# Computing Chromatic Adaptation

Sabine Süsstrunk

A thesis submitted for the Degree of
Doctor of Philosophy
in the
School of Computing Sciences,
University of East Anglia, Norwich.
July 2005

# Abstract

Most of today's *chromatic adaptation transforms* (CATs) are based on a modified form of the von Kries chromatic adaptation model, which states that chromatic adaptation is an independent gain regulation of the three photoreceptors in the human visual system. However, modern CATs apply the scaling not in cone space, but use "sharper" sensors, i.e. sensors that have a narrower shape than cones. The recommended transforms currently in use are derived by minimizing perceptual error over experimentally obtained corresponding color data sets.

We show that these sensors are still not optimally sharp. Using different computational approaches, we obtain sensors that are even more narrowband. In a first experiment, we derive a CAT by using spectral sharpening on Lam's corresponding color data set. The resulting Sharp CAT, which minimizes XYZ errors, performs as well as the current most popular CATs when tested on several corresponding color data sets and evaluating perceptual error. Designing a spherical sampling technique, we can indeed show that these CAT sensors are not unique, and that there exist a large number of sensors that perform just as well as CAT02, the chromatic adaptation transform used in CIECAM02 and the ICC color management framework. We speculate that in order to make a final decision on a single CAT, we should consider secondary factors, such as their applicability in a color imaging workflow. We show that sharp sensors are very appropriate for color encodings, as they provide excellent gamut coverage and hue constancy. Finally, we derive sensors for a CAT that provide stable color ratios over different illuminants, i.e. that only model physical responses, which still can predict experimentally obtained appearance data. The resulting sensors are sharp.

# Contents

# List of Figures

# List of Tables

# Publications

1. G.D. Finlayson and S. Süsstrunk, Spectral Sharpening and the Bradford Transform, *Proceedings of the Colour Imaging Symposium (CIS) 2000*, University of Derby, Colour Institute, pp. 236-243, 2000.

2. G.D. Finlayson and S. Süsstrunk, Performance of a Chromatic Adaptation Transform based on Spectral Sharpening, *Proceedings of IS&T/SID 8th Color Imaging Conference*, pp. 49-55, 2000.

3. S. Süsstrunk, J. Holm, and G.D. Finlayson, Chromatic Adaptation Behavior of Different RGB Sensors, *Proceedings of IS&T/SPIE Electronic Imaging 2001*, SPIE Vol. 4300, pp. 172-183, 2001.

4. G.D. Finlayson and S. Süsstrunk, Spherical Sampling and Color Transformations, *Proceedings of IS&T/SID 9th Color Imaging Conference*, pp. 321-325, 2001.

5. G.D. Finlayson and S. Süsstrunk, Color Ratios and Chromatic Adaptation, *Proceedings of IS&T First European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pp. 7-10, 2002.

6. S. Süsstrunk, C. Fredembach, J. Holm, and G.D. Finlayson, White-point independent RGB Primaries for Color Image Encodings, *Proceedings of the International Symposium on Imaging Science (ICIS 02)*, pp. 421-422, 2002.

7. G.D. Finlayson and S. Süsstrunk, Optimization for Hue Constant RGB Sensors, *Proceedings of IS&T/SID 10th Color Imaging Conference*, pp. 343-348, 2002.

8. S. Süsstrunk and G.D. Finlayson, Evaluating Chromatic Adaptation Transform Performance, accepted to *IS&T/SID 13th Color Imaging Conference*, November 2005.

Note that only my publications relevant to this thesis are listed here.

# Acknowledgements

First and foremost, my immense gratitude goes to my supervisor, Professor Graham Finlayson, who took me on as a part-time student for this adventure to the ”knowledge frontier.” Besides being a brilliant scholar, Graham is also a very nice man. I thoroughly enjoy our discussions about color, imaging, math, politics, education, and life. It is (*not* past tense) a great educational and personal experience working with Graham.

Many thanks also to my co-supervisor, Dr Steven Hordley. A great guy and very patient man, who after several times of “I don’t get it, explain it to me again,” still took me out for dinner. Dedicated to the last, he read a draft of my thesis on a Pacific beach and one just before mountain climbing. And I will never (let him) forget the Norwich walking tour.

I wish to thank my two examiners, Dr Michael H. Brill and Dr Mark Fisher, whose thoughtful comments certainly improved the clarity (and grammar) of this manuscript.

Many thanks also to the wonderful (past and current) members of the Colour Group at UEA, who kept me company during numerous lunches, dinners, and pub visits, as well as squash games and long walks. In alphabetical order: Dr Ali, Clement, Betty, Dr Gerald, Julia, Dr Jeff, Dr Peter and Dr Ruixia. They certainly made me often wish that I could stay longer in Norwich than just a few weeks every year.

Turning to Switzerland and EPFL, where I spent most of my working hours in the past six years, I am very grateful to Jocelyne Plantefol and Professor Martin Vetterli for their support (which means that they never quite let me forget that I needed to get this done...). In the (mistaken) belief that writing this thesis can compete with a spectacular view of Lake Geneva, Martin even let me use his house for three months.

Many, many thanks also to my family. Their unwavering support for whatever I choose to do is a constant in my life. I will never forget my Dad's words, so many years ago, when he told me: "I don't really understand what you do, but I am proud of you." Through the years, they have always encouraged me to follow my heart, with the consequence that I was not often around. They are great, I love them, and thus I usually spare them my cooking.

Not that they did not try to influence my decisions once in a while...as my late brother Christian once said, I would be an idiot not to take the EPFL position. So I did, and the rest is history. This thesis is dedicated to him.

I wish to acknowledge the late Professor Franz Tomamichel and his wife Dr Marianne Schönenberger Tomamichel, who were my first mentors before I realized what a mentor is. Without their encouragement and support, my path would certainly have taken another route.

And finally, my thanks and gratitude to all my friends on both sides of the Atlantic. David Grayson said it best: "Friendship is neither a formality nor a mode: it is rather a life."

# Chapter 1

# Introduction

*Chromatic Adaptation* is the ability of the human visual system to approximately preserve the appearance of an object color, no matter which light "color" illuminates a particular scene. We perceive a white piece of paper as white, regardless whether viewed under a bluish daylight or yellowish tungsten light source. Similar effects can be observed with other color hues.

Figure 1.1 illustrates this visual phenomenon. The left image is the same image as the middle image, except that it is overlayed by a bluish filter. While the image seems to have a color cast, i.e. we perceive all the colors to be too blue, we still can clearly identify the original hues. For example, the helmet still appears yellowish. In the right image, the filter was just overlayed on the helmet, which now appears greenish.

Figure 1.1: An example of chromatic adaptation. See text for explanation.

Image capturing systems, such as scanners and digital cameras, do not have the ability to adapt to an illumination source like the human visual system, as illustrated in Figure 1.2. The left image shows an object, the Macbeth color rendition chart [MMD76], captured under a tungsten light source. When we view an image, we adapt to the prelevant illuminant inherent in the viewing conditions. The image will appear to have a color cast if we are not adapted to the same light source the object was illuminated with when the image was taken. The right image was transformed to the viewing illuminant we usually adapt to when we view an image on a CRT monitor and appears more natural to us.

Figure 1.2: Left: An object imaged under a tungsten illuminant. Right: The same image transformed to appear correctly under monitor viewing conditions.

If the two helmets in the left and middle images of Figure 1.1 looked identical, we would have perfect *color constancy*. That is, we would be able to totally discount the effect of the blue filter, and the appearance of the helmet would depend only on its surface characteristics. However, as illustrated, that is not always the case. Indeed, the level of color constancy changes according to actual surface color, the viewing conditions, the scene composition, and the adaptation period. The appearance of a color under two different illuminants can therefore not simply be predicted by calculating only illumination change.

Figure 1.3 graphically illustrates this concept.[1] A surface color under one adapting illuminant can be described with coordinates $(ABC)^1$ in a coordinate system that is based on some three-dimensional sensor space. Changing the illuminant will result in coordinates $(ABC)^2$. However, the *appearance* of the surface color under the

---

[1] This figure is similar to a corresponding color illustration of R.W.G. Hunt, which he used in his keynote presentation at the 8th Color Imaging Conference in Scottsdale, AZ, 2000.

second illuminant is best described with coordinates $(ABC)^3$. The *corresponding color* coordinates of the surface color are $(ABC)^1$ and $(ABC)^3$, the distance between $(ABC)^2$ and $(ABC)^3$ is an indicator of the *color inconstancy* of the given surface color.

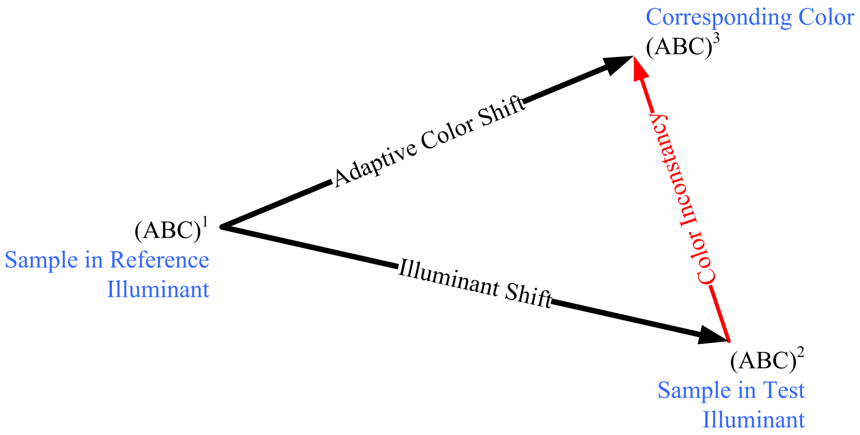


Figure 1.3: A graphical illustration of color constancy: If the human visual system is perfectly color constant, the effect of the illuminant could be discounted by simply modeling illumination change. However, most colors exhibit some degree of color inconstancy. The *corresponding color* of the sample under the reference illuminant $(ABC)^1$, i.e. the color coordinates that best describe its appearance under the test illuminant, are given by $(ABC)^3$.

To faithfully reproduce the appearance of image colors, it thus follows that all image processing systems need to apply a transform that converts the input colors captured under the input illuminant to the corresponding output colors under the output illuminant. This can be achieved by using a *chromatic adaptation transform* (CAT). Indeed, the prelevant color management framework used in color image applications, developed by the International Color Consortium (ICC) [ICC04], contains a chromatic adaptation transform. Similarly in color science, when the appearance of a color under a different illuminant needs to be predicted, a *color appearance model* [MFH$^+$02] containing a CAT is used. Basically, applying a chromatic adaptation transform to the color values under one adapting light source predicts the corresponding color values under another adapting light source.

In this thesis, we investigate different computational approaches to derive such chromatic adaptation transforms, which can be used in color science and color image processing. All our transforms are based on the von Kries model, which states that chromatic adaptation can be modeled as an independent gain control of three different sensor responses (see chapter 2). Mathematically, it can be implemented as a diagonal-matrix transform (DMT) that maps a color response matrix under one illuminant to a corresponding color response matrix under a different illuminant. Let $\mathbf{W}^a$ be the $(3 \times n)$ color response matrix under illuminant $a$ and $\mathbf{W}^b$ the $(3 \times n)$ color response matrix under illuminant $b$. $n$ is the number of color responses considered. With respect to Figure 1.3, $(\mathrm{ABC})^1$ is a column vector in $\mathbf{W}^a$, and $(\mathrm{ABC})^3$ is a column vector in $\mathbf{W}^b$. The von Kries chromatic adaptation model can be written as:

$$\mathbf{W}^b = \mathcal{D}^{b,a}\mathbf{W}^a \tag{1.1}$$

$\mathcal{D}^{b,a}$ is a diagonal transform that contains the scaling coefficients $g_1^{b,a}, g_2^{b,a}, g_3^{b,a}$, one for the response of each sensor:

$$\mathcal{D}^{b,a} = \begin{bmatrix} g_1^{b,a} & 0 & 0 \\ 0 & g_2^{b,a} & 0 \\ 0 & 0 & g_3^{b,a} \end{bmatrix}$$

Von Kries originally assumed that the color responses should correspond to cone responses, but subsequent research has found that a diagonal mapping in cone space is not accurate enough to predict corresponding colors under different illuminants. Thus, cone responses need first to be transformed to color responses in another sensor space before the scaling is applied. In modern CATs, this is done with a linear transform, and thus eq. 1.1 can be extended to:

$$\mathbf{W}^b = \mathbf{M}^{-1}\mathcal{D}^{b,a}\mathbf{M}\mathbf{W}^a \tag{1.2}$$

where $\mathbf{M}$ is a nonsingular (3x3) matrix linearly transforming cone responses to color responses in another sensor space. However, which sensor space (i.e. which sensor transformation matrix $\mathbf{M}$) is most appropriate is still under investigation, and we aim to provide some answers in this thesis.

In chapter 3, we derive a chromatic adaptation transform based on the *spectral sharpening* of Lam's corresponding color data [Lam85] and show that it performs just as well as other CAT's from the literature [Lam85, Fai01, LLRH02]. These CATs are all derived by minimizing perceptual error $\Delta E$ over one or more sets of corresponding color data. Spectral sharpening, on the other hand, is a mathematical approach that optimizes the transformation of these color response (XYZ) data sets, subject to a von Kries type DMT, by minimizing least-squares error [FDF94b].

There are two advantages to minimizing a physical error. First, the mathematics is more tractable and we arrive at an elegant closed form solution. Second, the problem of light change is physical and so, from an engineering perspective, we expect that the human visual system (HVS) must accomodate the scene physics. Thus, why not assume that the HVS can also deal with physics optimally?

The Sharp CAT that we derive uses sensors that are significantly different from the other CATs. The sensors are much "sharper," i.e. more narrow-band, than the cone sensitivities or the sensors used in the published CATs. Yet, we show that they perform statistically equivalently to the other CATs currently proposed in the literature and adopted in the color science and color imaging community.

The differences in sensor shape between those we derived for the Sharp CAT and those used in published CATs inspired us to investigate if there are other sensor spaces that might perform just as well. We develop a spherical sampling algorithm (see chapter 4) that lets us sample a three-dimensional space and find not just one, but all solutions for a given optimization criterion. We used the linear Bradford CAT [Lam85], the Sharp CAT derived in chapter 3, and CAT02 [MFH$^+$02], the chromatic adaptation transform used in the latest color appearance model CIECAM02 [MFH$^+$02] and the ICC color management specifications [ICC04] for performance comparison on Lam's corresponding color data. With our sampling technique, we found thousands of different sensor combinations, respectively thousands of different chromatic adaptation transforms, which performed equivalently based on a perceptual error criterion (RMS $\Delta E_{94}$). Evaluating statistical significance with respect to CAT02 still resulted in 1056 different CATs.

Finding such a large number of CATs that are similarly able to predict corresponding colors under different illuminants led us to investigate if other factors should be considered in the choice of the most appropriate chromatic adaptation transform, such as its relevance in a color image workflow. In chapter 5, we investigate if the different sensor combinations found in the previous chapter exhibit good gamut coverage and hue properties. If yes, these sensors could be used in color image encodings. These colorimetric RGB encodings are considered to be white-point independent, because the image RGB values are equal to post-adaptation sensor responses. Based on a gamut of surface colors [SSSS02], we found many sensor combinations that have better useful and encodable gamut coverage than ROMM RGB [ANS02b] and sRGB [IEC99] sensors. Evaluating those with respect to hue constancy, we found similar behavior compared to the other color sensors. We show that the best sensors in terms of chromatic adaptation, gamut coverage, and hue constancy are sharp.

In chapter 6, we derive CAT sensors that are based only on physical scene measurements and not psychophysically derived corresponding color data. Inspired by the retinex model of color vision that predicts that not absolute color appearance, but color response ratios calculated over scene colors remain constant, we derive sensors with the spherical sampling technique discussed in chapter 4 that optimally keep color ratios stable over a range of illuminants. We tested two reflectance data sets, Macbeth [MMD76] and Munsell [Mun76], under several illuminants ranging from A to D100. Comparing the corresponding CATs with CAT02 on Lam's corresponding color data, we found that their performance was statistically equivalent.

All the sensors we derive through our algorithms are thus much "sharper," i.e. have narrower shapes than the cone sensitivities or the sensors used in the published CATs. While we approach the problem of mapping corresponding colors from a mathematical perspective and thus our sensors are not based on visual experiments, they do seem to be psychophysically relevant and we refer to the appropriate literature throughout this thesis. Sharp sensors have also been found to be very effective in other color image processing tasks, such as color constancy algorithms, visual efficiency, and color image encodings. With our results, we thus present more evidence that sharp sensors are very appropriate when modeling visual tasks in machine vision systems.

## 1.1 Image Formation

Image formation is about a sensor's response to radiation. In this thesis, we are concerned only with the part of the radiation spectrum (approx. $360 - 780$ nm) that the human visual system is sensitive to. Equally, the sensors considered are sensitive only to such radiation.

We now introduce the basic models, sensors, and transforms of human vision, color science and color imaging related to the concepts described in the subsequent chapters. We start with an introduction into the notations and equations of the physical image formation model, followed by a discussion of the sensors associated with the human visual system: color matching functions, cone fundamentals, and opponent color sensitivities. The last section summarizes the different standards and recommendations pertaining to Colorimetry, i.e. the part of color science that deals with the measurement of physically defined color stimuli and their numerical representation.

### 1.1.1 Illuminants

Visible radiation, i.e. "light," can be characterized by its *Spectral Power Distribution* (SPD) $E(\lambda)$. An illuminant SPD denotes the radiant power given at each wavelength per wavelength interval of the visible spectrum, and can be measured with a radiometer. When modeling the human visual or an imaging system, however, we are often more interested in relative than absolute responses. Consequently, it suffices to use *relative* spectral power distributions that are normalized to 100 at 560 nm [WS82].

The *Commission International de L'Eclairage* (CIE) recommendations for *colorimetry* [CIE86] specify relative spectral power distributions of typical phases of daylight illuminants. These illuminant SPDs are calculated according to a method proposed by Judd et al. [JMW64], based on the mean and first two principal components of a series of daylight measurements.

It is common usage to call these illuminants by the letter "D" for daylight, and the first two numbers of the corresponding correlated color temperature. The term correlated color temperature is defined as the temperature of a blackbody (Planckian) radiator whose perceived color most closely resembles that of the given selective radiator at the same brightness and under specified viewing conditions [WS82]. For example, D55 is an illuminant calculated according to the method proposed in [JMW64], respectively [CIE86], with a correlated color temperature of 5,500 Kelvin. Figure 1.4 illustrates different daylight illuminants, ranging from D45 to D75.

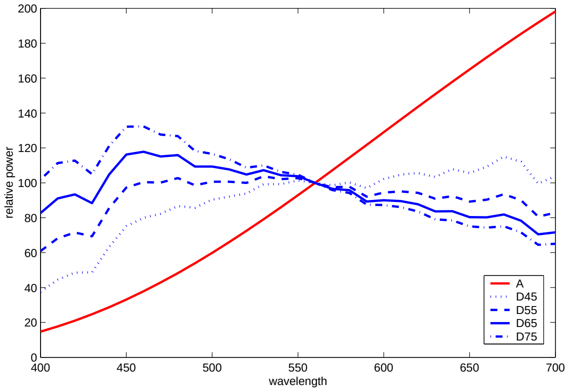


Figure 1.4: Relative spectral power distributions (SPD) of CIE daylight illuminants and standard colorimetry illuminants A and D65.

For the purpose of colorimetry (see section 1.3), the *International Organization for Standardization* (ISO) and the CIE have standardized two specific illuminants, D65 and A [ISO98]. D65 is a daylight illuminant with a correlated color temperature of 6,500 Kelvin, A is a tungsten-filament illuminant whose relative SPD is that of a Planckian radiator at a temperature of 2,856 K. With reference back to Figure 1.2, the top image was taken under a light source that approximated illuminant A, and rendered to illuminant D65 (bottom image).

### 1.1.2 Reflectance

The color of a (non-transparent) object is characterized by its *surface reflectance* $S(\lambda)$, ($S(\lambda) \in [0,1]$). For each wavelength or wavelength interval, a reflectance factor indicates how much of the incoming radiation is reflected. $S(\lambda) = 1$ means that all incoming radiation is reflected, while $S(\lambda) = 0$ indicates that all incoming radiation is absorbed. Figure 1.5 illustrates three reflectances of the Macbeth color checker rendition chart [MMD76].

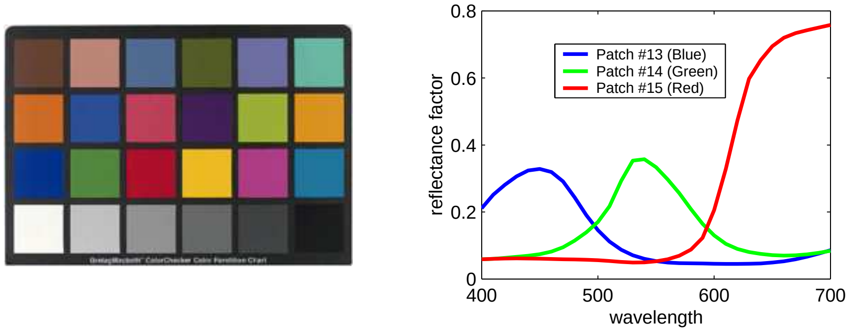


Figure 1.5: Left: the Macbeth ColorChecker. Right: reflectance spectra of color patches 13 (blue), 14 (green), and 15 (red).

### 1.1.3 Sensors

For the purpose of this thesis, a sensor is an entity that reacts to light. The term is also used to denote the sensitivity function $R(\lambda)$ that indicates the sensor's responsiveness to radiation at a given wavelength per wavelength interval. In general, sensors are considered to be physical entities, such as CCD or CMOS photo sites or the cones in the human retina (see section 1.2), which physically exist and give a positive response when radiation is detected. We assume throughout this document, however, that sensors do not need to be physical and will also allow sensors that have negative sensitivities. From a conceptual point of view, these sensors are the result of some "processing," either by the human visual system or an imaging system. For an extended discussion on HVS sensors, see section 1.2.

### 1.1.4 Physical Image Formation

In case of imaging an object, the radiation falling on the sensor is the product of the reflectance and relative illuminant SPD,[2] which is called the *color signal* or *color stimulus* $C(\lambda)$. The color response $\rho_k$ of a sensor $k$ with sensitivity $R_k(\lambda)$ at spatial position $\mathbf{x}$ can therefore be expressed as:

$$\rho_k(\mathbf{x}) = \int_\omega C(\mathbf{x}, \lambda) R_k(\lambda) d\lambda \tag{1.3}$$

or

$$\rho_k(\mathbf{x}) = \int_\omega S(\mathbf{x}, \lambda) E(\mathbf{x}, \lambda) R_k(\lambda) d\lambda \tag{1.4}$$

where $\omega$ indicates the visible spectrum. For the purpose of this thesis, $R_k(\lambda)$, $C(\mathbf{x}, \lambda)$, $E(\mathbf{x}, \lambda)$ and $S(\mathbf{x}, \lambda)$, can be adequately represented by samples taken at $\Delta\lambda = 10$nm intervals over the spectral range of 400-700 nm [SSS92]. The integral in eq. (1.3) can thus be replaced by summation:

$$\rho_k(\mathbf{x}) = \nu \sum_{\lambda=1}^{31} C(\mathbf{x}, \lambda) R_k(\lambda) = \nu \sum_{\lambda=1}^{31} S(\mathbf{x}, \lambda) E(\mathbf{x}, \lambda) R_k(\lambda) \tag{1.5}$$

where $\nu$ is a normalization factor, which usually either normalizes the response $\rho_k(\mathbf{x})$ depending on the number of samples used ($\nu = \frac{1}{31}$), or ensures that $\max(\rho_k(\mathbf{x})) = 1$. We will assume from now on that such a normalization factor is used, and not note it explicitly.

Using algebraic notations, color signal $C(\mathbf{x}, \lambda)$, reflectance $S(\mathbf{x}, \lambda)$, illumination $E(\mathbf{x}, \lambda)$ and sensor sensitivity $R_k(\lambda)$ can thus be expressed as $31 \times 1$ vectors $\mathbf{c}_\mathbf{x}$, $\mathbf{s}_\mathbf{x}$, $\mathbf{e}_\mathbf{x}$, and $\mathbf{r}_k$, respectively. Eq. (1.5) becomes:

$$\rho_k(\mathbf{x}) = \mathbf{c}_\mathbf{x}^T \mathbf{r}_k = \mathbf{s}_\mathbf{x}^T \text{diag}(\mathbf{e}_\mathbf{x}) \mathbf{r}_k \tag{1.6}$$

where $^T$ is the transpose and "diag" is an operator that turns $\mathbf{e}_\mathbf{x}$ into a diagonal matrix:

$$\text{diag}(\mathbf{e}_\mathbf{x}) = \begin{bmatrix} e_{1,\mathbf{x}} & 0 & \cdots & \cdots & 0 \\ 0 & e_{2,\mathbf{x}} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & e_{31,\mathbf{x}} \end{bmatrix}$$

---

[2] Note that in case of imaging "light," the reflectance factor $S(\lambda) = 1$ and $C(\lambda) = E(\lambda)$.

Any physical sensor can have a number of filters (or channels) with different sensitivities $R_k(\lambda)$. For the human visual system and trichromatic imaging systems, $k = 1, 2, 3$. Thus, the total color response at position $\mathbf{x}$ of a system with three sensors is a vector with three entries: $\boldsymbol{\rho}(\mathbf{x}) = [\rho_1(\mathbf{x}), \rho_2(\mathbf{x}), \rho_3(\mathbf{x})]^T$. In general, the letters $R, G, B$ will be used for color responses with sensors that have their peak sensitivities in the red (long), green (medium), or blue (short) wavelength part of the visible spectrum, respectively, $X, Y, Z$ when the sensors correspond to the CIE color matching functions, and $L, M, S$ when the sensors correspond to cone fundamentals (see section 1.2). The position index $\mathbf{x}$ will be used only when necessary to distinguish between two spatially different sensor responses.

The physical image formation model of eqs. (1.3-1.6) is a simplified model, as it does not take into account any physical illuminant, surface or sensor properties. In fact, it is applicable only to a “flat world” with no shadows, a single source illuminant, no surface reflectance interactions, and Lambertian surfaces that reflect incoming light equally in all directions.[3] However, this model is sufficient for the purpose of the simulations described in the subsequent chapters. Many of the corresponding color data sets we use for our computations (chapter 2), as well as the chromatic adaptation transforms derived based on stable color ratios (chapter 6) assume this kind of simplified model. For a discussion on extended image formation models, see [Hur98, Hor99, Mal01].

## 1.2 Sensors of the Human Visual System

For the purpose of modeling the human visual system (HVS), several different sensor sensitivities can be considered, derived from physiological and/or psychophysical properties of the human visual system. It will be evident from the subsequent discussion that several “sensors” are active in the visual processing chain. The “sensors” that we will derive in the subsequent chapters, namely those that allow us to model chromatic adaptation as a simple gain control, are not among them.

---

[3] Such scene arrangements are usually called *Mondrians*, so named by Edwin Land to describe his experimental set-up that resembled the paintings of Dutch artist Piet Mondrian.

However, some characteristics of these sensors can also be found in human vision models, as explained later.

Physiologically, the photoreceptors or light sensitive elements of the human visual system are the *rods* and *cones*. They contain light-sensitive photo-chemicals, converting light quanta into an electrical potential that is transmitted to the brain in a chain of neural interactions. They are located in the retina, the curved surface at the back of the eye (see Figure 1.6). The names are derived from their typical shape.

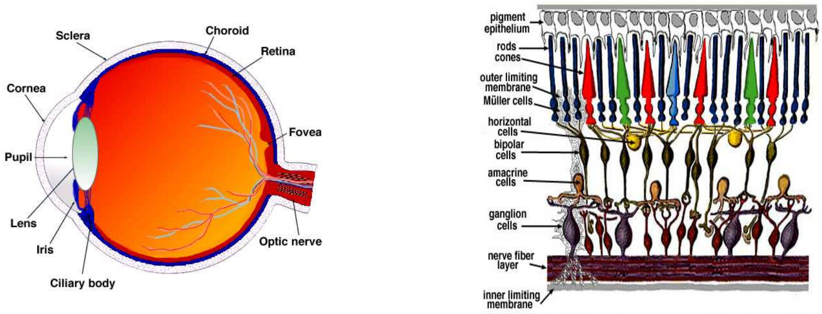


Figure 1.6: Left: cross-section of the human eye. Right: a close-up of the retinal cell layers (note that the light enters from the bottom). Both illustrations are taken from [KFN01].

Rods and cones are not uniformly distributed in the retina. Cones are primarily concentrated in the *fovea*, the area of the retina at the end of the optical axis that covers approximately 2° of visual angle. Beyond 10°, there are almost no cones. Rods, on the other hand, are primarily located outside of the fovea [Ost35, CSP$^+$87].

There are three types of cones, called $L$ (long), $M$ (middle) and $S$ (short) for their relative spectral positions of their peak sensitivities (see Figure 1.10). There is only one type of rod. The activities of the rods and cones are driven by the overall luminance level of a stimulus. *Photopic vision* ($> 10\ cd/m^2$) refers to visual sensations when only the cones are active, *scotopic vision* ($< 0.01\ cd/m^2$) when only the rods are active. *Mesopic vision* refers to the luminance range where both cones and rods are active [HF86]. Consequently, the ability to differentiate color signals is possible only in photopic or mesopic vision conditions, and we will always assume these conditions for the subsequent simulations.

The cones are connected through several cell layers with the ganglion cells, which transport the visual information through the optic nerve to the visual cortex (see Figure 1.6). Horizontal cells make connections to both photoreceptors and bipolar cells. Each horizontal cell is also connected to its neighbors. The response of any given horizontal cell is determined by the spatially weighted average of the signals of the cells surrounding it. Bipolar cells receive input from photoreceptors and horizontal cells and then produce a signal proportional to the difference between the two. Bipolar cells respond to light stimulation with either hyperpolarization (OFF-center) or depolarization (ON-center). Information from the bipolar cell passes through the amacrine cell layer to the ganglion cells.

The response of ganglion cells is usually described with *receptive fields.* A visual receptive field is defined as the area of the retina that influences the response of a single cell. Ganglion cells can respond to color stimuli in one of two ways: *color opponent responses* (P-cells or midget cells, and K-cells or giant bistratified ganglion cells) and *luminance responses* (M-cells or parasol cells) [Lee01]. Color opponent receptive fields each have centers and surrounds with separate color opponent properties: red ($L$) center with green ($M$) surround, green ($M$) center with red ($L$) surround, and blue ($S$) center with yellow ($L + M$) surround. A positive input in the center will inhibit the surrounding area (ON-center, OFF-surround), and vice versa (OFF-center, ON-surround). The receptive fields overlap, i.e. cones in the retina usually provide input to several neighboring receptive fields. Similarly, luminance receptive fields have an antagonism between high (white) and low (black) luminance.[4]

The separation of color opponent and luminance receptive fields can also be extended to the *Lateral Geniculate Nucleus* (LCN) and the *primary visual cortex* (also called V1 or striate cortex) [DSKK58, DKL84, LH88, Hub95]. P-cells project to the parvocellular layers, M-cells to the magnocellular layers, and K-cells to the

---

[4] Note that there is still discussion about the influence of the $S$ cone to the luminance signal. The distribution of the $S$ cone in the retina is very sparse compared to $L$ and $M$, and color vision modeling often assumes that luminance depends only on the $L$ and $M$ cone responses [Boy96].

koniocellular layers. These three pathways differ in the type of visual information they are sensitive to. Cells in the M-layers are orientation and movement selective but lack color sensitivity. Cells in the P-layers and K-layers are either color or brightness selective, but do not react to orientation. Many of them are also considered to be double-opponency, i.e. they have spatial antagonism between the same cone mechanism acting on the center and surround of their receptive fields. For example, $L$ cones responding to a stimulus excite the cell and $L$ cones responding to its background inhibit the cell. This inhibition normalizes the response over space because it creates a difference between the responses of different retinal areas [Gou01]. One can hypothesize that the ability of the human visual system to disregard illuminant gradients and to preserve local contrast, as assumed by the lightness algorithms, is due to these double-opponent cells in the visual cortex.

Intuitively, it is evident that accurate sensitivities of the human visual system are difficult to establish. It is very complex to "measure" the sensitivity at the different processing stages in a living human, although several studies have measured cone photocurrents [BNS87, SS95, SS01a] or neural responses at different stages of visual processing in animals [DSKK58, DKL84, Hub95]. The most common approach is to establish psychophysical testing procedures that isolate the responses of interest from other responses. In visual psychophysics, different stimuli pertaining to a HVS model are presented to human observers, and the responses of the observers are statistically evaluated. The following sections shortly describe two models of the human visual system, their corresponding psychophysical studies, and the resulting sensitivities that are used today in color science, color imaging, and human visual system modeling. Note that they do correspond to physiological mechanisms, as outlined above.

### 1.2.1 Trichromatic Theory of Color Vision

The *trichromatic* theory of color vision, also referred to as the *Young-Helmholtz three component theory* [You70, vH62], assumes that the signals generated in the three cone types, which are independent and have different spectral sensitivities, are transmitted directly to the brain where "color sensations" are experienced that correlate in a simple and direct way to the three cone signals. This theory has been found to hold in a series of *color matching* experiments (for an overview, see [WS82]). The experimental laws of color matching assume that for a given observation condition, test color stimulus $C^t(\lambda)$ can be matched completely by an additive mixture of three fixed primary stimuli $C^r(\lambda), C^g(\lambda), C^b(\lambda)$ with adjustable radiant power:

$$C^t(\lambda) = RC^r(\lambda) + GC^g(\lambda) + BC^b(\lambda) \tag{1.7}$$

$R, G, B$ are the relative intensities of $C^r(\lambda), C^g(\lambda), C^b(\lambda)$, respectively, and are called the *tristimulus values* of $C^t(\lambda)$. Any set of primaries can be used, as long as none of the primaries can be color matched with a mixture of the other two.

The results of color matches obey certain linearity laws, as first formulated by Grassman in 1853 [WS82]. If $C_1(\lambda)$, $C_2(\lambda)$, $C_3(\lambda)$, and $C_4(\lambda)$ are color stimuli and the symbol $\equiv$ has the meaning of "visual match," then:

- Symmetry Law: if $C_1(\lambda) \equiv C_2(\lambda)$, then $C_2(\lambda) \equiv C_1(\lambda)$
- Transitivity Law: if $C_1(\lambda) \equiv C_2(\lambda)$ and $C_2(\lambda) \equiv C_3(\lambda)$, then $C_1(\lambda) \equiv C_3(\lambda)$
- Proportionality Law: if $C_1(\lambda) \equiv C_2(\lambda)$, then $\alpha C_1(\lambda) \equiv \alpha C_2(\lambda)$, where $\alpha$ is a positive factor that increases or reduces the radiant power of the color stimulus while its relative SPD remains the same.
- Additivity Law: if $C_1(\lambda) \equiv C_2(\lambda)$ and $C_3(\lambda) \equiv C_4(\lambda)$, then $(C_1(\lambda) + C_3(\lambda)) \equiv (C_2(\lambda) + C_4(\lambda))$

These *generalized laws of trichromacy* ignore the dependence of color matches on the observational conditions, such as different radiant power, viewing eccentricity, stimulus surround, and adaptation to previous stimuli. To control viewing conditions, color matching experiments are therefore usually done with a *colorimeter.* A colorimeter is a device with a partitioned viewing area, where one half displays the reference color stimulus and the other half the mixture of the three primaries that can be adjusted by the observer to match the reference (see Figure 1.7).

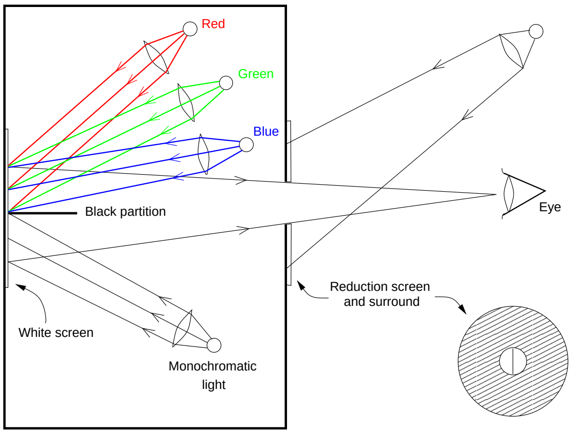


Figure 1.7: Typical colorimeter set-up for color matching experiments.

### Color Matching Functions

In 1931, the CIE (Commission Internationale de l'Eclairage) standardized a set of *Color Matching Functions* (CMFs) based on color matching experiments by Wright and Guild [WS82] using a colorimeter with a 2° bipartite field. Assuming that additivity holds and the luminous efficiency function $V(\lambda)$ of the HVS is a linear combination of the CMFs, they established a set of $\overline{r}$, $\overline{g}$, $\overline{b}$ color matching functions with "real" red $C^r(\lambda)$ ($\lambda = 700$ nm), green $C^g(\lambda)$ ($\lambda = 546.1$ nm), and blue $C^b(\lambda)$ ($\lambda = 435.8$ nm) monochromatic primaries based on the chromaticity coordinates of their experimental primaries. These $\overline{r}$, $\overline{g}$, $\overline{b}$ CMFs [CIE86] illustrate the relative amount $R, G, B$ of primaries $C^r$, $C^g$, and $C^b$ needed to additively mix a monochromatic light source at a given wavelength (see Figure 1.8).

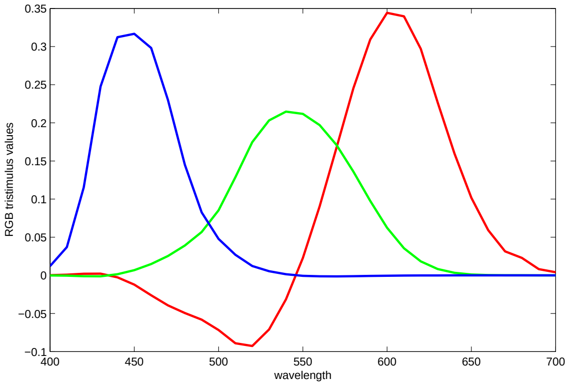


Figure 1.8: CIE 1931 2° $\overline{r}$, $\overline{g}$, $\overline{b}$ color matching functions.

The CIE additionally standardized a set of $\overline{x}$, $\overline{y}$, $\overline{z}$ color matching functions, based on imaginary $X, Y, Z$ primaries, which are a linear combination of the color matching functions derived from the original primaries. The transform was designed so that the $\overline{x}$, $\overline{y}$, $\overline{z}$ CMFs do not contain any negative values, primarily to design physical measuring devices, and that the $\overline{y}$ color matching function corresponds to $V(\lambda)$ [WS82, Hun98].

These $\overline{x}$, $\overline{y}$, $\overline{z}$ color matching functions are also called the *CIE 1931 standard observer* and are used for colorimetric calculations (see section 1.3) when the size of the stimulus does not extend 4° of visual angle. These CMFs are still an international standard today, even though Judd [Jud51] and later Vos [Vos78] proposed a modification based on a corrected luminous efficiency function, called $V_M(\lambda)$. The original $V(\lambda)$ of 1924 used in the derivation of the CIE 1931 CMFs underestimates the sensitivities at wavelength below 460 nm. Today, the color vision research community almost exclusively uses the Judd-Vos modified 2° CMFs [SS01b], while the color science and color imaging communities still use the original CIE 1931 2° CMFs (see Figure 1.9).

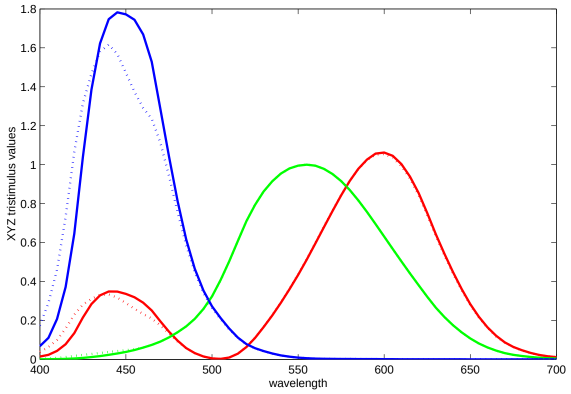


Figure 1.9: CIE 1931 2° $\overline{x}$, $\overline{y}$, $\overline{z}$ CMFs (solid line) and Judd-Vos modified 2° $\overline{x}$, $\overline{y}$, $\overline{z}$ CMFs (dotted line).

In 1964, the $\overline{x}_{10}$, $\overline{y}_{10}$, $\overline{z}_{10}$ color matching functions for the CIE 1964 supplementary standard observer were developed by Judd [WS82], based on experimental investigations by Stiles and Burch [SB59] and Speranskaya [Spe59] with stimuli sizes of 10°.

### Cone Fundamentals

The color matching functions described above are not cone sensitivities, i.e. absorption spectra of the cone pigments. They are based on color matching experiments, and their shape is determined by the choice of the primaries. However, if we assume that the basic principle of the trichromatic theory of color vision is correct, then cone responses also behave additively, and the cone sensitivities (also called *cone fundamentals*) are a linear combination of color matching functions (divided by intra-ocular medium absorption spectra). While this assumption is somewhat questionable when considering the complexity of the HVS, there are advantages when modeling visual processing. Physical measurements (see section 1.1.4) can easily be linearly transformed into cone responses.

Using color matching data and experimental data of color deficient observers, several sets of cone fundamentals that are linear combinations of either 2° or 10° CMFs were published (see [SS01b] for an overview and [SS03] for data). Vos and Walraven [VW71] and Smith and Pokorny [SP75] base their cone sensitivities on the modified 2° $X, Y, Z$ color matching functions. Stockman and Sharpe [SS00] base their 2° and 10° cone fundamentals on the Stiles and Burch 10° CMFs. Figure 1.10 illustrates the different 2° cone fundamentals.

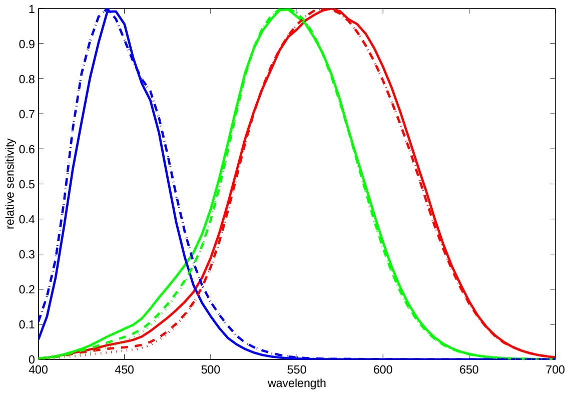


Figure 1.10: The normalized cone fundamentals of Stockman and Sharpe (solid line), Smith and Pokorny (dashed line), and Vos and Walraven (dotted line).

As can be seen in Figure 1.10, the $L$ and $M$ cone sensitivities are very correlated, i.e. their spectral distributions overlap significantly. Additionally, they have a broad base. From the point of view of quantum efficiency, broad-band sensors are able to capture more quanta and are thus overall more sensitive to radiation. However, from the point of view of coding efficiency, having two nearly identical sensors is inefficient as they both carry similar information [DB91]. As discussed below, the human visual system has found a way to de-correlate these sensor responses by its ability to encode the difference of the signals instead of the absolute responses.

### 1.2.2 Opponent Color Modulations

The theory of *opponent colors* is commonly attributed to Hering [Her78], although Goethe (1832) [vG91] previously discussed the concept. They both observed that certain colors are never perceived together, i.e. their names do not mix. We never see bluish-yellows or reddish-greens, where as bluish-greens (turquoise) and yellowish-reds (orange) are very common mixture descriptions. Hering also observed that there is a distinct pattern to the color of after-images. For example, if one looks at a unique red patch for a certain amount of time, and then switches to look at a homogeneous white area, one will perceive a green patch in the white area. Hering hypothesized that this antagonism between colors occurred in the retina, and that there are two major opponent classes of processing: spectrally opponent processes (red vs. green and yellow vs. blue) and spectrally non-opponent processes (black vs. white).

Experimental psychophysical support for Hering's theory was first provided by Jameson and Hurvich [JH55, HJ57]. They conducted a set of *hue-cancellation* experiments, where observers used monochromatic opponent light to "cancel" any hue that was not perceived as unique: red was canceled with green, blue was canceled with yellow, and vice versa. Repeating the experiment for all spectral lights and using the amount of opponent light needed as an indicator, they established opponent color curves over the visible spectrum.

Subsequent physiological experiments (see above, i.e. [DSKK58, DKL84, LH88, Hub95]) corroborated the presence of an opponent encoding mechanism in the human visual system. Additional psychophysical experiments [CKL75, WW82, SH88, PW93] have shown that such a representation correlates much better with experimental color discrimination and color appearance data than the additive theory of color vision. Figure 1.11 illustrates the opponent color responses Poirson and Wandell [PW93] derived from a color appearance experiment involving spatial patterns.

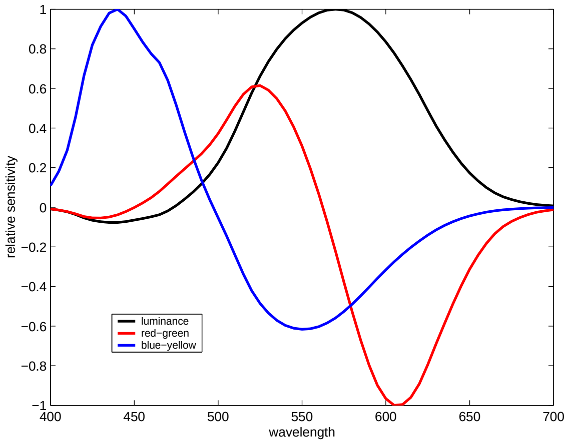


Figure 1.11: The normalized opponent sensitivities of Poirson and Wandell [PW93].

