Chapter 5

Perceiving Color

We interact with color all the time. It is a part and parcel of our lives, so much so, that we probably cannot appreciate it unless we lose our perception of color. Mr. I, who lost color perception due to an accident exclaimed with anguish, “My dog looks gray, tomato juice is black and color TV is an hodge podge”. Color not only adds beauty to our life, but serves important signalling functions. Natural world provide us with many signals to identify and classify objects. Many of these come in terms of color. For e.g. banana turns yellow when its ripe, the sky turns red when it is dawn and so on. Not only so, color plays an important role in our ability of perceptual organization where we group objects together or tell one object apart from another.

5.1 The Color Stimuli

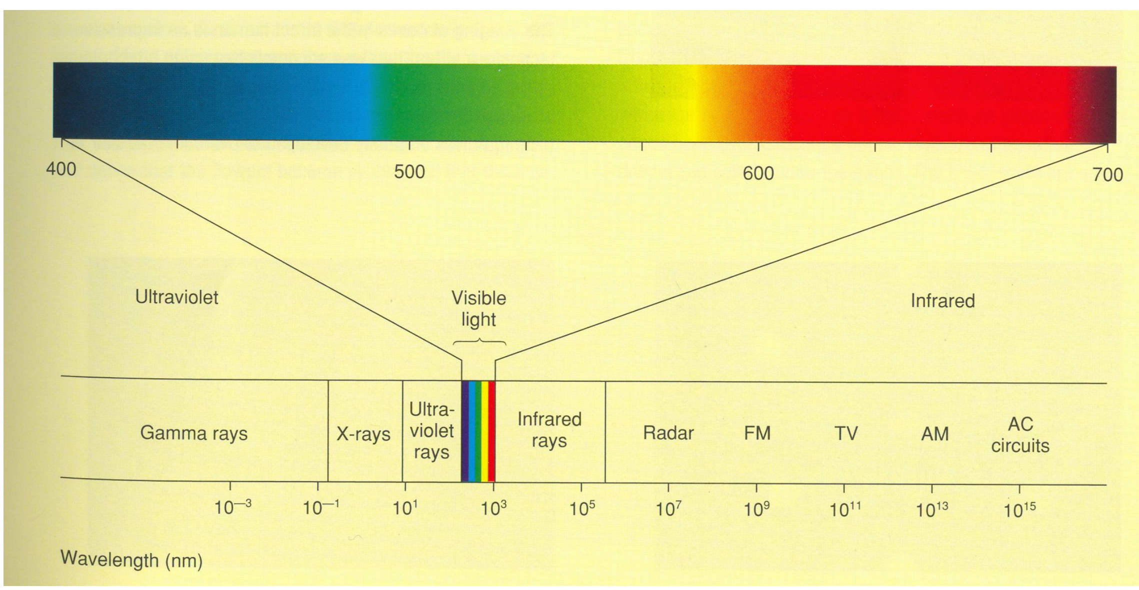

Figure 5.1: The electromagnetic spectrum of light.

This deals with how do we perceive color in a image. Newton explained light as tiny packets of energy called *photon*. These are actually packets of electromagnetic energy which has an associated wavelength. This wavelength is important for color vision. We have seen that the spectrum has wavelength that varies between 400 nm and 700 nm. Figure 5.1 shows the visible spectrum of colors. We have also seen that color is due to selective emission/transmission/reflection of certain wavelengths more than other. If all different wavelengths of light were emitted/transmitted/reflected equally, we would only see a gray 'Ganzfeld'. In fact, selective reflection of wavelengths by different objects is the not the only thing that is responsible for color vision. The second thing that is equally important is the eye's response to different wavelengths. This response can be different from species to species, and also shows a variance across individuals of same species. So, color perception is different for different species, even for individuals. That is the reason, color is often considered as an perception, and not reality!

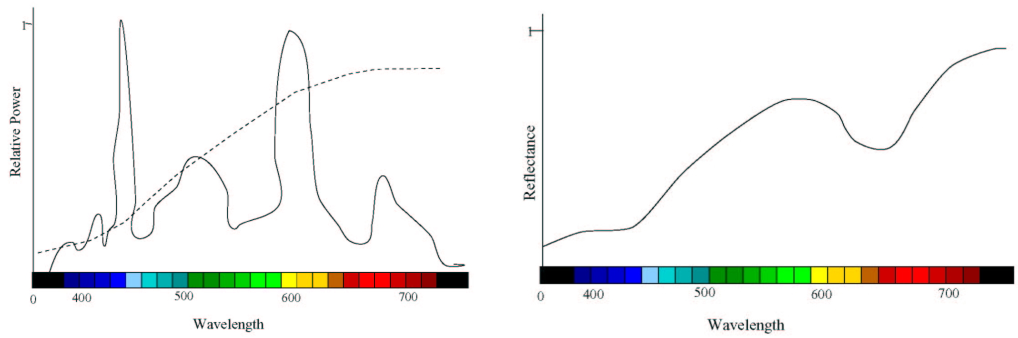

Figure 5.2: Left: The illumination spectrum ($I(\lambda)$) of a flouroscent and tungsten lamp. Right: The reflectance spectrum ($R(\lambda)$) of a red apple.

We will start from the illumination of a scene. A scene is usually lighted by some light source. This light source emits light differently at different wavelengths λ. This function, denoted by $I(\lambda)$ gives the illumination spectrum. Similarly, for a object, its relative reflectance at different wavelengths define its reflectance spectrum $R(\lambda)$. These spectra for a a couple of light sources and a red apple is illustrated in Figure 5.2.

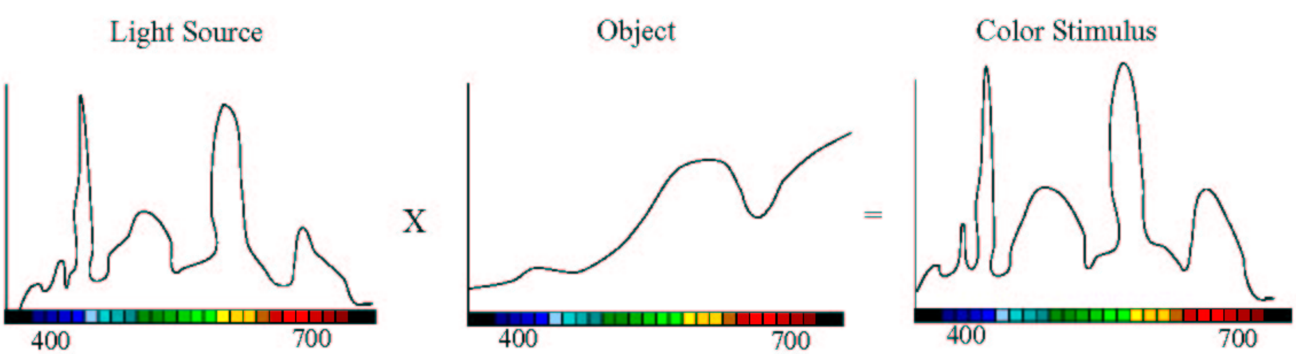

Figure 5.3: The product of the illumination and the reflectance generated the color stimuli.

