Color in Image Search Engines

Theo Gevers
Faculty of Science,
University of Amsterdam,
Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

Intelligent Sensory Information Systems
Faculty of Science
University of Amsterdam
The Netherlands

Published in
Principles of Visual Information Retrieval,
editor Michael S. Lew,
Springer-Verlag, London,
ISBN 1-852333-381-2, February 2001

Contents

Intelligent Sensory Information Systems
Faculty of Science
University of Amsterdam
Kruislaan 403
1098 SJ Amsterdam
The Netherlands
tel: +31 20 525 7516
fax: +31 20 525 7490
e-mail: gevers@wins.uva.nl

1 Introduction

Can you imagine an existence without color? Indeed, color does not only add beauty to objects but does give more information about objects as well. Further, color information often facilitates your life, like in traffic, or in sport to identify your favorite team when both teams wear dark shirts.

After more than three hundred years since Newton established the fundamentals of color in his "Opticks" (1704) [18], color has been involved in many fields ranging from pure scientific, to abstract art, and applied areas. For example, the scientific work on light and color resulted in the quantum mechanics started and elaborated by Max Planck, Albert Einstein and Niels Bohr. In painting, Albert Munsell provided the theoretical basis in his "A Color Notation" (1905) [16] on which most painters derived their notions about color ordering. The emotional and psychological influence of color on humans has been studied by Goethe in its famous book "Farbenlehre" (1840) [11]. Further, the value of the biological and therapeutic effect of light and color has been analyzed, and views on color from folklore, philosophy and language have been articulated by Descartes, Schopenhauer, Hegel and Wittgenstein.

Today, with the growth and popularity of the World Wide Web, a new application field is born through the tremendous amount of visual information, such as images and video's, which has been made accessible publicly. With this new application area, color has returned to the center of interest of a growing number of scientists, artists and companies. Aside from decorating and advertising potentials for Web-design, color information has already been used as a powerful tool in content-based image and video retrieval. Various color based image search schemes have been proposed based on various representation schemes such as color histograms, color moments, color edge orientation, color texture, and color correlograms, [1], [2], [5], [20], [27], [28]. These image representation schemes have been created on the basis of RGB, and other color systems such as HSI and $L^*a^*b^*$ [1], [2], [5], [20], [27]. In particular, the Picasso [3] and ImageRover [21] system use the $L^*u^*v^*$ color space for image indexing and retrieval. The QBIC system [5] evaluates similarity of global color properties using histograms based on a linear combination of the RGB color space. MARS [15] is based on the $L^*a^*b^*$ color space which is (like $L^*u^*v^*$) a perceptual uniform color space. The PicToSeek system [10] is based on color models robust to a change in viewing direction, object geometry and illumination. Hence, the choice of color systems is of great importance for the purpose of proper image retrieval. It induces the equivalence classes to the actual retrieval algorithm. However, no color system can be considered as universal, because color can be interpreted and modeled in different ways. Each color system has its own set of color models, which are the parameters of the color system. Color systems have been developed for different purposes: 1. display and printing processes: RGB, CMY; 2. television and video transmittion efficiency: YIQ, YUV; 3. color standardization: XYZ; 4. color uncorrelation: $I_1I_2I_3$; 5. color normalization and representation: rgb, xyz; 6. perceptual uniformity: $U^*V^*W^*$, $L^*a^*b^*$, $L^*u^*v^*$; 7. and intuitive description: HSI, HSV. With this large variety of color systems, the inevitable question arises which color system to use for which kind of image retrieval application. To this end, criteria are required to classify the various color systems for the purpose of content-based image retrieval. Firstly, an important criterion is that the color system is independent of the underlying imaging device. This is required when images in the image database are recorded by different imaging devices such as scanners, camera's and camrecorder (e.g. images on Internet). Another prerequisite is that the color system should exhibit perceptual uniformity meaning that numerical distances within the color space can be related to human perceptual differences. This is important when images are to be retrieved which should be visually similar (e.g. stamps, trademarks and paintings databases). Also, the transformation needed to compute the color system should be linear. A non-linear transformation may introduce instabilities with respect to noise causing poor retrieval accuracy.

Further, the color system should be composed of color models which are understandable and intuitive to the user. Moreover, to achieve robust and discriminative image retrieval, color invariance is an important criterion. In general, images and video's are taken from objects from different viewpoints. Two recordings made of the same object from different viewpoints will yield different shadowing, shading and highlighting cues changing the intensity data fields considerably. Moreover, large differences in the illumination color will drastically change the photometric content of images even when they are taken from the same object. Hence, a proper retrieval scheme should be robust to imaging conditions discounting the disturbing influences of a change in viewpoint, object pose, and illumination.

In this chapter, the aim is to provide a survey on the basics of color, color models and ordering systems, and the state-of-the-art on color invariance. For the purpose of color-based image retrieval, our aim is to provide a taxonomy on color systems composed according to the following criteria:

- Is the color system device independent
- Is the color system perceptual uniform
- Is the color system linear
- Is the color system intuitive
- Is the color system robust against varying imaging conditions
  - Invariant to a change in viewing direction;
  - Invariant to a change in object geometry;
  - Invariant to a change in the direction of the illumination;
  - Invariant to a change in the intensity of the illumination;
  - Invariant to a change in the spectral power distribution (SPD) of the illumination.

The color system taxonomy can be used to select the proper color system for a specific application. For example, consider an image database of textile printing samples (e.g. curtains). The application is to search for samples with similar color appearances. When the samples have been recorded under the same imaging conditions (i.e. camera, illumination and sample pose), a perceptual uniform color systems (e.g. $L^*a^*b^*$) is most suitable. When the lightning conditions are different between the recordings, a color invariant system is most appropriate eliminating the disturbing influences such as shading, shadows and highlights.

This chapter is organized as follows. Firstly, the fundamentals of color will be given in Section 2. As color appearance depends on the light, object and observer, we will study and analyze this triplet in detail in Section 3. Further, in Section 4, the standardization of the light-object-observer triplet will be presented. A survey on color invariance is presented in Section 5. The taxonomy of color systems is given in Section 6. Color and image search engines are discussed in Section 7. Conclusions will be drawn in Section 8. Further, the various color systems and their performance can be experienced within the PicToSeek and Pic2Seek systems on-line at: http://www.wins.uva.nl/research/isis/zomax/.

2 Color Fundamentals

Fundamentally, color is part of the electromagnetic spectrum with energy in the range from 380- to 780-nm wavelength. This is the part of the spectrum to which the human eyes are sensitive. For color measurements, it is often restricted to 400-700 nm. This visible part of the spectrum is perceived as the colors from violet through indigo, blue, green, yellow, and orange to red. This continuous spectrum of colors is obtained when a beam of sunlight is split through a glass prism. Wavelength is the physical difference between the various regions of the spectrum. The unit length of the wavelength is the nanometer (nm). Each wavelength value within the visible band corresponds to a distinct color. Most of the colors that we see do not correspond to one single wavelength but are a mixture of wavelengths from the electromagnetic spectrum, where the amount of energy at each wavelength is represented by a spectral energy distribution. For example, the energy emitted by a white-light source contains a nearly equal amount of all wavelengths, see Figure 1.a. When white light shines upon an object, some wavelengths are reflected and some are absorbed. For example, a green object reflects light with wavelengths primarily around the 500 nm range. The other wavelengths are absorbed, see Figure 1.b.

Color can be described in different ways than only physically by its wavelength characteristics. A system that describes color, is called a color system. Color can be defined and modeled in different ways and each color system has its own set of color models (usually three). The three color features which are usually taken to describe the visual (intuitive) sensation of color are hue, saturation and lightness. The **hue** corresponds with the dominant wavelength of the spectral energy distribution and is the color we see when viewing the light. For instance, if one sees a color with predominant high wavelengths then the color is interpreted as red. Thus, the hue is the kind of color, like red, green, blue, yellow, cyan, etc. **Saturation** corresponds to the excitation purity of the color and is defined as the proportion of pure light with respect to white light needed to produce the color. High saturation denotes little dilution. In other words, saturation is the richness of a hue and therefore denotes how pure a hue is. **Lightness** is related to the intensity (or energy) of the light reflected from objects. Further, **brightness** corresponds to the intensity of the light coming from light sources (i.e. self- luminous objects). The higher the emitted intensity, the brighter the color appears. The description of spectral information in terms of these three color models is derived as follows. Consider the energy emmited by a white light source as shown in Figure 1.a. All wavelengths contribute more or less equally to the total energy. When the color of the light source has a dominant wavelength around the 506 nm part of the spectrum, as shown in Figure 1.b, then the perceived color will be perceived as greenish. Let the power energy of the dominant wavelength be denoted by EH and the wavelengths contributing to produce white light of intensity by EW, see Figure 1.b. Hence, the hue is green corresponding to the dominant wavelength EH of 506 nm where the energy power is given by EH. Further, saturation depends on the difference between E_H and E_W: the larger the difference the more pure the color is. The lightness is equal to the area under the curve of the spectral power distribution (i.e. total energy).

In conclusion, the relative spectral power (light) gives complete information about the color (i.e. color fingerprints). However, the human eye is incapable of analyzing color into its spectral components. An intuitive approximation of the spectral information of color by humans is in terms of hue, saturation and lightness. Although, the approximation of the spectral curves by these three visual attributes suffice for a large number of problems, there are still a number of problems for which spectral information is essential and where human interpretation fails such as standards for color measurement, color matches under different illumination, and color mixture analysis.

The spectral power distribution of the light from objects (for example 1.b) is the result of different complex factors such as the light source and material characteristics. In the next section, these complex factors will be discussed.

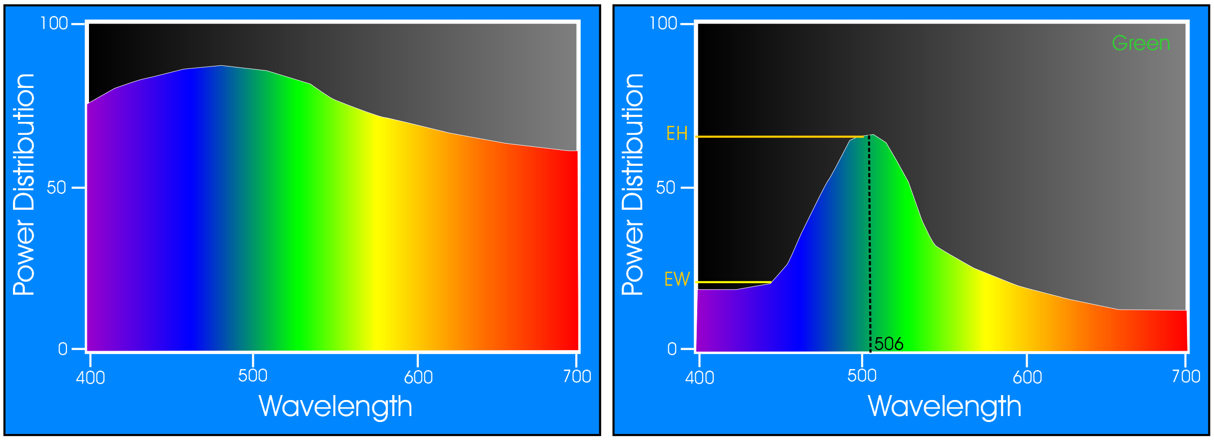

Figure 1: *a. Relative spectral power distribution of a white light source. b. Spectral power distribution of the light reflected from a green sample.*

3 Color Appearance

Let's follow the light emitted by a light source. Through gas discharge (e.g. fluorescent lamps) or by heating up a material (the filament of the lamp), the light source generates light illuminating an object. Some part of the spectrum of the light is reflected from the object and is subsequently measured by an observer such as our light-sensitive eyes or by a color camera. The measured light is then sent to our brain (or computer) where the color of the light is observed (interpreted). The light-object-observer triplet depends on very complex factors: the observed color may vary with a change in the intensity and energy distribution of the light source, material characteristics, mode of viewing, the observer and so on.

