Color Image Processing and Applications

Book · January 2000

DOI: 10.1007/978-3-662-04186-4

2 authors, including:

Konstantinos Plataniotis
University of Toronto
801 PUBLICATIONS **23,008** CITATIONS

SEE PROFILE

K.N. Plataniotis and A.N. Venetsanopoulos

Color Image Processing and Applications

Engineering – Monograph (English)

February 18, 2000

Springer-Verlag
Berlin Heidelberg NewYork
London Paris Tokyo
Hong Kong Barcelona
Budapest

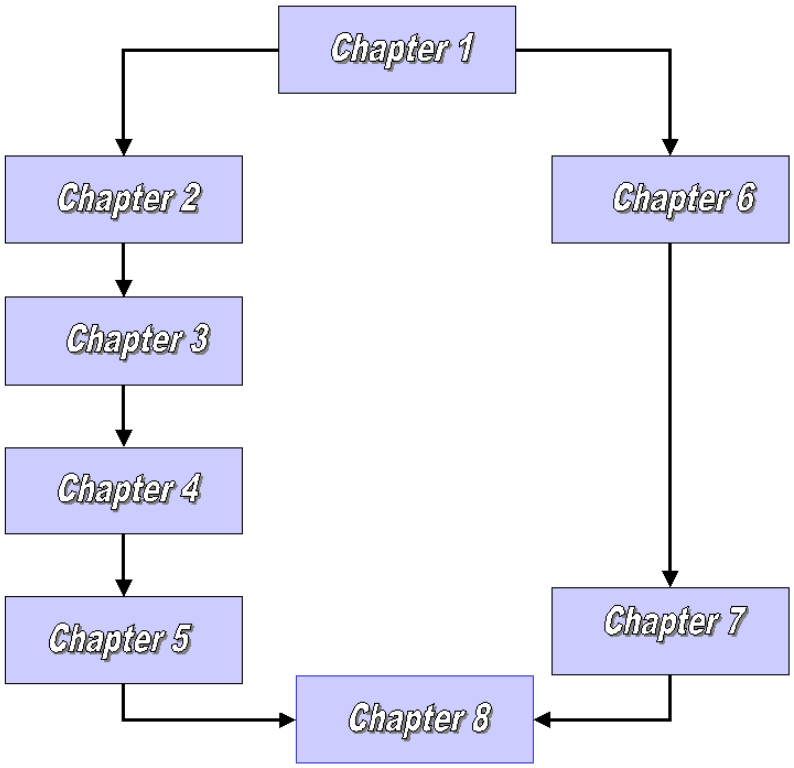
Preface

The perception of color is of paramount importance to humans since they routinely use color features to sense the environment, recognize objects and convey information. Color image processing and analysis is concerned with the manipulation of digital color images on a computer utilizing digital signal processing techniques. Like most advanced signal processing techniques, it was, until recently, confined to academic institutions and research laboratories that could afford the expensive image processing hardware needed to handle the processing overhead required to process large numbers of color images. However, with the advent of powerful desktop computers and the proliferation of image collection devices, such as digital cameras and scanners, color image processing techniques are now within the grasp of the general public.

This book is aimed at researchers and practitioners that work in the area of color image processing. Its purpose is to fill an existing gap in scientific literature by presenting the state of the art research in the area. It is written at a level which can be easily understood by a graduate student in an Electrical and Computer Engineering or Computer Science program. Therefore, it can be used as a textbook that covers part of a modern graduate course in digital image processing or multimedia systems. It can also be used as a textbook for a graduate course on digital signal processing since it contains algorithms, design criteria and architectures for processing and analysis systems.

The book is structured into four parts. The first, Chapter 1, deals with color principles and is aimed at readers who have very little prior knowledge of color science. Readers interested in color image processing may read the second part of the book (Chapters 2-5). It covers the major, although somewhat mature, fields of color image processing. Color image processing is characterized by a large number of algorithms that are specific solutions to specific problems, for example vector median filters have been developed to remove impulsive noise from images. Some of them are mathematical or content independent operations that are applied to each and every pixel, such as morphological operators. Others are algorithmic in nature, in the sense that a recursive strategy may be necessary to find edge pixels in an image.

The third part of the book, Chapters 6-7, deals with color image analysis and coding techniques. The ultimate goal of color image analysis is to enhance human-computer interaction. Recent applications of image analysis includes compression of color images either for transmission across the internetwork or coding of video images for video conferencing. Finally, the fourth part (Chapter 8) covers emerging applications of color image processing. Color is useful for accessing multimedia databases. Local color information, for example in the form of color histograms, can be used to index and retrieve images from the database. Color features can also be used to identify objects of interest, such as human faces and hand areas, for applications ranging from video conferencing, to perceptual interfaces and virtual environments. Because of the dual nature of this investigation, processing and analysis, the logical dependence of the chapters is somewhat unusual. The following diagram can help the reader chart the course.



Logical dependence between chapters

Acknowledgment

We acknowledge a number of individuals who have contributed in different ways to the preparation of this book. In particular, we wish to extend our appreciation to Prof. M. Zervakis for contributing the image restoration section, and to Dr. N. Herodotou for his informative inputs and valuable suggestions in the emerging applications chapter. Three graduate students of ours also merit special thanks. Shu Yu Zhu for her input and high quality figures included in the color edge detection chapter, Ido Rabinovitch for his contribution to the color image coding section and Nicolaos Ikonomakis for his valuable contribution in the color segmentation chapter. We also thank Nicolaos for reviewing the chapters of the book and helping with the Latex formating of the manuscript. We also grateful to Terri Vlassopoulos for proofreading the manuscript, and Frank Holzwarth of Springer Verlag for his help during the preparation of the book. Finally, we are indebted to Peter Androutsos who helped us tremendously on the development of the companion software.

Contents

1. Color Spaces 1
- 1.1 Basics of Color Vision 1
- 1.2 The CIE Chromaticity-based Models 4
- 1.3 The CIE-RGB Color Model 9
- 1.4 Gamma Correction 13
- 1.5 Linear and Non-linear RGB Color Spaces 16
  - 1.5.1 Linear RGB Color Space 16
  - 1.5.2 Non-linear RGB Color Space 17
- 1.6 Color Spaces Linearly Related to the RGB 20
- 1.7 The YIQ Color Space 23
- 1.8 The HSI Family of Color Models 25
- 1.9 Perceptually Uniform Color Spaces 32
  - 1.9.1 The CIE L*u*v* Color Space 33
  - 1.9.2 The CIE L*a*b* Color Space 35
  - 1.9.3 Cylindrical L*u*v* and L*a*b* Color Space 37
  - 1.9.4 Applications of L*u*v* and L*a*b* spaces 37
- 1.10 The Munsell Color Space 39
- 1.11 The Opponent Color Space 41
- 1.12 New Trends 42
- 1.13 Color Images 45
- 1.14 Summary 45

2. Color Image Filtering 51
- 2.1 Introduction 51
- 2.2 Color Noise 52
- 2.3 Modeling Sensor Noise 53
- 2.4 Modeling Transmission Noise 55
- 2.5 Multivariate Data Ordering Schemes 58
  - 2.5.1 Marginal Ordering 59
  - 2.5.2 Conditional Ordering 62
  - 2.5.3 Partial Ordering 62
  - 2.5.4 Reduced Ordering 63
- 2.6 A Practical Example 67
- 2.7 Vector Ordering 69

2.8 The Distance Measures 70
2.9 The Similarity Measures 72
2.10 Filters Based on Marginal Ordering 77
2.11 Filters Based on Reduced Ordering 81
2.12 Filters Based on Vector Ordering 89
2.13 Directional-based Filters 92
2.14 Computational Complexity 98
2.15 Conclusion 100

3. Adaptive Image Filters 107
3.1 Introduction 107
3.2 The Adaptive Fuzzy System 109
3.2.1 Determining the Parameters 112
3.2.2 The Membership Function 113
3.2.3 The Generalized Membership Function 115
3.2.4 Members of the Adaptive Fuzzy Filter Family 116
3.2.5 A Combined Fuzzy Directional and Fuzzy Median Filter 122
3.2.6 Comments 125
3.2.7 Application to *1-D* Signals 128
3.3 The Bayesian Parametric Approach 131
3.4 The Non-parametric Approach 137
3.5 Adaptive Morphological Filters 146
3.5.1 Introduction 146
3.5.2 Computation of the NOP and the NCP 152
3.5.3 Computational Complexity and Fast Algorithms 154
3.6 Simulation Studies 157
3.7 Conclusions 173

4. Color Edge Detection 179
4.1 Introduction 179
4.2 Overview Of Color Edge Detection Methodology 181
4.2.1 Techniques Extended From Monochrome Edge Detection 181
4.2.2 Vector Space Approaches 183
4.3 Vector Order Statistic Edge Operators 189
4.4 Difference Vector Operators 194
4.5 Evaluation Procedures and Results 197
4.5.1 Probabilistic Evaluation 198
4.5.2 Noise Performance 200
4.5.3 Subjective Evaluation 201
4.6 Conclusion 203

5. Color Image Enhancement and Restoration 209
5.1 Introduction 209
5.2 Histogram Equalization 210
5.3 Color Image Restoration 214
5.4 Restoration Algorithms 217
5.5 Algorithm Formulation 220
5.5.1 Definitions 220
5.5.2 Direct Algorithms 223
5.5.3 Robust Algorithms 227
5.6 Conclusions 229

6. Color Image Segmentation 237
6.1 Introduction 237
6.2 Pixel-based Techniques 239
6.2.1 Histogram Thresholding 239
6.2.2 Clustering 242
6.3 Region-based Techniques 247
6.3.1 Region Growing 248
6.3.2 Split and Merge 250
6.4 Edge-based Techniques 252
6.5 Model-based Techniques 253
6.5.1 The Maximum A-posteriori Method 254
6.5.2 The Adaptive MAP Method 255
6.6 Physics-based Techniques 256
6.7 Hybrid Techniques 257
6.8 Application 260
6.8.1 Pixel Classification 260
6.8.2 Seed Determination 262
6.8.3 Region Growing 267
6.8.4 Region Merging 269
6.8.5 Results 271
6.9 Conclusion 273

7. Color Image Compression 279
7.1 Introduction 279
7.2 Image Compression Comparison Terminology 282
7.3 Image Representation for Compression Applications 285
7.4 Lossless Waveform-based Image Compression Techniques 286
7.4.1 Entropy Coding 286
7.4.2 Lossless Compression Using Spatial Redundancy 288
7.5 Lossy Waveform-based Image Compression Techniques 290
7.5.1 Spatial Domain Methodologies 290
7.5.2 Transform Domain Methodologies 292
7.6 Second Generation Image Compression Techniques 304
7.7 Perceptually Motivated Compression Techniques 307

7.7.1 Modeling the Human Visual System 307
7.7.2 Perceptually Motivated DCT Image Coding 311
7.7.3 Perceptually Motivated Wavelet-based Coding 313
7.7.4 Perceptually Motivated Region-based Coding 317
7.8 Color Video Compression 319
7.9 Conclusion 324

8. Emerging Applications 329
8.1 Input Analysis Using Color Information 331
8.2 Shape and Color Analysis 337
8.2.1 Fuzzy Membership Functions 338
8.2.2 Aggregation Operators 340
8.3 Experimental Results 343
8.4 Conclusions 345

A. Companion Image Processing Software 349
A.1 Image Filtering 350
A.2 Image Analysis 350
A.3 Image Transforms 351
A.4 Noise Generation 351

Index 353

List of Figures

1.1 The visible light spectrum 1
1.2 The CIE XYZ color matching functions 7
1.3 The CIE RGB color matching functions 7
1.4 The chromaticity diagram 9
1.5 The Maxwell triangle 10
1.6 The RGB color model 11
1.7 Linear to Non-linear Light Transformation 18
1.8 Non-linear to linear Light Transformation 19
1.9 Transformation of Intensities from Image Capture to Image Display 19
1.10 The HSI Color Space 26
1.11 The HLS Color Space 31
1.12 The HSV Color Space 31
1.13 The L*u*v* Color Space 34
1.14 The Munsell color system 40
1.15 The Opponent color stage of the human visual system 42
1.16 A taxonomy of color models 46

3.1 Simulation I: Filter outputs (1^{st} component) 129
3.2 Simulation I: Filter outputs (2^{nd} component) 129
3.3 Simulation II: Actual signal and noisy input (1^{st} component) 130
3.4 Simulation II: Actual signal and noisy input (2^{nd} component) 131
3.5 Simulation II: Filter outputs (1^{st} component) 132
3.6 Simulation II: Filter outputs (2^{nd} component) 132
3.7 A flowchart of the NOP research algorithm 155
3.8 The adaptive morphological filter 157
3.9 'Peppers' corrupted by 4% impulsive noise 169
3.10 'Lenna' corrupted with Gaussian noise $\sigma = 15$ mixed with 2% impulsive noise 169
3.11 VMF of (3.9) using 3x3 window 170
3.12 $BVDF$ of (3.9) using 3x3 window 170
3.13 HF of (3.9) using 3x3 window 170
3.14 AHF of (3.9) using 3x3 window 170
3.15 $FVDF$ of (3.9) using 3x3 window 170
3.16 $ANNMF$ of (3.9) using 3x3 window 170
3.17 $CANNMF$ of (3.9) using 3x3 window 170

3.18 $BFMA$ of (3.9) using 3x3 window 170
3.19 VMF of (3.10) using 3x3 window 171
3.20 $BVDF$ of (3.10) using 3x3 window 171
3.21 HF of (3.10) using 3x3 window 171
3.22 AHF of (3.10) using 3x3 window 171
3.23 $FVDF$ of (3.10) using 3x3 window 171
3.24 $ANNMF$ of (3.10) using 3x3 window 171
3.25 $CANNMF$ of (3.10) using 3x3 window 171
3.26 $BFMA$ of (3.10) using 3x3 window 171
3.27 'Mandrill' - 10% impulsive noise 173
3.28 NOP-NCP filtering results 173
3.29 VMF using 3x3 window 173
3.30 Mutistage Close-opening filtering results 173

4.1 Edge detection by derivative operators 180
4.2 Sub-window Configurations 195
4.3 Test color image 'ellipse' 202
4.4 Test color image 'flower' 202
4.5 Test color image 'Lenna' 202
4.6 Edge map of 'ellipse': Sobel detector 203
4.7 Edge map of 'ellipse': VR detector 203
4.8 Edge map of 'ellipse': DV detector 203
4.9 Edge map of 'ellipse': DV_hv detector 203
4.10 Edge map of 'flower': Sobel detector 204
4.11 Edge map of 'flower': VR detector 204
4.12 Edge map of 'flower': DV detector 204
4.13 Edge map of 'flower': DVadap detector 204
4.14 Edge map of 'Lenna': Sobel detector 205
4.15 Edge map of 'Lenna': VR detector 205
4.16 Edge map of 'Lenna': DV detector 205
4.17 Edge map of 'Lenna': DVadap detector 205

5.1 The original color image 'mountain' 215
5.2 The histogram equalized color output 215

6.1 Partitioned image 250
6.2 Corresponding quad-tree 250
6.3 The HSI cone with achromatic region in yellow 261
6.4 Original image. Achromatic pixels: intensity < 10, > 90 262
6.5 Saturation < 5 262
6.6 Saturation < 10 262
6.7 Saturation< 15 262
6.8 Original image. Achromatic pixels: saturation < 10, intensity> 90 263
6.9 Intensity < 5 263
6.10 Intensity < 10 263

6.11 Intensity < 15 263
6.12 Original image. Achromatic pixels: saturation< 10, intensity< 10 . 264
6.13 Intensity > 85 264
6.14 Intensity > 90 264
6.15 Intensity > 95 264
6.16 Original image 265
6.17 Pixel classification with chromatic pixels in red and achromatic pixels in the original color 265
6.18 Original image 265
6.19 Pixel classification with chromatic pixels in tan and achromatic pixels in the original color 265
6.20 Artificial image with level 1, 2, and 3 seeds. 266
6.21 The region growing algorithm. 267
6.22 Original 'Claire' image 270
6.23 'Claire' image showing seeds with $VAR = 0.2$ 270
6.24 Segmented 'Claire' image (before merging), $T_{chrom} = 0.15$ 270
6.25 Segmented 'Claire' image (after merging), $T_{chrom} = 0.15$ and $T_{merge} = 0.2$ 270
6.26 Original 'Carphone' image 271
6.27 'Carphone' image showing seeds with $VAR = 0.2$ 271
6.28 Segmented 'Carphone' image (before merging), $T_{chrom} = 0.15$ 271
6.29 Segmented 'Carphone' image (after merging), $T_{chrom} = 0.15$ and $T_{merge} = 0.2$ 271
6.30 Original 'Mother-Daughter' image 272
6.31 'Mother-Daughter' image showing seeds with $VAR = 0.2$ 272
6.32 Segmented 'Mother-Daughter' image (before merging), $T_{chrom} = 0.15$ 272
6.33 Segmented 'Mother-Daughter' image (after merging), $T_{chrom} = 0.15$ and $T_{merge} = 0.2$ 272