Luminance and opponent color sensitivities are considered to be orthogonal, and are generally modeled as a de-correlating transformation of cone fundamentals [EMG01]. Using the image formation model of eqs. 1.3-1.6, this assumption allows us to derive opponent color representations from physical measurements. The red-green opponent channel is usually a function of $L - M$, the blue-yellow channel of $(L + M) - S$, and luminance of $L + M$.

However, the color opponent responses are not directly related to quantum catches of the cones, due to the neural interactions in the retina. Opponency works on *cone contrast*, i.e. the relative cone responses compared to the environment. Contrast can either be taken into account with a Weber-type contrast function where the difference of stimulus and environmental (background) stimulus is normalized by the environmental stimulus. These contrast signals are then linearly transformed to opponent signals. DKL, an opponent representation used by the color vision community [DKL84, Bra96], is an example of such an encoding. In color science and computer vision, contrast is usually modeled by a logarithmic or power function. Color responses are normalized by the color response of the environment (a white surface or the illuminant) and then non-linearly encoded to account for lightness perception before being transformed to opponent signals (CIELAB, see section 1.3 and YCC, see Appendix D.1.)

## 1.3 CIE Colorimetry

Colorimetry is the part of color science that deals with the measurement of physically defined color stimuli and their numerical representation. The CIE [CIE78, CIE86, CIE95, ISO98] has published several standards and recommendations pertaining to colorimetry that are summarized below. However, note that the basic principles of colorimetry remain the same, regardless if standardized or modified color matching functions are used (see section 1.2.1).

A color response can be characterized by its relative tristimulus values $X, Y, Z$ according to the image formation model of eqs. 1.3-1.6, using physical measurements $E(\lambda)$ and $S(\lambda)$ of illuminant and surface reflectance, respectively. Sensors are either the color matching functions of the CIE 1931 standard observers (2°) or the CIE 1964 supplementary standard observer (10°), dependent on the stimuli size. The CIE $X, Y, Z$ tristimulus values follow the trichromatic color matching laws described in section 1.2.1. For example, two stimuli with equal specification will look the same when viewed by observers with normal color vision under identical observation conditions, i.e. they color match.

For the purpose of emphasizing *relative* magnitudes of the tristimulus values, which are related to color attributes, the $X, Y, Z$ tristimulus values are often normalized by dividing by the sum of their components:

$$\begin{aligned} x &= \tfrac{X}{X+Y+Z} \\ y &= \tfrac{Y}{X+Y+Z} \\ z &= \tfrac{Z}{X+Y+Z} = 1 - x - y \end{aligned} \tag{1.8}$$

The $x, y, z$ *chromaticity values* therefore represent the percentage of $X, Y, Z$ of a particular color response. $x, y$ chromaticity values are often used to graphically represent tristimulus values. For example, the $x, y$ chromaticity coordinates of the CIE 1931 standard observer spectral $X, Y, Z$ tristimulus values graphically indicate the two-dimensional *gamut* of the human visual system, i.e. the color coordinates that are visually achievable. In the case of CIE 1931 or CIE 1964 $\overline{r}$, $\overline{g}$, $\overline{b}$ and $\overline{x}$, $\overline{y}$, $\overline{z}$ CMFs, the color gamut boundary is called the *spectral locus*. The gamut of any RGB sensors whose transformation from RGB to XYZ is known can, of course, also be plotted.

When modeling visual or imaging systems and evaluating psychophysical experiments, it is often more useful to measure or predict the *difference* of color responses rather than their actual or relative values. The $X, Y, Z$ and $x, y, z$ color representations are not *perceptually uniform*, i.e. equal Euclidean distances do not equate to equal perceptual color differences [Mac43, Mac44]. In 1976, the CIE therefore published additional, more perceptually uniform representations to facilitate the interpretation of color differences: CIE $u^{'}$,$v^{'}$ chromaticity diagram, CIE $L^*, u^*, v^*$ (CIELUV), and CIE $L^*, a^*, b^*$ (CIELAB) [CIE78].

The CIE $u^{'}$,$v^{'}$ chromaticity values are derived from $X, Y, Z$ and $x, y$ as follows [Hun98]:

$$\begin{aligned} u^{'} &= \tfrac{4X}{X+15Y+3Z} = \tfrac{4x}{-2x+12y+3} \\ v^{'} &= \tfrac{9Y}{X+15Y+3Z} = \tfrac{9y}{-2x+12y+3} \end{aligned} \tag{1.9}$$

In the CIE $u^{'}$,$v^{'}$ chromaticity diagram, perceptually equal color differences result in (almost) equal Euclidean distances. Figure 1.12 shows a comparison of small color differences in the $x, y$ and $u^{'}$,$v^{'}$ chromaticity diagrams, based on experimental data of Wright [Wri41] and illustrated by Hunt [Hun98]. Each line is three times the length of a distance representing a *just noticeable difference* (JND) in a 2° field.

Both CIELAB and CIELUV are opponent color spaces, where $L^*$ represents the lightness of a color response, $a^*$ or $u^*$ its red-greenness, and $b^*$ or $v^*$ its yellow-blueness. The CIELUV system is commonly used for lighting and display, where as the CIELAB system is more often used for reflecting stimuli, although the CIE did not specify any preferred usage [Rob90]. As the color differences in the subsequent chapters are calculated based on CIELAB, the reader is referred to [WS82, CIE86, Hun98] for further information on CIELUV.

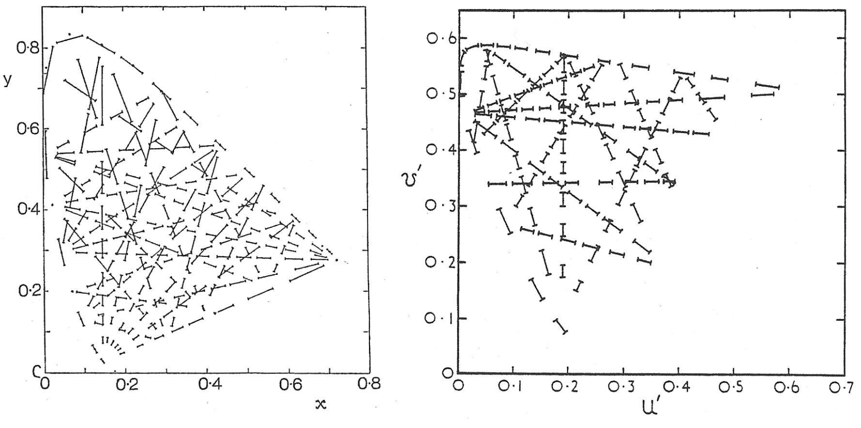


Figure 1.12: Equal color differences in the $x, y$ (left) and $u^{'}$,$v^{'}$ (right) chromaticity diagrams. The illustration is taken from [Hun98].

The transformation from $X, Y, Z$ tristimulus values to CIELAB values is as follows [Hun98]:

$$\begin{aligned} L^* &= 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \qquad \forall \frac{Y}{Y_n} > 0.008856 \\ L^* &= 903.3\frac{Y}{Y_n} \qquad \forall \frac{Y}{Y_n} \leq 0.008856 \\ a^* &= 500\left[\left(\frac{X}{X_n}\right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right] \\ b^* &= 200\left[\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}\right] \\ &\forall \frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} > 0.008856 \end{aligned} \tag{1.10}$$

where $X_n, Y_n, Z_n$ are the tristimulus values of the nominally white object-color stimulus, usually the illuminant normalized to $Y_n = 100$. If $\frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} \leq 0.008856$, then $\left(\frac{X}{X_n}\right)^{\frac{1}{3}}, \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}, \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}$ in eq. 1.10 are replaced with $7.787F + \frac{16}{116}$, where $F$ is $\frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n}$, respectively.

Color differences $\Delta E$ are then expressed as the Euclidean distance between the CIELAB coordinates:

$$\Delta E = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2} \tag{1.11}$$

or

$$\Delta E = \sqrt{(\Delta L)^2 + (\Delta H)^2 + (\Delta C)^2} \tag{1.12}$$

where $\Delta H$ is a measure of hue difference:

$$\Delta H = \sqrt{(\Delta a)^2 + (\Delta b)^2 - (\Delta C)^2}$$

The *hue angle* $h$ and *chroma* $C$ are defined as:

$$\begin{aligned} h &= arctan\left(\tfrac{b}{a}\right) \\ C &= \sqrt{a^2+b^2} \end{aligned} \tag{1.13}$$

In 1994, the CIE introduced a modified color difference formula, CIE $\Delta E_{94}$ [CIE95], which correlates better with visual perception of small color differences. It decreases the weights given to differences in $\Delta C$ and $\Delta H$ with increasing $C$. Equation (1.12) is modified as follows:

$$\Delta E_{94} = \sqrt{\left(\frac{\Delta L}{k_L S_L}\right)^2 + \left(\frac{\Delta C}{k_C S_C}\right)^2 + \left(\frac{\Delta H}{k_H S_H}\right)^2} \tag{1.14}$$

where $S_L = 1$, $S_C = 1 + 0.045\sqrt{C_1 C_2}$, $S_H = 1 + 0.015\sqrt{C_1 C_2}$ and $k_L = k_C = k_H = 1$. $C_1$ and $C_2$ refer to the chroma of the two color responses under consideration.

We use $\Delta E_{94}$ in the subsequent chapters as the preferred error metric.

## 1.4 Conclusions

In this chapter, the basic concepts in image formation, both from a computational and human visual system point of view, have been discussed. We will use the basic image formation model and the perceptual error metrics $\Delta E$ and $\Delta E_{94}$ in the calculations to derive the best chromatic adaptation transform. The HVS sensors encountered in the literature, derived either through physiological or psychophysical modeling, were introduced. While the "sharp" chromatic adaptation sensors that we will derive in the subsequent chapters, based on a simple von Kries scaling model, are not represented here, some physiological and psychophysical evidence exists that they are at least plausible.

In the next chapter, the literature related to chromatic adaptation is reviewed. We present the basic von Kries model, on which we base our chromatic adaptation transforms, and its parameters. Other models of chromatic adaptation, based on physiological models and psychophysical experiments are also introduced. Related physical models, mostly used in color constancy, are discussed as they are relevant for the experiments described in the subsequent chapters. We also summarize the experimentally derived color responses, i.e. corresponding colors, that describe the appearance of the same surface color viewed under two different illuminants.

# Chapter 2

# Chromatic Adaptation

Chromatic adaptation can be described as *the mechanism of the human visual system to discount the influence of the illuminant SPD and to approximately preserve the appearance of object colors* [Fai98]. Color constancy is achieved when *the appearance of an object is determined only by its surface spectral reflectance* [Mal01]. Thus, chromatic adaptation refers to a *process of the human visual system*, while color constancy refers to the *physical state* under which the process works. While we often assume that the visual system is "approximately color constant," given by our daily visual experiences, there are also instances when color constancy fails [LW96, BRK97]. Such an example is illustrated in Figures 1.1 and 1.3.

Chromatic adaptation research has therefore been primarily concerned with finding a chromatic adaptation model, based on physiological knowledge of the human visual system and psychophysical experiments. Chromatic adaptation models (or transforms) aim to predict the appearance of a surface color under different illuminant conditions. Color constancy algorithms, on the other hand, aim to recover the physical illuminant, reflectances and/or sensor characteristics necessary to achieve color constancy.

In this thesis, we aim to develop mathematical tools to model chromatic adaptation, i.e. to map color responses under a reference illuminant to color responses under a test illuminant, such that the resulting color response appears to come from the same object. We are not concerned with recovering illuminants or reflectances, as provided by color constancy algorithms. However, we base our results in part on color constancy research, which has provided elegant and closed form solutions for physical data, i.e. physical measurements of color responses under different illuminants. Our goal here is not to model psychophysical data, but to best account for scene physics and then to relate our algorithms to psychophysical data. We find that we are able to predict psychophysical measurements using our models.

In this chapter, we visit the chromatic adaptation and color constancy literature relevant to this thesis. We first briefly touch on how chromatic adaptation models are evaluated. Then, the original von Kries chromatic adaptation model is explained (section 2.2). Section 2.3 discusses how the the scaling coefficients of the von Kries model parameters were derived based on the experimental studies. Sections 2.4 and 2.5 then present the different chromatic adaptation transforms and models found in the literature. The discussion is restricted to "simple" von Kries linear and non-linear model extensions, we do not consider here two-stage adaptation models proposed by authors who studied color discrimination [HJ57, YSP93, WM94, AH01]. Section 2.6 discusses relational color constancy, which is based on the premise that the human visual system retains color ratios over changes in illuminant as opposed to absolute color measurements. Section 2.7 introduces spectral sharpening, a technique successfully used in color constancy algorithms that we will use for chromatic adaptation modeling. In section 2.8, we finally discuss the psychophysical studies that have resulted in experimental data (corresponding colors) used to investigate chromatic adaptation.

## 2.1 Performance Measure

In the following sections, we introduce the chromatic adaptation models that were derived based on different experimental methods, starting with the von Kries model. The goal of a chromatic adaptation model, equivalently a *chromatic adaptation transform* (CAT), is to predict as well as possible the corresponding color of a stimulus $\boldsymbol{\rho}^b$ under a test adaptation condition $\mathcal{A}^b$, given the experimental measurements $\boldsymbol{\rho}^a$ under a reference adaptation condition $\mathcal{A}^a$:

$$\boldsymbol{\rho}^b \approx \mathrm{CAT}(\boldsymbol{\rho}^a) \tag{2.1}$$

In other words, when applying a model or transform we obtain an estimate $\tilde{\boldsymbol{\rho}}^b$ of the corresponding colors under the test conditions:

$$\tilde{\boldsymbol{\rho}}^b = \mathrm{CAT}(\boldsymbol{\rho}^a) \tag{2.2}$$

The difference of $\tilde{\boldsymbol{\rho}}^b$ to $\boldsymbol{\rho}^b$ is an indicator of the appropriateness of a given model or transform. In general, the quality of a chromatic adaptation model or transform, i.e. the *prediction errors*, is evaluated using Euclidean distances in a perceptually uniform color space, either in CIELUV [BW92, Bäu95], CIE $u', v'$ [CW95, LW96, Bra98] or in CIELAB [Lam85, BBS97, Fai01, SHF01, LLRH02, MFH$^+$02]. We use CIELAB $\Delta E$ prediction error measures (see section 1.3) throughout this thesis.

## 2.2 Von Kries Chromatic Adaptation Model

*Johannes von Kries* is generally considered to be the first who attempted to describe the visual phenomenon of chromatic adaptation with a model. In [vK02], he investigates the additivity laws (see section 1.2) under different adaptation conditions and concludes that certain properties hold. He studied the effect of chromatic adaptation using an asymmetric matching experiment.

As illustrated in Figure 2.1, the basic concept of (simultaneous) asymmetric matching is to expose the observer to color stimuli under different adaptation conditions $\mathcal{A}^a$ and $\mathcal{A}^b$ [WS82]. The difference in adaptation conditions is usually limited to using a uniform, often neutral surround illuminated by two different illuminants, $E^a(\lambda)$ and $E^b(\lambda)$. The observer, using a colorimetric like device (see Figures 1.7) with two viewing areas matches under $\mathcal{A}^b$ the perceived color stimulus $C_1(\lambda)$ under $\mathcal{A}^a$. One eye is looking at the reference stimulus $C_1(\lambda)$ under $\mathcal{A}^a$, and the other eye is looking at a viewing area under $\mathcal{A}^b$, which contains a field with a mixture of red, green and blue primaries that are adjusted by the observer to match the reference stimulus. The resulting match is $C_2(\lambda)$. $C_1(\lambda)$ and $C_2(\lambda)$ are *corresponding color* signals.

Note that for such an experimental set-up, the color appearance mode is aperture color, i.e. primary lights are used to match a test light source. The experiment does not involve evaluating illuminant changes on surface reflectances.

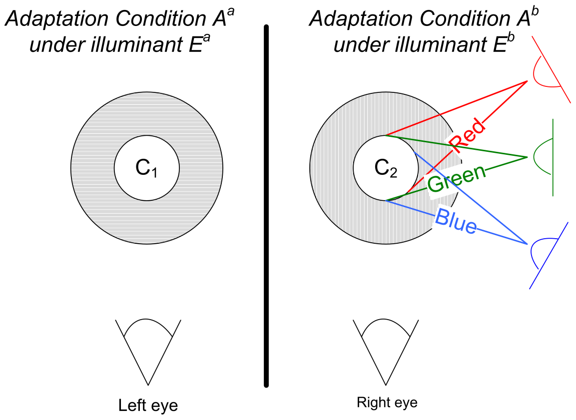


Figure 2.1: A simple asymmetric matching experiment. See text for explanation.

Von Kries observed that if $C_1(\lambda)\mathcal{A}^a$ under adaptation condition $\mathcal{A}^a$ has the same appearance as $C_2(\lambda)\mathcal{A}^b$ under adaptation condition $\mathcal{A}^b$, and equivalently for $C_3(\lambda)\mathcal{A}^a$ and $C_4(\lambda)\mathcal{A}^b$, then $(C_1(\lambda)+C_3(\lambda))\mathcal{A}^a$ has the same appearance as $(C_2(\lambda)+C_4(\lambda))\mathcal{A}^b$. Equally, $\alpha C_1(\lambda)\mathcal{A}^a$ matches $\alpha C_2(\lambda)\mathcal{A}^b$.

Comparing his *theorem of proportionality* [vK02] with the trichromacy laws (see section 1.2.1), von Kries deduced that additivity and proportionality hold for color responses independent of adaptation, if adaptation can be modeled by modifying the sensor sensitivities. He concluded that color changes caused by adaptation can be explained as a modification of the spectral sensitivities of the cone fundamentals, i.e. $\mathcal{A}^a \mapsto \{L^a, M^a, S^a\}$ and $\mathcal{A}^b \mapsto \{L^b, M^b, S^b\}$. He further assumes that the cone responses are independent from one another, each adapting exclusively to its own function. In [vK70], he states that the change of stimuli appearances can be modeled by a multiplicative coefficient that is a function of the higher or lower excitability of the individual cones. Von Kries does not explicitly state how the coefficients should be calculated, except to say that they are dependent on the adaptation condition. He also mentions that his model is a simplifications that might not hold for all conditions.

The von Kries's chromatic adaptation model can therefore be expressed as follows: the cone responses $L^a, M^a, S^a$ of a color stimulus under adaptation condition $\mathcal{A}^a$ can be mapped to its illuminant-invariant descriptors $d_1^o, d_2^o, d_3^o$ with three adaptation dependent gain factors $g_1^a, g_2^a, g_3^a$ that are independent for each channel.

Mathematically, it follows that:

$$\begin{aligned} d_1^o &= g_1^a L^a \\ d_2^o &= g_2^a M^a \\ d_3^o &= g_3^a S^a \end{aligned} \tag{2.3}$$

or, expressed as a *Diagonal Matrix Transformation* (DMT):

$$\begin{bmatrix} d_1^o \\ d_2^o \\ d_3^o \end{bmatrix} = \begin{bmatrix} g_1^a & 0 & 0 \\ 0 & g_2^a & 0 \\ 0 & 0 & g_3^a \end{bmatrix} \begin{bmatrix} L^a \\ M^a \\ S^a \end{bmatrix} \tag{2.4}$$

Equivalently, $L^a, M^a, S^a$ and $L^b, M^b, S^b$ of two color stimuli that appear identical when regarded under the two different adaptation conditions $\mathcal{A}^a$ and $\mathcal{A}^b$ are related by the ratio of their coefficients:

$$\begin{bmatrix} L^b \\ M^b \\ S^b \end{bmatrix} = \begin{bmatrix} \frac{g_1^a}{g_1^b} & 0 & 0 \\ 0 & \frac{g_2^a}{g_2^b} & 0 \\ 0 & 0 & \frac{g_3^a}{g_3^b} \end{bmatrix} \begin{bmatrix} L^a \\ M^a \\ S^a \end{bmatrix} \tag{2.5}$$

Von Kries assumes that his model is valid for cone responses. Thus, to use his model with any colorimetric measurements, i.e. tristimulus values XYZ (see section 1.3) or colorimetric RGB values (see Appendix D), eq. 2.4 can be extended to:

$$\begin{bmatrix} R^b \\ G^b \\ B^b \end{bmatrix} = \mathbf{M}^{-1}\mathcal{D}\mathbf{M} \begin{bmatrix} R^a \\ G^a \\ B^a \end{bmatrix} \tag{2.6}$$

where $\mathbf{M}$ is a nonsingular (3x3) matrix linearly transforming XYZ or RGB values to cone responses, $\mathbf{M}^{-1}$ its inverse and $\mathcal{D}$ the diagonal matrix containing the gain coefficients.

It is generally accepted today that a von Kries chromatic adaptation model is, at first approximation, able to model chromatic adaptation [BW92, Bäu94, Bäu95, CW95, Wue96, BBS97]. By defining all the parameters of the model described in eq. 2.6, it is possible to design a *chromatic adaptation transform* (CAT) that predicts color appearance under different illuminants.

However, there is no general agreement on how to define the specific parameters. Recall from eqs. 2.3-2.6, the mapping from color responses $\boldsymbol{\rho}^a = [R^a, G^a, B^a]^T$ under illuminant $E^a(\lambda)$ to illuminant independent descriptors $\mathbf{d^o} = [d_1^o, d_2^o, d_3^o]^T$ according to the von Kries model is as follows:

$$\mathbf{d^o} = \mathcal{D}\mathbf{M}\boldsymbol{\rho}^a \tag{2.7}$$

The parameters to define are the linear transform $\mathbf{M}$ and the scaling coefficients $g_1^a, g_2^a, g_2^a$ contained in the $3 \times 3$ diagonal matrix $\mathcal{D}$. Many chromatic adaptation transforms assume that the linear transform $\mathbf{M}$ is used to map from tristimulus values XYZ or colorimetric RGB values to cone responses. In other words, the sensors are broad band and are equal (or at least close) to cone fundamentals. While this makes sense from the point of view of human visual processing, we will show in chapters 3, 4, and 6 that other sensors might be more optimal. We delay the discussion of how $\mathbf{M}$ is derived to these chapters. Note that for the models discussed in the subsequent sections (2.3 - 2.5.1), the authors of these studies assume that the scaling is applicable in cone space if not otherwise noted.

## 2.3 Scaling Coefficients

As discussed in section 2.2, von Kries states that the scaling coefficients are dependent on the adaptation condition, but are *independent* for each visual channel. Since von Kries, many researchers have studied what these adaptation conditions $\mathcal{A}$ are. Depending on their experimental data, they have either based the chromatic adaptation model on illuminant information alone, or taken the surround of a stimulus into consideration. It is known that both have an effect on the color *appearance* of a stimulus, but are both adaptation phenomena, or is just the influence of the the illuminant important?

Ives [Ive12] early on hypothesized, without mentioning von Kries, that adaptation is a function of the illuminant “color.” Long exposure to one color decreases the cone sensitivity, while the other sensitivities increase. Land and McCann (see section 2.4.1) and West and Brill [WB82] follow this assumption by defining the scaling coefficients to be the inverse of the color response of a white patch within the scene. Thus, the illuminant independent descriptor $d_k^o(\mathbf{x})$ of a surface at position $\mathbf{x}$ with cone response $\rho_k^a(\mathbf{x})$ in channel $k$ under illuminant $E^a(\lambda)$ is calculated as follows:

$$d_k^o(\mathbf{x}) = \frac{1}{\rho_k^w}\rho_k^a(\mathbf{x}) \tag{2.8}$$

where $\rho_k^w$ is the cone response in channel $k$ of the white patch. If, for that surface, $S(\lambda) = 1$ is valid over the whole visible spectrum, the color response is dependent only on the SPD of the illuminant. This adaptation is often referred to as the von Kries-Ives model [Mal01] and is used in modern color science and color imaging CATs [Fai98, Fai01, CF01, ICC04, MFH$^+$02]. We will use these DMT coefficients when we derive new chromatic adaptation transforms in chapters 3, 4, and 6.

From a human visual processing point of view, however, this way of calculating the coefficients seems too simple. It is well known that the surround of a stimulus has a strong influence on its appearance, an effect known as *simultaneous contrast*. Thus, it seems unreasonable to neglect the effect of the surround. Helson [Hel34, Hel38] proposed that the coefficients should be calculated as the inverse of the average cone responses, independently for each cone class, of the visual field. Thus,

$$d_k^o(\mathbf{x}) = \frac{1}{\overline{\rho_k}}\rho_k^a(\mathbf{x}) \tag{2.9}$$

where $\overline{\rho_k}$ is the arithmetic mean of the color response in channel $k$. Note that if the average of the surround reflectances is neutral, then the difference in average cone response and the cone response of the adapting illuminant is a scale factor that cancels out if responses are mapped from one illuminant to another.[1]

Helson's hypothesis has been validated for simple scenes, such as found in simultaneous asymmetric matching experiments using a colorimeter with neutral and colored, but uniform surrounds [CW95, Wue96]. However, other studies have clearly indicated that the average of surround cannot alone account for chromatic adaptation. For example, Bäuml found that when the test stimulus was surrounded by other surfaces, calculating the scaling coefficients using the mean of the surrounding surface reflectances did not predict his experimental data, obtained by a achromatic matching experiment [Bäu94] or asymmetric matching experiment [Bäu95]. He concluded that the visual system's adjustment to illuminant changes does not depend

---

[1] In a later study using a memory matching experiment, Helson et al. [HJW52] use the von Kries-Ives model (i.e. illuminant-dependent coefficients) in their chromatic adapatation transform.

on the surface reflectances. Brainard [Bra98] found similar results in achromatic matching where the observers saw the stimuli in a three dimensional surround. Observers achieved good color constancy when the illuminant was varied, but not when the background surface against which the test surface is seen is changed.

Both the von Kries-Ives and von Kries-Helson model assume the coefficients are dependent only on signals from the same channel, and are not influenced by the responses of the other channels. In recent literature [BW92, DB00], this assumption is called the *strong* von Kries coefficient model (see Figure 2.2). Some studies, however, found better predictions if they relaxed the independence criterion and allowed for signals from other channels to influence the gain factors for one cone class. Several authors [WW82, Bäu95, DB00] found that there is a significant influence of L and M cone signals on the gain factor of the S cone signals. Similarly, Shevell and Humanski [SH88] showed that the red-green color appearance of a test light seen only by the L and M cones is modulated by changes in a background field that is visible only to the S cones.

Brainard and Wandell [BW92] therefore formulated a more general linear model, which they call the *weak* form of the von Kries coefficient model. As illustrated in Figure 2.2, they allow for the gain factors to be determined by signals from all three color channels. In their interpretation, the gain factors are only dependent on the illuminant's cone responses. More generally, the weak von Kries model can be expressed as:

$$d_k^o(\mathbf{x}) = \tfrac{1}{g_k^a}\rho_k^a(\mathbf{x}); \quad g_k^a = f(\rho_r^w, \rho_g^w, \rho_b^w) \tag{2.10}$$

where the scaling coefficient $g_k^a$ is a function of $\rho_r^w, \rho_g^w, \rho_b^w$, the color signals of the illuminant (see section 2.5.1).

In general, modern chromatic adaptation models and transforms thus assume that adaptation can be modeled by illuminant-dependent coefficients alone, either independent for each channel [Ive12, HJW52, Lam85, Fai01, LLRH02, MFH$^+$02], or dependent on all three color responses [BW92, LW93, Bäu95, LW96, BBS97, Bra98, DB00]. Other authors considered the surround and calculated the illuminant-independent descriptor based on a background factor [Hel34, Hel38, Mac63, NTS81, TSN84], or as a spatial interaction of the stimulus with many surround surfaces

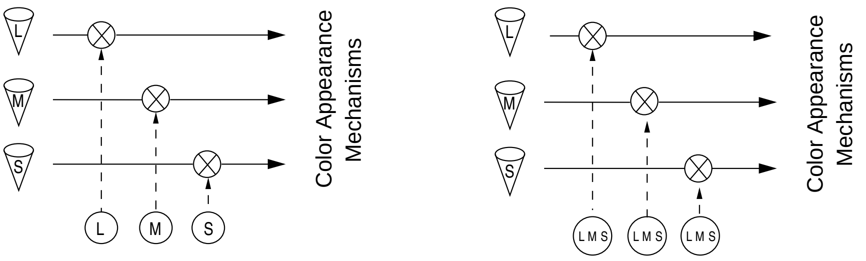


Figure 2.2: Strong and week von Kries coefficient models. Left: The *strong* von Kries model assumes that the gain is dependent only on signals from the same cone class. Right: The *weak* von Kries model assumes that the gain is dependent on signals from all cone classes. The illustrations are taken from [BW92].

[Lan77, Buc80, GJT88, FNC+97]. The corresponding chromatic adaptation models are discussed in the next sections.

The question posed in the first paragraph, i.e. if the scaling coefficients depend only on the illuminant or if the surround surface reflectances have to be taken into account, can probably be answered by considering the experimental conditions in use when studying chromatic adaptation. When using a haploscopic experimental set-up, the human visual system is in aperture color appearance mode and is able to isolate the effect of the surround without confounding it with the influence of the illuminant. In surface color appearance mode, when surface reflectances are evaluated under different illuminants, the surround can have an effect on appearance.

## 2.4 Strong von Kries Coefficient Model

The chromatic adaptation transforms used in color imaging applications to map from one adapting illuminant to another, such as in the color management framework developed by the International Color Consortium (ICC) [ICC04], are based on the strong von Kries-Ives coefficient model. The same applies to today's color appearance models [MFH+02]; when the appearance of a color under a different illuminant needs to be predicted, the chromatic adaptation transform applied is von Kries-Ives. However, the scaling is not applied to cone responses, but to modified cone responses. As mentioned above, we refer to chapter 3 for a more extensive discussion.

The theoretical limits of the strong von Kries coefficient model has primarily been studied in the context of *computational color constancy.* Computational color constancy algorithms either try to estimate the adapting illuminant from a scene and/or estimate surface reflectance [Buc80, MW86, For90, FHH01], to find appropriate sensors [FDF94b, FF96, DF00], or to render image appearance under different illuminants [Lan77, Lan86a, Hor74, Hur86, BB87].

Mathematically, the assumption that the exact illuminant SPD and spectral reflectances can be recovered from color responses infers that the product of a color response of an illuminant $\rho_k^E(\mathbf{x})$ and the color response of a surface $\rho_k^S(\mathbf{x})$ (under equi-energy illuminant, i.e. $E(\mathbf{x},\lambda) = 1$) are equal to the color response of their product:

$$\rho_k^{E,S}(\mathbf{x}) = \rho_k^E(\mathbf{x})\rho_k^S(\mathbf{x}) \tag{2.11}$$

However, as discussed by many authors [BW81, Lan83, Hur86, For90], this is true only when then sensor sensitivities are independent, implying very narrow-band filters that are in the extreme Delta functions responsive only at one wavelength. For example, if the sensor $k$ has non-zero response only at wavelength $\lambda_i$, eq. 1.3 can be written as:

$$\rho_k(\mathbf{x}) = \int_\omega S(\mathbf{x},\lambda)E(\mathbf{x},\lambda)\delta(\lambda_i)d\lambda = S(\mathbf{x},\lambda_i)E(\lambda_i) \tag{2.12}$$

In this case, a change in illuminant changes $\rho_k(\mathbf{x})$ only by a scale factor.

However, such sensor sensitivities apply neither for the human visual system, nor for any camera sensors that have quantum efficiency constraints. Therefore, additional assumptions about “the world” have been investigated to model color constancy. We will restrict the discussion to image appearance rendering, as we are not concerned in this thesis about reflectance or illuminant recovery. For an extended discussion on illuminant or surface reflectance recovery, see [Hur98, Hor99, Mal01].

### 2.4.1 Retinex

*Lightness algorithms* [Lan64, Lan77, Lan86b, Hor74, Hur86, BB87] presume that adaptation behaves according to eq. 2.3, i.e. that each receptor response is independently scaled by some scene-dependent factor. Edwin Land [Lan64, LM71, Lan74, Lan77, Lan83, Lan86a] pioneered the work on lightness algorithms with his *retinex* model of color vision. He observed during experiments with Mondrian scenes that the perceived colors of surface reflectances do not depend on the radiant energy of their stimuli, but that their *relative lightness rank-ordering* remained invariant independent of the illumination uniformity and color. For a review of his experiments and early algorithm, see Appendix A.

In general, the lightness value for a sensor $k$ is calculated by finding the average ratio between the color response of a position $\mathbf{x}_0$ and many surrounding positions. Early retinex algorithms calculated the ratios along a path [Lan64], while later implementations use a weighted average of the surround [Lan86a]. As discussed by McCann et al. [MMT76], there are some parallels to human visual processing. The retinex algorithm contains a local normalization due to the influence of the neighboring pixels, which could be related to the ganglion cell receptive fields. The spatial normalization over the whole visual field could be explained by the double-opponent fields found in the visual cortex.

Setting the paths' lengths and the number of paths to infinity is equivalent to normalizing the surface reflectance at position $\mathbf{x}_0$ by the geometric mean of all surface reflectances [BW86a]. As the triplet of computed lightness values for the three spectral channels should define the illuminant-independent descriptor of a color patch, the algorithm implicitly assumes then that the mean reflectance is the same in each channel and for every scene. This constraint, called *gray-world*, assumes that the (geometric) mean surface reflectance of each scene in each spectral channel is the same: gray. This assumption also appears in other color constancy algorithms [Buc80, DL86, GJT88], and has been used for many color reproduction algorithms [Hun95]. It is, in spirit, close to Helson's [Hel34] proposal that the von Kries coefficients $g_1^a, g_2^a, g_3^a$ are inversely proportional to the (arithmetic) average photoreceptor excitations within a cone class under adaptation condition $\mathcal{A}^a$.

The lightness values are not accurate illuminant-independent descriptors if the mean scene reflectance is not gray. Land and McCann [LM71, MMT76] therefore introduced a normalization into the path equation. To recover illuminant-independent lightness descriptors, this normalization assumes that the brightest patch in each channel is a perfect reflector, i.e. it reflects all illuminant energy. This *maxRGB* assumption is close to the von Kries-Ives model, which also predicts that the coefficient is based solely on the illuminant.

## 2.5 General Chromatic Adaptation Models

As mentioned above, the strong von Kries coefficient model is accurate only at first approximation in predicting chromatic adaptation. If a simple diagonal model mapping cone responses from one illuminant to another would suffice, then a plot of match versus test cone coordinates would always fall along a line through the origin: $g_k^b \rho_k^b = g_k^a \rho_k^a$. However, some authors observed substantial deviations. Either because a line does not adequately describe the data, or the best fitting line is offset from the origin. Thus, the von Kries model was extended to a 9 parameter coefficient model, with a power function to map observations non-linearly, by an additive factor, or some combination. We divide the discussion into two sections. The linear affine models that use a full 9 parameter coefficient model plus additive offset, which under some constraint can be reduced to a weak or even strong von Kries coefficient model, and the non-linear models.

### 2.5.1 Linear Chromatic Adaptation Models

The linear chromatic adaptation models are based on the assumption that surface reflectance and illuminant SPD can be approximated with a small number of weighted basis functions. Thus, $E(\lambda)$ and $S(\lambda)$ can be expressed as:

$$E(\lambda) = \sum_{i=1}^{m} \epsilon_i E_i(\lambda) \tag{2.13}$$

where $\epsilon_i$ are unique real numbers, and $E_i(\lambda)$ are the basis functions. Similarly, surface reflectances are expressed as:

$$S(\mathbf{x},\lambda) = \sum_{j=1}^{m} \sigma_j(\mathbf{x}) S_j(\lambda) \tag{2.14}$$

In matrix notation, illuminant $\mathbf{e}$ and surface reflectance $\mathbf{s}$ becomes:

$$\begin{aligned} \mathbf{e} &= \mathbf{B}_e \boldsymbol{\epsilon} \\ \mathbf{s} &= \mathbf{B}_s \boldsymbol{\sigma} \end{aligned} \tag{2.15}$$

The $n \times m$ matrix $\mathbf{B}_e$ contain the $m$ basis function for the illuminant, with $n$ samples over the visible spectrum. $\boldsymbol{\epsilon}$ is a $m \times 1$ vector containing the illuminant weights. Equivalently, the $n \times m$ matrix $\mathbf{B}_s$ contain the $m$ basis function for the surface reflectance and $\boldsymbol{\sigma}$ the surface weights. Using the image formation model of eq. 1.6, color responses can then be calculated as

$$\boldsymbol{\rho} = \mathbf{R}\mathrm{diag}(\mathbf{B}_e \boldsymbol{\epsilon}) \mathbf{B}_s \boldsymbol{\sigma} \tag{2.16}$$

where $\mathbf{R}$ is a $k \times n$ matrix containing the sensor sensitivities and $\boldsymbol{\rho}$ a $k$-dimensional vector containing the sensor responses.

Brainard and Wandell [BW92] proposed different linear chromatic adaptation models to predict asymmetric matching data collected through haploscopic experiments. The basic assumption is that the transform parameters are linear functions of the illuminant change. Let $\boldsymbol{\rho}^a$ be the $3 \times 1$ cone response vector of an object under standard illuminant $E^a(\lambda)$ and $\boldsymbol{\rho}^b$ the $3 \times 1$ cone response vector of the corresponding match under the test illuminant $E^b(\lambda)$. The change of responses can be calculated as $\Delta\boldsymbol{\rho} = \boldsymbol{\rho}^b - \boldsymbol{\rho}^a$. Let $\boldsymbol{\epsilon}^a$ be a $m \times 1$ vector containing the weights $\epsilon_i$ of the $m$ basis functions $E_i(\lambda)$ of the standard illuminant and $\boldsymbol{\epsilon}^b$ the equivalent vector of the test illuminant. The change in illuminant is given as $\Delta\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^b - \boldsymbol{\epsilon}^a$. The model predicts $\Delta\boldsymbol{\rho}$ as a function of $\boldsymbol{\rho}^a$ and $\Delta\boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon}^a$ fixed. Then, for any illuminant change $\Delta\boldsymbol{\epsilon}$, the change in cone responses can be calculated as:

$$\Delta\boldsymbol{\rho} = \mathbf{X}_{\Delta\boldsymbol{\epsilon}} + \mathbf{a}_{\Delta\boldsymbol{\epsilon}} \tag{2.17}$$

where $\mathbf{X}_{\Delta\boldsymbol{\epsilon}}$ is a $3 \times 3$ matrix that depends on $\Delta\boldsymbol{\epsilon}$ and $\mathbf{a}_{\Delta\boldsymbol{\epsilon}}$ a $3 \times 1$ vector also depending on $\Delta\boldsymbol{\epsilon}$. This general form of the model is called the *affine* model. Constraining the elements of vector $\mathbf{a}_{\Delta\boldsymbol{\epsilon}}$ to zero results in a *linear* model, and a *diagonal* model can be derived by constraining the off-diagonal terms of matrix $\mathbf{X}_{\Delta\boldsymbol{\epsilon}}$ to zero. This diagonal model resembles a weak von Kries model (see Figure 2.2), as it allows interaction of different channels on the scaling coefficients.

Bäuml [Bäu95], using a similar model, also found that the prediction error is minimized if the gain coefficients are calculated based on the input of all three channels, specifically on the cone responses of the illuminant. The three-dimensional vector $\mathbf{g}$ that contains the gain coefficient $g_k$ of the diagonal von Kries matrix are calculated with a $3 \times 3$ matrix $\mathbf{Y}$, such that

$$\mathbf{g} = \mathbf{Y}_{\boldsymbol{\epsilon}^b} \Delta\boldsymbol{\epsilon} \tag{2.18}$$

where $\mathbf{Y}_{\boldsymbol{\epsilon}^b}$ depends on the test illuminant, and $\Delta\boldsymbol{\epsilon}$ is a $3 \times 1$ vector. His model corresponds to a weak von Kries model of eq. 2.10.

Brainard et al. [BBS97] proposed an equivalent illuminant model based on the idea that the human visual system is capable of estimating the illuminant in a scene. If $\mathbf{B}_e$ and $\mathbf{B}_s$ are three dimensional models, the estimated surface reflectance weights $\hat{\boldsymbol{\sigma}}$ under the test illuminant can be calculated by inverting eq. 2.16.

$$\hat{\boldsymbol{\sigma}} = [\mathbf{R}\mathrm{diag}(\mathbf{B}_e \hat{\boldsymbol{\epsilon}}^a)\mathbf{B}_s]^{-1} \boldsymbol{\rho}^a \tag{2.19}$$

Under the matching conditions, the observer would estimate the illuminant to be $(\mathbf{B}_e \hat{\boldsymbol{\epsilon}}^b)$. Given these conditions, an observer would achieve color constancy if the surface reflectance, characterized by $\hat{\boldsymbol{\sigma}}$, has the same appearance under the illuminant $E^a(\lambda)$ and $E^b(\lambda)$ characterized by the illuminant weights $\hat{\boldsymbol{\epsilon}}^a$ and $\hat{\boldsymbol{\epsilon}}^b$, respectively. Thus,

$$\boldsymbol{\rho}^b = \mathbf{R}\mathrm{diag}(\mathbf{B}_e \hat{\boldsymbol{\epsilon}}^b)\mathbf{B}_s [\mathbf{R}\mathrm{diag}(\mathbf{B}_e \hat{\boldsymbol{\epsilon}}^a)\mathbf{B}_s]^{-1} \boldsymbol{\rho}^a \tag{2.20}$$

or

$$\boldsymbol{\rho}^b = \mathbf{Z}(\hat{\boldsymbol{\epsilon}}^a, \hat{\boldsymbol{\epsilon}}^b) \boldsymbol{\rho}^a \tag{2.21}$$

where $\mathbf{Z}(\hat{\boldsymbol{\epsilon}}^a, \hat{\boldsymbol{\epsilon}}^b)$ is a $3 \times 3$ matrix that depends only on the illuminant estimates $\hat{\boldsymbol{\epsilon}}^a$ and $\hat{\boldsymbol{\epsilon}}^b$, and not on the surface reflectance. The 9 parameter matrix $\mathbf{Z}$ is a linear chromatic adaptation model with six degrees of freedom: the 3 variables of $\hat{\boldsymbol{\epsilon}}^a$ and the 3 variables of $\hat{\boldsymbol{\epsilon}}^b$.

### 2.5.2 Non-linear Chromatic Adaptation Models

MacAdam [Mac61, Mac63], based on the results of monocular asymmetric matches, was the first to postulate a non-linear hypothesis to model chromatic adaptation. He found that the following relationship best described his experimental data and also those of Burnham et al. [BEN57] and Wassef [Was58, Was59]:

$$d_k^o(\mathbf{x}) = \alpha_k + (\beta_k \rho_k^a(\mathbf{x}))^{\gamma_k} \tag{2.22}$$

The variables $\alpha_k, \beta_k$ and $\gamma_k$ depend on the adaptation condition $\mathcal{A}^a$. He also tested the following logarithmic expression as alternative to the power function and found comparable results:

$$d_k^o(\mathbf{x}) = \alpha_k \log(\beta_k[\rho_k^a(\mathbf{x})] + \gamma_k) \tag{2.23}$$

Note that the variables $\alpha_k, \beta_k$ and $\gamma_k$ are calculated differently than in eq. 2.22. The same is also true for the variables in the following equation.

Power or logarithmic functions are a way of accounting for cone contrast (see section 1.2.2). This implies that the von Kries gain mechanism is active at a post-receptoral level, and thus chromatic adaptation is not solely a function of sensitivity regulation of the cone fundamentals. Therefore, the non-linear models, while still using a diagonal matrix transform, are not “true” von Kries models anymore in the sense of “just” modeling chromatic adaptation with a gain factor.

Nayatani et al. [NTS81, TSN84] also formulated a non-linear model, with

$$d_k^o(\mathbf{x}) = \alpha_k (\rho_k^{a*}(\mathbf{x}))^{\beta_k(\rho_k^a(\mathbf{x}_b))} \tag{2.24}$$

where

$$\rho_k^{a*}(\mathbf{x}) = \frac{\rho_k^a(\mathbf{x}) + \rho_{k,n}^a}{\rho_k^a(\mathbf{x}_b)} \tag{2.25}$$

$\rho_{k,n}^a$ is the noise component of the cone response mechanism, and $\rho_k^a(\mathbf{x}_b)$ is the response of the background.

Lam [Lam85], using the results of a memory matching experiment, simplified the MacAdam and Nayatani model by allowing only a non-linear power function in the blue visual channel, and neglecting the influence of the background:

$$d_b^o(\mathbf{x}) = \frac{1}{(\rho_b^w)^\beta} (\rho_k^a(\mathbf{x}))^\beta \tag{2.26}$$

where

$$\beta = (\rho_b^w)^{0.0834}$$

For the red and green channel, the illuminant independent descriptors are calculated according to the von Kries-Ives model of eq. 2.8. Such a model, i.e. only using a non-linearity in the blue channel, was first proposed by Bartleson [Bar79a, Bar79b]. To fit the model to his experimental data, Lam also derived response sensitivities (linear combination of color matching functions) that are slightly more de-correlated than cone fundamentals. The resulting *Bradford* chromatic adaptation transform was used in the color appearance model CIECAM97s [CIE98] and will be further discussed in chapter 3.

Similarly, Lucassen and Walraven [LW93] also found for haploscopic matching under different "colored" illuminants that a non-linear response function better predicts their experimental data for the S cone signals. Using our notation, their model can be expressed as:

$$d_b^o(\mathbf{x}) = (\rho_b^w)^{\beta} \log \left( 4.35 \frac{\rho_b^a(\mathbf{x})}{\rho_b^w} \right) \tag{2.27}$$

This model has some similarity to the retinex model discussed in section 2.4.1.

It is interesting that many experimental studies show a failure of the strong von Kries coefficient model in the S channel. However, as discussed above, different authors use different methods to correct for it, either by allowing an interaction of the L and M channel in the coefficient calculation, or by introducing a non-linearity.