Let's take a closer look at the triplet light-object-observer. As stated above, the traveling of light starts with the **light** source. It is known that the color of the light source has a large influence on the observed color. Examples are the spotlights installed in theaters or disco's. A substantial change in spotlight color will subsequently change the color of the dresses of the dancers. In contrast, people have a large degree of color constancy: the ability of observing the same color under different lightning conditions. For example, an object illuminated by daylight is perceived the same as when the object is illuminated by a candle light.

The second major component is the **object** or **sample** itself. The object color determines the observed color by reflecting (opaque or reflecting materials), absorbing (transparent materials), or absorbing and transmitting (translucent materials) parts of the incident light. Also the viewing mode influences the observed color. A change in viewing position of the camera will change the color and intensity field of the object color with respect to shading and highlighting cues. Further, material characteristics of the object, in terms of structure, gloss and metallic effects, influence the observed color.

Thirdly, the light-sensitive detectors of the **observer**, such as the human eyes or color (ccd) camera's, determine which color will be observed. It often occurs that different human observers will judge the color of an object differently even when the object is seen under the same viewing conditions. Similarly, two recording made from the same object with two different camera's (i.e. with different color filters) but under the same imaging conditions, may yield different pictures.

Hence, the observed color of an object depends on complex set of imaging conditions. To get more insight in the imaging process, it is essential to understand the basics of color measurement. Therefore, in this section, we will subsequently focus on the three major components: the light source in Section 3.1, the object characteristics and viewing mode in Section 3.2, and the observer (eye and camera) in Section 3.3.

3.1 The Light Source

The main light source is the sun. Further, artificial light sources exist generating light by gas discharge (neon, argon or xenon) such as fluorescent lamps, or by heating up material (e.g. the filament of a lamp). Light produced by different light sources may vary with respect to their spectral power distribution (SPD), which is the amount of radiant power at each wavelength of the visible spectrum. Instruments exist for measuring the radiation power at each wavelength. To make the SPD independent of the amount of light, the spectral power distribution is computed with respect to the radiation of light at a single wavelength (usually 560 nm which is given the value 100). The relative spectral power distribution of a light source is denoted by $E(\lambda)$.

Color temperature is commonly used to express the relative SPD of light sources. Color temperatures correspond to the temperature of a heated black body radiator. The color of the blackbody radiator changes with temperature. For example, the radiator changes from black at 0 K (Kelvin), to red at about 1000 K, white at 4500 K to bluish white at about 6500 K. An incandescent lamp would give a color temperature of 2900 K and a fluorescent lamp about 4100 K. Historically, color temperature has been introduced to match and describe (old fashioned) light sources such as candles and incandescent lamps, generating light in a similar way as the black body due their thermal radiation. However, modern light sources such as fluorescent lamps often do not properly match a color temperature. These light sources have so-called correlated color temperatures instead of exact color temperatures.

The most important light source is the **sun**. The color temperature of the sun may vary during the time of the day (e.g. reddish at sunrise and bluish at noon). Further, the weather plays also an important role in the determination of the color of the sun. For example, on a clear day the color temperature is about 5500 K. The color of the sun changes with the day and month of the year due to latitude and atmospheric conditions. Hence, the color of the sun may vary between color temperatures of 2000 K to 10000 K (indoor north sky daylight). For standardization of colorimetric measurements, in 1931, the international lighting commission (CIE), recommended that the average daylight has the color temperature of 6500 K and has been specified as the standard illuminant D65. Also other standard illuminants have been provided such as D50, D55 and D75. Note the difference between light sources and illuminants. In contrast to real light sources, illuminants may not be physically realisable but are given by numerical values determining the SPD. Before illuminant D65, standard illuminant C has been used as the average daylight. In contrast to the D series, illuminant C can be reproduced by filtering a tungsten filament lamp. However, real daylight has more ultraviolet then illuminant C. This ultraviolet band (380-400 nm) has been specified in D65. In Figure 2, the relative spectral power distribution are given for illuminant A, C and D65. The spectral power distribution of illuminant A (black body with a color temperature of 2856 K) is similar to that of a incandescent lamp.

Artificial light sources have been used in a large variety of places for illuminating, for example, indoor buildings and outdoor roads. The most important light sources are the fluorescent lamps based on gas-discharge of rare gasses such as neon and xenon, or metallic vapors such as mercury and sodium vapor. The filtered xenon arc lamp approximates the D65 illuminant (i.e. average daylight) the closest from all existing light sources. The xenon arc lamp has a high amount of continuum in contrast to other discharge lamps which have short peaks in their SPD corresponding to the characteristics of the excited molecules (gas dependent). Finally, standard illuminants have been recommended by the CIE denoted by illuminant F1-F12 corresponding to SPD of different type of fluorescent lamps.

In conclusion, color temperatures are used to describe the relative spectral power distribu-

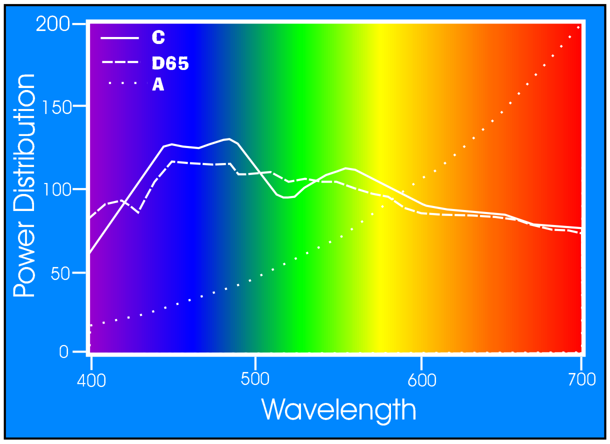

Figure 2: *a. Relative spectral power distribution of illuminant A, C, and D65.*

tion of light sources. The CIE recommended different SPD's as standards such as illuminant A (incandescent light), D65 and C (average daylight), and F1-F12 (fluorescent lamps). Manufactures are in search of lamps having a SPD similar to that of average daylight D65. Filtered xenon light is the closest approximation but is very expensive. Fluorescent lamps are cheap and mostly used for the lightning in stores. However, their SPD differ significantly from daylight. This causes the problem that the color of clothes may differ significantly if you wear it indoors or outdoors (under daylight).

3.2 The Object

In this chapter, colored materials are called objects or samples. Objects may consist of different materials such as plastics, glass, wood, textile, paper, coating, metals, brick, etc. Objects can be divided into three major groups. 1. Transparent objects, where some part of the light is absorbed and some part goes unscattered through the sample (e.g. sunglasses, windows, bottles etc), 2. Translucent objects, which absorb, transmit and scatter parts of the light (e.g. lamp panels and shades), 3. Opaque objects, where objects absorb and reflect a portion of the light and no light is transmitted (e.g. paper, textile and walls). Because most of the material that surround us is opaque, in the chapter, we will focus on opaque materials.

Objects are characterized by the percentage of light that they reflect (or transmit) at each wavelength of the visible spectrum. The amount of reflected light from an object depends on the color of the material. The amount of light reflected from an object is computed with respect to that of a white standard. The ratio of the two reflected values is the reflectance factor determining the so-called relative spectral reflectance curve, a number from 0 to 1 (or 0 to 100 in percentage). Thus the relative spectral reflectance curve is not absolute but relative to a white reference and is denoted by $S(\lambda)$. Spectral reflectance curves can be obtained by measurement instruments such as spectrophotometers. The reflectance curve can be seen as the fingerprint of the object's color. In Figure 3.a the spectral curves are shown for green and orange paint samples. The green paint sample absorbs the blue and red parts of the visible spectrum. Further, the orange paint sample absorbs the green and blue part of the visible spectrum. The SPD of a reflecting object is calculated by adding the product of the spectral power distribution of the illuminant and the reflectance (transmittance) of the object at each wavelength of the visible spectrum:
spectrum:

$$P(\lambda) = E(\lambda)S(\lambda) \tag{1}$$

where $P(\lambda)$ is the spectral power in the light reflected by the object with reflectance $S(\lambda)$ under illuminant $E(\lambda)$.

Figure 3 illustrates this concept. The spectral power distribution reflected from a green sample with reflectance given in Figure 3.a when the light illuminant has the spectral power distribution of Figure 3.b is shown in Figure 3.c.

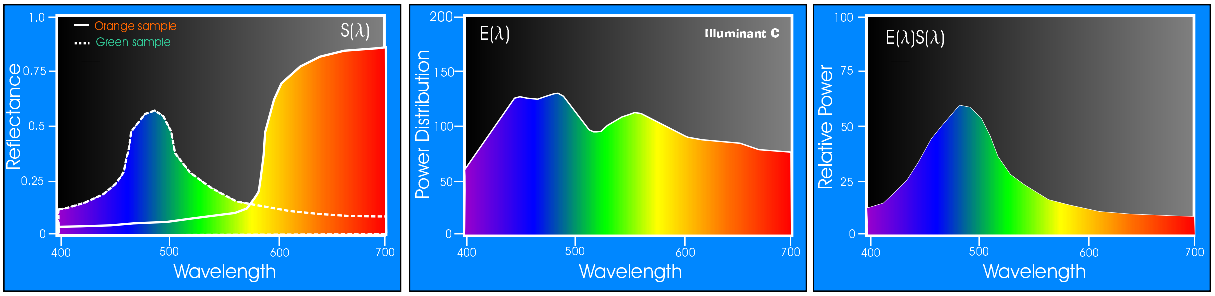

Figure 3: *a. Spectral reflectance curves of a green and orange sample. b. Relative spectral distribution of the power from illuminant C. c. Spectral power distribution of the light reflected from a green sample under illuminant C.*

In conclusion, the color of an object can be measured objectively (i.e. independent of the illuminant) and is given by the relative spectral reflectance curve. The color reflected from an object is the product of SPD of the illuminant and the spectral reflectance of the object and is computed by $P(\lambda) = E(\lambda)S(\lambda)$. In the next section, the measured SPD of $P(\lambda)$ will be transformed into three numbers corresponding to the numbers we use to describe the color of the object.

3.3 The Observer

The third component of the triplet is the observer. The observer measures light coming directly from a light source $E(\lambda)$ or light which has been reflected (or transmitted) from objects in the scene $P(\lambda)$. The observer can be a color camera or the human eyes. For the human eye, the retina contains two different types of light-sensitive receptors, called *rods* and *cones*. Rods are more sensitive to light and are responsible for vision in twilight. There are theories stating that rods also playing a little role in color vision. The cones are responsible for color vision and consist of three types of receptors sensitive to long (red), middle (green) and short (blue) wavelengths. The human eyes have been studied by many scientists. There are two theories which have been widely accepted and used in practice: the trichromacy and opponent theory. In fact, the two theories are completive. It has been recognized that both views have their value.

In this section, the trichromacy and opponent theories are outlined. Further, the spectral sensitivities of the different receptors (cones) of the eye are given.