7.1 The zig-zag scan 297
7.2 DCT based coding 298
7.3 Original color image 'Peppers' 299
7.4 Image coded at a compression ratio $5:1$ 299
7.5 Image coded at a compression ratio $6:1$ 299
7.6 Image coded at a compression ratio $6.3:1$ 299
7.7 Image coded at a compression ratio $6.35:1$ 299
7.8 Image coded at a compression ratio $6.75:1$ 299
7.9 Subband coding scheme 301
7.10 Relationship between different scale subspaces 302
7.11 Multiresolution analysis decomposition 303
7.12 The wavelet-based scheme 304
7.13 Second generation coding schemes 304
7.14 The human visual system 307
7.15 Overall operation of the processing module 318

7.16 MPEG-1: Coding module 322
7.17 MPEG-1: Decoding module 322

8.1 Skin and Lip Clusters in the RGB color space 333
8.2 Skin and Lip Clusters in the L*a*b* color space 333
8.3 Skin and Lip hue Distributions in the HSV color space 334
8.4 Overall scheme to extract the facial regions within a scene 337
8.5 Template for hair color classification = R_1 + R_2 + R_3 342
8.6 Carphone: Frame 80 344
8.7 Segmented frame 344
8.8 Frames 20-95 344
8.9 Miss America: Frame 20 345
8.10 Frames 20-120 345
8.11 Akiyo: Frame 20 345
8.12 Frames 20-110 345

A.1 Screenshot of the main CIPAView window at startup. 350
A.2 Screenshot of Difference Vector Mean edge detector being applied 351
A.3 Gray scale image quantized to 4 levels 352
A.4 Screenshot of an image being corrupted by Impulsive Noise. 352

List of Tables

1.1 EBU Tech 3213 Primaries 12
1.2 EBU Tech 3213 Primaries 13
1.3 Color Model 46

2.1 Computational Complexity 100

3.1 Noise Distributions 158
3.2 Filters Compared 159
3.3 Subjective Image Evaluation Guidelines 161
3.4 Figure of Merit 162
3.5 NMSE($x10^{-2}$) for the RGB 'Lenna' image, 3×3 window 164
3.6 NMSE(x10^{-2}) for the RGB 'Lenna' image, 5×5 window 165
3.7 NMSE(x10^{-2}) for the RGB 'peppers' image, 3×3 window 165
3.8 NMSE(x10^{-2}) for the RGB 'peppers' image, 5×5 window 166
3.9 NCD for the RGB 'Lenna' image, 3×3 window 166
3.10 NCD for the RGB 'Lenna' image, 5×5 window 167
3.11 NCD for the RGB 'peppers' image, 3×3 window 167
3.12 NCD for the RGB 'peppers' image, 5×5 window 168
3.13 Subjective Evaluation 168
3.14 Performance measures for the image Mandrill 172

4.1 Vector Order Statistic Operators 198
4.2 Difference Vector Operators 199
4.3 Numerical Evaluation with Synthetic Images 199
4.4 Noise Performance 201

6.1 Comparison of Chromatic Distance Measures 269
6.2 Color Image Segmentation Techniques 273

7.1 Storage requirements 280
7.2 A taxonomy of image compression methodologies: First Generation 283
7.3 A taxonomy of image compression methodologies: Second Generation 283
7.4 Quantization table for the luminance component 296
7.5 Quantization table for the chrominance components 296

7.6 The JPEG suggested quantization table 312
7.7 Quantization matrix based on the contrast sensitivity function for 1.0 min/pixel 312

8.1 Miss America (Width×Height=360×288):Shape & Color Analysis. 343

1. Color Spaces

1.1 Basics of Color Vision

Color is a sensation created in response to excitation of our visual system by electromagnetic radiation known as light [1], [2], [3]. More specific, color is the perceptual result of light in the visible region of the electromagnetic spectrum, having wavelengths in the region of $400nm$ to $700nm$, incident upon the retina of the human eye. Physical power or radiance of the incident light is in a spectral power distribution (SPD), often divided into 31 components each representing a $10nm$ band [4]-[13].

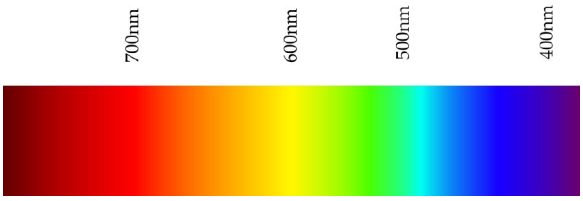


Fig. 1.1. The visible light spectrum

The human retina has three types of color photo-receptor cells, called cones , which respond to radiation with somewhat different spectral response curves [4]-[5]. A fourth type of photo-receptor cells, called roads , are also present in the retina. These are effective only at extremely low light levels, for example during night vision. Although rods are important for vision, they play no role in image reproduction [14], [15].

The branch of color science concerned with the appropriate description and specification of a color is called colorimetry [5], [10]. Since there are exactly three types of color photo-receptor cone cells, three numerical components are necessary and sufficient to describe a color, providing that appropriate spectral weighting functions are used. Therefore, a color can be specified by a tri-component vector. The set of all colors form a vector space called color space or color model. The three components of a color can be defined in many different ways leading to various color spaces [5], [9].

Before proceeding with color specification systems (color spaces), it is appropriate to define a few terms: Intensity (usually denoted I), brightness

(Br), luminance (Y), lightness (L^*), hue (H) and saturation (S), which are often confused or misused in the literature. The intensity (I) is a measure, over some interval of the electromagnetic spectrum, of the flow of power that is radiated from, or incident on a surface and expressed in units of watts per square meter [4], [18], [16]. The intensity (I) is often called a linear light measure and thus is expressed in units, such as watts per square meter [4], [5]. The brightness (Br) is defined as the attribute of a visual sensation according to which an area appears to emit more or less light [5]. Since brightness perception is very complex, the Commission Internationale de L'Eclairage (CIE) defined another quantity luminance (Y) which is radiant power weighted by a spectral sensitivity function that is characteristic of human vision [5]. Human vision has a nonlinear perceptual response to luminance which is called lightness (L^*). The nonlinearity is roughly logarithmic [4].

Humans interpret a color based on its lightness (L^*), hue (H) and saturation (S) [5]. Hue is a color attribute associated with the dominant wavelength in a mixture of light waves. Thus hue represents the dominant color as perceived by an observer; when an object is said to be red, orange, or yellow the hue is being specified. In other words, it is the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors: red, yellow, green and blue, or a combination of two of them [4], [5]. Saturation refers to the relative purity or the amount of white light mixed with a hue. The pure spectrum colors are fully saturated and contain no white light. Colors such as pink (red and white) and lavender (violet and white) are less saturated, with the degree of saturation being inversely proportional to the amount of white light added [1]. A color can be de-saturated by adding white light that contains power at all wavelengths [4]. Hue and saturation together describe the chrominance. The perception of color is basically determined by luminance and chrominance [1].

To utilize color as a visual cue in multimedia, image processing, graphics and computer vision applications, an appropriate method for representing the color signal is needed. The different color specification systems or color models (color spaces or solids) address this need. Color spaces provide a rational method to specify, order, manipulate and effectively display the object colors taken into consideration. A well chosen representation preserves essential information and provides insight to the visual operation needed. Thus, the selected color model should be well suited to address the problem's statement and solution. The process of selecting the best color representation involves knowing how color signals are generated and what information is needed from these signals. Although color spaces impose constraints on color perception and representation they also help humans perform important tasks. In particular, the color models may be used to define colors, discriminate between colors, judge similarity between color and identify color categories for a number of applications [12], [13].

Color model literature can be found in the domain of modern sciences, such as physics, engineering, artificial intelligence, computer science, psychology and philosophy. In the literature four basic color model families can be distinguished [14]:

1. **Colorimetric color models**, which are based on physical measurements of spectral reflectance. Three primary color filters and a photo-meter, such as the CIE *chromaticity diagram* usually serve as the initial points for such models.
2. **Psychophysical color models**, which are based on the human perception of color. Such models are either based on subjective observation criteria and comparative references (e.g. Munsell color model) or are built through experimentation to comply with the human perception of color (e.g. Hue, Saturation and Lightness model).
3. **Physiologically inspired color models**, which are based on the three *primaries*, the three types of *cones* in the human *retina*. The Red-Green-Blue (RGB) color space used in computer hardware is the best known example of a physiologically inspired color model.
4. **Opponent color models**, which are based on perception experiments, utilizing mainly pairwise opponent primary colors, such as the Yellow-Blue and Red-Green color pairs.

In image processing applications, color models can alternatively be divided into three categories. Namely:

1. **Device-oriented color models**, which are associated with input, processing and output signal devices. Such spaces are of paramount importance in modern applications, where there is a need to specify color in a way that is compatible with the hardware tools used to provide, manipulate or receive the color signals.
2. **User-oriented color models**, which are utilized as a bridge between the human operators and the hardware used to manipulate the color information. Such models allow the user to specify color in terms of perceptual attributes and they can be considered an experimental approximation of the human perception of color.
3. **Device-independent color models**, which are used to specify color signals independently of the characteristics of a given device or application. Such models are of importance in applications, where color comparisons and transmission of visual information over networks connecting different hardware platforms are required.

In 1931, the Commission Internationale de L'Eclairage (CIE) adopted standard color curves for a hypothetical standard observer. These color curves specify how a specific spectral power distribution (SPD) of an *external stimulus* (visible radiant light incident on the eye) can be transformed into a set of three numbers that specify the color. The CIE color specification system

is based on the description of color as the luminance component Y and two additional components X and Z [5]. The spectral weighting curves of X and Z have been standardized by the CIE based on statistics from experiments involving human observers [5]. The CIE XYZ *tristimulus values* can be used to describe any color. The corresponding color space is called the CIE XYZ color space. The XYZ model is a device independent color space that is useful in applications where consistent color representation across devices with different characteristics is important. Thus, it is exceptionally useful for color management purposes.

The CIE XYZ space is perceptually highly non uniform [4]. Therefore, it is not appropriate for quantitative manipulations involving color perception and is seldom used in image processing applications [4], [10]. Traditionally, color images have been specified by the non-linear red (R'), green (G') and blue (B') tristimulus values where color image storage, processing and analysis is done in this non-linear RGB ($R'G'B'$) color space. The red, green and blue components are called the primary colors . In general, hardware devices such as video cameras, color image scanners and computer monitors process the color information based on these primary colors. Other popular color spaces in image processing are the YIQ (North American TV standard), the HSI (Hue, Saturation and Intensity), and the HSV (Hue, Saturation, Value) color spaces used in computer graphics.

Although XYZ is used only indirectly it has a significant role in image processing since other color spaces can be derived from it through mathematical transforms. For example, the linear RGB color space can be transformed to and from the CIE XYZ color space using a simple linear three-by-three matrix transform. Similarly, other color spaces, such as non-linear RGB, YIQ and HSI can be transformed to and from the CIE XYZ space, but might require complex and non-linear computations. The CIE have also derived and standardized two other color spaces, called L*u*v* and L*a*b*, from the CIE XYZ color space which are perceptually uniform [5].

The rest of this chapter is devoted to the analysis of the different color spaces in use today. The different color representation models are discussed and analyzed in detail with emphasis placed on motivation and design characteristics.

1.2 The CIE Chromaticity-based Models

Over the years, the CIE committee has sponsored the research of color perception. This has lead to a class of widely used mathematical color models. The derivation of these models has been based on a number of color matching experiments, where an observer judges whether two parts of a visual stimulus match in appearance. Since the colorimetry experiments are based on a matching procedure in which the human observer judges the visual similarity of two areas the theoretical model predicts only matching and not perceived

colors. Through these experiments it was found that light of almost any spectral composition can be matched by mixtures of only three primaries (lights of a single wavelength). The CIE had defined a number of standard observer color matching functions by compiling experiments with different observers, different light sources and with various power and spectral compositions. Based on the experiments performed by CIE early in this century, it was determined that these three primary colors can be broadly chosen, provided that they are independent.

The CIE's experimental matching laws allow for the representation of colors as vectors in a three-dimensional space defined by the three primary colors. In this way, changes between color spaces can be accomplished easily. The next few paragraphs will briefly outline how such a task can be accomplished.

According to experiments conducted by Thomas Young in the nineteenth century [19], and later validated by other researchers [20], there are three different types of cones in the human retina, each with different absorption spectra: $S_1(\lambda)$, $S_2(\lambda)$, $S_3(\lambda)$, where $380 \leq \lambda \leq 780$ (nm). These approximately peak in the yellow-green, green and blue regions of the electromagnetic spectrum with significant overlap between S_1 and S_2. For each wavelength the absorption spectra provides the weight with which light of a given spectral distribution (SPD) contributes to the cone's output. Based on Young's theory, the color sensation that is produced by a light having SPD $C(\lambda)$ can be defined as:

$$\alpha_i(C) = \int_{\lambda_1}^{\lambda_2} S_i(\lambda) C(\lambda)\, d\lambda \tag{1.1}$$

for $i = 1, 2, 3$. According to (1.1) any two colors $C_1(\lambda)$, $C_2(\lambda)$ such that $\alpha_i(C_1) = \alpha_i(C_2)$, $i = 1, 2, 3$ will be perceived to be identical even if $C_1(\lambda)$ and $C_2(\lambda)$ are different. This well known phenomenon of spectrally different stimuli that are indistinguishable to a human observer is called *metamers* [14] and constitutes a rather dramatic illustration of the perceptual nature of color and the limitations of the color modeling process. Assume that three primary colors C_k , $k = 1, 2, 3$ with SPD $C_k(\lambda)$ are available and let

$$\int C_k(\lambda)\, d\lambda = 1 \tag{1.2}$$

To match a color C with spectral energy distribution $C(\lambda)$, the three primaries are mixed in proportions of β_k, $k = 1, 2, 3$. Their linear combination $\sum_{k=1}^{3} \beta_k C_k(\lambda)$ should be perceived as $C(\lambda)$. Substituting this into (1.1) leads to:

$$\alpha_i(C) = \int (\sum_{k=1}^{3} \beta_k C_k(\lambda)) S_i(\lambda)\, d\lambda = \sum_{k=1}^{3} \beta_k \int S_i(\lambda) C_k(\lambda)\, d\lambda \tag{1.3}$$

for $i = 1, 2, 3$.

The quantity $\int S_i(\lambda)C_k(\lambda)\,d\lambda$ can be interpreted as the i^{th}, $i = 1,2,3$ cone response generated by one unit of the k^{th} primary color:

$$\alpha_{i,k} = \alpha_i(C_k) = \int S_i(\lambda)C_k(\lambda)\,d\lambda \tag{1.4}$$

Therefore, the color matching equations are:

$$\sum_{k=1}^{3} \beta_k \alpha_{i,k} = \alpha_i(C) = \int S_i(\lambda)C(\lambda)\,d\lambda \tag{1.5}$$

assuming a certain set of primary colors $C_k(\lambda)$ and spectral sensitivity curves $S_i(\lambda)$. For a given arbitrary color, β_k can be found by simply solving (1.4) and (1.5).

Following the same approach w_k can be defined as the amount of the k^{th} primary required to match the reference white, providing that there is available a reference white light source with known energy distribution $w(\lambda)$.

In such a case, the values obtained through

$$T_k(C) = \frac{\beta_k}{w_k} \tag{1.6}$$

for $k = 1,2,3$ are called *tristimulus* values of the color C, and determine the relative amounts of primitives required to match that color. The tristimulus values of any given color $C(\lambda)$ can be obtained given the spectral tristimulus values $T_k(\lambda)$, which are defined as the tristimulus values of unit energy spectral color at wavelength λ. The spectral tristimulus $T_k(\lambda)$ provide the so-called *spectral matching curves* which are obtained by setting $C(\lambda) = \delta(\lambda - \lambda^*)$ in (1.5).

The spectral matching curves for a particular choice of color primaries with an approximately red, green and blue appearance were defined in the CIE 1931 standard [9]. A set of pure monochromatic primaries are used, blue ($435.8nm$), green ($546.1nm$) and red ($700nm$). In Figures 1.2 and 1.3 the Y-axis indicates the relative amount of each primary needed to match a stimulus of the wavelength reported on the X-axis. It can be seen that some of the values are negative. Negative numbers require that the primary in question be added to the opposite side of the original stimulus. Since negative sources are not physically realizable it can be concluded that the arbitrary set of three primary sources cannot match all the visible colors. However, for any given color a suitable set of three primary colors can be found.

