for the reference viewing condition.
2.3 Paramerism

In order to calculate an index of metamerism for a sample pair, it is necessary that the pair exactly match in a reference viewing condition. However, in real situations small color-differences exist between the pair. Strictly speaking, such a sample pair should not be classified as metamers according to the definition, as this case does not meet the assumption for the calculation of an index of metamerism. Many new terminologies for such a pair of samples were discussed (Rodrigues 1980, Billmeyer 1983, Kuehni 1983, Robertson 1983, and Fairman 1986). Mostly, the terms paramers, parameric, and paramerism are used in place of metamers, metameric, and metamerism. According to ASTM E 284, paramers are specimens having different spectrophotometric curves that produce approximately the same color sensation under the same illuminating and viewing conditions (ASTM E 284).

At present, there is no specific recommendation from the CIE on how to calculate the index of metamerism for paramers. For a parameric pair, the color difference consists of a simple color difference and a metameric color difference. So some techniques such as additive correction (AC) (Luo 1996) and multiplicative correction (MC) (Brockes 1970) methods were proposed to eliminate the simple color difference to calculate the real index of metamerism. Berns and Billmeyer (Berns 1983) proposed a method to calculate indices of metamerism with constant chromatic adaptation. Among these correction methods, parameric decomposition proposed by Fairman (Fairman 1987) is a good way to correct the spectral reflectance curve of one of the samples so that an exact
tristimulus match is achieved for the reference condition. Parameric decomposition is based on linear combination model of the stimulus (Berns 1994) and Wyszecki’s hypothesis (Wyszecki 1953).

2.3.1 Linear Mixing Model

In the classical color-matching experiment, a stimulus can be obtained by the linear combination of three process primaries. These three process primaries are also called three basis functions. This is similar to CRT displays. That is:

\[ N(\lambda) = c_1 p_1(\lambda) + c_2 p_2(\lambda) + c_3 p_3(\lambda) , \]

where \( N(\lambda) \) represents the spectral reflectance of the stimulus,

\[ \mathbf{P} = [ p1 \ p2 \ p3 ] = \begin{bmatrix} p_{1,1} & p_{2,1} & p_{3,1} \\ p_{1,2} & p_{2,2} & p_{3,2} \\ \vdots & \vdots & \vdots \\ p_{1,j} & p_{2,j} & p_{3,j} \end{bmatrix} \]

is a \( j \times 3 \) matrix representing the three process primaries, and \( \mathbf{C} = [ c1 \ c2 \ c3 ]' \) is the vector of scalars to the three corresponding process primaries and the prime symbol is the matrix-vector transpose operation.

Let a \( j \times 3 \) matrix \( \mathbf{A} \) be defined whose 3 columns are a set of illuminant-observer combination, \( \mathbf{A} = \mathbf{S}\mathbf{M} \), where \( j \) is the number of wavelengths, \( \mathbf{S} \) is a matrix with spectral radiant power distribution of the reference illuminant in its diagonal, and
M = \begin{bmatrix}
  \bar{x}_{i1} & \bar{y}_{i1} & \bar{z}_{i1} \\
  \bar{x}_{i2} & \bar{y}_{i2} & \bar{z}_{i2} \\
  \vdots & \vdots & \vdots \\
  \bar{x}_{im} & \bar{y}_{im} & \bar{z}_{im} 
\end{bmatrix}

is a set of color-matching functions.

The tristimulus values \( T \) of a stimulus under the reference condition are calculated as

\[ T = A'N = A'PC. \tag{2.4} \]

The scalars can be obtained by inverting the Equation (2.4):

\[ C = (A'P)^{-1}T. \tag{2.5} \]

From the scalars, the spectral reflectance of the stimulus is generated:

\[ N = P(A'P)^{-1}T. \tag{2.6} \]

**2.3.2 Wyszecki’s hypothesis**

According to the Wyszecki’s hypothesis, any stimulus is composed of two components---a fundamental stimulus (with tristimulus values equal to the stimulus) and a metameric black (with tristimulus values equal to zero). That is, the fundamental stimulus carries all information essential to human vision and the metameric black has a null stimulus not perceived to human vision, but always evokes black. By Wyszecki’s hypothesis, the spectral reflectance of a mixture can be expressed in matrix notation:

\[ N = N^* + B, \tag{2.7} \]
where \( N^* \), a \( j \times 1 \) column vector, represents the fundamental stimulus and \( B \), a \( j \times 1 \) column vector, is the metamer black.

Since the metamer black has tristimulus values of zero, it has the important feature of having both positive and negative values, and thus, it cannot be realized physically. However, it can be used very effectively for generating a set of metamers mathematically (Wyszecki 1982).

### 2.4 Parameric Decomposition

Given the spectral reflectance of a parameric pair, the standard \( N_{std} \) and the sample \( N_{spl} \), their tristimulus values under the reference condition are calculated by the following:

\[
T_{std} = A^T N_{std},
\]

\[
T_{spl} = A^T N_{spl}.
\]

Since the parameric pair does not exactly match under the reference condition, there is a difference in tristimulus values \( \Delta T \) between the standard and the sample. That is,

\[
\Delta T = T_{std} - T_{spl} = A^T (N_{std} - N_{spl}).
\]

Let \( \Delta N \) be the incremental correction to \( N_{spl} \) needed for a tristimulus match to \( N_{std} \). Then

\[
N_{corr} = N_{spl} + \Delta N,
\]
where \( N_{corr} \) represents the spectral reflectance of the corrected sample which is a tristimulus match to \( N_{std} \).

From Equation (2.6), we have

\[
\Delta N = P(A'P)^{-1} \Delta T .
\]  
(2.12)

Finally, combining Equation (2.11) and (2.12) the spectral reflectance of the corrected sample is generated by

\[
N_{corr} = N_{spl} + P(A'P)^{-1} A' (N_{std} - N_{spl}) .
\]  
(2.13)

Setting

\[
S = P(A'P)^{-1} A' ,
\]  
(2.14)

and rewriting Equation (2.13), it becomes

\[
N_{corr} = SN_{std} + (I - S)N_{spl} ,
\]  
(2.15)

where \( I \) is the identity matrix.

Now the tristimulus values of \( N_{corr} \) as well as the two components \( SN_{std} \) and \( (I - S)N_{spl} \), can be calculated

\[
A'SN_{std} = A'P(A'P)^{-1} A'N_{std} = A'N_{std} = T_{std} ,
\]  
(2.16)

\[
A'(I - S)N_{spl} = A'IN_{spl} - A'P(A'P)^{-1} A'N_{spl} = A'N_{spl} - A'N_{spl} = 0 ,
\]  
(2.17)

\[
A'N_{corr} = A'SN_{std} + A'(I - S)N_{spl} = T_{std} .
\]  
(2.18)
It can be seen that the corrected sample has the same tristimulus values as those of the standard as well as the component $SN_{std}$ while the component $(I - S)N_{spl}$ has tristimulus values of zero. Therefore, according to Wyszecki's hypothesis, the component $SN_{std}$ is just the fundamental stimulus of the standard, $N^*$, and the component $(I - S)N_{spl}$ is the metameric black, $B$, of the sample. That is, the spectral reflectance of the corrected sample is obtained by adding the fundamental stimulus of the standard and the metameric black of the sample. Mathematically,

$$\begin{align*}
N^*_{std} &= SN_{std}, \quad (2.19) \\
B_{spl} &= N_{spl} - N^*_{spl} = (I - S)N_{spl}, \quad (2.20) \\
N_{corr} &= N^*_{std} + B_{spl}, \quad (2.21)
\end{align*}$$

where $N^*_{std}$ and $N^*_{spl}$ is the fundamental stimulus of the standard and the sample, respectively, $B_{spl}$ represents the metameric black of the sample.

From Equation (2.19), the fundamental stimulus is the linear combination of the process primaries that match the standard in terms of the tristimulus values under the reference condition. Matrix $S$ is viewed as the projection operator performing the mapping from one spectrum to another (Fairman 1987).

Combining Equation (2.14) and (2.19), the fundamental stimulus of the standard can be rewritten as
\[ \mathbf{N}_{std}^* = \mathbf{P} (\mathbf{A}' \mathbf{P})^{-1} \mathbf{T}_{std} . \] (2.22)

From Equation (2.22), \( \mathbf{N}_{std}^* \) is dependent only on the basis functions \( \mathbf{P} \) and the tristimulus values \( \mathbf{T}_{std} \) of the standard under the reference condition \( \mathbf{A} \). Therefore, this component in the corrected sample ensures the tristimulus match to the standard. From the derivation of Equation (2.18), the metameric black of the sample ensures the spectral reflectance of the corrected sample has the similar shape with that of the original sample.

An important property of matrix \( \mathbf{S} \) is idempotency, that is,
\[ \mathbf{SS} = \mathbf{S} . \] (2.23)
Based on this property, it can be proven that
\[ \mathbf{SN}^* = \mathbf{SSN} = \mathbf{SN} = \mathbf{N}^* . \] (2.24)
So further reduction by \( \mathbf{S} \) gets the same unique fundamental stimulus.

In the field of colorant formulation and shading, the spectral reflectance of a mixture can be obtained by the linear combination of three “statistical colorants”, a set of colorants where the color-mixing model is linear in reflectance (Berns 1994). So parameric decomposition can also be viewed as batch correction using three “statistical colorants” to adjust the spectrum of the batch (sample) to achieve tristimulus equality. In this case, these three “colorants” correspond to the basis functions or the process primaries. There can be different choices of the process primaries derived from different sources such as a three-primary additive color system, a sampling of many pigment colors and any ensemble of objects such as Munsell color chips. If the process primaries
are the color matching functions multiplied by the reference illuminant spectral power distribution, \( P = A \), the approach of parameric decomposition known as the Matrix R method was developed by Cohen and Kappauf (Cohen 1982, 1985) and used by Fairman (Fariman 1987) for correcting parameric pairs. In this research, alternative process primaries were derived from dimensionality reductions techniques such as principal component analysis (PCA) and independent component analysis (ICA) for a pre-defined spectral dataset.

### 2.4.1 Matrix R Method

If \( P \) in Equation (2.14) is instead replaced by \( A \), the projection matrix \( S \) becomes the matrix \( R \) as shown in the equation below:

\[
R = A(A^T A)^{-1} A^T.
\]  

(2.25)

Then the fundamental stimulus and the metameric black of a stimulus are calculated by:

\[
N^* = RN,
\]  

(2.26)

\[
B = N - N^* = (I - R)N.
\]  

(2.27)

In addition to idempotency, matrix \( R \) is symmetrical. Each row (or column) is the scalar fundamental of a monochromatic, spectral color stimulus (Cohen 1985). That is, if \( N \) is a narrow-band light that has unit amplitude in \( j \)th wavelength band, and zero at others, its fundamental metamer is the \( j \)th row (or column) of matrix \( R \).

Another important property of matrix \( R \) is invariance under linear transformation. That is, matrix \( R \) is independent of the arbitrary primaries selected for
the color-matching experiment. More properties of matrix $R$ can be found in Cohen’s book (Cohen 2002).

Since color-matching functions weighted by the spectral power distribution of the illuminant, $A$, are used as the process primaries, matrix $R$ method is noted as CMFs-based method to perform parameric decomposition in this research.

2.4.2 PCA and ICA Method

Instead of direct color matching functions multiplied by the reference illuminant are used, these two methods are to find alternative process primaries using principal component analysis and independent component analysis for a pre-defined spectral dataset.

2.5 Principal Component Analysis

Principal component analysis (PCA) is concerned with explaining the variance-covariance structure through a few linear combinations of the original variables (Johnson 2002). Its goal is to find an uncorrelated representation of a set of correlated $n$-dimensional vectors. In color technology, PCA is used extensively for data reduction. Hundreds or thousands of reflectance spectra are described by a considerably smaller set of eigenvectors, termed “statistical colorants” by Tzeng and Berns (Tzeng 2005). Data reduction is accomplished by neglecting the unimportant directions in which a sample set’s variances are insignificant. Because the dominant sample variations are along several significant directions, the number of these directions approximates the dimensionality of the sample set. Given the spectral reflectance of one sample set, $X$, a
\( n \times q \) matrix with \( n \) spectral bands (wavelength) and \( q \) number of samples. \( \mathbf{X} \) is denoted as the collection of column vectors of \((x_1, x_2, x_3, \ldots x_n)\), and \( T \) is the matrix-vector transpose operation in the next calculation. The steps to perform PCA data reduction for this sample set are shown as follows:

1. Calculate the mean spectrum, \( \overline{\mathbf{X}} \), and variance-covariance matrix \( \mathbf{C} \), a \((n \times n)\) matrix, as follows:

\[
\overline{\mathbf{X}} = \frac{1}{q} \sum_{i=1}^{q} (x_1, x_2, x_3, \ldots x_n),
\]

\( (2.28) \)

\[
\mathbf{C} = \frac{1}{q} \sum_{i=1}^{q} (\mathbf{X}_i - \overline{\mathbf{X}})(\mathbf{X}_i - \overline{\mathbf{X}})^T.
\]

\( (2.29) \)

2. Calculate the total \( n \) eigenvalues \((\lambda_1, \lambda_2, \ldots, \lambda_n)\) and corresponding \( n \) eigenvectors \((\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)\) of \( \mathbf{C} \), as follows:

\[
|\mathbf{C} - \lambda_i \mathbf{I}| = 0,
\]

\( (2.30) \)

\[
\mathbf{C} \mathbf{e}_i = \lambda_i \mathbf{e}_i,
\]

\( (2.31) \)

where \( i = 1, 2, \ldots, n \), \( \mathbf{I} \) is the \((n \times n)\) identity matrix. The first eigenvector \( \mathbf{e}_1 \), having the largest associated eigenvalue \( \lambda_1 \), gives the direction where the variance of the original data is largest. The second eigenvector \( \mathbf{e}_2 \), having the second largest associated eigenvalue \( \lambda_2 \), gives the direction of maximum variance, subject to being uncorrelated to \( \mathbf{e}_1 \), and so forth. The first \( m \) \((m < n)\) eigenvectors form an orthogonal \((n \times m)\) matrix \( \mathbf{V} \) defining a \( m \)-dimensional subspace in the original \( n \)-dimensional space. Then these \( m \) eigenvectors are referred as the basis functions in dimensionality reduction. That is, the
first function accounts for as much of the variance as possible in the original data, and the second function accounts for the second largest portion of the variance in the original data.