3.3.1 Trichromacy Theory

The comprehension of color perception started with Newton in 1666, the year in which he dispersed (white) light with a glass prism. The result was the amazing discovery that white light is composed of a full range of colors. In 1801, the English scientist Thomas Young suggested the theory that just three primary additive colors were sufficient to produce all colors. The trichromacy theory has been expanded by the German scientist Hermann von Helmholtz and is usually called the Young-Helmholtz approach. They suggested that the human eye perceives color by the stimulation of three pigments in the cones of the retina. So the perceived color by the human eye has three degrees of freedom. By taking three arbitrary colors from the electromagnetic spectrum, the range of perceived colors which is obtained by the mixture of these three colors with corresponding intensities, is called the *color gamut* and the colors themselves are called *primary colors*. The trichromacy theory has been confirmed only in 1960, where three types of receptors have been identified in the retina. The maximum responses of these receptors correspond to blue (near 440 nm), green (near 540 nm) and red (near 590 nm).

The trichromacy theory has found its way in different applications such as color camera's, camrecorders and video monitors. For example, a T.V. tube is composed of a large array of triangular dot patterns of electron-sensitive phosphor. Each dot in a triad is capable of producing light in one of the primary colors. When each dot in a triad is activated by electrons it emits light, and the amount of light depends upon the strength of the electron stream. The three primary colors from each phosphor triad are added together and the result will be a color because the phosphor dots are very small. The color of a triangular dot pattern depends upon the ratio of the primary color intensities.

3.3.2 Opponent Theory

The opponent color theory started at about 1500 when Leonardo da Vinci came to the conclusion that colors are produced by the mixture of yellow and green, blue and red, white and black. Arthur Shopenhauer noted the same opposition of red-green, yellow-blue and white-black. This opponent color theory has been completed by Edwald Hering concluding that the working of the eye is based on the three kinds of opposite colors. An example of opponent color theory is the so-called after- image. Looking for a while at a green sample will cause a red after-image (excluding yellow and blue). Focusing on the chromatic channels (i.e. red-green and blue-yellow), they are opponent in two different ways. First, as mentioned above, no color seems to be a mixture of both members of any opponent pair. Secondly, each member of an opponent pair exhibits the other. In other words, by adding a balanced proportion of green and red, a hue will be produced which is neither greenish nor reddish. The opponent color theory has been confirmed in 1950 where opponent color signals were detected in the optical connection between eye and brain.

Modern theories combine the trichromacy and opponent color theory; the process starts by light entering the eye, which is detected by trichromacy cones in the retina, and is further processed into three opponent signals on their way to the brain. More recent developments are in the Retinex theory proposed by Edwin Land. Experiments show that people have a considerable amount of color constancy (i.e. color are perceived the same even under different illumination). Unfortunately, all theories discussed so far are constrained and are incapable of explaining the effects of, for example, influences of intensities/colors adjacent to the sample, intensity and color distribution of the light source, mode of viewing and so on.

3.3.3 Color Response of the Eye

We have seen that the light from a light source $E(\lambda)$ as well as light reflected by an object $S(\lambda)$ can be determined with the aid of physical measurement instruments. However, the sensation of a human observer can not be measured by an objective instrument. Therefore, the spectral sensitivities of the human eyes are measured only indirectly as follows.

Experiments have been conducted on human observers without any vision anomalies. The observers were asked to match a test light, consisting of only one wavelength, by adjusting the energy level of three separate primary lights. The three primary lights were additively mixed to match the test light of one wavelength. At each wavelength the amount of energy was recorded for the three primary colors yielding the so-called color matching functions. Different triplets of primaries can be used to get different color matching functions matching the visible spectrum. Color matching functions can be transformed to those which are obtained by other primaries. Historically, above experiments have been carried by W.D. Wright (7 observers) and J. Guild (10 observers) using different primary colors: 460, 530 and 650 nm (Wright) and 460, 543 and 630 nm (Guild). The viewing angle was 2^o, corresponding to the viewing of a sample of about 0.4 inches at reading distance of 10 inches, in which the light illuminates only the fovea. Later experiments have been conducted with a viewing angle of 10^o, corresponding to the viewing of a sample of about 1.9 inches at reading distance of 10 inches, by Stiles and Burch (about 50 observers). The tricolor functions of the human eye are shown in Figure 4.a.

3.3.4 The CIE 2^o Standard Observer

The complication of the color matching functions so far is that a negative amount of at least one of the primaries was necessary to produce the full set of spectral colors, see Figure 4.a. A negative amount was accomplished by adding one of the primaries to the test spectral light. However, a negative amount is inconvenient for computational reasons. To this end, the CIE recommended mathematical transformations based on three primary standards X, Y and Z. These primary standards are not real but imaginary. The reason to choose these imaginary primary standards are: 1. No negative values in the color matching functions, 2. The primaries X, Y and Z enclose all possible colors 3. The Y primary color corresponds with intensity.

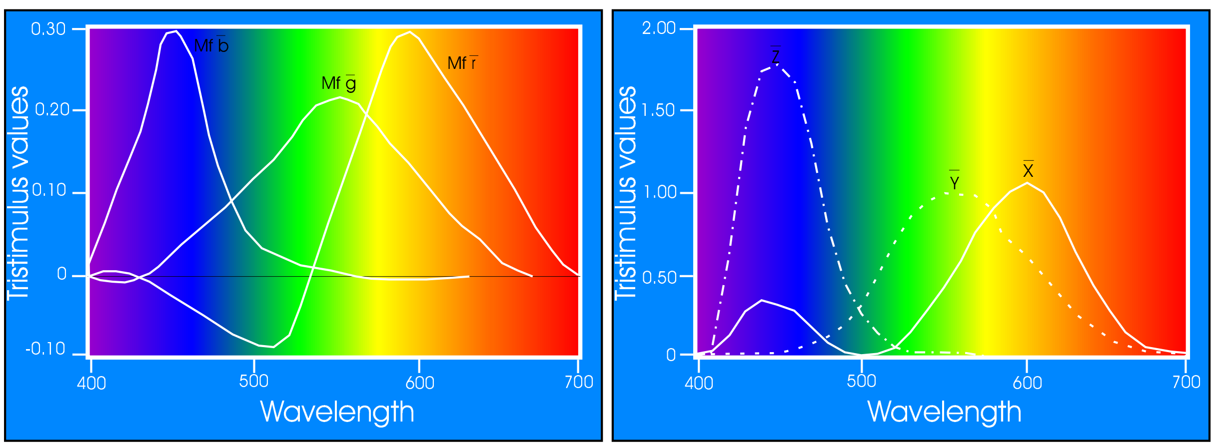

Figure 4: *a. Color matching functions $\overline{r}$, $\overline{g}$ and $\overline{b}$ of the human observer. b. Color matching functions $\overline{x}$, $\overline{y}$ and $\overline{z}$ from the standard observer 2^o (CIE 1931).*

This resulted in the CIE color matching functions denoted by $\overline{x}$, $\overline{y}$ and $\overline{z}$ giving the amount of each of the primary colors required to match a color of one watt of radiant power for each wavelength, see Figure 4.b. The curves correspond to the spectral sensitivity curves of the three receptors in the human eyes. Note that $\overline{x}$ correspond to the spectral luminance factor of the eye. Depending on the viewing angle, again two versions exist: the standard observer - CIE 1931, 2^o (shorten to the 2^o standard observer), and the standard observer - CIE 1931 10^o (cf. 10^o standard observer). The standard observer - CIE 1931 10^o has been introduced to correspond better to the visual matching of larger samples.

3.3.5 Summary

In conclusion, the spectral sensitivities of the different cone types of the human retina have been determined indirectly by testing on the additive mixture of three primary colors. The

CIE recommends two sets of standard spectral curves (also called color matching functions) depending on the angle an observer views a sample: 2^o standard observer (corresponding to the viewing of a sample of about 0.4 inches at reading distance of 10 inches) and 10^o standard observer (corresponding to the viewing of a sample of about 1.9 inches at reading distance of 10 inches). In the next section, the matching functions are used to express a color spectrum into three numbers.

4 Colorimetry

Standardization of the light-object-observer triplet is required to measure objectively the imaging process. As discussed in Section 3.1, different light sources exist with different spectral power distributions $E(\lambda)$. To get comparable results, CIE recommended various illuminants. Further, in Section 3.2, the light reflected by objects $S(\lambda)$ can be measured. Then, the reflected color of the object can be determined by $P(\lambda) = E(\lambda)S(\lambda)$. Further, the observer perceives color in terms of three color signals based on the trichromacy theory can be modeled by:

$$R = \int_\lambda E(\lambda)S(\lambda)f_R(\lambda)d\lambda \tag{2}$$

$$G = \int_\lambda E(\lambda)S(\lambda)f_G(\lambda)d\lambda \tag{3}$$

$$B = \int_\lambda E(\lambda)S(\lambda)f_B(\lambda)d\lambda \tag{4}$$

where the tristimulus values are obtained by adding the product of the SPD of the light source $E(\lambda)$, the reflectance (or transmittance) factor of the object $S(\lambda)$ and the color matching functions $f_C(\lambda)$ for $C \in \{R, G, B\}$ of the observer (eye or camera) at each wavelength of the visible spectrum. Depending on the color matching functions ($\overline{r}$, $\overline{g}$ and $\overline{b}$ or $\overline{x}$, $\overline{y}$ and $\overline{z}$), the sum of each equation is equal to the amount of primaries to match the sample.

This section is outlined as follows. First, in Section 4.1 the calculation of the tristimulus values is discussed for the CIE XYZ color system. Further, chromatity coordinates are discussed to graphically represent color. In Section 4.2, calculation of the tristimulus values is derived from color matching functions from a RGB system (e.g. color camera). Different color system are then discussed and defined in terms of R, G and B coordinates. In Section 4.5, color order systems are presented computing color differences approximating human perception.

4.1 XYZ System

Having three color matching functions of the standard observer, we are now be able to compute three numbers (called the tristimulus values) equivalent to what a standard observer perceives:

$$X = \int_\lambda E(\lambda)S(\lambda)\overline{x}(\lambda)d\lambda \tag{5}$$

$$Y = \int_\lambda E(\lambda)S(\lambda)\overline{y}(\lambda)d\lambda \tag{6}$$

$$Z = \int_\lambda E(\lambda)S(\lambda)\overline{z}(\lambda)d\lambda \tag{7}$$

where $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ $\overline{z}(\lambda)$ are the CIE color matching functions of the 2^o CIE standard observer, see Figure 4.b.

In this way, we are able to compute the color tristimuli values equivalent to a CIE 1931 2^o standard observer viewing an object of green paint illuminated by light source D65. Consider the reflectance graph of a green sample. Further, assume that the light source is illuminant D65, then the X, Y and Z tristimulus values is computed by adding up the product of light, object and matching functions at each wavelength.

The tristimulus values do not have a rather comprehensive meaning for human beings. For example, if we have two colors $X_1 = 34$, $Y_1 = 45$ and $Z_1 = 102$, and $X_2 = 64$, $Y_2 = 90$ and $Z_2 = 202$ respectively, then it can be deducted that the first color is twice that intense the second color. However, it is difficult to describe the chromaticity of the color. Therefore, the notion of the chromaticity of the color is accomplished by defining the so-called chromaticity coordinates:

$$x = \frac{X}{X+Y+Z} \tag{8}$$

$$y = \frac{Y}{X+Y+Z} \tag{9}$$

$$z = \frac{Z}{X+Y+Z} \tag{10}$$

It is obvious that the intensity information is factored out of the system, because the chromaticity coordinates reflect only the ratio of the three standard primary colors.