Based on the assumption that the human visual system behaves linearly, the CIE had defined spectral matching curves in terms of virtual primaries. This constitutes a linear transformation such that the spectral matching curves are all positive and thus immediately applicable for a range of practical situations. The end results are referred to as the CIE 1931 standard observer matching curves and the individual curves (functions) are labeled

$\bar{x}$, $\bar{y}$, $\bar{z}$ respectively. In the CIE 1931 standard the matching curves were selected so that $\bar{y}$ was proportional to the human luminosity function, which was an experimentally determined measure of the perceived brightness of monochromatic light.

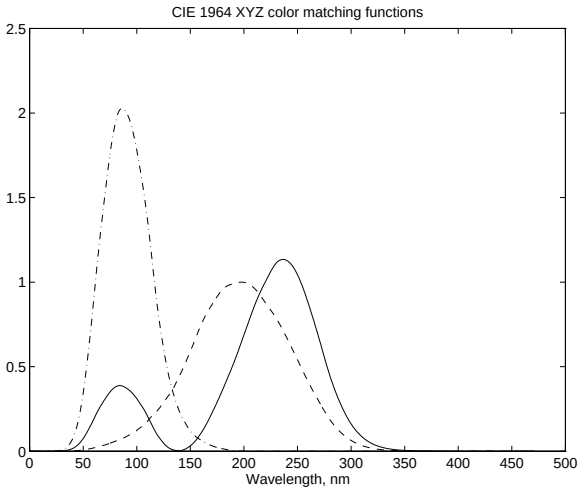


Fig. 1.2. The CIE XYZ color matching functions

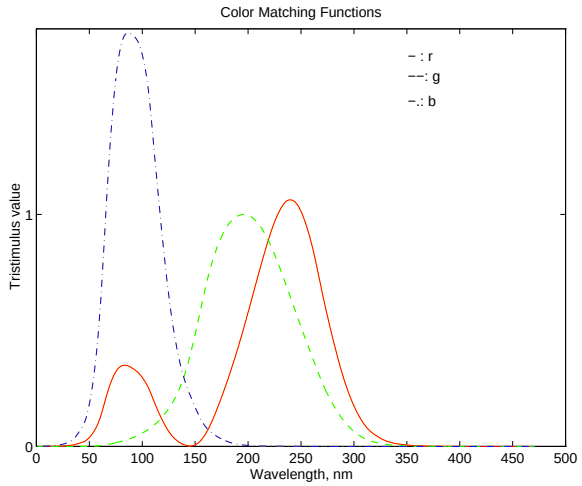


Fig. 1.3. The CIE RGB color matching functions

If the spectral energy distribution $C(\lambda)$ of a stimulus is given, then the chromaticity coordinates can be determined in two stages. First, the tristimulus values X, Y, Z are calculated as follows:

$$X = \int \bar{x}(\lambda) C(\lambda)\, d\lambda \tag{1.7}$$

$$Y = \int \bar{y}(\lambda) C(\lambda)\, d\lambda \tag{1.8}$$

$$Z = \int \bar{z}(\lambda) C(\lambda)\, d\lambda \tag{1.9}$$

The new set of primaries must satisfy the following conditions:

1. The XYZ components for all visible colors should be non-negative.
2. Two of the primaries should have zero luminance.
3. As many spectral colors as possible should have at least one zero XYZ component.

Secondly, normalized tristimulus values, called chromaticity coordinates, are calculated based on the primaries as follows:

$$x = \frac{X}{X+Y+Z} \tag{1.10}$$

$$y = \frac{Y}{X+Y+Z} \tag{1.11}$$

$$z = \frac{Z}{X+Y+Z} \tag{1.12}$$

Clearly $z = 1 - (x + y)$ and hence only two coordinates are necessary to describe a color match. Therefore, the chromaticity coordinates project the $3-D$ color solid on a plane, and they are usually plotted as a parametric $x-y$ plot with z implicitly evaluated as $z = 1 - (x + y)$. This diagram is known as the chromaticity diagram and has a number of interesting properties that are used extensively in image processing. In particular,

1. The chromaticity coordinates (x, y) jointly represent the chrominance components of a given color.
2. The entire color space can be represented by the coordinates (x, y, T), in which $T = constant$ is a given chrominance plane.
3. The chromaticity diagram represents every physically realizable color as a point within a well defined boundary. The boundary represents the primary sources. The boundary vertices have coordinates defined by the chromaticities of the primaries.
4. A white point is located in the center of the chromaticity diagram. More saturated colors radiate outwards from white. Complementary pure colors can easily be determined from the diagram.
5. In the chromaticity diagram, the color perception obtained through the superposition of light coming from two different sources, lies on a straight line between the points representing the component lights in the diagram.

6. Since the chromaticity diagram reveals the range of all colors which can be produced by means of the three primaries (gamut), it can be used to guide the selection of primaries subject to design constraints and technical specifications.
7. The chromaticity diagram can be utilized to determine the hue and saturation of a given color since it represents chrominance by eliminating luminance. Based on the initial objectives set out by CIE, two of the primaries, X and Z, have zero luminance while the primary Y is the luminance indicator determined by the light-efficiency function $V(\lambda)$ at the spectral matching curve $\bar{y}$. Thus, in the chromaticity diagram the dominant wavelength (hue) can be defined as the intersection between a line drawn from the reference white through the given color to the boundaries of the diagram. Once the hue has been determined, then the purity of a given color can be found as the ratio $r = \frac{wc}{wp}$ of the line segments that connect the reference white with the color (wc) to the line segment between the reference white and the dominant wavelength/hue (wp).

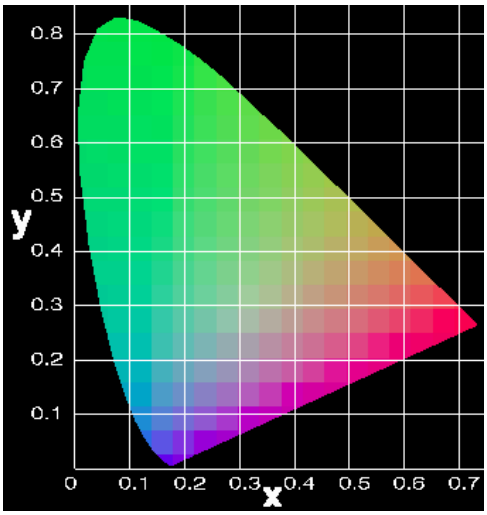


Fig. 1.4. The chromaticity diagram

1.3 The CIE-RGB Color Model

The fundamental assumption behind modern colorimetry theory, as it applies to image processing tasks, is that the initial basis for color vision lies in the different excitation of three classes of photo-receptor cones in the retina. These include the red, green and blue receptors, which define a trichromatic

space whose basis of primaries are pure colors in the short, medium and high portions of the visible spectrum [4], [5], [10].

As a result of the assumed linear nature of light, and due to the principle of superposition, the colors of a mixture are a function of the primaries and the fraction of each primary that is mixed. Throughout this analysis, the primaries need not be known, just their tristimulus values. This principle is called additive reproduction. It is employed in image and video devices used today where the color spectra from red, green and blue light beams are physically summed at the surface of the projection screen. Direct view color CRT's (cathode ray tube) also utilize additive reproduction. In particular, the CRT's screen consists of small dots which produce red, green and blue light. When the screen is viewed from a distance the spectra of these dots add up in the retina of the observer. In practice, it is possible to reproduce a large number of colors by additive reproduction using the three primaries: red, green and blue. The colors that result from additive reproduction are completely determined by the three primaries.

The video projectors and the color CRT's in use today utilize a color space collectively known under the name RGB, which is based on the red, green and blue primaries and a white reference point. To uniquely specify a color space based on the three primary colors the chromaticity values of each primary color and a white reference point need to be specified. The gamut of colors which can be mixed from the set of the RGB primaries is given in the (x, y) chromaticity diagram by a triangle whose vertices are the chromaticities of the primaries (Maxwell triangle) [5], [20]. This is shown in Figure 1.5.

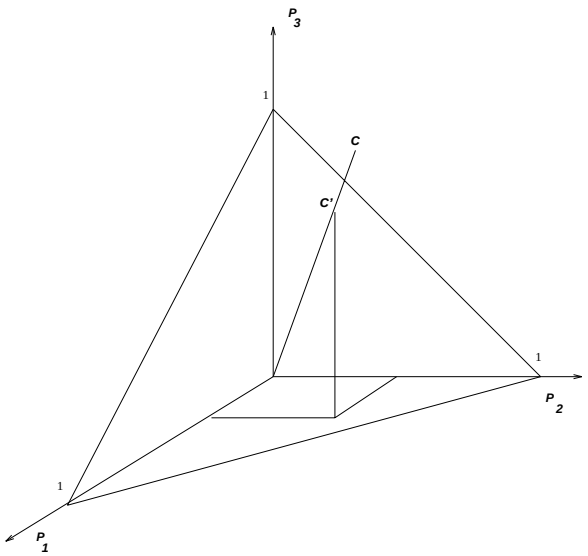


Fig. 1.5. The Maxwell triangle

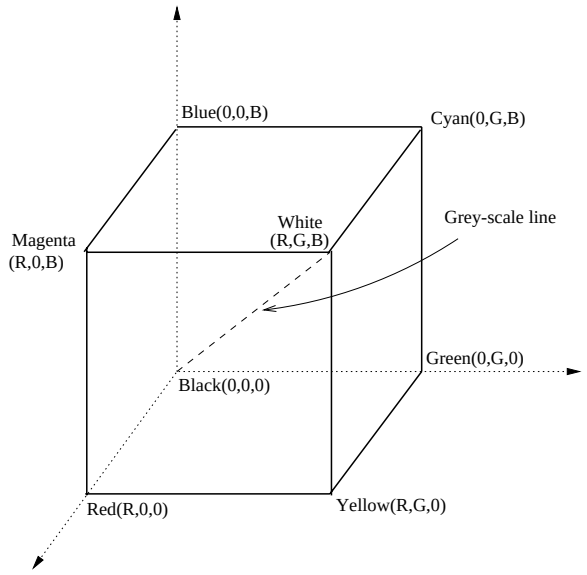


Fig. 1.6. The RGB color model

In the red, green and blue system the color solid generated is a bounded subset of the space generated by each primary. Using an appropriate scale along each primary axis, the space can normalized, so that the maximum is 1. Therefore, as can be seen in Figure 1.6 the RGB color solid is a cube, called the RGB cube. The origin of the cube, defined as $(0, 0, 0)$ corresponds to black and the point with coordinates $(1, 1, 1)$ corresponds to the system's brightest white.

In image processing, computer graphics and multimedia systems the RGB representation is the most often used. A digital color image is represented by a two dimensional array of three variate vectors which are comprised of the pixel's red, green and blue values. However, these pixel values are relative to the three primary colors which form the color space. As it was mentioned earlier, to uniquely define a color space, the chromaticities of the three primary colors and the reference white must be specified. If these are not specified within the chromaticity diagram, the pixel values which are used in the digital representation of the color image are meaningless [16].

In practice, although a number of RGB space variants have been defined and are in use today, their exact specifications are usually not available to the end-user. Multimedia users assume that all digital images are represented in the same RGB space and thus use, compare or manipulate them directly no matter where these images are from. If a color digital image is represented in the RGB system and no information about its chromaticity characteristics is available, the user cannot accurately reproduce or manipulate the image.

Although in computing and multimedia systems there are no standard primaries or white point chromaticities, a number of color space standards

have been defined and used in the television industry. Among them are the Federal Communication Commission of America (FCC) 1953 primaries, the Society of Motion Picture and Television Engineers (SMPTE) 'C' primaries, the European Broadcasting Union (EBU) primaries and the ITU-R BT.709 standard (formerly known as CCIR Rec. 709) [24]. Most of these standards use a white reference point known as CIE D65 but other reference points, such as the CIE illuminant E are also be used [4].

In additive color mixtures the white point is defined as the one with equal red, green and blue components. However, there is no unique physical or perceptual definition of white, so the characteristics of the white reference point should be defined prior to its utilization in the color space definition.

In the CIE illuminant E, or equal-energy illuminant, white is defined as the point whose spectral power distribution is uniform throughout the visible spectrum. A more realistic reference white, which approximates daylight has been specified numerically by the CIE as illuminant D65. The D65 reference white is the one most often used for color interchange and the reference point used throughout this work.

The appropriate red, green and blue chromaticities are determined by the technology employed, such as the sensors in the cameras, the phosphors within the CTR's and the illuminants used. The standards are an attempt to quantify the industry's practice. For example, in the FCC-NTSC standard, the set of primaries and specified white reference point were representative of the phosphors used in color CRTs of a certain era.

Although the sensor technology has changed over the years in response to market demands for brighter television receivers, the standards remain the same. To alleviate this problem, the European Broadcasting Union (EBU) has established a new standard (EBU Tech 3213). It is defined in Table 1.1.

Table 1.1. EBU Tech 3213 Primaries

Colorimetry	Red	Green	Blue	White D65
x	0.640	0.290	0.150	0.3127
y	0.330	0.600	0.060	0.3290
z	0.030	0.110	0.790	0.3582

Recently, an international agreement has finally been reached on the primaries for the High Definition Television (HDTV) specification. These primaries are representative of contemporary monitors in computing, computer graphics and studio video production. The standard is known as ITU-R BT.709 and its primaries along with the D65 reference white are defined in Table 1.2.

The different RGB systems can be converted amongst each other using a linear transformation assuming that the white references values being used are known. As an example, if it is assumed that the D65 is used in both

Table 1.2. EBU Tech 3213 Primaries

Colorimetry	Red	Green	Blue	White D65
x	0.640	0.300	0.150	0.3127
y	0.330	0.600	0.060	0.3290
z	0.030	0.100	0.790	0.3582

systems, then the conversion between the ITU-R BT.709 and SMPTE 'C' primaries is defined by the following matrix transformation:

$$\begin{bmatrix} R_{709} \\ G_{709} \\ B_{709} \end{bmatrix} = \begin{bmatrix} 0.939555 & 0.050173 & 0.010272 \\ 0.017775 & 0.9655795 & 0.016430 \\ -0.001622 & -0.004371 & 1.005993 \end{bmatrix} \begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} \tag{1.13}$$

where R_{709}, G_{709}, B_{709} are the linear red, green and blue components of the ITU-R BT.709 and R_c, G_c, B_c are the linear components in the SMPTE 'C' system. The conversion should be carried out in the linear voltage domain, where the pixel values must first be converted into linear voltages. This is achieved by applying the so-called *gamma correction*.

1.4 Gamma Correction

In image processing, computer graphics, digital video and photography, the symbol γ represents a numerical parameter which describes the nonlinearity of the intensity reproduction. The cathode-ray tube (CRT) employed in modern computing systems is nonlinear in the sense that the intensity of light reproduced at the screen of a CRT monitor is a nonlinear function of the voltage input. A CRT has a power law response to applied voltage. The light intensity produced on the display is proportional to the applied voltage raised to a power denoted by γ [4], [16], [17]. Thus, the produced intensity by the CRT and the voltage applied on the CRT have the following relationship:

$$I_{int} = (v')^{\gamma} \tag{1.14}$$

The relationship which is called the 'five-halves' power law is dictated by the physics of the CRT electron gun. The above function applies to a single electron gun of a gray-scale CRT or each of the three red, green and blue electron guns of a color CRT. The functions associated with the three guns on a color CRT are very similar to each other but not necessarily identical. The actual value of γ for a particular CRT may range from about 2.3 to 2.6 although most practitioners frequently claim values lower than 2.2 for video monitors.

The process of pre-computing for the nonlinearity by computing a voltage signal from an intensity value is called *gamma correction*. The function required is approximately a 0.45 power function. In image processing applications, gamma correction is accomplished by analog circuits at the camera.

In computer graphics, gamma correction is usually accomplished by incorporating the function into a frame buffer lookup table. Although in image processing systems gamma was originally used to refer to the nonlinearity of the CRT, it is generalized to refer to the nonlinearity of an entire image processing system. The γ value of an image or an image processing system can be calculated by multiplying the γ's of its individual components from the image capture stage to the display.

The model used in (1.14) can cause wide variability in the value of gamma mainly due to the black level errors since it forces the zero voltage to map to zero intensity for any value of gamma. A slightly different model can be used in order to resolve the black level error. The modified model is given as:

$$I_{int} = (voltage + \epsilon)^{2.5} \tag{1.15}$$

By fixing the exponent of the power function at 2.5 and using the single parameter to accommodate black level errors the modified model fits the observed nonlinearity much better than the variable gamma model in (1.14).