When an objected is illuminated by a light source, the amount of light that is reflected from that object at different wavelength is given by the product of $I(\lambda)$ and $R(\lambda)$. Since this is the spectrum that stimulates the vision, this is called the *color stimuli*, denoted by $C(\lambda)$. Thus,

$$C(\lambda) = I(\lambda) \times R(\lambda) \tag{5.1}$$

as illustrated in Figure 5.3.

Color stimuli can be of different types as shown in Figure 5.4. When it has light of only one wavelength, it is called *monochromatic*, e.g. a laser beam. When relative amount of light from all wavelengths is equal, then it is

Figure 5.4: Different types of color stimuli: monochromatic (left), achormatic (middle) and polychromatic (right). Each spectrum is represented in the color that it generates.

called *achromatic*. The sunlight is close to achromatic in the day time. Finally, if the stimuli has different amounts of light from different wavelengths, it is called *polychromatic*.

While defining color formally, we need to define a few properties. These are *brightness/intensity*, *luminance*, *hue* and *saturation*. The hue and saturation together is called the chrominance of a color. The brightness/intensity of the color is the amount of light energy contained in the spectrum of the color and can be estimates as the area under the spectrum. The luminance is defined as the 'perceived brightness' and can be defined as the multiplication of the color spectrum with the luminous sensitivity function of the eye. The chrominance gives us the sense of colorfulness. The hue is defined by a *dominant wavelength* which can be thought of as the weighted mean of all the wavelengths weighted by their relative amount in the spectrum. This tells that the color produces a sensation of hue similar to the hue of its dominant wavelength. The saturation gives color its vibrancy and is the amount of achromatic color present in a color. This can be thought of as the variance of the spectrum from the dominant wavelength. For e.g pink has a dominant wavelength of red. But, it is a unsaturated version of red since white is combined with red to get pink. However, the saturation of a vibrant pink is more than that of light baby pink. Another concept that is useful to introduce here is that of *lightness*. This is the relative amount of light reflected. For example, a black object reflects 10% of all wavelengths. Now, if it is in sunlight, which has intensity of $100\ cd/m^2$, then the black object will reflect $10\ cd/m^2$. However, in a dimly lit room with illumination of $10\ cd/m^2$, it will reflect $1\ cd/m^2$. Note that the brightness of the object has changed, but the lightness has remained same. Because of this lightness, the black object continues to look black in both bright and dim lights, and does not change to gray in bright light. For example, the plots we have showed so far has relative power in their y axis. So basically, they are providing information about the lightness. When we know the absolute brightness present, then we will be able to fix the scale factor.

5.1.1 Color Mixtures

Having considered the visual appearance of both achromatic and chromatic colors, now we turn towards theory of combination of multiple colors or color mixtures. In this chapter, we are going to discuss additive color mixtures which is how our eye combines colors. However, we will see while studying color reproduction that subtractive color systems (in case of media like inks or dyes) is also an significantly important way of combining colors.

Additive Mixture of Colors

Colors are mixed in the eye in a fashion in which bands of wavelengths are added to each other. This is called *additive mixture of colors*. Thus, the wavelengths present in the spectrum of the color formed by superposition of multiple colors is given by their *union* the wavelengths present in the spectrums which are being added. However, the response of the new stimuli at each wavelength is given by a simple addition of responses.

Displays produce color in an additive fashion. The surface of a color display is made up of hundreds of tiny dots of phosphor. Phosphors are compounds that emit light when bombarded with electrons and the amount of light

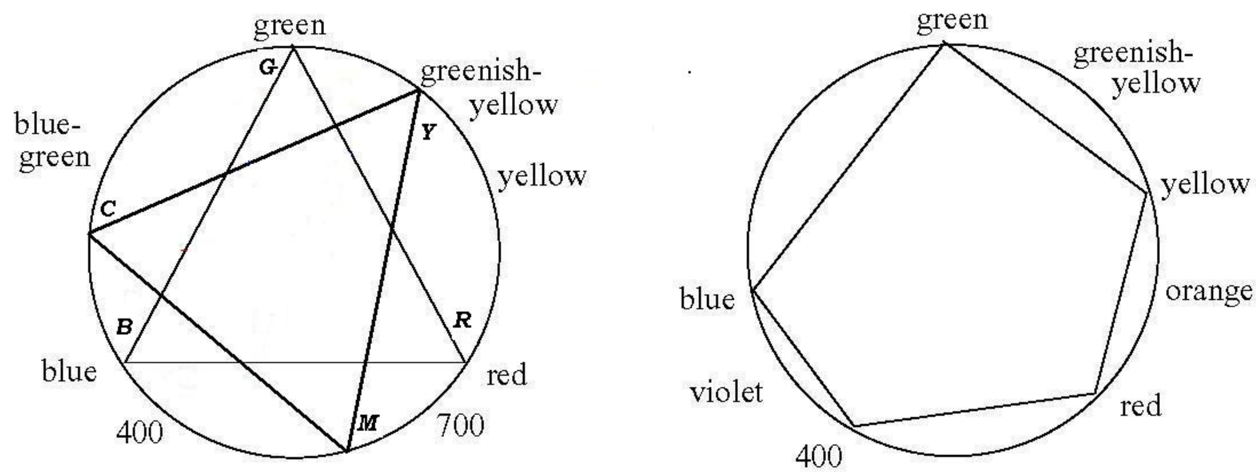

Figure 5.5: Newton's color wheel and additive color mixtures.

given off depends on the strength of the electron beam. The phosphors on the screen are in groups of three, with one phosphor emitting longer wavelengths (red), one emitting the middle wavelengths (green) and one emitting the shorter wavelength (blue). All the three phosphors together produce a very broad band containing all of the visible wavelengths. Varying the intensity levels of the phosphors produce the different levels of lightness. Thus *red*, *green* and *blue* are called the primaries for the additive mixture of colors.