Since the sum of the chromaticity coordinates equals unity, two of these three quantities are sufficient to describe a color. When the x and y values are represented in a plane, the *C.I.E.* 1931 *chromaticity diagram* is obtained, see Figure 5. Spectral colors are lying on the tongue-shaped curve according to their wavelengths. In the middle of the diagram, we find neutral white at $x = 0.3333$ and $y = 0.3333$ with corresponding color temperature of 6000 K. For illuminant D65, we obtain $x = 0.3127$ and $y = 0.3290$. Illuminant C has chromaticity coordinates $x = 0.3101$ and $y = 0.3163$, see Figure 5.a. A third dimension can be imagined related to the brightness which is perpendicular to the chromaticity diagram. Instead of using X, Y and Z, a color is more easily understood when it is specified in terms of x, y and Y (i.e. intensity).

Hue and saturation are defined in the chromaticity diagram as follows, see Figure 5.b. First a reference white-light point should be defined. This point is defined as a point in the chromaticity diagram that approximately represents the average daylight, for example illuminant C or D65. For a color $G1$, the hue is defined as the wavelength on the spectral curve that intersects the line from reference white through $G1$ to the spectral curve, which is $G2$ at 523 nm. When $||G1||$ is the distance from $P1$ to the white-point, then the saturation is the relative distance to the white-point, given by $\frac{||G1||}{||G2||}$, where $||G2||$ is the distance from $G2$ to the white point. Hence the most saturated colors are the spectral colors.

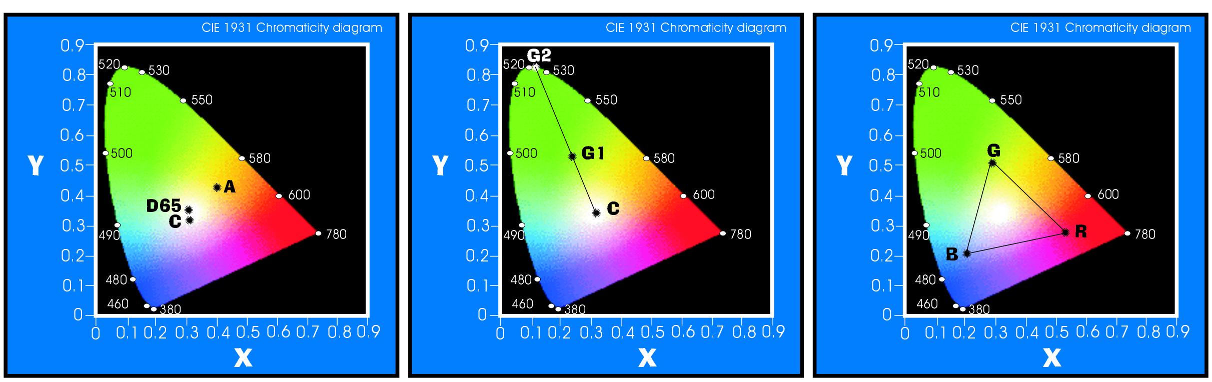

Figure 5: *a. Illuminant A, C and D65 plotted in the xy-plane. b. Definition of hue and saturation under illuminants A and D65. c. Color gamuts.*

Color gamuts are represented in the chromaticity diagram by lines joining the points defining the color gamut, see Figure 5.c. All colors along the line joining R and G can be obtained by mixing the amounts, corresponding to the distances from R and G, of the colors represented by the endpoints. The color gamut for three points R, G, and B is the triangle formed by the three vertices of the three colors. Any color within the triangle can be produced by a weighted sum of the three colors of the triangle, but no colors represented by points outside the triangle could be produced by these colors. It is now obvious why no set of three primary colors can produce all colors, since no triangle within the diagram can encompass all colors. It also explains why there are many different standards possible for the primary colors, because any triangle in the diagram, defined by its vertices, could be taken as a standard.

In conclusion, we are now able to assign precise numerical values to the color sensation of a standard observer in terms of intuitive attributes hue, saturation and brightness in a total objective manner. In fact, the XYZ system introduced by CIE is the scientific basis of objective color measurement. The XYZ system allows us to compute tristimulus values describing the sensation of a human being, given by matching functions $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ and $\overline{z}(\lambda)$, viewing an object $S(\lambda)$, illuminated by a light source $E(\lambda)$). Although, any set of primary colors can be taken, in the next section, the RGB primaries are given of a color camera. The purpose of the RGB system is to derive mathematical formulae to define color systems in terms of RGB coordinates directly coming from a color camera.

4.2 RGB System

As discussed above, a linear function of the tristimulus values converts a set of color primaries into another set. The standard RGB established by CIE in 1931 with three monochromatic primaries at wavelengths 700 nm (R)ed, 546.1 nm (G)reen, and 435.8 nm (B)lue is the RGB Spectral Primary Color Coordinate System corresponding to the 2^o standard observer. The color matching functions have already been shown in Figure 4.b. Another set of primaries have been recommended by the National Television Systems Committee (NTSC) for phosphor standardization. As the NTSC primaries are not pure monochromatic sources of radiation, the color gamut produced by the NTSC primaries is smaller than available from the spectral primaries. Digital camera's and video's mostly use RGB NTSC.

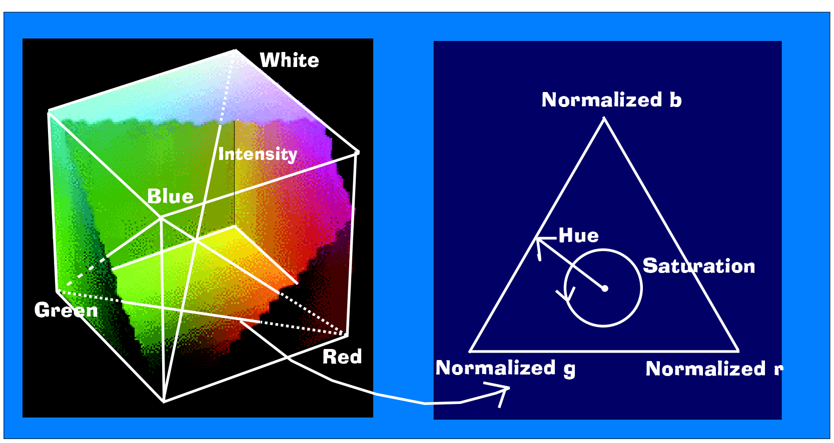

Figure 6: *a. RGB-color space b. Definition of hue and saturation in the chromaticity plane.*

To represent the RGB-color space, a cube can be defined on the R, G, and B axes, see Figure 6.a. White is produced when all three primary colors are at M, where M is the maximum light intensity, say $M = 255$. The axis connecting the black and white corners defines the intensity:

$$I(R,G,B) = R + G + B \tag{11}$$

All points in a plane perpendicular to the grey axis of the color cube have the same intensity. The plane through the color cube at points $R = G = B = M$ is one such plane. The projection of RGB points on the rgb chromaticity triangle is defined by:

$$r(R,G,B) = \frac{R}{R+G+B} \tag{12}$$

$$g(R,G,B) = \frac{G}{R+G+B} \tag{13}$$

$$b(R,G,B) = \frac{B}{R+G+B} \tag{14}$$

yielding the rgb color space which is normalized with respect to intensity and graphically represented by Figure 6.a and 6.b where the intensity axis I in RGB-space is projected onto $r = g = b = 1/3$ in the chromaticity plane.

4.3 HSI System

Hue and saturation are defined in the chromaticity triangle in the standard way as follows. Similar to Section 4.1 for the xy diagram, the reference white point is first defined in the rg chromaticity diagram. Let's assume a white light source and hence the reference point is represented by $r = g = b = 1/3$. Then saturation S_{rgb} is defined as the radial distance of a point from the reference white- light point mathematically specified as:

$$S_{rgb}(r,g,b) = \sqrt{(r-1/3)^2 + (g-1/3)^2 + (b-1/3)^2} \tag{15}$$

and graphically represented by Figure 6.b.

H_{rgb} is a function of the angle between a reference line (e.g. horizontal axis) and the color point:

$$H_{rgb}(r,g,b) = \arctan\left(\frac{r-1/3}{g-1/3}\right) \tag{16}$$

see Figure 6.c as well.

The transformation from RGB used here to compute H in terms of R, G and B is given by [14]:

$$H(R,G,B) = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right) \tag{17}$$

and S measuring the relative white content of a color as having a particular hue by:

$$S(R,G,B) = 1 - \frac{\min(R,G,B)}{R+G+B} \tag{18}$$

Note that rgb, H and S are undefined for achromatic colors (i.e. $R = G = B = 0$).

In conclusion, hue, saturation and intensity are calculated from the original R, G, B values from the corresponding red, green, and blue images provided by the color camera.

4.4 YIQ and YUV System

Various camera's provide images in terms of the YIQ NTSC transmittion color coordinate system. The National Television Systems Committee (N.T.S.C.) developed the three color attributes Y, I, and Q for transmission efficiency. The tristimulus value Y corresponds to the luminance of a color. I and Q correspond closely the hue and saturation of a color. By reducing the spatial bandwidth of I and Q without noticeable image degradation, efficient color transmittion is obtained. The PAL and SECAM standards used in Europe, the Y, U, and V tristimulus values are used. The I and Q color attributes are related to U and V by a simple rotation of the color coordinates in color space. The conversion matrix to compute the YIQ values from the original RGB NTSC tristimulus values is given by: $Y = 0.299 * R + 0.587 * G + 0.114 * B$, $I = 0.596 * R - 0.274 * G - 0.312 * B$, and $Q = 0.211 * R - 0.523 * G + 0.312 * B$.

4.5 Color Order Systems

A color system is visual uniform when numerical distances can be related to human perceptual differences: the closer a color is to another color in the color space, the more similar they are. It is known that XYZ and RGB are not visual uniform. To achieve visual uniformity, color order systems have been proposed. A color order system is a multi-dimensional (usually three) space arranging the gamut of colors of visual sensation. This geometric ordering relates a notation to each color corresponding to its position. Each color is given a number yielding an objective classification criterion. Hence, the goal of a color order system is to represent and classify colors objectively, and to provide color differences. Color order systems are often accompanied by atlases and catalogues containing real samples from specific position within the space. Color systems are usually based on perceptual uniformity where the Euclidean distance between two position in the space closely corresponds to the difference between the two colors as perceived by a human being. In this section, the Munsell's color system is discussed in Section 4.5.1. The CIELAB colorimetric space will be given in Section 4.5.2. Color difference will be discussed in Section 4.5.3

4.5.1 Munsell's Color Order System

Albert H. Munsell proposed in 1905, in his "A Color Notation", a geometric ordering space to classify colors. The publication was accompanied with an Atlas containing charts, derived from psychological and experimental measurements, designed to produce a uniform color space. The Munsell color space is based on three perceptual attributes. First, he developed the notation of *Value* ranging from white to black with perceptual equal steps. Further, a *Hue* circle has been designed corresponding to a certain *Chroma* value. The hue circle contains the property that steps between two different hues closely correspond to their perceptual differences. To graphically illustrate the Munsell color space, the color order system can be represented by a cylinder from which the center axis relates to Value. Further, Chroma corresponds to the distance from this central axis of constant Hue forming circles around the Value axis. In this way, Munsell provided a visually balanced ordering of colors to yield an objective color identification scheme. The Munsell system has been updated by modern instruments. Today, it is possible to compute the Munsell notations from the CIE XYZ tristimulus values.