3. Calculate the first $m$ ($m < n$) principal components coordinates discarding and using the mean spectrum $\bar{X}$, as the following:

\[
Y = V^T X, \tag{2.32}
\]

\[
Y^* = V^T (X - \bar{X}), \tag{2.33}
\]

where $Y = (Y_1, Y_2, ..., Y_m)^T$ are the first $m$ principal components coordinates of the uncentered data $X$; $Y^*$ are the first $m$ principal components coordinates of the mean-centered data $X - \bar{X}$. $V$ contains the first $m$ eigenvectors. It can be seen that a principal component is the projection (linear combination) of the original data on the direction represented by the associated eigenvector. The first principal component is the linear combination with maximum variance (Johnson 2002).

4. Reconstruct the spectral reflectance $\hat{X}$ and $\hat{X}^*$, discarding and using the mean spectrum, respectively, as the following:

\[
\hat{X} = VY = VV^T X, \tag{2.34}
\]

\[
\hat{X}^* = \bar{X} + VV^T (X - \bar{X}). \tag{2.35}
\]

**2.6 Independent Component Analysis**

Independent component analysis (ICA) is another technique for dimensionality reduction. Its goal is to produce basis functions that give rise to maximum statistical independence.
of the data (Hyvarinen 2000). That is, it seeks directions in feature such that the resulting signals show independence. When applying dimensionality reduction of spectral data (Hyvarinen 2001), the model of ICA is given by

\[ X = AS, \tag{2.36} \]

where \( X \) is a \( q \times n \) matrix with \( n \) spectral bands (wavelength) and \( q \) is the number of samples representing the observed mixtures (a linear combination of the original source signals and a mixing matrix). \( S \) is a \( m \times n \) matrix with \( m \) independent components (ICs) representing the original source signals. For the intent of the dimensionality reduction, \( m \) is less than \( q \). \( A \) is a \( q \times m \) scalar matrix of mixing coefficients to construct \( X \) from the various independent components. There are two approaches to perform ICA: ICA using the mean-subtracted data and ICA using the original data.

### 2.6.1 ICA Using the Mean-Subtracted Data

For the ICA approach using the mean-subtracted data, before attempting to estimate \( A \) and \( S \), both “centering” and “whitening” preprocessing are performed (Hyvarinen 2001). First, the obtained data \( X \) are centered by subtracting their mean value \( \bar{X} \) to make \( X \) zero-centered, noted as \( \bar{X} \) in this research.

\[ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} (x_1, x_2, x_3, \ldots x_q), \tag{2.37} \]

\[ \bar{X} = X - \bar{X}, \tag{2.38} \]
where $\bar{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are $1 \times q$ vectors. It should be noted that in this case the mean vector is not the mean spectrum as in PCA since $q$ represents the number of the samples in the dataset. That is, it does not have a physical interpretation in terms of spectral reflectance. Then the centered data are whitened, which means they are linearly transformed so that the components are uncorrelated and have unit variance. “Whitening” can be performed via eigenvalue decomposition of the covariance matrix $E\{\mathbf{XX}^T\} = \mathbf{UDU}^T$. $\mathbf{U}$ is here the orthogonal matrix of eigenvectors of $E\{\mathbf{XX}^T\}$ and $\mathbf{D}$ is the diagonal matrix of its eigenvalues. Whitening can now be done by

$$\tilde{\mathbf{X}} = \mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{X},$$

(2.39)

where $\mathbf{D}^{-1/2}$ refers to taking each element in $\mathbf{D}$ and raising it to the $-1/2$ power.

In addition to simplifying the ICA algorithm, whitening can also perform the dimensionality reduction at the same time (Laamanen 2001). When calculating the whitening matrix $\mathbf{D}^{-1/2}\mathbf{U}^T$, the first $m$ largest eigenvalues and corresponding eigenvector of $E\{\mathbf{XX}^T\}$ are selected and the rest are discarded, as in often done in PCA. That is, $\mathbf{D} = \text{diag}(d_1, d_2, \ldots d_m)$ and $\mathbf{U}$ now is the orthogonal matrix with $m$ eigenvectors. So far, we can see that PCA is the preprocessing of ICA. In this case, $\tilde{\mathbf{X}}$ is the centered and whitened data, which is a $m \times n$ matrix. The number of dimensions is reduced from $q$ to $m$ rather than from $n$ to $m$ in PCA.

After the preprocessing of the data, the ICA algorithm is performed. There are many different algorithms such as CoBliss, FastICA, and JADE. The JADE algorithm
Joint Approximate Diagonalization of Eigenvalues was used in this research (Cardoso 1998). The direct output of the JADE algorithm is the separation matrix $B$, which is the inverse of mixing matrix $A$. The independent components $m \times n$ matrix $S$ is estimated by

$$S = B\bar{X}.$$  \hspace{1cm} (2.40)

In this approach, the independent components are used as the $m$ basis functions to form the new coordinate system. The original data could be reconstructed by

$$\hat{X} = S^T S\bar{X} + \bar{X}.$$  \hspace{1cm} (2.41)

2.6.2 ICA Using the Original Data

Considering the fact that the mean vector has no physical interpretation in terms of spectral reflectance, the original data without a mean vector offset can be used to perform ICA in this new approach. The "whitening" preprocessing is also required to perform the dimensionality reduction, as shown in the following:

$$\bar{X} = D^{-1/2}U^T X,$$  \hspace{1cm} (2.42)

where $X$ is the original data, $D = \text{diag}(d_1,d_2,...d_m)$ and $U$ is the orthogonal matrix with $m$ eigenvectors coming from $E\{XX^T\}$.

Applying the same JADE algorithm as the approach of ICA with the mean, the independent components $m \times n$ matrix $S$ is estimated by

$$S = BX.$$  \hspace{1cm} (2.43)

The spectral reconstruction can be performed by
\[ \hat{X} = S^T S X. \]  

(2.44)

### 2.7 Additive Correction and Multiplicative Correction

In addition to the parameric decomposition method, the additive and multiplicative correction methods are two other metameric correction methods to eliminate the residual color difference between a parameric pair under the reference illuminant required in order to calculate the special index of metamerism.

#### 2.7.1 Additive Correction

Additive Correction (AC) is a metameric correction wherein the directed difference in each colorimetric axis between the standard and the batch in the reference illuminant are added to the difference between the standard and the batch in the test illuminant (Fairman 1991). The resulting CIELAB lightness difference \( \Delta L' \) under the test illuminant is

\[ \Delta L' = L'_{b,t} - L'_{s,r} - (L'_{b,r} - L'_{s,r}), \]  

(2.45)

where the subscript \( b \) and \( s \) represent the batch and the standard, respectively; the subscript \( t \) and \( r \) represent the test illuminant and the reference illuminant, respectively. \( \Delta a' \) and \( \Delta b' \) are calculated in similar fashion.

#### 2.7.2 Multiplicative Correction

The multiplicative Correction (MC) is a metameric correction wherein the tristimulus values of the batch in the test illuminant are, in turn, multiplied by the ratio of the
corresponding tristimulus values of the standard in the reference illuminant to the corresponding tristimulus values of the batch in the reference illuminant (Fairman 1991).

Mathematically (Brockes 1970), the corrected tristimulus value $X_{corr}$ of the batch is calculated:

$$X_{corr} = X_{b,t} \frac{X_{s,r}}{X_{b,r}},$$

(2.46)

where the subscript $b$ and $s$ represent the batch and the standard, respectively; the subscript $t$ and $r$ represent the test illuminant and the reference illuminant, respectively. $Y_{corr}$ and $Z_{corr}$ are calculated in similar fashion.

### 2.8 Kubelka-Munk theory

In the field of colorant formation and shading, Kubelka-Munk theory is used as the mathematical treatment to reveal the relationship of colorant and color in this kind of media, that is, to predict internal reflectance of the sample from background reflectance, absorption and scattering properties of the colorant layer, and the thickness of the colorant layer (Berns 2000; Allen 1980). For opaque samples, a pair of simplified equations were derived:

$$R_{\lambda,i} = 1 + \left( \frac{K}{S} \right)_{\lambda} - \left[ \left( \frac{K}{S} \right)_{\lambda}^2 + 2 \left( \frac{K}{S} \right)_{\lambda} \right]^2,$$

(2.47)

$$\left( \frac{K}{S} \right)_{\lambda} = \frac{(1 - R_{\lambda,i})^2}{2R_{\lambda,i}},$$

(2.48)
where $R_{\lambda t}$ is the internal reflectance of sample; $\left( \frac{K}{S} \right)_\lambda$ is the ratio of absorption $K_\lambda$ and scattering $S_\lambda$.

In the two equations above, absorption $K_\lambda$ and scattering $S_\lambda$ appear only as a ratio. For a mixture, the $\left( \frac{K}{S} \right)_\lambda$ ratio is the additive combination of each colorant’s unit absorptivity, $k_\lambda$, and unit scattering, $s_\lambda$, scaled by effective concentration, $c$, plus the absorption and scattering of the substrate (Berns 2000). Mathematically,

$$\left( \frac{K}{S} \right)_\lambda = \frac{k_{\lambda,t} + c_1 k_{\lambda,1} + c_2 k_{\lambda,2} + \cdots + c_n k_{\lambda,n}}{s_{\lambda,t} + c_1 s_{\lambda,1} + c_2 s_{\lambda,2} + \cdots + c_n s_{\lambda,n}},$$

(2.49)

where $k_{\lambda,t}$ and $s_{\lambda,t}$ respectively represent the absorption coefficient and scattering coefficient of the substrate without coloration, $c_1, c_2$ are the concentrations of the various colorants, and $k_{\lambda,1}, k_{\lambda,n}$ and $s_{\lambda,1}, s_{\lambda,n}$ represent their respective unit absorption and scattering coefficients. This approach requires the use of both the unit absorption coefficients and unit scattering coefficients of the colorants. It is called two-constant Kubelka-Munk theory.

For paint systems, there is no substrate and concentration is defined as a ratio compared with white. In this case, Equation (2.49) becomes

$$\left( \frac{K}{S} \right)_\lambda = \frac{c_1 k_{\lambda,1} + c_2 k_{\lambda,2} + \cdots + c_n k_{\lambda,n}}{c_1 s_{\lambda,1} + c_2 s_{\lambda,2} + \cdots + c_n s_{\lambda,n}},$$

and

$$\sum_{i=1}^{n} c_i = 1.0.$$
Suppose a mixture concluded white and three chromatics. In this case,

\[
\left( \frac{K}{S} \right)_\lambda = \frac{c_1 k_{\lambda,1} + c_2 k_{\lambda,2} + c_3 k_{\lambda,3} + (1 - c_1 - c_2 - c_3) k_{\lambda,w}}{c_1 s_{\lambda,1} + c_2 s_{\lambda,2} + c_3 s_{\lambda,3} + (1 - c_1 - c_2 - c_3) s_{\lambda,w}}.
\]

(2.51)
3. Spectral Database and Dimensionality Reduction

In order to get the alternative process primaries to perform parameric decomposition, a database containing the spectral properties of DuPont pigments was a prerequisite. In this chapter, a spectral database will be developed. Based on this database, two spectral reflectance dimensionality reduction techniques, principal component analysis (PCA) and independent component analysis (ICA) will be explored to derive possible process primaries.

3.1 DuPont Spectral Database Definition

The spectral database was developed using an automotive finish paint system and two-constant Kubelka-Munk turbid-media theory (Kubelka 1931, Berns 2000). Totally there were thirty-eight pigments including reference white and reference black. Table 3-1 shows the pigments names, Ingrd_clr_number and a description.
### Table 3-1. DuPont pigments used in pigment mixing.