## 2.6 Relational Color Constancy

*Relational* color constancy is defined as the invariance of *perceived relations* between the colors of surfaces under different illuminants, as opposed to "normal" color constancy where we assume that the perceived colors are invariant [FNC$^+$97]. In other words, the difference is in the invariance of color relation versus color percept. The human visual system encodes and maintains the *ratio* of photoreceptor excitations to different surfaces to achieve color constancy, as opposed to absolute values. These ratios are calculated within a cone type, and are independent of the other cone types. In concept, this is very much related to the retinex algorithm described above.

Foster and Nascimento [FN94] have shown that for a large class of pigmented surfaces and for surfaces with random spectral reflectances, cone excitation ratios are statistically almost invariant under changes in illumination. Lately, they have demonstrated that the ratio discrimination is indeed independent for each color channel and not solely based on luminance processing of the human visual system [NF00]. Other authors have also corroborated the importance of cone ratios to account for color constancy. Tiplitz Blackwell and Buchsbaum [BB88] found through experimentation that color constancy is almost perfect if the ratios of $L$, $M$, and $S$ of the stimulus center to the $L$, $M$, and $S$ of the surround remains constant as the illuminant changes. If the ratio changes, color constancy fails. Dannemiller [Dan93] studied the rank ordering of photon catches from natural objects illuminated with daylight and tungsten light for a model human fovea. He found that the observed rank orderings remained nearly stable across illuminant changes for all three cone classes. Brill and West [BW81] have used color ratios in theoretical studies to set constraints on illuminant and surface reflectance spectra for color constancy.

Foster et al. [FNC$^+$97] pointed out that stable cone ratios is a consequence of assuming a (weak) von Kries coefficient model. Assume that $\rho_k^a(\mathbf{x}_m)$ and $\rho_k^a(\mathbf{x}_n)$ are cone responses of cone type $k$ for surfaces $\mathbf{x}_m$ and $\mathbf{x}_n$ under illuminant $E^a(\lambda)$, and $d_k^o(\mathbf{x}_m)$ and $d_k^o(\mathbf{x}_n)$ its illuminant independent descriptors. Similarly, $\rho_k^b(\mathbf{x}_m)$ and $\rho_k^b(\mathbf{x}_n)$ are the cone responses for the same surfaces and cone type under illuminant $E^b(\lambda)$. According to the von Kries coefficient rule (eq. 2.3):

$$\begin{array}{ll} d_k^o(\mathbf{x}_m) = g_k^a \rho_k^a(\mathbf{x}_m) & d_k^o(\mathbf{x}_n) = g_k^a \rho_k^a(\mathbf{x}_n) \\ d_k^o(\mathbf{x}_m) = g_k^b \rho_k^b(\mathbf{x}_m) & d_k^o(\mathbf{x}_n) = g_k^b \rho_k^b(\mathbf{x}_n) \end{array} \tag{2.28}$$

where $g_k^a$ and $g_k^b$ are the illuminant dependent $(E^a(\lambda), E^b(\lambda))$ gain coefficients for channel $k$. By rearranging eq. 2.28, it becomes evident that cone response ratios remain invariant:

$$\frac{d_k^o(\mathbf{x}_m)}{d_k^o(\mathbf{x}_n)} = \frac{\rho_k^a(\mathbf{x}_m)}{\rho_k^a(\mathbf{x}_n)} = \frac{\rho_k^b(\mathbf{x}_m)}{\rho_k^b(\mathbf{x}_n)} \tag{2.29}$$

as $g_k^a$ and $g_k^b$ are canceled.

As mentioned in [FNC$^+$97], Hurlbert noted that these invariant ratios are not necessarily computed at immediate post-receptoral levels, they could be computed at some higher level in the visual system as part of a more general accommodation to the natural structure of the visual environment. Rinner and Gegenfurtner [RG99], in a study about the time course of adaptation, found that the mechanism for instantaneous adaptation is probably governed by long-range spatial interactions found in the visual cortex. We use this argument in chapter 6 to derive sensors for a chromatic adaptation transform that keep color ratios stable.

## 2.7 Spectral Sharpening

For the lightness algorithms, the general model based chromatic adaptation transforms, and relational color constancy research discussed in the previous three sections, the assumption is that the sensors are broadband and are (or resemble) the cone fundamentals of the human visual system. West and Brill [WB82, BW86b] established the theoretical limits that reflection functions can have assuming cone sensitivities. They found that certain reflection functions can satisfy the strong von Kries coefficient model of eq. 2.4, but that these reflection functions are restricted and do not resemble natural reflectances.

However, Finlayson, Drew, and Funt [FDF94b, FDF94a, Fin95, FF96, DF00] have shown that a *generalized coefficient model* can be successfully applied in color constancy. In their model, the sensors sensitivities are first transformed by a linear (3x3) matrix to a new set before the diagonal transform is applied. Equivalently, the transform can be directly applied to the color responses. With reference back to eq. 2.6, the generalized coefficient model can be written as

$$\begin{bmatrix} d_1^o \\ d_2^o \\ d_3^o \end{bmatrix} = \begin{bmatrix} g_1^a & & \\ & g_2^a & \\ & & g_3^a \end{bmatrix} \begin{bmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{2,2} & t_{3,3} \end{bmatrix} \begin{bmatrix} R^a \\ G^a \\ B^a \end{bmatrix} \tag{2.30}$$

Equivalently:

$$\mathbf{d^o} = \mathcal{D}\mathbf{T}\boldsymbol{\rho}^a \tag{2.31}$$

Note that the fundamental difference between eq. 2.6 and eq. 2.30 is that the von Kries model assumes adaptation is a function of cone responses, while Finlayson et al. do not put any restrictions on the sensors.

Inspired by the problem outlined in the introduction of section 2.4 that exact illuminant SPD and surface reflectances can be recovered only when the sensor sensitivities approach Delta functions (eq. 2.12), Finlayson et al. [FDF94b, Fin95] derived *sharpening* transforms that rendered the original sensors more narrow-band. They present two methods on how to achieve this, called *sensor-based* sharpening and *data-based* sharpening. We limit the discussion here to sensor based sharpening, as data-based sharpening is revisited in chapter 3.

### 2.7.1 Sensor-based Sharpening

Sensor-based sharpening finds linear combinations of broadband sensors, for example cone fundamentals, that are maximally sensitive in a given wavelength interval $[\lambda_1, \lambda_2]$, independent of the illuminant. Given a $k$-dimensional sensor sensitivity matrix $\mathbf{R}$ ($k \times n$, usually $k = 3$), the sensor $\mathbf{R}^T\mathbf{c}$ is the most sensitive in $[\lambda_1, \lambda_2]$ if the percentage of its norm is maximized with respect to this interval:

$$\max \frac{|\mathbf{R}^T_{[\lambda_1,\lambda_2]}\mathbf{c}|}{|\mathbf{R}^T\mathbf{c}|} \tag{2.32}$$

$\mathbf{c}$ is a $k$-dimensional coefficient vector. Solving for $\mathbf{c}$ can be achieved using a Lagrange multiplier and minimizing:

$$I = \sum_{\upsilon}(\mathbf{R}^T\mathbf{c})^2\Delta\lambda + \mu\left\{\sum_{\omega}(\mathbf{R}^T\mathbf{c})^2\Delta\lambda - 1\right\} \tag{2.33}$$

where $\upsilon$ denotes wavelengths outside of the sharpening interval $[\lambda_1, \lambda_2]$, $\omega$ the visible spectrum, and $\mu$ the Lagrange multiplier. By differentiating with respect to $\mathbf{c}$ and equating to the zero vector, we obtain:

$$\frac{1}{2}\frac{\delta I}{\delta \mathbf{c}} = \sum_{\upsilon}\mathbf{R}\mathbf{R}^T\mathbf{c}\Delta\lambda + \mu\left\{\sum_{\omega}\mathbf{R}\mathbf{R}^T\mathbf{c}\Delta\lambda\right\} = 0 \tag{2.34}$$

Differentiating with respect to $\mu$ results in $\sum_{\omega}(\mathbf{R}^T\mathbf{c})^2\Delta\lambda = 1$. Thus, the norm of the sharpened sensor is equal to unity, independent of the original length of the sensor set $\mathbf{R}$.

We can define two $k \times k$ matrices, such that $\Xi_\upsilon = \sum_\upsilon \mathbf{R}\mathbf{R}^T\Delta\lambda$, respectively $\Xi_\omega = \sum_\omega \mathbf{R}\mathbf{R}^T\Delta\lambda$. Thus, eq. 2.34 can be rearranged to:

$$\Xi_\upsilon\mathbf{c} = -\mu\Xi_\omega\mathbf{c} \tag{2.35}$$

Equivalently:

$$[\Xi_\omega]^{-1}\Xi_\upsilon \mathbf{c} = -\mu \mathbf{c} \tag{2.36}$$

Disregarding the trivial solutions $\mathbf{c} = 0$ and $\mu = 0$, it is evident from eq. 2.36 that $\mathbf{c}$ is an eigenvector of $[\Xi_\omega]^{-1}\Xi_\upsilon$. Solving for $\mathbf{c}$ for each of the $k$ wavelength intervals, such that $\sum_\upsilon (\mathbf{R}^T\mathbf{c})^2\Delta\lambda$ is minimized, results in the transformation matrix $\mathbf{T}$ of eq. 2.31.

The interval $[\lambda_1, \lambda_2]$ can be chosen arbitrarily. Finlayson et al. [FDF94b] used spectral intervals corresponding approximately to the "blue" (400-480nm), "green" (520-560nm), and "red" (580-650nm) parts of the spectrum. This ensures that the whole visible spectrum is sampled and that the peak sensitivities of the new sensors are close to those of the cone fundamentals. Sharpening the Vos and Walraven [VW71] cone fundamentals, they obtained sensors that are "sharper," i.e. more narrow-band than the cones, and the maximum sensitivity in the "red" interval moved to a longer wavelength. Figure 2.3 shows the original Vos and Walraven cone fundamental and the spectrally sharpened sensors. They validated their sensors by comparing actual and predicted physical color responses for surface reflectances under several illuminants and showed that on average, using sharpened sensors with a DMT is better able to predict the color response than using Vos and Walraven cone fundamentals. We found a similar result, using actual corresponding colors, when comparing the Hunt-Pointer-Estevez chromatic adaptation transform to a "sharp" chromatic adaptation transform (see chapter 3).

### 2.7.2 Applications of Spectral Sharpening

Applying their results to other color constancy algorithms proved that the generalized diagonal transform model is able to achieve good color constancy, provided that illuminant and reflectance can be modeled by small finite linear combinations [FDF94a], such as is assumed by Maloney and Wandell [MW86] and by Forsyth [For90]. Hubel and Finlayson [HF98] also validated sharp sensors in a color appearance experiment. Using a similar set-up as Braun and Fairchild ([BF96], see section 2.8.1), they compared images (print to monitor) rendered under different illuminants using several chromatic adaptation transforms. Especially for large

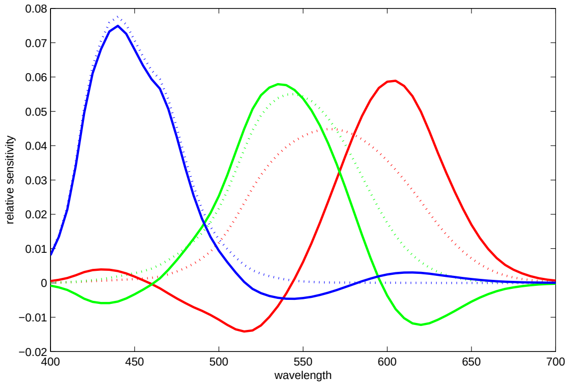


Figure 2.3: Original Vos and Walraven cone fundamentals (dotted line) and sensor-based sharpened sensors (solid line).

illumination shifts, a transform based on sharp sensors performed better than the others. Hubel et al. [HHFD97] have used sharp sensors to determine the best transform from digital camera response values to color encoding values. Drew and Finlayson [DF00] have also shown how to derive all positive sharpened sensors, or sensors that give all positive responses. Sharpened sensors can have negative values, as shown in Figure 2.3. Positive sensors, if they are physically realizable, are useful in digital cameras that *white-balance* in sensor space. White-balancing is a color image processing step where the response of a given channel is normalized with respect to the illuminant and the different quantum efficiencies of the real filter and sensor combination:

$$R^o = \frac{R^a}{g_R}, \qquad G^o = \frac{G^a}{g_G}, \qquad B^o = \frac{B^a}{g_B} \tag{2.37}$$

Here, the factors $g_R, g_G, g_B$ indicate a multiplication of illuminant response and quantum efficiency gain factors. Eq. 2.37 is a DMT, so sharpened sensors improve the color constancy of a camera compared to any other with arbitrary filter responses whose best solution might be a full 9 parameter matrix.

As some of the approaches used in this thesis are based on spectral sharpening, we will revisit the literature the subsequent chapters.

## 2.8 Corresponding Colors

Before we can evaluate the von Kries model or develop new models for chromatic adaptation, we must first have experimental measurements of pairs of stimuli (*corresponding colors*) that match in appearance with respect to different adaptation conditions. Corresponding colors can be described as a pair of tristimulus values ($XYZ$) or cone responses ($LMS$), based on one physical stimulus that appear to be the same color when viewed under two different illumination sources [Hun95]. Below we describe the psychophysical experiments conducted to collect such stimuli pairs.

### 2.8.1 Experimental Methods

Chromatic adaptation is a *color appearance* phenomenon, and it is well known that color appearance could depend on many factors, such as local surround, spatial and chromatic scene arrangement, adaptive luminance, etc. [Fai98]. It is very difficult to "distill" the effect of chromatic adaptation from all other color appearance mechanisms, and as such define an experimental protocol for it. Thus, several experimental methods have been proposed that are briefly described below.

#### Asymmetric Matching

Many studies investigating chromatic adaptation models used *asymmetric matching* experiments. The basic concept of asymmetric matching is to expose the observer to color stimuli under different illuminants $E^a(\lambda)$ and $E^b(\lambda)$. The observer then matches under $E^b(\lambda)$ the perceived color under $E^a(\lambda)$ [WS82]. *Simultaneous* asymmetric matching refers to the observer being adapted to two different illuminants at the same time. Some early experiments [Wri34, Hun50, BEN57] use *haploscopic* (also called dichoptic) color matching where the observers use a colorimeter like device (see Figures 1.7 and 2.1) with two viewing areas. One eye is looking at a test light surrounded by an area illuminated by $E^a(\lambda)$, and the other eye is looking at a reference light, surrounded by an area illuminated by $E^b(\lambda)$. The viewing area under $E^b(\lambda)$ contains a field with a mixture of red, green and blue primaries that can be adjusted to match the test stimulus. The uniform (and often neutral) surround is adjusted for different luminance conditions. Breneman [Bre87] and Chichilnisky and Wandell [CW95] used a similar set-up, but with a complex surround (natural scene) or a colored surround, respectively.

The advantages of this method is its simplicity and high precision. Disadvantages are that both eyes are adapted to different illuminants. As signals of both eyes are mixed at the level of the LGN, it is not clear what the effect of this interaction is. Thus, MacAdam [Mac56] used a single bipartite viewing field where both halves show different adapting conditions. The observer fixates to the partition line. In this case, the same retinal area in both eyes is viewing the same adaptation conditions. Such a set-up is called *monocular* matching.

Wuerger [Wue96] had the observer fix a point on a CRT monitor, but additionally varied the temporal frequency of the stimuli. To allow for a more natural viewing experience, other studies [AR86, ARSG91, BBS97, DB00] allowed movements of both eyes over the two adapting fields. Again, it is difficult to predict how the observers' adaptation state behaves if they can move their eyes freely from one region to the other.

In *successive* asymmetric color matching, the observer views a single scene first under $E^a(\lambda)$ and then $E^b(\lambda)$. The observer adjusts a colored patch under $E^b(\lambda)$ to appear the same as under $E^a(\lambda)$. Brainard and Wandell [BW92] and Bäuml [Bäu95] used two CRT monitors calibrated to different illuminants, and Braun and Fairchild [BF96] used a CRT monitor to match hardcopy. Obviously, the accuracy of these experiments depend on the ability of the observers to remember color accurately. McCann et al. [MMT76] and Lucassen et al. [LW93, LW96] used a combination of successive and haploscopic matching: one eye was adjusted to one viewing conditions, while the other eye successively matched the stimuli.

### Achromatic Matching

In achromatic matching, the observers adjust a specified color to appear neutral [FL92, Bäu94, Bra98, KMB02]. This method is less exact in predicting the remapping of colors due to illumination change. It provides coordinates of points in the observer's achromatic locus, but not exactly which point. However, Speigle and Brainard [SB99] reported after comparing the two different methods that achromatic adjustments, together with a gain-control model, can be used to make accurate predictions of the chromaticity of asymmetric matches.

### Memory Matching and Magnitude Estimation

Helson et al. [HJW52] and Lam [Lam85] used a *memory matching* experiment were observers first learned to describe colors in terms of hue, lightness and chroma of the Munsell color ordering system [WS82, Mun76]. The training is time intensive, but a trained observer is reasonably accurate in predicting Munsell coordinates. For a detailed description of Lam's experiment, see section 3.2. Arend [Are93] used a combination memory-achromatic matching experiment. He trained the observers on the Munsell color system, and then had them create neutrals or unique hues under different illuminants on a CRT monitor, as they remembered from the Munsell coordinates.

In *magnitude estimation* (or color naming) experiments, the observers are asked to rate the lightness, colorfulness and hue of a stimulus [Bar79a, Bre87, UUB89, LCR$^{+}$91, KLB95, SB96]. These values are then projected to a chromaticity diagram, from where tristimulus values can be derived at equal luminance, i.e. equal tristimulus value $Y$.

## 2.9 Conclusions

The literature most relevant for the subsequent chapters has been summarized. The different studies and resulting models emphasize that chromatic adaptation is still a rich field of research, both from a computer vision and human vision system modeling point of view. We will concentrate on the former in the subsequent chapters to define algorithms useful in color science and color imaging applications.

# Chapter 3

# Spectral Sharpening

## 3.1 Introduction

Chromatic adaptation transforms (CATs) map color responses under a reference light source to color responses for a target light source such that the corresponding color predicts the appearance of a given surface reflectance under the target light source. In color science and color appearance models [MFH$^+$02], CATs are used to predict the appearance of a surface color under illuminants with different spectral power distributions (SPD). In color imaging, they are used in color management systems to transform pixel values from one color image encoding to another, if the color image encodings specify different white-points [ISO04].

Considering the wide spread applications, several CATs have lately been proposed in the literature. The implication of all these recently proposed chromatic adaptation transforms [Lam85, Fai01, LLRH02] is that color correction for illumination takes place not in cone space but rather in a "narrowed" cone space, if we assume that the strong von Kries chromatic adaptation model (see section 2.4) holds, which predicts that chromatic adaptation can be modeled as an independent gain control in each channel. The CAT sensors have their sensitivity more narrowly concentrated than the cones. But are these sensors optimally narrow for predicting corresponding colors subject to a strong von Kries DMT-transform? Recent work in color constancy has shown that it is possible to sharpen sensors to a much greater extent [FDF94b, FDF94a] when physical scene measurements (XYZ or camera RGBs) need to be mapped under different illuminants.

Generally in this thesis, we want to explore if the mathematical techniques used in modeling physical data, i.e. measurements of color responses under different illuminants, is also applicable to experimentally derived data, i.e. corresponding colors. In this chapter, we derive a CAT based on spectral sharpening and compare the perceptual errors between actual appearance and predicted appearance of a color under different illuminants to the different CATs recently proposed in the literature, since it is perceptual error that these CATs minimize. We begin with Lam's original experiments (section 3.2) that resulted in the Bradford transform, which was extensively used in color science and color imaging.

We compare the performance of the linear Bradford transform, the linear CAT proposed by Fairchild [Fai01], and CAT02 implemented in CIECAM02 [LLRH02, MFH$^{+}$02] (section 3.3) with that of an adaptation transform that is based on sharp sensors. We describe how we derive a sharp chromatic adaptation transform based on corresponding color data (section 3.4). In section 3.5, perceptual errors in CIELAB $\Delta E$, $\Delta E_{CIE94}$, and $\Delta E_{CMC(1:1)}$ are calculated for several corresponding color data sets and analyzed for their statistical significance. The results are found to be similar for all transforms, with the exception of the "original" von Kries applied in cone space. The sharp transform performs equally well as CAT02, the newly proposed chromatic adaptation transform for CIECAM02: there is no statistically significant difference in performance for most data sets. We discuss our findings in relation with other studies in section 3.6.

## 3.2 Lam's Experiment

In his experiment to derive a chromatic adaptation transform, Lam [Lam85] used 58 dyed wool samples. His main objectives when choosing the colors were that the samples represent a reasonable gamut of chromaticities corresponding to ordinary collections of object colors (see Figure 3.1), and that the samples have various degrees of color constancy with regard to change of illuminant from D65 to A (see Appendix B).

To evaluate the samples, Lam used a memory matching experiment where observers are asked to describe the color appearance of stimuli in relation with a memorized color ordering system. Lam trained the observers on the Munsell system. Each observer was asked to describe the appearance of the samples in Munsell hue, chroma and value terms. The observers were fully adapted to the illuminant before they began the ordering. He used five observers with each observer repeating the experiment twice, resulting in ten color descriptions for each surface and for each illuminant, respectively.

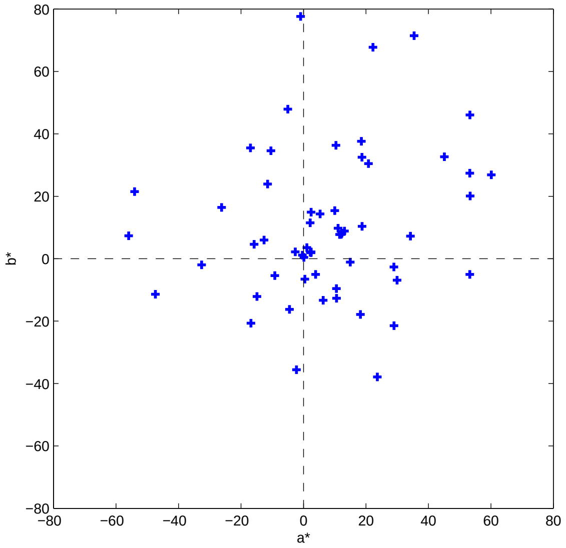


Figure 3.1: Distribution of Lam's 58 samples in CIE $a^*, b^*$, measured under D65.

Lam converted the average Munsell coordinates of each sample under illuminant D65 and A to CIE 1931 Y, x and y values so that a color difference formula can be applied to the data. He calculated tristimulus values using the 1931 CIE equivalents of Munsell samples under illuminant C [NNJ43]. To calculate Munsell equivalent values under D65, he used the Helson et al. [HJW52] chromatic adaptation transform to correct for the illuminant change from C to D65.

Lam was now in a position to derive a chromatic adaptation transform, i.e. to find a mapping that related his corresponding color data. He derived a chromatic adaptation transform minimizing $\Delta E$ between actual and predicted corresponding colors under the following set of constraints: (1) the transform should maintain achromatic constancy for all neutral samples, (2) it should work with different adapting illuminants, and (3) it should be reversible (i.e. when a particular color is transformed from A to D65, and back to A again, the tristimulus values before transformation and after transformation back to A should be the same). These constraints have subsequently been generally adapted when developing CATs [FS00a, Fai01, LLRH02].

The Bradford chromatic adaptation transform, called KING1 in his thesis, is based on a simplified MacAdam [Mac61, Mac63] or Nayatani transform [NTS81], and is similar to the Bartleson transform [Bar79a, Bar79b] as it also contains a non-linearity in the blue channel (see section 2.5.2). The individual steps are as follows:

*Step 1:* Transformation from $X^a, Y^a, Z^a$ to $R^a, G^a, B^a$ under the reference illuminant $E(\lambda)^a$:

$$\begin{bmatrix} R^a \\ G^a \\ B^a \end{bmatrix} = \mathbf{M}_{BFD} \begin{bmatrix} X^a/Y^a \\ Y^a/Y^a \\ Z^a/Y^a \end{bmatrix} \tag{3.1}$$

where

$$M = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}$$

*Step 2:* Transformation from $R^a, G^a, B^a$ to $R^b, G^b, B^b$ under the test illuminant $E(\lambda)^b$.

$$\begin{bmatrix} R^b \\ G^b \\ B^b \end{bmatrix} = \begin{bmatrix} \frac{R^b_w}{R^a_w} & & \\ & \frac{G^b_w}{G^a_w} & \\ & & \frac{B^b_w}{(B^a_w)^p} \end{bmatrix} \begin{bmatrix} R^a \\ G^a \\ (B^a)^p \end{bmatrix} \tag{3.2}$$

where

$$p = (B^a_w/B^b_w)^{0.0834}$$

Quantities $R^a_w, G^a_w, B^a_w$ and $R^b_w, G^b_w, B^b_w$ are computed from the tristimulus values of the reference and test illuminants, respectively, through eq. 3.1.

*Step 3:* Transformations from $R^b, G^b, B^b$ to $X^b, Y^b, Z^b$.

$$\begin{bmatrix} X^b \\ Y^b \\ Z^b \end{bmatrix} = \mathbf{M}_{BFD}^{-1} \begin{bmatrix} R^b Y^a \\ G^b Y^a \\ B^a Y^a \end{bmatrix} \tag{3.3}$$

Like other studies [Mac61, Mac63, NTS81, WW82, SH88, Bäu95, LW93, DB00], Lam found that the strong von Kries model was not adequate to predict his corresponding color data, and chose a non-linear chromatic adaptation model to define his transform. A strong von Kries model assumes that the gain is dependent only on signals from the same channel, or one cone class. However, as opposed to MacAdams and Nayatani et al., he restricted the non-linearity to the blue channel. A similar transform was later developed by Lucassen and Walraven [LW93], who found that a log-based transform better predicts corresponding color responses in the blue channel.

One implication of the Bradford chromatic adaptation transform is that color correction for illumination takes place not in cone space but rather in a "narrowed" cone space. The Bradford sensors (the linear combination of XYZs defined in the Bradford transform) have their sensitivity more narrowly concentrated than the cones (see Figure 3.2). Additionally, the long Bradford sensor has its peak sensitivity shifted to longer wavelength.

## 3.3 Linear Chromatic Adaptation Transforms

For a few years, the Bradford chromatic adaptation transform was the preferred CAT for color science applications. It was standardized as CMCCAT97 [LH98a] by the Colour Measurement Committee (CMC) of the Society of Dyers and Colourists as the preferred chromatic adaptation transform to use and by the CIE in the color appearance model CIECAM97s [LH98b, CIE98]. However, due to problems in CIECAM97s model reversibility [Fai01] and computational speed, the Bradford CAT was not adopted in the imaging community where CATs are used in color management applications [SW00, ICC04]. The non-linear correction in the blue of the Bradford transform was considered negligible and is not encoded. The linear

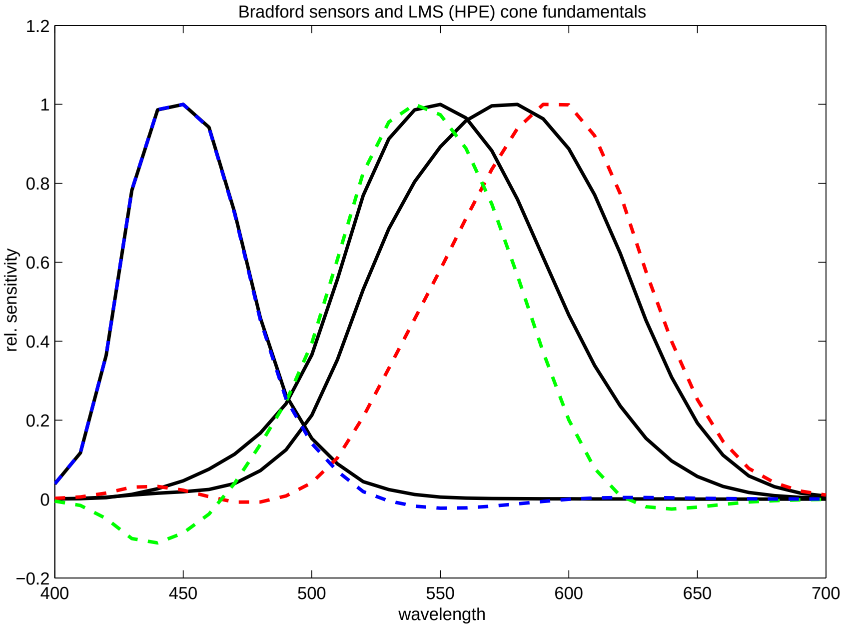


Figure 3.2: The normalized Bradford sensors (dash) compared to the normalized LMS (HPE) cone fundamentals (solid).

Bradford transform thus becomes a Diagonal Matrix Transform (DMT)-type von Kries-Ives [Mal01] transform:

$$\begin{bmatrix} X^b \\ Y^b \\ Z^b \end{bmatrix} = \mathbf{M}_{BFD}^{-1} \mathcal{D} \mathbf{M}_{BFD} \begin{bmatrix} X^a \\ Y^a \\ Z^a \end{bmatrix} \tag{3.4}$$

where

$$\mathcal{D} = \begin{bmatrix} \frac{R_w^b}{R_w^a} & & \\ & \frac{G_w^b}{G_w^a} & \\ & & \frac{B_w^b}{B_w^a} \end{bmatrix}$$

Quantities $R_w^a, G_w^a, B_w^a$ and $R_w^b, G_w^b, B_w^b$ are computed from the tristimulus values of the reference and test illuminants by multiplying the corresponding XYZ vectors by $\mathbf{M}_{BFD}$.

Subsequent studies also proposed linear CATs in a narrower cone space. Fairchild [Fai01], in an effort to simplify CIECAM97s and make it reversible, proposed a linear CAT that most closely performed to the original Bradford transform in terms of perceptual error. He used Munsell samples to calculate corresponding colors under illuminants A and D65 using the non-linear Bradford CAT. He then developed a linear CAT by minimizing the CIELAB differences to the predictions of the Bradford CAT on this corresponding color data set. Li et al. [LLRH02] derived a new CMCCAT2000 based on minimizing perceptual error ($\Delta E$) over a set of corresponding color data compiled by Luo and Rhodes [LR99]. A slightly different transform was derived by excluding the data set of McCann et al. [MMT76]. This latter CAT was adopted by the CIE as the chromatic adaptation transform CAT02 for a new color appearance model CIECAM02 [MFH$^+$02]. We use the same data set in this study to test the performance of the CAT we derive in this chapter. The linear transformations $\mathbf{M}$ are as follows:

$$\mathbf{M}_{Fai} = \begin{bmatrix} 0.8562 & 0.3372 & -0.1934 \\ -0.8360 & 1.8327 & 0.0033 \\ 0.0357 & -0.0469 & 1.0112 \end{bmatrix}$$

$$\mathbf{M}_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6974 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix}$$

The corresponding sensors are illustrated in Figure 3.3. The three transforms, linear Bradford, Fairchild and CAT02 result in very similar sensors, slightly more narrow-band than the cone responses, and containing some negative values. For comparison, the Hunt-Pointer-Estevez (HPE) [Hun98] cone fundamentals are also plotted. The linear transform from XYZ to LMS is as follows:

$$\mathbf{M}_{HPE} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0 & 0 & 1 \end{bmatrix}$$

The HPE cone fundamentals have been used by Hunt to develop color vision models [Hun98], and are also part of Nayatani et al. [NTS81] and Fairchild's [Fai96] color appearance models [Fai98]. They are derived as a linear combination of the CIE 1931 XYZ color matching functions that most closely match a cone quanta absorption study of Estevez for 2 degree observers [Est79].

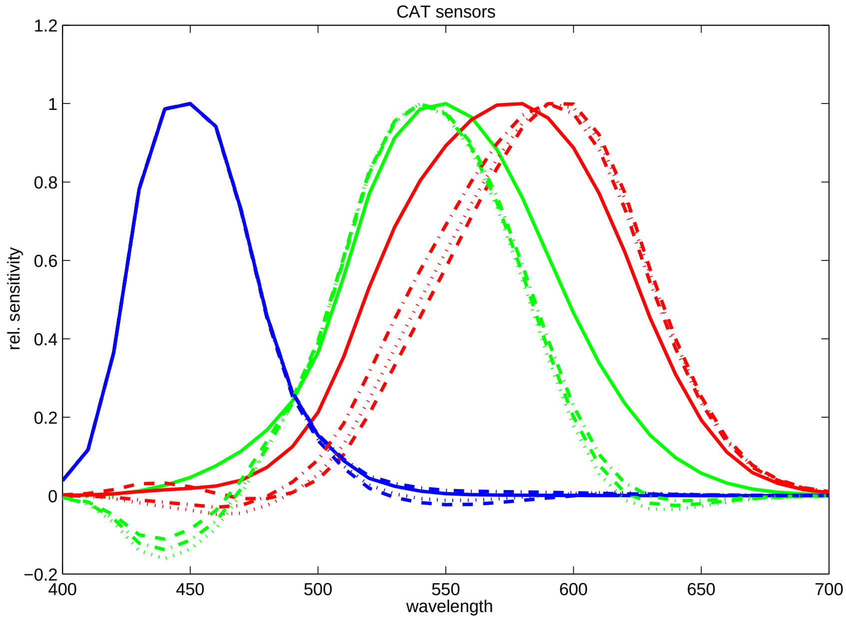


Figure 3.3: The normalized sensors resulting from different chromatic adaptation transforms: HPE (solid color), Bradford (dash), Fairchild (dotted), CAT02 (dash-dot).

## 3.4 The Sharp Adaptation Transform

The Sharp adaptation transform [FS00b, FS00a] used in this experiment is derived with a data-based sharpening algorithm, first described by Finlayson et al. [FDF94b]. The performance of the diagonal-matrix transformations that are used in many color constancy algorithms can be improved if the two data sets are first transformed by a sharpening transform $\mathbf{T}$. They showed that the resulting "sharp" sensors are the most appropriate basis for modeling and/or computing adaptation of physical quantities (raw XYZs) across illuminants, i.e. for solving the non-perceptual adaptation problems when treating XYZs as the important units. Data-based sharpening was initially developed to validate sensor-based sharpening and to evaluate if the sharpening transform varies drastically with different illuminant and surface reflections.

The data-based sharpening transform $\mathbf{T}$ is derived by finding the best transform between two color response data sets under different illuminants [FDF94b]. For our purpose of deriving a chromatic adaptation transform, we assume that the data sets contain corresponding color data. We can thus derive a linear chromatic adaptation transform based on data-based spectral sharpening [FDF94b] as follows.

Let $\mathbf{W}^a$ $3 \times n$ be a set of tristimulus values under illuminant $E^a(\lambda)$ and $\mathbf{W}^b$ a set of tristimulus values under illuminant $E^b(\lambda)$. According to von Kries, the two sets should be approximately equal, subject to a diagonal matrix transform (DMT):

$$\mathbf{W}^b \approx \mathbf{\Lambda}^{b,a}\mathbf{W}^a \tag{3.5}$$

where $\mathbf{\Lambda}^{b,a}$ is a diagonal transform. The idea of spectral sharpening is that the residual error in the mapping can be reduced if both data sets are first linearly transformed by a sharpening matrix $\mathbf{T}$:

$$\mathbf{T}\mathbf{W}^b \approx \mathbf{\Lambda}^{b,a}\mathbf{T}\mathbf{W}^a \tag{3.6}$$

Solving for $\mathbf{\Lambda}^{b,a}$ requires minimizing some error criterion. Using least-squares, $\mathbf{\Lambda}^{b,a}$ can be solved for by the Moore-Penrose inverse:

$$\mathbf{\Lambda}^{b,a} = \mathbf{T}\mathbf{W}^b[\mathbf{T}\mathbf{W}^a]^+ \tag{3.7}$$

where $[]^+$ denotes the pseudoinverse [GvL96]. The solution for $\mathbf{T}$ must ensure that $\mathbf{\Lambda}^{b,a}$ is diagonal. Rearranging eq. 3.7, we obtain:

$$[\mathbf{T}]^{-1}\mathbf{\Lambda}^{b,a}\mathbf{T} = \mathbf{W}^b[\mathbf{W}^a]^+ \tag{3.8}$$

The solution for $\mathbf{T}$ is therefore in the eigenvector decomposition of:

$$\mathbf{W}^b[\mathbf{W}^a]^+ = \mathbf{U}^{b,a}\mathbf{\Lambda}^{b,a}[\mathbf{U}^{b,a}]^{-1}$$

Thus, $\mathbf{T}$ is equal to $[\mathbf{U}^{b,a}]^{-1}$.

Note that the best general transform that maps $\mathbf{W}^a$ to $\mathbf{W}^b$ in the least-squares sense is found by:

$$\mathbf{B}^{b,a} = \mathbf{W}^b[\mathbf{W}^a]^+ = \mathbf{W}^b[((\mathbf{W}^a)^T(\mathbf{W}^a))^{-1}(\mathbf{W}^a)^T] \tag{3.9}$$

Thus, eq. 3.8 can be interpreted as simply being the eigenvector decomposition of the general transform $\mathbf{B}^{b,a}$. What is important to note though is that if the sharpening transformation $\mathbf{T}$ is first applied to the color responses $\mathbf{W}^b, \mathbf{W}^a$, then the optimal transformation *is* the diagonal transform $\mathbf{\Lambda}^{b,a}$.

Thus, using Lam's experiment, the prediction of the corresponding colors under D65 using a von Kries-Ives DMT chromatic adaptation model should approximately equal

$$\mathbf{T}\mathbf{W}^b \approx \mathcal{D}^{b,a}\mathbf{T}\mathbf{W}^a \tag{3.10}$$

where $\mathbf{W}^b$ is a 3 x 58 matrix of corresponding color XYZs under illuminant D65, $\mathbf{W}^a$ is a 3 x 58 matrix of the measured XYZs under illuminant A and $\mathcal{D}^{b,a}$ is the diagonal matrix formed from the ratios of the two sharpened white-point vectors $[R_w^{D65}, G_w^{D65}, B_w^{D65}]$ and $[R_w^A, G_w^A, B_w^A]$, derived by multiplying vectors $[X_w^{D65}, Y_w^{D65}, Z_w^{D65}]$ and $[X_w^A, Y_w^A, Z_w^A]$ with $\mathbf{T}$.

The matrix $\mathbf{T}$ is derived from the matrix $\mathbf{B}^{b,a}$ that best maps $\mathbf{W}^a$ to $\mathbf{W}^b$ minimizing least-squares error, as shown above. However, while $\mathbf{B}^{b,a}$ calculated using eq. 3.9 results in the smallest mapping error, it will not fulfill the requirement that particular colors are mapped without error, e.g. preserving achromaticity for neutral colors. Therefore, $\mathbf{B}^{b,a}$ was derived using a white point preserving least-squares regression algorithm [FD97a, FD97b]. The intent is to map the values in $\mathbf{W}^a$ to corresponding values in $\mathbf{W}^b$ so that the RMS error is minimized subject to the constraint that, as an artifact of the minimization, the achromatic scale is correctly mapped. In order to preserve white:

$$\mathbf{B}^{b,a} = \mathbf{D} + (\mathbf{Z}\mathbf{N}) \tag{3.11}$$

where $\mathbf{D}$ is the diagonal matrix formed from the ratios of the two white point vectors $[X_w^{D65}, Y_w^{D65}, Z_w^{D65}]$ and $[X_w^A, Y_w^A, Z_w^A]$, respectively. $\mathbf{Z}$ is a $3 \times 2$ matrix composed of any two vectors orthogonal to the $[X_w^A, Y_w^A, Z_w^A]$ vector. $\mathbf{N}$ is obtained by substituting $\mathbf{Z}$, $\mathbf{N}$ and $\mathbf{D}$ in eq. 3.9 and solving for $\mathbf{N}$. The sharpening transform $\mathbf{T}$ can then be derived through eigenvector decomposition of the general transform $\mathbf{B}^{b,a}$:

$$\mathbf{B}^{b,a} = \mathbf{U}^{b,a}\mathbf{\Lambda}^{b,a}[\mathbf{U}^{b,a}]^{-1} \tag{3.12}$$

where $\mathbf{T}$ is equal to $[\mathbf{U}^{b,a}]^{-1}$.

The predicted corresponding colors under illuminant D65 of Lam's 58 samples, using this Sharp transform, are calculated with eq. 3.10.

Numerically, the linear transformation $\mathbf{T}$, which we will call $\mathbf{M}_{Sharp}$ [FS00b, FS00a], is equal to:

$$\mathbf{M}_{Sharp} = \begin{bmatrix} 1.2694 & -0.0988 & -0.1706 \\ -0.8364 & 1.8006 & 0.0357 \\ 0.0297 & -0.0315 & 1.0018 \end{bmatrix} \tag{3.13}$$

The corresponding sensors are illustrated in Figure 3.4.

Note that for the rest of this thesis, when we refer to the *Sharp* (capital $S$) sensors, we refer to the sensors illustrated in Figure 3.4, which are derived by multiplying the CIE XYZ color matching function with eq. 3.13. When we refer to the *Sharp* (capital $S$) chromatic adaptation transform, we refer to a von Kries-Ives DMT that is applied in the Sharp sensor space. When 'sharp' is not captialized, we refer to sensors that resemble the Sharp sensors of Figure 3.4, i.e. they are more narrow-band than cone fundamentals.

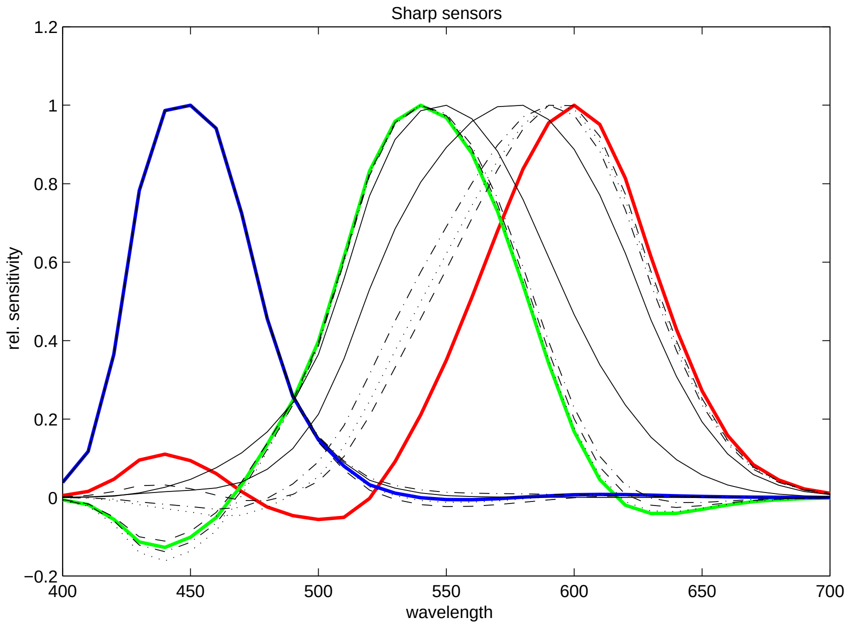


Figure 3.4: The Sharp sensors (solid color) compared to the other CAT sensors: HPE (solid), Bradford (dash), Fairchild (dotted), CAT02 (dash-dot).

## 3.5 Comparison of the Sharp CAT with other linear CATs

Applying the resulting Sharp CAT, derived via data-based sharpening of the corresponding colors of the 58 Lam samples under illuminants A and D65 minimizes the RMS error between corresponding XYZs. It also yields sensors that are visibly sharper than those implied by the Bradford transform (see Figure 3.4). However, what we are most interested in is to compare the *perceptual* error between actual appearance and predicted appearance of a color under different illuminants using the Sharp and the other linear transforms, an we will thus compare errors in a more perceptual color space encoding.

Several corresponding color data sets were used to compare the performance of the Sharp transform, the linear Bradford, the CAT02 and the Fairchild transform, as well as the original von Kries transform using HPE cone fundamentals. Together they form a set accumulated by Luo and Rhodes for the purpose of deriving and evaluating color appearance models and chromatic adaptation transforms [LR99]. The same data set was also used to derive CAT02 [LLRH02]. Table 3.1 lists the characteristics of the data sets used in this study.

The actual and predicted XYZ values were converted to CIELAB space using eq. 1.10. Three perceptual error prediction methods, $\Delta E_{Lab}$ (eq. 1.11), $\Delta E_{94}$ (1.14), and $\Delta E_{CMC(1:1)}$ [Hun98] were applied. One-tail student-t tests [WMM98] for matched pairs were used to compare the results to find if the variations in errors are statistically significant. For each corresponding color data set, the resulting $p$-values were calculated using the best performing CAT in terms of mean perceptual error as one input, and the other CATs as the other input. The null hypothesis was that the mean of the perceptual errors of the best performing CAT and the other CAT is equal to zero. The alternative hypothesis was that the best performing CAT has a smaller mean. A large $p$-value supports the null hypothesis, and a small $p$-value rejects it. $p$-values equal or larger than 0.05 and 0.01 indicate that the means are equal (or that the null hypothesis cannot be rejected) at a confidence level of 95 and 99 percent, respectively.