The voltage-to-intensity function defined in (1.15) is nearly the inverse of the luminance-to-brightness relationship of human vision. Human vision defines luminance as a weighted mixture of the spectral energy where the weights are determined by the characteristics of the human retina. The CIE has standardized a weighting function which relates spectral power to luminance. In this standardized function, the perceived luminance by humans relates to the physical luminance (proportional to intensity) by the following equation:

$$L^* = \begin{cases} 116(\frac{Y}{Y_n})^{\frac{1}{3}} - 16 \text{ if } \frac{Y}{Y_n} > 0.008856 \\ 903.3(\frac{Y}{Y_n})^{\frac{1}{3}} \quad \text{if } \frac{Y}{Y_n} \leq 0.008856 \end{cases} \tag{1.16}$$

where Y_n is the luminance of the reference white, usually normalized either to 1.0 or 100. Thus, the lightness perceived by humans is, approximately, the cubic root of the luminance. The lightness sensation can be computed as intensity raised, approximately to the third power. Thus, the entire image processing system can be considered linear or almost linear.

To compensate for the nonlinearity of the display (CRT), gamma correction with a power of $(\frac{1}{\gamma})$ can be used so that the overall system γ is approximately 1.

In a video system, the gamma correction is applied to the camera for precomputing the nonlinearity of the display. The gamma correction performs the following transfer function:

$$voltage' = (voltage)^{\frac{1}{\gamma}} \tag{1.17}$$

where *voltage* is the voltage generated by the camera sensors. The gamma corrected value is the reciprocal of the gamma resulting in a transfer function with unit power exponent.

To achieve subjectively pleasing images, the end-to-end power function of the overall imaging system should be around 1.1 or 1.2 instead of the mathematically correct linear system.

The REC 709 specifies a power exponent of 0.45 at the camera which, in conjunction with the 2.5 exponent at the display, results in an overall exponent value of about 1.13. If the γ value is greater than 1, the image appears sharper but the scene contrast range, which can be reproduced, is reduced. On the other hand, reducing the γ value has a tendency to make the image appear soft and washed out.

For color images, the linear values R, G, and B values should be converted into nonlinear voltages R', G' and B' through the application of the gamma correction process. The color CRT will then convert R', G' and B' into linear red, green and blue light to reproduce the original color.

The ITU-R BT.709 standard recommends a gamma exponent value of 0.45 for the High Definition Television. In practical systems, such as TV cameras, certain modifications are required to ensure proper operation near the dark regions of an image, where the slope of a pure power function is infinite at zero. The red tristimulus (linear light) component may be gamma-corrected at the camera by applying the following convention:

$$R'_{709} = \begin{cases} 4.5R & \text{if } R \leq 0.018 \\ 1.099R^{0.45} - 0.099 & \text{if } 0.018 < R \end{cases} \tag{1.18}$$

with R denoting the linear light and R'_{709} the resulting gamma corrected value. The computations are identical for the G and B components.

The linear R, G, and B are normally in the range $[0, 1]$ when color images are used in digital form. The software library translates these floating point values to 8-bit integers in the range of 0 to 255 for use by the graphics hardware. Thus, the gamma corrected value should be:

$$R' = 255R^{\frac{1}{\gamma}} \tag{1.19}$$

The constant 255 in (1.19) is added during the A/D process. However, gamma correction is usually performed in cameras, and thus, pixel values are in most cases nonlinear voltages. Thus, intensity values stored in the frame-buffer of the computing device are gamma corrected on-the-fly by hardware look up tables on their way to the computer monitor display. Modern image processing systems utilize a wide variety of sources of color images, such as images captured by digital cameras, scanned images, digitized video frames and computer generated images. Digitized video frames usually have a gamma correction value between 0.5 and 0.45. Digital scanners assume an output gamma in the range of 1.4 to 2.2 and they perform their gamma correction accordingly. For computer generated images the gamma correction value is usually unknown. In the absence of the actual gamma value the recommended gamma correction is 0.45.

In summary, pixel values alone cannot specify the actual color. The gamma correction value used for capturing or generating the color image is needed. Thus, two images which have been captured with two cameras operating under different gamma correction values will represent colors differently even if the same primaries and the same white reference point are used.

1.5 Linear and Non-linear RGB Color Spaces

The image processing literature rarely discriminates between linear RGB and non-linear (R′G′B′) gamma corrected values. For example, in the JPEG and MPEG standards and in image filtering, non-linear RGB (R′G′B′) color values are implicit. Unacceptable results are obtained when JPEG or MPEG schemes are applied to linear RGB image data [4]. On the other hand, in computer graphics, linear RGB values are implicitly used [4]. Therefore, it is very important to understand the difference between linear and non-linear RGB values and be aware of which values are used in an image processing application. Hereafter, the notation R′G′B′ will be used for non-linear RGB values so that they can be clearly distinguished from the linear RGB values.

1.5.1 Linear RGB Color Space

As mentioned earlier, intensity is a measure, over some interval of the electromagnetic spectrum, of the flow of power that is radiated from an object. Intensity is often called a linear light measure. The linear R value is proportional to the intensity of the physical power that is radiated from an object around the 700 nm band of the visible spectrum. Similarly, a linear G value corresponds to the 546.1 nm band and a linear B value corresponds to the 435.8 nm band. As a result the linear RGB space is device independent and used in some color management systems to achieve color consistency across diverse devices.

The linear RGB values in the range [0, 1] can be converted to the corresponding CIE XYZ values in the range [0, 1] using the following matrix transformation [4]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1.20}$$

The transformation from CIE XYZ values in the range [0, 1] to RGB values in the range [0, 1] is defined by:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2405 & -1.5372 & -0.4985 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0573 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1.21}$$

Alternatively, tristimulus XYZ values can be obtained from the linear RGB values through the following matrix [5]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.490\ 0.310\ 0.200 \\ 0.117\ 0.812\ 0.011 \\ 0.000\ 0.010\ 0.990 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1.22}$$

The linear RGB values are a physical representation of the chromatic light radiated from an object. However, the perceptual response of the human visual system to radiate red, green, and blue intensities is non-linear and more complex. The linear RGB space is, perceptually, highly non-uniform and not suitable for numerical analysis of the perceptual attributes. Thus, the linear RGB values are very rarely used to represent an image. On the contrary, non-linear R′G′B′ values are traditionally used in image processing applications such as filtering.

1.5.2 Non-linear RGB Color Space

When an image acquisition system, e.g. a video camera, is used to capture the image of an object, the camera is exposed to the linear light radiated from the object. The linear RGB intensities incident on the camera are transformed to non-linear RGB signals using gamma correction. The transformation to non-linear R′G′B′ values in the range [0, 1] from linear RGB values in the range [0, 1] is defined by:

$$\begin{aligned} R' &= \begin{cases} 4.5R, & \text{if } R \leq 0.018 \\ 1.099R^{\frac{1}{\gamma_C}} - 0.099, & \text{otherwise} \end{cases} \\ G' &= \begin{cases} 4.5G, & \text{if } G \leq 0.018 \\ 1.099G^{\frac{1}{\gamma_C}} - 0.099, & \text{otherwise} \end{cases} \\ B' &= \begin{cases} 4.5B, & \text{if } B \leq 0.018 \\ 1.099B^{\frac{1}{\gamma_C}} - 0.099, & \text{otherwise} \end{cases} \end{aligned} \tag{1.23}$$

where γ_C is known as the *gamma factor* of the camera or the acquisition device. The value of γ_C that is commonly used in video cameras is $\frac{1}{0.45}$ ($\simeq 2.22$) [4]. The above transformation is graphically depicted in Figure 1.7. The linear segment near low intensities minimizes the effect of sensor noise in practical cameras and scanners.

Thus, the digital values of the image pixels acquired from the object and stored within a camera or a scanner are the R′G′B′ values usually converted to the range of 0 to 255. Three bytes are then required to represent the three components, R', G', and B' of a color image pixel with one byte for each component. It is these non-linear R′G′B′ values that are stored as image data files in computers and are used in image processing applications. The RGB symbol used in image processing literature usually refers to the R′G′B′

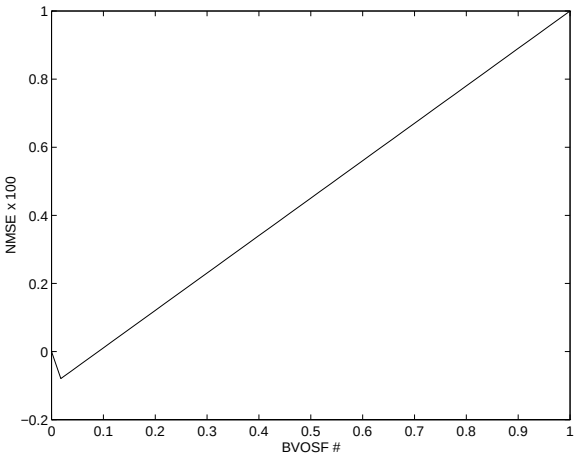


Fig. 1.7. Linear to Non-linear Light Transformation

values and, therefore, care must be taken in color space conversions and other relevant calculations.

Suppose the acquired image of an object needs to be displayed in a display device such as a computer monitor. Ideally, a user would like to see (perceive) the exact reproduction of the object. As pointed out, the image data is in R′G′B′ values. Signals (usually voltage) proportional to the R′G′B′ values will be applied to the red, green, and blue guns of the CRT (Cathode Ray Tube) respectively. The intensity of the red, green, and blue lights generated by the CRT is a non-linear function of the applied signal. The non-linearity of the CRT is a function of the electrostatics of the cathode and the grid of the electron gun. In order to achieve correct reproduction of intensities, an ideal monitor should invert the transformation at the acquisition device (camera) so that the intensities generated are identical to the linear RGB intensities that were radiated from the object and incident in the acquisition device. Only then will the perception of the displayed image be identical to the perceived object.

A conventional CRT has a power-law response, as depicted in Figure 1.8. This power-law response, which inverts the non-linear (R′G′B′) values in the range [0, 1] back to linear RGB values in the range [0, 1], is defined by the following power function [4]:

$$
\begin{aligned}
R &= \begin{cases} \frac{R'}{4.5}, & \text{if } R' \leq 0.018 \\ \left(\frac{R' + 0.099}{1.099}\right)^{\gamma_D} & \text{otherwise} \end{cases} \\
G &= \begin{cases} \frac{G'}{4.5}, & \text{if } G' \leq 0.018 \\ \left(\frac{G' + 0.099}{1.099}\right)^{\gamma_D} & \text{otherwise} \end{cases}
\end{aligned} \tag{1.24}
$$

$$B = \begin{cases} \frac{B'}{4.5}, & \text{if } B' \leq 0.018 \\ \left(\frac{B' + 0.099}{1.099}\right)^{\gamma_D} & \text{otherwise} \end{cases}$$

The value of the power function, γ_D, is known as the *gamma* factor of the display device or CRT. Normal display devices have γ_D in the range of 2.2 to 2.45. For exact reproduction of the intensities, *gamma* factor of the display device must be equal to the *gamma* factor of the acquisition device ($\gamma_C = \gamma_D$). Therefore, a CRT with a *gamma* factor of 2.2 should correctly reproduce the intensities.

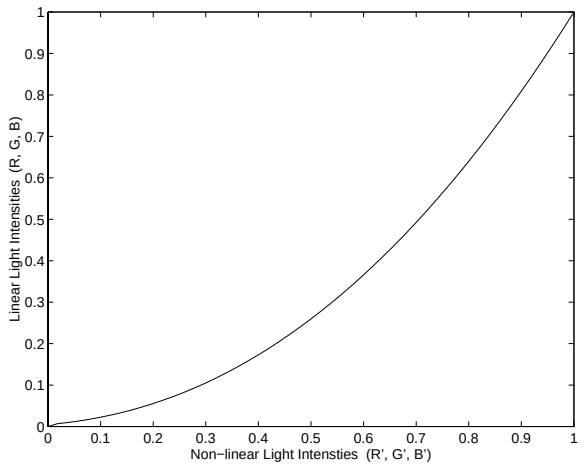


Fig. 1.8. Non-linear to linear Light Transformation

The transformations that take place throughout the process of image acquisition to image display and perception are illustrated in Figure 1.9.

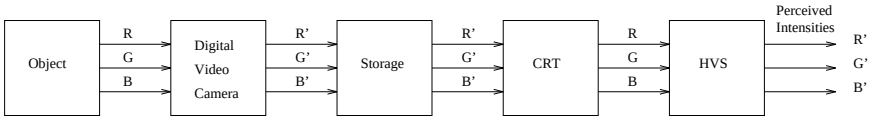


Fig. 1.9. Transformation of Intensities from Image Capture to Image Display

It is obvious from the above discussion that the R′G′B′ space is a device dependent space. Suppose a color image, represented in the R′G′B′ space, is displayed on two computer monitors having different *gamma* factors. The red, green, and blue intensities produced by the monitors will not be identical and the displayed images might have different appearances. Device dependent spaces cannot be used if color consistency across various devices, such as display devices, printers, etc., is of primary concern. However, similar devices

(e.g. two computer monitors) usually have similar *gamma* factors and in such cases device dependency might not be an important issue.

As mentioned before, the human visual system has a non-linear perceptual response to intensity, which is roughly logarithmic and is, approximately, the inverse of a conventional CRT's non-linearity [4]. In other words, the perceived red, green, and blue intensities are approximately related to the R′G′B′ values. Due to this fact, computations involving R′G′B′ values have an approximate relation to the human color perception and the R′G′B′ space is less perceptually non-uniform relative to the CIE XYZ and linear RGB spaces [4]. Hence, distance measures defined between the R′G′B′ values of two color vectors provide a computationally simple estimation of the error between them. This is very useful for real-time applications and systems in which computational resources are at premium.

However, the R′G′B′ space is not adequately uniform, and it cannot be used for accurate perceptual computations. In such instances, perceptually uniform color spaces (e.g. L*u*v* and L*a*b*) that are derived based on the attributes of human color perception are more desirable than the R′G′B′ space [4].

1.6 Color Spaces Linearly Related to the RGB

In transmitting color images through a computer-centric network, all three primaries should be transmitted. Thus, storage or transmission of a color image using RGB components requires a channel capacity three times that of gray scale images. To reduce these requirements and to boost bandwidth utilization, the properties of the human visual system must be taken into consideration. There is strong evidence that the human visual system forms an achromatic channel and two chromatic color-difference channels in the retina. Consequently, a color image can be represented as a wide band component corresponding to brightness, and two narrow band color components with considerably less data rate than that allocated to brightness.

Since the large percentage (around 60%) of brightness is attributed to the green primary, then it is advantageous to base the color components on the other two primaries. The simplest way to form the two color components is to remove them by subtraction, (e.g. the brightness from the blue and red primaries). In this way the unit RGB color cube is transformed into the luminance Y and two color difference components $B - Y$ and $R - Y$ [33], [34]. Once these color difference components have been formed, they can be sub-sampled to reduce the bandwidth or data capacity without any visible degradation in performance. The color difference components are calculated from nonlinear gamma corrected values R',G',B' rather than the tristimulus (linear voltage) R, G, B primary components.

According to the CIE standards the color imaging system should operate similarly to a gray scale system, with a CIE luminance component Y formed

as a weighted sum of RGB tristimulus values. The coefficients in the weighted sum correspond to the sensitivity of the human visual system to each of the RGB primaries. The coefficients are also a function of the chromaticity of the white reference point used. International agreement on the REC. 709 standard provides a value for the luminance component based on the REC. 709 primaries [24]. Thus, the Y'_{709} luminance equation is:

$$Y'_{709} = 0.2125R'_{709} + 0.7154G'_{709} + 0.0721B'_{709} \tag{1.25}$$

where R'_{709}, B'_{709} and G'_{709} are the gamma-corrected (nonlinear) values of the three primaries. The two color difference components $B'_{709} - Y'_{709}$ and $R'_{709} - Y'_{709}$ can be formed on the basis of the above equation.

Various scale factors are applied to the basic color difference components for different applications. For example, the $Y'P_RP_B$ is used for component analog video, such as BetaCam, and $Y'C_BC_R$ for component digital video, such as studio video, JPEG and MPEG. Kodak's YCC (PhotoCD model) uses scale factors optimized for the gamut of film colors [31]. All these systems utilize different versions of the $(Y'_{709}, B'_{709}-Y'_{709}, R'_{709}-Y'_{709})$ which are scaled to place the extrema of the component signals at more convenient values.