<table>
<thead>
<tr>
<th>Pigment No.</th>
<th>Ingrd_clr_number</th>
<th>Description</th>
<th>Pigment No.</th>
<th>Ingrd_clr_number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10131</td>
<td>Ref White</td>
<td>20</td>
<td>1253</td>
<td>Weak Black</td>
</tr>
<tr>
<td>2</td>
<td>10253</td>
<td>Ref Black</td>
<td>21</td>
<td>1362</td>
<td>Ls Red Oxide</td>
</tr>
<tr>
<td>3</td>
<td>259</td>
<td>Secondary Black</td>
<td>22</td>
<td>1630</td>
<td>Ls Yellow Oxide</td>
</tr>
<tr>
<td>4</td>
<td>451</td>
<td>Brown Transoxide</td>
<td>23</td>
<td>1811</td>
<td>Red</td>
</tr>
<tr>
<td>5</td>
<td>468</td>
<td>Transoxide Red</td>
<td>24</td>
<td>3501</td>
<td>Blue</td>
</tr>
<tr>
<td>6</td>
<td>568</td>
<td>Violet</td>
<td>25</td>
<td>3606</td>
<td>Orange</td>
</tr>
<tr>
<td>7</td>
<td>588</td>
<td>Violet</td>
<td>26</td>
<td>3609</td>
<td>Yellow</td>
</tr>
<tr>
<td>8</td>
<td>616</td>
<td>Orange</td>
<td>27</td>
<td>3612</td>
<td>Orange</td>
</tr>
<tr>
<td>9</td>
<td>634</td>
<td>Transoxide Yellow</td>
<td>28</td>
<td>6102</td>
<td>Fine Aluminum</td>
</tr>
<tr>
<td>10</td>
<td>639</td>
<td>Yellow</td>
<td>29</td>
<td>6304</td>
<td>Red</td>
</tr>
<tr>
<td>11</td>
<td>643</td>
<td>Yellow</td>
<td>30</td>
<td>6333</td>
<td>Red</td>
</tr>
<tr>
<td>12</td>
<td>656</td>
<td>Green Gold</td>
<td>31</td>
<td>6501</td>
<td>Blue</td>
</tr>
<tr>
<td>13</td>
<td>690</td>
<td>Yellow</td>
<td>32</td>
<td>6516</td>
<td>Blue</td>
</tr>
<tr>
<td>14</td>
<td>724</td>
<td>Green</td>
<td>33</td>
<td>6546</td>
<td>Blue</td>
</tr>
<tr>
<td>15</td>
<td>805</td>
<td>Red</td>
<td>34</td>
<td>6602</td>
<td>Yellow</td>
</tr>
<tr>
<td>16</td>
<td>823</td>
<td>Red</td>
<td>35</td>
<td>6812</td>
<td>Red</td>
</tr>
<tr>
<td>17</td>
<td>839</td>
<td>Red</td>
<td>36</td>
<td>10630</td>
<td>Hs Yellow Oxide</td>
</tr>
<tr>
<td>18</td>
<td>853</td>
<td>Red</td>
<td>37</td>
<td>10730</td>
<td>Green</td>
</tr>
<tr>
<td>19</td>
<td>1131</td>
<td>Weak White</td>
<td>38</td>
<td>1730</td>
<td>L/S green</td>
</tr>
</tbody>
</table>

The absorption (unit \( k \)) and scattering (unit \( s \)) properties of each pigment were provided by DuPont. The wavelength range was from 400 nm to 700 nm with 10 nm
increments. For the Saunderson correction, $K_1$ and $K_2$ were set to 0.04 and 0.5, respectively.

In this research, there were two approaches to create mixtures using four-pigment formulation. One approach used three chromatics plus the reference white. The other used two chromatics plus reference white and reference black. For the first approach, there were \( \frac{36!}{(36-3)!3!} = 7,140 \) possible pigment combinations and three concentrations, 4\%, 6\%, and 10\%, were designed for the selected three pigments. For each combination, the concentration of the white was \( 100\% - (4\% + 6\% + 10\%) = 80\% \). So in this case, totally \( 7,140 \times 6 = 42,840 \) recipes were designed. For the second approach, there were \( \frac{36!}{(36-2)!2!} = 630 \) possible pigment combinations and two concentrations, 15\% and 25\%, for the selected two pigments. For each combination, the concentration of the black was 10\% and the white 50\%. In this case, \( 630 \times 2 = 1,260 \) recipes were designed. Finally there were \( 7,140 + 630 = 7,770 \) pigment combinations and \( 42,840 + 1,260 = 44,100 \) recipes to define the Dupont spectral dataset. For each recipe, spectral reflectance factor was estimated using the two-constant method of Kubelka-Munk theory. Figure 3-1, 3-2, and 3-3 show the colorimetric histograms of all the samples in the defined Dupont dataset. Figure 3-4, 3-5, and 3-6 show the color gamut of this dataset that are viewed from the top of the \( a^* - b^* \), \( L^* - a^* \), and \( L^* - b^* \) plane, respectively. The CIELAB
colorimetric attributes were calculated for illuminant D65 and the 1931 standard observer.

Figure 3-1. The chroma distribution, $C_{ab}^*$, of the defined DuPont 44,100 samples.

Figure 3-2. The hue distribution, $h_{ab}$, of the defined DuPont 44,100 samples.
Figure 3-3. The lightness distribution, $L^*$, of the defined DuPont 44,100 samples.

Figure 3-4. Pigment color gamut of the defined DuPont database in CIELAB $a^*$-$b^*$ plane.
Figure 3-5. Pigment color gamut of the defined DuPont database in CIELAB L*-a* plane.

Figure 3-6. Pigment color gamut of the defined DuPont database in CIELAB L*-b* plane.
From the figures above, it is clear that these samples gave a good coverage of color space. From Figure 3-2, it can be found that the samples were weighted more heavily in the red part of the color space. This is not surprising since there were eight pigments showing red attribute in the given pigment list, as shown in Table 3-1.

3.2 Principal Component Analysis (PCA)

Principal component analysis was applied to the defined DuPont spectral dataset. The scree plot, cumulative variance plot, and the spectra of the first seven eigenvectors are shown in Figure 3-7. The variances associated with the *i*th eigenvector derived in the reflectance space are listed in Table 3-2. (As many as 31 eigenvectors can be shown.)
Figure 3-7. The first six eigenvectors and the mean derived from the reflectance spectra of the defined DuPont spectral data set.
Table 3-2. The eigenvalue, % variance, and cumulative % variance associated with the ith eigenvector derived in the spectral reflectance space of the defined DuPont dataset.

<table>
<thead>
<tr>
<th>Eigenvector number</th>
<th>Eigenvalue</th>
<th>% Variance</th>
<th>Cumulative % Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5779</td>
<td>71.69</td>
<td>71.69</td>
</tr>
<tr>
<td>2</td>
<td>0.1366</td>
<td>16.94</td>
<td>88.62</td>
</tr>
<tr>
<td>3</td>
<td>0.0622</td>
<td>7.71</td>
<td>96.33</td>
</tr>
<tr>
<td>4</td>
<td>0.0136</td>
<td>1.69</td>
<td>98.02</td>
</tr>
<tr>
<td>5</td>
<td>0.0068</td>
<td>0.85</td>
<td>98.87</td>
</tr>
<tr>
<td>6</td>
<td>0.0046</td>
<td>0.57</td>
<td>99.44</td>
</tr>
<tr>
<td>7</td>
<td>0.0024</td>
<td>0.30</td>
<td>99.74</td>
</tr>
<tr>
<td>8</td>
<td>0.0010</td>
<td>0.12</td>
<td>99.86</td>
</tr>
<tr>
<td>9</td>
<td>0.0005</td>
<td>0.06</td>
<td>99.92</td>
</tr>
<tr>
<td>10</td>
<td>0.0002</td>
<td>0.03</td>
<td>99.95</td>
</tr>
<tr>
<td>11</td>
<td>0.0002</td>
<td>0.02</td>
<td>99.97</td>
</tr>
<tr>
<td>12</td>
<td>0.0001</td>
<td>0.01</td>
<td>99.98</td>
</tr>
<tr>
<td>13</td>
<td>0.0001</td>
<td>0.01</td>
<td>99.99</td>
</tr>
<tr>
<td>14</td>
<td>0.0000</td>
<td>0.01</td>
<td>99.99</td>
</tr>
<tr>
<td>15</td>
<td>0.0000</td>
<td>0.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

The first eigenvector explained most of the variance in the sample set and points out the weighted average direction along which the samples are distributed. The shape of the first eigenvector and the sample mean indicates that the samples in the defined data set were weighted more heavily in the red part of the color space. Similarly, the shapes of the second eigenvector and the third eigenvector reveal that some specific pigments exist in the pigment mixing processing. These three eigenvectors explained 96.33% of the total variance. That is, the majority of samples varied along the directions of these three statistical dimensions. Such eigenvectors were termed "statistical colorants" by Tzeng and Berns (Tzeng 2005). It can be viewed as if there existed a set of colorants whose spectral properties coincide with the directions of the major sample variation, therefore,
the whole sample variations are the exact combinations of the set of imaginary colorants, which are the eigenvectors derived by PCA.

Although the 4th and 5th eigenvectors also have some significant spectral information, they would not be used as the “statistical colorants” since the purpose of this research is to derive three basis functions.

The existence of the sample mean poses a question: whether the sample mean should be included or not. If yes, it essentially reduces the rank of the data set by one. The answer depends on whether these two approaches have the approximate spectral reconstruction accuracy and whether the sample mean can be well estimated by a linear combination of the selected eigenvectors. The colorimetric and spectral accuracy are shown in Table 3-3. The spectral fits for an arbitrary sample using the first three eigenvectors with and without the mean are plotted in Figure 3-8. The colorimetric accuracy was calculated using CIEDE2000 for the 1931 standard observer under illuminant D65. The degree of metamerism MI is expressed by a special index of metamerism that consisted of both a parametric correction (Fariman 1987) for illuminant D65 and the use of CIEDE2000 for the 1931 standard observer under illuminant A. The spectral accuracy is expressed by the RMS of spectral reflectance.
Figure 3-8. An example of the spectral fits using the three approaches of PCA.

Table 3-3. The colorimetric and spectral accuracy of the spectral reconstruction of the entire DuPont dataset using the three approaches of PCA.

<table>
<thead>
<tr>
<th></th>
<th>2 eigenvectors + mean</th>
<th>3 eigenvectors</th>
<th>3 eigenvectors+ mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔE₀₀</td>
<td>Spectral RMS</td>
<td>MI</td>
</tr>
<tr>
<td>Mean</td>
<td>8.61</td>
<td>1.718</td>
<td>0.05</td>
</tr>
<tr>
<td>95% Percentile</td>
<td>18.51</td>
<td>4.42</td>
<td>0.10</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>5.38</td>
<td>1.346</td>
<td>0.02</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.05</td>
<td>0.011</td>
<td>0.01</td>
</tr>
<tr>
<td>Maximum</td>
<td>42.08</td>
<td>13.039</td>
<td>0.16</td>
</tr>
</tbody>
</table>
When the sample mean is counted as a basis function, the approach of two eigenvectors plus the sample mean would have three basis functions to reconstruct the spectra as well as the approach of three eigenvectors without the sample mean. In this case, the latter has the significant improvement in the spectral reconstruction as shown in Figure 3-6 and Table 3-3. Although the approach of three eigenvectors with the sample mean has a closer fit to the data than the approach of three eigenvectors without the sample mean, the difference between them is very small. That is, using the first three eigenvectors, the approach without the sample mean has the approximate spectral reconstruction accuracy with the approach with the sample mean, which requires an extra parameter. Henceforth, considering the goal of finding three basis functions (statistical colorants) of the defined DuPont spectral dataset, the sample mean should not be counted as a basis function. That is, the first three eigenvectors derived in the spectral space of the defined DuPont spectral data set will be used as the process primaries to perform the parameric decomposition.
3.3 Independent Component Analysis (ICA)

The three independent components of the DuPont spectral dataset from the two ICA approaches are shown in Figures 3-9, and 3-10, respectively.

Figure 3-9. The three basis functions of the mean-subtracted DuPont dataset using ICA.

Figure 3-10. The three basis functions of the original DuPont dataset using ICA.
For PCA, since the covariance matrix of the mean-centered data and un-mean-centered data is identical, the two approaches have the same eigenvectors. For ICA, it’s goal is to seek directions in feature such that the resulting signals show independence, so the mean-centered data would result in different independence components from that of the data without a mean offset. Comparing Figure 3-9 and 3-10, the two sets of basis functions are different. For the sake of finding the possible process primaries, both sets of basis functions will be used to perform parameric decomposition in Chapter Five.

One should keep in mind that the variance and order of the independent components cannot be determined (Hyvarinen 2001). So, one should not refer to the lines labeled as $s_l$ in Figure 3-9 and 3-10 the first independent component. This is different from PCA in which the variance and order of the principal component is determined by the corresponding eigenvalue. Interestingly, the basis functions from ICA without the mean subtracted are very similar to that of PCA. However, by definition, they are not equivalent.

The colorimetric and spectral accuracy are shown in Table 3-4. The spectral fits for an arbitrary sample using first three eigenvectors with and without the mean are plotted in Figure 3-9. The colorimetric and spectral accuracy are shown in Table IV. The colorimetric accuracy was calculated using CIEDE2000 for the 1931 standard observer under illuminant D65. The degree of metamerism MI is expressed by a special index of metamerism that consisted of both a parameric correction (Fariman 1987) for illuminant
D65 and the use of CIEDE2000 for the 1931 standard observer under illuminant A. The spectral accuracy is expressed by the RMS of spectral reflectance.

Table 3-4. The colorimetric and spectral accuracy of the spectral reconstruction of the entire DuPont dataset using two approaches of ICA.

<table>
<thead>
<tr>
<th></th>
<th>3 independent components</th>
<th>3 independent components + mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta E_{00} )</td>
<td>Spectral RMS</td>
</tr>
<tr>
<td>Mean</td>
<td>4.63</td>
<td>1.595</td>
</tr>
<tr>
<td>95% Percentile</td>
<td>8.16</td>
<td>4.56</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.23</td>
<td>1.060</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.12</td>
<td>0.001</td>
</tr>
<tr>
<td>Maximum</td>
<td>18.49</td>
<td>8.434</td>
</tr>
</tbody>
</table>
Figure 3-11. An example of the spectral fits using the two approaches of ICA.

Comparing Table 3-3 with Table 3-4 and Figure 3-6 with Figure 3-9, for the mean-centered data, ICA had the better spectral reconstruction performance than PCA, and for the data without a mean offset, ICA also had greater spectral reconstruction accuracy than PCA.

Similarly with PCA, the spectral reconstruction performance of ICA with the mean was better than that without the mean. However the mean vector in ICA is not a spectrum, which was described in Chapter Two. That is, this mean vector has no physical
meaning. Therefore, in this case, there is not an issue whether the mean should be counted as a basis function or not.

3.4 Conclusions

In this chapter, a spectral database containing the spectral properties of DuPont pigments was defined using two-constant Kubelka-Munk turbid-media theory. Two spectral reflectance dimensionality reduction techniques (PCA and ICA) were applied to this database.