Table 3.1: Characteristics of the corresponding color data sets used in this study [Lam85, HJW52, MSKI91, LCR$^+$91, KLB95, BF96, Bre87].

| Data Sets | No. of Samples | Approx. Test Illuminant | Approx. Reference Illuminant | Sample Size | Medium | Experimental Method |
|---|---|---|---|---|---|---|
| Lam | 58 | D65 | A | L | Refl. | Memory |
| Helson | 59 | D65 | A | S | Refl. | Memory |
| CSAJ | 87 | D65 | A | S | Refl. | Haploscopic |
| Lutchi | 43 | D65 | A | S | Refl. | Magnitude |
| Lutchi D50 | 44 | D65 | D50 | S | Refl. | Magnitude |
| Lutchi WF | 41 | D65 | WF | S | Refl. | Magnitude |
| Kuo&Luo | 40 | D65 | A | L | Refl. | Magnitude |
| Kuo&Luo TL84 | 41 | D65 | TL84 | S | Refl. | Magnitude |
| Braun&Fairchild 1 | 17 | D65 | D93 | S | Monitor & Refl. | Asymmetric |
| Braun&Fairchild 2 | 16 | D65 | D93 | S | Monitor & Refl. | Asymmetric |
| Braun&Fairchild 3 | 17 | D65 | D30 | S | Monitor & Refl. | Asymmetric |
| Braun&Fairchild 4 | 16 | D65 | D30 | S | Monitor & Refl. | Asymmetric |
| Breneman 1 | 12 | D65 | A | S | Trans. | Magnitude |
| Breneman 8 | 12 | D65 | A | S | Trans. | Magnitude |
| Breneman 4 | 12 | D65 | A | S | Trans. | Magnitude |
| Breneman 6 | 11 | D55 | A | S | Trans. | Magnitude |

Table 3.2: The number of times a transform performed best or was statistically the same ($p$-values $\geq 0.05$, 95% confidence) as the best transform.

| Error Metric | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{HPE}$ |
|---|---|---|---|---|---|
| $\Delta E$ | 10 | 9 | 14 | 11 | 6 |
| $\Delta E_{CIE94}$ | 12 | 8 | 11 | 11 | 6 |
| $\Delta E_{CMC(1:1)}$ | 12 | 9 | 12 | 9 | 6 |

Table 3.3: The number of times a transform performed best or was statistically the same ($p$-values $\geq 0.01$, 99% confidence) as the best transform.

| Error Metric | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{HPE}$ |
|---|---|---|---|---|---|
| $\Delta E$ | 15 | 12 | 14 | 14 | 7 |
| $\Delta E_{CIE94}$ | 14 | 11 | 14 | 12 | 9 |
| $\Delta E_{CMC(1:1)}$ | 14 | 11 | 16 | 12 | 8 |

The results for each error prediction method, corresponding color data set, and transform are listed in Tables C.1, C.2 and C.3 in Appendix C. To summarize, Table 3.2 and Table 3.3 lists the number of times for the total number of datasets a CAT performed best or was statistically the same as the best transform at a 95 and 99 percent confidence level, respectively. The maximum score for each error metric is 16, as 16 corresponding color data sets were tested.

## 3.6 Discussion

Using 95% confidence, the CAT02 performs best considering only $\Delta E$ as error metric. That is not surprising, considering that CAT02 was optimized over the same data set used in this study minimizing $\Delta E$. However, for $\Delta E_{CIE94}$ and $\Delta E_{CMC(1:1)}$, which are considered to be more precise error metrics [Hun98], the performance of Sharp CAT, CAT02 and the Fairchild CAT are approximately the same. All three were able to predict corresponding colors for almost all data sets as well as the best performing transform. The linear Bradford CAT performed slightly worse, but not significantly when the confidence interval is increased to 99%.

These findings have been experimentally corroborated. Calabria and Fairchild [CF01] tested different CATs in a psychovisual study on real and simulated images and found that there was no visual difference between linear CATs in a narrowed cone space and the non-linear Bradford transform. Hubel and Finlayson [HF98] even found a slightly better performance of a sharp CAT when comparing printed images to monitor images.

The "original" von Kries transform, scaling cone responses, does perform less well than the other four linear transforms in this study as well as the visual studies of [HF98, CF01]. The von Kries sensors (i.e. the cone fundamentals) have distinctly different peaks for the green and red sensors than the other CATs (see Figure 3.3), which leads to the conclusion that chromatic adaptation is not just a function of the cone responses of the human visual system. This corresponds well with many psychophysical studies that also could not find an appropriate mapping with a DMT in cone space (see chapter 2). However, if we allow that scaling does not take place in cone space, but an "intermediate" cone space, a linear transformation seems to be a good enough approximation to model the effect of illumination change on color appearance.

The Bradford, CAT02 and Fairchild CATs are derived by minimizing perceptual error, i.e. $\Delta E$. The Sharp CAT is derived by optimizing the mapping from XYZ under one illuminant to the XYZs under another illuminant, subject to a DMT transform. The resulting sensors (see Figure 3.3) are much "sharper," i.e. more narrow-band than the other CAT sensors.

Though perceptual data was not used to derive spectrally sharpened sensors, spectral sharpening does appear to be psychophysically relevant. Indeed, sharp sensors have been discovered in many different psychophysical studies. Foster [Fos81] observed that when field spectral sensitivities of the red and green response of the human eye are determined in the presence of a small background field, the resulting curves are more narrow and de-correlated than the regular cone responses. These sharpened curves tend to peak at wavelengths of 530 nm and 605 nm, respectively.

Poirson and Wandell [PW90] studied the color discrimination ability of the visual system when targets are briefly presented in a complex display. The spectral sensitivities derived from their experimental data peak relatively sharply around 530 and 610 nm. Thornton [Tho99] postulated that the visual response to spectral lights consists of sharp sensors with peak wavelength around 450, 533, and 611 nm, as the conventional CIE color matching functions (CMFs) do not transform appropriately to predict color matches by the same observer using another set of primaries. He proposed that the real color matching functions that approximate the visual sensitivity to spectral lights are much narrower sensitivities than the current CIE CMFs (see section 1.2.1). Brill et al. [BFHT98] discussed prime-color wavelengths of 450, 540, and 605 nm. They proved that monitor primaries based on these wavelengths induce the largest gamut size, and that these monitors are visually very efficient. The color matching functions derived from these primaries, when linearly related to the CIE 1931 color matching functions, are sharp and de-correlated. Finlayson and Morovic [FM01] found that using a sharp CAT is better able to predict metameric color matches under different illuminants than other CATs.

Thus, it is plausible that a gain control model in a "sharpened" color space is able to predict chromatic adaptation to a first approximation. This is not to say that the corresponding sharp sensors have any physiological meaning, but that the combination of individual processing sites in the HVS might result in sensors that are narrower than the cones.

From an imaging point of view, sharp sensors also have the advantage that they are close to sRGB [IEC99] and other RGB encoding color matching functions [SHF01]. So basing chromatic adaptation transforms on sharp sensors meshes well with standard color correction methods used in digital color cameras (see section 2.7.2), and might reduce noise in a color image processing chain.

## 3.7 Conclusions

We have shown that using a computational approach to chromatic adaptation yields the same results as the more empirically derived chromatic adaptation transforms present in the literature. By first transforming the color responses under one illuminant to color responses in a more narrow-band sensor space, which is derived by optimizing the diagonal mapping between the corresponding colors, we can predict as accurately as the other CATs the appearance of the colors under a second illuminant.

The sensors that we derive as being optimal to model illumination change, subject to a strong von Kries-Ives DMT-type chromatic adaptation model, are quite different from the sensors of the other CATs. The are sharper, i.e. they have narrower sensitivities and the peak sensitivity for red is at longer wavelength. Such sensors have been found in the literature in color matching and color discrimination tasks. However, the difference in sensors of the different CATs, while still mostly producing the same results, do question the assumption that only one chromatic adaptation transform should be used in all color science and color imaging applications. In the next chapter, we will explore if there are not other sensors that could also be used.

# Chapter 4

# Spherical Sampling

## 4.1 Introduction

We have seen in the previous chapter that there are several chromatic adaptation transforms (CATs) that can accurately predict corresponding color appearance under different illuminants, using the strong von Kries-Ives coefficient model. All operate in a sensor space that is narrower than cone space. However, there is still significant differences between them, which can be accounted for in the way they were derived. The Bradford transform is a linear version of the original transform found by Lam [Lam85], who minimized perceptual error ($\Delta E$) on a set of corresponding color data he obtained from a memory matching experience. CAT02 is derived by minimizing perceptual error over several sets of corresponding colors [LR99, MFH$^{+}$02]. The Sharp CAT was derived by minimizing XYZs over Lam's corresponding color data set (see section 3.4).

Considering that the shape of the sensors differs between the CATs, but their performance in predicting corresponding color over several different data sets [LR99] remains mostly the same, lead us to investigate if there are not other sensors that would perform equally well. This might be especially relevant in lieu of the standardization efforts of chromatic adaptation transforms in the color science and color management communities [ICC04], who would like to standardize one chromatic adaptation transform for all applications.

Recall that the strong von Kries-Ives DMT chromatic adaptation model is as follows:

$$[X^b, Y^b, Z^b]^T = \mathbf{M}_{CAT}^{-1} \mathcal{D}^{b,a} \mathbf{M}_{CAT} [X^a, Y^a, Z^a]^T \tag{4.1}$$

where $X^b, Y^b, Z^b$ and $X^a, Y^a, Z^a$ are the tristimulus values of a corresponding color pair under the test and reference illuminant, respectively. $\mathcal{D}^{b,a}$ is a diagonal transform:

$$\mathcal{D}^{b,a} = \begin{bmatrix} \frac{R_w^b}{R_w^a} & & \\ & \frac{G_w^b}{G_w^a} & \\ & & \frac{B_w^b}{B_w^a} \end{bmatrix}$$

Quantities $R_w^a, G_w^a, B_w^a$ and $R_w^b, G_w^b, B_w^b$ are computed from the tristimulus values of the test and reference illuminants by multiplying the corresponding XYZ vectors by $\mathbf{M}_{CAT}$. $\mathbf{M}_{CAT}$ is the $(3 \times 3)$ non-singular matrix that converts the XYZ values to sensor responses. Von Kries originally assumed that the sensors correspond to cone fundamentals, but that has been shown not to be accurate (see chapter 3). Thus, modern CATs define specific linear transforms, such as $\mathbf{M}_{CAT02}$ for the chromatic adaptation transform used in the color appearance model CIECAM02 [MFH+02], that convert cone responses into specific sensor responses. Indeed, the difference in all the CATs mentioned in the previous chapter are the coefficients of $\mathbf{M}_{CAT}$.

To find other chromatic adaptation transforms that are capable of predicting corresponding color data, subject to the strong von Kries-Ives coefficient model of eq. 4.1, means thus to find sensors that are a linear combination of the CIE XYZ color matching functions (CMFs), and by definition the cone fundamentals[1]. A plausible sensor set could therefore be defined as *all* sensors that are within a linear combination of the CMFs. Considering that the CAT sensors span a three-dimensional space, we can approach this by assuming that all possible sensors can be illustrated as vectors with their end-points on the surface of a sphere, if the lengths of the vectors are normalized to unity. The axes of the sphere are orthonormal sensors that are within a linear transformation of the CMFs. Sampling this sphere will allow to find all possible sensors that exhibit favorable chromatic adaptation transform behavior.

---

[1] Recall that we assume that cone fundamentals are a linear combination of CIE XYZ color matching functions, see section 1.2.

In this chapter, we thus develop a spherical sampling technique that is able to find a large number of solutions that fulfill some criteria [FS01b]. Our algorithm lets us find all possible RGB sensors that exhibit favorable CAT behavior when tested on Lam's corresponding color data set, subject to a $\Delta E_{94}$ error criterion. It is easier to find a set of solutions, and to be sure to find all possible solutions, if the solution space is *sampled*. Optimization through sampling implies that all possible combinations are tried, and the best solution is retained that corresponds to the global minimum. Sampling also allows retaining all possible solutions that fulfill some criteria if the result is not unique. While we apply the spherical sampling technique to finding RGB sensors optimal for von Kries-Ives DMT chromatic adaptation transforms of eq. 4.1, the technique can be used to find any three-dimensional sensors of equal magnitude.

We found that there are a number of sensors that meet the criterion, and that the Bradford, Sharp (of eq. 3.13), and CAT02 sensors are not unique. Thus, it becomes evident that the overall 'winning' adaptation transform may best be chosen by looking at secondary factors, such as transform plausibility [FM01] or similarities to other sensors found in an imaging workflow [HHF99, SHF01].

## 4.2 Spherical Sampling

In the case of trichromatic (RGB and XYZ) imaging applications, the basis functions span a three-dimensional space. If the lengths of the vectors are normalized to unity, then different vector combinations can be illustrated with their end-points that lie on the surface of a sphere (see Figure 4.1). Trying all possible combinations of three points distributed over the surface of the sphere allows us to find all possible solutions to a given problem. To determine the sample points, the surface of the sphere has therefore to be sampled at a predefined distance that depends on the application. That leads to the question of how to uniformly distribute a large number of points ($N$) on the surface of a sphere.

There has been significant research done on this problem, and a variety of algorithms have been proposed [Ale72, CS93, AWR$^{+}$94]. We have chosen to use the generalized spiral set method proposed by Rakhmanov, Saff, and Zhou [RSZ94]. Using spherical coordinates $(\phi, \theta)$, $0 \leq \phi \leq \pi$ and $0 \leq \theta \leq 2\pi$, the coordinates of the $N$ points can be calculated as follows:

$$\begin{aligned} \phi_n &= \arccos(h_n), \qquad h_n = -1 + \tfrac{2(n-1)}{(N-1)}, \quad 1 \leq n \leq N \\ \theta_n &= \left[\theta_{n-1} + \tfrac{3.6}{\sqrt{N}} \tfrac{1}{\sqrt{1-h_n^2}}\right], \quad 2 \leq n \leq N-1, \quad \theta_1 = \theta_N = 0 \end{aligned} \tag{4.2}$$

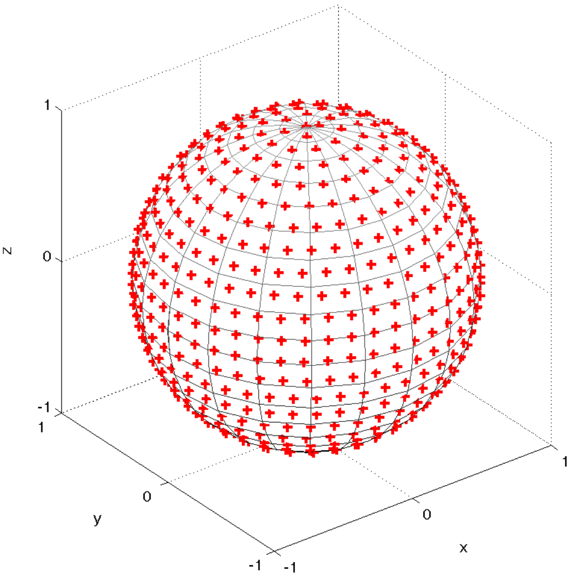


Figure 4.1: Evenly distributed points on a sphere ($N = 700$), using the generalized spiral set method.

Figure 4.1 illustrates the distribution of the points for $N = 700$. The cartesian coordinate vectors **p** (x,y,z) that correspond to each sample point can then be calculated as follows ($\zeta = 1$):

$$\begin{aligned} x &= \zeta \cos\theta \sin\phi \\ y &= \zeta \sin\theta \sin\phi \\ z &= \zeta \cos\phi \end{aligned} \tag{4.3}$$

Each point in the sphere corresponds to some linear combination of the XYZ color matching functions. However, the XYZ functions themselves are not orthonormal. In order to use the points on the sphere as shown in Figure 4.1 to represent sensors, we need to first map the XYZ color matching functions to a new set of orthonormal functions, i.e new coordinate axes.

Orthonormal sensors can be found using singular value decomposition (SVD). If $\mathbf{C}$ denotes the $(m \times 3)$ XYZ color matching function matrix, then we can write $\mathbf{C}$ as:

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{4.4}$$

$\mathbf{U}$ is an orthonormal $m \times 3$ matrix, $\boldsymbol{\Sigma}$ is a diagonal matrix $(3 \times 3)$ of singular values, and $\mathbf{V}^T$ is an orthonormal $3 \times 3$ matrix. $\mathbf{U}$ is the orthonormal basis that we seek. By multiplying $\mathbf{U}$ with a linear transformation matrix $\mathbf{P}$ $(3 \times 3)$ that consists of three sample point vectors $\mathbf{p}_i$, $\mathbf{p}_j$, $\mathbf{p}_k$, a new matrix $\boldsymbol{\mathcal{C}}$ $(m \times 3)$ of color values can be derived that corresponds to a new RGB sensor set:

$$\boldsymbol{\mathcal{C}} = \mathbf{U}\mathbf{P}, \quad \mathbf{P} = [\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k] \tag{4.5}$$

Note that these newly derived RGB sensor sets $\boldsymbol{\mathcal{C}}$ are only orthonormal in wavelength if the vectors of matrix $\mathbf{P}$ are orthonormal.

If we take every triplet of points $\mathbf{p}_i$, $\mathbf{p}_j$, $\mathbf{p}_k$, on the sphere and post-multiply $\mathbf{U}$, we will generate a set of sensitivities that are uniformly distributed on a sphere in sensor space, resulting in a corresponding set of evenly distributed sensors. With reference back to equation 4.1, if we start with XYZ vectors, then $\mathbf{M}_{CAT}$ will equal to:

$$\mathbf{M}_{CAT}^T = [\boldsymbol{\Sigma}\mathbf{V}^T]^{-1}\mathbf{P} \tag{4.6}$$

While the sampling technique was derived starting with XYZ color matching functions, the algorithms can, of course, be adapted to using any kind of RGB sensors. It is important to point out that the sampling technique returns three sensors that have equal magnitude. In contrast, cone sensitivities are known to have different sensitivities: the short-wave mechanism is much less responsive than the long- and medium- wave mechanisms. Here we can avoid sensor magnitude because this variable is not important in the context of adaptation transforms (we are looking for scalars relating sensor responses across lighting conditions and these relative scalings are independent of the absolute magnitude of the sensors).

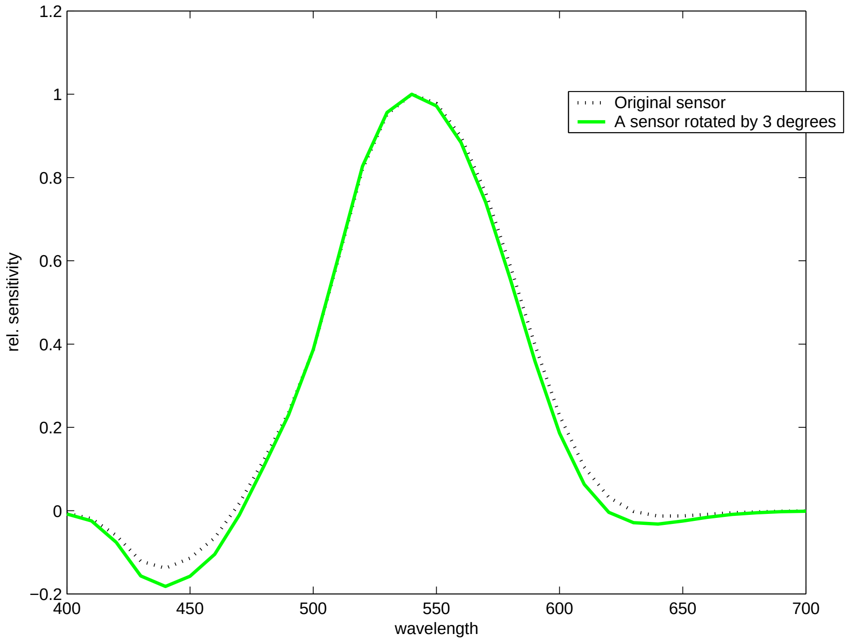


Figure 4.2: The corresponding spectral sensor sensitivities of two neighboring surface point vectors that are 3 degrees apart.

## 4.3 Sampling Experiment

We use the spherical sampling technique described in the previous section to find appropriate sensors for chromatic adaptation transforms. First, we determined the number of surface points necessary to uniformly sample the sphere. We found for $N = 5,000$, the angle between two neighboring surface point vectors varies between $3 - 5$ degrees, which we deemed sufficiently dense. Figure 4.2 illustrates the corresponding spectral sensitivities of two neighboring vectors that are 3 degrees apart. Therefore, the necessary combinations to check are equal to:

$$\frac{N!}{m!(N-m)!} = 2.08 \times 10^{10}, \quad N = 5,000, \quad m = 3 \tag{4.7}$$

Out of computational considerations, we can reduce the number of combinations by assuming that the sample points giving a positive result are located around the points that describe the Bradford, Sharp, and CAT02 transform. We determine the location of the vectors for the three transforms, and retain only the sample points that fall within 20 degrees of those points (see Figure 4.3). For $N = 5,000$, 240 “red,” 167 “green,” and 163 “blue” vectors fulfill that criterion, resulting in

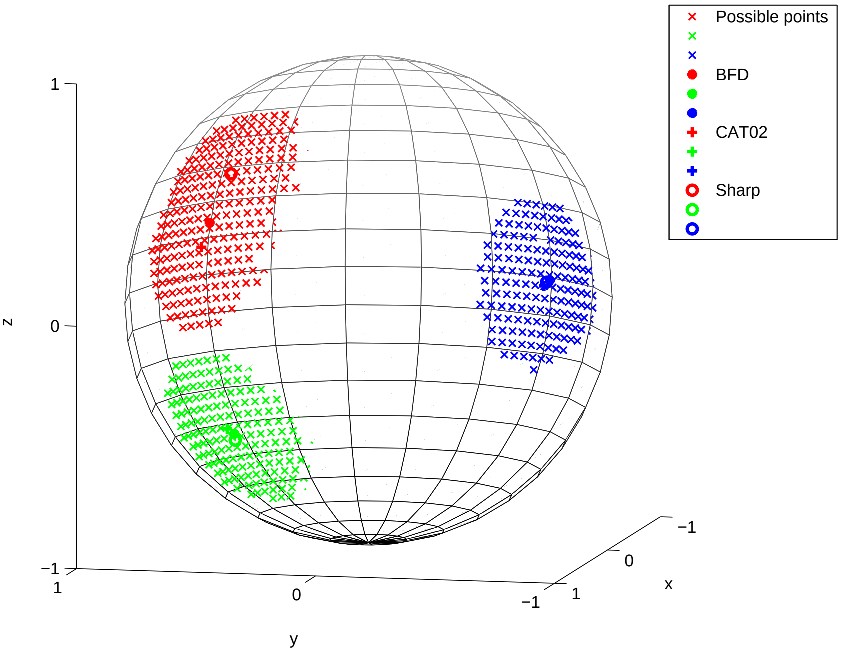


Figure 4.3: All sample points within a 20 degree radius of Bradford, Sharp and CAT02 (for $N = 5,000$).

$6.53 \times 10^6$ different sensor combinations $\boldsymbol{\mathcal{C}}$ (see eq. 4.5) to be tested. Note that these sensor combinations are not orthonormal in wavelength.

To evaluate if the resulting RGB sensors have good chromatic adaptation transform behavior, we use Lam's corresponding color data (see section 3.2) as input data. We chose this data set as it has a reasonable number of corresponding color pairs, and has been shown to be quite stable, i.e. CAT sensors derived solely from Lam's experimental data where able to predict corresponding colors obtained from other studies when used with a von Kries-Ives DMT-transform [LLR00, FS00a].

We chose the error criterion as follows. We have seen in the previous chapter that the Sharp CAT, Bradford CAT, and CAT02 RMS $\Delta E_{94}$ prediction errors for the same data set are 3.40, 3.54, and 3.45, respectively (see Appendix C, Table C.2). So for a first error criterion, we assume that all chromatic adaptation transforms resulting in a $\Delta E_{94} \leq 4$ for Lam's data set are adequate for predicting corresponding colors. Recall from Table C.2 that the RMS $\Delta E_{94}$ error is equal to 5.01 for von Kries operating on cone responses (HPE).

Of the 6.53 million possible chromatic adaptation transforms evaluated, 14,025 fulfilled the error criterion. See Figure 4.4 for the corresponding sampling points and Figure 4.5 for the corresponding RGB sensors. For clarity of illustration, the red, green, and blue sensors are plotted in different windows.

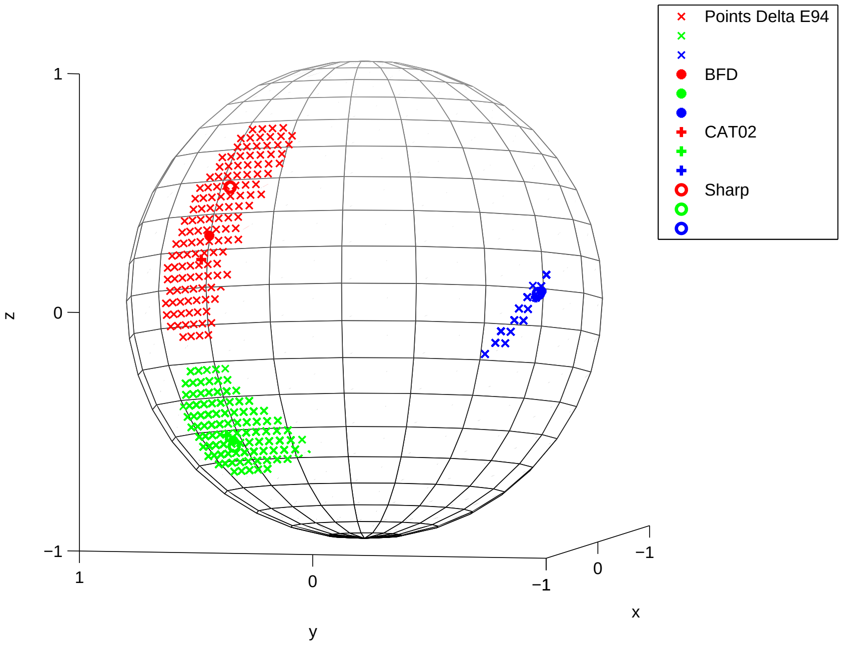


Figure 4.4: All sample points that result in sensor combinations with a RMS CIE $\Delta E_{94} \leq 4$ prediction error.

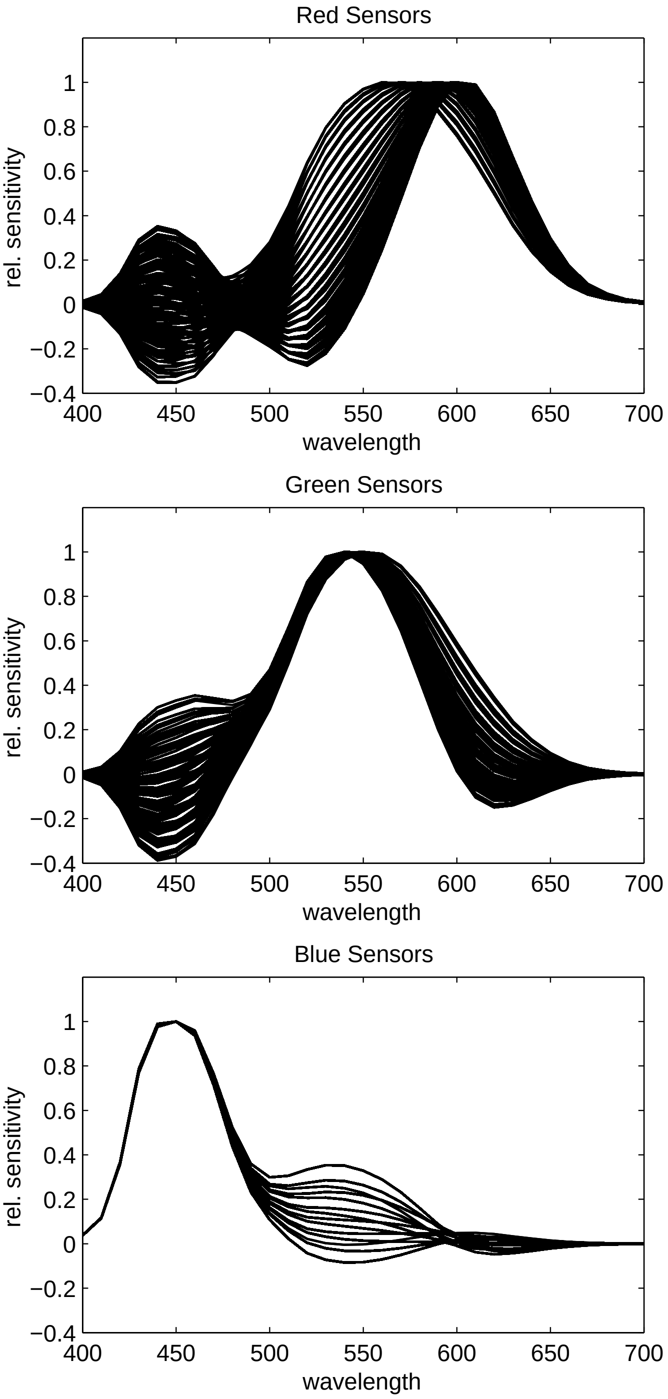
It is evident from Figure 4.5 that the largest sensor variability is in the red, followed by green and then blue. This can be illustrated using isometric surfaces, which indicate the sample points that correspond to sensors resulting in different $\Delta E_{94}$ (see Figure 4.6). To isometric surfaces were created by projecting the sphere surface points onto a plane. The distance between two points is an indicator of angular distance.

In Figure 4.6, the black line illustrates the convex hull of all sample points within 20 degree radius of Bradford, Sharp and CAT02 sensors. The other iso-contours are the convex hulls of $\Delta E_{94} = 4$ (blue -), $\Delta E_{94} = 3.9$ (green -), $\Delta E_{94} = 3.8$ (red -), $\Delta E_{94} = 3.7$ (yellow -), $\Delta E_{94} = 3.6$ (blue - -), $\Delta E_{94} = 3.5$ (green - -), $\Delta E_{94} = 3.4$ (red - -). Note that for the red, and partly for the green sensors, the sampling radius of 20 degrees was not sufficient to capture all possible sensors. For the red sensors, the isometric surface for $\Delta E_{94} = 3.7$ extends to the border of the isometric surface of all samples considered. For the green sensors, the limit is $\Delta E_{94} = 3.8$. Thus, the actual number of sensor combinations that result in a $\Delta E_{94} \leq 4$ is probably even higher than the 14,025 found here.

In Figure 4.6 and 4.7, the position of the Bradford sensors are indicated with a $\bullet$, the CAT02 sensors with a $+$, and the Sharp sensors with a $\circ$. The sample point corresponding to the best performing $\mathbf{M}_{CAT}$ is illustrated with a $\mathbf{x}$. It is interesting to note that Sharp sensors are closer to the best performing sensors for the red and green sensitivities, where as CAT02 is closer for blue. However, the RMS $\Delta E_{94}$ is almost identical for Sharp and CAT02.

The best sensor combination found through spherical sampling has a RMS error of $\Delta E_{94} = 3.37$. The best sensors are plotted in Figure 4.8, and the corresponding $\mathbf{M}_{CAT-Best}$ is:

$$\mathbf{M}_{CAT-Best} = \begin{bmatrix} 1.1083 & 0.0631 & -0.1714 \\ -0.8044 & 1.7798 & 0.0246 \\ 0.0000 & 0.0151 & 0.9849 \end{bmatrix}.$$


Red Sensors

Green Sensors

Blue Sensors

Figure 4.5: All RGB sensors that result in 14,025 different combinations with a RMS CIE $\Delta E_{94} \leq 4$ prediction error.

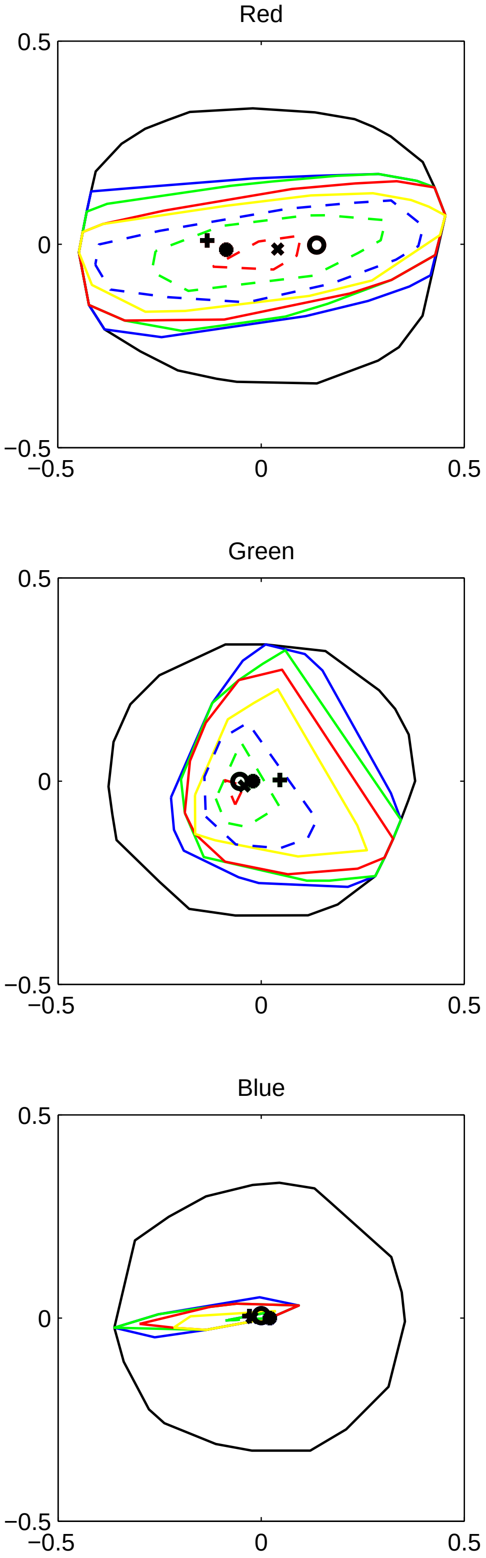


Figure 4.6: Isometric surfaces projected on a plane for different RMS CIE $\Delta E_{94}$ thresholds.

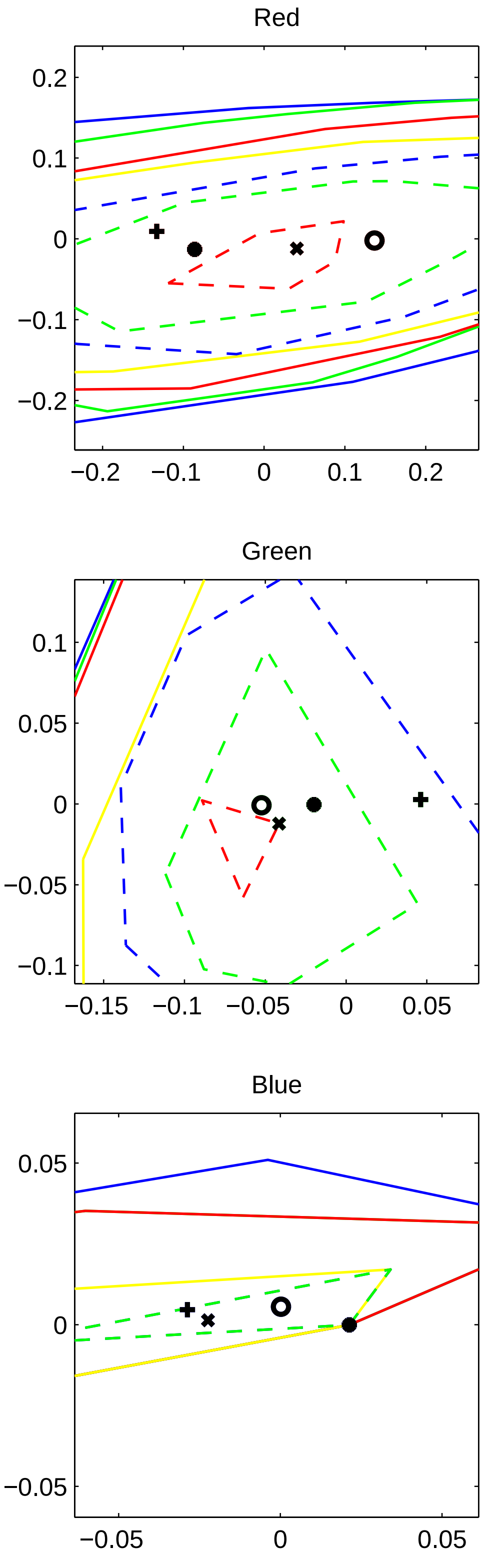


Figure 4.7: Close-up of the 4.6. See text for explanation.

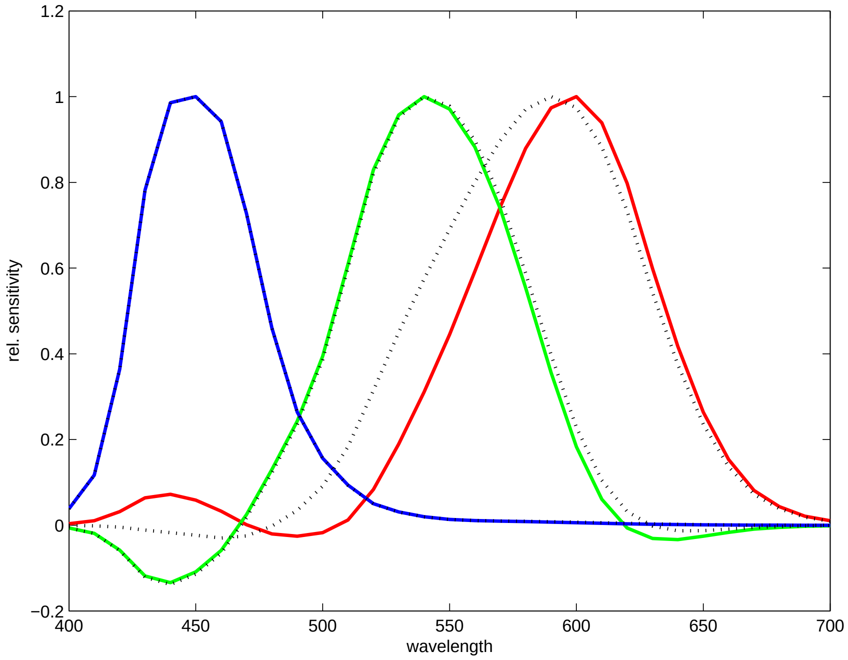


Figure 4.8: The best sensors found through spherical sampling (solid lines) compared to the CAT02 sensors (dotted lines).

## 4.4 Statistical Evaluation

The sampling criterion above was RMS $\Delta E_{94}$ prediction error for Lam's corresponding color data. While we can find a minimum that performs better than the existing CATs (RMS $\Delta E_{94}$=3.37), a slight difference in RMS error might not be statistically significant. It is of more interest to study how many other sensors perform statistically equivalently. Thus, one-tail student t-tests for matched pairs [WMM98] were additionally calculated to evaluate how many of the 14,025 RGB sensors resulted in a chromatic adaptation transform that was not statistically different from the CAT02 transform. We chose CAT02 for comparison, as it was implemented in the latest color appearance model CIECAM02 [MFH$^+$02] and the latest ICC color management profile specifications [ICC04] and thus is widely used. The null hypothesis was that the mean of the $\Delta E_{94}$ prediction errors of the sampled sensor CAT and those of CAT02 is equal to zero. The alternative hypothesis was that CAT02 has a smaller mean. A large $p$-value supports the null hypothesis, and a small $p$-value rejects it.

$p$-values equal or larger than 0.05 indicate that the means are equal (or that the null hypothesis cannot be rejected) at a confidence level of 95. At the 95 percent confidence interval, 1,056 RGB sensor sets remained. They result from a combination of 59 red, 31 green and 3 blue sampled sensors. The resulting sample points are shown in Figure 4.9, and the corresponding sensor sets in Figure 4.10. Again for clarity, the red, green, and blue sensors are plotted in different windows.

We also tested if CAT02 is statistically significantly different from the CAT with the transform $\mathbf{M}_{CAT-Best}$ that resulted in minimum RMS $\Delta E_{94}$ error (RMS $\Delta E_{94}$=3.37). We found that the null hypothesis could not be rejected ($p = 0.07$). Thus, corresponding colors clearly do not support a single unique von Kries type chromatic adaptation transform. Looking at Lam's data alone, there are probably at least 1,056 sensor sets to consider.

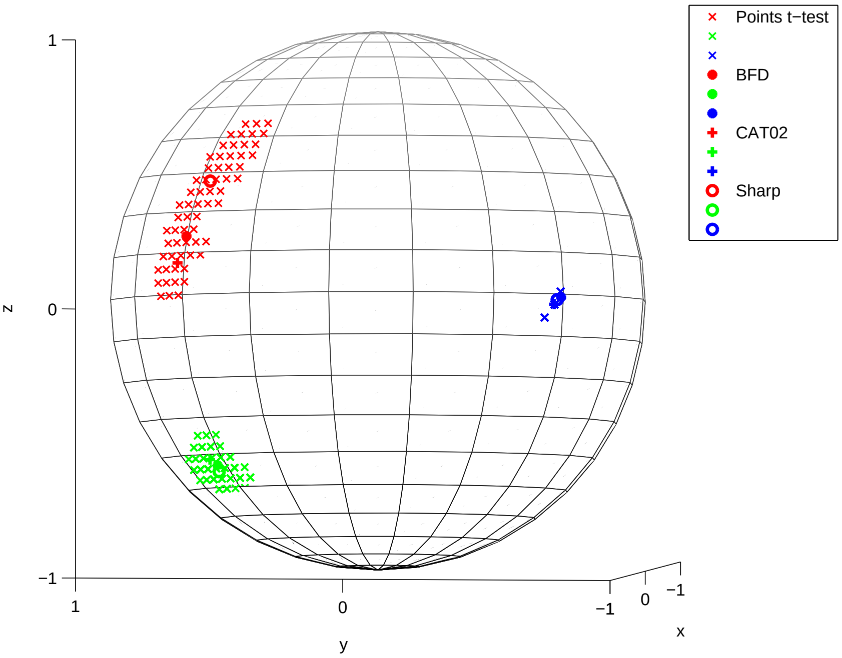


Figure 4.9: All sample points that result in sensor combinations that are not statistically significantly different from CAT02 at 95 percent confidence.

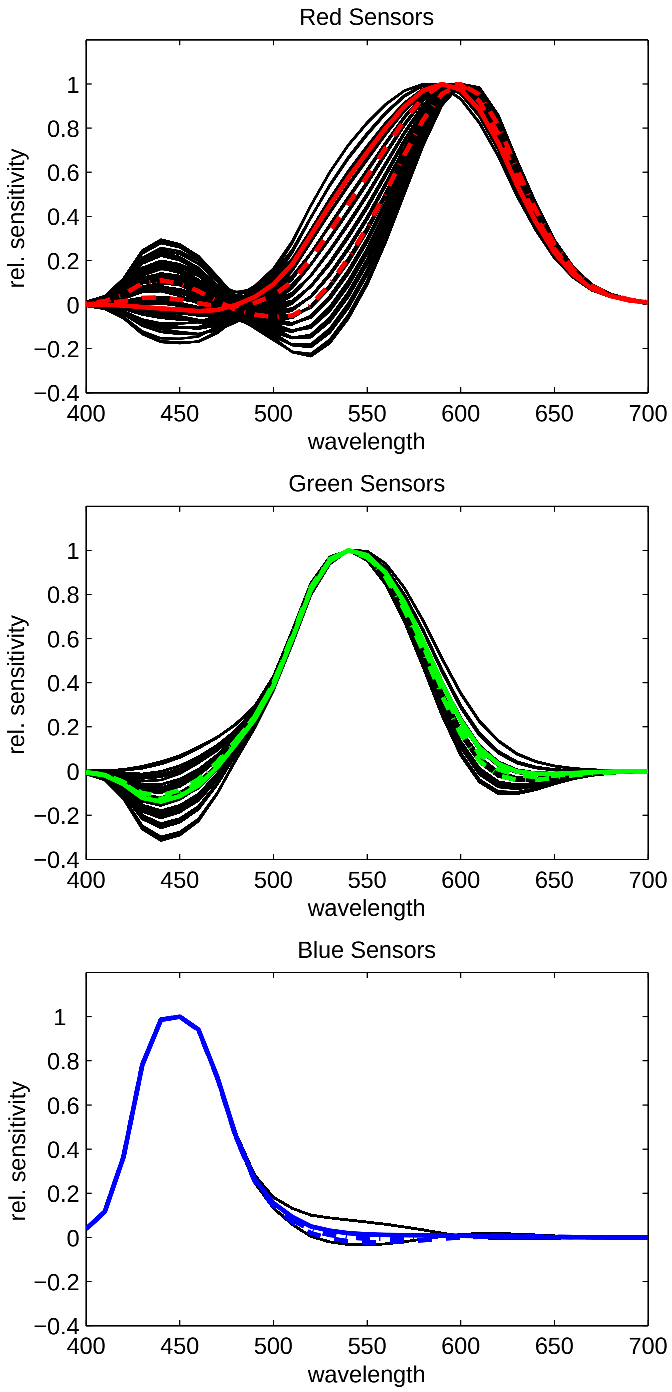


Figure 4.10: All RGB sensors (59 red, 31 green, and 3 blue for a total of 1,056 combinations) that are not statistically significantly different from CAT02 at 95 percent confidence. For comparison, the Sharp sensors, the Bradford sensors and the CAT02 sensors are also plotted.

## 4.5 Conclusions

We present a spherical sampling technique that can be used to evaluate and/or find linear color transformations. It has the advantage over other optimization techniques that it not only can easily find a global minimum, it can also return a set of solutions if so required. While we applied the technique to find possible chromatic adaptation sensors, it can easily be applied to find any three-dimensional sensors with equal magnitude.

We also show that the Bradford, Sharp, and CAT02 sensors are not unique. There is a number of other RGB sensors that exhibit the same favorable chromatic adaptation behavior. If the sampling distance is further decreased, the number would increase even more. This leads to the conclusion that there is either too much noise in the corresponding data set used to evaluate CATs, that the von Kries model used to implement chromatic adaptation transforms is too much of a simplification, or that there are many possible solutions to chromatic adaptation transforms and it is not critical which one is used.

As discussed in Appendix B, the design of the experiment can influence the corresponding color data one obtains. In other words, color constancy varies according to experimental conditions. This has lead many researchers to either propose extended chromatic adaptation transforms that include non-linearities [Mac63, NTS81, TSN84, Lam85, LW93], or models that determine the scaling coefficients by taking into account the responses of all cone classes [WW82, BW92, Bäu95, DB00]. However, many of the same authors agree that even though the strong von Kries DMT model cannot fully account for chromatic adaptation, it is at first approximation accurate enough to model colors under different illuminants, a result that was collaborated using images by Calabria and Fairchild [CF01].