4.5.2 CIELAB Color Order Systems

As stated above, a color system is visual uniform when numerical distances can be related to human perceptual differences. Hence, the closer a color is to another color in the color space, the more similar they are. MacAdam proved the visual non-uniformity of the XYZ color system by doing experiments. These experiments are based on the *just noticeable difference* (JND) concept.

The JND is defined as the minimal visual difference of two colors: let Q_0 be a color then color Q_1 is just noticeable different from Q_0, if there is no color Q_2, lying on the line from Q_0 through Q_1, which is closer to and noticeable different from Q_0. The results of the experiments made by MacAdam is that colors with coordinates (x_i, y_i) in the xy chromaticity diagram, which are just noticeable different from color center (x_0, y_0), are lying on an ellipsis with center (x_0, y_0) and all colors which are not (just) noticeable different from (x_0, y_0) are lying outside the ellipsis. The ellipses of an arbitrary number of colors were calculated and analyzed. MacAdam proved that the differences between the axis-diameters of the ellipses were rather (too) large and therefore non-uniformity of the XYZ color system can easily be seen.

Over the last decades various attempts have been made to develop perceptual uniform spaces and color difference equations. To this end, in 1976, the CIE recommended a new system $L^*a^*b^*$, computed from the XYZ color system, having the property that the closer a point (representing a color) is to another point (also representing a color -possibly the same-), the more visual similar the colors (represented by the points) are. In other words, the magnitude of the perceived color difference of two colors corresponds to the Euclidean distance between the two colors in the color system.

The $L^*a^*b^*$ system is based on the three dimensional coordinate system based on the opponent theory using black-white L^*, red-green a^*, and yellow-blue b^* components. The L^* axis corresponds to the lightness where $L^* = 100$ is white and $L^* = 0$ is black. Further, a^* ranges from red $+a^*$ to green $-a^*$ while b^* ranges from yellow $+b^*$ to blue $-b^*$.

The L^*, a^* and b^* coordinates are computed from the X, Y and Z tristimulus values as follows:

$$L^* = 116 f(\frac{Y}{Y_n})16 \tag{19}$$

$$a^* = 500[f(\frac{X}{X_n} - \frac{Y}{Y_n})] \tag{20}$$

$$b^* = 500[f(\frac{Y}{Y_n} - \frac{Z}{Z_n})] \tag{21}$$

where X_n, Y_n and Z_n are the tristimulus values of the nominal white stimulus. For example, for D65/10^o: $X_n = 94.81$m $Y_n = 100.00$ and $Z_n = 107.304$. Further, $f(\frac{X}{X_n}) = (\frac{X}{X_n})^{1/3}$ for values $\frac{X}{X_n}$ greater than 0.008856 and $f(\frac{X}{X_n}) = 7.787(\frac{X}{X_n}) + 16/116$ for values $\frac{X}{X_n}$ equal or less than 0.008856. Equal arguments hold for Y and Z.

The intuitive visual attributes hue and chroma (saturation) can be expressed by:

$$C^* = (a^{*2} + b^{*2})^{1/2} \tag{22}$$

$$h = \arctan(\frac{b^*}{a^*}) \tag{23}$$

Note that the CIE $L^*a^*b^*$ chroma is not the same as the Munsell Chroma. Further, the CIE $L^*a^*b^*$ is still not perfectly visual uniform. However, the system allows us to define a color difference by measuring the geometric distance between two colors in the color space. The color difference ruler is discussed in the following section.

4.5.3 Color Difference

Many image retrieval applications must provide the ability to retrieve visual similar images satisfying the expectation of the operator. Therefore, images should be retrieved within the limits of acceptable variations of color. For example, consider a database of images taken from paper samples. The operator searches for samples with a specific color. Then, color differences are essential for the retrieval task at hand. Note the difference between perceptibility and acceptability. The color difference is perceptible if the difference can be seen by a human operator. Even when the difference is perceptible for different tasks, the color difference can still be acceptable. As discussed above, the $L^*a^*b^*$ is approximately visual uniform and provides us the ability to define a perceptual color difference in the Euclidean way.

First, the differences are subtracted from the L^*, a^*, and b^* components of the trial from the standard:

$$\Delta L^* = {L^*}_{\text{trial}} - {L^*}_{\text{standard}} \tag{24}$$

$$\Delta a^* = {a^*}_{\text{trial}} - {a^*}_{\text{standard}} \tag{25}$$

$$\Delta b^* = {b^*}_{\text{trial}} - {b^*}_{\text{standard}} \tag{26}$$

where a positive value of ΔL^* implies that the sample is lighter than the standard one. Further, a positive value of Δa^* and Δb^* and indicates that the sample is redder and yellow than the standard.

The total CIE $L^*a^*b^*$ is given by:

$$\Delta E^*_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{27}$$

The difference in saturation is given by:

$$\Delta C^* = {C^*}_{\text{trial}} - {C^*}_{\text{standard}} \tag{28}$$

Due to the circular nature of hue, the hue difference is given by:

$$\Delta H^*_{ab} = \sqrt{(\Delta E^*_{ab})^2 + (\Delta C^*)^2 + (\Delta L^*)^2} \tag{29}$$

4.5.4 Summary

Color order systems are required to identify colors into a geometric structure. The Munsell space was one of the first color order system with the aim to represent color objectively. Due to the perceptual nonuniformity of the CIE XYZ system, the CIE recommended the $L^*a^*b^*$ color order system where the Euclidean distance between two colors closely corresponds to the visual color difference.

5 Color Invariance

In the previous section, the scientific measurement of color has been outlined to measure objectively the light-object-observer process yielding basic notions on color, color models and ordering systems. The advantage of using, for example, the $L^*a^*b^*$ system, is that it corresponds with human perception which is useful when retrieving images which are visually similar. However, it is known that the $L^*a^*b^*$ and RGB color system are dependent on the imaging conditions. Therefore, for general image retrieval, an important color property is color invariance. A color invariant system contains color invariant models which are more or less insensitive to the varying imaging conditions such as variations in illumination (e.g. lamps having different spectral power distribution) and object pose (changing shadows, shading and highlighting cues). Color invariants for the purpose of viewpoint invariant image retrieval is given in [10]. In this section, an overview is given of color invariants.

This section is outlined as follows. First, in Section 5.1, a general reflection model is discussed to model the interaction between light and material. Then, in Section 5.2, assuming white illumination, basic color systems are analyzed with respect to their invariance. In Section 5.3, color constant color systems are given. Color constant color systems are independent of the relative spectral power distribution of the light source.

5.1 Reflection from Inhomogeneous Dielectric Materials

To arrive at a uniform description it is common to divide opaque materials into two classes on the basis of their optical properties: optically homogeneous and inhomogeneous materials.

Optically homogeneous materials have a constant index of refraction throughout the material. This means that when light is incident upon the surface of an homogeneous materials, some fraction of it is reflected. A smooth surface reflects light only in the direction such that the angle of incidence equals the angle of reflection. The properties of this reflected light are determined by the optical and geometric properties of the surface. Metals are homogeneous materials having a larger specular component than other materials.

Optically inhomogeneous materials are composed of a vehicle with many embedded colorant particles that differ optically from the vehicle. The fraction of the incident light that is not reflected from the surface enters the body of the material. The body is composed of a vehicle and many colorant particles. When light encounters a colorant particle, some portion of it is reflected. After many reflections, the light is diffused, and a significant fraction can exit back through the surface in a wide range of directions. Some examples of inhomogeneous materials are plastics, paper, textiles, and paints.

The Dichromatic Reflection Model describes the light, which is reflected from a point on a dielectric, inhomogeneous material as a mixture of the light reflected at the material surface and the light reflected from the material body.

Let $E(\vec{x}, \lambda)$ be the spectral power distribution of the incident (ambient) light at the object surface at $\vec{x}$, and let $L(\vec{x}, \lambda)$ be the spectral reflectance function of the object at $\vec{x}$. The spectral sensitivity of the kth sensor is given by $F_k(\lambda)$. Then ρ_k, the sensor response of the kth channel, is given by:

$$\rho_k(\vec{x}) = \int_\lambda E(\vec{x}, \lambda) L(\vec{x}, \lambda) F_k(\lambda) d\lambda \tag{30}$$

where λ denotes the wavelength, and $L(\vec{x}, \lambda)$ is a complex function based on the geometric and spectral properties of the object surface. The integral is taken from the visible spectrum (e.g. 380-700 nm).

Further, consider an opaque inhomogeneous dielectric object, then the geometric and surface reflection component of function $L(\vec{x}, \lambda)$ can be decomposed in a body and surface reflection component as described by Shafer [24]:

$$\phi_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s}) \int_\lambda E(\vec{x}, \lambda) B(\vec{x}, \lambda) F_k(\lambda) d\lambda +$$

$$G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v}) \int_\lambda E(\vec{x}, \lambda) S(\vec{x}, \lambda) F_k(\lambda) d\lambda \tag{31}$$

giving the kth sensor response. Further, $B(\vec{x}, \lambda)$ and $S(\vec{x}, \lambda)$ are the surface albedo and Fresnel reflectance at $\vec{x}$ respectively. $\vec{n}$ is the surface patch normal, $\vec{s}$ is the direction of the illumination source, and $\vec{v}$ is the direction of the viewer. Geometric terms G_B and G_S denote the geometric dependencies on the body and surface reflection component respectively.

5.2 Reflectance with White Illumination

Considering the neutral interface reflection (NIR) model (assuming that $S(\vec{x}, \lambda)$ has a nearly constant value independent of the wavelength) and approximately white illumination, then $S(\vec{x}, \lambda) = S(\vec{x})$, and $E(\vec{x}, \lambda) = E(\vec{x})$. Then, the measured sensor values are given by [9]:

$$\omega_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x}) \int_\lambda B(\vec{x}, \lambda)F_k(\lambda)d\lambda +$$

$$G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v})E(\vec{x})S(\vec{x}) \int_\lambda F_k(\lambda)d\lambda \tag{32}$$

giving the kth sensor response of an infinitesimal surface patch under the assumption of a white light source.

If the integrated white condition holds (i.e. the area under the sensor spectral functions is approximately the same):

$$\int_\lambda F_i(\lambda)d\lambda = \int_\lambda F_j(\lambda)d\lambda \tag{33}$$

the reflection from inhomogeneous dielectric materials under white illumination is given by [9]:

$$\omega_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x}) \int_\lambda B(\vec{x}, \lambda)F_k(\lambda)d\lambda +$$

$$G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v})E(\vec{x})S(\vec{x})F \tag{34}$$

If $\omega(\vec{x})$ is not dependent on $\vec{x}$, we obtain:

$$\omega_k = G_B(\vec{n}, \vec{s})E \int_\lambda B(\lambda)F_k(\lambda)d\lambda + G_S(\vec{n}, \vec{s}, \vec{v})ESF \tag{35}$$

5.2.1 Invariance for Matte Surfaces

Consider the body reflection term of equation (34):

$$\beta_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x}) \int_\lambda B(\vec{x}, \lambda)F_k(\lambda)d\lambda \tag{36}$$

giving the kth sensor response of an infinitesimal *matte* surface patch under the assumption of a white light source.

The body reflection component describes the way light interacts with a dull surface. The light spectrum E falls on a surface B. The geometric and photometric properties of the body reflection depends on many factors. If we assume a random distribution of the pigments, the light exits in random directions from the body. In this case, the distribution of exiting light can be described by Lambert's law. Lambertian reflection models dull, matte surfaces which appear equally bright regardless from angle they are viewed. They reflect light with equal intensity in all directions.