It was obvious that both PCA and ICA built effective linear models for spectral reflectance dimensionality reduction. That is, using a small number of basis functions, the spectral reflectance of the dataset was reconstructed with tolerable accuracy. The comparison between PCA and ICA shows ICA has slightly better performance than PCA using not only the mean-subtracted data but also the original data.
4. Metameric and Parameric Pairs Preparation

In the last chapter, a spectral database containing the spectral properties of DuPont pigments was developed. In this chapter, a number of color centers having strong metameric effect will be selected from this database. A number of metameric and parameric batches surrounding these color centers will be synthesized.

4.1 Selection of the Neutral Samples

In order to generate highly metameric recipes, neutral samples ($C_{ab}^*<20$) were sampled as the aim colors from the defined Dupont database. In this research, 1000 samples with $C_{ab}^*<20$ were selected to generate their corresponding metameric pairs. The CIELAB plots of these samples are shown in the following figures. The CIELAB colorimetric attributes were calculated for illuminant D65 and the 1931 standard observer.
Figure 4-1. The chroma distribution of the selected neutral samples.

Figure 4-2. The hue distribution of the selected neutral samples.
4.2 Generation of the Metameric Pairs

Treating each neutral sample as the standard, a combinatorial tristimulus matching algorithm (Allen 1966, Berns 2000) was used to determine its highly metameric recipe. As described in Chapter Three, there were 7,770 different four-pigment combinations. The tristimulus matching algorithm tested every combination to see whether this combination can match the tristimulus values of the standard under the reference condition (Illuminant D65 and the 1931 CIE standard observer). A list of candidate recipes ranked by illuminant metamerism index for the secondary illuminant (illuminant A) was generated. Finally, the recipe with the highest metamerism was selected to generate the batch matching the standard.

From the generated metameric pairs, 24 metameric pairs having largest degree of metamerism and a reasonable sampling in CIELAB space were selected. That is, there
were 24 color centers in this research. The recipes and color difference under four illuminants of these twenty-four metameric pairs are listed in Table 4-1. In this table, for each color center, the first row represents the standard and the second row represents the metameric batch. As described in Chapter Three, in the four-pigment formulation used in this research, three pigments were mixed with the reference white. In Table 4-1, P1, P2, and P3 represent the selected three DuPont pigment number and C1, C2, and C3 represent the corresponding concentration of each pigment, respectively. The concentration of reference white can be expressed as the difference of the total concentration of the three pigments from unity. The CIELAB colorimetric plots of these twenty-four color centers are shown in Figure 4-4. The CIELAB colorimetric attributes were calculated under illuminant D65 and the 1931 standard observer.
Figure 4-4. Colorimetric plots of the selected 24 color centers.

It was worth noting that when selecting these twenty-four metameric pairs, that higher MI was more important than CIELAB position. Finally, although the scatter of these 24 positions in CIELAB space is not very uniform, they have high indices of metamerism as listed in Table 4-1. It can be seen that the difference of colorimetric
attributions of these 24 color centers are not very big. The reason is that they are all neutral samples, which have highly metameric recipes.

Table 4-1. The recipes, and the color difference of the selected 24 metameric pairs.

<table>
<thead>
<tr>
<th>Pair No.</th>
<th>Pigments</th>
<th>Concentrations</th>
<th>CIE94</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1</td>
<td>P2</td>
<td>P3</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>17</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>17</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>15</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>24</td>
<td>38</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>25</td>
<td>29</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>11</td>
<td>22</td>
</tr>
</tbody>
</table>
Table 4-1. (Continued) The recipes, and the color difference of the selected 24 metameric pairs.

<table>
<thead>
<tr>
<th>Pair No.</th>
<th>Pigments</th>
<th>Concentrations</th>
<th>CIE94</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1</td>
<td>P2</td>
<td>P3</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>33</td>
<td>35</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>17</td>
<td>33</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>10</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>23</td>
<td>9</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>23</td>
<td>37</td>
</tr>
<tr>
<td>24</td>
<td>6</td>
<td>22</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>33</td>
<td>35</td>
</tr>
</tbody>
</table>

The spectral reflectances of the 24 metameric pairs are plotted in Figure 4-5.
Figure 4-5. The spectral reflectance of the twenty-four metameric pairs. The green lines represent the standard and the red lines represent the metameric batch. Note that each plot has a different range of the reflectance factor.

From Figure 4-5, each metameric pair has the spectral reflectance that cross at least three times across the visible spectrum. This is in reasonable agreement with Stiles and Wyszecki’s theoretical verification (Stiles 1968).
4.3 Generation of the Parameric Pairs

In this phase, the objective is to develop a number of paramers surrounding each of the 24 color centers. The paramers were generated by perturbing the pigments’ concentrations of the metameric batch to create the varying residual color-difference under the reference illuminant (D65). The perturbations were made along eight component difference (ΔL* Δa* Δb*) directions. For each direction, there were eight levels of component difference: 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0. Table 4-2 illustrates the eight directions with the component difference of 1.5 CIELAB units.

Table 4-2 The eight parameric batches along eight directions with component difference of 1.5 from the standard under illuminant D65.

<table>
<thead>
<tr>
<th>Direction</th>
<th>ΔL*</th>
<th>Δa*</th>
<th>Δb*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>2</td>
<td>-1.5</td>
<td>-1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>-1.5</td>
<td>1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>4</td>
<td>-1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
<td>-1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>-1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>7</td>
<td>1.5</td>
<td>1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

So for each color position, there were 8 x 6=48 parameric pairs with varying color-difference directions and levels. The eight directions ensured that positive and negative sampling were balanced; the varying color difference levels allow the evaluation of the effect of residual color difference under the reference illuminant on the performance of the parameric decomposition. Finally, 24 x 48=1,152 parameric pairs were generated. The positions of the members of the parameric set for the 9th color center
in the CIELAB space under illuminant D65 are shown in Figure 4-6. The reflectance curves of these parameters as well as the metameric pair are shown in Figure 4-7. Their recipes, CIELAB values under illuminant D65, and the color difference under illuminant D65 and illuminant A are listed in Table 4-3.

Figure 4-6. A set of parameters in CIELAB space. The red point represents the standard and the green points represent the 48 parametric batches.
Figure 4-7. The reflectance curves of a set of paramers and the metameric pair.

Table 4-3. The recipes and color attributes of the paramers and the metamer for the 9th color center.

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>L*</th>
<th>a*</th>
<th>b*</th>
<th>ΔE°94 (D65)</th>
<th>ΔE°94 (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>4</td>
<td>24</td>
<td>38</td>
<td>0.100</td>
<td>0.060</td>
<td>0.040</td>
<td>59.40</td>
<td>1.66</td>
<td>-4.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metamer</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.046</td>
<td>0.215</td>
<td>0.020</td>
<td>59.40</td>
<td>1.66</td>
<td>-4.18</td>
<td>0.00</td>
<td>5.10</td>
</tr>
<tr>
<td>Paramer 1</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.203</td>
<td>0.021</td>
<td>58.90</td>
<td>1.16</td>
<td>-4.68</td>
<td>0.82</td>
<td>4.83</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.042</td>
<td>0.206</td>
<td>0.019</td>
<td>58.90</td>
<td>1.16</td>
<td>-3.68</td>
<td>0.78</td>
<td>4.70</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.046</td>
<td>0.203</td>
<td>0.018</td>
<td>58.90</td>
<td>2.16</td>
<td>-4.68</td>
<td>0.78</td>
<td>5.59</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.048</td>
<td>0.226</td>
<td>0.018</td>
<td>58.90</td>
<td>2.16</td>
<td>-3.68</td>
<td>0.82</td>
<td>5.56</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.039</td>
<td>0.180</td>
<td>0.019</td>
<td>59.90</td>
<td>1.16</td>
<td>-4.68</td>
<td>0.82</td>
<td>4.77</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.039</td>
<td>0.192</td>
<td>0.018</td>
<td>59.90</td>
<td>1.16</td>
<td>-3.68</td>
<td>0.78</td>
<td>4.66</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.036</td>
<td>0.158</td>
<td>0.014</td>
<td>59.90</td>
<td>2.16</td>
<td>-4.68</td>
<td>0.78</td>
<td>5.55</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.041</td>
<td>0.194</td>
<td>0.015</td>
<td>59.90</td>
<td>2.16</td>
<td>-3.68</td>
<td>0.82</td>
<td>5.54</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.202</td>
<td>0.024</td>
<td>58.40</td>
<td>0.66</td>
<td>-5.18</td>
<td>1.63</td>
<td>4.76</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.040</td>
<td>0.209</td>
<td>0.020</td>
<td>58.40</td>
<td>0.66</td>
<td>-3.18</td>
<td>1.57</td>
<td>4.39</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.048</td>
<td>0.202</td>
<td>0.017</td>
<td>58.40</td>
<td>2.66</td>
<td>-5.18</td>
<td>1.56</td>
<td>6.12</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.047</td>
<td>0.225</td>
<td>0.015</td>
<td>58.40</td>
<td>2.66</td>
<td>-3.18</td>
<td>1.65</td>
<td>6.13</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.033</td>
<td>0.151</td>
<td>0.018</td>
<td>60.40</td>
<td>0.66</td>
<td>-5.18</td>
<td>1.63</td>
<td>4.61</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.039</td>
<td>0.202</td>
<td>0.020</td>
<td>60.40</td>
<td>0.66</td>
<td>-3.18</td>
<td>1.57</td>
<td>4.29</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.035</td>
<td>0.147</td>
<td>0.012</td>
<td>60.40</td>
<td>2.66</td>
<td>-5.18</td>
<td>1.56</td>
<td>6.05</td>
</tr>
</tbody>
</table>
Table 4-3. The recipes and color attributes of the parameters and the metamer for the 9th color center.

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>L* (D65)</th>
<th>a* (D65)</th>
<th>b* (D65)</th>
<th>ΔE&lt;sub&gt;94&lt;/sub&gt; (D65)</th>
<th>ΔE&lt;sub&gt;94&lt;/sub&gt;* (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.042</td>
<td>0.197</td>
<td>0.013</td>
<td>60.40</td>
<td>2.66</td>
<td>-3.18</td>
<td>1.65</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.043</td>
<td>0.200</td>
<td>0.026</td>
<td>57.90</td>
<td>0.16</td>
<td>-5.68</td>
<td>2.44</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.040</td>
<td>0.223</td>
<td>0.022</td>
<td>57.90</td>
<td>0.16</td>
<td>-2.68</td>
<td>2.36</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.053</td>
<td>0.213</td>
<td>0.017</td>
<td>57.90</td>
<td>3.16</td>
<td>-5.68</td>
<td>2.34</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.050</td>
<td>0.241</td>
<td>0.014</td>
<td>57.90</td>
<td>3.16</td>
<td>-2.68</td>
<td>2.49</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.036</td>
<td>0.163</td>
<td>0.022</td>
<td>60.90</td>
<td>0.16</td>
<td>-5.68</td>
<td>2.44</td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.033</td>
<td>0.182</td>
<td>0.018</td>
<td>60.90</td>
<td>0.16</td>
<td>-2.68</td>
<td>2.36</td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.172</td>
<td>0.014</td>
<td>60.90</td>
<td>3.16</td>
<td>-5.68</td>
<td>2.34</td>
</tr>
<tr>
<td>24</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.040</td>
<td>0.188</td>
<td>0.011</td>
<td>60.90</td>
<td>3.16</td>
<td>-2.68</td>
<td>2.49</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.038</td>
<td>0.177</td>
<td>0.026</td>
<td>57.40</td>
<td>-0.34</td>
<td>-6.18</td>
<td>3.24</td>
</tr>
<tr>
<td>26</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.035</td>
<td>0.206</td>
<td>0.021</td>
<td>57.40</td>
<td>-0.34</td>
<td>-2.18</td>
<td>3.17</td>
</tr>
<tr>
<td>27</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.051</td>
<td>0.196</td>
<td>0.015</td>
<td>57.40</td>
<td>3.66</td>
<td>-6.18</td>
<td>3.11</td>
</tr>
<tr>
<td>28</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.045</td>
<td>0.220</td>
<td>0.011</td>
<td>57.40</td>
<td>3.66</td>
<td>-2.18</td>
<td>3.32</td>
</tr>
<tr>
<td>29</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.033</td>
<td>0.149</td>
<td>0.022</td>
<td>61.40</td>
<td>-0.34</td>
<td>-6.18</td>
<td>3.24</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.028</td>
<td>0.159</td>
<td>0.016</td>
<td>61.40</td>
<td>-0.34</td>
<td>-2.18</td>
<td>3.17</td>
</tr>
<tr>
<td>31</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.039</td>
<td>0.145</td>
<td>0.011</td>
<td>61.40</td>
<td>3.66</td>
<td>-6.18</td>
<td>3.11</td>
</tr>
<tr>
<td>32</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.038</td>
<td>0.183</td>
<td>0.009</td>
<td>61.40</td>
<td>3.66</td>
<td>-2.18</td>
<td>3.32</td>
</tr>
<tr>
<td>33</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.036</td>
<td>0.164</td>
<td>0.026</td>
<td>56.90</td>
<td>-0.84</td>
<td>-6.68</td>
<td>4.04</td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.272</td>
<td>0.028</td>
<td>56.90</td>
<td>-0.84</td>
<td>-1.68</td>
<td>3.99</td>
</tr>
<tr>
<td>35</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.057</td>
<td>0.209</td>
<td>0.015</td>
<td>56.90</td>
<td>4.16</td>
<td>-6.68</td>
<td>3.89</td>
</tr>
<tr>
<td>36</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.051</td>
<td>0.252</td>
<td>0.010</td>
<td>56.90</td>
<td>4.16</td>
<td>-1.68</td>
<td>4.15</td>
</tr>
<tr>
<td>37</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.033</td>
<td>0.148</td>
<td>0.025</td>
<td>61.90</td>
<td>-0.84</td>
<td>-6.68</td>
<td>4.04</td>
</tr>
<tr>
<td>38</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.032</td>
<td>0.197</td>
<td>0.021</td>
<td>61.90</td>
<td>-0.84</td>
<td>-1.68</td>
<td>3.99</td>
</tr>
<tr>
<td>39</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.043</td>
<td>0.151</td>
<td>0.011</td>
<td>61.90</td>
<td>4.16</td>
<td>-6.68</td>
<td>3.89</td>
</tr>
<tr>
<td>40</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.039</td>
<td>0.187</td>
<td>0.007</td>
<td>61.90</td>
<td>4.16</td>
<td>-1.68</td>
<td>4.15</td>
</tr>
<tr>
<td>41</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.202</td>
<td>0.035</td>
<td>56.40</td>
<td>-1.34</td>
<td>-7.18</td>
<td>4.83</td>
</tr>
<tr>
<td>42</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.032</td>
<td>0.211</td>
<td>0.022</td>
<td>56.40</td>
<td>-1.34</td>
<td>-1.18</td>
<td>4.84</td>
</tr>
<tr>
<td>43</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.058</td>
<td>0.202</td>
<td>0.014</td>
<td>56.40</td>
<td>4.66</td>
<td>-7.18</td>
<td>4.66</td>
</tr>
<tr>
<td>44</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.053</td>
<td>0.265</td>
<td>0.009</td>
<td>56.40</td>
<td>4.66</td>
<td>-1.18</td>
<td>4.98</td>
</tr>
<tr>
<td>45</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.029</td>
<td>0.129</td>
<td>0.024</td>
<td>62.40</td>
<td>-1.34</td>
<td>-7.18</td>
<td>4.83</td>
</tr>
<tr>
<td>46</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.028</td>
<td>0.181</td>
<td>0.020</td>
<td>62.40</td>
<td>-1.34</td>
<td>-1.18</td>
<td>4.84</td>
</tr>
<tr>
<td>47</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.044</td>
<td>0.147</td>
<td>0.010</td>
<td>62.40</td>
<td>4.66</td>
<td>-7.18</td>
<td>4.66</td>
</tr>
<tr>
<td>48</td>
<td>6</td>
<td>9</td>
<td>33</td>
<td>0.040</td>
<td>0.194</td>
<td>0.005</td>
<td>62.40</td>
<td>4.66</td>
<td>-1.18</td>
<td>4.98</td>
</tr>
</tbody>
</table>
4.4 Conclusions