We are going to deal with additive color systems since they are easier to model. The additive color system was first found by Newton by his famous color wheel experiment. He painted a circular wheel with all the colors on it as shown in Figure 5.5. Then he rotated the wheel very fast and found that the colors disappear and only a sense of white remains. From this he found that if we add any two color on this wheel, we can reproduce a third color which is in the wheel. So Newton's color wheel (Figure 5.5) was designed. The periphery represents all the monochromatic saturated colors and the interior represents saturated versions of these color. Newton showed that a color formed by adding two colors lies on the straight line joining the two colors on the color wheel. Where it lies on this straight line is given the proportion of the intensity of each color used. Thus, with just two colors, we can only reproduce colors on a straight line. However, if we would like to reproduce a reasonable portion of this wheel, we need to have at least three primaries, as shown by the triangle RGB in Figure 5.5. That is the reason, as we will see, most visual sensors (both manmade like cameras or natural like our eye) uses three primaries to reproduce all different colors.

Also note that different sets of primaries can be chosen, as shown by the triangle CYM in Figure 5.5. And the same color can be reproduced by a different combination of this different set of primaries. Note that the two vectors of the triangle formed by a set of primaries acts like a coordinate system within which the coordinates of any color indicates the proportions of the three colors that should be mixed to generate the color. When the set of primaries are changed, a new coordinate system is generated where the proportion of the primaries needed to generate the color changes and hence a change in its coordinates also. However, since both the coordinate systems lie on a plane, they can be transformed to each other by simple linear transformations. Hence, a set of primaries can be transformed *linearly* to generate a different set of primaries. The same transformation can be used to generate a transformation of the coordinates of a color from one to another.

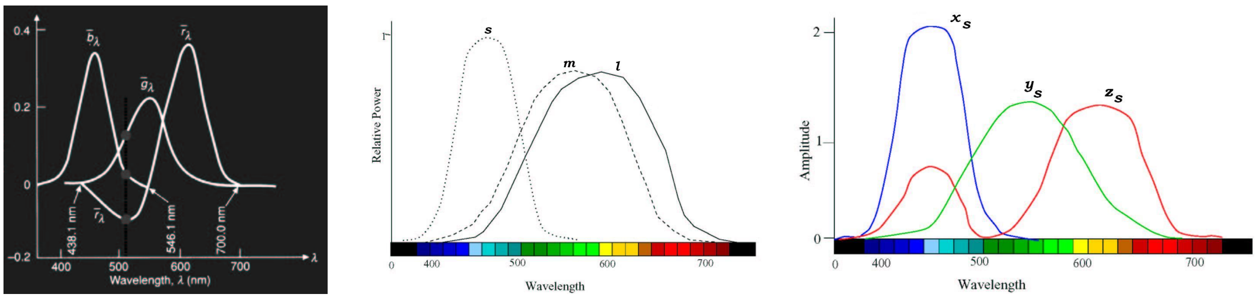

Figure 5.6: Left: The color matching functions, $\bar{r}$, $\bar{g}$ and $\bar{b}$ for Helmholtz's experiment. The primaries used were of wavelengths 438, 546 and 700 nm respectively. Middle: The response of the three different cones of the humans given by $\bar{x}$, $\bar{y}$ and $\bar{z}$. Right: The CIE color matching functions for a standard observer.

5.2 Trichromatic Theory in the Eye

This idea that colors can be reproduced by three primaries is true for our eyes also and was proved beyond doubt by Helmholtz's famous *color matching experiment*. In this experiment, the subjects were allowed to observe a screen that was divided in two parts, A and B. On part A, a monochromatic light was projected. On part B, three different primaries, red , green and blue, were projected. The subjects were given three dials to adjust the levels of the three different primaries in B such that a color matching the monochromatic color in A is produced. This process was repeated for every wavelength in the visible spectrum. Of course, there were wavelengths that cannot be matched by adjusting the primaries at all. In that case, some amounts of the primaries were added in part A itself and it was registered as negative amount of the primary required in part B to match the color of that wavelength. Three curves were obtained, one for each primary, by plotting the contribution of the primary in creating the color of a particular wavelength. These three curves, denoted by $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$are called the color matching functions (CMF) and is shown in Figure 5.6.

Note that there is a different way to look at the color matching functions. Each of them can be thought of as the response of a sensor to a particular wavelength of light. And the total response of three such sensors can create the perception of the color of that wavelength. This is exactly what happens in the eye. Eye has three different types of cones who respond differently to different wavelengths and the sum of these three responses create the perception of color. The response of these three cones are denoted by $s(\lambda)$, $m(\lambda)$ and $l(\lambda)$. Note however that when talking about sensors, negative values do not make sense. So, the response of the different cones does not have any negative regions. In fact, Figure 5.6 shows the response of the three sensors in the human eye. This fact that humans have three sensors to create the sensation of color was first proposed by Helmholtz and is called the *trichromatic theory*. Later, physiological studies revealed the response curves shown in Figure 5.6. Thus the response of these cones to a color stimuli $C(\lambda)$ is given by

$$S(\lambda) = C(\lambda)s(\lambda)$$

$$M(\lambda) = C(\lambda)m(\lambda)$$

$$L(\lambda) = C(\lambda)l(\lambda)$$

.

Note that one important side effect of the trichromatic theory of color is what we call *metamerism*. Note that since the sensors respond selectively to different wavelength, two different stimuli can produce similar response. This phenomenon is called *metamerism* and the two stimuli which produce similar response is called the *metamers*. So while reproducing colors in a device or sensors, producing metamers of actual color stimuli is sufficient. Reproducing the exact stimuli is not required.