As a consequence, a uniformly colored surface which is curved (i.e. varying surface orientation) gives rise to a broad variance of RGB values. The same argument holds for intensity I.

In contrast, rgb is insensitive to surface orientation, illumination direction and illumination intensity mathematically specified by substituting equation (36) in equation (12) - (14):

$$r(R_b, G_b, B_b) = \frac{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})k_R}{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})(k_R + k_G + k_B)} = \frac{k_R}{k_R + k_G + k_B} \tag{37}$$

$$g(R_b, G_b, B_b) = \frac{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})k_G}{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})(k_R + k_G + k_B)} = \frac{k_G}{k_R + k_G + k_B} \tag{38}$$

$$r(R_b, G_b, B_b) = \frac{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})k_B}{G_B(\vec{x}, \vec{n}, \vec{s})E(\vec{x})(k_R + k_G + k_B)} = \frac{k_B}{k_R + k_G + k_B} \tag{39}$$

factoring out dependencies on illumination and object geometry and hence only dependent on the sensors and the surface albedo.

Because S corresponds to the radial distance from the color to the main diagonal in the RGB-color space, S is an invariant for matte, dull surfaces illuminated by white light cf. equation (36) and equation (18):

$$S(R_b, G_b, B_b) = 1 - \frac{\min(G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})k_R, G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})k_G, G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})k_B)}{G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})(k_R + k_G + k_B)} =$$

$$1 - \frac{\min(k_R, k_G, k_B)}{(k_R + k_G + k_B)} \tag{40}$$

only dependent on the sensors and the surface albedo.

Similarly, H is an invariant for matte, dull surfaces illuminated by white light cf. equation (36) and equation (17):

$$H(R_b, G_b, B_b) = \arctan(\frac{\sqrt{3}G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})(k_G - k_B)}{G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})((k_R - k_G) + (k_R - k_B))}) =$$

$$\arctan(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) + (k_R - k_B)}) \tag{41}$$

Obviously, in practice, the assumption of objects composed of matte, dull surfaces is not always realistic. To that end, the effect of surface reflection (highlights) is discussed in the following section.

5.2.2 Invariance for Shiny Surfaces

Consider the surface reflection term of equation (34):

$$C_s = G_S(\vec{x},\vec{n},\vec{s},\vec{v})E(\vec{x})S(\vec{x})F \tag{42}$$

for $C_s \in \{R_s, G_s, B_s\}$ giving the red, green and blue sensor response for a highlighted infinitesimal surface patch with white illumination.

When light hits the surface of a dielectric inhomogeneous material, it must first pass through the interface between the surrounding medium (e.g. air) and the material. Since the refraction index of the material is generally different from that of the surrounding medium, some percentage of the incident light is reflected at the surface of the material.

Several models have been developed in the physics and computer graphics communities to describe the geometric properties of the light reflected from rough surfaces. Issues in modeling this process involve the roughness scale of the surface, compared to the wavelengths of the incident light, and self-shadowing effects on the surface which depend on the viewing direction and the direction of the incident light. The common used surface reflection model is described by a function with a sharp peak around the angle of perfect mirror reflection. For smooth surfaces, surface (specular) reflections are in the direction of $\vec{r}$, which is $\vec{s}$ mirrored about $\vec{n}$. For non-perfect rough reflectors, the maximum surface reflection occurs when α, the angle between $\vec{r}$ and $\vec{e}$ is zero; and falls off rapidly as α increases. The falloff is approximated by $\cos^n \alpha$, where n is the surface's surface-reflection exponent or degree of shininess.

Because H is a function of the angle between the main diagonal and the color point in RGB-sensor space, all possible colors of the same (shiny) surface region (i.e. with fixed albedo) have to be of the same hue as follows from substituting equation (42) in equation (17):

$$H(R_w, G_w, G_w) = \arctan(\frac{\sqrt{3}(G_w - B_w)}{(R_w - G_w) + (R_w - B_w)}) =$$

$$\arctan(\frac{\sqrt{3}G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})(k_G - k_B)}{G_B(\vec{x},\vec{n},\vec{s})E(\vec{x})((k_R - k_G) + (k_R - k_B))}) =$$

$$\arctan(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) + (k_R - k_B)}) \quad (43)$$

factoring out dependencies on illumination, object geometry, viewpoint and specular reflection coefficient and hence only dependent on the sensors and the surface albedo. Note that $C_w = G_B(\vec{x},\vec{n},\vec{s})E(\vec{x}) \int_\lambda B(\vec{x},\lambda)F_C(\lambda)d\lambda + G_S(\vec{x},\vec{n},\vec{s},\vec{v})E(\vec{x})S(\vec{x})F$ for $C = \{R, G, B\}$.

Obviously other color features depend on the contribution of the surface reflection component and hence are sensitive to highlights.

5.3 Color Constancy

Existing color constancy methods require specific a priori information about the observed scene (e.g. the placement of calibration patches of known spectral reflectance in the scene) which will not be feasible in practical situations, [6], [7], [13] for example. To circumvent these problems, simple and effective illumination-independent color ratio's have been proposed by Funt and Finlayson [8] and Nayar and Bolle [17]. In fact, these color models are based on the ratio of surface albedos rather then the recovering of the actual surface albedo itself. However, these color models assume that the variation in spectral power distribution of the illumination can be modeled by the coefficient rule or von Kries model, where the change in the illumination color is approximated by a 3x3 diagonal matrix among the sensor bands and is equal to the multiplication of each RGB-color band by an independent scalar factor. The diagonal model of illumination change holds exactly in the case of narrow-band sensors. Consider the the body reflection term of the dichromatic reflection model:

$$C_c = G_B(\vec{x},\vec{n},\vec{s})E(\vec{x}) \int_\lambda B(\vec{x},\lambda)F_k(\lambda)d\lambda \quad (44)$$

for $C = \{R, G, B\}$, where $C_c = \{R_c, G_c, B_c\}$ gives the red, green and blue sensor response of a matte infinitesimal surface patch of an inhomogeneous dielectric object under unknown spectral power distribution of the illumination.

Suppose that the sensor sensitivities of the color camera are narrow-band with spectral response be approximated by delta functions $f_K(\lambda) = \delta(\lambda - \lambda_K)$, then the measured sensor values are:

$$C_K = G_B(\vec{x},\vec{n},\vec{s})E(\vec{x},\lambda_K)B(\vec{x},\lambda_K) \quad (45)$$

The color ratio's proposed by Nayar and Bolle is given by [17]:

$$N(C^{\vec{x}_1}, C^{\vec{x}_2}) = \frac{C^{\vec{x}_1} - C^{\vec{x}_2}}{C^{\vec{x}_2} + C^{\vec{x}_1}} \quad (46)$$

and by Funt and Finlayson by [8]:

$$F(C^{\vec{x}_1}, C^{\vec{x}_2}) = \frac{C^{\vec{x}_1}}{C^{\vec{x}_2}} \quad (47)$$

expressing color ratio's between two neighboring image locations, for $C \in \{R, G, B\}$, where $\vec{x}_1$ and $\vec{x}_2$ denote the image locations of the two neighboring pixels. Note that the set $\{R, G, B\}$ must be colors from narrow-band sensor filters and that they are used in defining the color ratio because they are immediately available from a color camera, but any other set of narrow-band colors derived from the visible spectrum will do as well. Although standard video camera's are not equipped with narrow-band filters, spectral sharpening could be applied [4] to achieve this to a large extend.

Assuming that the color of the illumination is locally constant (i.e. $E(\vec{x}_1, \lambda_K) = E(\vec{x}_2, \lambda_K)$) and that neighboring points have the same surface orientation (i.e. $G_B(\vec{x}_1, \vec{n}, \vec{s}) = G_B(\vec{x}_2, \vec{n}, \vec{s})$), then the color ratio N is independent of the illumination intensity and color as shown by substituting equation (46) in equation (44):

$$\frac{G_B(\vec{x}_1, \vec{n}, \vec{s})E(\vec{x}_1, \lambda_K)B(\vec{x}_1, \lambda_K) - G_B(\vec{x}_2, \vec{n}, \vec{s})E(\vec{x}_2, \lambda_K)B(\vec{x}_2, \lambda_K)}{G_B(\vec{x}_1, \vec{n}, \vec{s})E(\vec{x}_1, \lambda_K)B(\vec{x}_1, \lambda_K) + G_B(\vec{x}_2, \vec{n}, \vec{s})E(\vec{x}_2, \lambda_K)B(\vec{x}_2, \lambda_K)} = \frac{B(\vec{x}_1, \lambda_K) - B(\vec{x}_2, \lambda_K)}{B(\vec{x}_1, \lambda_K) + B(\vec{x}_2, \lambda_K)} \tag{48}$$

Equal arguments hold for the color ratio F by substituting equation (47) in equation (44):

$$\frac{G_B(\vec{x}_1, \vec{n}, \vec{s})E(\vec{x}_1, \lambda_K)B(\vec{x}_1, \lambda_K)}{G_B(\vec{x}_2, \vec{n}, \vec{s})E(\vec{x}_2, \lambda_K)B(\vec{x}_2, \lambda_K)} = \frac{B(\vec{x}_1, \lambda_K)}{B(\vec{x}_2, \lambda_K)} \tag{49}$$

However, it is assumed that the neighboring points, from which the color ratio's are computed, have the same surface normal. Therefore, the method depends on varying surface orientation of the object (i.e. the geometry of the objects) affecting negatively the recognition performance. To this end, a color constant color ratio has been proposed not only independent of the illumination color but also discounting the object's geometry [9]:

$$m(C_1^{\vec{x}_1}, C_1^{\vec{x}_2}, C_2^{\vec{x}_1}, C_2^{\vec{x}_2}) = \frac{C_1^{\vec{x}_1} C_2^{\vec{x}_2}}{C_1^{\vec{x}_2} C_2^{\vec{x}_1}}, C_1 \neq C_2 \tag{50}$$

expressing the color ratio between two neighboring image locations, for $C_1, C_2 \in \{R, G, B\}$ where $\vec{x}_1$ and $\vec{x}_2$ denote the image locations of the two neighboring pixels.

The color ratio is independent of the illumination intensity and color, and also to a change in viewpoint, object geometry, and illumination direction as shown by substituting equation (50) in equation (44):

$$\frac{(G_B(\vec{x_1}, \vec{n}, \vec{s})E(\vec{x_1}, \lambda_K)B(\vec{x_1}, \lambda_{C_1}))(G_B(\vec{x_2}, \vec{n}, \vec{s})E(\vec{x_2}, \lambda_K)B(\vec{x_2}, \lambda_{C_2}))}{(G_B(\vec{x_2}, \vec{n}, \vec{s})E(\vec{x_2}, \lambda_K)B(\vec{x_2}, \lambda_{C_1}))(G_B(\vec{x_1}, \vec{n}, \vec{s})E(\vec{x_1}, \lambda_K)B(\vec{x_1}, \lambda_{C_2}))} = \frac{B(\vec{x_1}, \lambda_{C_1})B(\vec{x_2}, \lambda_{C_2})}{B(\vec{x_2}, \lambda_{C_1})B(\vec{x_1}, \lambda_{C_2})} \tag{51}$$

factoring out dependencies on object geometry and illumination direction $G_B(\vec{x}_1, \vec{n}, \vec{s})$ and $G_B(\vec{x_2}, \vec{n}, \vec{s})$, and illumination for $E(\vec{x_1}, \lambda_{C_2})) = E(\vec{x_2}, \lambda_{C_2})$, and hence only dependent on the ratio of surface albedos, where $\vec{x}_1$ and $\vec{x}_2$ are two neighboring locations on the object's surface not necessarily of the same orientation.