In particular, the $Y'P_RP_B$ system used in component analog equipment is defined by the following set:

$$\begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418686 & -0.081312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{1.26}$$

and

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1. & 0. & 1.402 \\ 1. & -0.344136 & -0.714136 \\ 1. & 1.772 & 0. \end{bmatrix} \begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} \tag{1.27}$$

The first row comprises the luminance coefficients which sum to unity. For each of the other two columns the coefficients sum to zero, a necessity for color difference formulas. The 0.5 weights reflect the maximum excursion of P_B and P_R for the blue and the red primaries.

The $\mathrm{Y'C}_B\mathrm{C}_R$ is the Rec ITU-R BT. 601-4 international standard for studio quality component digital video. The luminance signal is coded in 8 bits. The Y' has an excursion of 219 with an offset of 16, with the black point coded at 16 and the white at code 235. Color differences are also coded in 8-bit forms with excursions of 112 and offset of 128 for a range of 16 through 240 inclusive.

To compute $\mathrm{Y'C}_B\mathrm{C}_R$ from nonlinear $\mathrm{R'G'B'}$ in the range of $[0, 1]$ the following set should be used:

$$\begin{bmatrix} Y'_{601} \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.0 \\ 112.0 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{1.28}$$

with the inverse transform

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 0.00456821 & 0.0 & 0.00625893 \\ 0.00456621 & -0.00153632 & -0.00318811 \\ 0.00456621 & 0.00791071 & 0.0 \end{bmatrix} \cdot \left(\begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} - \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \right) \tag{1.29}$$

When 8-bit R′G′B′ are used, black is coded at 0 and white is at 255. To encode $\mathrm{Y'C}_B\mathrm{C}_R$ from R′G′B′ in the range of $[0, 255]$ using 8-bit binary arithmetic the transformation matrix should be scaled by $\frac{256}{255}$. The resulting transformation pair is as follows:

$$\begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \frac{1}{256} \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.0 \\ 112.0 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R'_{255} \\ G'_{255} \\ B'_{255} \end{bmatrix} \tag{1.30}$$

where R'_{255} is the gamma-corrected value, using a gamma-correction lookup table for $\frac{1}{\gamma}$. This yields the RGB intensity values with integer components between 0 and 255 which are gamma-corrected by the hardware. To obtain R′G′B′ values in the range $[0, 255]$ from $\mathrm{Y'C}_B\mathrm{C}_R$ using 8-bit arithmetic the following transformation should be used:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \frac{1}{256} \begin{bmatrix} 0.00456821 & 0.0 & 0.00625893 \\ 0.00456621 & -0.00153632 & -0.00318811 \\ 0.00456621 & 0.00791071 & 0.0 \end{bmatrix} \left(\begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} - \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \right) \tag{1.31}$$

Some of the coefficients when scaled by $\frac{1}{256}$ may be larger than unity and, thus some clipping may be required so that they fall within the acceptable RGB range.

The Kodak YCC color space is another example of a predistorted color space, which has been designed for the storage of still color images on the Photo-CD. It is derived from the predistorted (gamma-corrected) R′G′B′ values using the ITU-R BT.709 recommended white reference point, primaries, and gamma correction values. The YCC space is similar to the $\mathrm{Y'C}_B\mathrm{C}_R$ discussed, although scaling of $B' - Y'$ and $R' - Y'$ is asymmetrical in order to accommodate a wide color gamut, similar to that of a photographic film. In particular the following relationship holds for Photo-CD compressed formats:

$$Y' = \frac{255}{1.402} Y'_{601} \tag{1.32}$$

$$C_1 = 156 + 111.40(B' - Y') \tag{1.33}$$

$$C_2 = 137 + 135.64(R' - Y') \tag{1.34}$$

The two chrominance components are compressed by factors of 2 both horizontally and vertically. To reproduce predistorted R′G′B′ values in the range of $[0,1]$ from integer PhotoYCC components the following transform is applied:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \frac{1}{256} \begin{bmatrix} 0.00549804 & 0.0 & 0.0051681 \\ 0.00549804 & -0.0015446 & -0.0026325 \\ 0.00549804 & 0.0079533 & 0.0 \end{bmatrix} \left(\begin{bmatrix} Y' \\ C_1 \\ C_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 156 \\ 137 \end{bmatrix} \right) \tag{1.35}$$

The $B' - Y'$ and $R' - Y'$ components can be converted into polar coordinates to represent the perceptual attributes of hue and saturation. The values can be computed using the following formulas [34]:

$$H = tan^{-1}(\frac{B' - Y'}{R' - Y'}) \tag{1.36}$$

$$S = ((B' - Y')^2 + (R' - Y')^2)\frac{1}{2} \tag{1.37}$$

where the saturation S is the length of the vector from the origin of the chromatic plane to the specific color and the hue H is the angle between the $R' - Y'$ axis and the saturation component [33].

1.7 The YIQ Color Space

The YIQ color specification system, used in commercial color TV broadcasting and video systems, is based upon the color television standard that was adopted in the 1950s by the National Television Standard committee (NTSC) [10], [1], [27], [28]. Basically, YIQ is a recoding of non-linear R′G′B′ for transmission efficiency and for maintaining compatibility with monochrome TV standards [1], [4]. In fact, the Y component of the YIQ system provides all the video information required by a monochrome television system.

The YIQ model was designed to take advantage of the human visual system's greater sensitivity to change in luminance than to changes in hue or saturation [1]. Due to these characteristics of the human visual system, it is useful in a video system to specify a color with a component representative of luminance Y and two other components: the in-phase I, an orange-cyan axis, and the quadrature Q component, the magenta-green axis. The two chrominance components are used to jointly represent hue and saturation .

With this model, it is possible to convey the component representative of luminance Y in such a way that noise (or quantization) introduced in transmission, processing and storage is minimal and has a perceptually similar effect across the entire tone scale from black to white [4]. This is done by allowing more bandwidth (bits) to code the luminance (Y) and less bandwidth (bits) to code the chrominance (I and Q) for efficient transmission and storage purposes without introducing large perceptual errors due to quantization [1]. Another implication is that the luminance (Y) component of an image can be processed without affecting its chrominance (color content). For instance, histogram equalization to a color image represented in YIQ format can be done simply by applying histogram equalization to its Y component [1]. The relative colors in the image are not affected by this process.

The ideal way to accomplish these goals would be to form a luminance component (Y) by applying a matrix transform to the linear RGB components and then subjecting the luminance (Y) to a non-linear transfer function to achieve a component similar to lightness L^*. However, there are practical reasons in a video system why these operations are performed in the opposite order [4]. First, gamma correction is applied to each of the linear RGB. Then, a weighted sum of the nonlinear components is computed to form a component representative of luminance Y. The resulting component (luma) is related to luminance but is not the same as the CIE luminance Y although the same symbol is used for both of them.

The nonlinear RGB to YIQ conversion is defined by the following matrix transformation [4], [1]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{1.38}$$

As can be seen from the above transformation, the blue component has a small contribution to the brightness sensation (luma Y) despite the fact that human vision has extraordinarily good color discrimination capability in the blue color [4]. The inverse matrix transformation is performed to convert YIQ to nonlinear R′G′B′.

Introducing a cylindrical coordinate transformation, numerical values for hue and saturation can be calculated as follows:

$$H_{YIQ} = tan^{-1}(\frac{Q}{I}) \tag{1.39}$$

$$S_{YIQ} = (I^2 + Q^2)^{\frac{1}{2}} \tag{1.40}$$

As described it, the YIQ model is developed from a perceptual point of view and provides several advantages in image coding and communications applications by decoupling the luma (Y) and chrominance components (I and Q). Nevertheless, YIQ is a perceptually non-uniform color space and thus not appropriate for perceptual color difference quantification. For example,

the Euclidean distance is not capable of accurately measuring the perceptual color distance in the perceptually non-uniform YIQ color space. Therefore, YIQ is not the best color space for quantitative computations involving human color perception.

1.8 The HSI Family of Color Models

In image processing systems, it is often convenient to specify colors in a way that is compatible with the hardware used. The different variants of the RGB monitor model address that need. Although these systems are computationally practical, they are not useful for user specification and recognition of colors. The user cannot easily specify a desired color in the RGB model. On the other hand, perceptual features, such as perceived luminance (intensity), saturation and hue correlate well with the human perception of color . Therefore, a color model in which these color attributes form the basis of the space is preferable from the users point of view. Models based on lightness, hue and saturation are considered to be better suited for human interaction. The analysis of the user-oriented color spaces starts by introducing the family of intensity, hue and saturation (HSI) models [28], [29]. This family of models is used primarily in computer graphics to specify colors using the artistic notion of tints, shades and tones. However, all the HSI models are derived from the RGB color space by coordinate transformations. In a computer centered image processing system, it is necessary to transform the color coordinates to RGB for display and vice versa for color manipulation within the selected space.

The HSI family of color models use approximately cylindrical coordinates. The saturation (S) is proportional to radial distance, and the hue (H) is a function of the angle in the polar coordinate system. The intensity (I) or lightness (L) is the distance along the axis perpendicular to the polar coordinate plane. The dominant factor in selecting a particular HSI model is the definition of the lightness, which determines the constant-lightness surfaces, and thus, the shape of the color solid that represents the model. In the cylindrical models, the set of color pixels in the RGB cube which are assigned a common lightness value (L) form a constant-lightness surface. Any line parallel to the main diagonal of the color RGB cube meets the constant-lightness surface at most in one point.

The HSI color space was developed to specify, numerically, the values of hue, saturation, and intensity of a color [4]. The HSI color model is depicted in Figure 1.10. The hue (H) is measured by the angle around the vertical axis and has a range of values between 0 and 360 degrees beginning with red at 0°. It gives a measure of the spectral composition of a color. The saturation (S) is a ratio that ranges from 0 (i.e. on the I axis), extending radially outwards to a maximum value of 1 on the surface of the cone. This component refers to the proportion of pure light of the dominant wavelength and indicates how

far a color is from a gray of equal brightness. The intensity (I) also ranges between 0 and 1 and is a measure of the relative brightness. At the top and bottom of the cone, where $I = 0$ and 1 respectively, H and S are undefined and meaningless. At any point along the I axis the Saturation component is zero and the hue is undefined. This singularity occurs whenever $R = G = B$.

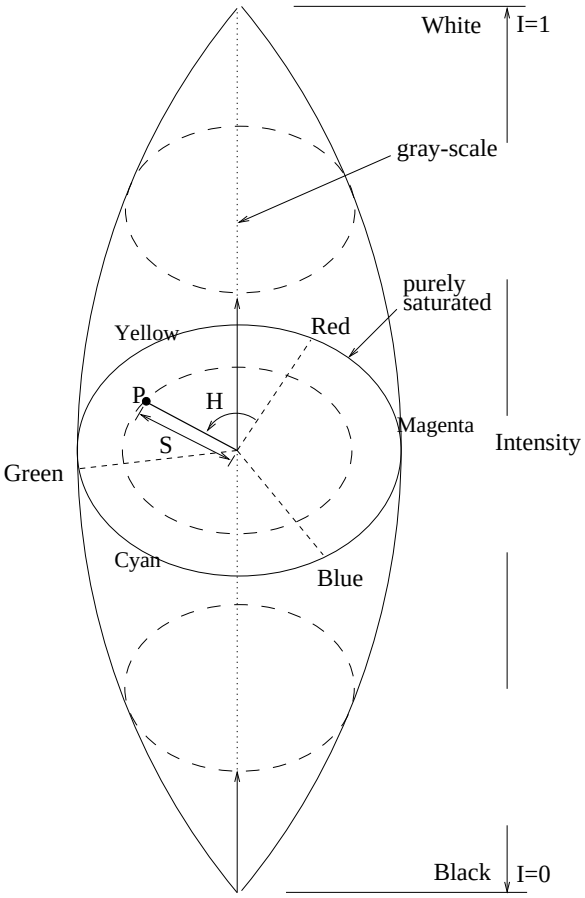


Fig. 1.10. The HSI Color Space

The HSI color model owes its usefulness to two principal facts [1], [28]. First, like in the YIQ model, the intensity component I is decoupled from the chrominance information represented as hue H and saturation S. Second, the hue (H) and saturation (S) components are intimately related to the way in which humans perceive chrominance [1]. Hence, these features make the HSI an ideal color model for image processing applications where the chrominance is of importance rather than the overall color perception (which is determined by both luminance and chrominance). One example of the usefulness of the

HSI model is in the design of imaging systems that automatically determine the ripeness of fruits and vegetables [1]. Another application is color image histogram equalization performed in the HSI space to avoid undesirable shifts in image hue [10].

The simplest way to choose constant-lightness surfaces is to define them as planes. A simplified definition of the perceived lightness in terms of the R,G,B values is $L = \frac{R'+G'+B'}{3}$, where the normalization is used to control the range of lightness values. The different constant-lightness surfaces are perpendicular to the main diagonal of the RGB cube and parallel to each other. The shape of a constant lightness surface is a triangle for $0 \leq L \leq \frac{M}{3}$ and $\frac{2M}{3} \leq L \leq M$ with $L \in [0, M]$ and where M is a given lightness threshold.

The theory underlying the derivation of conversion formulas between the RGB space and HSI space is described in detail in [1], [28]. The image processing literature on HSI does not clearly indicate whether the linear or the non-linear RGB is used in these conversions [4]. Thus the non-linear (R′G′B′), which is implicit in traditional image processing, shall be used. But this ambiguity must be noted.

The conversion from R′G′B′ (range [0, 1]) to HSI (range [0, 1]) is highly nonlinear and considerably complicated:

$$H = \cos^{-1}\left[\frac{\frac{1}{2}[(R' - G') + (R' - B')]}{[(R' - G')^2 + (R' - B')(G' - B')]^{\frac{1}{2}}}\right] \tag{1.41}$$

$$S = 1 - \frac{3}{(R' + G' + B')}[\min(R',\ G',\ B')] \tag{1.42}$$

$$I = \frac{1}{3}(R' + G' + B') \tag{1.43}$$

where $H = 360° - H$, if $(B'/I) > (G'/I)$. Hue is normalized to the range [0, 1] by letting $H = H/360°$. Hue (H) is not defined when the saturation (S) is zero. Similarly, saturation (S) is undefined if intensity (I) is zero.

To transform the HSI values (range [0, 1]) back to the R′G′B′ values (range [0, 1]), then the H values in [0, 1] range must first be converted back to the un-normalized $[0^o, 360^o]$ range by letting $H = 360°(H)$. For the $R'G'$ (red and green) sector ($0° < H \leq 120°$), the conversion is:

$$B' = I\ (1 - S) \tag{1.44}$$

$$R' = I\left[1 + \frac{S\ \cos H}{\cos\,(60° - H)}\right] \tag{1.45}$$

$$G' = 3I - (R' + B') \tag{1.46}$$

The conversion for the $G'B'$ (green and blue) sector ($120° < H \leq 240°$) is given by:

$$H = H - 120° \tag{1.47}$$

$$R' = I\,(1 - S) \tag{1.48}$$

$$G' = I\left[1 + \frac{S\,\cos H}{\cos(60^\circ - H)}\right] \tag{1.49}$$

$$B' = 3I - (R' + G') \tag{1.50}$$

Finally, for the $B'R'$ (blue and red) sector ($240^\circ < H \leq 360^\circ$), the corresponding equations are:

$$H = H - 240^\circ \tag{1.51}$$

$$G' = I\,(1 - S) \tag{1.52}$$

$$B' = I\left[1 + \frac{S\,\cos H}{\cos(60^\circ - H)}\right] \tag{1.53}$$

$$R' = 3I - (G' + B') \tag{1.54}$$

Fast versions of the transformation, containing fewer multiplications and avoiding square roots, are often used in hue calculations. Also, formulas without trigonometric functions can be used. For example, hue can be evaluated using the following formula [44]:

1. If $B' = \min(R',\ G',\ B')$ then

$$H = \frac{G' - B'}{3(R' + G' - 2B'} \tag{1.55}$$

2. If $R' = \min(R',\ G',\ B')$ then

$$H = \frac{B' - R'}{R' + G' - 2B'} + \frac{1}{3} \tag{1.56}$$

3. If $G' = \min(R',\ G',\ B')$ then

$$H = \frac{B' - R'}{R' + G' - 2B'} + \frac{2}{3} \tag{1.57}$$

Although the HSI model is useful in some image processing applications, the formulation of it is flawed with respect to the properties of color vision. The usual formulation makes no clear reference to the linearity or nonlinearity of the underlying RGB and to the lightness perception of human vision [4]. It computes the brightness as $(R' + G' + B')/3$ and assigns the name intensity I. Recall that the brightness perception is related to luminance Y. Thus, this computation conflicts with the properties of color vision [4].