In this chapter, 24 color centers with strong metameric effects were selected from the pre-defined Dupont dataset. For each color center, the metameric batch was formulated using Kubelka-Munk theory and a combinatorial tristimulus matching algorithm. Finally, 48 parameric pairs surrounding each color center were formulated. These parameric pairs will be corrected to metameric pairs using various process primaries in the next Chapter.
5. Parameric Decomposition

In the last chapter, a number of parameters surrounding 24 color centers were formulated with varying residual color difference levels under the reference illuminant D65. In this chapter, these parameric pairs were corrected to be metameric pairs using parameric decomposition.

5.1 Different Process Primaries

In Chapter Three, principal component analysis (PCA) and independent component analysis (ICA) were performed on the reflectances of the pre-defined DuPont spectral dataset. In addition, as shown in Appendix A, PCA and ICA were performed on 1,269 reflectance spectra of the chips in the Munsell Book of Color-Matte Finish Collection. The derived first three eigenvectors and three independent components using both the original data and the mean-subtracted data were viewed as the possible “statistical colorants” of the dataset. Therefore, in this research these basis functions were used as the process primaries to perform the parameric decomposition as well as the CIE 1931 color-matching functions multiplied by the spectral power distribution of illuminant D65. The latter were used as the process primaries in Fairman’s matrix R method. Therefore, there were seven sets of process primaries. They are shown in Figure 5-1.
Figure 5-1. Seven sets of process primaries to perform parametric decomposition.
It should be noted that, for both DuPont dataset and Munsell dataset, the process primaries from ICA using the original data are very similar to that of PCA. In addition, for both PCA and ICA, the basis functions of DuPont dataset are different from that of Munsell dataset. This indicates that the “statistical colorants” of these two sets of dataset are different. As discussed in Chapter Two, it is possible that these different process primaries would generate different projection operator matrix S unless there is the linear transform relationship between different process primaries. The computer renderings of the generated seven projection operators are displayed in Figure 5-2.
Figure 5-2. Seven generated projection operators from seven sets of process primaries.
From Figure 5-2, the matrix S from the CIE 1931 color matching functions weighted by illuminant D65 (matrix R) is symmetrical, as proved by Cohen. However, the other three matrixes are not symmetrical. Since all the matrices are computed from different process primaries weighted by the CIE 1931 color matching functions, they have the characteristics of color matching functions, as shown in Figure 5-2. That is, they have peaks in the color-matching functions peak wavelengths. The diagonal of matrix S shown in Figure 5-3 further confirms this.

Figure 5-3. The diagonal of Matrix S.
Except for the diagonal of matrix S from ICA using the mean-subtracted data, the other five diagonals have the three peak positions near the peak wavelengths of the normal human visual system. This is the reason that the weighted root mean square error (RMS) using the diagonal of the matrix R is a good spectral curve difference metric. It should be noted that the diagonals of the matrix S from PCA using Munsell dataset, ICA using the original Munsell dataset, and ICA using the original DuPont dataset are almost identical.

5.2 Fundamental Stimuli

In a given illuminant-observer space, the fundamental stimulus is unique to a certain set of tristimulus values. However, the choice of different process primaries will lead to different fundamental and metameric black components of the parameric samples. Figure 5-4 demonstrates the fundamental stimuli using the seven sets of process primaries and the metamers of the 24 color centers.
Figure 5-4. The fundamental stimuli from seven sets of process primaries and the metamers of the 24 color centers.
For the process primaries from ICA using the mean-subtracted data, the generated fundamental stimuli do not satisfy the reflectance factor’s limit, which is from zero to unity. That is, they are not physically possible and natural. Therefore, these two sets of process primaries were not used to perform the further parameric decomposition. In the following section of this thesis, for convenience, the approach of ICA using the original data is denoted as ICA.

From Figure 5-4, it is clear that the fundamental stimuli from PCA using Munsell dataset, ICA using the original Munsell dataset, and ICA using the original DuPont dataset are very similar. The reason of this similarity is that these three methods have almost identical matrix S as shown in Figure 5-2 and 5-3. It should be noted that for each set of process primaries, the 24 fundamental stimuli have similar shapes. For the sake of convenience of comparison between the five sets of process primaries, the fundamental stimuli of the 9th color center are plotted in Figure 5-5, as an example.

![Image of Figure 5-5](image_url)

Figure 5-5. The fundamental stimuli and the metameric pair of the 9th color center.
5.3 Metameric Correction

The generated 1,152 parameric batches for the 24 standards were corrected to be metameric matches to the standards using the parameric decomposition method with five sets of process primaries. In order to evaluate the parameric decomposition performance of each set of process primaries, the metameric batch for each color center should be formulated using two-constant Kubelka-Munk turbid-media theory (Kubelka, 1931) with the same DuPont pigments as the corresponding parameric batches. Then the formulated spectra are compared with the corrected spectra using the various process primaries. The reason of this kind of comparison is that Kubelka-Munk turbid-media theory is viewed as the best theory in the field of colorant formulation and shade. In fact, the formulation of the metameric batch was accomplished in the step of the generation of the metameric pairs. The generated metameric spectrum for each color center would be as the reference to evaluate the corrected spectra from the three sets of process primaries.

Figure 5-6 shows the corrected spectra using five sets of process primaries as well as the formulated metameric spectrum using Kubelka-Munk theory for a parameric pair.
Figure 5-6. An example of the corrected spectra using five sets of process primaries, the metameric spectrum using Kubelka-Munk theory, the parameric batch, and the standard.

The seven curves except for the parameric batch in Figure 5-6 are now the metamers which have the same tristimulus values for the reference illuminant (D65) and CIE 1931 2° standard observer. That is, each of the six metameric batches (the colored lines in Figure 5-6) and the standard (the black solid line in Figure 5-6) become a metameric pair from a parameric pair (the two black lines), which has the residual color difference of 3.39 $\Delta E^*_{04}$. From Figure 5-6, it should be noted that the corrected spectrum
using PCA-based process primaries is almost identical to that from ICA-based process primaries. By comparison, the corrected spectra from ICA and PCA-based process primaries are a closer fit to the metameric spectrum from Kubelka-Munk theory than that from the CMFs-based process primaries (Matrix R method).

The spectral difference between the corrected spectra and the formulated metameric spectra using Kubelka-Munk theory are plotted in Figure 5-7, 5-8, 5-9, 5-10, and 5-11 for five sets of process primaries, respectively.

![Figure 5-7. Difference plot of corrected spectra using matrix R method from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters.](image-url)
Figure 5-8. Difference plot of corrected spectra using DuPont PCA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters.

Figure 5-9. Difference plot of corrected spectra using DuPont ICA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters.
Figure 5-10. Difference plot of corrected spectra using Munsell PCA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters.

Figure 5-11. Difference plot of corrected spectra using Munsell ICA-based process primaries from formulated spectra using Kubelka-Munk theory for the entire set of 1,152 parameters.
It can be found that difference plot for PCA-based and ICA-based process primaries look very similar. Relatively, the corrected spectra using matrix R method presented larger deviation from the formulated spectra using Kubelka-Munk theory.

The root-mean-square (RMS) spectral error between the corrected spectrum and that formulated using Kubelka-Munk theory was calculated as a goodness of fit metric for the performance of each set of process primaries. Figure 5-12 shows the mean spectral RMS errors for the five sets of process primaries.

![Figure 5-12. The mean spectral RMS error between the corrected spectra from five sets of process primaries and the formulated metameric spectra from Kubelka-Munk theory for the 24 color centers.](image)

For all of the color centers, the corrected spectra using CMFs-based process primaries (matrix R) presented much more deviation from the formulated spectra using Kubelka-Munk theory than that using the other four sets of process primaries. For several
color centers, PCA-based approach performed better than ICA-based approach. However, in terms of overall performance, they produced very similar spectra comparing with the formulated spectra using Kubelka-Munk theory, especially for the process primaries from PCA for Munsell dataset, ICA for the original Munsell dataset, and ICA for the original DuPont dataset. The latter three process primaries produced slightly closer spectra to the formulated spectra using Kubelka-Munk theory than that from PCA for DuPont dataset.

The statistical RMS error between the corrected spectrum and that formulated using Kubelka-Munk theory for the pooled data of all color centers is listed in Table 5.1.

Table 5-1. The spectral RMS error between the corrected spectrum using five sets of process primaries and that formulated using Kubelka-Munk theory for the pooled data of all color centers.

<table>
<thead>
<tr>
<th>Spectral RMS Error</th>
<th>Matrix R</th>
<th>PCA DuPont</th>
<th>ICA DuPont</th>
<th>PCA Munsell</th>
<th>ICA Munsell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0124</td>
<td>0.0052</td>
<td>0.0045</td>
<td>0.0044</td>
<td>0.0044</td>
</tr>
<tr>
<td>95% Percentile</td>
<td>0.0317</td>
<td>0.0129</td>
<td>0.0110</td>
<td>0.0111</td>
<td>0.0106</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0094</td>
<td>0.0039</td>
<td>0.0036</td>
<td>0.0036</td>
<td>0.0035</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.0006</td>
<td>0.0004</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0004</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.0525</td>
<td>0.0266</td>
<td>0.0265</td>
<td>0.0271</td>
<td>0.0259</td>
</tr>
</tbody>
</table>

5.4 The Estimation of Metameric Indices

After the parameric decomposition, the meaningful metameric indices can be calculated from conventional formulas. In this research, the metameric indices between the standards and the formulated metameric spectra using Kubelka-Munk theory are assumed to be the "true" values indicating the degree of metamerism. Therefore, the goodness of fit of the estimated metameric indices using five sets of process primaries to that using
Kubelka-Munk theory were used to evaluate the performance of various sets of process primaries.

As described in Chapter Two, there are two types of metameric indices, special index of metamerism and general index of metamerism.

5.4.1 Special Index of Metamerism: Change in Illuminant

The special indexes of metamerism of the corrected parameric pairs were calculated using CIE94 for three test illuminants, illuminant A, F2, and F11. Table 5-3 shows the estimated MI values of a corrected parameric pair using five sets of process primaries as well as the estimated MI values using additive correction (AC) and multiplicative correction (MC). The recipes and color-differences of this pair as well as the formulated metameric batch using Kubelka-Munk theory are listed in Table 5-2. The color-difference between the standard and the metameric batch under each of the three test illuminants is viewed as the “true” special index of metamerism of the corresponding parameric pair.

| Table 5-2 The recipes and the color difference of a representative parameric pair and the corresponding formulated metameric pair using Kubelka-Munk theory. |
|---|---|---|---|---|
| | Recipe | \( \Delta E^*_{94} \) (D65 2°) | \( \Delta E^*_{94} \) (A 2°) | \( \Delta E^*_{94} \) (F2 2°) | \( \Delta E^*_{94} \) (F11 2°) |
| Pigments | Concentrations | |
| Standard | 1 | 4 | 24 | 38 | 0.800 | 0.100 | 0.060 | 0.040 |
| Parameric Batch | 1 | 6 | 9 | 33 | 0.558 | 0.038 | 0.177 | 0.026 | 3.24 | 5.22 | 5.17 | 3.69 |
| Metameric Batch | 1 | 6 | 9 | 33 | 0.720 | 0.046 | 0.215 | 0.020 | 0.00 | 5.10 | 2.42 | 3.05 |
Table 5-3. The estimated special index of metamerism of a representative corrected parameric pair using seven correction techniques and the formulated metameric pair using Kubelka-Munk theory for three test illuminants.