Note that the color ratio's do not require any specific a priori information about the observed scene, as the color model is an illumination-invariant surface descriptor based on the ratio of surface albedos rather then the recovering of the actual surface albedo itself. Also the intensity and spectral power distribution of the illumination is allowed to vary across the scene (e.g. multiple light sources with different SPD's), and a certain amount of object occlusion and cluttering is tolerated due to the local computation of the color ratio.

6 Color System Taxonomy

The purpose of this section is to give a compact formulation on the relevance of the color systems for the purpose of color based image retrieval. Therefore, a survey is given on the color models. Each color system is briefly discussed and transformations are given in terms of RGB NTSC tristimulus color coordinate system. Further, the following criteria are used to classify the color systems: 1. Is the color system device independent; 2. Is the color system perceptual uniform; 3. Is the color system non-linear; 4. Is the color system intuitive; 5. Is the color system robust against varying imaging conditions. The section will end up with a table summarizing the main and important characteristics of the color systems.

6.1 Grey-Value System

The color model $GREY$ or $INTENSITY$ is calculated from the original R, G, B NTSC tristimulus values from the corresponding red, green, and blue images provided by a ccd color camera and hence dependent on the imaging device. Grey is not perceptual uniform as a just noticeable brighter grey-value does not correspond with a difference between two successive grey-values. Grey is heavily influenced by the imaging conditions.

- COLOR FEATURE: $GREY$;
- TRANSFORMATION:

$$GREY = 0.299R + 0.587G + 0.144B; \tag{52}$$

- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - Linear
  - Intuitive
  - Dependent on viewing direction, object geometry, direction of the illumination, intensity and color of the illumination
- REMARKS: Grey value information.

6.2 RGB Color System

The RGB color system represents the color (R)ed, (G)reen and (B)lue color. In this chapter, the R, B, and B color features correspond to the primary colors where $R = 700$ nm, $G = 546.1$ nm, and $B = 435.8$ nm. Similar to grey-value, the RGB color system is not perceptual uniform and dependent on the imaging conditions. Using RGB values for image retrieval cause problems when the query and target image are recorded under different imaging conditions.

- COLOR MODELS: R, G, and B;
- TRANSFORMATION: no transformation;
- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - None
  - Not intuitive
  - Dependent on viewing direction, object geometry, direction of the illumination, intensity and color of the illumination
- REMARKS: No transformation required.

6.3 rgb Color System

The rgb color system has three color features: r, g, and b. These color models are called normalized colors, because each of them is calculated by dividing the values of respectively R, G, and B by their total sum. Because the r, g, and b coordinates depend only on the ratio of the R, G, and B coordinates (i.e. factoring luminance out of the system), they have the important property that they are not sensitive to surface orientation, illumination direction and illumination intensity cf. equations (37- 39). An other important property is that it is convenient to represent these features in the chromaticity diagram. Normalized colors become unstable and meaningless when the intensity is small [12].

- COLOR MODELS: r, g, and b;
- TRANSFORMATION:

$$r = \frac{R}{(R+G+B)} \tag{53}$$

$$g = \frac{G}{(R+G+B)} \tag{54}$$

$$b = \frac{B}{(R+G+B)} \tag{55}$$

- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - Nonlinear: become unstable when intensity is small
  - Not intuitive
  - Dependent on highlights and a change in the color of the illumination.
- REMARKS: conveniently represented in the chromaticity diagram.

6.4 XYZ Color System

This color system is based on the additive mixture of three imaginary primaries X, Y, and Z introduced by CIE. These primaries can not be seen by a human eye or produced, because they are too saturated. The fact that these primaries are imaginary colors is not important, since any perceived color can be described mathematically by the amounts of these primaries. An other important property is that the luminance is determined only by the Y value. Because XYZ system is a linear combination of R, G and B values, the XYZ color system inherits all the dependencies on the imaging conditions from the RGB color system. Note that the color system is device independent as the X, Y and Z values are objective in their interpretation. Further, the conversion matrix given below is based on the RGB NTSC color coordinates system.

- COLOR MODELS: X, Y, Z;
- TRANSFORMATION:

$$X = 0.607R + 0.174G + 0.200B \tag{56}$$

$$Y = 0.299R + 0.587G + 0.114B \tag{57}$$

$$Z = 0.000R + 0.066G + 1.116 \tag{58}$$

- CHARACTERISTICS:
  - Device independent
  - Not perceptual uniform
  - Linear transformation
  - Not intuitive
  - Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination
- REMARKS: all colors are described mathematically by three color features, luminance is based on the Y value alone.

6.5 xyz Color System

The color features of this system are the chromaticity coordinates: x, y, and z. Two of those chromaticity coordinates are enough to provide a complete specification and can be represented in the chromaticity diagram. Similar to rgb, this system cancels intensity out yielding independence of surface orientation, illumination direction and illumination intensity. Again problems arise when intensity is low.

- COLOR MODELS: x, y, z;
- TRANSFORMATION:

$$x = \frac{X}{(X+Y+Z)} \tag{59}$$

$$y = \frac{Y}{(X+Y+Z)} \tag{60}$$

$$z = \frac{Z}{(X+Y+Z)} \tag{61}$$

- FEATURE PROBLEMS:
  - Device independent
  - Not perceptual uniform
  - Nonlinear transformation: become unstable when intensity is small
  - Not intuitive
  - Dependent highlights and a change in the color of the illumination
- REMARKS: conveniently represented in the chromaticity diagram.

6.6 $U^*V^*W^*$ Color System

C.I.E. introduced the $U^*V^*W^*$ color system which has three color features: U^*, V^*, and W^*. The color model W^* is based on the scaling of luminance. The luminance of a color is determined only by its Y value. Scaling luminance between 0 (black) and 100(white), the scaling method starts with black and selects a *just noticeable brighter* grey-value. Taking this just noticeable brighter grey-value the next just noticeable brighter grey-value is selected. This process continues until white is reached. This process is a method to scale brightness visually and the curve obtained by the scaling method can be approximated by the following formula:

$$W^* = \begin{cases} 116(\frac{Y}{Y_0})^{\frac{1}{3}} - 16 & , \text{ if } \frac{Y}{Y_0} > 0.008856 \\ \\ 903.3(\frac{Y}{Y_0}) & , \text{ if } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \tag{62}$$

and Y_0 is a nominally white object-color stimulus.

The other two color features solve the problem of large difference of the axis-diameters of the ellipses in the chromaticity diagram, where colors which are not noticeable different for a particular color are lying on the ellipses and all colors which are (just) noticeable different are lying outside the ellipses. The system is visual uniform, because a luminance difference corresponds with the same noticed luminance difference and the ellipses in the adjusted chromaticity diagram have constant axis-diameters. The U^* and V^* color models become unstable and meaningless when intensity is small.

- COLOR MODELS: U^*, V^*, W^*,
- TRANSFORMATION:

$$U^* = 13W^*(u - u_0), \tag{63}$$

$$V^* = 13W^*(v - v_0), \tag{64}$$

$$W^* = \begin{cases} 116(\frac{Y}{Y_0})^{\frac{1}{3}} - 16 & , \text{ if } \frac{Y}{Y_0} > 0.008856 \\ \\ 903.3(\frac{Y}{Y_0}) & , \text{ if } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \tag{65}$$

$$u = \frac{4X}{(X + 15Y + 3Z)}, \tag{66}$$

$$v = \frac{6Y}{(X + 15Y + 3Z)}, \tag{67}$$

$$u_0 = \frac{4X_0}{(X_0 + 15Y_0 + 3Z_0)}, \tag{68}$$

$$v_0 = \frac{6Y_0}{(X_0 + 15Y_0 + 3Z_0)}, \tag{69}$$

- CHARACTERISTICS:
  - Device independent
  - Perceptual uniform
  - Nonlinear transformation: become unstable when intensity is small
  - Not intuitive
  - Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination
- REMARKS: Visual uniform.

X_0, Y_0, Z_0 are values of a nominally white object-color stimulus.

6.7 $L^*a^*b^*$ Color System

An other kind of visual uniform color system proposed by $C.I.E.$ is the $L^*a^*b^*$ color system. The color feature L^* correlates with the perceived luminance and corresponds to W^* of the $U^*V^*W^*$ color system. Color feature a^* correlates with the red-green content of a color and b^* reflects the yellow-blue content.

- COLOR MODELS: L^*, a^*, b^*,
- TRANSFORMATION:

$$L^* = \begin{cases} 116(\frac{Y}{Y_0})^{\frac{1}{3}} - 16 & , \text{ if } \frac{Y}{Y_0} > 0.008856 \\ \\ 903.3(\frac{Y}{Y_0}) & , \text{ if } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \tag{70}$$

$$a^* = 500[(\frac{X}{X_0})^{\frac{1}{3}} - (\frac{Y}{Y_0})^{\frac{1}{3}}], \tag{71}$$

$$b^* = 200[(\frac{Y}{Y_0})^{\frac{1}{3}} - (\frac{Z}{Z_0})^{\frac{1}{3}}]; \tag{72}$$

- CHARACTERISTICS:
  - Device independent
  - Perceptual uniform
  - Nonlinear transformation: become unstable when intensity is small
  - Not intuitive
  - Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination

6.8 $I1I2I3$ Color System

When RGB images are highly correlated, it might be desirable to down-weight this correlation. This can be achieved by computing the Karhunen-Loève transformation. This transformation is calculated from the covariance matrix and calculates an uncorrelated basis.

Three color models I_1, I_2 and I_3 have been presented by doing experiments and by analyzing the results of eight color scenes by [19] The color scenes were digitalized with 256*256 spatial resolution and 6-bit intensity resolution for each R, G, and B. Ohta et al. calculated the eigenvectors of for each of the eight color scenes. From the analysis, three orthogonal color models were derived: $I1 = (R + G + B)$, $I2 = (R - G)/2$ and $I3 = (2G - R - B)/4$. Note that $I1$ corresponds to intensity.

- COLOR MODELS:

  $I1$, $I2$, $I3$;
- TRANSFORMATION:

$$I1 = \frac{(R + G + B)}{3}, \tag{73}$$

$$I2 = \frac{(R - B)}{2}, \tag{74}$$

$$I3 = \frac{(2G - R - B)}{2}, \tag{75}$$

- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - Linear
  - Intuitive
  - Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination
- REMARKS: Uncorrelation based on the Karhunen-Loève transformation of eight different color images.

6.9 YIQ and YUV

The National Television Systems Committee (N.T.S.C.) developed the three color attributes Y, I, and Q for transmission efficiency. The tristimulus value Y corresponds to the luminance of a color. I and Q correspond closely the hue and saturation of a color. By reducing the spatial bandwidth of I and Q without noticeable image degradation, efficient color transmittion is obtained. The PAL and SECAM standards used in Europe, the Y, U, and V tristimulus values are used. The I and Q color attributes are related to U and V by a simple rotation of the color coordinates in color space.

- COLOR MODELS: Y, I, Q;
- TRANSFORMATION:

$$Y = 0.299R + 0.587G + 0.114B \tag{76}$$

$$I = 0.596R - 0.274G - 0.312B \tag{77}$$

$$Q = 0.211R - 0.523G + 0.312B \tag{78}$$

- CHARACTERISTICS:
  - Device independent
  - Not perceptual uniform
  - Linear transformation
  - Not intuitive
  - Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination
- REMARKS: Y is the luminance of a color.