In addition to this, there is a discontinuity in the hue at 360^o and thus, the formulation introduces visible discontinuities in the color space. Another major disadvantage of the HSI space is that it is not perceptually uniform.

Consequently, the HSI model is not very useful for perceptual image computation and for conveyance of accurate color information. As such, distance measures, such as the Euclidean distance, cannot estimate adequately the perceptual color distance in this space.

The model discussed above is not the only member of the family. In particular, the double hexcone HLS model can be defined by simply modifying the constant-lightness surface. It is depicted in Figure 1.11. In the HLS model the lightness is defined as:

$$L = \frac{\max(R', G', B') + \min(R', G', B')}{2} \tag{1.58}$$

If the maximum and the minimum value coincide then $S = 0$ and the hue is undefined. Otherwise based on the lightness value, saturation is defined as follows:

1. If $L \leq 0.5$ then $S = \frac{(Max - Min)}{(Max + Min)}$
2. If $L > 0.5$ then $S = \frac{(Max - Min)}{(2 - Max - Min)}$

where $Max = \max(R', G', B')$ and $Min = \min(R', G', B')$ respectively.

Similarly, hue is calculated according to:

1. If $R' = Max$ then
$$H = \frac{G' - B'}{Max - Min} \tag{1.59}$$
2. If $G' = Max$ then
$$H = \frac{B' - R'}{Max - Min} \tag{1.60}$$
3. If $B' = Max$ then
$$H = 4 + \frac{R' - G'}{Max - Min} \tag{1.61}$$

The backward transform starts by rescaling the hue angles into the range $[0, 6]$. Then, the following cases are considered:

1. If $S = 0$, hue is undefined and $(R', G', B') = (L, L, L)$
2. Otherwise, $i = Floor(H)$ (the $Floor(X)$ function returns the largest integer which is not greater than X), in which i is the sector number of the hue and $f = H - i$ is the hue value in each sector. The following cases are considered:

- if $L \leq L_{critical} = \frac{255}{2}$ then

$$Max = L(1 + S) \tag{1.62}$$

$$Mid1 = L(2fS + 1 - S) \tag{1.63}$$

$$Mid2 = L(2(1-f)S + 1 - S) \tag{1.64}$$

$$Min = L(1-S) \tag{1.65}$$

- if $L > L_{critical} = \frac{255}{2}$ then

$$Max = L(1-S) + 255S \tag{1.66}$$

$$Mid1 = 2((1-f)S - (0.5-f)Max) \tag{1.67}$$

$$Mid2 = 2(fL - (f-0.5)Max) \tag{1.68}$$

$$Min = L(1+S) - 255S \tag{1.69}$$

Based on these intermediate values the following assignments should be made:

1. if $i = 0$ then $(R', G', B') = (Max, Mid1, Min)$
2. if $i = 1$ then $(R', G', B') = (Mid2, Max, Min)$
3. if $i = 2$ then $(R', G', B') = (Min, Max, Mid1)$
4. if $i = 3$ then $(R', G', B') = (Min, Mid2, Max)$
5. if $i = 4$ then $(R', G', B') = (Mid1, Min, Max)$
6. if $i = 5$ then $(R', G', B') = (Max, Min, Mid2)$

The HSV (hue, saturation, value) color model also belongs to this group of hue-oriented color coordinate systems which correspond more closely to the human perception of color. This user-oriented color space is based on the intuitive appeal of the artist's tint, shade, and tone. The HSV coordinate system, proposed originally in Smith [36], is cylindrical and is conveniently represented by the hexcone model shown in Figure 1.12 [23], [27]. The set of equations below can be used to transform a point in the RGB coordinate system to the appropriate value in the HSV space.

$$H_1 = \cos^{-1}\{\frac{\frac{1}{2}[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}}\} \tag{1.70}$$

$$H = H_1 \quad , \text{ if } \quad B \leq G \tag{1.71}$$

$$H = 360^\circ - H_1 \quad , \text{ if } \quad B > G \tag{1.72}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \tag{1.73}$$

$$V = \frac{\max(R,G,B)}{255} \tag{1.74}$$

Here the RGB values are between 0 and 255. A fast algorithm used here to convert the set of RGB values to the HSV color space is provided in [23].

The important advantages of the HSI family of color spaces over other color spaces are:

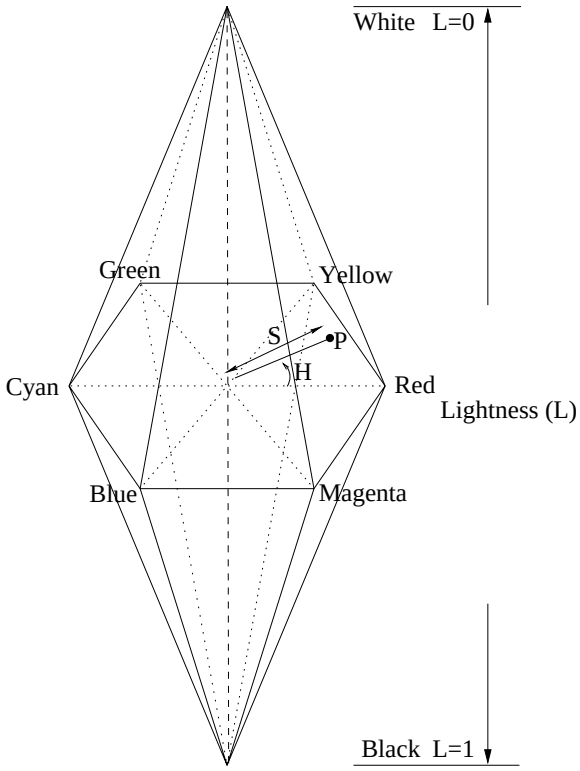


Fig. 1.11. The HLS Color Space

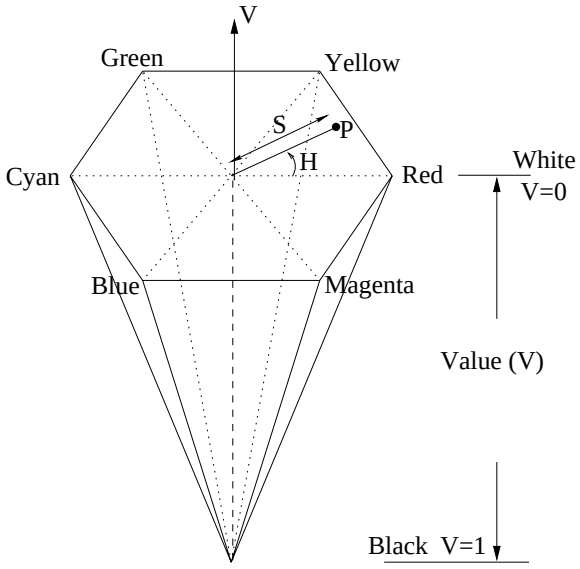


Fig. 1.12. The HSV Color Space

- Good compatibility with human intuition.
- Separability of chromatic values from achromatic values.
- The possibility of using one color feature, hue, only for segmentation purposes. Many image segmentation approaches take advantage of this. Segmentation is usually performed in one color feature (hue) instead of three, allowing the use of much faster algorithms.

However, hue-oriented color spaces have some significant drawbacks, such as:

- singularities in the transform, e.g. undefined hue for achromatic points
- sensitivity to small deviations of RGB values near singular points
- numerical instability when operating on hue due to the angular nature of the feature.

1.9 Perceptually Uniform Color Spaces

Visual sensitivity to small differences among colors is of paramount importance in color perception and specification experiments. A color system that is to be used for color specification should be able to represent any color with high precision. All systems currently available for such tasks are based on the CIE XYZ color model. In image processing, it is of particular interest in a perceptually uniform color space where a small perturbation in a component value is approximately equally perceptible across the range of that value. The color specification systems discussed until now, such as the XYZ or RGB tristimulus values and the various RGB hardware oriented systems are far from uniform. Recalling the discussion of YIQ space earlier in this chapter, the ideal way to compute the perceptual components representative of luminance and chrominance is to appropriately form the matrix of linear RGB components and then subject them to nonlinear transfer functions based on the color sensing properties of the human visual system. A similar procedure is used by CIE to formulate the L*u*v* and L*a*b* spaces. The linear RGB components are first transformed to CIE XYZ components using the appropriate matrix.

Finding a transformation of XYZ which transforms this color space into a reasonably perceptually uniform color space consumed a decade or more at the CIE and in the end, no single system could be agreed upon [4], [5]. Finally, in 1976, CIE standardized two spaces, L*u*v* and L*a*b*, as perceptually uniform. They are slightly different because of the different approaches to their formulation [4], [5], [25], [30]. Nevertheless, both spaces are equally good in perceptual uniformity and provide very good estimates of color difference (distance) between two color vectors.

Both systems are based on the perceived lightness L^* and a set of opponent color axes, approximately red-green versus yellow-blue. According to

the CIE 1976 standard, the perceived lightness of a standard observer is assumed to follow the physical luminance (a quantity proportional to intensity) according to a cubic root law. Therefore, the lightness L^* is defined by the CIE as:

$$L^* = \begin{cases} 116(\frac{Y}{Y_n})^{\frac{1}{3}} - 16 \text{ if } \frac{Y}{Y_n} > 0.008856 \\ 903.3(\frac{Y}{Y_n})^{\frac{1}{3}} \quad \text{if } \frac{Y}{Y_n} \leq 0.008856 \end{cases} \tag{1.75}$$

where Y_n is the physical luminance of the white reference point . The range of values for L^* is from 0 to 100 representing a black and a reference white respectively. A difference of unity between two L^* values, the so-called ΔL^* is the threshold of discrimination.

This standard function relates perceived lightness to linear light luminance. Luminance can be computed as a weighted sum of red, green and blue components. If three sources appear red, green and blue and have the same power in the visible spectrum, the green will appear the brightest of the three because the luminous efficiency function peaks in the green region of the spectrum. Thus, the coefficients that correspond to contemporary CRT displays (ITU-R BT. 709 recommendation) [24] reflect that fact, when using the following equation for the calculation of the luminance:

$$Y_{709} = 0.2125R + 0.7154G + 0.0721B \tag{1.76}$$

The u^* and v^* components in L*u*v* space and the the a^* and b^* components in L*a*b* space are representative of chrominance. In addition, both are device independent color spaces. Both these color spaces are, however, computationally intensive to transform to and from the linear as well as nonlinear RGB spaces. This is a disadvantage if real-time processing is required or if computational resources are at a premium.

1.9.1 The CIE L*u*v* Color Space

The first uniform color space standardized by CIE is the L*u*v* illustrated in Figure 1.13. It is derived based on the CIE XYZ space and white reference point [4], [5]. The white reference point [X_n, Y_n, Z_n] is the linear RGB = [1, 1, 1] values converted to the XYZ values using the following transformation:

$$\begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{1.77}$$

Alternatively, white reference points can be defined based on the Federal Communications Commission (FCC) or the European Broadcasting Union (EBU) RGB values using the following transformations respectively [35]:

$$\begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix} = \begin{bmatrix} 0.607\ 0.174\ 0.200 \\ 0.299\ 0.587\ 0.114 \\ 0.000\ 0.066\ 1.116 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{1.78}$$

$$\begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix} = \begin{bmatrix} 0.430\ 0.342\ 0.178 \\ 0.222\ 0.702\ 0.071 \\ 0.020\ 0.130\ 0.939 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{1.79}$$

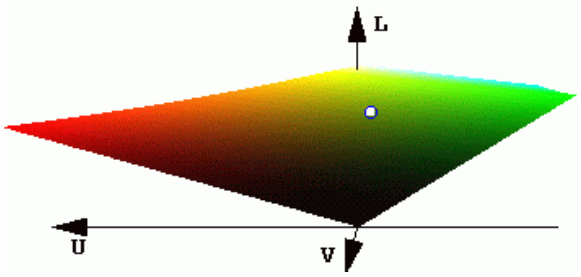


Fig. 1.13. The L*u*v* Color Space

The lightness component L^* is defined by the CIE as a modified cube root of luminance Y [4], [31], [37], [32]:

$$L^* = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3\left(\frac{Y}{Y_n}\right) & E\text{otherwise} \end{cases} \tag{1.80}$$

The CIE definition of L^* applies a linear segment near black for $(Y/Y_n) \leq 0.008856$. This linear segment is unimportant for practical purposes [4]. L^* has a range [0, 100], and a L^* of unity is roughly the threshold of visibility [4].

Computation of u^* and v^* involves intermediate u', v', u'_n, and v'_n quantities defined as:

$$u' = \frac{4X}{X + 15Y + 3Z} \quad v' = \frac{9Y}{X + 15Y + 3Z} \tag{1.81}$$

$$u'_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n} \quad v'_n = \frac{9Y_n}{X_n + 15Y_n + 3Z_n} \tag{1.82}$$

with the CIE XYZ values computed through (1.20) and (1.21).

Finally, u^* and v^* are computed as:

$$u^* = 13L^*(u' - u'_n) \tag{1.83}$$

$$v^* = 13L^*(v' - v'_n) \tag{1.84}$$

Conversion from L*u*v* to XYZ is accomplished by ignoring the linear segment of L^*. In particular, the linear segment can be ignored if the luminance variable Y is represented with eight bits of precision or less.

Then, the luminance Y is given by:

$$Y = \left(\frac{L^* + 16}{116}\right)^3 Y_n \tag{1.85}$$

To compute X and Z, first compute u' and v' as:

$$u' = \frac{u^*}{13L^*} + u'_n \quad v' = \frac{v^*}{13L^*} + v'_n \tag{1.86}$$

Finally, X and Z are given by:

$$X = \frac{1}{4}\left(\frac{u'(9.0 - 15.0\ v')\ Y}{v'} + 15.0\ u'\ Y\right) \tag{1.87}$$

$$Z = \frac{1}{3}\left(\frac{(9.0 - 15.0\ v')\ Y}{v'} - X\right) \tag{1.88}$$

Consider two color vectors $\mathbf{x}_{L^*u^*v^*}$ and $\mathbf{y}_{L^*u^*v^*}$ in the L*u*v* space represented as:

$$\mathbf{x}_{L^*u^*v^*} = [x_{L^*},\ x_{u^*},\ x_{v^*}]^T \quad \text{and} \quad \mathbf{y}_{L^*u^*v^*} = [y_{L^*},\ y_{u^*},\ y_{v^*}]^T \tag{1.89}$$

The perceptual color distance in the L*u*v* space, called the total color difference ΔE^*_{uv} in [5], is defined as the Euclidean distance (L_2 norm) between the two color vectors $\mathbf{x}_{L^*u^*v^*}$ and $\mathbf{y}_{L^*u^*v^*}$:

$$\Delta E^*_{uv} = ||\mathbf{x}_{L^*u^*v^*} - \mathbf{y}_{L^*u^*v^*}||_{L_2}$$

$$\Delta E^*_{uv} = \left[(x_{L^*} - y_{L^*})^2 + (x_{u^*} - y_{u^*})^2 + (x_{v^*} - y_{v^*})^2\right]^{\frac{1}{2}} \tag{1.90}$$

It should be mentioned that in a perceptually uniform space, the Euclidean distance is an accurate measure of the perceptual color difference [5]. As such, the color difference formula ΔE^*_{uv} is widely used for the evaluation of color reproduction quality in an image processing system, such as color coding systems.

1.9.2 The CIE L*a*b* Color Space

The L*a*b* color space is the second uniform color space standardized by CIE. It is also derived based on the CIE XYZ space and white reference point [5], [37].

The lightness L^* component is the same as in the L*u*v* space. The L^*, a^* and b^* components are given by:

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \tag{1.91}$$

$$a^* = 500\left[\left(\frac{X}{X_n}\right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right] \tag{1.92}$$

$$b^* = 200\left[\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}\right] \tag{1.93}$$

with the constraint that $\frac{X}{X_n}$, $\frac{Y}{Y_n}$, $\frac{Z}{Z_n} > 0.01$. This constraint will be satisfied for most practical purposes [4]. Hence, the modified formulae described in [5] for cases that do not not satisfy this constraint can be ignored in practice [4], [10].