<table>
<thead>
<tr>
<th></th>
<th>MI (D65-&gt;A)</th>
<th>MI (D65-&gt;F2)</th>
<th>MI(D65-&gt;F11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix R</td>
<td>4.15</td>
<td>1.98</td>
<td>2.56</td>
</tr>
<tr>
<td>PCA DuPont</td>
<td>5.63</td>
<td>2.82</td>
<td>3.45</td>
</tr>
<tr>
<td>ICA DuPont</td>
<td>5.08</td>
<td>2.44</td>
<td>3.00</td>
</tr>
<tr>
<td>PCA Munsell</td>
<td>4.93</td>
<td>2.34</td>
<td>2.98</td>
</tr>
<tr>
<td>ICA Munsell</td>
<td>4.98</td>
<td>2.36</td>
<td>2.99</td>
</tr>
<tr>
<td>AC</td>
<td>5.25</td>
<td>3.28</td>
<td>3.07</td>
</tr>
<tr>
<td>MC</td>
<td>4.90</td>
<td>2.98</td>
<td>2.83</td>
</tr>
<tr>
<td>K-M</td>
<td>5.10</td>
<td>2.42</td>
<td>3.05</td>
</tr>
</tbody>
</table>

Some statistical tests could be performed in order to establish whether the residual error between the MI values using Kubelka-Munk theory and the estimated MI values using various metameric correction methods are significantly different. The F-test is a parametric test which is designed to test if two population variances are equal (Cui 2005). It is based on certain assumptions of normality about the data. So before using it, it is important to ensure that the data have a normal distribution. Figure 5.13 shows the distribution histogram of the residual errors between the MI values using Kubelka-Munk theory and the estimated MI values using matrix R method for all the parameters under illuminant A.
Figure 5-13. Testing MI residual data against a normal distribution.

The histogram in Figure 5.13 indicates that the residual errors between the MI values using Kubelka-Munk theory and the estimated MI values using matrix R method have a reasonable normal distribution. Similar distribution was found in the data from other metameric correction method. So F-test was performed to see whether the residual error between the MI values using Kubelka-Munk theory and the estimated MI values using various metameric correction methods are significantly different. The F-test hypothesis is described below.

1) Formulate the null and alternate hypotheses (two-tailed)

\[ H_0: V_A = V_B \] (Two correction methods without significant difference)

\[ H_a: V_A \neq V_B \] (Two correction methods with significant difference)
where $V_A$ and $V_B$ represent the variances of the residuals for correction method A and B, respectively.

2) Calculate the $F$ value: $F = V_A / V_B$

3) Reject the hypothesis ($H_0$) if $F > F_c$ or if $F < 1 / F_c$ where $F_c = 1.123 (1 / F_c = 0.891)$ is the critical value depending on the number of samples ($N=1,152$ in this research) and the assumed significance level ($\alpha=0.05$ in this research). That is, the upper and lower critical values to accept the hypothesis $H_0$ are 1.123 and 0.891 of the 95% confidence level, respectively.

The F-test results for the MI residuals from various metameric correction methods are shown in Table 5-3, 5-4 and 5-5, respectively, for illuminant A, F2, and F11.

<table>
<thead>
<tr>
<th></th>
<th>Matrix R</th>
<th>PCA DuPont</th>
<th>ICA DuPont</th>
<th>PCA Munsell</th>
<th>ICA Munsell</th>
<th>AC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix R</td>
<td></td>
<td>4.030</td>
<td>7.797</td>
<td>7.664</td>
<td>8.064</td>
<td>1.247</td>
<td>2.942</td>
</tr>
<tr>
<td>PCA DuPont</td>
<td>4.030</td>
<td></td>
<td>1.935</td>
<td>1.902</td>
<td>2.001</td>
<td>3.230</td>
<td>1.369</td>
</tr>
<tr>
<td>ICA DuPont</td>
<td>7.797</td>
<td>1.935</td>
<td></td>
<td>1.017</td>
<td>1.034</td>
<td>6.251</td>
<td>2.651</td>
</tr>
<tr>
<td>PCA Munsell</td>
<td>7.664</td>
<td>1.902</td>
<td>1.017</td>
<td></td>
<td>1.052</td>
<td>6.144</td>
<td>2.605</td>
</tr>
<tr>
<td>ICA Munsell</td>
<td>8.064</td>
<td>2.001</td>
<td>1.034</td>
<td>1.052</td>
<td></td>
<td>6.465</td>
<td>2.741</td>
</tr>
<tr>
<td>MC</td>
<td>2.942</td>
<td>1.369</td>
<td>2.651</td>
<td>2.605</td>
<td>2.741</td>
<td>2.358</td>
<td></td>
</tr>
</tbody>
</table>
Table 5-5. Significance of the difference between the MI residuals from various metameric correction methods for test illuminant F2.

<table>
<thead>
<tr>
<th></th>
<th>Matrix R</th>
<th>PCA DuPont</th>
<th>ICA DuPont</th>
<th>PCA Munsell</th>
<th>ICA Munsell</th>
<th>AC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix R</td>
<td>3.741</td>
<td>6.183</td>
<td>5.969</td>
<td>6.559</td>
<td>2.347</td>
<td>1.418</td>
<td></td>
</tr>
<tr>
<td>PCA DuPont</td>
<td>3.741</td>
<td>1.653</td>
<td>1.596</td>
<td>1.753</td>
<td>8.780</td>
<td>2.637</td>
<td></td>
</tr>
<tr>
<td>ICA DuPont</td>
<td>6.183</td>
<td>1.653</td>
<td>1.035</td>
<td>1.061</td>
<td>14.509</td>
<td>4.359</td>
<td></td>
</tr>
<tr>
<td>PCA Munsell</td>
<td>5.969</td>
<td>1.596</td>
<td>1.035</td>
<td>1.099</td>
<td>6.144</td>
<td>4.208</td>
<td></td>
</tr>
<tr>
<td>ICA Munsell</td>
<td>6.559</td>
<td>1.753</td>
<td>1.061</td>
<td>1.099</td>
<td>15.39</td>
<td>4.624</td>
<td></td>
</tr>
<tr>
<td>MC</td>
<td>1.418</td>
<td>2.637</td>
<td>4.359</td>
<td>4.208</td>
<td>4.624</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-6. Significance of the difference between the MI residuals from various metameric correction methods for test illuminant F11.

<table>
<thead>
<tr>
<th></th>
<th>Matrix R</th>
<th>PCA DuPont</th>
<th>ICA DuPont</th>
<th>PCA Munsell</th>
<th>ICA Munsell</th>
<th>AC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix R</td>
<td>2.290</td>
<td>3.695</td>
<td>3.662</td>
<td>3.631</td>
<td>1.139</td>
<td>1.363</td>
<td></td>
</tr>
<tr>
<td>PCA DuPont</td>
<td>2.290</td>
<td>1.613</td>
<td>1.599</td>
<td>1.586</td>
<td>2.608</td>
<td>1.680</td>
<td></td>
</tr>
<tr>
<td>ICA DuPont</td>
<td>3.695</td>
<td>1.613</td>
<td>1.009</td>
<td>1.017</td>
<td>4.208</td>
<td>2.711</td>
<td></td>
</tr>
<tr>
<td>PCA Munsell</td>
<td>3.662</td>
<td>1.599</td>
<td>1.009</td>
<td>1.008</td>
<td>4.170</td>
<td>2.687</td>
<td></td>
</tr>
<tr>
<td>ICA Munsell</td>
<td>3.631</td>
<td>1.586</td>
<td>1.017</td>
<td>1.008</td>
<td>4.136</td>
<td>2.665</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>1.139</td>
<td>2.608</td>
<td>4.208</td>
<td>4.170</td>
<td>4.136</td>
<td></td>
<td>1.552</td>
</tr>
<tr>
<td>MC</td>
<td>1.363</td>
<td>1.680</td>
<td>2.711</td>
<td>2.687</td>
<td>2.665</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As expected, the F values between PCA Munsell, ICA DuPont, and ICA Munsell fell inside the range from 0.891 to 1.123. This indicates that the difference between these three methods is not statistically significant to the 95% confidence level. Comparing five parameric decomposition methods, ICA-based methods and PCA-based methods have similar performance (the F-values are relatively close to the critical value), while there are relatively significant differences between these two approaches and matrix R method. This is consistent with the comparison of the corrected spectra using these five parameric
decomposition methods as shown in Figure 5-6 ~ 5-12 and Table 5-1. Interestingly, the matrix R method has relatively similar performance with additive and multiplicative correction methods (AC and MC). The larger F value appears in the comparison between the alternative process primaries method (PCA and ICA) and AC method. This indicates that these two approaches have the most significant difference in estimating the MI values.

In addition to F-test, the linear-fit lines and the correlation coefficients between the MI values using Kubelka-Munk theory and the estimated MI values for all the parameters using various correction methods are shown in Figure 5-14, 5-15, and 5-16, respectively, for illuminant A, F2, and F11.
Figure 5-14. Correlation of MI (D65->A) between Kubelka-Munk theory and seven metamerter correction methods.
Figure 5-15. Correlation of MI (D65→F2) between Kubelka-Munk theory and seven metameric correction methods.
Figure 5-16. Correlation of MI (D65->F11) between Kubelka-Munk theory and seven metameric correction methods.
Comparing the performance of various metameric correction methods, the same trend was seen for the three test illuminants. The MI values from PCA and ICA-based process primaries have the better agreement with that from Kubelka-Munk theory than matrix R method, AC method, and MC method. The AC method produced very similar MI values with that from matrix R method.

5.4.2 Index for Metamerism Potential

The index for metamerism potential of each corrected pair using the five sets of process primaries was calculated according to Equation 2-1. That is, the index for metamerism potential was expressed by weighted root mean square (WRMS) error using the diagonal of matrix R. Similarly, the spectral WRMS error between the standard and the formulated metameric batch using Kubelka-Munk theory was calculated to be the “true” index for metamerism potential. Matrix R was generated for the reference illuminant D65 and CIE 1931 2° standard observer. The diagonal of it can be found in Figure 5-3.

Similarly, F-test was performed in order to establish whether the residual error between the WRMS values using Kubelka-Munk theory and the estimated WRMS values using the five set of process primaries are significantly different. The F-test results for the WRMS residuals from the five process primaries are shown in Table 5-7.
Table 5-7. Significance of the difference between the WRMS residuals using the five sets of process primaries.

<table>
<thead>
<tr>
<th></th>
<th>Matrix R</th>
<th>PCA DuPont</th>
<th>ICA DuPont</th>
<th>PCA Munsell</th>
<th>ICA Munsell</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA DuPont</td>
<td></td>
<td>1.362</td>
<td>1.328</td>
<td>1.418</td>
<td></td>
</tr>
<tr>
<td>ICA DuPont</td>
<td>4.623</td>
<td></td>
<td>1.025</td>
<td>1.041</td>
<td></td>
</tr>
<tr>
<td>PCA Munsell</td>
<td>4.509</td>
<td>1.328</td>
<td></td>
<td>1.068</td>
<td></td>
</tr>
<tr>
<td>ICA Munsell</td>
<td>4.814</td>
<td>1.418</td>
<td>1.041</td>
<td></td>
<td>1.068</td>
</tr>
</tbody>
</table>

Table 5-7 reveals the same trend as the comparison of special index of metamerism, as shown in Table 5-4~5-6 for the five sets of process primaries.

**5.5 Conclusions**

In this chapter, five sets of process primaries were shown. The generated 1,152 parameric pairs were corrected to metameric pairs through parameric decomposition using these five sets of process primaries. It was found that the choice of different process primaries led to different fundamental and metameric black components of the parameric samples, which in turn resulted in different corrected spectra of the parameric samples. The corrected spectra using the five sets of process primaries were compared with the formulated metameric spectra using Kubelka-Munk theory. The results showed that CMFs-based process primaries (matrix R method) generated the metameric spectra with the largest deviation from that using Kubelka-Munk theory. The metameric spectra from ICA-based and PCA-based process primaries using Munsell dataset were almost identical to those from ICA-based process primaries using DuPont dataset. The corrected spectra
using these three sets of process primaries were slightly closer to those using Kubelka-Munk theory than those from PCA-based process primaries using DuPont dataset.

The metameric indices of the corrected parameric pairs using the five sets of process primaries were calculated. The special index of metamerism, change in illuminant, was expressed by the CIE94 color-difference equation under three test illuminants and the index of metamerism potential was calculated using the weighted spectral RMS error using the diagonal of matrix R for the reference illuminant D65 and CIE 1931 2° standard observer. The correlation between these values and those from Kubelka-Munk theory were computed. As expected, the estimated metameric indices from CMFs-based process primaries present relatively poor correlation to those from Kubelka-Munk theory. The ICA-based process primaries showed very similar performance in estimation of metameric indices with PCA-based process primaries, especially for the process primaries derived from Munsell dataset. This result was consistent with the comparison of the corrected spectra.

The results above are not surprising. As mentioned in Chapter Two, parameric decomposition can be viewed as batch correction using three “colorants” where the color-mixing model is linear in reflectance (Berns 1994). The CMFs-based process primaries are not the colorant primaries with which the real specimens can be moved in color space (Fairman 1991). Alternatively, the PCA-based and ICA-based process primaries were derived from Munsell and the pre-defined DuPont spectral dataset. They represent the
"statistical colorants" of the dataset. When they are used to perform batch correction, it is expected that there is a better approximation of a color match to tristimulus equality.