Consequently, we speculate that in order to make a final choice on chromatic adaptation transforms, other secondary factors should be examined. The argument that sharp sensors are plausible from a human vision modeling point of view was made in the previous chapter. Camera sensors that are optimal with respect to color constancy are sharp [FD96, DF00]. Sharp sensors are also fairly close to RGB sensors that are the basis for colorimetric color image encodings [HHF99, SHF01].

Thus, sharp sensors might fit better within a color image workflow than the Bradford or CAT02 sensors, which do not have such correspondences.

# Chapter 5

# White-Point Independent RGB Sensors

## 5.1 Introduction

In the previous chapter, we have found a large number of sensor combinations that statistically result in the same chromatic adaptation performance than the most widely used CAT, CAT02. We speculate that in order for the color science and color imaging community to decide on chromatic adaptation transforms, secondary factors should be considered. In this chapter, we thus explore if the sensors found in the previous chapter could be used in color encodings[1]. These sensors are considered to be *white-point independent*, because the image RGB values are equal to the post-adaptation sensor responses (see section 5.2).

Two necessary characteristics for color encodings to be used in color imaging applications are good *gamut coverage* and *hue constancy.* Gamut coverage can be illustrated in a two-dimensional color space, such as $x, y$ or the more perceptually uniform $u^{'}, v^{'}$, or in a three-dimensional color space, such as CIE LAB. Figure 5.1 illustrates the $x, y$ and $u^{'}, v^{'}$ of two RGB primary sets with an encoding range [0,1], and Figure 5.2 the corresponding three-dimensional CIELAB gamut. The gamut boundaries correspond to the convex hull of all color responses possible with a given

---

[1] For the reader who is not familiar with color encodings, Appendix D provides a short introduction.

sensor set in a given encoding range. Thus, all color values contained within the gamut boundaries can be encoded, where as all color values outside of the gamut border cannot.

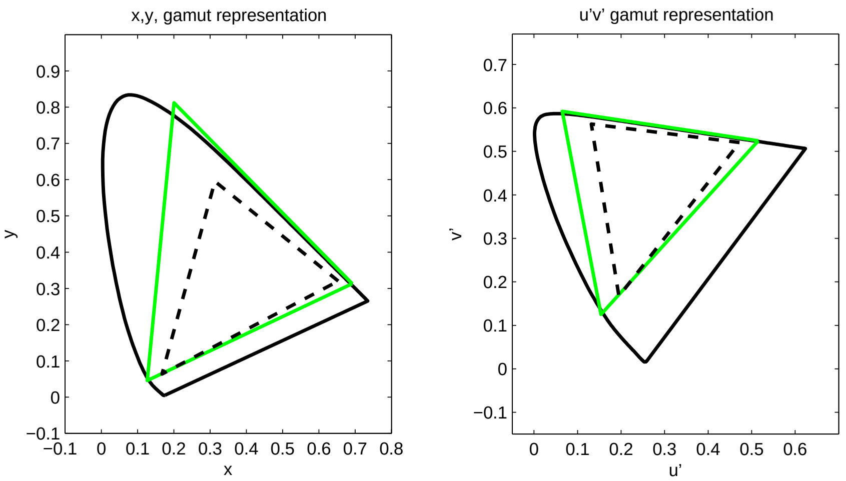


Figure 5.1: The $x, y$ and $u^{'}, v^{'}$ chromaticities of two RGB sensor sets in the encoding range of [0,1].

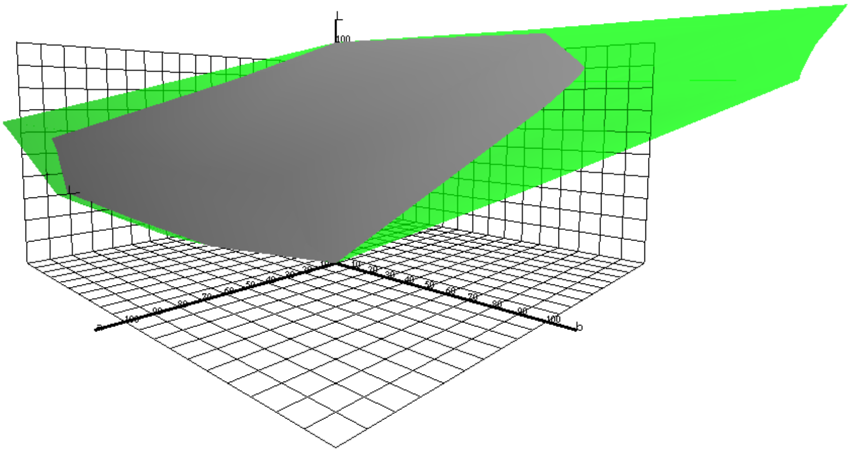


Figure 5.2: The CIE LAB gamut of two RGB sensor sets in the encoding range of [0,1].

In section 5.3, we present a method to classify the gamut coverage of the RGB sensors retained in section 4.4, namely those who are not statistically significantly different than the CAT02 sensors in predicting Lam's corresponding color data subject to using a strong von Kries-Ives coefficient model. We base the classification on the surface gamut compiled by Steingrimsson et al. [SSSS02]. Implicitly we assume that the sensors will be used as the basis for defining a color encoding intended for print applications, as other applications might strive for a different gamut coverage. The resulting best RGB sensors in terms of surface gamut coverage are sharp, i.e. narrow-band, emphasizing again that for imaging applications, the choice of a chromatic adaptation transform based on sharp sensors might be most appropriate [SFHF02].

Our second evaluation criterion is *hue constancy*, as RGB color encodings used in color imaging applications should have a high degree of hue constancy. In other words, a color ramp created by varying the encoding values to create different sensations of lightness or chroma (saturation) should still result in the same hue over the whole ramp. In section 5.4, we thus present an optimization technique to find hue constant RGB sensors [FS02b]. The hue representation is based on a log RGB opponent color space that is invariant to brightness and gamma.

In section 5.5, we combine both experiments and compare how the best sensors found in the gamut coverage experiment perform in terms of hue constancy. Using our measures, we find that they perform comparatively to current color image encoding sensors, and thus could be used in color encodings. We thus show that sharp sensors are optimal not only in terms of chromatic adaptation modeling, but do fulfill secondary factors that are important in color imaging applications.

## 5.2 White-point Independence

RGB encodings based on sensors that exhibit good chromatic adaptation behavior can be considered to be *white-point independent*, because the image RGB values are equal to the post-adaptation sensor responses. From the standpoint of chromatic adaptation, the linear image RGB values are an appearance description. With respect to eq. 2.3, they are the illuminant invariant descriptors of a surface color. When transforming these RGB values to the corresponding color XYZ values for a particular adopted white-point, the scaling for the destination adopted white is applied to the matrix that transforms from RGB to XYZ. This matrix is obtained by pre-multiplying matrix $\mathbf{M}^{-1}$ by the RGB scaling from an equi-energy adopted illuminant ($E(\lambda) = 1\ \forall\lambda$) to the destination adopted illuminant. Note that $\mathbf{M}^{-1}$ here is equal to $\mathbf{M}_{CAT}^{-1}$ of eq. 4.1.

For example, rendering RGB to XYZ under illuminant D65 is as follows:

$$\begin{bmatrix} X^{D65} \\ Y^{D65} \\ Z^{D65} \end{bmatrix} = (\mathbf{M}^{-1}\mathcal{D}^{D65}) \begin{bmatrix} R_{lin} \\ G_{lin} \\ B_{lin} \end{bmatrix} \tag{5.1}$$

where

$$\mathcal{D}^{D65} = \begin{bmatrix} R_w^{D65} & & \\ & G_w^{D65} & \\ & & B_w^{D65} \end{bmatrix}$$

$R_w^{D65}, G_w^{D65}, B_w^{D65}$ are obtained by multiplying $X_w^{D65}, Y_w^{D65}, Z_w^{D65}$ by $\mathbf{M}$. The $R_w^e, G_w^e, B_w^e$ of an equi-energy illuminant are by definition equal to 1.

Likewise, converting XYZ values for a particular adopted white to the RGB appearance description involves applying the inverse of the matrix $(\mathbf{M}^{-1}\mathcal{D}^{D65})$ that is used to transform from the RGB values to the adopted white XYZ values. It can be assumed that color image processing would become computationally "cheaper," and quantization errors would be minimized if images were encoded based on RGB sensors that exhibit such favorable chromatic adaptation behavior.

## 5.3 Gamut Coverage Optimization

Through the spherical sampling technique described in chapter 4, we found a large number (1,056) of XYZ to RGB transforms ($\mathbf{M}$) that are not statistically significantly different from CAT02 in predicting corresponding colors. The goal here is to find among them the RGB to XYZ transforms ($\mathbf{M}^{-1}$) and the corresponding RGB sensors that wholly enclose the gamut of optimal surface colors while minimizing the "waste" of encoding colors outside of the surface gamut.

As discussed in Appendix D.2, a color encoding is always based on a specific sensor space, but additionally includes a digital encoding method. Integer digital encodings linearly specify the *digital code value range* associated with the *color space* *range*. The color space range defines the maximum and minimum digital values that are represented in the digital encoding. Most RGB color space ranges will typically be defined as $[0, 1]$, and values outside this range are clipped to the maximum and minimum values. The digital code value range defines the minimum and maximum integer digital code values corresponding to the minimum and maximum color space values. For example, an 8-bit per channel encoding for an RGB color space with range $[0, 1]$ will associate the digital code value 0 to the color space value 0, and 255 to 1, respectively. If the gamut of a color encoding is large, or the color space range is larger than $[0, 1]$, more bits need be associated with the color space range to avoid quantization artifacts when editing and/or displaying an image. However, increasing the number of bits per channel increases the file size of an image.

Output-referred color image encodings are based on color encodings and are usually optimized for output on a real or virtual output device. As such, only colors that can be displayed on that device need to be encoded. The digital values beyond the color gamut are "wasted," i.e. they will never be displayed. In case of output-referred color image encodings intended for print, the gamut surface color is a reasonable criterion for comparison; a color that is not a surface color will never be printed. Defining a color encoding that has approximately the same gamut coverage as the surface colors will minimize the number of bits needed to encode the surface gamut, and thus minimize the image file size, while ensuring that most printable colors are encodable.

The gamut obtainable with surface colors used in this study was derived by Steingrimsson et al. [SSSS02] by calculating the convex hull from real surface reflectances (Pointer [Poi80], Pantone, Munsell [Mun76], and real printer ink reflectances) and then converting them to CIE L*a*b* under D50 (see section 1.3). The data consists of gamut boundary points at different hue angles for different lightness values, ranging from L*=15 and L*=90 in intervals of 5 L*, for a total of 16 L* levels.

Using surface colors to define the target gamut and evaluating the gamut under a D50 white-point implies that the resulting color encoding is optimized for print reproduction. However, the method presented here works with any target gamut, and can be used to find the optimal RGB sensors for any encoding intent.

### 5.3.1 Experiment

A matrix $\mathbf{S}$ $(3 \times 37,000)$ of RGB values between $[0,1]^2$ describing the RGB cube surface was transformed to a matrix $\mathbf{P}$ $(3 \times 37,000)$ of XYZ under D50 through $\mathbf{M}^{-1}$, which was pre-multiplied by the RGB scaling $\mathcal{D}^{D50}$ to XYZ under D50:

$$\mathbf{P} = (\mathbf{M}^{-1}\mathcal{D}^{D50})\mathbf{S} \tag{5.2}$$

This operation was done 1056 times, for each of the sensor combinations found in 4.4.

For all $\mathbf{P}$, the resulting XYZ values were transformed to CIELAB under D50 using eq. 1.10, resulting in 1056 $\mathbf{J}_{\mathrm{LAB}}$ $(3 \times 37,000)$ matrices. For the same L* values and intervals as supplied by Steingrimsson et al.'s data for the surface gamut, we extracted the subset of corresponding $L^*, a^*, b^*$ values from our $\mathbf{J}$ matrices. For example for L*=50:

$$\mathbf{J}_{50\mathrm{L}^*} = \{L^*, a^*, b^*\} \quad \forall L^* = 50, a^*, b^* \in \mathbf{J}_{\mathrm{LAB}} \tag{5.3}$$

For all intervals of L*=5, between L*=15 and L*=90, the corresponding (1056) $\mathbf{J}_{\mathrm{L}^*}$ matrices were calculated. Finally, we calculated the convex hull for each $\mathbf{J}_{\mathrm{L}^*}$, which gives us a gamut boundary for each level of L*. The polygon area $A_{\mathrm{L}^*,RGB}$ enclosed within the convex hull corresponds to the gamut of color responses that a given RGB sensor set can encode at that L* value. The surface gamut $A_{\mathrm{L}^*,surface}$ is provided by the convex hull coordinates supplied by Steingrimsson et al. Figure 5.3 illustrates for L*=50 the areas contained within the convex hulls of a RGB sensor combination and the surface colors.

We now have to derive a measure that evaluates the applicability of a sensor set in terms of gamut coverage. Considering that our color coordinates are expressed in a perceptual color encoding (CIELAB), where we assume that equal Euclidean distances correspond to equal perceived color differences, an intuitive measure would

---

[2] By using this range of code values, we assume that the useful encoding range is between $[0,1]$.

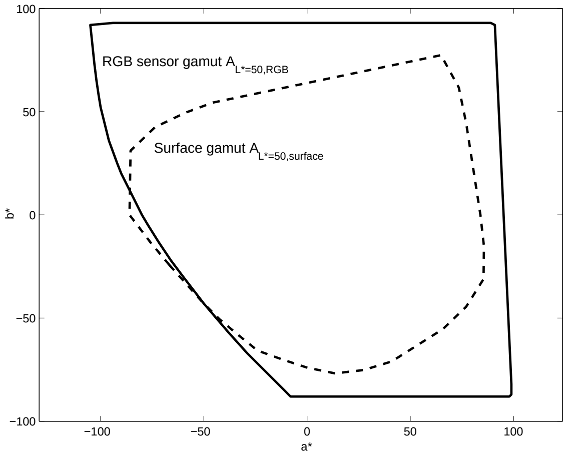


Figure 5.3: The gamut area $A_{\mathrm{L}^*,RGB}$ contained within the convex hull of a RGB sensor set (solid) and the and gamut area $A_{\mathrm{L}^*,surface}$ of the surface colors (dash) at L*=50.

be to consider the percentage of gamut area in common. We thus first define the *encodable gamut* that describes the surface colors that can be encoded with a given RGB sensor set as the proportion of surface colors contained within the RGB sensor gamut area, compared to the whole surface gamut:[3]

$$G_{en,\mathrm{L}^*} = \frac{A_{\mathrm{L}^*,surface} \cap A_{\mathrm{L}^*,RGB}}{A_{\mathrm{L}^*,surface}} \tag{5.4}$$

where

$$A_{\mathrm{L}^*,surface} \cap A_{\mathrm{L}^*,RGB} =$$
$$\{a^*_{\mathrm{L}^*,surface}, b^*_{\mathrm{L}^*,surface} | a^*_{\mathrm{L}^*,surface}, b^*_{\mathrm{L}^*,surface} \in A_{\mathrm{L}^*,surface} \cdots$$
$$\ldots \wedge a^*_{\mathrm{L}^*,surface}, b^*_{\mathrm{L}^*,surface} \in A_{\mathrm{L}^*,RGB}\}$$

$G_{en,\mathrm{L}^*}$ will take a value between $[0,1]$; zero if no surface color responses are encodable, one if all the surface colors are contained within the convex hull of the RGB sensors. With respect to Figure 5.3, the encodable gamut is 0.96.

However, the encodable gamut is not the only measure that is important. Recall that each color encoding does specify not only a color space range per channel, $[0,1]$ in our experiment, but also a digital code value range (see Appendix D.2), i.e. 8-bits ($[0 \cdots 255]$), 10-bits ($[0 \cdots 1023]$), etc. The number of code values determines the

---

[3] The intersections of surface and RGB sensor gamut areas were obtained using the Spatial and Geometric Analysis Toolbox (SaGA) for Matlab developed by Pankratov [Pan].

quantization of the color space range. Not associating enough code values to encode the color space range can result in visible quantization artifacts. Thus, large gamut encodings define a higher number of bits per channel (10-bits or more), whereas small gamut encodings are usually limited to 8-bits per channel. The number of bits chosen to encode a color response has, of course, a direct influence on the file size of an image.

If the encodable gamut is large, we thus need to quantize it with more bits. However, if the encodable gamut is much larger than the *useful gamut*, i.e. the color responses that need to be encoded for a specific application, then code values are "wasted." In other words, images rendered to that specific application will not use any code values beyond those that describe the intended gamut, in our case the surface gamut. Thus, we also define a measure for the useful gamut, which is the proportion of the surface gamut contained within the RGB sensor gamut. A larger useful gamut implies that fewer code values outside of the surface gamut are wasted.

$$G_{us,\mathrm{L}^*} = \frac{A_{\mathrm{L}^*,surface} \cap A_{\mathrm{L}^*,RGB}}{A_{\mathrm{L}^*,RGB}} \tag{5.5}$$

Now we have two measures that define the applicability of a RGB sensor encoding to describe surface colors for different L* values. To get single measures over all L* values, we average the results of the individual $G_{en,\mathrm{L}^*}$ and $G_{us,\mathrm{L}^*}$ over the 16 L* levels:

$$G_{.} = \frac{1}{16} \sum_{\mathrm{L}^*=15:5:90} G_{.,\mathrm{L}^*} \tag{5.6}$$

### 5.3.2 Results

In Figure 5.4, the gamuts of the 20 best transforms, i.e. the ones that resulted in the largest useful gamuts, are plotted at different values of L*. In addition, the surface gamuts and the gamuts obtained by the ROMM/RIMM and ITU-R.BT 709 sensors are also illustrated. The ROMM/RIMM RGB sensors are the basis for the ROMM/RIMM color image encodings [ANS02b, ANS02a], and the ITU-R.BT 709 sensors for the sRGB, e-sRGB and scRGB [IEC99, ANS01, IEC03] color image encodings, which are widely used in imaging applications. The transform matrices corresponding to the 20 best sensors are listed in Appendix E. For reference, the

$x', y'$ gamuts are also plotted in Figure 5.5, and in Figure 5.6 the corresponding RGB sensors.

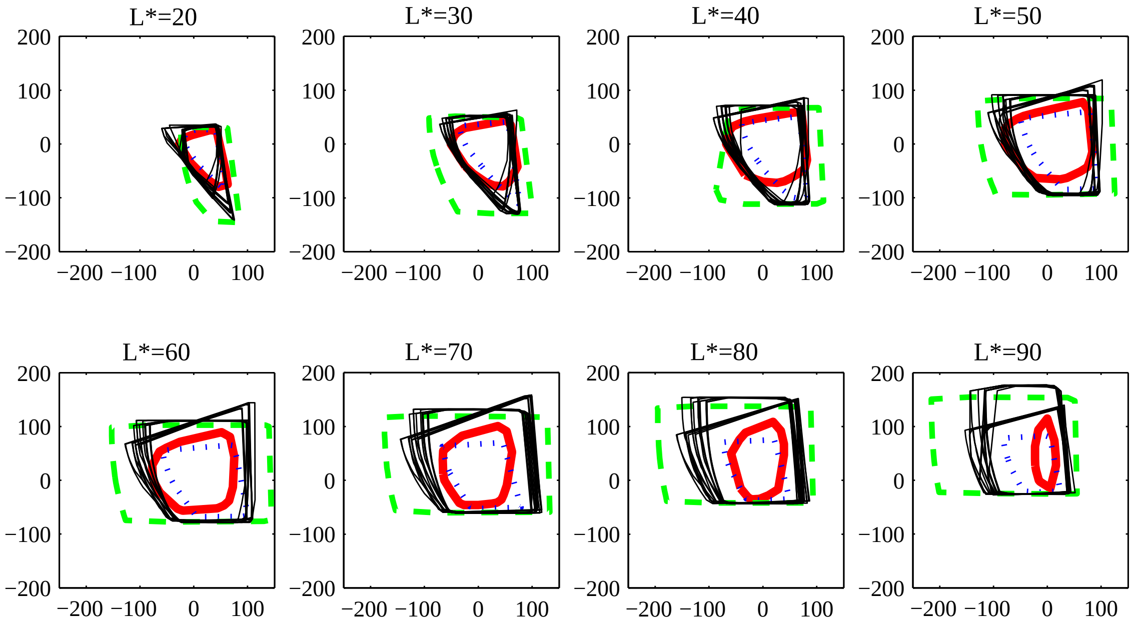


Figure 5.4: $a*, b*$ gamuts at different values of L*: surface (red bold solid line), 20 best RGB (black solid lines), ROMM/RIMM (green dashed line), ITU-R.BT 709 (blue dotted line).

As can be seen in table 5.1, the best RGB sensor set can generally encode almost all surface colors, but also have "waste," i.e. they can encode RGB values that do not have corresponding surface reflectances. For comparison, the table also contains the useful and encodable gamuts of the ROMM/RIMM RGB sensors and ITU-R.BT 709 sensors. ROMM sensors, which were optimized for print reproduction, performed as well as our sensors in terms of encodable gamut. However, they performed worse (by approx. 20 percent) in terms of useful gamut. Conversely, the ITU-R.BT 709 sensors, which are mostly used for monitor color image encodings, have a high percentage of useful gamut but cannot encode as many surface colors. CATs based on these sensors have previously been considered as chromatic adaptation transforms, and it was found that they do not perform as well as the most popular CATs [SHF01].

Table 5.1: The encodable $G_{en}$ and useful $G_{us}$ gamut percentages of the best 20 transforms.

| Sensors (Transforms) | % of encodable gamut $G_{en}$ | % of useful gamut $G_{us}$ |
|---|---|---|
| $\mathbf{M}_1$ | 98.88 | 52.56 |
| $\mathbf{M}_2$ | 98.93 | 52.18 |
| $\mathbf{M}_3$ | 95.90 | 53.86 |
| $\mathbf{M}_4$ | 96.71 | 52.53 |
| $\mathbf{M}_5$ | 94.60 | 53.87 |
| $\mathbf{M}_6$ | 98.55 | 51.11 |
| $\mathbf{M}_7$ | 96.35 | 52.61 |
| $\mathbf{M}_8$ | 93.73 | 54.32 |
| $\mathbf{M}_9$ | 98.15 | 51.31 |
| $\mathbf{M}_{10}$ | 97.94 | 51.27 |
| $\mathbf{M}_{11}$ | 95.83 | 52.71 |
| $\mathbf{M}_{12}$ | 95.05 | 53.23 |
| $\mathbf{M}_{13}$ | 95.58 | 52.66 |
| $\mathbf{M}_{14}$ | 97.05 | 51.41 |
| $\mathbf{M}_{15}$ | 98.52 | 50.43 |
| $\mathbf{M}_{16}$ | 99.43 | 49.72 |
| $\mathbf{M}_{17}$ | 96.10 | 51.40 |
| $\mathbf{M}_{18}$ | 91.61 | 54.55 |
| $\mathbf{M}_{19}$ | 98.27 | 49.89 |
| $\mathbf{M}_{20}$ | 98.86 | 49.46 |
| $\mathbf{M}_{ROMM/RIMM}$ | 99.97 | 30.81 |
| $\mathbf{M}_{ITU-R.BT709}$ | 73.82 | 79.43 |
| $\mathbf{M}_{Sharp}$ | 99.82 | 40.77 |
| $\mathbf{M}_{CAT02}$ | 100.00 | 18.44 |

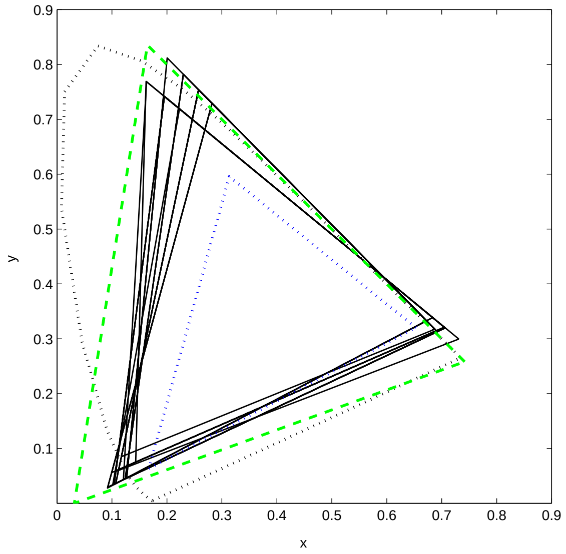


Figure 5.5: $x, y$ color gamuts of the 20 best transforms, compared to ROMM/RIMM (green dashed line) and ITU-R.BT 709 (blue dotted line).

As expected, the mean encodable gamuts of the Sharp (eq. 3.13) and CAT02 sensors are sufficient to encode almost or all surface colors, but the useful gamut is relatively small. Indeed, the useful gamut of the Sharp sensors is even slightly larger than that of the ROMM/RIMM RGB transform. While the CAT02 sensors are white-point independent, they do not have good gamut coverage properties. The ROMM and ITU-R.BT 709 sensors, which have optimized gamut coverage for their respective usage, are not white-point independent.

All the sensors found in this experiment are white-point independent and have better gamut properties than CAT02. Compared to CAT02, the optimal RGB sensors in terms of gamut coverage are much more narrow-band than the CAT02 sensors (see Figure 5.6), and resemble the Sharp sensors (see Figure 3.4). This is expected when comparing the percentages of useful and encodable gamuts of the Sharp sensors, the CAT02 sensors and the best sensors (see Table 5.1).

## 5.4 Hue Optimization

The second characteristic necessary for color encodings we investigate is *hue constancy*. Hue is an important property of a color. As defined by the CIE, hue is

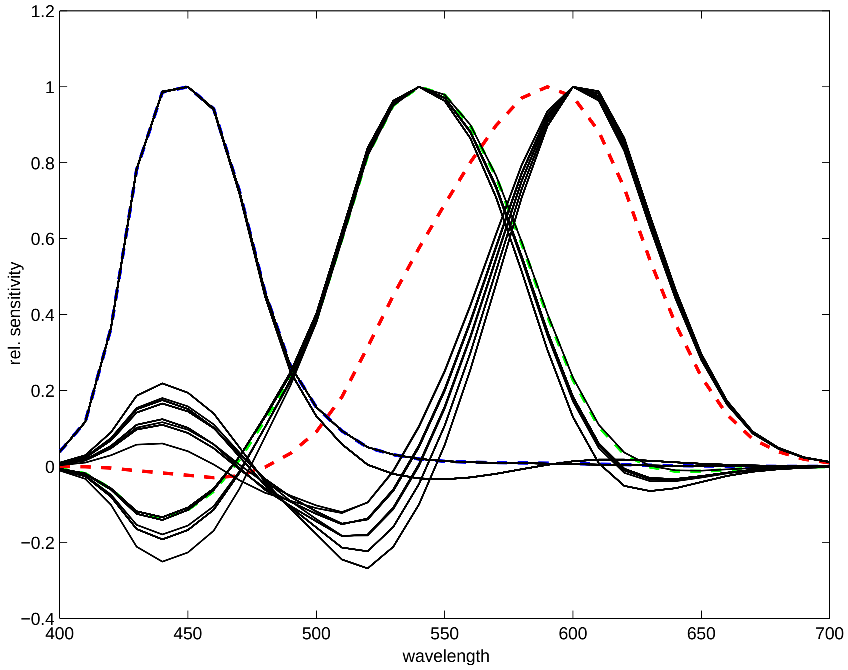


Figure 5.6: 20 RGB sensor sets that resulted in the largest useful gamuts, compared to CAT02.

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 [CIE87]. In other words, it is the "name" of a color, and is one of the perceptual correlates like saturation and brightness.

Hue is the attribute of a color that people generally find easiest to identify. It is most often used to "describe" a color. Thus, segmenting images according to hue has been widely used for object segmentation, object recognition and image retrieval [PK94, MKNM95, GS99]. In the computer vision community, image RGB values are converted to a hue-based representation using a color transformation like HSV, HLS and IHS [SH98]. These transformations do not require a-priori knowledge about the RGB values, i.e. an exact definition linking RGB and tristimulus values (XYZ) is not needed. While not knowing the exact encoding parameters precludes the calculations of the CIE definition of hue (see section 1.3), the results are acceptable for the applications mentioned above.

In the color imaging community, it has been noted that RGB color image encodings should have a high degree of hue constancy. In other words, a color ramp created by varying the encoding values to create different sensations in lightness or chroma (saturation) should still result in the same hue over the whole ramp. Additionally, simple non-linear channel editing should not affect the hue of a color [Mor01]. In effect, hue constancy was one of the optimization criteria used in the development of the ROMM RGB color encoding [SWG01]. Their optimization was based on the CIE LAB hue definition (see eq. 1.13).

We investigate if a hue based representation that is invariant to brightness and gamma [FS01a], developed for the computer vision community, can be used to evaluate the hue constancy of RGB sensors. We use spherical sampling (see chapter 4) to find a sensor set that maximizes straight hue lines for psychophysically derived constant hue data [HB95] in a log RGB opponent color space. Using this definition of hue constancy, we additionally calculate the hue behavior of the ROMM, ITU-BT.R 709, HPE, Sharp (eq. 3.13, CAT02, and Bradford sensors [FS02b].

The hue definition of the log RGB opponent color space is similar to the hue definitions in other color spaces. It has been successfully applied to image retrieval applications [FS01a]. It is gamma invariant, i.e. the power function usually applied to any RGB encoding cancels out. However, XYZ based hue definitions, such as $h_{a*b*}$ (see section 1.3) and $H_{PT}$ of the IPT opponent encoding [EF98a], are calculated with a power function. The power functions are equal to 1/3 and 0.43 for CIE LAB and IPT, respectively. Consequently, LAB $h_{a*b*}$ and IPT $H_{PT}$ are strictly speaking applicable only to one given "contrast" encoding. While this contrast encoding is based on the encoding of the human visual system, it does not necessarily reflect other color image encodings that might be appropriate for certain applications like computer graphics.

We are not claiming here that the opponent log RGB hue definition is the best measure for defining visual hue constancy. However, it might be an appropriate tool to test hue constancy in terms of defining RGB color image encodings, i.e. it might be suitable for color engineering purposes as opposed to color vision modeling.

### 5.4.1 Brightness and Gamma Invariant Hue

The log RGB opponent color encoding [FS01a] was developed on the following premise. In imaging applications, linear RGB signals captured by the digitizing device are usually encoded with a power function to compensate for system nonlinearity such as the monitor transfer function. Therefore, the RGB vectors encoded for each pixel are equal to:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \alpha R_{lin}^{\gamma} \\ \alpha G_{lin}^{\gamma} \\ \alpha B_{lin}^{\gamma} \end{bmatrix} \tag{5.7}$$

$\alpha$ is a scalar that compensates for the illuminance. Applying a log transform to the RGB values removes the $\gamma$ term from the exponent and turns them into multiplicative scalars:

$$\log \begin{bmatrix} \alpha R_{lin}^{\gamma} \\ \alpha G_{lin}^{\gamma} \\ \alpha B_{lin}^{\gamma} \end{bmatrix} = \begin{bmatrix} \log(\alpha) + \gamma \log R_{lin}^{\gamma} \\ \log(\alpha) + \gamma \log G_{lin}^{\gamma} \\ \log(\alpha) + \gamma \log B_{lin}^{\gamma} \end{bmatrix} \tag{5.8}$$

Brightness $\alpha$ becomes an additive rather than a multiplicative term. Taking differences of color channels, i.e. projecting orthogonal to the unitary vector (1,1,1) allows to remove the brightness term:

$$\begin{bmatrix} \log(\alpha) + \gamma \log R_{lin} \\ \log(\alpha) + \gamma \log G_{lin} \\ \log(\alpha) + \gamma \log B_{lin} \end{bmatrix} \mapsto \begin{bmatrix} \gamma \log(R_{lin}) - \gamma \log(G_{lin}) \\ \gamma \log(R_{lin}) + \gamma \log(G_{lin}) - 2\gamma \log(B_{lin}) \end{bmatrix} \tag{5.9}$$

Note that the definitions of the above differences describe coordinates in an opponent color representation. They are similar to the opponent color axes used by the human visual system and could be regarded as having perceptual relevance (see section 1.2.2).

Finally, ratios of the opponent color coordinates are formed to cancel $\gamma$:

$$\frac{\gamma \log(R_{lin}) - \gamma \log(G_{lin})}{\gamma \log(R_{lin}) + \gamma \log(G_{lin}) - 2\gamma \log(B_{lin})} = \frac{\log(R_{lin}) - \log(G_{lin})}{\log(R_{lin}) + \log(G_{lin}) - 2\log(B_{lin})} \tag{5.10}$$

Hue is defined as the inverse tangent of the ratio of eq. 5.10:

$$H = \tan^{-1} \frac{\log(R_{lin}) - \log(G_{lin})}{\log(R_{lin}) + \log(G_{lin}) - 2\log(B_{lin})} \tag{5.11}$$

The definition of hue is very similar to the CIE $h_{a*b*}$ definition, which defines hue as the inverse tangent of the ratio of the opponent color coordinates $b*$ (blue-yellow) and $a*$ (red-green), see eq. 1.13. What is noteworthy about equation 5.11 is that it follows only from the engineering imperative to cancel brightness and gamma. There was no attempt made to model visual response. Rather, mathematically deriving gamma and brightness independence led to a visual system like definition of hue.

### 5.4.2 Experiment

The goal of the experiment is to find RGB sensors that keep hue constant and to evaluate the hue constancy of known sensors. As initial data sets, we used the hue-constant data published by Hung and Berns [HB95] (see Figure 5.7). They derived XYZ tristimulus values, four values at each hue angle, for 12 hue angles, through a psychophysical experiment on a CRT monitor. The original data set was defined under illuminant C. We used the Sharp chromatic adaptation transform ($\mathbf{M}_{Sharp}$, see eq. 3.13) to calculate the corresponding colors under illuminant D65, so that the results better correspond to previous hue-constancy experiments [Mor01, EF98b].

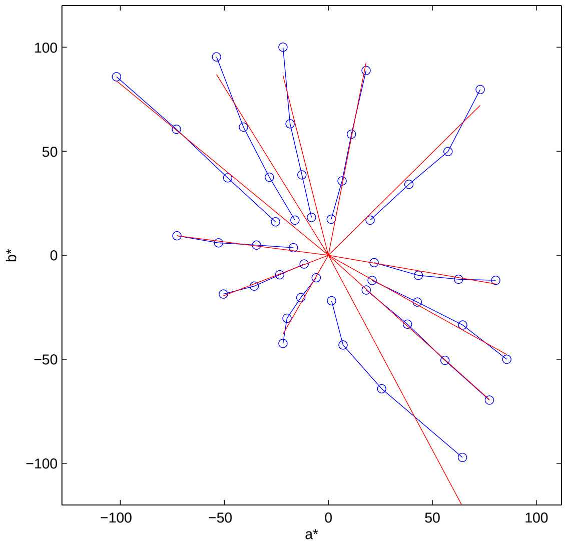


Figure 5.7: The Hung and Berns [HB95] constant hue data, plotted in $a*, b*$ under D65. This Figure illustrates the hue inconstancy of CIE Lab. If the data is hue constant, it should lie on a straight line through the origin.

The spherical sampling technique as described in chapter 4 was applied to find the best sensor that keeps hue most constant. Out of computational consideration, we tested only sensors that had their endpoints located within 30 degrees of the ITU BT.R. 709 and ROMM red, green, and blue sensors, resulting in ∼ 6 million different RGB sensor sets **R** to be tested (see Figure 5.8). The ITU BT.R. 709 sensors and the ROMM sensors are considered to be somewhat hue constant [Mor01, WSG02].

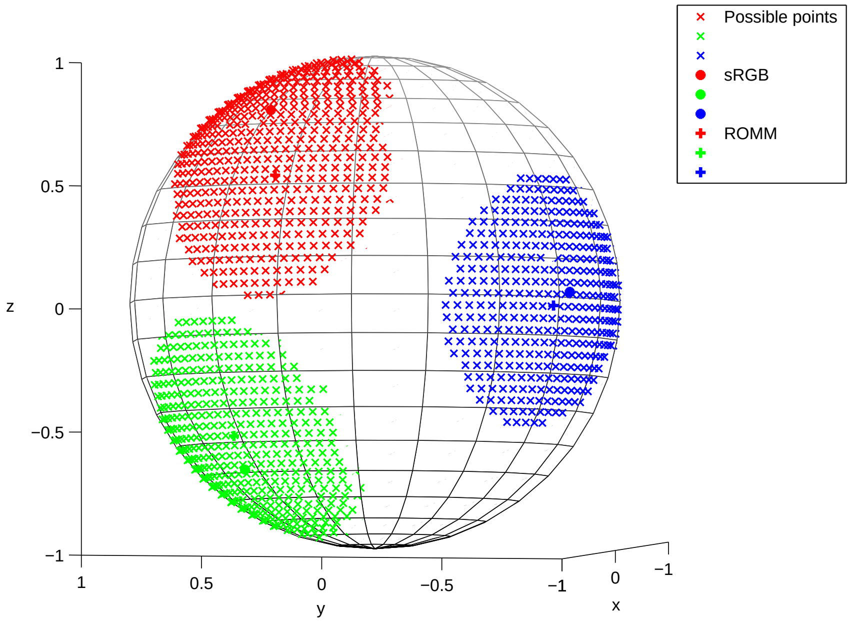


Figure 5.8: Sample points around 30 degrees of the ITU-R.BT 709 and ROMM sensors considered in the optimization.

The hue-constant XYZ values were converted to linear RGB using the (3x3) linear transforms **M** found through the spherical sampling technique. The Hung and Berns data set consists of 12 hue angles, with four points for each hue, resulting in 48 XYZ values. If **P** is a (48x3) matrix of normalized XYZ values, and **S** is a (48x3) matrix of linear RGB values, then:

$$\mathbf{S} = \mathbf{P}\mathbf{M}^T \tag{5.12}$$

Not all transforms result in positive R, G, and B values. Before converting to log space, the values need to be offset to render them all positive. This offset can be considered as adding brightness to the RGB channels, or defining a different encoding range. Once all the RGB values are rendered positive, the log RGB opponent color matrix $\mathbf{O}$ (48x2) can then be calculated as described in eqs. 5.8 to 5.10.

Perfectly hue constant data in an opponent color encoding should have the same hue angle. Data points with equal hue should thus lie on a line going through the origin (see Figure 5.7). Calculating the deviation from the line, i.e. how far the data points actually are from the line representing the hue angle, gives an indication of hue constancy. Therefore, the problem of finding the most hue constant RGB sensor can be expressed as a line-fitting problem.

Using singular value decomposition, we can find a line that best fits a set of data points minimizing least squares error. However, we have the additional constraint that the line needs to go through the origin. Adding a "mirrored" point to each of the opponent log RGB matrix entries will result in a hue matrix $\mathbf{H}$ (96x2) with a mean equal to zero:

$$\mathbf{H} = [\mathbf{O}; \mathbf{O} \times (-1)] \tag{5.13}$$

For better comparisons between the different sensor results, $\mathbf{H}$ was transformed linearly to a new matrix $\tilde{\mathbf{H}}$ that is "white," i.e. its components are uncorrelated and covariances equal zero. From eq. 5.13 it is clear that $\mathbf{H}$ is already centered, i.e. it has zero mean. Therefore, the "whitening" transformation is based on the eigenvalue decomposition of the covariance matrix ($\mathbf{H}^T\mathbf{H}$ of $\mathbf{H}$):

$$\mathbf{H}^T\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \tag{5.14}$$

where $\mathbf{E}$ is a matrix of eigenvectors, and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues. $\tilde{\mathbf{H}}$ is calculated as follows:

$$\tilde{\mathbf{H}} = \mathbf{H}(\mathbf{U}\mathbf{\Lambda}^{-1/2})^T \tag{5.15}$$

As the best fitting line has to be calculated for each hue angle separately, $\tilde{\mathbf{H}}$ was divided into $n$ ($n = 12$) matrices $\tilde{\mathbf{H}}_n$ (8x2), each containing four hue points and four mirrored points that should lie on the same hue line. The singular value decomposition of $\tilde{\mathbf{H}}_n$ can then be written as:

$$\tilde{\mathbf{H}}_n = \mathbf{U}_n\mathbf{\Sigma}_n\mathbf{V}_n^T \tag{5.16}$$

where $\mathbf{U}_n$ and $\mathbf{V}_n$ are singular vector matrices and $\mathbf{\Sigma}_n$ is a diagonal matrix of singular values:

$$\mathbf{\Sigma}_n = \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}$$

The second singular value, $\sigma_2$, is the residual error, i.e. the distance of the actual points to the line that best fits the data.

The mean residual error $\epsilon$ of one sensor set was calculated by averaging over the individual residual errors of the 12 hue lines, such that:

$$\epsilon = \frac{1}{n} \sum_{n=1}^{12} \sigma_2(n) \tag{5.17}$$

Minimizing the mean residual error $\epsilon$ derived the best sensor of the original sensor sets $\mathbf{R}$ found through the spherical sampling technique:

$$\mathbf{r}_{opt} = \min_{\mathbf{r}_{opt} \in \mathbf{R}} (\epsilon) \tag{5.18}$$

### 5.4.3 Results

The result of the spherical optimization is illustrated in Figure 5.9, and the corresponding sensors $\mathbf{r}_{opt}$ in Figure 5.10. The sensors have relatively large negative values and are probably not specifically suited to imaging applications. As a comparison, the hue constancy of the ROMM RGB, ITU-BT.R 709, HPE, Sharp, CAT02, and Bradford sensors (see chapter 3) are also plotted in Figures F.1 - F.6 in Appendix F. The residual errors are listed in Table 5.2.

The HPE sensors are closest to the optimal hue constancy sensors $\mathbf{r}_{opt}$, followed by the CAT sensors (Bradford, Sharp, and CAT02). The encoding sensors, ROMM RGB and ITU-BT.R 709, are last. Note, however, the there is no perceptual meaning to the error $\epsilon$. That is, we cannot make any predictions about how "good" or "bad" the hue constancy of the individual color spaces, and by extension the individual color image encodings are. We provide here only a measure of "better." The hue constancy of ROMM and ITU-BT.R 709 based image encodings have been investigated in other studies, specifically using image rendering operations [Mor01, WSG02]. The first study [Mor01] found the performance of ITU-BT.R 709

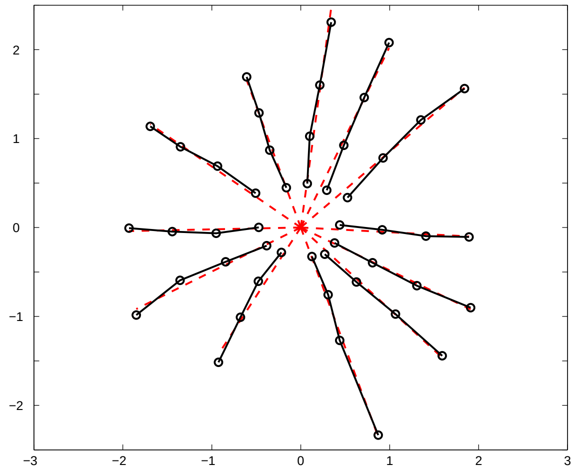


Figure 5.9: Best hue constancy found with spherical sampling. The mean residual error is equal to 0.0884.

based encodings to be slightly better, whereas the other study [WSG02] concluded that ROMM based encodings are preferred.

The experimental procedure presented in this section has many degrees of freedom. The offset term used to avoid negative RGB values has an influence on the final result, as well as the whitening. The error term to be minimized could be refined to take the lengths of the hue vectors or the correlation of the RGB sensors into account. However, the method is very flexible, and can be used for hue definition independent of gamma and white-point encoding. Note that both Lab $H_{a*b*}$ and IPT $H_{PT}$ are contrast dependent, and IPT additionally requires a white-point of D65. Indeed, this work could be regarded as a continuation of the work of Ebner and Fairchild on IPT to find a generalized model for evaluating hue constancy for color engineering purposes.

Additionally, other hue constant test data should also be evaluated. The Hung and Berns data has a limited number of tristimulus values. The "blue" hue vector pointing to the x-axis in Figures 5.12 and F.1 to F.6 is not very straight for almost all the sensors, which adds a significant amount to the residual error. It is not clear from this study if this is due to the sensors or noise in the test data.

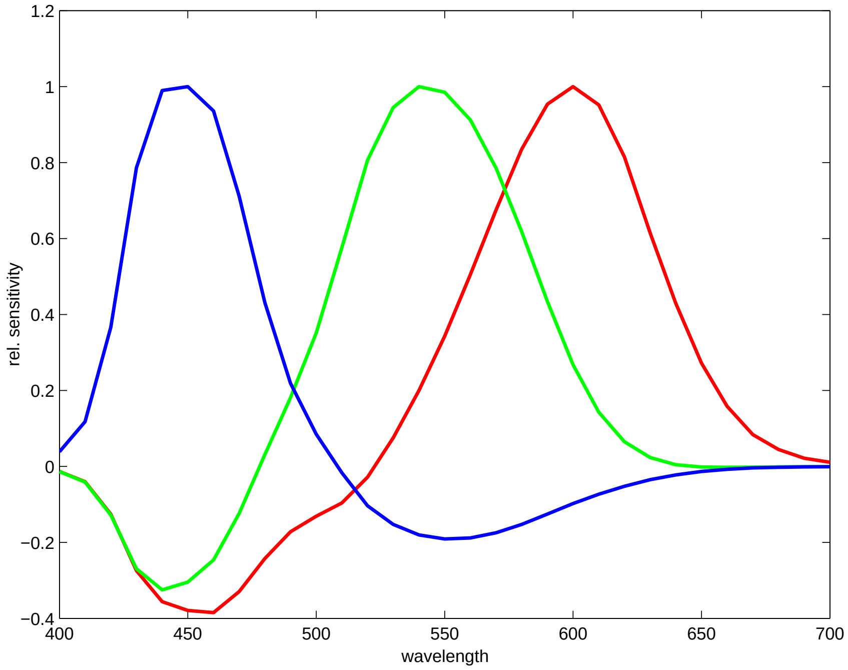


Figure 5.10: Sensors that result in the best hue constancy found with spherical sampling.