5.2.1 XYZ Color Space

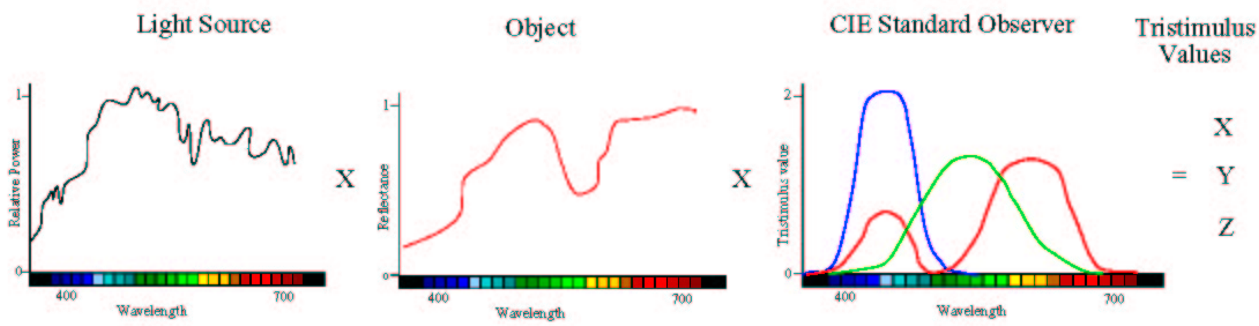

Figure 5.7: Tristimulus Values

Our perception of color involves properties like luminance, brightness, hue and saturation which cannot be directly related to the response of the three cones of the eye. In fact, to formally define all these perceptions and come up with the consistent perceptual organization of color, the CIE defined a set of *standard* color matching functions, $x_s(\lambda)$, $y_s(\lambda)$ and $z_s(\lambda)$. These are called the CIE color matching functions for standard observer. How these relate to the response of the three cones will follow shortly. But it is important to understand that all these sets of CMFs like $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$; $\bar{s}(\lambda)$, $\bar{m}(\lambda)$, and $\bar{l}(\lambda)$; and $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ can be related to each other by linear transformations. In fact, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ does not have any correspondence of using real color primaries. These were found mathematically from linear transformations of differ primaries (represented by 3×3 matrix) and designed to be always positive. These CIE color matching functions, $x_s(\lambda)$, $y_s(\lambda)$ and $z_s(\lambda)$ for a standard observer is shown in Figure 5.6.

Now, we will see the relationship between the cone sensitivity functions $\bar{s}(\lambda)$, $\bar{m}(\lambda)$, and $\bar{l}(\lambda)$ and the color matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$. Color matching means that the test monochromatic light and the comparable mixture of primaries generate identical response in the cones. Hence, the following are true.

$$\bar{l}_R\bar{r}(\lambda) + \bar{l}_G\bar{g}(\lambda) + \bar{l}_B\bar{b}(\lambda) = \bar{l}(\lambda)$$

$$\bar{m}_R\bar{r}(\lambda) + \bar{m}_G\bar{g}(\lambda) + \bar{m}_B\bar{b}(\lambda) = \bar{m}(\lambda)$$

$$\bar{s}_R\bar{r}(\lambda) + \bar{s}_G\bar{g}(\lambda) + \bar{s}_B\bar{b}(\lambda) = \bar{s}(\lambda)$$

where $\bar{l}_R, \bar{l}_G$ and $\bar{l}_B$ are, respectively, the L-cone sensitivities to the R, G and B primary lights; and similarly $\bar{m}_R, \bar{m}_G$ and $\bar{m}_B$ are the M-cone sensitivities to the primary lights; and $\bar{s}_R, \bar{s}_G$ and $\bar{s}_B$ are the S-cone sensitivities. We know $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and we assume for the red R primary, that $\bar{s}_R$ is effectively zero, since the S-cones are insensitive to long-wavelength lights. (The intensity of the spectral light λ, which is also known, is equal in energy units throughout the spectrum, and so is discounted from the above equations). There are, therefore, only eight unknowns required for the following linear transformation.

$$\begin{pmatrix} \bar{l}_R & \bar{l}_G & \bar{l}_B \\ \bar{m}_R & \bar{m}_G & \bar{m}_B \\ \bar{s}_R & \bar{s}_G & \bar{s}_B \end{pmatrix} \begin{pmatrix} \bar{r}(\lambda) \\ \bar{g}(\lambda) \\ \bar{b}(\lambda) \end{pmatrix} = \begin{pmatrix} \bar{l}(\lambda) \\ \bar{m}(\lambda) \\ \bar{s}(\lambda) \end{pmatrix}$$

However, because we are usually unconcerned about the absolute sizes of $\bar{s}(\lambda)$, $\bar{m}(\lambda)$, the eight unknowns collapse to just five as follows.

$$\begin{pmatrix} \frac{\bar{l}_R}{\bar{l}_B} & \frac{\bar{l}_G}{\bar{l}_B} & 1 \\ \frac{\bar{m}_R}{\bar{m}_B} & \frac{\bar{m}_G}{\bar{m}_B} & 1 \\ 0 & \frac{\bar{s}_G}{\bar{s}_B} & 1 \end{pmatrix} \begin{pmatrix} \bar{r}(\lambda) \\ \bar{g}(\lambda) \\ \bar{b}(\lambda) \end{pmatrix} = \begin{pmatrix} k_l\bar{l}(\lambda) \\ k_m\bar{m}(\lambda) \\ k_b\bar{s}(\lambda) \end{pmatrix}$$

where the absolute values of k_l, k_m, or k_b remain unknown, but are typically chosen to scale three functions in some way: for example, so that $k_s\bar{s}(\lambda)$, $k_m\bar{m}(\lambda)$, and $k_l\bar{l}(\lambda)$ peak at unity. In the well-known solution of the above equation by Smith & Pokorny (1975), $k_m\bar{m}(\lambda) + k_l\bar{l}(\lambda)$ sum to $\bar{y}(\lambda)$ of the CIE standard observer color matching functions. This is called the luminosity function and gives the luminance (perceived brightness) of a color. Note that this defines one of the linear function using which $\bar{s}(\lambda)$, $\bar{m}(\lambda)$, and $\bar{l}(\lambda)$ can be used to derive $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$.

The tristimulus values of a color stimuli is defined by the tristimulus values X, Y and Z where

$$X = \int_\lambda C(\lambda)x_s(\lambda) = \sum_{\lambda=400}^{700} C(\lambda)x_s(\lambda)$$

$$Y = \int_\lambda C(\lambda)y_s(\lambda) = \sum_{\lambda=400}^{700} C(\lambda)y_s(\lambda)$$

$$Z = \int_\lambda C(\lambda)z_s(\lambda) = \sum_{\lambda=400}^{700} C(\lambda)z_s(\lambda)$$