In addition, for the special index of metamerism, change in illuminant, additive correction (AC) and multiplicative correction (MC) were compared with the five parametric decomposition methods using the five sets of process primaries above. It was found that these two methods performed worse than ICA-based and PCA-based process primaries method. Interestingly, the CMFs-based process primaries method (matrix R method) has similar performance in estimation of metameric indices with the AC method.
6. Summary and Conclusions

6.1 Summary

A spectral database containing the spectral properties of DuPont pigments was defined using DuPont automotive finish paint system and two-constant Kubelka-Munk turbid-media theory. The samples in this database gave a good coverage of CIELAB color space. Two spectral reflectance dimensionality reduction techniques, principal component analysis (PCA) and independent component analysis (ICA) were performed to reconstruct the spectra and derive the “statistical colorants” of this database. For these two techniques, two approaches using the mean-centered and original data, respectively, were explored.

24 color centers generating a large degree of metamerism and a reasonable sampling in CIELAB space were selected from the defined DuPont database. For each color center, a metameric batch was formulated using Kubelka-Munk theory and a combinatorial tristimulus matching algorithm. Finally, 48 parameric pairs surrounding each color center were formulated to present varying color-difference directions and levels for the reference illuminant.

Four sets of alternative process primaries were derived from PCA and ICA for Munsell and the pre-defined DuPont spectral dataset. The generated paramers were
corrected to metamers through parameric decomposition method using these four sets of process primaries as well as CMFs-based process primaries (matrix R method). For the corrected parameters, special index of metamerism, change in illuminant, and general index of metamerism were calculated. The special index of metamerism, change in illuminant, additive correction (AC) and multiplicative correction (MC) were also calculated using additive correction method (AC) and multiplicative correction method (MC).

6.2 Conclusions

The spectral reconstruction accuracy showed that both PCA and ICA built effective linear models for spectral reflectance dimensionality reduction. That is, using a small number of basis functions, the spectral reflectance of the dataset was reconstructed with tolerable accuracy. For both PCA and ICA, the spectral reconstruction performance with the mean was better than that without the mean using the same number of basis functions. The comparison between PCA and ICA showed ICA had slightly better performance than PCA using not only the mean-centered data but also the data excluding the mean offset.

The choice of different process primaries led to different fundamental and metameric black components of the parameric samples that in turn resulting in different corrected parameric pairs. The reason was that the corrected spectrum was viewed as the linear combination of the process primaries. However, the corrected spectra using PCA Munsell, ICA DuPont, and ICA Munsell are almost identical. The corrected spectra using five sets of process primaries were compared with the formulated metameric
spectra using Kubelka-Munk theory. The results showed that CMFs-based process primaries generated metameric spectra with the largest deviation from that using Kubelka-Munk theory. The process primaries derived from PCA for Munsell dataset, ICA for Munsell dataset, and ICA for DuPont dataset had almost identical performance in correcting the spectra of the parameters. The metameric spectra from these three sets of process primaries were slightly closer to those using Kubelka-Munk theory than those from the process primaries derived from PCA for DuPont dataset.

The F-test results showed that the differences between PCA Munsell, ICA DuPont, and ICA Munsell were insignificant with 95% confidence level. Comparing five parametric decomposition methods, ICA-based methods and PCA-based methods had similar performance (the F-values are relatively close to the critical value), while there were relatively significant differences between these two approaches and matrix R method. The matrix R method had relatively similar performance with additive and multiplicative correction methods (AC and MC).

### 6.3 Future Research

In this research, the metameric indices between the standards and the formulated metameric spectra using Kubelka-Munk theory are assumed to be the “true” values indicating the degree of metamerism to evaluate the performance of each metameric correction method. Another way is to conduct a psychophysical experiment to quantify visually the degree of metamerism. The key issue of this kind of experiment is to get the
visual data representing the degree of illuminant metamerism, not just color difference under multiple sources.
Appendix A

A Comparison of PCA and ICA for Data Reduction of the Munsell Book of Color, Matte Edition

A.1 Introduction

Cohen was the first to analyze the characteristic spectra of the Munsell colors using Principal Component Analysis (PCA) (Cohen 1964). Parkkinen, et al. (Parkkinen 1989) and Fairman, et al. (Fairman 2004) used similar methods for a larger dataset with 1,257 Munsell color chips. Laamanen, et al. (Laamanen 2001) compared Independent Component Analysis (ICA) and PCA in color recognition using 1,269 reflectance spectra of the Munsell Color chips. Their results showed that ICA had better reconstruction performance than PCA with the same number of dimensions. Ramanath, et al. (Ramanath 2004) used the same 1,269 Munsell samples to perform the spectral dimensionality reduction using ICA, PCA, and Neural networks. They found that PCA performed better than ICA in reproducing the spectra of the samples. Given the inconsistent results, these analyses were repeated in this section. In addition, considering the fact that the traditional PCA and ICA used the mean-centered data while in some applications, the mean is discarded (Tzeng 2005), the linear models of PCA and ICA without a mean offset are evaluated.
A.2 Comparison of PCA and ICA

As described in Chapter Two, PCA derives a transformation that makes variables uncorrelated while ICA derives a transformation that makes variables as independent as possible. The statistical independence takes into consideration higher order moments and is a stronger statistical property than uncorrelatedness (the second order statistic used in PCA). So in connection with ICA, PCA is a useful preprocessing step. To illustrate the difference between PCA and ICA, consider a set of two-dimensional samples. The scatter plot of $x_1$ and $x_2$ is shown in Figure A.1.

![Figure A.1. Scatter plot of a two-dimensional sample set.](image)

PCA finds an intrinsic orthogonal coordinate system for the observation data that preserves the maximum amount of variance. Figure A.2 demonstrates this procedure of PCA.
Figure A. 2. The procedure of PCA to a two-dimensional sample set.

The left plot in Figure A.2 indicates strongly that direction $e_1$ accounts for most significant variance, whereas the orthogonal direction $e_2$ has less variance. After applying PCA, the sample set is represented in the new eigenvector coordinate system labeled as $(Y_1, Y_2)$, as shown in the right plot in Figure A.2. In this new coordinate system, the variable $Y_1$ represents the first principal component coordinate along eigenvector $e_1$. The variable $Y_2$ represents the second principal component coordinate along eigenvector $e_2$. That is, the two principal components are the projections of the original sample set on the directions of the two eigenvectors, respectively. So the two eigenvectors, $e_1$ and $e_2$, are the two basis functions resulting from PCA. However, dimensionality reduction is not possible.

ICA finds an independent but not orthogonal coordinate system where the representations of the observation data are independent of each other. Similarly, Figure A.3 demonstrates the procedure of ICA.
Figure A. 3. The procedure of ICA to a two-dimensional sample set.

In the left side of Figure A.3, the red line and green line, labeled as $S_1$ and $S_2$, respectively, represent the directions of the two basis functions being looked for through ICA. After the projection of the original sample set onto the two basis functions, as shown in the right side of Figure A.3, it is clear that the represented variables $S_1$ and $S_2$ in the new coordinate system are independent of each other, that is, knowing the value of $S_1$ does not help in determining the value of $S_2$. It can be seen that the new coordinate system is not the rotated result of the original coordinate system since for ICA the transformation matrix is not an orthogonal matrix. This feature is different from PCA.
A.2 Results and Discussion

In this study, 1,269 reflectance spectra of the chips in the Munsell Book of Color-Matte Finish Collection defined the data set. The Munsell spectra database is available at: http://cs.joensuu.fi/~spectral/databases/download/munsell_spec_matt.htm.

The wavelength range was from 400 nm to 700 nm with 10 nm intervals. That is, the number of spectral bands $n$ equaled 31 and the number of samples $q$ was 1,269. Figure A.3, A.4, and A.5 show the color gamut of the Munsell dataset that are viewed from the top of the $a^*-b^*$, $L^*-a^*$, and $L^*-b^*$ plane, respectively. The CIELAB colorimetric attributes were calculated for illuminant D65 and the 1931 standard observer.

Figure A.4. Color gamut of the Munsell database in CIELAB $a^*-b^*$ plane.
Figure A. 5. Color gamut of the Munsell database in CIELAB L*-a* plane.

Figure A. 6. Color gamut of the Munsell database in CIELAB L*-b* plane.
The three basis functions of the Munsell data set from PCA and the two ICA approaches are shown in Figure A.7, A.8, and A.9, respectively.

Figure A. 7. The three basis functions of the Munsell spectral dataset from PCA and the mean spectrum of the Munsell spectral dataset.
ICA using the mean-subtracted data

Figure A. 8. The three basis functions of the Munsell spectral dataset from ICA using the mean-subtracted data.

ICA using the original data

Figure A. 9. The three basis functions of the Munsell spectral dataset from ICA using the original data.
The three basis functions of the Munsell spectral dataset from PCA and ICA using the mean-subtracted data are consistent with that in the literature but the opposite signs in some basis functions and different normalization. For PCA, since the covariance matrix of the mean-subtracted data and the original data is identical, two approaches have the same eigenvectors. For ICA, its goal is to seek directions in feature such that the resulting signals show independence, so the mean-subtracted data would result in different independence components from that of the data without a mean offset. Comparing Figure A.8 and A.9, two sets of basis functions are different. Interestingly, the basis functions from ICA using the original data are very similar to that of PCA.

The average colorimetric and spectral accuracy of spectral reconstruction using the various dimensionality reduction techniques is listed in Table A.1. The colorimetric accuracy was calculated using CIEDE2000 for 1931 standard observer under illuminant D65. The spectral accuracy is expressed by the RMS of spectral reflectance (Berns 2000). The spectral reconstruction for an arbitrary sample in the Munsell data set using the derived three basis functions from PCA and ICA, are plotted in Figure A.10.
Table A. 1. The average colorimetric and spectral accuracy of the spectral reconstruction for the Munsell dataset with different number of basis functions from PCA and ICA.

<table>
<thead>
<tr>
<th>Function number</th>
<th>PCA using the mean-subtracted data</th>
<th>PCA using the original data</th>
<th>ICA using the mean-subtracted data</th>
<th>ICA using the original data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DE2000 (D65,2°)</td>
<td>DE2000 (D65,2°)</td>
<td>DE2000 (D65,2°)</td>
<td>DE2000 (D65,2°)</td>
</tr>
<tr>
<td>1</td>
<td>16.2 0.077</td>
<td>17.43 0.08</td>
<td>12.02 0.044</td>
<td>16.54 0.077</td>
</tr>
<tr>
<td>2</td>
<td>12.01 0.042</td>
<td>12.47 0.043</td>
<td>1.97 0.023</td>
<td>12.13 0.042</td>
</tr>
<tr>
<td>3</td>
<td>2.18 0.019</td>
<td>1.91 0.021</td>
<td>1.7 0.016</td>
<td>1.86 0.02</td>
</tr>
<tr>
<td>4</td>
<td>1.22 0.013</td>
<td>1.37 0.015</td>
<td>0.64 0.012</td>
<td>1.23 0.014</td>
</tr>
<tr>
<td>5</td>
<td>0.56 0.009</td>
<td>0.76 0.012</td>
<td>0.24 0.008</td>
<td>0.58 0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.56 0.008</td>
<td>0.76 0.011</td>
<td>0.23 0.006</td>
<td>0.46 0.008</td>
</tr>
<tr>
<td>7</td>
<td>0.12 0.006</td>
<td>0.16 0.007</td>
<td>0.13 0.005</td>
<td>0.17 0.006</td>
</tr>
<tr>
<td>8</td>
<td>0.1 0.004</td>
<td>0.15 0.007</td>
<td>0.1 0.003</td>
<td>0.12 0.004</td>
</tr>
<tr>
<td>9</td>
<td>0.09 0.003</td>
<td>0.12 0.004</td>
<td>0.08 0.003</td>
<td>0.09 0.003</td>
</tr>
<tr>
<td>10</td>
<td>0.06 0.002</td>
<td>0.08 0.003</td>
<td>0.08 0.002</td>
<td>0.06 0.003</td>
</tr>
<tr>
<td>11</td>
<td>0.05 0.002</td>
<td>0.07 0.002</td>
<td>0.02 0.002</td>
<td>0.05 0.002</td>
</tr>
<tr>
<td>12</td>
<td>0.05 0.002</td>
<td>0.08 0.002</td>
<td>0.01 0.002</td>
<td>0.05 0.002</td>
</tr>
<tr>
<td>13</td>
<td>0.01 0.001</td>
<td>0.01 0.002</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
</tr>
<tr>
<td>14</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
</tr>
<tr>
<td>15</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
</tr>
<tr>
<td>16</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
<td>0.01 0.001</td>
</tr>
<tr>
<td>17</td>
<td>0.00 0.001</td>
<td>0.00 0.001</td>
<td>0.00 0.001</td>
<td>0.00 0.001</td>
</tr>
</tbody>
</table>
Figure A. 10. An example of the spectral reconstruction using three basis functions from PCA and ICA for Munsell dataset.

It is obvious that both PCA and ICA built effective linear models for spectral reflectance dimensionality reduction. That is, using a small number of basis functions, the spectral reflectance of the dataset can be reconstructed with tolerable accuracy.

For both PCA and ICA, the spectral reconstruction performance using the mean-subtracted data was better than that using the original data with the same number of basis functions. However, if the mean spectrum was treated as a basis function, the results of PCA using the original data would be better than that of PCA using the mean-subtracted
data with the same number of basis functions. However, when the number of functions is larger than two, the difference becomes very small.

For the mean-subtracted data, ICA has the better spectral reconstruction performance than PCA. This conclusion is consistent with the result of Laamanen, et al, however, opposite to that of Ramanath, et al. For the data without a mean offset, ICA also has smaller spectral reconstruction accuracy than PCA.

A significant result is that ICA using the original data has very close performance to PCA using the mean-subtracted data. The latter needs to know the mean spectrum of the dataset before reconstructing it. This indicates that in order to obtain the same spectral reconstruction, there is an extra parameter to know for PCA using the mean-subtracted data comparing with ICA using the original data.
Appendix B

Table B.1. The derived first three eigenvectors and three independent components for Munsell data.