The back conversion to the XYZ space from the L*a*b* space is done by first computing the luminance Y, as described in the back conversion of L*u*v*, followed by the computation of X and Z:

$$Y = \left(\frac{L^* + 16}{116}\right)^3 Y_n \tag{1.94}$$

$$X = \left(\frac{a^*}{500} + \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right)^3 X_n \tag{1.95}$$

$$Z = \left(-\frac{b^*}{200} + \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right)^3 Z_n \tag{1.96}$$

The perceptual color distance in the L*a*b* is similar to the one in the L*u*v*. The two color vectors $\mathbf{x}_{L^*a^*b^*}$ and $\mathbf{y}_{L^*a^*b^*}$ in the L*a*b* space can be represented as:

$$\mathbf{x}_{L^*a^*b^*} = [x_{L^*},\, x_{a^*},\, x_{b^*}]^T \quad \text{and} \quad \mathbf{y}_{L^*a^*b^*} = [y_{L^*},\, y_{a^*},\, y_{b^*}]^T \tag{1.97}$$

The perceptual color distance (or total color difference) in the L*a*b* space, ΔE^*_{ab}, between two color vectors $\mathbf{x}_{L^*u^*v^*}$ and $\mathbf{y}_{L^*u^*v^*}$ is given by the Euclidean distance (L_2 norm):

$$\begin{aligned}\Delta E^*_{ab} &= ||\mathbf{x}_{L^*a^*b^*} - \mathbf{y}_{L^*a^*b^*}||_{L_2} \\ &= \left[(x_{L^*} - y_{L^*})^2 + (x_{a^*} - y_{a^*})^2 + (x_{b^*} - y_{b^*})^2\right]^{\frac{1}{2}}\end{aligned} \tag{1.98}$$

The color difference formula ΔE^*_{ab} is applicable to the observing conditions normally found in practice, as in the case of ΔE^*_{uv}. However, this simple difference formula values color differences too strongly when compared to experimental results. To correct the problem a new difference formula was recommended in 1994 by CIE [25], [31]. The new formula is as follows:

$$\Delta E^*_{ab_{94}} = [\frac{(x_{L^*} - y_{L^*})^2}{K_L S_L} + \frac{(x_{a^*} - y_{a^*})^2}{K_c S_c} + \frac{(x_{b^*} - y_{b^*})^2}{K_H S_H}]^{\frac{1}{2}} \tag{1.99}$$

where the factors K_L, K_c, K_H are factors to match the perception of the background conditions, and S_L, S_c, S_H are linear functions of the differences in chroma. Standard reference values for the calculation for $\Delta E^*_{ab_{94}}$ have been

specified by the CIE. Namely, the values most often in use are $K_L = K_c = K_H = 1$, $S_L = 1$, $S_c = 1 + 0.045((x_{a^*} - y_{a^*})$ and $S_H = 1 + 0.015((x_{b^*} - y_{b^*})$ respectively. The parametric values may be modified to correspond to typical experimental conditions. As an example, for the textile industry, the K_L factor should be 2, and the K_c and K_H factors should be 1. For all other applications a value of 1 is recommended for all parametric factors [38].

1.9.3 Cylindrical L*u*v* and L*a*b* Color Space

Any color expressed in the rectangular coordinate system of axes L*u*v* or L*a*b* can also be expressed in terms of cylindrical coordinates with the perceived lightness L^* and the psychometric correlates of chroma and hue [37]. The chroma in the L*u*v* space is denoted as C^*_{uv} and that in the L*a*b* space C^*_{ab}. They are defined as [5]:

$$C^*_{uv} = \left((u^*)^2 + (v^*)^2\right)^{\frac{1}{2}} \tag{1.100}$$

$$C^*_{ab} = \left((a^*)^2 + (b^*)^2\right)^{\frac{1}{2}} \tag{1.101}$$

The hue angles are useful quantities in specifying hue numerically [5], [37]. Hue angle h_{uv} in the L*u*v* space and h_{ab} in the L*a*b* space are defined as [5]:

$$h_{uv} = \arctan\left(\frac{v^*}{u^*}\right) \tag{1.102}$$

$$h_{ab} = \arctan\left(\frac{b^*}{a^*}\right) \tag{1.103}$$

The saturation s^*_{uv} in the L*u*v* space is given by:

$$s^*_{uv} = \frac{C^*_{uv}}{L^*} \tag{1.104}$$

1.9.4 Applications of L*u*v* and L*a*b* spaces

The L*u*v* and L*a*b* spaces are very useful in applications where precise quantification of perceptual distance between two colors is necessary [5]. For example in the realization of perceptual based vector order statistics filters. If a degraded color image has to be filtered so that it closely resembles, in perception, the un-degraded original image, then a good criterion to optimize is the perceptual error between the output image and the un-degraded original image. Also, they are very useful for evaluation of perceptual closeness or perceptual error between two color images [4]. Precise evaluation of perceptual closeness between two colors is also essential in color matching systems used in various applications such as multimedia products, image arts, entertainment, and advertisements [6], [14], [22].

$L^*u^*v^*$ and $L^*a^*b^*$ color spaces are extremely useful in imaging systems where exact perceptual reproduction of color images (color consistency) across the entire system is of primary concern rather than real-time or simple computing. Applications include advertising, graphic arts, digitized or animated paintings etc. Suppose, an imaging system consists of various color devices, for example video camera/digital scanner, display device, and printer. A painting has to be digitized, displayed, and printed. The displayed and printed versions of the painting must appear as close as possible to the original image. $L^*u^*v^*$ and $L^*a^*b^*$ color spaces are the best to work with in such cases. Both these systems have been successfully applied to image coding for printing [4], [16].

Color calibration is another important process related to color consistency. It basically equalizes an image to be viewed under different illumination or viewing conditions. For instance, an image of a target object can only be taken under a specific lighting condition in a laboratory. But the appearance of this target object under normal viewing conditions, say in ambient light, has to be known. Suppose, there is a sample object whose image under ambient light is available. Then the solution is to obtain the image of the sample object under the same specific lighting condition in the laboratory. Then a correction formula can be formulated based on the images of the sample object obtained and these can be used to correct the target object for the ambient light [14]. Perceptual based color spaces, such as $L^*a^*b^*$, are very useful for computations in such problems [31], [37]. An instance, where such calibration techniques have great potential, is medical imaging in dentistry.

Perceptually uniform color spaces, with the Euclidean metric to quantify color distances, are particularly useful in color image segmentation of natural scenes using histogram-based or clustering techniques.

A method of detecting clusters by fitting to them some circular-cylindrical decision elements in the $L^*a^*b^*$ uniform color coordinate system was proposed in [39], [40]. The method estimates the clusters' color distributions without imposing any constraints on their forms. Boundaries of the decision elements are formed with constant lightness and constant chromaticity loci. Each boundary is obtained using only 1-D histograms of the $L^*H^\circ C^*$ cylindrical coordinates of the image data. The cylindrical coordinates $L^*H^\circ C^*$ [30] of the $L^*a^*b^*$ color space known as lightness, hue, and chroma, are given by:

$$L^* = L^* \tag{1.105}$$

$$H^\circ = \arctan(b^*/a^*) \tag{1.106}$$

$$C^* = (a^{*2} + b^{*2})^{1/2} \tag{1.107}$$

The $L^*a^*b^*$ space is often used in *color management systems* (CMS). A color management system handles the color calibration and color consistency issues. It is a layer of software resident on a computer that negotiates color reproduction between the application and color devices. Color management systems perform the color transformations necessary to exchange accurate

color between diverse devices [4], [43]. A uniform, based on CIE L*u*v*, color space named TekHVC was proposed by Tektronix as part of its commercially available CMS [45].

1.10 The Munsell Color Space

The Munsell color space represents the earliest attempt to organize color perception into a color space [5], [14], [46]. The Munsell space is defined as a comparative reference for artists. Its general shape is that of a cylindrical representation with three dimensions roughly corresponding to the perceived lightness, hue and saturation. However, contrary to the HSV or HSI color models where the color solids were parameterized by hue, saturation and perceived lightness, the Munsell space uses the method of the *color atlas*, where the perception attributes are used for sampling.

The fundamental principle behind the Munsell color space is that of equality of visual spacing between each of the three attributes. Hue is scaled according to some uniquely identifiable color. It is represented by a circular band divided into ten sections. The sections are defined as red, yellow-red, yellow, green-yellow, green, blue-green, blue, purple-blue, purple and red-purple. Each section can be further divided into ten subsections if finer divisions of hue are necessary. A chromatic hue is described according to its resemblance to one or two adjacent hues. Value in the Munsell color space refers to a color's lightness or darkness and is divided into eleven sections numbered zero to ten. Value zero represents black while a value of ten represent white. The chroma defines the color's strength. It is measured in numbered steps starting at one with weak colors having low chroma values. The maximum possible chroma depends on the hue and the value being used. As can be seen in Fig. (1.14), the vertical axis of the Munsell color solid is the line of V values ranging from black to white. Hue changes along each of the circles perpendicular to the vertical axis. Finally, chroma starts at zero on the V axis and changes along the radius of each circle.

The Munsell space is comprised of a set of 1200 color chips each assigned a unique hue, value and chroma component. These chips are grouped in such a way that they form a three dimensional solid, which resembles a warped sphere [5]. There are different editions of the basic Munsell book of colors, with different finishes (glossy or matte), different sample sizes and a different number of samples. The glossy finish collection displays color point chips arranged on 40 constant-hue charts. On each constant-hue chart the chips are arranged in rows and columns. In this edition the colors progress from light at the top of each chart to very dark at the bottom by steps which are intended to be perceptually equal. They also progress from achromatic colors, such as white and gray at the inside edge of the chart, to chromatic colors at the outside edge of the chart by steps that are also intended to be

perceptually equal. All the charts together make up the color atlas, which is the color solid of the Munsell system.

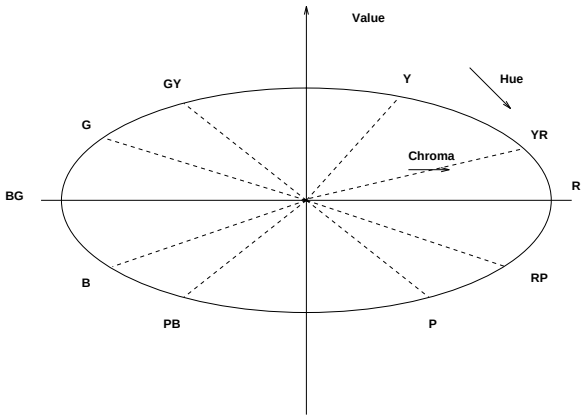


Fig. 1.14. The Munsell color system

Although the Munsell book of colors can be used to define or name colors, in practice is not used directly for image processing applications. Usually stored image data, most often in RGB format, are converted to the Munsell coordinates using either lookup tables or closed formulas prior to the actual application. The conversion from the RGB components to the Munsell hue (H), value (V) corresponding to luminance and chroma (C) corresponding to saturation, can be achieved by using the following mathematical algorithm [47]:

$$x = 0.620R + 0.178G + 0.204B$$

$$y = 0.299R + 0.587G + 0.144B$$

$$z = 0.056G + 0.942B \tag{1.108}$$

A nonlinear transformation is applied to the intermediate values as follows:

$$p = f(x) - f(y) \tag{1.109}$$

$$q = 0.4(f(z) - f(y)) \tag{1.110}$$

where $f(r) = 11.6r^{\frac{1}{3}} - 1.6$. Further the new variables are transformed to:

$$s = (a + bcos(\theta))p \tag{1.111}$$

$$t = (c + dsin(\theta))q \tag{1.112}$$

where $\theta = tan^{-1}(\frac{p}{q})$, $a = 8.880$, $b = 0.966$, $c = 8.025$ and $d = 2.558$. Finally, the requested values are obtained as:

$$H = \arctan(\frac{s}{t}) \tag{1.113}$$

$$V = f(y) \tag{1.114}$$

and

$$C = (s^2 + t^2)^{\frac{1}{2}} \tag{1.115}$$

Alternatively, conversion from RGB, or other color spaces, to the Munsell color space can be achieved through look-up tables and published charts [5].

In summary, the Munsell color system is an attempt to define color in terms of hue, chroma and lightness parameters based on subjective observations rather than direct measurements or controlled perceptual experiments. Although it has been found that the Munsell space is not as perceptually uniform as originally claimed and, despite the fact that it cannot directly integrate with additive color schemes, it is still in use today despite attempts to introduce colorimetric models for its replacement.

1.11 The Opponent Color Space

The opponent color space family is a set of physiologically motivated color spaces inspired by the physiology of the human visual system. According to the theory of color vision discussed in [48] the human vision system can be expressed in terms of opponent hues, yellow and blue on one hand and green and red on the other, which cancel each other when superimposed. In [49] an experimental procedure was developed which allowed researchers to quantitatively express the amounts of each of the basic hues present in any spectral stimulus. The color model of [50], [51], [52], [44] suggests the transformation of the RGB 'cone' signals to three channels, one achromatic channel (I) and two opponent color channels (RG, YB) according to:

$$RG = R - G \tag{1.116}$$

$$YB = 2B - R - G \tag{1.117}$$

$$I = R + G + B \tag{1.118}$$

At the same time a set of effective color features was derived by systematic experiments of region segmentation [53]. According to the segmentation procedure of [53] the color which has the deep valleys on its histogram and has the largest discriminant power to separate the color clusters in a given region need not be the R, G, and B color features. Since a feature is said to have large discriminant power if its variance is large, color features with large discriminant power were derived by utilizing the Karhunen-Loeve (KL) transformation. At every step of segmenting a region, calculation of the new color features is done for the pixels in that region by the KL transform of R, G, and B signals. Based on extensive experiments [53], it was concluded

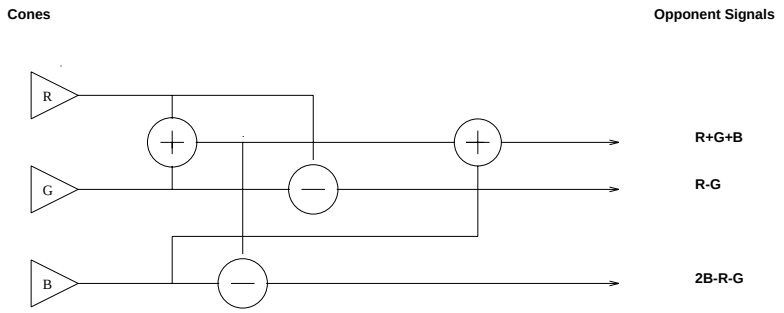


Fig. 1.15. The Opponent color stage of the human visual system

that three color features constitute an effective set of features for segmenting color images, [54], [55]:

$$I1 = \frac{(R+G+B)}{3} \tag{1.119}$$

$$I2 = (R-B) \tag{1.120}$$

$$I3 = \frac{(2G-R-B)}{2} \tag{1.121}$$

In the opponent color space hue could be coded in a circular format ranging through blue, green, yellow, red and black to white. Saturation is defined as distance from the hue circle making hue and saturation speciable with in color categories. Therefore, although opponent representation are often thought as a linear transforms of RGB space, the opponent representation is much more suitable for modeling perceived color than RGB is [14].

1.12 New Trends

The plethora of color models available poses application difficulties. Since most of them are designed to perform well in a specific application, their performance deteriorates rapidly under different operating conditions. Therefore, there is a need to merge the different (mainly device dependent) color spaces into a single standard space. The differences between the monitor RGB space and device independent spaces, such as the HVS and the CIE L*a*b* spaces impose problems in applications, such as multimedia database navigation and face recognition primarily due to the complexity of the operations needed to support the transform from/to device dependent color spaces.

To overcome such problems and to serve the needs of network-centric applications and WWW-based color imaging systems, a new standardized color space based on a colorimetric RGB (sRGB) space has recently been proposed [56]. The aim of the new color space is to complement the current color space

management strategies by providing a simple, yet efficient and cost effective method of handling color in the operating systems, device drivers and the Web using a simple and robust device independent color definition.

Since most computer monitors are similar in their key color characteristics and the RGB space is the most suitable color space for the devices forming a modern computer-based imaging systems, the colorimetric RGB space seems to be the best candidate for such a standardized color space.

In defining a colorimetric color space, two factors are of paramount importance:

- the viewing environment parameters with its dependencies on the Human Visual System
- the standard device space colorimetric definitions and transformations [56]

The viewing environment descriptions contain all the necessary transforms needed to support conversions between standard and target viewing environments. On the other hand, the colorimetric definitions provide the transforms necessary to convert between the new sRGB and the CIE-XYZ color space.