X, Y and Z are called the tristimulus values of a color stimulus C and defines a three dimensional space within which different colors can be organized. This is illustrated in Figure 5.7. Note that metameric colors have same tristimulus values. Also, the real colors will occupy only a subset of the XYZ space since it is defined by the imaginary primaries. The volume thus spanned by the real colors in the XYZ space is shown in Figure 5.8. This space makes is easy to compare or operate with colors since we do not need to deal with the spectrums but can operate in the 3D XYZ space. For example, to find the color C_n resulting from adding two colors, $C_1 = (X_1, Y_1, Z_1)$ and $C_2 = (X_2, Y_2, Z_2)$, we can just add their tristimulus values to get

$$C = (X_n, Y_n, Z_n) = (X_1 + X2, Y_1 + Y_2, Z_1 + Z_2). \tag{5.2}$$

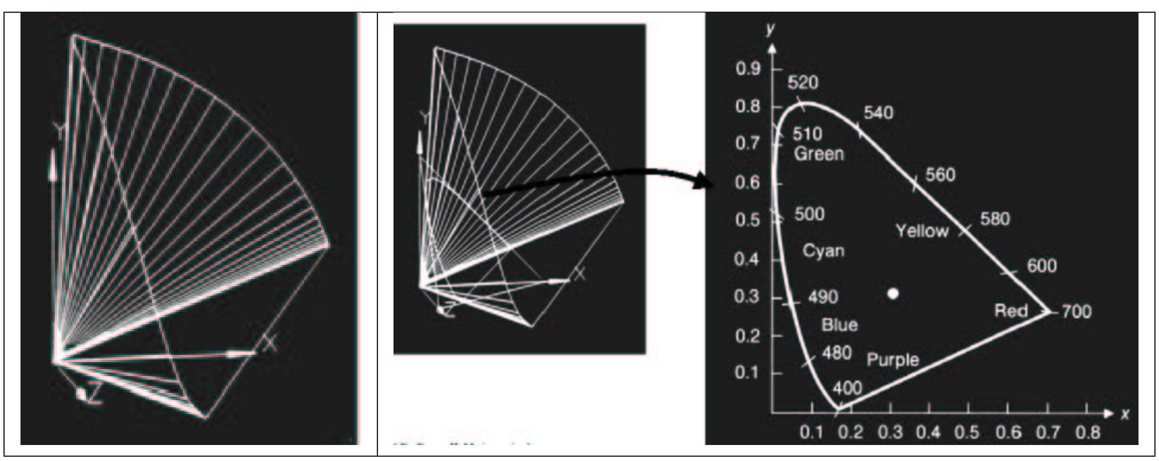

Figure 5.8: Left: The volume spanned by the real colors in the XYZ space. Right: Generation of the chromaticity diagram from the XYZ space.

But there is one problem with the XYZ space. It does not give us the feel for the perceptually important properties of color like luminance, hue and saturation. For example, one cannot get answers for the questions, how are colors of similar hue arranged in this space? Or, How are saturated colors arranged in this space? So, let us see how we can get a feel for all these properties of colors that help in our perceptual organization from the XYZ space.

For this the 3D space defining the real colors are projected on a 2D plane given by $X + Y + Z = \alpha$, to define the chromaticity coordinates (x, y) of a color as

$$x = \frac{X}{X+Y+Z}; \quad y = \frac{Y}{X+Y+Z} \tag{5.3}$$

This is shown in Figure 5.8. *Brightness* is related to the area under the stimuli spectrum. Hence $X + Y + Z$ provides a good approximation of the absolute brightness of a color. Since (X, Y, Z) and (kX, kY, kZ) generate the same chromaticity coordinates (Equation 5.3), a vector originating from 0 of the XYZ space has colors with same chrominance but with different brightness. So, the 2D projection into the xy plane essentially phases out the brightness mapping colors with same brightness at the same chromaticity coordinates. The chromaticity coordinates, x and y, can range from 0.0 to 1.0. However, all values of x and y again do not represent real colors. Only a subset enclosed by the the horse shoe shaped curve represent real colors. The real colors span a horse shoe shaped area called the *chromaticity diagram*.

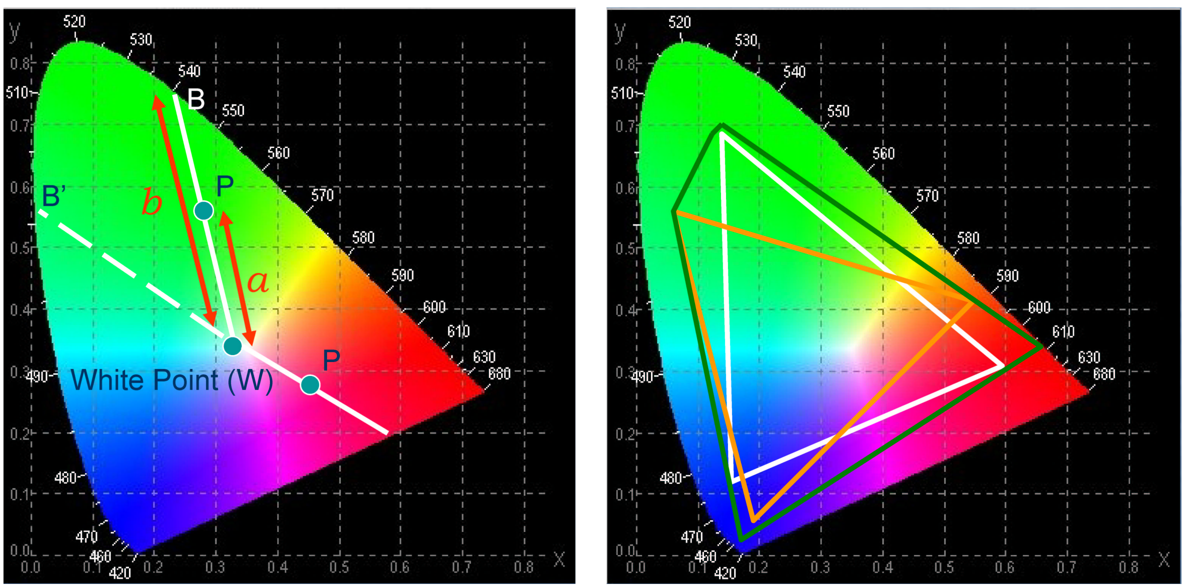

Figure 5.9: The Chromaticity Diagram. Left: Showing how colors of different hue and saturation are arranged. Right: Color Gamuts

Now we will study the chromaticity diagram to see how the colors or different hue and saturation are arranged on it. Figure 5.9 shows the arrangements of the different hues in both saturated and unsaturated form on the chromaticity diagram. An achromatic hue is generated when $X = Y = Z$ and hence $x = y = 0.33$ (Equation 5.3. Thus, the point $W = (0.33, 0.33)$ is called the white point on the chromaticity coordinate. All the monochromatic colors in the visible spectrum, called the spectral colors, form the boundary curve of the horse shoe region. The arrangement of the different wavelength, which in turn organizes the hues, are shown in Figure 5.9. However note that the line joining two ends of the curve represents purples which are not spectral colors. Note that our visible spectrum does not form a closed loop, but this one does. The purples can be thought of as the colors completing the loop. This, this boundary is called the non-spectral boundary of the chromaticity diagram. For any color P on a straight line joining W and any point on the boundary B has the same hue as the monochromatic color B. The saturation of the color increases as we come from B towards W. The wavelength of B is designated as the *dominant wavelength* of P and the measure of saturation is given by the ratio of $a = |P\bar{W}|$ and $b = |W\bar{B}|$. Thus,