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>PCA</th>
<th>ICA (using the original data)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>400</td>
<td>-0.09483</td>
<td>-0.11571</td>
</tr>
<tr>
<td>410</td>
<td>-0.13535</td>
<td>-0.17545</td>
</tr>
<tr>
<td>420</td>
<td>-0.14416</td>
<td>-0.19512</td>
</tr>
<tr>
<td>430</td>
<td>-0.14455</td>
<td>-0.20431</td>
</tr>
<tr>
<td>440</td>
<td>-0.14471</td>
<td>-0.21357</td>
</tr>
<tr>
<td>450</td>
<td>-0.14426</td>
<td>-0.22174</td>
</tr>
<tr>
<td>460</td>
<td>-0.14547</td>
<td>-0.23037</td>
</tr>
<tr>
<td>470</td>
<td>-0.14844</td>
<td>-0.23984</td>
</tr>
<tr>
<td>480</td>
<td>-0.15138</td>
<td>-0.24656</td>
</tr>
<tr>
<td>490</td>
<td>-0.15409</td>
<td>-0.24615</td>
</tr>
<tr>
<td>500</td>
<td>-0.15917</td>
<td>-0.23393</td>
</tr>
<tr>
<td>510</td>
<td>-0.16671</td>
<td>-0.20303</td>
</tr>
<tr>
<td>520</td>
<td>-0.17236</td>
<td>-0.16461</td>
</tr>
<tr>
<td>530</td>
<td>-0.17494</td>
<td>-0.13271</td>
</tr>
<tr>
<td>540</td>
<td>-0.1787</td>
<td>-0.10312</td>
</tr>
<tr>
<td>550</td>
<td>-0.18646</td>
<td>-0.06287</td>
</tr>
<tr>
<td>560</td>
<td>-0.1945</td>
<td>-0.01346</td>
</tr>
<tr>
<td>570</td>
<td>-0.1986</td>
<td>0.028693</td>
</tr>
<tr>
<td>580</td>
<td>-0.20264</td>
<td>0.066536</td>
</tr>
<tr>
<td>590</td>
<td>-0.20568</td>
<td>0.10536</td>
</tr>
<tr>
<td>600</td>
<td>-0.2066</td>
<td>0.13967</td>
</tr>
<tr>
<td>610</td>
<td>-0.20613</td>
<td>0.16515</td>
</tr>
<tr>
<td>620</td>
<td>-0.20504</td>
<td>0.18118</td>
</tr>
<tr>
<td>630</td>
<td>-0.20501</td>
<td>0.19007</td>
</tr>
<tr>
<td>640</td>
<td>-0.20429</td>
<td>0.19277</td>
</tr>
<tr>
<td>650</td>
<td>-0.20422</td>
<td>0.19352</td>
</tr>
<tr>
<td>660</td>
<td>-0.20381</td>
<td>0.19251</td>
</tr>
<tr>
<td>670</td>
<td>-0.20306</td>
<td>0.19191</td>
</tr>
<tr>
<td>680</td>
<td>-0.20272</td>
<td>0.19335</td>
</tr>
<tr>
<td>690</td>
<td>-0.2023</td>
<td>0.19402</td>
</tr>
<tr>
<td>700</td>
<td>-0.20287</td>
<td>0.19553</td>
</tr>
</tbody>
</table>
## Appendix C

Table B. 2. The derived first three eigenvectors and three independent components for DuPont data.

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>PCA</th>
<th>ICA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>400</td>
<td>-0.00425</td>
<td>-0.0225</td>
</tr>
<tr>
<td>410</td>
<td>-0.01748</td>
<td>-0.05111</td>
</tr>
<tr>
<td>420</td>
<td>-0.02578</td>
<td>-0.07548</td>
</tr>
<tr>
<td>430</td>
<td>-0.03133</td>
<td>-0.10023</td>
</tr>
<tr>
<td>440</td>
<td>-0.03652</td>
<td>-0.12688</td>
</tr>
<tr>
<td>450</td>
<td>-0.04061</td>
<td>-0.15061</td>
</tr>
<tr>
<td>460</td>
<td>-0.04325</td>
<td>-0.17165</td>
</tr>
<tr>
<td>470</td>
<td>-0.04504</td>
<td>-0.1899</td>
</tr>
<tr>
<td>480</td>
<td>-0.04455</td>
<td>-0.20891</td>
</tr>
<tr>
<td>490</td>
<td>-0.04172</td>
<td>-0.22811</td>
</tr>
<tr>
<td>500</td>
<td>-0.03634</td>
<td>-0.24489</td>
</tr>
<tr>
<td>510</td>
<td>-0.02803</td>
<td>-0.26335</td>
</tr>
<tr>
<td>520</td>
<td>-0.01608</td>
<td>-0.28103</td>
</tr>
<tr>
<td>530</td>
<td>5.27E-05</td>
<td>-0.28968</td>
</tr>
<tr>
<td>540</td>
<td>0.019062</td>
<td>-0.29217</td>
</tr>
<tr>
<td>550</td>
<td>0.0362</td>
<td>-0.29245</td>
</tr>
<tr>
<td>560</td>
<td>0.057033</td>
<td>-0.29097</td>
</tr>
<tr>
<td>570</td>
<td>0.0817</td>
<td>-0.28473</td>
</tr>
<tr>
<td>580</td>
<td>0.1131</td>
<td>-0.26227</td>
</tr>
<tr>
<td>590</td>
<td>0.15925</td>
<td>-0.20685</td>
</tr>
<tr>
<td>600</td>
<td>0.20762</td>
<td>-0.13822</td>
</tr>
<tr>
<td>610</td>
<td>0.25061</td>
<td>-0.07166</td>
</tr>
<tr>
<td>620</td>
<td>0.28127</td>
<td>-0.02464</td>
</tr>
<tr>
<td>630</td>
<td>0.29668</td>
<td>0.001759</td>
</tr>
<tr>
<td>640</td>
<td>0.30334</td>
<td>0.020608</td>
</tr>
<tr>
<td>650</td>
<td>0.30578</td>
<td>0.0354</td>
</tr>
<tr>
<td>660</td>
<td>0.30689</td>
<td>0.047166</td>
</tr>
<tr>
<td>670</td>
<td>0.30703</td>
<td>0.052928</td>
</tr>
<tr>
<td>680</td>
<td>0.30799</td>
<td>0.055206</td>
</tr>
<tr>
<td>690</td>
<td>0.3097</td>
<td>0.055681</td>
</tr>
<tr>
<td>700</td>
<td>0.31228</td>
<td>0.055154</td>
</tr>
</tbody>
</table>
Appendix D

Matlab Code for ICA

%ICA of the reflectance spectra of 1269 matt Munsell color
munsell_data=load('Munsell_data.txt');%400nm-700nm with 10nm increments, In %=this
case, it is a 31x1269 matrix

%------------------ICA using the original data---------------------------------------------
m=3 %look for three basis functions
B=jadeR(munsell_data ',m) % apply JADE algorithm
S=B* munsell_data '; %calculate ICs.
for i=1:m;S(i,:)=S(i,:)./norm(S(i,:));end % Normalize ICs to have length equal to %one.
rec_munsell=S'*S*(munsell_data); %Spectral reconstruction
%---------------------------------------------

%------------------ICA using the mean-subtracted data-------------------------------------
M=mean(munsell_data); %the mean vector of the data, it is a 1269x1 vector.
M_munsell=munsell_data-ones(31,1)*M; % calculate the Mean-subtracted data
m=3 %look for three basis functions
B=jadeR(M_munsell ',m) % apply JADE algorithm
S=B* M_munsell '; %calculate ICs.
for i=1:m;S(i,:)=S(i,:)./norm(S(i,:));end
rec_munsell=S'*S* M_munsell +ones(31,1)*M; %Spectral reconstruction
%---------------------------------------------
function B = jadeR(X,m)
% this code is available from http://sig.enst.fr/~cardoso/stuff.html

% Finding the number of sources
[n,T] = size(X);

% whitening & projection onto signal subspace
% ===========================================
if verbose, fprintf('jade -> Whitening the data\n'); end
[U,D] = eig((X*X')/T);
[puiss,k] = sort(diag(D));
rangeW = n-m+1:n ; % indices to the m most significant directions
scales = sqrt(puiss(rangeW)); % scales
W = diag(1./scales) * U(1:n,k(rangeW))' ; % whitener
iW = U(1:n,k(rangeW)) * diag(scales) ; % its pseudo-inverse
X = W*X;

% Estimation of the cumulant matrices.
% ===========================================
if verbose, fprintf('jade -> Estimating cumulant matrices\n'); end

dimsymm = (m*(m+1))/2; % Dim. of the space of real symm matrices
nbcm = dimsymm ; % number of cumulant matrices
CM = zeros(m,m*nbcm); % Storage for cumulant matrices
R = eye(m); %
Qij = zeros(m); % Temp for a cum. matrix
Xim = zeros(1,m); % Temp
Xjm = zeros(1,m); % Temp
scale = ones(m,1)/T ; % for conven

Range = 1:m ; % will index the columns of CM where to store the cum. mats.
for im = 1:m
    Xim = X(im,:);
    Qij = (scale*(Xim*Xim)).*X' - R - 2 * R(:,im)*R(:,im)';
    CM(:,Range) = Qij;
    Range = Range + m;
for jm = 1:im-1
    Xjm = X(jm,:);
    Qij = (scale*(Xim*Xjm)).*X' - R(:,im)*R(:,jm)' - R(:,jm)*R(:,im)';
    CM(:,Range) = sqrt(2)*Qij;
end
Range = Range + m;
end;
end;

%%% joint diagonalization of the cumulant matrices

%%% Init
if 1,  %%% Init by diagonalizing a *single* cumulant matrix.

    if verbose, fprintf(jade -> Initialization of the diagonalization\n'); end
    [V,D] = eig(CM(:,1:m));  % For instance, this one
    for u=1:m:m*nbcm,    % updating accordingly the cumulant set given the init
        CM(:,u:u+m-1) = CM(:,u:u+m-1)*V;
    end;
    CM = V'*CM;

else,  %%% The dont-try-to-be-smart init
    V = eye(m);  % la rotation initiale
end;

seuil = 1/sqrt(T)/100;  % A statistically significant threshold
encore = 1;
swEEP = 0;
updates = 0;
g  = zeros(2,nbcm);
gg = zeros(2,2);
G  = zeros(2,2);
c  = 0 ;
s  = 0 ;
ton = 0 ;
toff = 0 ;
theta = 0 ;

%%% Joint diagonalization proper
if verbose, fprintf(jade -> Contrast optimization by joint diagonalization\n'); end

while encore, encore=0;

    if verbose, fprintf(jade -> Sweep #%d\n',sweep); end
    sweep=sweep+1;

for p=1:m-1,
for q=p+1:m,

Ip = p:m:m*nbcm ;
Iq = q:m:m*nbcm ;

%%% computation of Givens angle

    g   = [ CM(p,Ip)-CM(q,Iq) ; CM(p,Iq)+CM(q,Ip) ];
    gg  = g*g';
    ton = gg(1,1)-gg(2,2);
    toff = gg(1,2)+gg(2,1);
    theta = 0.5*atan2( toff , ton+sqrt(ton*ton+toff*toff) );

%%% Givens update

if abs(theta) > seuil, encore = 1 ;
    updates = updates + 1;
    c = cos(theta);
    s = sin(theta);
    G = [ c -s ; s c ] ;

    pair  = [p;q] ;
    V(:,pair) = V(:,pair)*G ;
    CM(pair,:) = G * CM(pair,:)
    CM(:,[Ip Iq]) = [ c*CM(:,Ip)+s*CM(:,Iq) -s*CM(:,Ip)+c*CM(:,Iq) ] ;

%%% fprintf(’jade -> %3d %3d %12.8f\n’,p,q,s);
end

end
end
end if verbose, fprintf(’jade -> Total of %d Givens rotations\n’,updates); end

%%% A separating matrix
\[ B = V^*W; \]

if verbose, fprintf(\texttt{jade} \rightarrow \texttt{Sorting the components\'\,updates}); end

\[ A = iW^*V; \]

\[ [\text{vars,keys}] = \text{sort}(\text{sum}(A.*A)); \]

\[ B = B(\text{keys},:); \]

\[ B = B(m:-1:1,:); \]

\% Signs are fixed by forcing the first column of B to have
\% non-negative entries.

\[ b = B(:,1); \]

\[ \text{signs} = \text{sign}(\text{sign}(b) + 0.1); \]

\[ B = \text{diag}(\text{signs})*B; \]

return ;
References

[Allen 1980]

[Alfvin 1995]

[ASTM E 284]

[Berns 1983]

[Berns 1988]

[Berns 1994]

[Berns 2000]

[Billmeyer 1983]
[Billmeyer 1991]

[Cardoso 1998]

[Choudhury 1996]

[Choudhury 2000]

[Chow 1999]

[CIE 1986]
CIE Colorimetry. CIE Pubisication No. 15.2. Vienna: Central Bureau of the CIE; 1986.

[CIE 1995]

[Cohen 1964]

[Cohen 1982]
[Cohen 1985]

[Cohen 2001]

[Fairchild 2005]

[Fairman 1986]

[Fairman 1987]

[Fairman 1991]

[Fairman 2004]

[Hyvarinen 2000]

[Hyvarinen 2001]
[Imai 2002]

[JADE algorithm]
The JADE algorithm in Matlab is available at: http://www.tsi.enst.fr/icacentral/Algos/cardoso/

[Johnson 2002]

[Kettler 2005]

[Kuehni 1983]

[Laamanen 2001]

[Luo 1996]

[Luo 1996]

[Luo 2001]
[Mohammadi 2004]

[Nimeroff 1965]

[Parkkinen 1989]

[Ramanath 2004]

[Robertson 1983]

[Shaw 1999]

[Rodrigues 1980]

[Tzeng 1999]

[Tzeng 2005]
[Wyszecki 1958]


[Wyszecki 1982]