## 5.5 White-point independence, Gamut coverage, and Hue constancy

We can combine the two experiments described above and run the sensors with best gamut coverage obtained in section 5.3 to evaluate their hue constancy. Figure 5.11 shows the plots for all 20 sensors, Figure 5.12 for the best sensors in terms of mean residual error ($\epsilon = 0.1601$, see Table 5.2), and Figure 5.13 illustrates the corresponding RGB sensors. Figure 5.14 plots the same sensors and additionally also the Sharp, CAT02, ROMM and ITU-BT.R 709 sensors. The matrix $\mathbf{M}_{\text{best of 20}}$ is one of the 20 matrices found in section 5.3.1 and listed in Appendix E that have best gamut coverage, but which additionally has the smallest residual error $\epsilon$ according to eqs. 5.12 to 5.18. The residual error $\epsilon$ is equal to 0.1601 (see Table 5.2). Thus,

$$\mathbf{M}_{\text{best of 20}} = \begin{bmatrix} 1.6351 & -0.4071 & -0.2280 \\ -0.8044 & 1.7798 & 0.0246 \\ 0.0000 & 0.0152 & 0.9848 \end{bmatrix}$$

Table 5.2: The mean residual errors in hue constancy for well known sensors.

| Transforms | mean residual error $\epsilon$ |
|---|---|
| $\mathbf{M}_{\mathbf{r}_{opt}}$ | 0.0884 |
| $\mathbf{M}_{ROMM}$ | 0.1686 |
| $\mathbf{M}_{ITU-R.BT709}$ | 0.1653 |
| $\mathbf{M}_{HPE}$ | 0.1039 |
| $\mathbf{M}_{Sharp}$ | 0.1504 |
| $\mathbf{M}_{CAT02}$ | 0.1563 |
| $\mathbf{M}_{BFD}$ | 0.1423 |
| $\mathbf{M}_{\text{best of 20}}$ | 0.1601 |

which corresponds to $\mathbf{M}_{20}$ in Table 5.1 with 98.86% of encodable gamut and 49.46% of useful gamut.

The best sensors perform slightly better than the encoding sensors, ROMM RGB and ITU-BT.R 709. As discussed above, our error metric does not allow for a definitive statement about how good or bad the hue constancy is, but can give a relative indication. Thus, we can certainly draw the conclusion that the RGB sensors shown in Figure 5.13 will perform equally well as ROMM and ITU-BT.R 709 in terms of hue constancy. This result has been corroborated by G. Woolfe using the experimental conditions of [WSG02][4]. In terms of surface gamut coverage, the best sensors are able to encode almost all surface colors given by the Steingrimsson et al.'s data set, as illustrated in Figure 5.15.

## 5.6 Conclusions

There are a number of RGB color image encodings with different characteristics proposed and/or standardized by the imaging industry today [IEC99, ANS01, ANS02b, ANS02a], optimized for different purposes within the color imaging workflow. Most define a fixed adopted white-point, i.e. illuminant, usually D50 or D65. However, the

---

[4] private communication

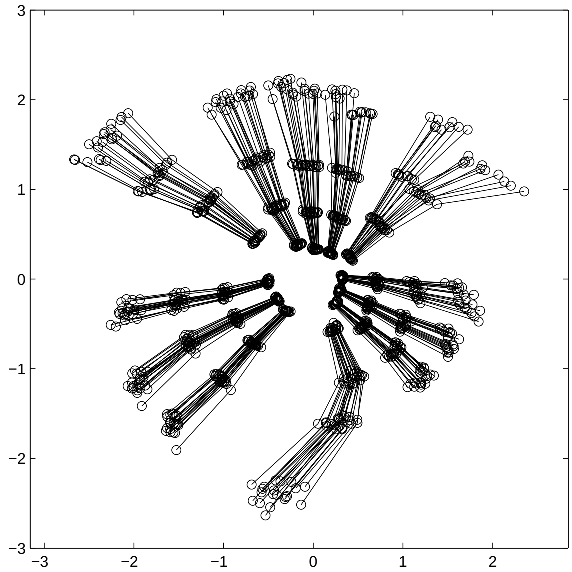


Figure 5.11: Hue constancy of the 20 best sensors optimized for gamut coverage. The mean residual error (over all 20 sensors) is equal to 0.1800.

adapted white-point of scenes can vary greatly due to different lighting conditions, and images are viewed under different illuminants. It is therefore necessary to apply chromatic adaptation transforms to represent images in these different encodings, and to accommodate for different viewing conditions. As a result, several chromatic adaptation transforms might be applied to an image before it is finally displayed or printed. Defining color image encodings that are white-point independent, i.e. that can accommodate different white-points with a simple scaling of linear RGB values, might thus simplify color image processing tasks.

We have presented two methods to evaluate if RGB sensors, found to be equivalent to CAT02 in predicting corresponding colors and thus white-point independent, could form the basis for such color image encodings. We investigated two important properties, namely gamut coverage and hue constancy. Gamut coverage was evaluated based on the gamut of surface colors, which makes the resulting RGB sensors suitable for output-referred color image encoding intended for print reproduction (see Appendix D). The best sensors in terms of encoding efficiency have more optimal useful gamut coverage than the ROMM and ITU-BT.R 709 sensors, which are also not white-point independent [SHF01]. The CAT02 sensors, on the other hand,

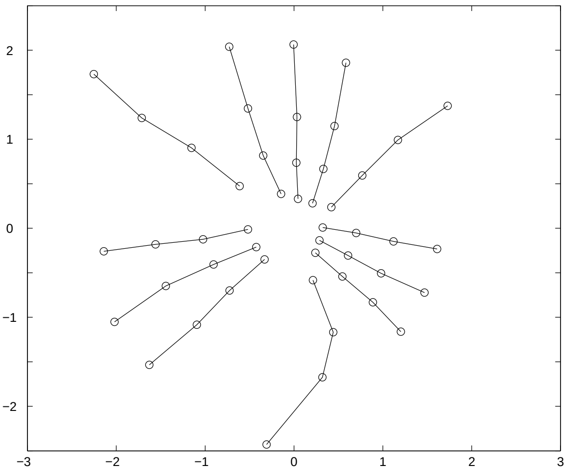


Figure 5.12: Hue constancy of the best RGB sensor set in terms of residual error of Figure 5.11. The mean residual error is equal to 0.1601.

have a very small useful gamut, making them unsuitable for color image encodings based on surface colors. The best sensors are sharp.

We also investigated hue constancy, an other important factor in the design of color image encodings. We have presented an optimization technique that evaluates hue constancy of RGB sensors in a log RGB opponent color encoding. We find that the best RGB sensor combinations in terms of gamut coverage have equivalent hue constancy as the ROMM and ITU-BT.R 709 sensors.

While there is more investigation needed if this log RGB representation can be used to define hue constancy for color imaging applications, we can already make the following observations. A log RGB opponent encoding would allow evaluating hue constancy independent of any contrast (gamma) corrections. Sharp sensors seem to be almost as hue constant as broad-band sensors, such as HPE, in this representation. This emphasizes again the results of our previous investigations, where sharp sensors were found to be optimal in predicting corresponding color data. There are many indicators that sharp sensors should play a role in color engineering.

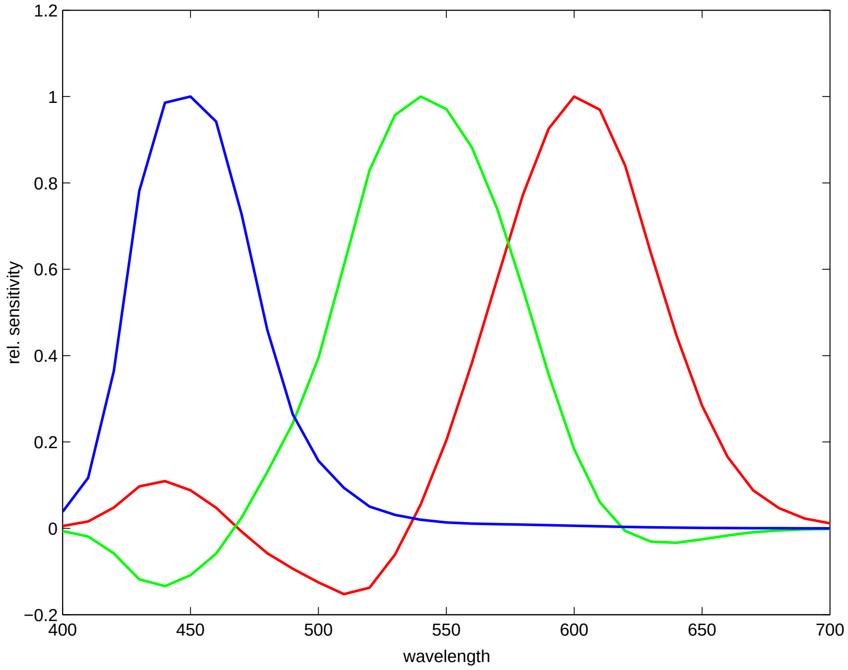


Figure 5.13: The best RGB sensors, culled for white-point independence, then for gamut coverage, and finally for hue constancy.

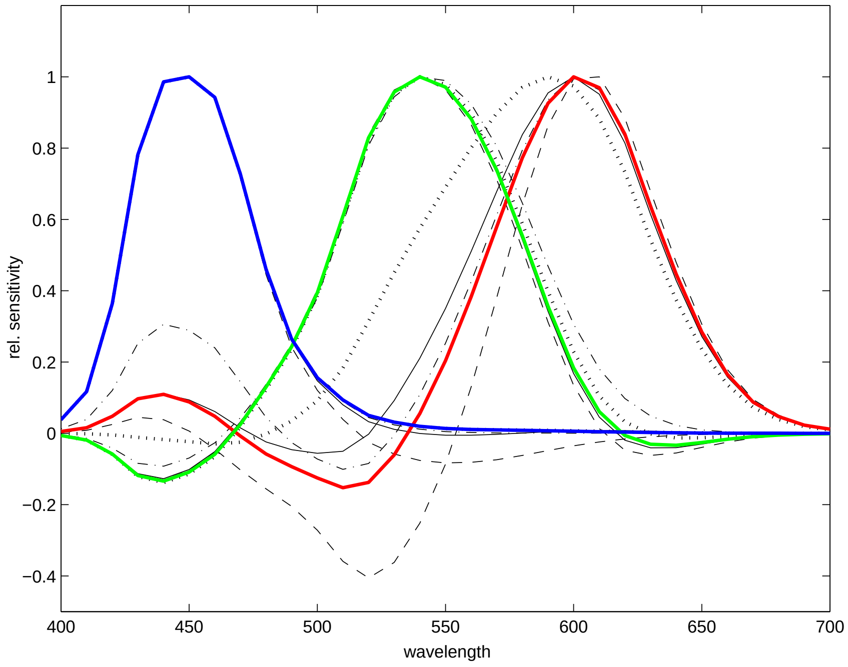


Figure 5.14: The best RGB sensors (solid color) of Fig. 5.13, compared to Sharp (black solid), CAT02 (dot), ROMM (dash dot) and ITU-BT.R 709 (dash) sensors.

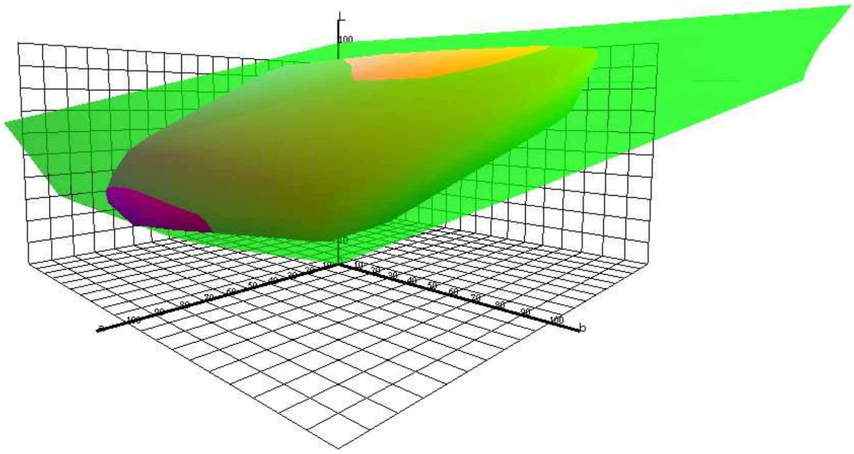


Figure 5.15: The CIE LAB gamuts of the RGB sensor set compared to the surface colors.

# Chapter 6

# Stable Color Ratios

## 6.1 Introduction

In chapter 3, we derived sensors for a strong von Kries-Ives chromatic adaptation transforms based on data-based sharpening of Lam's corresponding color data, which was collected using a memory matching experiment. Data-based sharpening is a technique to minimize least-squares error in the mapping of color responses under two illuminants, subject to a diagonal matrix transform (DMT), by pre-multiplying the data sets by a sharpening matrix **T** (see section 3.4 and eq. 3.6). Applying data-based sharpening to Lam's corresponding color data, whose color coordinates are given in XYZ, we will thus obtain a sharpening matrix **T** that minimizes XYZ mapping errors. If we apply this transformation to the XYZ color matching functions, we obtain sensors that are considerably sharper, i.e. more narrow-band, than the cone fundamentals or the sensors found in other chromatic adaptation transforms (CATs).

In this chapter, we investigate a different approach to find sensors applicable in a strong von Kries-Ives DMT CAT. Inspired by the work on Retinex and relational color constancy (see sections 2.4.1 and 2.6), we derive sensors that keep color response ratios stable. Instead of obtaining sensors by minimizing *absolute* error as in 3.4, we derive our sensors by minimizing *ratio* error [FS02a].

Recall that relational color constancy (see section 2.6) is defined as the invariance of *perceived relations* between the colors of surfaces under different illuminants, as opposed to "normal" color constancy where we assume that the perceived colors are invariant [FNC<sup>+</sup>97]. In other words, the difference is in the invariance of color relation versus color percept. The human visual system encodes and maintains the *ratio* of photoreceptor excitations to different surfaces to achieve color constancy, as opposed to absolute values. These ratios are determined within rather than between color classes.

Hurlbert [FNC+97] noted that these invariant ratios are not necessarily computed at immediate post-receptoral levels; they could be computed at some higher level in the visual system as part of a more general accommodation to the natural structure of the visual environment. We use this argument to derive color ratio stable sensors that are capable of predicting corresponding colors when used in a von Kries type chromatic adaptation transform. As opposed to Foster et al. [FNC+97], we do not assume that the *cone* ratios need to be invariant, but that such an "intermediate" sensor space is used.

The intuition that we apply is as follows: if a CAT is used, then sensor responses are (independently) scaled in some RGB space to account for illuminant change. It follows then that color ratios, computed within a single response channel (R, G or B) must cancel this scaling factor. That is, by looking for sensors that have good ratio stability, we must also be finding reasonable candidates on which to base a chromatic adaptation transform: stable ratios implies a von Kries CAT and vice versa [FS02a].

Using the spherical sampling technique described in section 4.2 and different surface reflectance and illuminant data, we can derive optimal sensors that provide stable color ratios over different illuminants. When using the sensors in a strong von Kries-Ives chromatic adaptation transform of eq. 4.1, we found that there is no statistical difference at the 95 percent confidence level between CAT02, which is used in the CIECAM02 [MFH+02] color appearance model and was derived by optimizing perceptual error (CIELAB $\Delta E$) over sets of corresponding color data, and the CATs based on our sensors. We evaluated the perceptual errors $\Delta E_{94}$ between actual and predicted corresponding colors of the CATs on Lam's experimentally derived corresponding color data set. We find that the resulting sensors are sharp.

This result is remarkable since it implies that optimizing a low level vision criterion, i.e. ratio stability, leads to sharp sensors, which in turn accounts for corresponding color data. Or, put another way, accounting for the physical measurements gives us insight into perception. This result is one of the clearest found to date that suggests that the visual system is based on scene physics.

## 6.2 Experiment

We aim to find the best RGB sensors that result in minimal ratio error between sensor responses of a given set of reflectance data over a range of illuminants. The experiment was done individually for each of the three sensors, under the assumption that ratio stability within one sensor response is independent from the other two.

The "color" or sensor response $\rho$ for any given reflectance under any illuminant for any sensor can be calculated as follows (see section 1.1.4):

$$\rho = \mathbf{s}^T \mathrm{diag}(\mathbf{e})\mathbf{r} \tag{6.1}$$

where $\mathbf{s}$, $\mathbf{e}$ and $\mathbf{r}$ are vectors of the reflectance factors, the illuminant's spectral power distribution, and the sensor sensitivity, respectively. $^T$ is the transpose and "diag" is an operator that turns $\mathbf{e}$ into a diagonal matrix.

Let $\mathbf{b}$ be a $(m \times 1)$ vector containing the colors for a set of $m$ reflectances under the main illuminant with a given sensor. The vector of color ratios $\mathbf{a}$ is calculated as follows:

$$\mathbf{a} = \left[\tfrac{\rho_i}{\rho_j}; \tfrac{\rho_i}{\rho_{j+1}}; \dots\right]; \quad \rho_i, \rho_j \in \mathbf{b}, \quad \rho_i \neq \rho_j \tag{6.2}$$

$\mathbf{a}$ is a component vector of $(m/2) \times (m-1)$ entries. If $\mathbf{a}^e$ is a ratio vector of the same set of reflectances under a different illuminant, then the total ratio error $\epsilon$ is given by:

$$\epsilon = \frac{1}{n} \sum_{e=1}^{n} \frac{|\mathbf{a} - \mathbf{a}^e|}{|\mathbf{a}|} \tag{6.3}$$

where $n$ is the number of illuminants tested other than the main illuminant. By minimizing $\epsilon$, we find the optimal sensor $\mathbf{r}_{opt}$ that keeps color ratios most stable:

$$\mathbf{r}_{opt} = \min_{\mathbf{r}_{opt} \in \mathbf{R}} (\epsilon) \tag{6.4}$$

The initial sensor set $\mathbf{R}$ was determined individually for each color response (R, G, and B) using the spherical sampling technique as described in section 4.2. Recall from eq. 4.5 that sensors can be calculated from the sampling points that lie on the surface of a sphere by:

$$\mathbf{R} = \mathbf{U}\boldsymbol{\mathcal{P}} \tag{6.5}$$

here $\mathbf{U}$ is the orthonormal $m \times 3$ matrix and $\boldsymbol{\mathcal{P}}$ is a $3 \times n$ containing the $n$ surface point coordinates within 20 degrees of either the red, green, or blue HPE, Bradford, CAT02, or Sharp (eq. 3.13) sensor coordinates. Using no constraint, the solution (or optimal sensor) for all three color responses would converge to the blue sensor, as the lowest ratio error $\epsilon$ is in the blue channel (see Table 6.1).

The color responses $\rho$ were calculated for seven different illuminants, the main illuminant D65, and six other illuminants: A, D45, D55, D75, D85, and D100. The illuminants cover the range of lights usually encountered in normal image capture. Two different reflectance data sets were used, the Macbeth Color Checker patches (24 reflectances) [MMD76] and the Munsell chips (462 reflectances) [Mun76]. We chose the Macbeth reflectances because the Macbeth chart is considered to be representative of reflectances encountered in real scenes [MMD76]. The Munsell reflectances were chosen because Lam trained his observers on the Munsell hue, chroma, and value terms when experimentally deriving his corresponding colors ([Lam85], see section 3.2). In section 6.3, we will evaluate how the CATs we define here, which have sensors that keep color ratio most stable, to Lam's corresponding color data.

The best sensors from the initial sensor set $\mathbf{R}$ were derived for each color response and reflectance data set individually using eqs. 6.1 to 6.4. However, these sensors might still not be optimal, and so we additionally applied a simplex search method [LRWW98] implemented in Matlab to find the points on the sphere that result in local minima $\epsilon$ for each sensor:

$$\mathbf{R}_{opt} = \mathbf{U}\mathbf{P}_{opt}; \mathbf{P}_{opt} = \min_{\mathbf{p}_{red},\mathbf{p}_{green},\mathbf{p}_{blue} \in \boldsymbol{\mathcal{S}}}(\epsilon) \tag{6.6}$$

where $\mathbf{R}_{opt}$ are the optimal red, green, and blue sensors that result in the minimum ratio errors ($\epsilon$), and $\boldsymbol{\mathcal{S}}$ is the set of all surface point coordinates on the sphere.

Table 6.1: The ratio errors $\epsilon$ for the HPE, Bradford, CAT02, Sharp and $\mathbf{R}_{opt}$ sensors, using the Macbeth reflectances.

| Ratio Error | HPE | BFD | CAT02 | Sharp | $\mathbf{r}_{opt,\text{Macbeth}}$ |
|---|---|---|---|---|---|
| R channel | 0.0411 | 0.0358 | 0.0305 | 0.0502 | 0.0168 |
| G channel | 0.0398 | 0.0180 | 0.0265 | 0.0236 | 0.0164 |
| B channel | 0.0222 | 0.0117 | 0.0320 | 0.0278 | 0.0116 |

Table 6.2: The ratio errors $\epsilon$ for the HPE, Bradford, CAT02, Sharp and $\mathbf{R}_{opt}$ sensors, using the Munsell reflectances.

| Ratio Error | HPE | BFD | CAT02 | Sharp | $\mathbf{r}_{opt,\text{Munsell}}$ |
|---|---|---|---|---|---|
| R channel | 0.0508 | 0.0394 | 0.0398 | 0.0420 | 0.0178 |
| G channel | 0.0459 | 0.0302 | 0.0442 | 0.0319 | 0.0220 |
| B channel | 0.0252 | 0.0179 | 0.0313 | 0.0261 | 0.0157 |

Tables 6.1 and 6.2 list the ratio errors $\epsilon$ for the different sensors, color responses and reflectance data sets, respectively. The resulting sensors $\mathbf{R}_{opt}$ that keep color ratios over changes in illuminants most constant are illustrated in Figure 6.1.

## 6.3 Comparison with CAT02

Now that we have derived sensors that optimize ratio stability, we wish to evaluate their appropriateness in accounting for corresponding color data. In order to evaluate the optimally stable ratio sensors in the context of chromatic adaptation according to the strong von Kries-Ives model (4.1), we calculate the linear transform mapping XYZs to the ratio stable color responses. With respect to eq. 4.6,

$$\mathbf{M}_{CAT}^{T} = [\boldsymbol{\Sigma}\mathbf{V}^{T}]^{-1}\mathbf{P}_{opt} \quad \mathbf{P} = [\mathbf{p}(\mathbf{r}_{opt,red}), \mathbf{p}(\mathbf{r}_{opt,green}), \mathbf{p}(\mathbf{r}_{opt,blue})] \tag{6.7}$$

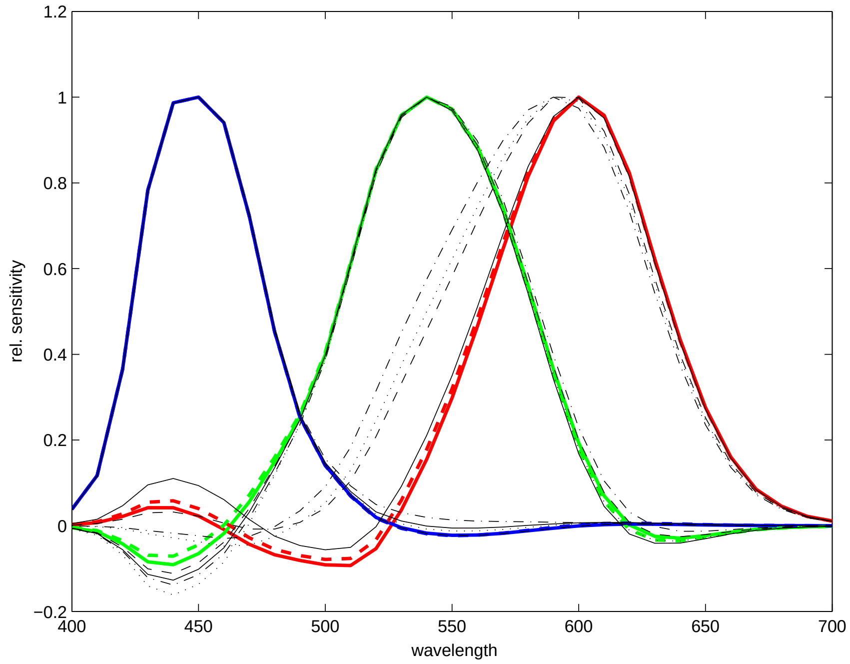


Figure 6.1: The sensors $\mathbf{R}_{opt}$ found by minimizing color ratio errors for the Macbeth (solid color) and the Munsell (dash color) reflectance data sets. For comparison, the Sharp (solid black), CAT02 (dash dot), Bradford (dash) and HPE (dot) sensors are also plotted.

The corresponding $\mathbf{M}_{CAT}$ linear transforms are as follows:

$$\mathbf{M}_{Macbeth} = \begin{bmatrix} 1.4642 & -0.2092 & -0.2550 \\ -0.7545 & 1.6993 & 0.0552 \\ 0.0382 & -0.0667 & 1.0285 \end{bmatrix}$$

$$\mathbf{M}_{Munsell} = \begin{bmatrix} 1.3912 & -0.1620 & -0.2292 \\ -0.7720 & 1.6970 & 0.0750 \\ 0.0448 & -0.0714 & 1.0266 \end{bmatrix}$$

We now applied these new chromatic adaptation transform to Lam's corresponding color data set. The predicted and actual XYZ values were converted to CIELAB (eq. 1.10), so that the perceptual prediction errors $\Delta E_{94}$ could be considered. Two-tail student t-tests for matched pairs [WMM98] were calculated to evaluate if the CATs are statistically different from the CAT02 transform. The results are summarized in Table 6.3.

Table 6.3: Mean CIE $\Delta E_{94}$ values for Lam's data set, and probability $p$-values resulting from the t-test evaluation.

| | Mean CIE $\Delta E_{94}$ | $p$-values |
|---|---|---|
| CAT02 | 2.98 | |
| Macbeth | 3.29 | 0.08 |
| Munsell | 3.20 | 0.16 |

At the 95 % confidence level ($p \geq 0.05$), the ratio optimal sensors derived from the Macbeth and Munsell data set deliver the same chromatic adaptation performance as the CAT02 sensors. However, the ratio optimal sensors are significantly more peaked than CAT02 (see Figure 6.1) The ratio optimal sensors are close to sharp sensors.

## 6.4 Absolute versus Ratio Errors

This result is interesting when viewed in the context of theories of human color vision. Retinex and relational color constancy assume that ratios play a key role in perception. The results here deliver sensors that optimize ratio stability.

Recall that there are, in general, two ways to determine how closely two vectors $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ match. Absolute error is concerned with the *distance* between $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$. Suitable distance metrics are $\sum_{k=1}^{3} |\rho_k - \rho_k'|$ or $\sum_{k=1}^{3} (\rho_k - \rho_k')^2$. The absolute error is often normalized by dividing by the magnitude of $\boldsymbol{\rho}$ or $\boldsymbol{\rho}'$. Relative errors compare the *ratios* of $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ to the unit vector, i.e. the vector components are divided: $\sum_{k=1}^{3} |\frac{\rho_k}{\rho_k'}|$

The sharpening technique discussed in chapter 3 minimizes the absolute error between actual and predicted observations. The sharpening transform is found through eigenvector decomposition of the general transform that best maps a set of observations under one illuminant to observations under another illuminant minimizing least-square errors. The resulting "shape" of the sensors are optimized to give minimal absolute errors for large sensor responses. Where sensor responses are

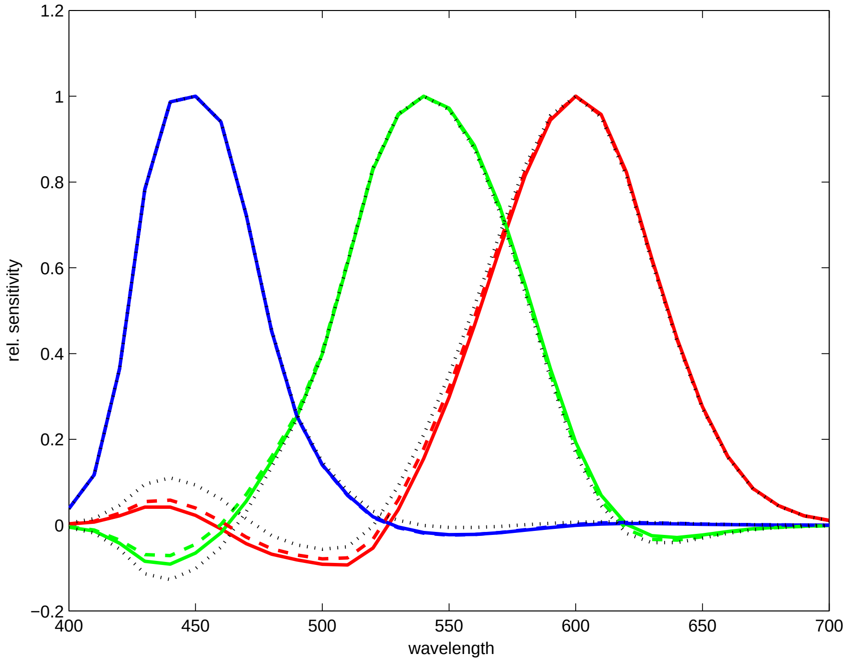


Figure 6.2: A comparison of the the sharp sensors (dot), found by minimizing RMS XYZ error over corresponding color data (eq. 3.13), and the Macbeth (solid) and Munsell (dash) sensors, found by minimizing ratio errors over different illuminants.

small, absolute errors will always be small. Accordingly, the sensors derived through spectral sharpening will be narrow-band, and might include negative values.

However, if we are concerned about maintaining color ratio stability, small sensor responses contribute to large errors. Suppose a sensor response lies in the range of [0,1]. The absolute error between responses 0.025 and 0.05 is small, i.e. 0.025. However, the relative error is very large, i.e. 200%. But the same absolute error at sensor responses of 0.925 and 0.95 will result in a ratio error of only 2.7%. In other words, relative error between fitted variables can be large only if the variables have small values.

Sharpened sensor are thus not optimal in terms of minimizing ratio errors. The negative sensitivities of the sharpened sensors increase the likelihood of small responses, which results in a larger ratio error. This is evident from the ratio errors listed in Tables 6.1 and 6.2. Intuitively, we would therefore not expect sharpened sensors to perform well when the goal is to maintain ratio stability over different illuminants. However, the ratio stable sensors obtained from our experiment, as shown in Figures 6.1 and 6.2, are considerably more sharp and narrow-band than CAT02, but similar to the sharp sensors derived through data-based sharpening of Lam's corresponding color data (see eq. 3.13), as illustrated in Figure 6.2.

## 6.5 Conclusions

We have shown that sensors that keep color ratios stable over different illuminants are able to predict experimentally derived corresponding color data. There is no statistical difference at the 95 percent confidence level between CAT02 and the chromatic adaptation transform based on sensors that have best color ratio stability for the Macbeth and Munsell reflectances under different illuminants. In effect, the different chromatic adaptation transforms will perform equally well.

This is in strong support of the relational color constancy and retinex theories of human visual system, which are based on the assumption that the HVS keeps spatial color relations within a scene invariant Specifically, the *ratio* of color excitations produced by light from different surfaces is retained and kept constant, rather than absolute excitation values. As a result, due to the multiplicative effect of the spectral power distribution of the light source on the color response of a surface reflectance, the illuminant cancels out.

Perhaps more importantly, we have demonstrated that there is a match between minimizing a physical variable, i.e. observations that can be modeled with linear image formation, and psychovisual data, i.e. corresponding color that are experimentally derived. One explanation for this result could be that during the psychovisual experiments, the subjects *do* encode the appearance of the color as the ratio between color and surround. This would support retinex and relational color constancy theories, and explain the good performance of our ratio stable sensors over corresponding color.

The Macbeth and Munsell sensors with stable color ratios are much "sharper," i.e. more narrow-band, than the CAT02 sensors, which were obtained by optimizing perceptual error over sets of corresponding color data. However, the CAT02 sensors also contain negative values, which has an influence on the ratio errors (see Tables

6.1 and 6.2). Thus, the Macbeth and Munsell sensors with narrower sensitivities, but less negative values, can perform just as well in predicting corresponding color. This also supports the use of sharpened, all positive sensors derived through spectral sharpening [DF00] for imaging applications. These sensors are physically realizable and have no negative lobes, which should make them ideal for digital cameras. If they are optimized to keep color ratios stable, white-balancing and transformation to color encoding specific illuminants can be done with simple scaling.

Twenty years ago, Lam [Lam85] derived a chromatic adaptation transform based on a memory matching experiment, where he trained the observers on the Munsell system's hue, chroma, and value terms (see section 3.2). The Bradford chromatic adaptation transform he derived from his experimental data based on minimizing $\Delta E$ perceptual error was the first published CAT that did not assume that scaling takes place in cone space, but rather in a narrower sensor space. For the next fifteen years, the Bradford CAT was extensively applied in color science and color image processing. His experimental corresponding color data has also been used in many other chromatic adaptation transform studies, including ours. It is perhaps fitting to end the research chapters of this thesis with the remark that while we are appreciative to have a stable corresponding color data set to evaluate chromatic adaptation, the *physical responses* of the same Munsell system can be used to derive mathematically a chromatic adaptation transform that performs just as well.

# Chapter 7

# Conclusions

Over a century ago, Johannes von Kries [vK02] proposed a model to account for chromatic adaptation, i.e. the ability of the human visual system to discount the color of the illuminant and to approximately preserve the appearance of an object. He proposed that illumination change can be modeled by an independent gain control of the photoreceptor responses. Since then, many researchers have investigated his model. Based on experimentally obtained corresponding color data or physiological models of the human visual system, these studies have proposed different extensions to the basic model, as discussed in chapter 2. However, most agree that the von Kries diagonal matrix transform is adequate as a first approximation to model chromatic adaptation.

In this thesis, we have taken a mathematical approach to model chromatic adaptation, i.e. to find accurate prediction of corresponding colors. Our first premise is that the strong von Kries-Ives diagonal matrix transform (DMT) is correct, which states that color responses are independently scaled in each channel by a coefficient that is dependent only on the illuminant's response in that channel. However, as opposed to many previous studies, we do not impose that the scaling takes place in cone space. Indeed, which sensor space (or spaces) is most appropriate for mapping corresponding colors as a diagonal transform is the core research question that this thesis addresses.

In chapter 3, we use data-based spectral sharpening, a technique that was successfully used in color constancy algorithms to map physical measurements ("raw"

XYZs) under different illuminants subject to a DMT, to derive sensors that can be used in a chromatic adaptation transform (CAT). We apply the technique to corresponding color data, i.e. experimentally derived measurements of matching color responses under different illuminants. Color responses are first transformed to a intermediate sensor space by a "sharpening" transform. This transform in effect changes the sensor sensitivities. The resulting sensors are "sharp," i.e. more narrow-band than the cone fundamentals or the sensors used in modern chromatic adaptation transforms. However, when applying a CAT based on these sensors to a number of corresponding color data sets, we find that there is no statistically significant difference in $\Delta E$ prediction error between our Sharp CAT and other modern CATs, which were all derived by minimizing perceptual error over these data sets.

The difference in sensor shape between the Sharp CAT and other CATs led us to investigate in chapter 4 if there are not other sensors, not yet considered in the chromatic adaptation literature, which could be used in a strong von Kries-Ives CAT. We designed a spherical sampling algorithm that is able to retain all solutions that fulfill a performance criterion. We show that there are indeed thousands of sensor combinations that are not statistically significantly different from CAT02 in predicting corresponding color. The best sensor combination has sharp (i.e. narrow-band) sensors. We thus conclude that sharp transforms, i.e. von Kries-Ives diagonal matrix transforms applied in a color space that has narrow-band sensors, improves illuminant-invariance with respect to cone fundamentals.

Experimentally deriving color measurements that describe color appearance of a stimulus under different illuminants is not trivial, as discussed in section 2.8 and Appendix B. The data might be noisy, and thus there might not be a unique optimal solution. On the other hand, all these sensors result in CATs that give acceptable visual results [CF01] (see Appendix G). We can conclude that for the color science and color imaging communities to choose sensors for a 'best' CAT, secondary factors should be considered.

We investigate such secondary factors in chapter 5, specifically two properties of color encodings: gamut coverage and hue constancy. The sensors derived in the previous two chapters are sharp, and so are most sensors used in color encodings

[SHF01]. Thus, we investigate if the sensors that can map corresponding colors as well as CAT02 found in the previous chapter could be used for color encodings. These sensors are considered to be white-point independent, because the image RGB values are equal to the post-adaptation sensor responses.

Designing an experiment that compares the encodable and useful gamuts of our sensors with a gamut of surface colors, we find a number of sensors that have better gamut coverage then ROMM and ITU-BT.R 709, two sensor combinations that are used in a number of color image encodings. Hue constancy we evaluated based on a log RGB opponent representation, and we can find sensor combinations that have better hue constancy than the encoding sensors. Combining the two experiments, we can define a set of narrow-band RGB sensitivities that are white-point independent, have good gamut properties and hue constancy, on which we can base output-referred color image encodings for print reproduction.

In chapter 6, we derive sensors from physical color measurements, as opposed to corresponding colors. Based on the insight of retinex and relational color constancy research, we found sensors that keep color ratios stable over several illuminants, and show that these sensors are also able to predict color appearance under different illuminants. We thus show that chromatic adaptation transforms do not have to be derived using experimental data at all. The resulting sensors are again more sharp (i.e. more narrow-band) than cone fundamentals.

The research presented in this thesis has two fundamental results. First, we show that a mathematical approach with closed-form solutions is capable of modeling experimentally derived corresponding color data. Second, we show that sharp sensors that have much narrower sensitivities than cone fundamentals or color matching functions do play a role in color appearance, and should be explored further for application in color science and color image processing.

## 7.1 Future Work

The research presented in all the chapters can be further expanded. We could, for example, derive sensors using spectral sharpening from corresponding colors other than Lam's data set, and see how they compare to each other. We might be able to draw conclusions about the quality of a data set depending on the sensors we derive and comparing their performance across data sets. Also, we might explore if data sets that have more color inconstant corresponding colors make a difference. If yes, is that due to partial adaptation, where a different chromatic adaptation model needs to be explored, or is it due to the experimental set-up? Similarly for chapter 4, we could try different data sets and error criteria to find sensor sets. The sensor combinations found in chapter 5 should be further tested in real imaging applications to judge their suitability in an output-referred image encoding. As for chapter 6, more reflectance data and corresponding color data sets could be evaluated. We can also define all positive sensors that keep color ratios most stable and investigate their use in color image capture applications.

However, the most interesting extension of this research is to study the applicability of sharp sensors in other color appearance context. While we have investigated the use of sharp sensors in modeling chromatic adaptation, sharp sensors might also be appropriate in color similarity and color discrimination tasks. Recall that the log RGB opponent encoding of section 5.4 gives us a definition of hue, which is an appearance correlate. Testing if our log RGB opponent color definition has any perceptual meaning, and under which conditions, i.e. for which sensor responses, would be an interesting research topic.

# Bibliography

[AH01] D. Alleysson and J. Hérault. Variability in color discrimination data explained by a generic model with non linear and adaptive processing. *COLOR Research and Application*, 26:S225–S229, 2001.

[Ale72] R. Alexander. On the sum of distances between $N$ points on a sphere. *Acta Math. Acad. Sci. Hungar.*, 23:443–448, 1972.

[ANS01] ANSI/I3A IT10.7667. Electronic still picture imaging-extended sRGB color encoding e-sRGB, 2001.

[ANS02a] ANSI/I3A IT10.7466. Electronic still picture imaging-reference input medium metric RGB color encoding RIMM-RGB, 2002.

[ANS02b] ANSI/I3A IT10.7666. Electronic still picture imaging-reference output medium metric RGB color encoding ROMM-RGB, 2002.

[AR86] L.E. Arend and A. Reeves. Simultaneous color constancy. *Journal of the Optical Society of America, A*, 3(10):1743–1751, 1986.

[Are93] L.E. Arend. How much does illuminant color affect unattributed colors? *Journal of the Optical Society of America, A*, 10(10):2134–2147, 1993.

[ARSG91] L.E. Arend, A. Reeves, J. Schirillo, and R. Goldstein. Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America, A*, 8:661–672, 1991.

[AWR$^+$94] E.L. Altschuler, T.J. Williams, E.R. Ratner, F. Dowla, and F. Wooten. Method of constrained global optimization. *Phys. Rev. Lett.*, 72:2671–2674, 1994.

[Bar79a] C.J. Bartleson. Changes in color appearance with variations in chromatic adaptation. *COLOR Research and Application*, 4:119–138, 1979.

[Bar79b] C.J. Bartleson. Predicting corresponding colors with changes in adaptation. *COLOR Research and Application*, 4:143–155, 1979.

[Bäu94] K.H. Bäuml. Color appearance: Effects of illuminant changes under different surface collections. *Journal of the Optical Society of America, A*, 11(2):531–542, 1994.

[Bäu95] K.H. Bäuml. Illuminant changes under different surface collections: examining some principles of color appearance. *Journal of the Optical Society of America, A*, 12:261–271, 1995.

[BB87] G. Brelstaff and A. Blake. Computing lightness. *Pattern Recognition Letters*, 5:129–138, 1987.

[BB88] K. Tiplitz Blackwell and G. Buchsbaum. Quantitative studies of color constancy. *Journal of the Optical Society of America, A*, 5(10):1772–1780, 1988.

[BBS97] D.H. Brainard, W.A. Brunt, and J.M. Speigle. Color constancy in the nearly natural image. I. Asymmetric matches. *Journal of the Optical Society of America, A*, 14(9):2091–2110, 1997.

[BEN57] R.W. Burnham, R.M. Evans, and S.M. Newhall. Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America*, 47:35–42, 1957.

[BF96] K.M. Braun and M.D. Fairchild. Psychophysical generation of matching images for cross-media color reproduction. In *Proceedings of IS&T/SID 4th Color Imaging Conference*, pages 214–220, 1996.

[BFHT98] M.H. Brill, G.D. Finlayson, P.M. Hubel, and W.A. Thornton. Prime colours and colour imaging. In *Proceedings of IS&T/SID 6th Color Imaging Conference*, pages 33–42. IS&T/SID, 1998.

[BNS87] D.A. Baylor, B.J. Nunn, and J.L. Schnapf. Spectral sensitivity of cone of the monkey Macaca Fascicularis. *Journal of Physiology*, 39:145–160, 1987.

[Boy96] R.M. Boynton. History and current status of a physiologically based system of photometry and colorimetry. *Journal of the Optical Society of America, A*, 13(8):1609–1621, 1996.

[Bra96] D.H. Brainard. Cone contrast and opponent modulation color spaces. In Kaiser and R.M. Boynton, editors, *Human Color Vision*, pages 563–579. Optical Society of America, Washington, DC, 2 edition, 1996.

[Bra98] D.H. Brainard. Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America, A*, 15(2):307–325, 1998.

[Bre87] E.J. Breneman. Corresponding chromaticities for different states of adaptation to complex visual fields. *Journal of the Optical Society of America, A*, 4:1115–1129, 1987.

[BRK97] D.H. Brainard, M.D. Rutherford, and J.M. Kraft. Color constancy compared: Experiments with real images and color monitors. *Investigative Ophthalmology & Visual Science (Suppl.)*, 38:476, 1997.

[Buc80] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310:1–26, 1980.

[BW81] M.H. Brill and G. West. Contributions to the theory of invariance of color under the condition of varying illumination. *Journal of Mathematical Biology*, 11:337–350, 1981.

[BW86a] D.H. Brainard and B.A. Wandell. Analysis of the retinex theory of color vision. *Journal of the Optical Society of America, A*, 3(10):1651–1661, 1986.

[BW86b] M.H. Brill and G. West. Chromatic adaptation and color constancy: A possible dichotomy. *COLOR Research and Application*, 11:196–204, 1986.

[BW92] D.H. Brainard and B.A. Wandell. Asymmetric color-matching: How color appearance depends on the illuminant. *Journal of the Optical Society of America, A*, 9(9):1433–1448, 1992.

[CF01] A.J. Calabria and M.D. Fairchild. Herding CATS: A comparison of linear chromatic-adaptation transforms for CIECAM97s. In *Proceedings of IS&T/SID 9th Color Imaging Conference*, pages 174–178, 2001.

[CIE78] CIE Supplement No 2 to Publication No 15. *Recommendations on uniform colour spaces, color difference equations, and psychometric terms.* Central Bureau of the CIE, 1978.

[CIE86] CIE Publication No 15.2. *Colorimetry.* Central Bureau of the CIE, 2nd edition, 1986.

[CIE87] CIE Publication No 17.4. *International Lighting Vocabulary.* Central Bureau of the CIE, 1987.

[CIE95] CIE Publication No 116. *Industrial Colour Difference Evaluation.* Central Bureau of the CIE, 1995.

[CIE98] CIE Publication No 131. *The CIE 1997 interim Colour Appearance Model (simple version).* Central Bureau of the CIE, 1998.

[CKL75] C.M. Cicerone, D.H. Krantz, and J. Larimer. Opponent process additivity-III. Effect of moderate chromatic adaptation. *Vision Research*, 15:1125–1135, 1975.

[CS93] J.H. Conway and N.J.A Sloane. *Sphere Packings, Lattices and Groups.* Springer-Verlag, New York, 2 edition, 1993.