The reference viewing environment parameters can be found in [56] with the sRGB tristimulus values calculated from the CIE-XYZ values according to the following transform:

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1.122}$$

In practical image processing systems negative sRGB tristimulus values and sRGB values greater than 1 are not retained and typically removed by utilizing some form of clipping. In the sequence, the linear tristimulus values are transformed to nonlinear sR′G′B′ as follows:

1. If $R_{sRGB}, G_{sRGB}, B_{sRGB} \leq 0.0034$ then

$$sR' = 12.92 R_{sRGB} \tag{1.123}$$

$$sG' = 12.92 G_{sRGB} \tag{1.124}$$

$$sB' = 12.92 B_{sRGB} \tag{1.125}$$

2. else if $R_{sRGB}, G_{sRGB}, B_{sRGB} > 0.0034$ then

$$sR' = 1.055 {R_{sRGB}}^{\frac{1.0}{2.4}} - 0.055 \tag{1.126}$$

$$sG' = 1.055 {G_{sRGB}}^{\frac{1.0}{2.4}} - 0.055 \tag{1.127}$$

$$sB' = 1.055 {B_{sRGB}}^{\frac{1.0}{2.4}} - 0.055 \tag{1.128}$$

The effect of the above transformation is to closely fit a straightforward γ value of 2.2 with a slight offset to allow for invertibility in integer mathematics. The nonlinear R′G′B′ values are then converted to digital values with a black digital count of 0 and a white digital count of 255 for 24-bit coding as follows:

$$sR_d = 255.0sR' \tag{1.129}$$

$$sG_d = 255.0sG' \tag{1.130}$$

$$sB_d = 255.0sB' \tag{1.131}$$

The backwards transform is defined as follows:

$$sR' = sR_d + 255.0 \tag{1.132}$$

$$sG' = sG_d + 255.0 \tag{1.133}$$

$$sB' = sB_d + 255.0 \tag{1.134}$$

and

1. if $R_{sRGB}, G_{sRGB}, B_{sRGB} \leq 0.03928$ then

$$R_{sRGB} = sR' + 12.92 \tag{1.135}$$

$$G_{sRGB} = sG' + 12.92 \tag{1.136}$$

$$B_{sRGB} = sB' + 12.92 \tag{1.137}$$

2. else if $R_{sRGB}, G_{sRGB}, B_{sRGB} > 0.03928$ then

$$R_{sRGB} = \left(\frac{sR' + 0.055}{1.055}\right)^{2.4} \tag{1.138}$$

$$G_{sRGB} = \left(\frac{sG' + 0.055}{1.055}\right)^{2.4} \tag{1.139}$$

$$B_{sRGB} = \left(\frac{sB' + 0.055}{1.055}\right)^{2.4} \tag{1.140}$$

with

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124\ 0.3576\ 0.1805 \\ 0.2126\ 0.7152\ 0.0722 \\ 0.0193\ 0.1192\ 0.9505 \end{bmatrix} \begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} \tag{1.141}$$

The addition of a new standardized color space which supports Web-based imaging systems, device drivers, printers and monitors complementing the existing color management support can benefit producers and users alike by presenting a clear path towards an improved color management system.

1.13 Color Images

Color imaging systems are used to capture and reproduce the scenes that humans see. Imaging systems can be built using a variety of optical, electronic or chemical components. However, all of them perform three basic operations, namely: (i) image capture, (ii) signal processing, and (iii) image formation. Color-imaging devices exploit the trichromatic theory of color to regulate how much light from the three primary colors is absorbed or reflected to produce a desired color.

There are a number of ways to acquiring and reproducing color images, including but not limited to:

- **Photographic film**. The film which is used by conventional cameras contains three emulation layers, which are sensitive to red and blue light, which enters through the camera lens.
- **Digital cameras**. Digital cameras use a CCD to capture image information. Color information is captured by placing red, green and blue filters before the CCD and storing the response to each channel.
- **Cathode-Ray tubes**. CRTs are the display device used in televisions and computer monitors. They utilize a extremely fine array of phosphors that emit red, green and blue light at intensities governed by an electron gun, in accordance to an image signal. Due to the close proximity of the phosphors and the spatial filtering characteristics of the human eye, the emitted primary colors are mixed together producing an overall color.
- **Image scanners**. The most common method of scanning color images is the utilization of three CCD's each with a filter to capture red, green and blue light reflectance. These three images are then merged to create a copy of the scanned image.
- **Color printers**. Color printers are the most common method of attaining a printed copy of a captured color image. Although the trichromatic theory is still implemented, color in this domain is subtractive. The primaries which are used are usually cyan, magenta and yellow. The amount of the three primaries which appear on the printed media govern how much light is reflected.

1.14 Summary

In this chapter the phenomenon of color was discussed. The basic color sensing properties of the human visual system and the CIE standard color specification system XYZ were described in detail. The existence of three types of spectral absorption cones in the human eyes serves as the basis of the trichromatic theory of color, according to which all visible colors can be created by combining three . Thus, any color can be uniquely represented by a three dimensional vector in a color model defined by the three primary colors.

Table 1.3. Color Model

Color System	Transform (from RGB)	Component correlation
RGB	-	highly correlated
R′G′B′	non linear	
XYZ	linear	correlated
YIQ	linear	uncorrelated
YCC	linear	uncorrelated
I1I2I3	linear	correlated
HSV	non linear	correlated
HSI	non linear	correlated
HLS	non linear	correlated
L*u*v*	non linear	correlated
L*a*b*	non linear	correlated
Munsell	non linear	correlated

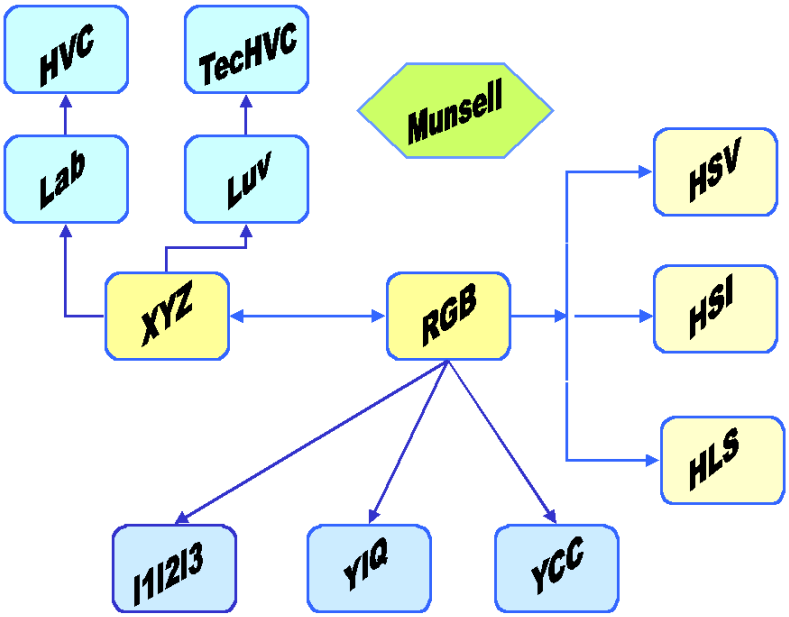


Fig. 1.16. A taxonomy of color models

Color specification models are of paramount importance in applications where efficient manipulation and communication of images and video frames are required. A number of color specification models are in use today. Examples include color spaces, such as the RGB, R′G′B′, YIQ, HSI, HSV, HLS,L*u*v*, and L*a*b*. The color model is a mathematical representation of spectral colors in a finite dimensional vector space. In each one of them the actual color is reconstructed by combining the basis elements of the vector

Color Spaces		
Models		**Applications**
Colorimetric	XYZ	colorimetric calculations
Device-oriented	- non-uniform spaces RGB, YIQ, YCC	storage, processing, analysis coding, color TV, storage (CD-ROM)
	- uniform spaces L*a*b*, L*u*v*	color difference evaluation analysis, color management systems
User-oriented	HSI, HSV, HLS, I1I2I3	human color perception multimedia, computer graphics
Munsell		human visual system

spaces, the so called primary colors. By defining different primary colors for the representation of the system different color models can be devised. One important aspect is the color transformation, the change of coordinates from one color system to another (see Table 1.3). Such a transformation associates to each color in one system a color in the other model. Each color model comes into existence for a specific application in color image processing. Unfortunately, there is no technique for determining the optimum coordinate model for all image processing applications. For a specific application the choice of a color model depends on the properties of the model and the design characteristics of the application. Table 1.14 summarizes the most popular color systems and some of their applications.

References

1. Gonzalez, R., Woods, R.E. (1992): Digital Image Processing. Addisson Wesley, Reading MA.
2. Robertson, P., Schonhut, J. (1999): Color in computer graphics. IEEE Computer Graphics and Applications, **19**(4), 18-19.
3. MacDonald, L.W. (1999): Using color effectively in computer graphics. IEEE Computer Graphics and Applications, **19**(4), 20-35.
4. Poynton, C.A. (1996): A Technical Introduction to Digital Video. Prentice Hall, Toronto, also available at http://www.inforamp.net/~poynton/Poynton–Digital-Video.html .
5. Wyszecki, G., Stiles, W.S. (1982): Color Science, Concepts and Methods, Quantitative Data and Formulas. John Wiley, N.Y. , 2^{nd} Edition.
6. Hall, R.A. (1981): Illumination and Color in Computer Generated Imagery. Springer Verlag, New York, N.Y.
7. Hurlbert, A. (1989): The Computation of Color. Ph.D Dissertation, Massachusetts Institute of Technology.
8. Hurvich, Leo M. (1981): Color Vision. Sinauer Associates, Sunderland MA.
9. Boynton, R.M. (1990): Human Color Vision. Halt, Rinehart and Winston.
10. Gomes, J., Velho, L. (1997): Image Processing for Computer Graphics. Springer Verlag, New York, N.Y., also available at http://www.springer-ny.com/catalog/np/mar97np/DATA/0-387-94854-6.html .

11. Fairchild, M.D. (1998): Color Appearance Models. Addison-Wesley, Readings, MA.
12. Sharma, G., Yrzel, M.J., Trussel, H.J. (1998): Color imaging for multimedia. Proceedings of the IEEE, **86**(6): 1088–1108.
13. Sharma, G., Trussel, H.J. (1997): Digital color processing. IEEE Trans. on Image Processing, **6**(7): 901-932.
14. Lammens, J.M.G. (1994): A Computational Model for Color Perception and Color Naming. Ph.D Dissertation, State University of New York at Buffalo, Buffalo, New York.
15. Johnson, G.M., Fairchild, M.D. (1999): Full spectral color calculations in realistic image synthesis. IEEE Computer Graphics and Applications, **19**(4), 47-53.
16. Lu, Guoyun (1996): Communication and Computing for Distributed Multimedia Systems. Artech House Publishers, Boston, MA.
17. Kubinger, W., Vincze, M., Ayromlou, M. (1998): The role of gamma correction in colour image processing. in Proceedings of the European Signal Processing Conference, **2**: 1041–1044.
18. Luong, Q.T. (1993): Color in computer vision. in Handbook of Pattern Recognition and Computer Vision, Word Scientific Publishing Company): 311–368.
19. Young, T. (1802): On the theory of light and colors. Philosophical Transactions of the Royal Society of London, **92**: 20–71.
20. Maxwell, J.C. (1890): On the theory of three primary colors. Science Papers 1, Cambridge University Press: 445–450.
21. Padgham, C.A., Saunders, J.E. (1975): The Perception of Light and Color. Academic Press, New York, N.Y.
22. Judd, D.B., Wyszecki, G. (1975): Color in Business, Science and Industry. John Wiley, New York, N.Y.
23. Foley, J.D., vanDam, A., Feiner, S.K., Hughes, J.F. (1990): Fundamentals of Interactive Computer Graphics. Addison Wesley, Reading, MA.
24. CCIR (1990): CCIR Recommendation 709. Basic parameter values for the HDTV standard for studio and for international program exchange. Geneva, Switcherland.
25. CIE (1995): CIE Publication 116. Industrial color-difference evaluation. Vienna, Austria.
26. Poynton, C.A. (1993): Gamma and its disguises. The nonlinear mappings of intensity in perception, CRTs, film and video. SMPTE Journal: 1099–1108.
27. Kasson M.J., Ploaffe, W. (1992): An analysis of selected computer interchange color spaces. ACM Transaction of Graphics, **11**(4): 373-405.
28. Shih, Tian-Yuan (1995): The reversibility of six geometric color spaces. Photogrammetric Engineering and Remote Sensing, **61**(10): 1223–1232.
29. Levkowitz H., Herman, G.T. (1993): GLHS: a generalized lightness, hue and saturation color model. Graphical Models and Image Processing, **CVGIP-55**(4): 271–285.
30. McLaren, K. (1976): The development of the CIE $L^*a^*b^*$ uniform color space. J. Soc. Dyers Colour, 338–341.
31. Hill, B., Roer, T., Vorhayen, F.W. (1997): Comparative analysis of the quantization of color spaces on the basis of the CIE-Lab color difference formula. ACM Transaction of Graphics, **16**(1): 110–154.
32. Hall, R. (1999): Comparing spectral color computation methods. IEEE Computer Graphics and Applications, **19**(4), 36-44.
33. Hague, G.E., Weeks, A.R., Myler, H.R. (1995): Histogram equalization of 24 bit color images in the color difference color space. Journal of Electronic Imaging, **4**(1), 15-23.

34. Weeks, A.R. (1996): Fundamentals of Electronic Image Processing. SPIE Press, Piscataway, New Jersey.
35. Benson, K. B. (1992): Television Engineering Handbook. McGraw-Hill, London, U.K.
36. Smith, A.R. (1978): Color gamut transform pairs. Computer Graphics (SIGGRAPH'78 Proceedings), **12**(3): 12–19.
37. Healey, C.G., Enns, J.T. (1995): A perceptual color segmentation algorithm. Technical Report, Department of Computer Science, University of British Columbia, Vancouver.
38. Luo, M. R. (1998): Color science. in Sangwine, S.J., Horne, R.E.N. (eds.), The Colour Image Processing Handbook, 26–52, Chapman & Hall, Cambridge, Great Britain.
39. Celenk, M. (1988): A recursive clustering technique for color picture segmentation. Proceedings of the Int. Conf. on Computer Vision and Pattern Recognition, **1**: 437–444.
40. Celenk, M. (1990): A color clustering technique for image segmentation. Computer Vision, Graphics, and Image Processing, **52**: 145–170.
41. Cong, Y. (1998): Intelligent Image Databases. Kluwer Academic Publishers, Boston, Ma.
42. Ikeda, M. (1980): Fundamentals of Color Technology. Asakura Publishing, Tokyo, Japan.
43. Rhodes, P. A. (1998): Colour management for the textile industry. in Sangwine, S.J., Horne, R.E.N. (eds.), The Colour Image Processing Handbook, 307-328, Chapman & Hall, Cambridge, Great Britain.
44. Palus, H. (1998): Colour spaces. in Sangwine, S.J., Horne, R.E.N. (eds.), The Colour Image Processing Handbook, 67–89, Chapman & Hall, Cambridge, Great Britain.
45. Tektronix (1990): TekColor Color Management System: System Implementers Manual. Tektronix Inc.
46. Birren, F. (1969): Munsell: A Grammar of Color. Van Nostrand Reinhold, New York, N.Y.
47. Miyahara, M., Yoshida, Y. (1988): Mathematical transforms of (R,G,B) colour data to Munsell (H,V,C) colour data. Visual Communications and Image Processing, **1001**, 650–657.
48. Hering, E. (1978): Zur Lehe vom Lichtsinne. C. Gerond's Sohn, Vienna, Austria.
49. Jameson, D., Hurvich, L.M. (1968): Opponent-response functions related to measured cone photo pigments. Journal of the Optical Society of America, **58**: 429–430.
50. de Valois, R.L., De Valois, K.K. (1975): Neural coding of color. in Carterette, E.C., Friedman, M.P. (eds.), Handbook of Perception. Volume 5, Chapter 5, 117–166, Academic Press, New York, N.Y.
51. de Valois, R.L., De Valois, K.K. (1993): A multistage color model. Vision Research **33**(8): 1053–1065.
52. Holla, K. (1982): Opponent colors as a 2-dimensional feature within a model of the first stages of the human visual system. Proceedings of the 6^{th} Int. Conf. on Pattern Recognition, **1**: 161–163.
53. Ohta, Y., Kanade, T., Sakai, T. (1980): Color information for region segmentation. Computer Graphics and Image Processing, **13**: 222–241.
54. von Stein, H.D., Reimers, W. (1983): Segmentation of color pictures with the aid of color information and spatial neighborhoods. Signal Processing II: Theories and Applications, **1**: 271–273.
55. Tominaga S. (1986): Color image segmentation using three perceptual attributes. Proceedings of CVPR'86, **1**: 628-630.

56. Stockes, M., Anderson, M., Chandrasekar, Sri., Motta, Ricardo (1997): A standard default color space for the Internet sRGB. International Color Consortium (ICC), contributed document electronic reprint (http://www.color.org).