6.10 HSI Color System

The human color perception is conveniently represented by the following set of color features: I(ntensity), S(aturation), and H(ue). I is an attribute in terms of which a light or surface color may be ordered on a scale from dim to bright. S denotes the relative white content of a color and H is the color aspect of a visual impression.

The problem of hue is that it becomes unstable when S is near zero due to the non-removable singularities in the nonlinear transformation, which a small perturbation of the input can cause a large jump in the transformed values.

- COLOR MODELS: I, H, S;
- TRANSFORMATION:

$$I = \frac{(R+G+B)}{3}, \tag{79}$$

$$H(R,G,B) = \arctan(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}) \tag{80}$$

$$S = 1 - 3\min(r,g,b); \tag{81}$$

- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - I is linear. Saturation is nonlinear: becomes unstable when intensity is near zero. Hue is nonlinear: becomes unstable when intensity and saturation are near zero
  - Intuitive
  - Intensity I: dependent on viewing direction, object geometry, direction of the illumination, intensity and color of the illumination.
  - Saturation S: dependent on highlights and a change in the color of the illumination.
  - Hue H: dependent on the color of the illumination.

6.11 Color Ratio's

Color constant color ratio's have been proposed by Funt and Finlayson [8], Nayar and Bolle [17] and Gevers and Smeulders [9]. These color constant models are based on the ratio of surface albedos rather then the recovering of the actual surface albedo itself. The constraint is that the illumination can be modeled by the coefficient rule. The coefficient model of illumination change holds exactly in the case of narrow-band sensors. Although standard video camera's are not equipped with narrow-band filters, spectral sharpening could be applied [4] to achieve this to a large extend.

- COLOR MODELS: The color ratio's proposed by Nayar and Bolle are given by [17]:

$$N(C^{\vec{x}_1}, C^{\vec{x}_2}) \tag{82}$$

Funt and Finlayson by [8]:

$$F(C^{\vec{x}_1}, C^{\vec{x}_2}) \tag{83}$$

expressing color ratio's between two neighboring image locations, for $C \in \{R, G, B\}$, where $\vec{x}_1$ and $\vec{x}_2$ denote the image locations of the two neighboring pixels.

The color ratio's of Gevers and Smeulders [9] are given by:

$$m(C_1^{\vec{x}_1}, C_1^{\vec{x}_2}, C_2^{\vec{x}_1}, C_2^{\vec{x}_2}) \tag{84}$$

expressing the color ratio between two neighboring image locations, for $C_1, C_2 \in \{R, G, B\}$ where $\vec{x}_1$ and $\vec{x}_2$ denote the image locations of the two neighboring pixels.

- TRANSFORMATION:

  Nayar and Bolle is given by [17]:

$$N(C^{\vec{x}_1}, C^{\vec{x}_2}) = \frac{C^{\vec{x}_1} - C^{\vec{x}_2}}{C^{\vec{x}_2} + C^{\vec{x}_1}} \tag{85}$$

  Funt and Finlayson by [8]:

$$F(C^{\vec{x}_1}, C^{\vec{x}_2}) = \frac{C^{\vec{x}_1}}{C^{\vec{x}_2}} \tag{86}$$

  and Gevers and Smeulders [9]:

$$m(C_1^{\vec{x}_1}, C_1^{\vec{x}_2}, C_2^{\vec{x}_1}, C_2^{\vec{x}_2}) = \frac{C_1^{\vec{x}_1} C_2^{\vec{x}_2}}{C_1^{\vec{x}_2} C_2^{\vec{x}_1}} \tag{87}$$

- CHARACTERISTICS:
  - Device dependent
  - Not perceptual uniform
  - Nonlinear: Becomes unstable when intensity is near zero.
  - Not intuitive
  - N and F: dependent on the object geometry. m no dependencies.

7 Color and Image Search Engines

Very large digital image archives have been created and used in a number of applications including archives of images of postal stamps, textile patterns, museum objects, trademarks and logos, and views from everyday life as it appears in home videos and consumer photography. Moreover, with the growth and popularity of the World Wide Web, a tremendous amount of visual information is made accessible publicly. As a consequence, there is a growing demand for search methods retrieving pictorial entities from large image archives. Attempts have been made to develop general purpose image retrieval systems based on multiple features (e.g. color, shape and texture) describing the image content [1], [2], [5], [20], [27], for example. Further, a number of systems are available for retrieving images from the World Wide Web, for example [21], [22], [23], [26]. Aside from different representation schemes, these systems retrieve images on the basis color. The Picasso [3] and ImageRover [21] system the $L^*u^*v^*$ color space has been used for image retrieval. The QBIC system [5] evaluates similarity of global color properties using histograms based on a linear combination of the RGB. MARS [15] uses the the $L^*a^*b^*$ color space because the color space consists of perceptually uniform colors, which better matches the human perception of color. The PicToSeek system [10] is based on color models robust to a change in viewing direction, object geometry and illumination.

We have seen that each color system has its own characteristics. A number of systems are linear combinations of the R, G and B values, such as the XYZ and the $I1I2I3$ color system, or normalized with respect to intensity, such as the rgb and the xyz color system. The $U^*V^*W^*$

and the $L^*a^*b^*$ color systems have distances which reflect the perceived similarity. Each image retrieval application demands a specific color system. In Figure 7, a taxonomy is given of the different color systems. Further, the various color systems and their performance can be experienced within the PicToSeek and Pic2Seek systems on-line at:
http://www.wins.uva.nl/research/isis/zomax/.

Color system	Device indep.	Perc. Uniform	Linear	Intuitive	View point	Object shape	Highlights	Illum. Intensity	Illum. SPD
RGB	-	-	+	-	-	-	-	-	-
XYZ	+	-	+	-	-	-	-	-	-
Norm. rgb	-	-	-	-	+	+	-	+	-
Norm. xyz	+	-	-	-	+	+	-	+	-
L*a*b*	+	+	-	-	-	-	-	-	-
U*V*W*	+	+	-	-	-	-	-	-	-
I1I2I3	-	-	+	-	-	-	-	-	-
YIQ	-	-	+	-	-	-	-	-	-
YUV	-	-	+	-	-	-	-	-	-
Intensity	-	-	+	+	-	-	-	-	-
Hue	-	-	-	+	+	+	+	+	-
Saturation	-	-	-	+	+	+	-	+	-
F, N	-	-	-	-	+	-	-	+	+
M	-	-	-	-	+	+	-	+	+

Figure 7: *a. Overview of the dependencies differentiated for the various color systems. + denotes that the condition is satisfied - denotes that the condition is not satisfied.*

8 Conclusion

In this chapter, a survey on the basics of color has been given. Further, color models and ordering systems were discussed, and the state-of-the-art on color invariance has been presented. For the purpose of color-based image retrieval, a taxonomy on color systems has been provided. The color system taxonomy can be used to select the proper color system for a specific application.

References

[1] Proceedings of First International Workshop on Image Databases and Multi Media Search, IDB-MMS '96, Amsterdam, The Netherlands, 1996.

[2] Proceedings of IEEE Workshop on Content-based Access and Video Libraries, CVPR, 1997.

[3] A. Del Bimbo, M. Mugnaini, P. Pala, F. Turco, *Visual Querying by Color Perceptive Regions*, Pattern recognition, Vol. 31, No. 9, pp. 1241-1253, 1998.

[4] Finlayson, G.D., Drew, M.S., and Funt, B.V., *Spectral Sharpening: Sensor Transformation for Improved Color Constancy*, JOSA, 11, pp. 1553-1563, May, 1994.

[5] Flickner, M. et al, *Query by Image and Video Content: the QBIC system*, IEEE Computer, Vol. 28, No. 9, 1995.

[6] Forsyth, D., *A Novel Algorithm for Color Constancy*, International Journal of Computer Vision, Vol. 5, pp. 5-36, 1990.

[7] Funt, B. V. and Drew, M. S., *Color Constancy Computation in Near-Mondrian Scenes*, In Proceedings of the CVPR, IEEE Computer Society Press, pp. 544-549, 1988.

[8] Funt, B. V. and Finlayson, G. D., *Color Constant Color Indexing*, IEEE PAMI, 17(5), pp. 522-529, 1995.

[9] T. Gevers and Arnold W.M. Smeulders, *Content-based Image Retrieval by Viewpoint-invariant Image Indexing*, Image and Vision Computing, (17)7, 1999.

[10] T. Gevers and Arnold W.M. Smeulders, *Color Based Object Recognition*, Pattern Recognition, 32, pp. 453-464, March, 1999.

[11] Goethe, *Farbenlehre*, 1840.

[12] Kender, J. R., *Saturation, Hue, and Normalized Colors: Calculation, Digitization Effects, and Use*, Technical Report, Department of Computer Science, Carnegie-Mellon University, 1976.

[13] E. H. Land, The Retinex Theory of Color Vision, Scientific Aamerican, 218(6), pp. 108-128, 1977.

[14] Levkowitz, H. and Herman G. T., *GLHS: A Generalized Lightness, Hue, and Saturation Color Model*, CVGIP: Graphical Models and Image Processing, Vol. 55, No. 4, pp. 271-285, 1993.

[15] S. Servetto, Y. Rui, K. Ramchandran, and T. S. Huang, *A Region-Based Representation of Images in MARS,* Journal on VLSI Signal Processing Systems, Volume 20, Issues 2, pp. 137-150, October 1998.

[16] A. H. Munsell, A Color Notation, 1st edn (Geo. H. Ellis Co., Boston); 11th edn (edited and rearranged)(Munsell Color Company, Baltimore, 1946); Now available from the Munsell Color Group, New Windsor, New York; 1905.

[17] S. K. Nayar, and R. M. Bolle, Reflectance Based Object Recognition, International Journal of Computer Vision, Vol. 17, No. 3, pp. 219-240, 1996

[18] I. Newton, Opticks, (w. Innys, London; reprinted: Dover, New York, 1952), 1704.

[19] Otha, Y., *Knowledge-Based Interpretation of Outdoor Natural Scenes*, Pitman Publishing, London, 1985.

[20] Pentland, A., Picard, R. W. and Sclaroff, S., *Photobook: Tools for Content-based Manipulation of Image Databases*, In Proceedings of Storage and Retrieval for Image and Video Databases II, 2, 185, SPIE, Bellingham, Wash. pp. 34-47, 1994

[21] Sclaroff, S., Taycher, L., La Cascia, M., *ImageRover: A Content-based Image Browser for the World Wide Web*, In Proceedings of IEEE Workshop on Content-based Access and Video Libraries, 1997.

[22] C. Frankel, M. Swain and Athitsos Webseer: An Image Search Engine for the World Wide Web, *TR-96-14*, U. Chicago, 1996.

[23] Smith, J. R. and Chang S.-F., *Visualseek: a Fully Automated Content-based Image Query System*, In Proceedings of ACM Multimedia 1996, 1996.

[24] Shafer, S. A., *Using Color to Separate Reflection Components*, COLOR Res. Appl., 10(4), pp 210-218, 1985.

[25] Swain, M. J. and Ballard, D. H., *Color Indexing*, International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32, 1991.

[26] Gupta, A., *Visual Information Retrieval Technology: A Virage Perspective*, TR 3A, Virage Inc., 1996.

[27] Proceedings of Visual97 Information Systems: The Second International Conference on Visual Information Systems, San Diego, USA, 1997.

[28] Proceedings of Visual99 Information Systems: The Third International Conference on Visual Information Systems, Amsterdam, The Netherlands, 1999.