[CSP$^+$87] C.A. Curcio, K.R. Sloan, O. Packer, A.E. Hendrickson, and R.E. Kalina. Distribution of cones in human and monkey retina: individual variability and radial asymmetry. *Science*, 236:579–582, 1987.

[CW95] E.J. Chichilnisky and B.A. Wandell. Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching. *Vision Research*, 35(2):239–254, 1995.

[Dan93] J.L. Dannemiller. Rank-orderings of photoreceptor photon catches from natural objects are nearly illuminant-invariant. *Vision Research*, 33:131–140, 1993.

[DB91] J.B. Derrico and G. Buchsbaum. A computational model of spatiochromatic image coding in early vision. *Journal of Visual Communication and Image Representation*, 2:31–38, 1991.

[DB00] P.B. Delahunt and D.H. Brainard. Control of chromatic adaptation: signals from separate cone classes interact. *Vision Research*, 40:2885–2903, 2000.

[DF00] M. Drew and G.D. Finlayson. Spectral sharpening with positivity. *Journal of the Optical Society of America, A*, 17(8):1361–1370, 2000.

[dFCL$^+$01] M.D. de Fez, P. Capilla, M.J. Luque, J. Perez-Carpinell, and J.C. del Pozo. Asymmetric colour matching: Memory matching versus simultaneous matching. *COLOR Research and Application*, 26(6):458–468, 2001.

[DKL84] A.M. Derrington, J. Krauskopf, and P. Lennie. Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357:241–265, 1984.

[DL86] M. D'Zmura and P. Lennie. Mechanisms of color constancy. *Journal of the Optical Society of America, A*, 3(10):1662–1671, 1986.

[DSKK58] R.L. DeValois, C.J. Smith, S.T. Kitai, and A.J. Karoly. Responses of single cells in different layers of the primate lateral geniculate nucleus to monochromatic light. *Science*, 127:238–239, 1958.

[EB72] A.A. Eastman and S.A. Brecher. The subjective measurements of color shifts with and without chromatic adaptation. *Journal of Illuminating Engineering Society*, 2:239–246, 1972.

[EF98a] F. Ebner and M.D. Fairchild. Development and testing of a color space (IPT) with improved hue uniformity. In *Proceedings of IS&T/SID 6th Color Imaging Conference*, pages 8–13. IS&T, 1998.

[EF98b] F. Ebner and M.D. Fairchild. Finding constant hue surfaces in color space. In *IS&T/SPIE Electronic Imaging 1998*, volume 3300, pages 107–117. SPIE, 1998.

[EMG01] R.T. Eskew, J.S. McLellan, and F. Giulianini. Chromatic detection and discrimination. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision: From Genes to Perception*. Cambridge University Press, 2001.

[Est79] O. Estevez. *On the fundamental data-base of normal and dichromatic colour vision*. PhD thesis, University of Amsterdam, 1979.

[Fai96] M.D. Fairchild. Refinement of the RLAB color space. *COLOR Research and Application*, 21:438–346, 1996.

[Fai98] M.D. Fairchild. *Color Appearance Models*. Addison-Wesely, Reading MA, 1998.

[Fai01] M.D. Fairchild. A revision of CIECAM97s for practical applications. *COLOR Research and Application*, 26(6):418–427, 2001.

[FD96] G.D. Finlayson and M.S. Drew. The maximum ignorance assumption with positivity. In *Proceedings of IS&T/SID 4th Color Imaging Conference*, pages 202–205. IS&T/SID, November 1996.

[FD97a] G.D. Finlayson and M.S. Drew. Constrained least-squares regression in color space. *Journal of Electronic Imaging*, 6(4):484–493, 1997.

[FD97b] G.D. Finlayson and M.S. Drew. White-point preserving color correction. In *Proceedings of IS&T/SID 5th Color Imaging Conference*, pages 258–261, 1997.

[FDF94a] G.D. Finlayson, M.S. Drew, and B.V. Funt. Color constancy: generalized diagonal transforms suffice. *Journal of the Optical Society of America, A*, 11(11):3011–3019, 1994.

[FDF94b] G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: Sensor transformations for improved color constancy. *Journal of the Optical Society of America, A*, 11(5):1553–1563, 1994.

[FF96] G.D. Finlayson and B.V. Funt. Coefficient channels: Derivation and relationship to other theoretical studies. *COLOR Research and Application*, 21(2):87–96, 1996.

[FHH01] G.D. Finlayson, S.D. Hordley, and P.M. Hubel. Color by correlation: A simple unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001.

[FHM04] G.D. Finlayson, S. Hordley, and P. Morovic. A multi-spectral image database and an application to image rendering across illumination. In *Proceedings of the Third International Conference on Image and Graphics*, pages 394–397. IEEE, 2004.

[Fin95] G.D. Finlayson. *Coefficient Colour Constancy*. PhD thesis, Simon Fraser University, 1995.

[FL92] M.D. Fairchild and P. Lennie. Chromatic adaptation to natural and incandescent illuminants. *Vision Research*, 32:2077–2085, 1992.

[FM01] G.D. Finlayson and P. Morovic. Is the Sharp adaptation transform more plausible than CMCCAT2000? In *Proceedings of IS&T/SID 9th Color Imaging Conference*, pages 310–315, 2001.

[FN94] D.H. Foster and S.M.C. Nascimento. Relational colour constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society of London, Series B*, 257:115–121, 1994.

[FNC+97] D.H. Foster, S.M.C. Nascimento, B.J. Craven, K.J. Linnell, F.W. Cornelissen, and E. Brenner. Four issues concerning color constancy and relational color constancy. *Vision Research*, 37(10):1341–1345, 1997.

[For90] D.A. Forsyth. A novel algorithm for colour constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.

[Fos81] D.H. Foster. Changes in field spectral sensitivities of red-, green- and blue-sensitive colour mechanisms obtained on small background fields. *Vision Research*, 21:1433–1455, 1981.

[FS00a] G.D. Finlayson and S. Süsstrunk. Performance of a chromatic adaptation transform based on spectral sharpening. In *Proceedings of IS&T/SID 8th Color Imaging Conference*, pages 49–55, 2000.

[FS00b] G.D. Finlayson and S. Süsstrunk. Spectral sharpening and the Bradford transform. In *Proceedings of the Colour Imaging Symposium (CIS) 2000*, pages 236–243. University of Derby, Colour Institute, 2000.

[FS01a] G.D. Finlayson and G. Schaefer. Hue that is invariant to brightness and gamma. In *Proceedings of the British Machine Vision Conference*, pages 303–312, 2001.

[FS01b] G.D. Finlayson and S. Süsstrunk. Spherical sampling and color transformations. In *Proceedings of IS&T/SID 9th Color Imaging Conference*, pages 321–325. IS&T, 2001.

[FS02a] G.D. Finlayson and S. Süsstrunk. Color ratios and chromatic adaptation. In *IS&T First European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pages 7–10. IS&T, 2002.

[FS02b] G.D. Finlayson and S. Süsstrunk. Optimization for hue constant rgb sensors. In *Proceedings of IS&T/SID 10th Color Imaging Conference*, pages 343–348. IS&T, 2002.

[GJT88] R. Gershon, A.D. Jepson, and J.K Tsotsos. From [R,G,B] to surface reflectance: Computing color constant descriptors in images. *Perception*, pages 755–758, 1988.

[Gou01] P. Gouras. Color vision. In H. Kolb, E. Fernandez, and R. Nelson, editors, *Webvision: The organization of the retina and visual system.* 2001. http://webvision.med.utah.edu/index.html.

[GS99] T. Gevers and A.W.M. Smeulders. Color based object recognition. *Pattern Recognition*, 32:453–464, 1999.

[GvL96] G.H. Golub and C.F. van Loan. *Matrix Computations.* John Hopkins University Press, 1996.

[HB95] P.C. Hung and R.S. Berns. Determination of constant hue loci for a crt gamut and their predictions using color appearance spaces. *COLOR Research and Application*, 20(5):285–295, 1995.

[Hel34] H. Helson. Some factors and implications of color constancy. *Journal of the Optical Society of America*, 33:555–567, 1934.

[Hel38] H. Helson. Fundamental problems in color vision. I. The principle governing changes in hue saturation and lightness of non-selective samples in chromatic illumination. *Journal of Experimental Psychology*, 23:439–476, 1938.

[Her78] E. Hering. *Zur Lehre vom Lichtsinne.* Carl Gerolds & Sohn, 1878.

[HF86] D.C. Hood and M.A. Finkelstein. Sensitivity to light. In K.R. Boff, L. Kaufman, and J.P. Thomas, editors, *Handbook of perception and human performance, Volume I: Sensory processes and perception.* John Wiley and Sons, 1986.

[HF98] P.M. Hubel and G.D. Finlayson. Sharp transformations for color appearance. In *Proceedings of IS&T/SPIE Electronic Imaging 1998*, volume 3300, pages 159–164. SPIE, 1998.

[HHF99] P.M. Hubel, J. Holm, and G.D. Finlayson. Illumination estimation and colour correction. In L.W. MacDonald and M.R. Luo, editors, *Colour Imaging: Vision and Technolog.* John Wiley & Sons, 1999.

[HHFD97] P.M. Hubel, J. Holm, G.D. Finlayson, and M.S. Drew. Matrix calculations for digital photography. In *Proceedings of IS&T/SID 5th Color Imaging Conference*, pages 105–111. IS&T/SID, November 1997.

[HJ57] L.M. Hurvich and D. Jameson. An opponent-process theory of color vision. *Psychological Review*, 64(10):384–404, 1957.

[HJW52] H. Helson, D.B. Judd, and M.H. Warren. Object-color changes from daylight to incandescent filament illumination. *Illuminating Engineer*, 47:221–233, 1952.

[Hor74] B.K.P. Horn. Determining lightness from an image. *Computer Graphics and Image Processing*, 3:277–299, 1974.

[Hor99] S.D. Hordley. *The Theory and Practice of Gamut Mapping Colour Constancy.* PhD thesis, University of Derby, 1999.

[Hub95] D.H. Hubel. *Eye, Brain and Vision.* Scientific American Library, 1995.

[Hun50] R.W.G. Hunt. The effects of daylight and tungsten light adaptation to color perception. *Journal of the Optical Society of America*, 40:362, 1950.

[Hun95] R.W.G. Hunt. *The Reproduction of Colour.* Fountain Press, 5th edition, 1995.

[Hun98] R.W.G. Hunt. *Measuring Colour.* Fountain Press, England, 3rd edition, 1998.

[Hur86] A.C. Hurlbert. Formal connections between lightness algorithms. *Journal of the Optical Society of America, A*, 3(10):1684–1693, 1986.

[Hur98] A.C. Hurlbert. Computational models of color constancy. In V. Walsh and J. Kulikowski, editors, *Perceptual Constancy.* Cambridge University Press, 1998.

[ICC04] ICC.1:2004-10. File format specification for color profiles (version 4.2.0), 2004.

[IEC99] IEC 61966 2-1:1999. Multimedia systems and equipment - colour measurment and management - Part 2-1: colour management-default RGB colour space - sRGB, 1999.

[IEC03] IEC/ISO 61966 2-2:2003. Multimedia systems and equipment - colour measurment and management - Part 2-2: colour management - extended RGB colour space - scRGB, 2003.

[ISO98] ISO 10526:1999/CIE S005/E-1998. *CIE Standard Illuminants for Colorimetry.* Central Bureau of the CIE, 1998.

[ISO04] ISO 22028-1:2004. Photography and graphic technology - extended colour encodings for digital image storage, manipulation and interchange - Part 1: architecture and requirements, 2004.

[Ive12] H.E. Ives. The relation between the color of the illuminant and the color of the illuminated object. *Trans. Illum. Eng. Soc.*, 7:62–72, 1912.

[JH55] D. Jameson and L.M Hurvich. Some quantitative aspects of an opponent-colors theory. I. Chromatic responses and spectral saturation. *Journal of the Optical Society of America*, 45(7):546–552, 1955.

[JMW64] D.B. Judd, D.L. MacAdam, and G. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, 54:1031–1042, 1964.

[Jud51] D.B. Judd. Report of U.S. Secretariat committee on colorimetry and artificial daylight. In *Proceedings of the Twelfth Session of the CIE*, volume 1. Central Bureau of the CIE, 1951.

[KFN01] Cone pathways through the retina. In H. Kolb, E. Fernandez, and R. Nelson, editors, *Webvision: The organization of the retina and visual system.* 2001. http://webvision.med.utah.edu/index.html.

[KLB95] W.G. Kuo, M.R. Luo, and H.E. Bez. Various chromatic adaptation transforms tested using new colour appearance data in textiles. *COLOR Research and Application*, 20(5):313–327, 1995.

[KMB02] J.M. Kraft, S.I. Maloney, and D.H. Brainard. Surface-illuminant ambiguity and color constancy: Effects of scene complexity and depth cues. *Perception*, 31:247–263, 2002.

[Lam85] K.M. Lam. *Metamerism and Colour Constancy.* PhD thesis, University of Bradford, 1985.

[Lan64] E.H. Land. The Retinex. *American Scientist*, 52(2):247–264, 1964.

[Lan74] E.H. Land. The Retinex theory of colour vision. In *Proceedings of the Royal Institute of Great Britain*, volume 47, pages 23–57. Applied Science Publishers, 1974.

[Lan77] E.H. Land. The Retinex theory of color vision. *Scientific American*, 237(6):108–128, 1977.

[Lan83] E.H. Land. Recent advances in retinex theory and some implications for cortical computations. *Proc. Natl. Acad. Sci. USA*, 80:5163–5169, 1983.

[Lan86a] E.H. Land. An alternative technique for the computation of the designator in the retinex theory of color vision. *National Academy of Sciences of the United States of America*, 83(10):3078–3080, 1986.

[Lan86b] E.H. Land. Recent advances in retinex theory. *Vision Research*, 26(1):7–21, 1986.

[LCR$^+$91] M.R. Luo, A.A. Clarke, P.A. Rhodes, S.A.R. Scrivener, A. Shappo, and C.J. Tait. Quantifying colour appearance. Part I. LUTCHI colour appearance data. *COLOR Research and Application*, 16:166–180, 1991.

[Lee01] B.B. Lee. Receptor inputs to primate ganglion cells. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision: From Genes to Perception.* Cambridge University Press, 2001.

[LH88] M.S. Livingston and D.H. Hubel. Segregation of form, color, movement and depth: Anatomy, physiology and perception. *Science*, 240:740–749, 1988.

[LH98a] M.R. Luo and R.W.G. Hunt. A chromatic adaptation transform and a colour inconstancy index. *COLOR Research and Application*, 23(3):154–158, 1998.

[LH98b] M.R. Luo and R.W.G. Hunt. The structure of the CIE 1997 colour appearance model (CIECAM97s). *COLOR Research and Application*, 23:138–146, 1998.

[LLR00] C. Li, M.R. Luo, and B. Rigg. Simplification of the CMCCAT97. In *CIC8*, pages 56–60, 2000.

[LLRH02] C. Li, M.R. Luo, B. Rigg, and R.W.G. Hunt. CMC 2000 chromatic adaptation transform: CMCCAT2000. *COLOR Research and Application*, 27(1):49–58, 2002.

[LM71] E.H. Land and J.J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971.

[LR99] M.R. Luo and P.A. Rhodes. Corresponding-colour datasets. *COLOR Research and Application*, 24(4):295–296, 1999.

[LRWW98] J.C. Lagaries, J.A. Reeds, M.H. Wright, and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

[LW93] M.P. Lucassen and J. Walraven. Quantifying color constancy: Evidence for non-linear processing of cone-specific contrast. *Vision Research*, 33(5/6):739–757, 1993.

[LW96] M.P. Lucassen and J. Walraven. Color constancy under natural and artificial illumination. *Vision Research*, 37(17):2699–2711, 1996.

[Mac43] D.L. MacAdam. Specification of small chromaticity differences. *Journal of the Optical Society of America*, 33(1):18–26, 1943.

[Mac44] D.L. MacAdam. On the geometry of color space. *Journal of the Franklin Institute*, 238(3):195–210, 1944.

[Mac56] D.L. MacAdam. Chromatic adaptation. *Journal of the Optical Society of America*, 46(7):500–513, 1956.

[Mac61] D.L. MacAdam. A non-linear hypothesis for chromatic adaptation. *Vision Research*, 1(7):9, 1961.

[Mac63] D.L. MacAdam. Chromatic adaptation. II. Non-linear hypothesis. *Journal of the Optical Society of America*, 53:1441, 1963.

[Mal01] L.T. Maloney. Physics-based approaches to modeling surface color perception. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision: From Genes to Perception*. Cambridge University Press, 2001.

[MFH$^+$02] N. Moroney, M.D. Fairchild, R.W.G. Hunt, C. Li, M.R. Luo, and T. Newman. The CIECAM02 color appearance model. In *Proceedings of IS&T/SID 10th Color Imaging Conference*, pages 23–27, 2002.

[MKNM95] B.M. Mehtre, M.S. Kankanhalli, A.D. Narasimhalu, and G.C. Man. Color matching for image retrieval. *Pattern Recognition Letters*, 16(3):325–331, 1995.

[MMD76] C.S. McCamy, H. Marcus, and J.G. Davidson. A color-rendition chart. *Journal of Applied Photographic Engineering*, 2:95–99, 1976.

[MMT76] J.J. McCann, S.P. McKee, and T.H. Taylor. Quantitative studies in Retinex theory. *Vision Research*, 16:445–458, 1976.

[Mor01] N. Moroney. Hue constancy of RGB spaces. In *Proceedings of IS&T/SID 9th Color Imaging Conference*, pages 163–167, 2001.

[MSKI91] L. Mori, H. Sobagaki, H. Komatsubara, and K. Ikeda. Field trials on the CIE chromatic adaptation formula. In *Proceedings of the CIE 22nd Session*. CIE, 1991.

[Mun76] Munsell Color. *Munsell Book of Color - Matte Finish Collection.* Baltimore, Md., 1976.

[MW86] L.T. Maloney and B.A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America, A*, 3(1):29–33, 1986.

[NBC57] S.M. Newhall, R.W. Burnham, and J.R. Clark. Comparison of successive with simultaneous colour matching. *Journal of the Optical Society of America*, 47:43, 1957.

[NF00] S.M.C. Nascimento and D.H. Foster. Relational color constancy in achromatic and isoluminant images. *Journal of the Optical Society of America, A*, 17(2):225–231, 2000.

[NNJ43] S.M. Newhall, D. Nickerson, and D.B. Judd. Final report of the OSA subcomittee on the spacing of the Munsell colours. *Journal of the Optical Society of America*, 33:385–418, 1943.

[NTS81] Y. Nayatani, K. Takahama, and H. Sobagaki. Formulation of a nonlinear model of chromatic adaptation. *COLOR Research and Application*, 6(3):161–171, 1981.

[Ost35] G. Osterberg. Topography of the layer of rods and cones in the human retina. *Acta Ophthal.(suppl.)*, 6:1–103, 1935.

[Pan] K. K. Pankratov. Spatial and geometric analysis toolbox. http://puddle.mit.edu/ glenn/kirill/saga.html.

[PK94] F. Perez and C. Koch. Hue color segmentation determines object boundaries. *International Journal of Computer Vision*, 12, 1994.

[Poi80] M.R. Pointer. The gamut of real surface colors. *COLOR Research and Application*, 5(3):145–155, 1980.

[PW90] A.B. Poirson and B.A. Wandell. Task-dependent color discrimination. *Journal of the Optical Society of America, A*, 7:776–782, 1990.

[PW93] A.B. Poirson and B.A. Wandell. Appearance of colored patterns: pattern-color separability. *Journal of the Optical Society of America, A*, 10(12):2458–2470, 1993.

[RG99] O. Rinner and K.R. Gegenfurtner. Time course of chromatic adaptation for color appearance and discrimination. Technical report, Max-Planck-Institute für biologische Kybernetik, October 1999. No. 69.

[RJ02] M. Rabbani and R. Joshi. An overview of the JPEG2000 still image compression standard. *Signal Processing: Image Communication*, 17(1):3–48, 2002.

[Rob90] A.R. Robertson. Historical development of CIE recommended color difference equations. *COLOR Research and Application*, 15:167–170, 1990.

[RSZ94] E.A. Rahkmanov, E.B. Saff, and Y.M. Zhou. Minimal discrete energy on the sphere. *Mathematical Research Letters*, 1:647–662, 1994.

[SB59] W.S. Stiles and J.M. Burch. N.P.L. colour-matching investigation: final report. *Optica Acta*, 6:1–26, 1959.

[SB96] J.M. Speigle and D.H. Brainard. Is color constancy task independent? In *Proceedings of the IS&T/SID 4th Color Imaging Conference*, pages 167–172. IS&T, 1996.

[SB99] J.M. Speigle and D.H. Brainard. Predicting color from gray: the relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America, A*, 16(10):2370–2376, 1999.

[SFHF02] S. Süsstrunk, C. Fredembach, J. Holm, and G.D. Finlayson. White-point independent RGB primaries for color image encodings. In *Proceedings of the International Symposium on Imaging Science (ICIS 02)*, pages 421–422, 2002.

[SH88] S.K. Shevell and R.A. Humanski. Color perception under chromatic adaptation: red/green equilibria with adapted short-wavelength-sensitive cones. *Vision Research*, 28:1345–1356, 1988.

[SH98] J. Sangwine and R.E.N. Horne. *The Colour Image Processing Handbook*. Chapman & Hall, 1998.

[SHF01] S. Süsstrunk, J. Holm, and G.D. Finlayson. Chromatic adaptation behavior of different RGB sensors. In *Proceedings IS&T/SPIE Electronic Imaging 2001*, volume 4300, pages 172–183. SPIE, 2001.

[SP75] V.C. Smith and J. Pokorny. Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15:161–171, 1975.

[Spe59] N.I. Speranskaya. Determination of spectrum color co-ordinates for twenty-seven normal observers. *Optics and Spectroscopy*, 7:424, 1959.

[SS95] D.M. Schneeweis and J.L. Schnapf. Photovoltage of rods and cones in the macaque retina. *Science*, 268:1053–1055, 1995.

[SS00] A. Stockman and L. T. Sharpe. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40:1711–1737, 2000.

[SS01a] J.L. Schnapf and D.M. Schneeweis. Electrophysiology of cone photoreceptors in the primate retina. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision: From Genes to Perception*. Cambridge University Press, 2001.

[SS01b] A. Stockman and L.T. Sharpe. Cone spectral sensitivities and color matching. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision: From Genes to Perception*. Cambridge University Press, 2001.

[SS03] A. Stockman and L.T. Sharpe. Color & vision database, 2003. http://cvrl.ioo.ucl.ac.uk/.

[SSS92] B. Smith, C. Spiekermann, and R. Sember. Numerical methods for colorimetric calculations: Sampling density requirements. *COLOR Research and Application*, 17(6):394–401, 1992.

[SSSS02] U. Steingrimsson, K. Simon, W. Steiger, and K. Schläpfer. The gamut obtainable with surface colors. In *Proceedings of IS&T's First European Conference on Colour in Graphics, Imaging, and Vision (CGIV02)*, pages 287–291, 2002.

[SW00] S. Swen and L. Wallis. Chromatic adaptation tag proposal, 2000. ICC Votable Proposal Submission No 8.2.

[SWG01] K.E. Spaulding, G.J. Woolfe, and E.J. Giorgianni. Optimized extended gamut color encoding for scene-referred and output-referred image states. *Journal of Imaging Science and Technology*, 45(5):418–426, 2001.

[Tho99] W.A. Thornton. Spectral sensitivities of the normal human visual system, color-matching functions and their principles, and how and why the two sets should coincide. *COLOR Research and Application*, 24(2):139–156, 1999.

[TSN84] K. Takahama, H. Sobagaki, and Y. Nayatani. Formulation of a non-linear model of chromatic adaptation for a light-grey background. *COLOR Research and Application*, 9:106–115, 1984.

[UUB89] K. Uchikawa, H. Uchikawa, and R.M. Boynton. Partial color constancy of isolated surface colors examined by a color-naming method. *Perception*, 18:83–91, 1989.

[vG91] J.W. von Goethe. *Zur Farbenlehre (1832).* M. Wenzel, Deutscher Klassiker Verlag, 1991.

[vH62] H. von Helmholtz. *Treatise on Physiological Optics (1924).* Dover Publications, New York, 1962. Translated by J. P. Southall.

[vK02] J. von Kries. Theoretische Studien über die Umstimmung des Sehorgans. *Festschrift der Albrecht-Ludwig Universität*, pages 145–158, 1902.

[vK70] J. von Kries (1905). Influence of adaptation on the effects produced by luminous stimuli. In D. L. MacAdam, editor, *Sources of Color Science*, pages 109–119. The MIT Press, Cambridge MA, 1970.

[Vos78] J.J. Vos. Colorimetric and photometric properties of a 2° fundamental observer. *COLOR Research and Application*, 3:125–128, 1978.

[VW71] J.J. Vos and P.L. Walraven. On the derivation of the foveal receptor primaries. *Vision Research*, 11:799–818, 1971.

[Was58] E.G.T. Wassef. Investigations into the theory of predictions of the appearance of colours and its bearing on the theory of colour vision. *Optica Acta*, 5:101, 1958.

[Was59] E.G.T. Wassef. Linearity of the relationship between the tristimulus values of corresponding colours seen under different conditions of chromatic adaptation. *Optica Acta*, 6:378, 1959.

[WB82] G. West and M.H. Brill. Necessary and sufficient conditions for von Kries chromatic adaptation to give color constancy. *Journal of Mathematical Biology*, 15:249–258, 1982.

[WM94] M.A. Webster and J.D. Mollon. The influence of contrast adaptation on color appearance. *Vision Research*, 34:1993–2020, 1994.

[WMM98] R.E. Walpole, R.H. Myers, and S.L. Myers. *Probability and Statistics for Engineers and Scientists*. Prentice Hall International, Upper Saddle River, NJ, 6th edition, 1998.

[Wri34] W.D. Wright. The measurement and analysis of colour adaptation phenomena. *Proceedings of the Royal Society (London)*, 115B:49, 1934.

[Wri41] W.D. Wright. The sensitivity of the eye to small color differences. *Proceedings of the Physical Society (London)*, 53:93, 1941.

[WS82] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas*. New York: Wiley, 2nd edition, 1982.

[WSG02] G. Woolfe, K. Spaulding, and E. Giorgianni. Hue preservation in rendering operations - An evaluation of RGB color encodings. In *Proceedings of IS&T/SID 10th Color Imaging Conference*, pages 317–324, 2002.

[Wue96] S.M. Wuerger. Color appearance changes resulting from iso-luminant chromatic adaptation. *Vision Research*, 36(19):3107–3118, 1996.

[WW82] J.S. Werner and J. Walraven. Effect of chromatic adaptation on the achromatic locus: the role of contrast, luminance and background color. *Vision Research*, 22:929–944, 1982.

[You70] Th. Young. On the theory of light and colors (1802). In D. L. MacAdam, editor, *Sources of Color Science*, pages 51–53. The MIT Press, Cambridge MA, 1970.

[YSP93] T. Yeh, V.C. Smith, and J. Pokorny. Chromatic discrimination with variation in chromaticity and luminance: data and theory. *Vision Research*, 33(13):1835–1845, 1993.

# Appendix A

# Land's Experiments

For the derivation of his vision model, Land performed various experiments, usually with black-and-white or color Mondrians. We summarize his experiments by describing two. In one experiment [Lan64, LM71], he created a black-and-white Mondrian and illuminated it with a projector. The observers' sensations varied from white to gray to black in various parts of the display. He then inserted a gray filter wedge into the projection path, whose transmission varied monotonically in one image direction. The observers saw very little change. All the various white, gray and black rectangles looked nearly as they did before (see Figure A.1). However, measuring the energy reflected from the different areas revealed that the energy coming from a perceived black rectangle at one part of the Mondrian was equal to the energy of a perceived white rectangle in another part. Repeating the experiment with different colored filters and obtaining the same result, he concluded that the human visual system actually encodes for each channel a "lightness" map whose rank-ordering remains invariant to (slow) spatial changes in illuminant radiant energy as well as illuminant "color."

In another experiment [Lan74, Lan77], Land uniformly illuminated a color Mondrian with three projectors emitting narrow-band short, medium, and long wavelength radiation. He measured the color signal of the white patch with a photometer. Considering that he used narrow-band illumination, he could equate one energy measure per spectral band to one cone quantum catch, i.e. one long, one medium, and one short wave energy measure. In a second color Mondrian, he randomly

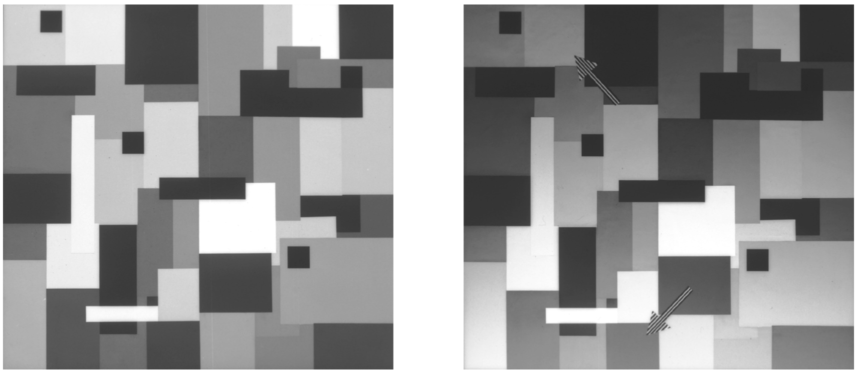


Figure A.1: Black-and-white Mondrian. Left: uniformly illuminated. Right: spatially slow varying illumination (lower right hand to upper left hand corner). The two patches indicated by the arrows have the same luminance. Both photos are scanned from film supplied in [LM71].

selected a "colored" patch ($S(\mathbf{x}, \lambda) \neq 1$), and adjusted the radiant power of the three projectors so that the color signal equaled the color signal of the white patch in the first Mondrian (see Figure A.2). Consequently, the quantum catch of the cones was identical for both stimuli, which should lead to the same color appearance if the trichromatic theory of color vision is correct (see section 1.2.1). However, when observers looked at both Mondrians together, the did not perceive the same color. The white patch in the first Mondrian looked white, while the colored patch in the second Mondrian retained its "color," i.e. blue remained blue, green remained green, etc. The observers were able to "extract" the original reflectances from the color signal.

Based on his experiments, Land therefore concluded that color appearance is not solely dependent on cone quantum catches, and that some cortical processing is responsible for appearance. Thus, he coined the term *retinex* from *retina* and *cortex*. He proposed a number of algorithms intended to model how the HVS computes these illuminant independent lightness values for each cone channel.

One of the early Land retinex algorithm is as follows [BW86a, Hur98]. The lightness value for a sensor $k$ is calculated by finding the average ratio between the color response of a position $\mathbf{x}_0$ and many surrounding positions. As the non-linear lightness response of the HVS is often modeled with a log function, the ratios can

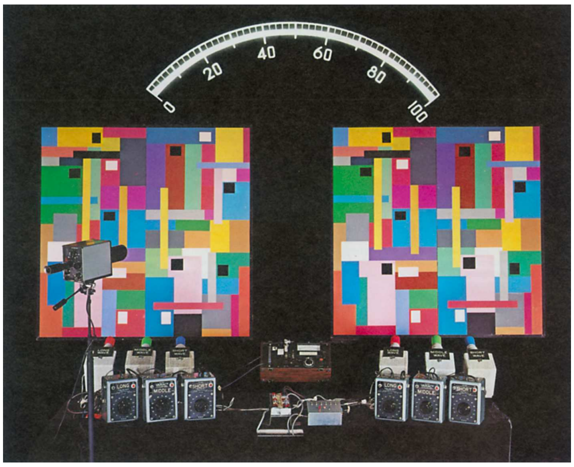


Figure A.2: The set-up used in the color Mondrian experiment. The illustration was taken from [Lan77].

be expressed as subtractions. Let $\mathbf{x}_0$ be the location where the lightness $l_k(\mathbf{x}_0)$ is to be computed. Let $p_k(\mathbf{x}_i), p_k(\mathbf{x}_{i+1})$ be color responses of subsequent locations on an arbitrary path $n$ through the scene, with end point response $p_k(\mathbf{x}_w^n)$:

$$l_k^n(\mathbf{x}_0, \mathbf{x}_w^n) = \sum_{i=0}^{w} \mathcal{T}\left[\log \rho_k(\mathbf{x}_{i+1}) - \log \rho_k(\mathbf{x}_i)\right] \tag{A.1}$$

$l_k^n(\mathbf{x}_0, \mathbf{x}_w^n)$ is the lightness at $\mathbf{x}_0$ relative to endpoint $\mathbf{x}_w^n$. $\mathcal{T}$ represents a thresholding operation that allows to disregard small changes due to small illuminant changes. Recall that he used Mondrians, so the response $\rho_k(\mathbf{x}_i)$ varies sharply with discrete patch boundaries. However, the effective illuminant varies smoothly across the entire scene, and so induces only small changes in $\rho_k(\mathbf{x}_i)$. Thus, $\mathcal{T}$ ensures that such small ratios are ignored:

$$\mathcal{T} = \begin{cases} 0 & \mid \log \rho_k(\mathbf{x}_{i+1}) - \log \rho_k(\mathbf{x}_i) \mid < \text{threshold} \\ 1 & \text{otherwise} \end{cases} \tag{A.2}$$

Therefore, eq. A.1 can be approximated as:

$$l_k^n(\mathbf{x}_0, \mathbf{x}_w) \approx \sum_{i=0}^{w}[\log \rho_k^S(\mathbf{x}_{i+1}) - \log \rho_k^S(\mathbf{x}_i)] = [\log \rho_k^S(\mathbf{x}_0) - \log \rho_k^S(\mathbf{x}_w^n)] \tag{A.3}$$

Spectral normalization is then achieved by averaging the relative lightness $l_k^n(\mathbf{x}_0, \mathbf{x}_w^n)$ over all paths $N$:

$$l_k(\mathbf{x}_0) = \frac{1}{N}\sum_{n=1}^{N} l_k^n(\mathbf{x}_0, \mathbf{x}_w^n) \approx \log \rho_k^S(\mathbf{x}_0) - \frac{1}{N}\sum_{n=1}^{N} \log \rho_k^S(\mathbf{x}_w^n). \tag{A.4}$$

# Appendix B

# Color Constancy of Corresponding Colors

If a diagonal chromatic adaptation model and the difference in adapting illuminants, i.e. $E^a(\lambda)$ and $E^b(\lambda)$ alone could account for the difference in tristimulus values of two corresponding colors, then their CIELAB values would be the same, as CIELAB is derived by normalizing by the tristimulus values of the illuminant (see eq. 1.10). This normalization can be regarded as an adaptation with respect the von Kries chromatic adaptation model of eq. 2.5, each scaling coefficient being the inverse of the tristimulus values of the illuminant, which in fact corresponds to the von Kries-Ives model of eq. 2.8. Thus, the sensor space in which the scaling takes place is given by the CIE XYZ color matching functions.

Figure B.1 illustrates the $a^*, b^*$ coordinates of the corresponding colors of Lam's memory matching experiment under illuminant A and D65. Colors close to the center of the $a^*, b^*$ diagram, i.e. achromatic colors, are generally more "color constant" than colors with high chroma, which corresponds to the findings of West and Brill [WB82] who found almost achromatic reflection spectra when defining the necessary conditions for the von Kries adaptation model to hold using broad band sensors. For high chroma colors, however, there is quite a substantial difference in CIELAB values, as illustrated in Figure B.1. Also, there seems to be a randomness in the hue and chroma direction of the color inconstancy that is not easily interpretable.

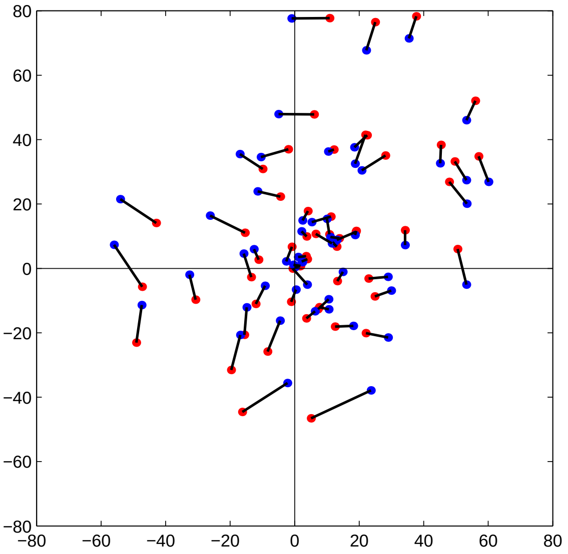


Figure B.1: The $a^*, b^*$ chromaticity coordinates of Lam's [Lam85] corresponding color pairs under illuminant A (red) and D65 (blue). The lines give an indication about their difference in chromaticity.

Several studies have tried to quantify color constancy under different experimental conditions. For example, Eastman and Brecher [EB72] reported that successive haploscopic matching yielded better color constancy than binocular matching. Fez et al. [dFCL$^+$01] compared results from haploscopic matching and memory matching. They found for both experiments that observers had a tendency to select more saturated colors, and the color matches in hue were better along the red-green axis than the yellow-blue axis, which confirms previous findings by other authors [WW82, LW93, Bäu95, Wue96, DB00]. Lam [Lam85] observed systematic discrepancies between the measured sample values under D65 and those obtained from visual inspection under D65. Newhall et al. [NBC57] found similar effects in their comparisons of successive (i.e. memory) matching with simultaneous color matching experiments.

Using a successive haploscopic matching experiment, Lucassen and Walraven [LW96] confirmed that the human visual system achieves better color constancy under broad band than narrow band illuminants, which suggests that chromatic adaptation models based on linear models for illuminant SPD and surface reflectance could be adequate (see section 2.5.1). They tested color constancy under two "natural" daylights and two metameric lights of only two primaries, and found better constancy under natural light. Tiplitz Blackwell and Buchsbaum [BB88] studied the surround of the stimuli and found color constancy differs with the background arrangement. Other studies found the influence of the background negligible [Bäu94, Bäu95, Bra98, KMB02]. Brainard et al. [BRK97] found better constancy for "real live" stimuli than metamerically identical stimuli on a computer screen.

Thus, depending on the task and conditions at hand, the human visual system might not strive for perfect color constancy. There is always a degree of inconstancy, and how much so is not dependent only on the experimental scene statistics, but also which task the observers need to solve [ARSG91]. Lucassen and Walraven [LW96] even concluded that it would therefore make sense to look for a chromatic adaptation model that is intrinsically incapable of giving perfect color constancy. Brainard et al [BBS97] used a similar argument to derive an equivalent illuminant model (see section 2.5.1).

### Color Inconstancy Index

Luo and Hunt [LH98a] proposed a *color inconstancy index* to characterize the constancy of corresponding colors. It is based on a perceptually uniform color space, such as CIELAB, and measures the Euclidean distance between the coordinates of the stimulus under the reference illuminant and the *color-matched* or *predicted* coordinates of the corresponding color under the reference illuminant. Color-matched stimuli are the ones obtained through a psychophysical experiment (see section 2.8). Predicted coordinates are the values obtained using a chromatic adaptation model that maps the test stimulus colors under the test illuminant to the predicted values under the reference illuminant. Large distances indicate the corresponding colors are not very constant, small distances indicate the colors are more constant (see Figures 1.3 and B.1).