B has a saturation of 1.0 as expected for a spectral color. If the straight line through W and P meets a non-spectral color, the dominant wavelength cannot be determined. Instead, this line is extended in the opposite direction. The wavelength of the spectral color B' that this extended line meets defines the *complementary* wavelength for P. Complementary wavelength defines a hue which when combined with the hue of the dominant wavelength of P will give achromatic color. One important point to note here is that the perceived brightness of a color, called luminance, is different from its brightness. It is given by the value Y. So, given Y, x and y, the tristimulus values of the color can be found uniquely using Equation 5.3. Now, we know how the luminance, hue and saturation can be related to the XYZ space. Thus, a more perceptual sense, color is often represented by (x, y, Y). Note that from these, the tristimulus values can be recovered as

$$X = \frac{xy}{Y}; \;\; Y = Y; \;\; Z = \frac{(1-x-y)y}{Y}$$

As we mentioned before, adding colors in the XYZ space is just adding their coordinates in this space. However, note that we still do not know how the different components like brightness, hue and saturation are interacting to create the luminance, hue and saturation of the new color. Let us take the case of generating the color $C_n = (X_n, Y_n, Z_n)$ by adding two colors $C_1 = (X_1, Y_1, Z_1)$ and $C_2 = (X_2, Y_2, Z_2)$, as in Equation 5.2. Note that the brightness of C_n, denoted by L_n is given by adding its XYZ coordinates as

$$\begin{aligned} B_n &= X_1 + Y_1 + Z_1 + X_2 + Y_2 + Z_2 && (5.4) \\ &= B_1 + B_2 && (5.5) \\ & && (5.6) \end{aligned}$$

This says that the brightness of the new color is the addition of the brightness of the colors that were added to produce it. In fact, this result can be easily generalized to more than two colors. Next, the chromaticity coordinates of C_n are

$$\begin{aligned} (x_n, y_n) &= (\frac{X_1 + X_2}{B_n}, \frac{Y_1 + Y_2}{B_n}) && (5.7) \\ &= (\frac{X_1}{B_n} + \frac{X_2}{B_n}, \frac{Y_1}{B_n} + \frac{Y_2}{B_n}) && (5.8) \\ &= (x_1\frac{B_1}{B_n} + x_2\frac{B_2}{B_n}, y_1\frac{B_1}{B_n} + y_2\frac{B_2}{B_n}) && (5.9) \\ &= (x_1, y_1)\frac{B_1}{B_n} + (x_2, y_2)\frac{B_2}{B_n} && (5.10) \\ & && (5.11) \end{aligned}$$

This shows that the chromaticity coordinates of the colors are added in proportion of the brightness to create the chromaticity coordinate of the new color. Now we know the relationship between how the different components of the colors are mixed to get the new color.

5.2.2 Phenomena Explained by Trichromatic Theory

There is one physiological phenomenon that we will study now that can be explained by the trichromatic theory. This is called *color deficiency*. People who can see all colors are called *trichromats*. There are many people who cannot see colors properly. In fact, color deficiency in humans can be categorized in three different classes.

1. *Monochromat:* This people can see no color. The world is always different shades of gray to them.

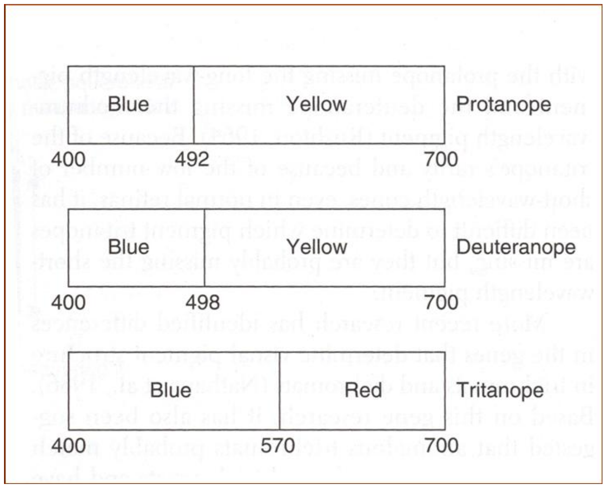

Figure 5.10: Different types of dichromats.

2. *Dichromat:* These people can see a few colors, not all. Their are three types, as explained in Figure 5.10. There color sensation with different wavelengths is shown in Figure 5.10. The protanopes can see the blue well. It starts from blue becoming grayer and grayer. Ultimately it becomes completely neutral at 492nm. Then the sensation of yellow predominates in the rest of the visible range. For the deuteranope, it is similar, but the neutral is reached at 498 nm. Finally, the tritanope can see only blue and red with the neutral at 570 nm.

3. *Color weakness:* Normal trichromats are also not all similar. Some have weaknesses for some region of the color space.

It was found that monochromats do not have any cones. Dichromats have only two types of cones: protanopes lack S cones, deuteranopes lack M and tritanope lacks L. Also color weaknesses are caused by reduced number of some kinds of cones or by reduced sensitivity of some kinds of cones. This proves that trichromatic theory can explain color blindness. However, note that another form of color deficiency has been found where people fail to see color even when their cones are all fine. This has been traced to damages in certain regions of visual cortex. This is called cerebral achromatopsia.

5.2.3 Phenomena Cannot be Explained by Trichromatic Theory

However, following are some phenomena that cannot be explained by trichromatic theory. These are as follows.

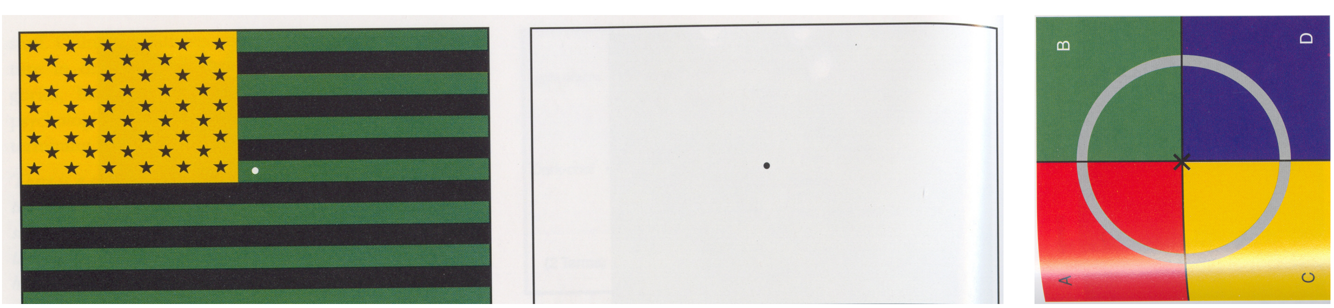

Figure 5.11: Left: Complementary After Image. Right: Simultaneous color contrast.

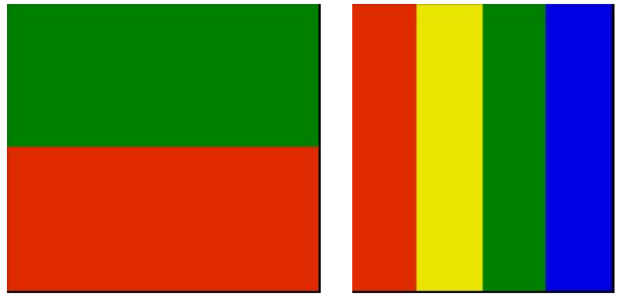

Figure 5.12: Chromatic adaptation.

1. Do the following experiment on Figure 5.11. Fix your gaze at the dot on the green-yellow flag. Keep looking at the green-yellow flag for 30 seconds or more. Then move your gaze to the white region by the side fixating on the black dot on the white box. You will be able to see a faint after image of the american flag. The green regions will be occupied by red, the yellow regions by blue and the black regions by white. This is called the complementary afterimage.

2. Notice the gray circle in Figure 5.11. The circle is of same gray but are surrounded by different regions by different colors. Note how this leads to tints in gray in different region. It is pinkish in green region, greenish in red region, bluish in yellow region and yellowish in blue region. This is called *simultaneous color contrast*.

3. Do the following experiment on the left pattern of Figure 5.12. Close your left eye. Stare at the pattern for 30 seconds or more with your right eye only. After 30 seconds, open your left eye. See how saturated the colors look from your left eye but washed out from your right eye. After your left eye comes back to normal, do the following experiment. Stare at the left pattern (now with both eyes open) for 30 seconds. Then shift your gaze to the right pattern. You will see that the top part of red bar looks more saturated than bottom and the bottom part of the green bar looks more saturated then the top. The yellow and blue appear more greenish in the bottom and reddish in the top. This is called *chromatic adaptation*. When the eye is exposed to a color for a long time, it loses some of its sensitivity towards that color.

4. Try visualizing mixing green and red. Though the color is yellow, often we feel comfortable thinking of yellow as the fourth primary color. More difficulty will be faced if you try to visualize mixing blue and yellow. While same is not true for red and blue or red and yellow.

5.3 Opponent Color Theory in Higher Visual Areas

All the above phenomena cannot be explained by the trichromatic theory. So another theory was proposed simultaneously with the trichromatic theory by Hering. This is called the opponent color theory. He proposed that there are six types of sensors in the eye as shown in Figure 5.13. The cells which are excited by red will be inhibited by green and vice versa. Similarly cells which are excited by blue and black will inhibit yellow and white respectively and vice versa. The sensitivities of these cells with different wavelengths is shown in Figure 5.13.

For many years there was struggle between the proponents of trichromatic theory and the opponent theory. Each can explain a few phenomenon, but not the other. Recently, in 1965, the issue was resolved by Valois and Valois. Cells were found in LGN which showed similar behavior as proposed by the opponent theory. These were called the opponent cells. This showed that both the theory existed by in different areas of the visual pathway.

The next question then is how the responses from S, M and L cones are connected to the opponent cells in the LGN. In fact, a theory for that was proposed and verified. Figure 5.14 shows the excitatory and inhibitory

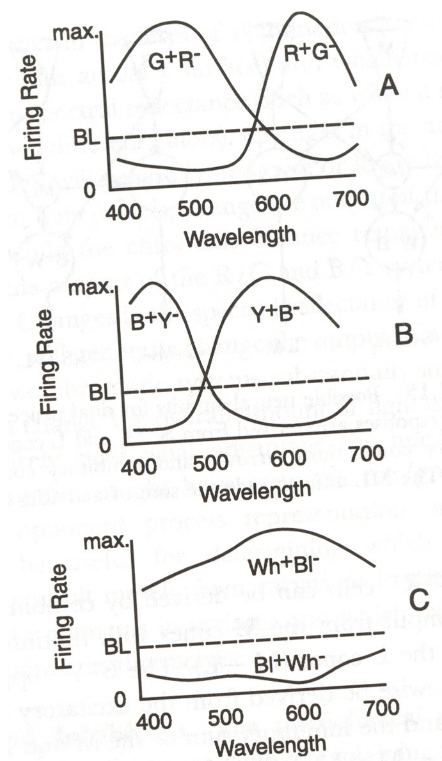

Figure 5.13: Six types of sensors proposed by opponent color theory.

connections from the cones to create the opponent characteristics of the opponent cells. This says that the three color-separated images are not transmitted directly to the brain. Instead, the neurons in the retina encode them as opponent signals. The output of all cones are summed weighted by their relative population size to produce an response that matches the combined cone response curve. Two weighted differences allows the construction of red-green (L-M+S) and the yellow-blue (L+M-S) signals. The naming is similar to lateral inhibition. In former, L is positive and M is negative and hence the nomenclature red-green. In the latter, both L and M (creates yellow) is positive and S is negative, and hence the name yellow-green. Similarly, complementary cells of green-red and blue-yellow have also present. These are illustrated in Figure 5.14(a). Conversion from LMS to these opponent signals serve to decorrelate the color information transmitted through the three channels reducing noise and hence increasing their efficiency.

5.3.1 Phenomena Explained by Opponent and Adaptation Theory

Now, we can explain complementary afterimages and chromatic adaptation by this opponent theory.

When we gaze at an image with red image, the R^+G^- adapts and fires less while the G^+R^- cells fire more. Hence the reds look unsaturated and greens saturated during chromatic adaptation. Same is true for blue and yellow. This imbalance causes a hue shift towards red or green in the yellow and the blue.

Similarly, when we have a prolonged exposure to green the G^+R^- cells are fire less. When we shift the gaze to neutral region the R^+G^- cells are firing relatively more and hence the red afterimage. Same is true for blue and yellow.

However, note that this still cannot explain the simultaneous color contrast. This was explained by the discovery of he double opponent cells in the visual cortex. These are opponent cells arranged in a center-surround fashion in the visual cortex. They have center excitatory and surround inhibitory receptive field as shown in Figure 5.14(b).