# Appendix C

# Prediction Errors of linear CATs

Table C.1: RMS, mean, min, and max $\Delta E$ color difference of actual and predicted color appearance of different linear chromatic adaptation transforms. The $p$-values indicate the confidence levels. $p$-values $\geq 0.05$ ($p$-values $\geq 0.01$) indicate that there is a 95% (99%) confidence that a given transform performs as well as the best transform for a given data set.

| Lam Data Set | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 5.08 | 5.25 | 5.07 | 4.85 | 7.69 |
| mean $\Delta E$ | 4.46 | 4.44 | 4.40 | 4.24 | 6.50 |
| min $\Delta E$ | 0.43 | 0.42 | 0.45 | 0.44 | 0.44 |
| max $\Delta E$ | 11.67 | 11.00 | 11.43 | 10.23 | 16.43 |
| $p$-value | 0.0105 | 0.1291 | 0.0901 |  | 0.0000 |

| Helson | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 6.15 | 6.72 | 6.02 | 5.98 | 8.11 |
| mean $\Delta E$ | 5.33 | 5.55 | 5.22 | 5.12 | 6.89 |
| min $\Delta E$ | 0.91 | 0.66 | 0.78 | 0.95 | 0.70 |
| max $\Delta E$ | 16.37 | 17.32 | 16.65 | 16.41 | 19.79 |
| $p$-value | 0.0143 | 0.0417 | 0.2494 |  | 0.0001 |

| CSAJ | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 5.56 | 5.91 | 5.45 | 5.47 | 7.46 |
| mean $\Delta E$ | 5.12 | 5.36 | 5.02 | 5.03 | 6.63 |
| min $\Delta E$ | 0.45 | 1.22 | 0.99 | 1.25 | 1.13 |
| max $\Delta E$ | 11.50 | 13.43 | 10.94 | 10.95 | 17.57 |
| $p$-value | 0.1412 | 0.0054 |  | 0.4341 | 0.0000 |

| Lutchi | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 7.64 | 7.59 | 6.77 | 7.11 | 8.39 |
| mean $\Delta E$ | 6.77 | 6.90 | 6.07 | 6.38 | 7.05 |
| min $\Delta E$ | 0.89 | 2.36 | 0.99 | 1.66 | 0.72 |
| max $\Delta E$ | 16.92 | 13.53 | 14.47 | 15.09 | 18.48 |
| $p$-value | 0.0310 | 0.0207 |  | 0.1667 | 0.0684 |

| Lutchi D50 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 6.84 | 6.88 | 6.65 | 6.75 | 6.56 |
| mean $\Delta E$ | 6.28 | 6.32 | 6.05 | 6.17 | 5.82 |
| min $\Delta E$ | 1.86 | 1.62 | 1.44 | 1.85 | 0.35 |
| max $\Delta E$ | 15.91 | 15.88 | 16.01 | 15.82 | 14.88 |
| $p$-value | 0.0165 | 0.0017 | 0.1110 | 0.0253 |  |

| LutchiWF | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 8.72 | 9.88 | 8.22 | 8.79 | 11.72 |
| mean $\Delta E$ | 7.80 | 8.87 | 7.27 | 7.91 | 10.55 |
| min $\Delta E$ | 1.32 | 2.22 | 1.12 | 1.56 | 3.17 |
| max $\Delta E$ | 21.37 | 19.88 | 20.17 | 20.16 | 22.91 |
| $p$-value | 0.0005 | 0.0001 |  | 0.0003 | 0.0000 |

| Kuo&Luo | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 7.66 | 7.02 | 8.04 | 7.41 | 10.19 |
| mean $\Delta E$ | 6.93 | 6.37 | 7.02 | 6.68 | 9.10 |
| min $\Delta E$ | 0.84 | 1.57 | 1.53 | 1.54 | 2.01 |
| max $\Delta E$ | 15.22 | 12.54 | 20.28 | 16.14 | 21.86 |
| $p$-value | 0.0776 |  | 0.1069 | 0.1800 | 0.0000 |

| Kuo&Luo TL84 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 4.73 | 5.04 | 4.81 | 4.82 | 6.37 |
| mean $\Delta E$ | 4.30 | 4.61 | 4.27 | 4.36 | 5.76 |
| min $\Delta E$ | 1.09 | 1.24 | 1.16 | 1.45 | 1.88 |
| max $\Delta E$ | 8.80 | 10.21 | 11.74 | 10.49 | 13.97 |
| $p$-value | 0.4014 | 0.0018 |  | 0.0453 | 0.0000 |

| Braun&Fairchild1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 4.02 | 3.89 | 4.06 | 4.02 | 4.12 |
| mean $\Delta E$ | 3.76 | 3.59 | 3.82 | 3.77 | 3.60 |
| min $\Delta E$ | 1.46 | 1.21 | 1.76 | 1.73 | 1.26 |
| max $\Delta E$ | 5.80 | 6.53 | 6.01 | 6.04 | 9.12 |
| $p$-value | 0.0981 |  | 0.0062 | 0.0110 | 0.4825 |

| Braun&Fairchild2 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 6.63 | 6.65 | 6.82 | 6.81 | 6.82 |
| mean $\Delta E$ | 5.90 | 5.96 | 6.14 | 6.12 | 6.30 |
| min $\Delta E$ | 1.92 | 2.51 | 2.53 | 2.38 | 2.58 |
| max $\Delta E$ | 12.73 | 12.35 | 12.61 | 12.56 | 11.51 |
| $p$-value |  | 0.2242 | 0.0043 | 0.0091 | 0.1182 |

| Braun&Fairchild3 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 7.25 | 7.36 | 7.59 | 7.50 | 9.67 |
| mean $\Delta E$ | 7.06 | 7.07 | 7.24 | 7.14 | 9.24 |
| min $\Delta E$ | 3.77 | 3.56 | 3.23 | 3.00 | 5.15 |
| max $\Delta E$ | 10.20 | 11.36 | 12.28 | 12.17 | 13.55 |
| $p$-value |  | 0.4756 | 0.2030 | 0.3744 | 0.0002 |

| Braun&Fairchild4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 6.04 | 5.84 | 6.09 | 5.99 | 7.03 |
| mean $\Delta E$ | 5.92 | 5.73 | 5.92 | 5.83 | 6.72 |
| min $\Delta E$ | 3.71 | 3.82 | 3.26 | 3.27 | 4.39 |
| max $\Delta E$ | 8.11 | 7.51 | 8.42 | 8.36 | 13.00 |
| $p$-value | 0.0818 |  | 0.0905 | 0.2738 | 0.0138 |

| Breneman1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 10.85 | 9.89 | 10.12 | 9.82 | 12.08 |
| mean $\Delta E$ | 10.53 | 9.10 | 9.68 | 9.39 | 10.72 |
| min $\Delta E$ | 6.51 | 2.61 | 4.95 | 5.45 | 3.12 |
| max $\Delta E$ | 15.44 | 16.30 | 15.31 | 15.16 | 20.38 |
| $p$-value | 0.0612 |  | 0.2745 | 0.3530 | 0.0608 |

| Breneman 8 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E$ | 14.01 | 16.11 | 13.31 | 13.65 | 19.13 |
| mean $\Delta E$ | 12.05 | 14.04 | 11.25 | 11.86 | 16.32 |
| min $\Delta E$ | 2.55 | 2.10 | 2.26 | 4.11 | 4.08 |
| max $\Delta E$ | 24.58 | 28.07 | 24.10 | 24.51 | 38.98 |
| $p$-value | 0.0548 | 0.0636 |  | 0.0800 | 0.0080 |

| Breneman 4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|:-:|:-:|:-:|:-:|:-:|
| RMS $\Delta E$ | 14.93 | 17.06 | 14.43 | 14.74 | 20.20 |
| mean $\Delta E$ | 12.27 | 14.67 | 11.60 | 12.27 | 17.35 |
| min $\Delta E$ | 2.21 | 1.49 | 2.47 | 2.61 | 3.99 |
| max $\Delta E$ | 26.53 | 29.89 | 26.10 | 26.44 | 40.41 |
| $p$-value | 0.0600 | 0.0351 | | 0.0819 | 0.0020 |

| Breneman6 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|:-:|:-:|:-:|:-:|:-:|
| RMS $\Delta E$ | 8.27 | 8.18 | 7.05 | 7.53 | 8.31 |
| mean $\Delta E$ | 7.92 | 7.73 | 6.83 | 7.17 | 7.38 |
| min $\Delta E$ | 5.07 | 4.98 | 3.83 | 4.34 | 0.88 |
| max $\Delta E$ | 14.23 | 14.63 | 9.54 | 12.75 | 14.31 |
| $p$-value | 0.0117 | 0.095 | | 0.1575 | 0.3636 |

Table C.2: RMS, mean, min, and max $\Delta E_{94}$ color difference of actual and predicted color appearance of different linear chromatic adaptation transforms. The $p$-values indicate the confidence levels. $p$-values $\geq 0.05$ ($p$-values $\geq 0.01$) indicate that there is a 95% (99%) confidence that a given transform performs as well as the best transform for a given data set.

| Lam Data Set | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 3.40 | 3.54 | 3.45 | 3.40 | 5.01 |
| mean $\Delta E_{94}$ | 2.93 | 3.00 | 2.98 | 2.94 | 4.31 |
| min $\Delta E_{94}$ | 0.43 | 0.41 | 0.44 | 0.44 | 0.41 |
| max $\Delta E_{94}$ | 8.44 | 8.44 | 8.41 | 8.46 | 11.63 |
| $p$-value |  | 0.2281 | 0.3333 | 0.4363 | 0.0000 |

| Helson | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.04 | 4.17 | 4.06 | 3.95 | 5.26 |
| mean $\Delta E_{94}$ | 3.42 | 3.49 | 3.45 | 3.31 | 4.52 |
| min $\Delta E_{94}$ | 0.73 | 0.63 | 0.75 | 0.72 | 0.51 |
| max $\Delta E_{94}$ | 11.87 | 12.55 | 12.15 | 11.92 | 14.45 |
| $p$-value | 0.0173 | 0.0138 | 0.0037 |  | 0.0000 |

| CSAJ | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.08 | 4.23 | 4.02 | 4.03 | 5.22 |
| mean $\Delta E_{94}$ | 3.72 | 3.84 | 3.66 | 3.66 | 4.71 |
| min $\Delta E_{94}$ | 0.32 | 0.87 | 0.76 | 0.91 | 0.96 |
| max $\Delta E_{94}$ | 9.18 | 9.27 | 9.01 | 9.26 | 11.52 |
| $p$-value | 0.0783 | 0.0000 | 0.4626 |  | 0.0000 |

| Lutchi | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.48 | 4.01 | 3.65 | 4.14 | 3.98 |
| mean $\Delta E_{94}$ | 4.03 | 3.71 | 3.35 | 3.81 | 3.47 |
| min $\Delta E_{94}$ | 0.42 | 1.57 | 0.84 | 1.24 | 0.52 |
| max $\Delta E_{94}$ | 9.65 | 7.55 | 7.54 | 8.36 | 8.50 |
| $p$-value | 0.0005 | 0.0286 |  | 0.0001 | 0.3754 |

| Lutchi D50 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 3.97 | 3.91 | 3.86 | 3.90 | 3.64 |
| mean $\Delta E_{94}$ | 3.59 | 3.52 | 3.45 | 3.51 | 3.11 |
| min $\Delta E_{94}$ | 1.39 | 1.07 | 0.95 | 1.27 | 0.22 |
| max $\Delta E_{94}$ | 9.95 | 9.89 | 9.97 | 9.91 | 9.25 |
| $p$-value | 0.0002 | 0.0000 | 0.0011 | 0.0002 |  |

| LutchiWF | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.22 | 4.75 | 4.15 | 4.40 | 6.18 |
| mean $\Delta E_{94}$ | 3.97 | 4.44 | 3.85 | 4.11 | 5.61 |
| min $\Delta E_{94}$ | 0.78 | 1.31 | 0.79 | 1.02 | 1.83 |
| max $\Delta E_{94}$ | 6.90 | 8.79 | 7.25 | 7.62 | 12.48 |
| $p$-value | 0.0963 | 0.0000 |  | 0.0000 | 0.0000 |

| Kuo&Luo | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.23 | 4.13 | 4.32 | 4.27 | 5.77 |
| mean $\Delta E_{94}$ | 3.98 | 3.87 | 3.99 | 4.01 | 5.26 |
| min $\Delta E_{94}$ | 0.68 | 1.13 | 1.21 | 1.10 | 1.56 |
| max $\Delta E_{94}$ | 6.38 | 6.47 | 8.68 | 7.62 | 10.99 |
| $p$-value | 0.2376 |  | 0.2657 | 0.1161 | 0.0000 |

| Kuo&Luo TL84 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 2.93 | 3.01 | 2.91 | 2.97 | 3.66 |
| mean $\Delta E_{94}$ | 2.71 | 2.79 | 2.64 | 2.74 | 3.34 |
| min $\Delta E_{94}$ | 0.90 | 1.14 | 1.04 | 1.17 | 1.48 |
| max $\Delta E_{94}$ | 4.96 | 4.72 | 5.42 | 5.01 | 6.61 |
| $p$-value | 0.2214 | 0.0021 |  | 0.0002 | 0.0000 |

| Braun&Fairchild1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 2.99 | 2.91 | 2.93 | 2.91 | 3.05 |
| mean $\Delta E_{94}$ | 2.79 | 2.68 | 2.71 | 2.68 | 2.72 |
| min $\Delta E_{94}$ | 1.06 | 0.78 | 0.81 | 0.66 | 0.80 |
| max $\Delta E_{94}$ | 5.30 | 5.32 | 5.26 | 5.24 | 6.04 |
| $p$-value | 0.0737 | 0.4906 | 0.0339 |  | 0.4142 |

| Braun&Fairchild2 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 5.17 | 5.18 | 5.29 | 5.28 | 5.24 |
| mean $\Delta E_{94}$ | 4.50 | 4.53 | 4.63 | 4.62 | 4.73 |
| min $\Delta E_{94}$ | 1.72 | 1.39 | 1.70 | 1.61 | 1.74 |
| max $\Delta E_{94}$ | 9.62 | 9.32 | 9.51 | 9.47 | 8.67 |
| $p$-value |  | 0.3294 | 0.0101 | 0.0328 | 0.1443 |

| Braun&Fairchild3 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.45 | 4.75 | 4.84 | 4.74 | 6.18 |
| mean $\Delta E_{94}$ | 4.27 | 4.53 | 4.59 | 4.45 | 5.99 |
| min $\Delta E_{94}$ | 2.26 | 2.13 | 2.12 | 2.10 | 2.65 |
| max $\Delta E_{94}$ | 6.34 | 7.24 | 7.72 | 7.72 | 8.64 |
| $p$-value |  | 0.0086 | 0.0186 | 0.0799 | 0.0001 |

| Braun&Fairchild4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.13 | 4.21 | 4.29 | 4.19 | 5.00 |
| mean $\Delta E_{94}$ | 3.97 | 4.03 | 4.11 | 4.03 | 4.76 |
| min $\Delta E_{94}$ | 2.52 | 2.42 | 2.22 | 2.27 | 2.65 |
| max $\Delta E_{94}$ | 6.58 | 6.37 | 6.47 | 6.07 | 9.51 |
| $p$-value |  | 0.2608 | 0.1452 | 0.2800 | 0.0147 |

| Breneman1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 5.90 | 5.62 | 5.52 | 5.71 | 6.47 |
| mean $\Delta E_{94}$ | 5.57 | 5.02 | 5.13 | 5.36 | 5.46 |
| min $\Delta E_{94}$ | 2.98 | 1.94 | 3.06 | 2.82 | 1.67 |
| max $\Delta E_{94}$ | 9.37 | 10.02 | 9.41 | 9.30 | 12.36 |
| $p$-value | 0.0398 |  | 0.3295 | 0.0798 | 0.2001 |

| Breneman 8 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 7.87 | 8.42 | 7.63 | 7.87 | 10.27 |
| mean $\Delta E_{94}$ | 6.83 | 7.17 | 6.47 | 6.81 | 8.48 |
| min $\Delta E_{94}$ | 1.85 | 1.11 | 1.93 | 1.93 | 2.59 |
| max $\Delta E_{94}$ | 13.10 | 14.53 | 12.85 | 13.20 | 20.87 |
| $p$-value | 0.1254 | 0.0553 |  | 0.0554 | 0.0152 |

| Breneman 4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 8.91 | 9.60 | 8.78 | 8.96 | 11.60 |
| mean $\Delta E_{94}$ | 7.24 | 7.86 | 7.01 | 7.31 | 9.45 |
| min $\Delta E_{94}$ | 1.11 | 1.08 | 1.92 | 1.61 | 2.91 |
| max $\Delta E_{94}$ | 17.45 | 18.94 | 17.19 | 17.50 | 22.90 |
| $p$-value | 0.1880 | 0.0228 |  | 0.0532 | 0.0047 |

| Breneman6 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{94}$ | 4.90 | 4.37 | 4.24 | 4.63 | 4.29 |
| mean $\Delta E_{94}$ | 4.66 | 4.17 | 4.08 | 4.41 | 3.68 |
| min $\Delta E_{94}$ | 2.85 | 2.81 | 2.58 | 2.83 | 0.47 |
| max $\Delta E_{94}$ | 7.93 | 6.59 | 6.21 | 6.97 | 6.53 |
| $p$-value | 0.1628 | 0.2591 | 0.3124 | 0.2078 | |

Table C.3: RMS, mean, min, and max $\Delta E_{CMC(1:1)}$ color difference of actual and predicted color appearance of different linear chromatic adaptation transforms. The $p$-values indicate the confidence levels. $p$-values $\geq 0.05$ ($p$-values $\geq 0.01$) indicate that there is a 95% (99%) confidence that a given transform performs as well as the best transform for a given data set.

| Lam Data Set | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 4.20 | 4.30 | 4.12 | 4.16 | 5.95 |
| mean $\Delta E_{CMC(1:1)}$ | 3.54 | 3.57 | 3.49 | 3.50 | 5.10 |
| min $\Delta E_{CMC(1:1)}$ | 0.55 | 0.48 | 0.61 | 0.59 | 0.53 |
| max $\Delta E_{CMC(1:1)}$ | 10.68 | 10.70 | 10.67 | 10.73 | 12.49 |
| $p$-value | 0.3149 | 0.2663 |  | 0.4202 | 0.0000 |

| Helson | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 4.68 | 4.85 | 4.66 | 4.58 | 6.16 |
| mean $\Delta E_{CMC(1:1)}$ | 3.96 | 4.06 | 3.94 | 3.83 | 5.22 |
| min $\Delta E_{CMC(1:1)}$ | 0.94 | 0.69 | 0.55 | 0.86 | 0.56 |
| max $\Delta E_{CMC(1:1)}$ | 12.16 | 12.71 | 12.30 | 12.21 | 14.64 |
| $p$-value | 0.0195 | 0.0263 | 0.0428 |  | 0.0002 |

| CSAJ | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 4.46 | 4.70 | 4.38 | 4.40 | 6.07 |
| mean $\Delta E_{CMC(1:1)}$ | 4.10 | 4.30 | 4.05 | 4.06 | 5.44 |
| min $\Delta E_{CMC(1:1)}$ | 0.39 | 1.05 | 0.86 | 1.07 | 1.19 |
| max $\Delta E_{CMC(1:1)}$ | 9.40 | 9.54 | 9.25 | 9.48 | 13.09 |
| $p$-value | 0.1935 | 0.0002 |  | 0.3800 | 0.0000 |

| Lutchi | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 5.16 | 4.52 | 4.29 | 4.77 | 4.70 |
| mean $\Delta E_{CMC(1:1)}$ | 4.62 | 4.22 | 3.91 | 4.34 | 4.10 |
| min $\Delta E_{CMC(1:1)}$ | 0.45 | 1.88 | 0.81 | 1.27 | 0.50 |
| max $\Delta E_{CMC(1:1)}$ | 11.02 | 8.71 | 8.17 | 9.68 | 10.15 |
| $p$-value | 0.0005 | 0.0950 |  | 0.0019 | 0.3349 |

| Lutchi D50 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 4.44 | 4.36 | 4.34 | 4.37 | 3.99 |
| mean $\Delta E_{CMC(1:1)}$ | 4.11 | 4.03 | 3.96 | 4.02 | 3.53 |
| min $\Delta E_{CMC(1:1)}$ | 1.50 | 1.15 | 1.02 | 1.38 | 0.25 |
| max $\Delta E_{CMC(1:1)}$ | 9.57 | 9.53 | 9.60 | 9.53 | 8.91 |
| $p$-value | 0.0003 | 0.0001 | 0.0019 | 0.0004 |  |

| LutchiWF | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 5.22 | 5.99 | 5.13 | 5.45 | 7.81 |
| mean $\Delta E_{CMC(1:1)}$ | 4.82 | 5.46 | 4.68 | 4.99 | 6.95 |
| min $\Delta E_{CMC(1:1)}$ | 0.89 | 1.51 | 0.88 | 1.14 | 2.05 |
| max $\Delta E_{CMC(1:1)}$ | 10.83 | 14.00 | 10.49 | 11.63 | 16.46 |
| $p$-value | 0.0701 | 0.0000 |  | 0.0000 | 0.0000 |

| Kuo&Luo | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 4.91 | 4.70 | 4.92 | 4.88 | 6.90 |
| mean $\Delta E_{CMC(1:1)}$ | 4.57 | 4.40 | 4.52 | 4.55 | 6.25 |
| min $\Delta E_{CMC(1:1)}$ | 0.72 | 1.31 | 1.47 | 1.31 | 1.89 |
| max $\Delta E_{CMC(1:1)}$ | 8.54 | 7.05 | 9.57 | 8.42 | 14.56 |
| $p$-value | 0.1951 |  | 0.3027 | 0.1853 | 0.0000 |

| Kuo&Luo TL84 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 3.46 | 3.58 | 3.47 | 3.51 | 4.39 |
| mean $\Delta E_{CMC(1:1)}$ | 3.14 | 3.30 | 3.12 | 3.20 | 4.04 |
| min $\Delta E_{CMC(1:1)}$ | 0.98 | 1.08 | 0.98 | 1.26 | 1.79 |
| max $\Delta E_{CMC(1:1)}$ | 7.15 | 6.71 | 7.06 | 7.14 | 8.92 |
| $p$-value | 0.3812 | 0.0040 |  | 0.0015 | 0.0001 |

| Braun&Fairchild1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 3.67 | 3.52 | 3.60 | 3.58 | 3.61 |
| mean $\Delta E_{CMC(1:1)}$ | 3.39 | 3.19 | 3.30 | 3.27 | 3.14 |
| min $\Delta E_{CMC(1:1)}$ | 1.07 | 0.70 | 0.92 | 0.77 | 1.04 |
| max $\Delta E_{CMC(1:1)}$ | 6.64 | 6.64 | 6.66 | 6.67 | 7.32 |
| $p$-value | 0.2644 | 0.4333 | 0.3190 | 0.3474 |  |

| Braun&Fairchild2 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 6.37 | 6.35 | 6.49 | 6.48 | 6.41 |
| mean $\Delta E_{CMC(1:1)}$ | 5.49 | 5.50 | 5.64 | 5.62 | 5.75 |
| min $\Delta E_{CMC(1:1)}$ | 1.49 | 1.56 | 1.74 | 1.65 | 1.96 |
| max $\Delta E_{CMC(1:1)}$ | 11.80 | 11.85 | 11.93 | 11.94 | 11.78 |
| $p$-value |  | 0.4399 | 0.0060 | 0.0242 | 0.1978 |

| Braun&Fairchild3 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 5.66 | 6.03 | 6.01 | 5.90 | 8.19 |
| mean $\Delta E_{CMC(1:1)}$ | 5.42 | 5.79 | 5.74 | 5.58 | 8.03 |
| min $\Delta E_{CMC(1:1)}$ | 2.84 | 2.69 | 2.72 | 2.50 | 3.15 |
| max $\Delta E_{CMC(1:1)}$ | 8.54 | 8.67 | 8.72 | 8.70 | 10.02 |
| $p$-value |  | 0.0008 | 0.0185 | 0.1338 | 0.0000 |

| Braun&Fairchild4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 5.35 | 5.30 | 5.39 | 5.39 | 5.92 |
| mean $\Delta E_{CMC(1:1)}$ | 5.04 | 5.00 | 5.10 | 5.10 | 5.63 |
| min $\Delta E_{CMC(1:1)}$ | 2.79 | 2.75 | 2.48 | 2.47 | 3.01 |
| max $\Delta E_{CMC(1:1)}$ | 8.55 | 8.44 | 8.29 | 8.37 | 10.27 |
| $p$-value | 0.3356 |  | 0.0674 | 0.1657 | 0.0390 |

| Breneman1 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 7.14 | 6.65 | 6.65 | 6.86 | 7.54 |
| mean $\Delta E_{CMC(1:1)}$ | 6.71 | 5.87 | 6.17 | 6.36 | 6.61 |
| min $\Delta E_{CMC(1:1)}$ | 3.51 | 2.20 | 3.75 | 2.96 | 2.54 |
| max $\Delta E_{CMC(1:1)}$ | 11.75 | 12.00 | 11.62 | 11.76 | 12.74 |
| $p$-value | 0.0343 |  | 0.2379 | 0.0902 | 0.1310 |

| Breneman 8 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 9.28 | 9.96 | 8.92 | 9.27 | 12.05 |
| mean $\Delta E_{CMC(1:1)}$ | 7.93 | 8.54 | 7.46 | 7.89 | 10.23 |
| min $\Delta E_{CMC(1:1)}$ | 1.84 | 1.27 | 1.66 | 2.21 | 2.81 |
| max $\Delta E_{CMC(1:1)}$ | 15.43 | 17.21 | 15.19 | 15.57 | 23.86 |
| $p$-value | 0.0900 | 0.0563 |  | 0.0422 | 0.0153 |

| Breneman 4 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---|---|---|---|---|---|
| RMS $\Delta E_{CMC(1:1)}$ | 10.57 | 11.43 | 10.33 | 10.61 | 13.72 |
| mean $\Delta E_{CMC(1:1)}$ | 8.50 | 9.51 | 8.18 | 8.56 | 11.55 |
| min $\Delta E_{CMC(1:1)}$ | 1.27 | 1.19 | 1.71 | 1.80 | 3.25 |
| max $\Delta E_{CMC(1:1)}$ | 21.12 | 22.95 | 20.84 | 21.19 | 27.36 |
| $p$-value | 0.1347 | 0.0197 |  | 0.0389 | 0.0066 |

| Breneman6 | $\mathbf{M}_{Sharp}$ | $\mathbf{M}_{BFD}$ | $\mathbf{M}_{CAT02}$ | $\mathbf{M}_{Fai01}$ | $\mathbf{M}_{vonKries}$ |
|---:|:---:|:---:|:---:|:---:|:---:|
| RMS $\Delta E_{CMC(1:1)}$ | 6.40 | 5.69 | 5.62 | 6.09 | 5.26 |
| mean $\Delta E_{CMC(1:1)}$ | 5.86 | 5.23 | 5.17 | 5.54 | 4.54 |
| min $\Delta E_{CMC(1:1)}$ | 3.16 | 3.00 | 2.71 | 2.96 | 0.68 |
| max $\Delta E_{CMC(1:1)}$ | 10.08 | 9.37 | 8.92 | 9.78 | 9.10 |
| $p$-value | 0.1484 | 0.235 | 0.2763 | 0.1944 | |

# Appendix D

# Color Spaces and Color Encodings

To unambiguously communicate color information, the color imaging community has defined several *color spaces*, *color encodings* and *color image encodings*. The following review of the definitions is based on ISO 22028 [ISO04] intended for the imaging community, and might differ from conventially used terminology. In chapter 5, we will follow the terminology defined below.

## D.1 Color Spaces

According to the CIE [CIE87], a color space is a "geometric representation of colors in space, usually of three dimensions." They can be broadly categorized into three types: *colorimetric*, *color appearance*, and *device-dependent*.

For colorimetric color spaces, the relationship between the color space and CIE colorimetry is clearly defined. Besides CIEXYZ, CIELAB, and CIELUV, *additive RGB color spaces* also fall into this category. They are defined by a set of *additive RGB primaries*, a *color space white-point* and a *color component transfer function*. The additive RGB sensors are a linear combination of the XYZ color matching functions:

$$\begin{bmatrix} R_{lin} \\ G_{lin} \\ B_{lin} \end{bmatrix} = \mathbf{M} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{D.1}$$

where $\mathbf{M}$ is a $(3 \times 3)$ non-singular matrix mapping XYZ values to linear RGB values. The *RGB primaries* associated with these sensors are the XYZ tristimulus values that correspond to pure red, green, and blue:

$$\begin{bmatrix} X_{red} & X_{green} & X_{blue} \\ Y_{red} & Y_{green} & Y_{blue} \\ Z_{red} & Z_{green} & Z_{blue} \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{D.2}$$

A color space white-point is the color stimulus to which the values are normalized, usually a CIE daylight illuminant such as D50 or D65. For example, a color space with a white point of D65 $(X_w^{D65}, Y_w^{D65}, Z_w^{D65})$ ensures that all achromatic colors, i.e. all scalings of $(X_w^{D65}, Y_w^{D65}, Z_w^{D65})$, are mapped to equal code values. For example for the tristimulus values of the white-point itself:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} X_w^{D65} \\ Y_w^{D65} \\ Z_w^{D65} \end{bmatrix} \tag{D.3}$$

A color component transfer function is a function that accounts for the non-linear response to luminance of the human visual system or to the non-linearity of a device. The function used to model device non-linearities is usually called a *gamma* ($\gamma$) function. In case of CRT monitors, this *gamma* function approximates a power function. Color component transfer functions are thus usually modeled with a logarithmic (see section 1.2.2) or power function.[1] A simple gamma function could take the form of:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \alpha R_{lin}^{\gamma} \\ \alpha G_{lin}^{\gamma} \\ \alpha B_{lin}^{\gamma} \end{bmatrix} \tag{D.4}$$

Luma-chroma color spaces derived from additive RGB spaces are also considered to be colorimetric color spaces. These spaces linearly transform the RGB values to more de-correlated, opponent representations. Generically, they are often referred to as YCC or YCrCb spaces. The resulting luminance and chrominance values are only loosely related to "true" perceptual luminance and chroma and depend on the

---

[1] Note that using these definitions, CIEXYZ calculated under illuminant D65 is a different color space compared to CIEXYZ under illuminant A. Similarly, CIEXYZ under D65 and CIEXYZ under D65 with a logarithmic transfer function are also different.

additive primaries and the transform. Such spaces are generally the bases for color image encodings used in compression [RJ02].

Color appearance color spaces are the output of *color appearance models*, such as CIECAM97s [CIE98] and CIECAM02 [MFH$^+$02]. They are generally based on CIE colorimetry and include parameters and non-linear transforms related to the stimulus surround and viewing environment. The color appearance color space values describe perceptual attributes such as hue, lightness, brightness, colorfulness, chroma, and saturation [Fai98]. Note that all color appearance models contain a chromatic adaptation transform to map a color response from an input illuminant to a illuminant-independent descriptor.

Device-dependent color spaces do not have a direct relationship with CIE colorimetry, but are defined by the characteristics of an input or output device. For input-device dependent color spaces, the spectral characteristics and color component transfer function of an actual or idealized input device is required, as well as a white-point. For output-device dependent color spaces, such as $CMYK$, the relationship between the control signals of a reference output device and the corresponding output image is specified either using output spectra, output colorimetry, or output density.

## D.2 Color Space Encodings

A color encoding is always based on a specific color space, but additionally includes a digital encoding method. Integer digital encodings linearly specify the *digital code value range* associated with the *color space range*. The color space range defines the maximum and minimum digital values that are represented in the digital encoding. Most RGB color space ranges will typically be defined as $[0, 1]$, while CIELAB may range from $[0, 100]$ for $L^*$ and $[-150, 150]$ for $a^*$ and $b^*$, respectively. The digital code value range defines the minimum and maximum integer digital code values corresponding to the minimum and maximum color space values. For example, an 8-bit per channel encoding for an RGB color space with range $[0, 1]$ will associate the digital code value 0 to the color space value 0, and 255 to 1, respectively.

Consequently, by varying the digital code value range and/or color space range, one can derive a family of color space encodings based on a single color space. *sRGB* [IEC99, ANS01, IEC03] and *ROMM/RIMM RGB* [ANS02b, ANS02a] are two such examples.

Using the color space primaries and the color space range, color encoding gamuts can be visually represented in $x, y$ or the more perceptually uniform $u', v'$ chromaticity diagrams. Figure D.1 illustrates the color encoding gamuts of the sRGB and ROMM RGB primaries with a color space range of $[0, 1]$ in $x, y$ and $u', v'$ coordinates. Both are additive RGB color spaces based on a linear transformation of the CIE 1931 CMFs. It can be seen that the encoding gamut of sRGB is smaller than that of ROMM RGB, i.e. more visible colors can be encoded in ROMM RGB than sRGB. The sRGB sensors were optimized to encompass a CRT monitor gamut for an encoding range of $[0, 1]$, while ROMM RGB was intended to cover the gamut of most printing colors. Note that the ROMM RGB gamut goes beyond the spectral locus, i.e. it encompasses chromaticity values that are not visible to the human eye. Thus, digital code values associated with the maximum color range values do not have perceptual meaning, and so are not "used" in the encoding of image data based on visible radiation.

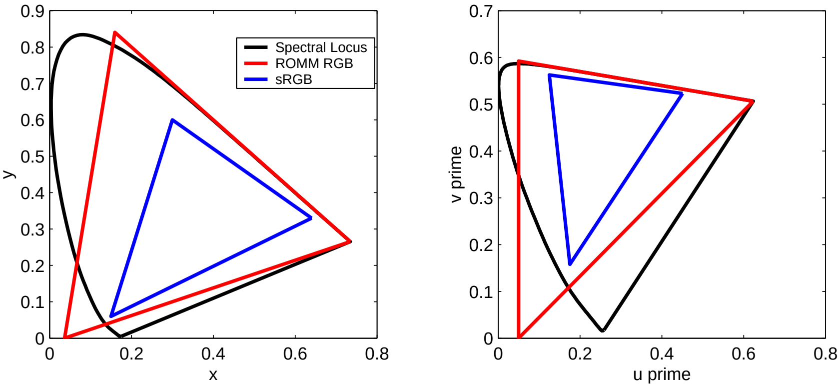


Figure D.1: Left: $x, y$ chromaticities of sRGB and ROMM color encodings. Right: same in $u', v'$ chromaticity coordinates.

## D.3 Color Image Encodings

Color image encodings are based on a specific color space encoding, but additionally define the parameters necessary to properly interpret the color values, such as the *image state* and the *reference viewing environment*. Image state refers to the color rendering of the encoded image. *Scene-referred* color encodings are representations of the estimated color space coordinates of the elements of the original scene. *Output-referred color encodings* are representations of the color space coordinates of image data that is rendered for a specific real or virtual output device and viewing conditions. Reference viewing conditions need to be associated with these image states so that the color appearance can be interpreted. Generally, image surround, adapted white-point, luminance of adapting field, and viewing flare is specified. In case of output-referred color encodings, a reference imaging medium, either a real or idealized monitor or print, also needs to be characterized by its medium white-point, medium black-point and target gamut.

Note that in theory, a color image encoding could be based on any color space encoding. In practice, color space encodings are usually optimized for a given image state by defining an application specific digital code value range and color space range. For example, the sRGB color image encoding for output-referred image representation is optimized for typical CRT monitor gamut and has a limited dynamic range. It is thus unsuitable for most scene-referred image data.

# Appendix E

# Linear transforms for color encodings

The 20 best transformation matrices found to have suitable color gamut encoding properties (see section 5.3)

$$\mathbf{M}_1 = \begin{bmatrix} 1.6351 & -0.4071 & -0.2280 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_2 = \begin{bmatrix} 1.6857 & -0.5079 & -0.1778 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_3 = \begin{bmatrix} 1.6351 & -0.4071 & -0.2280 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_4 = \begin{bmatrix} 1.7518 & -0.5185 & -0.2334 \\ -0.9422 & 1.9043 & 0.0379 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_5 = \begin{bmatrix} 1.8299 & -0.6281 & -0.2018 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_6 = \begin{bmatrix} 1.5741 & -0.4000 & -0.1741 \\ -0.9422 & 1.9043 & 0.0379 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_7 = \begin{bmatrix} 1.8299 & -0.6281 & -0.2018 \\ -0.9422 & 1.9043 & 0.0379 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_8 = \begin{bmatrix} 1.9241 & -0.7471 & -0.1770 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_9 = \begin{bmatrix} 1.7518 & -0.5185 & -0.2334 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{10} = \begin{bmatrix} 1.8299 & -0.6281 & -0.2018 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{11} = \begin{bmatrix} 1.5638 & -0.3067 & -0.2571 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_{12} = \begin{bmatrix} 1.7518 & -0.5185 & -0.2334 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{13} = \begin{bmatrix} 1.9241 & -0.7471 & -0.1770 \\ -0.9422 & 1.9043 & 0.0379 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{14} = \begin{bmatrix} 1.9241 & -0.7471 & -0.1770 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{15} = \begin{bmatrix} 1.5024 & -0.3006 & -0.2018 \\ -0.9422 & 1.9043 & 0.0379 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_{16} = \begin{bmatrix} 1.5741 & -0.4000 & -0.1741 \\ -0.8044 & 1.7798 & 0.0247 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_{17} = \begin{bmatrix} 1.5741 & -0.4000 & -0.1741 \\ -0.8584 & 1.8823 & -0.0238 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_{18} = \begin{bmatrix} 1.8299 & -0.6281 & -0.2018 \\ -0.9254 & 2.0028 & -0.0775 \\ 0.0000 & 0.0151 & 0.9848 \end{bmatrix}$$

$$\mathbf{M}_{19} = \begin{bmatrix} 1.5741 & -0.4000 & -0.1741 \\ -0.7166 & 1.7464 & -0.0298 \\ 0.0873 & -0.1096 & 1.0223 \end{bmatrix}$$

$$\mathbf{M}_{20} = \begin{bmatrix} 1.6351 & -0.4071 & -0.2280 \\ -0.8044 & 1.7798 & 0.0246 \\ 0.0000 & 0.0152 & 0.9848 \end{bmatrix}$$

# Appendix F

# Hue Constancy Plots

These figures illustrate the hue constancy of different white-point preserving sensors (see section 5.4)

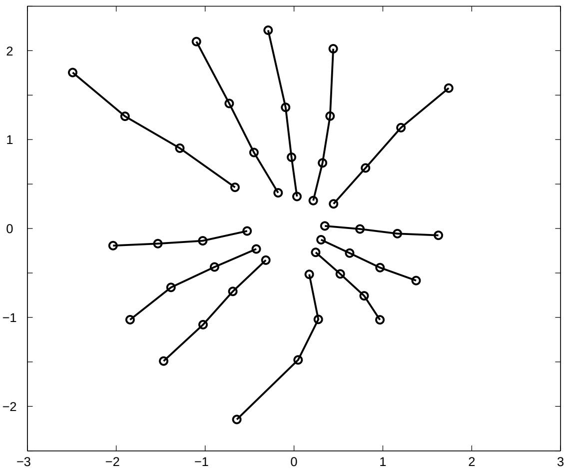


Figure F.1: Hue constancy for the ROMM sensors. The mean residual error is equal to 0.1686.

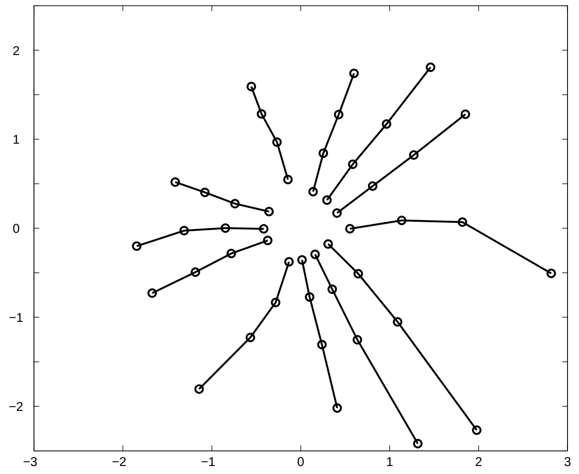


Figure F.2: Hue constancy for the 709 sensors. The mean residual error is equal to 0.1653.

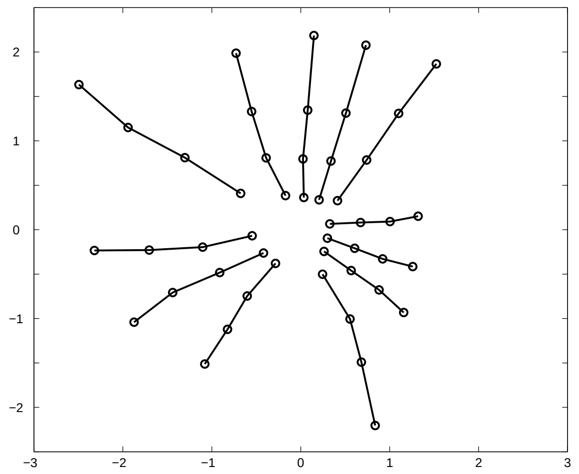


Figure F.3: Hue constancy for the HPE sensors. The mean residual error is equal to 0.1039.

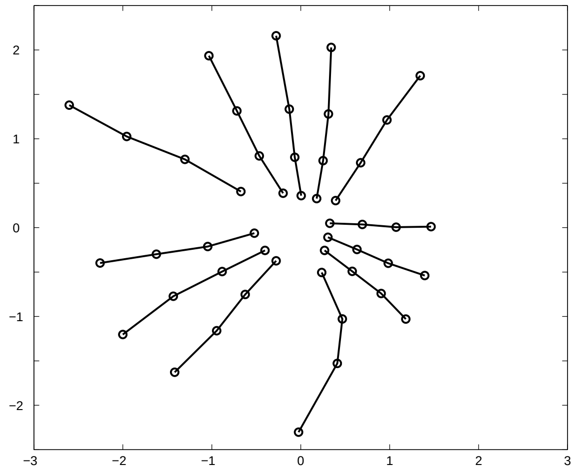


Figure F.4: Hue constancy for the Sharp CAT sensors. The mean residual error is equal to 0.1504.

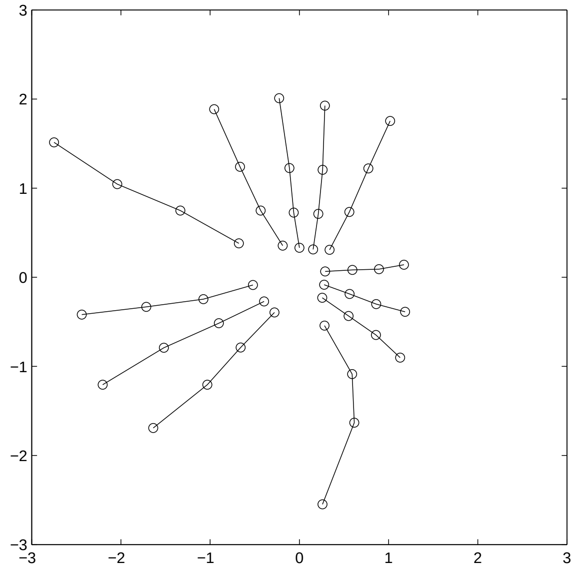


Figure F.5: Hue constancy for the CAT02 sensors. The mean residual error is equal to 0.1563.

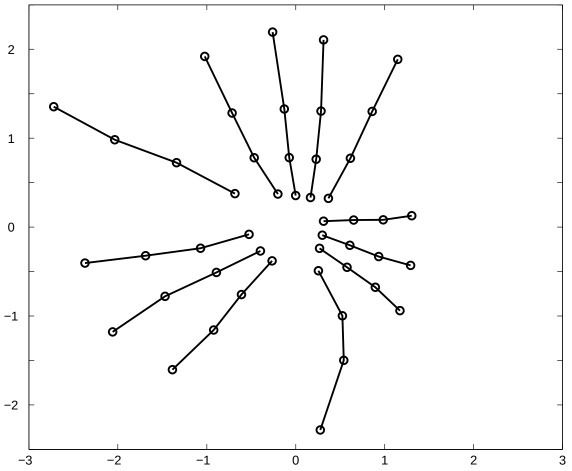


Figure F.6: Hue constancy for the Bradford sensors. The mean residual error is equal to 0.1423.

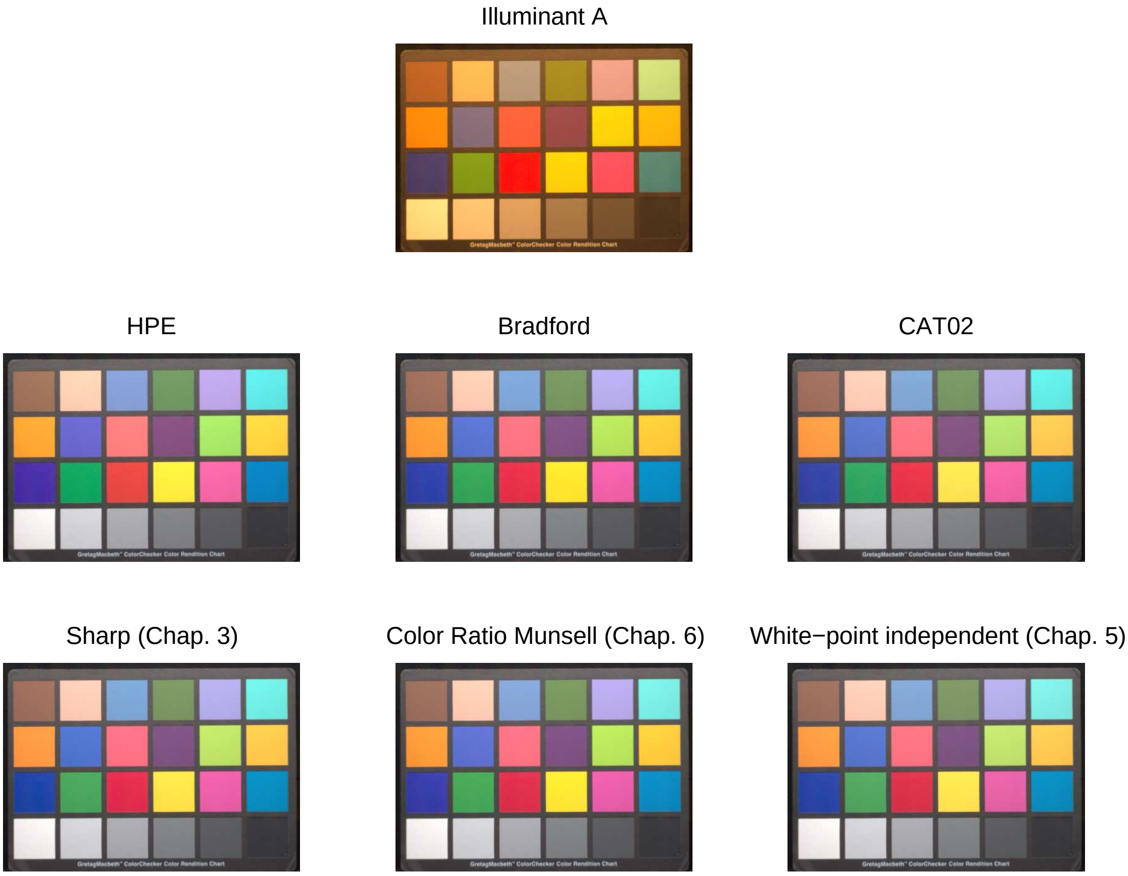
# Appendix G

# Visual Examples

The images on the following pages are from the multispectral image data set compiled by Finlayson et al. [FHM04]. The data is in reflectance factors per spectral interval (400-700nm, $\Delta\lambda = 10nm$). Using the image formation model of eq. 1.6, the multispectral data was rendered to XYZs under illuminant A and D100, respectively. Using a von Kries-Ives DMT chromatic adaptation transforms with the different sensors discussed in this thesis, the images were then rendered to illuminant D65 and converted to sRGB color image encoding for display.

Note that it is almost impossible to tell with these two images which of the transforms performs best. In the Kelloggs image G.3, the HPE sensors show a noticeable difference (the yellow turns slightly greenish, and the blue slightly purple), but the visual difference between the other images is neglectable.

These examples provide some visual proof that there are many sensors that can be used in a DMT-type CAT to render color under different illuminants, as discussed in in this thesis.



Illuminant A

HPE

Bradford

CAT02

Sharp (Chap. 3)

Color Ratio Munsell (Chap. 6)

White−point independent (Chap. 5)

Figure G.1: The Macbeth chart “captured” under illuminant A and rendered to D65 using the different tranforms discussed in this thesis.

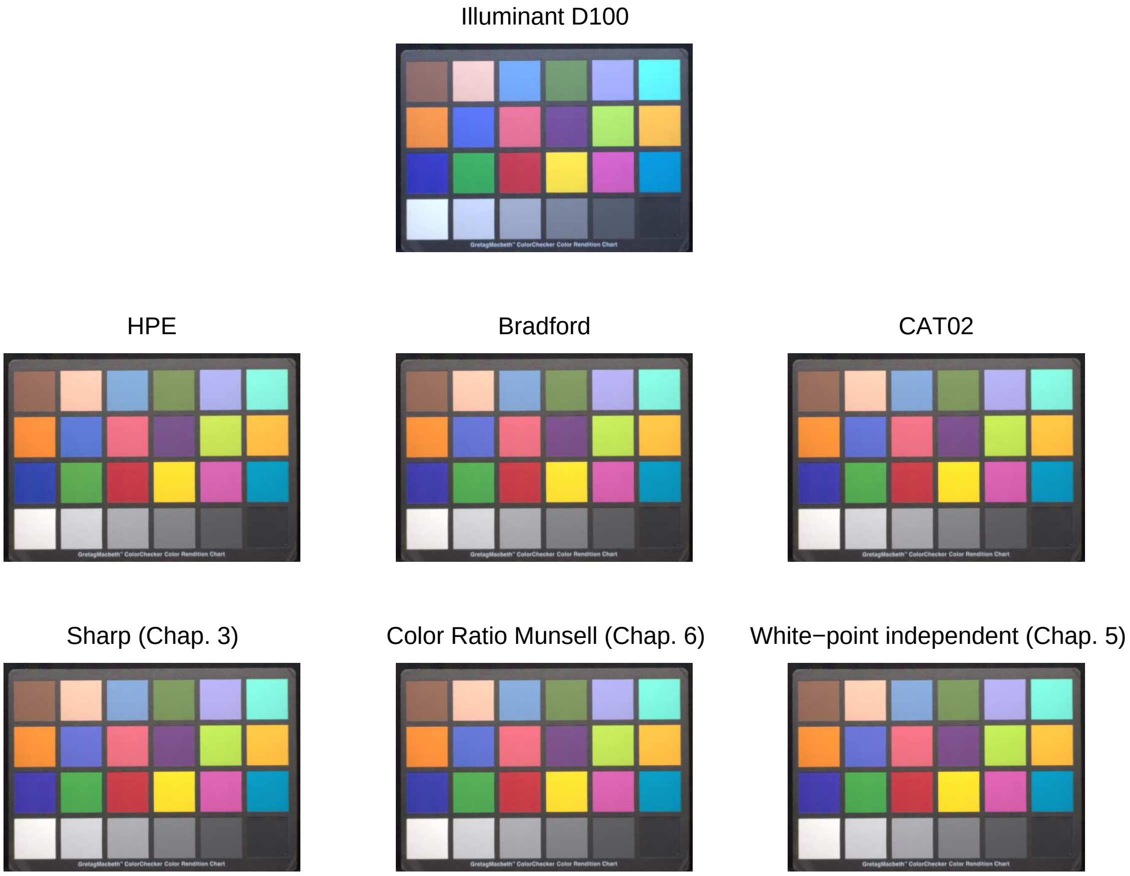


Figure G.2: The Macbeth chart “captured” under illuminant D100 and rendered to D65 using the different tranforms discussed in this thesis.

Illuminant A

HPE

Bradford

CAT02

Sharp (Chap. 3)

Color Ratio Munsell (Chap. 6)

White−point independent (Chap. 5)

Figure G.3: An image “captured” under illuminant A and rendered to D65 using the different tranforms discussed in this thesis.

Illuminant D100

HPE

Bradford

CAT02

Sharp (Chap. 3)

Color Ratio Munsell (Chap. 6)

White−point independent (Chap. 5)

Figure G.4: An image "captured" under illuminant D100 and rendered to D65 using the different tranforms discussed in this thesis.