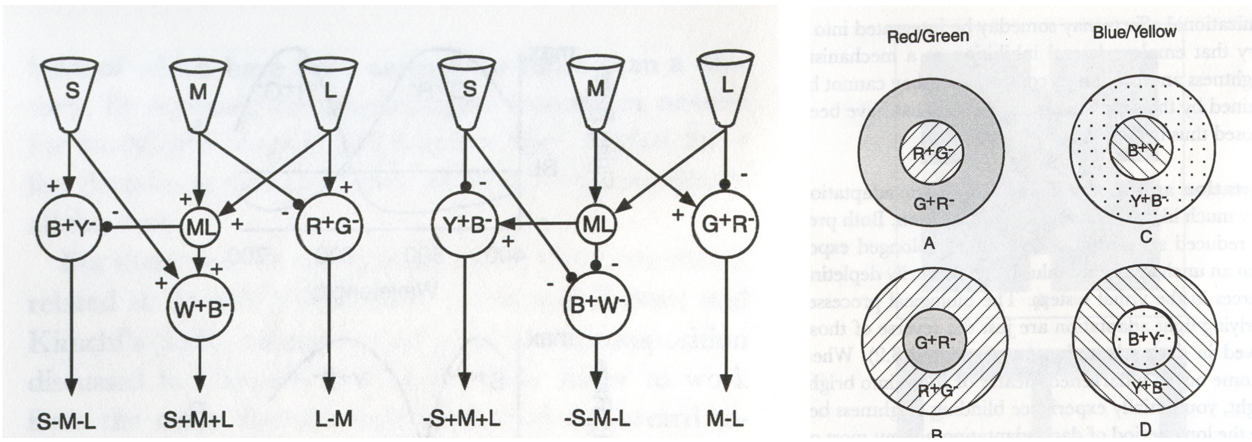

Figure 5.14: Left: Connection of the S, M and L cones to the opponent cells. Right: Receptive fields of the double opponent cells in the visual cortex.

These can explain the simultaneous color contrast just as lateral inhibition explained simultaneously contrast.

5.4 Adaptation Theories

However, the opponent theory alone still cannot explain the two important properties of visual system, called *lightness* and *color constancy*. Imagine you have a white paper with black letters. The black ink reflects only 10% of the photons and the white paper reflects 90%. In this case, let us assume that there are 100 photons falling on this page. Therefore, 10 photons will be reflected from the letters and 90 photons from the white paper. But now lets go out in the sun where $10,000$ photons are available. Now, the paper reflects 9000 photons and the ink reflects 1000 photons. Note that the black ink in the sunlight reflects more photons that the white paper indoor. But still the black ink does not cease to look black. This is called lightness constancy. How can we explain this?

Similarly, when we view objects under different illuminations conditions, their color perception seems to remain same. For example, a red object does not cease to be red if illuminated by sunlight instead of a tungsten lamp. This is color constancy.

This brings us to the discussion of adaptation theories. Adaptation is a dynamic mechanism that serve to optimize the visual response to the particular viewing environment in which we are immersed. Following are the different mechanisms of dark, light and chromatic adaptation.

Light Adaptation: Dark adaptation is the case of your eye adapting to darkness, for e.g. when entering a dark room on a bright sunny day and we have discussed this in details when studying the visual system before. Light adaptation is the inverse of this process, but it is important to consider is separately since its visual properties differ. When leaving a darkened room to walk into a bright sunny environment, the visual system should become less sensitive since much more visible energy is present. Though the mechanisms of this is similar to dark adaptation, there is an asymmetry in the time it takes. As opposed to 30 minutes of dark adaptation, light adaptation happens in the order of 5 minutes.

The capabilities provided by this is better illustrated in Figure 5.15. The visual system has a limited dynamic range of 100:1 available. However, the world we live in has a dynamic range that covers at least 10 orders of magnitude (starlit night to bright sunny afternoon). Fortunately, the whole of this huge range is usually not present at the same time. If a single response function was to be used for this huge dynamic range, the response curve will look like the dashed curve in Figure 5.15. But in this situation, the perceived contrast of the scene will be very low and also the visual sensitivity will be severely degraded due to low signal to noise ratio, especially for

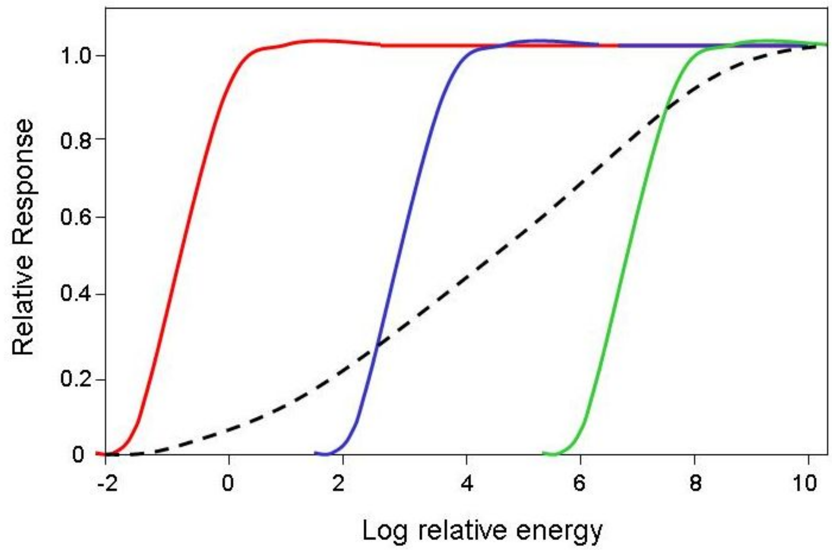

Figure 5.15: Illustration of light adaption by which a very large dynamic range can be mapped to relatively limited dynamic response of the eye. Solid curves show a family of adapted responses. The dotted curve how the response would look in the absence of any adaptation.

darker areas of the scene. But due to light adaptation, instead our eye responds in a fashion represented by family of solid curves. These curves map the relevant or useful visible range of illumination to the dynamic range offered by the eye. This results in the best possible visual perception for each situation. Thus, the visual response curve slides along the illumination level axis until the optimal viewing conditions for a given scene is reached. This is analogous to exposure control in a camera.

Chromatic Adaptation: Chromatic Adaptation can be observed by examining an white object under varying illumination conditions (like daylight, incandescent, and fluorescent). Daylight is much richer than fluorescent light in shorter wavelengths and fluorescent light is much richer than daylight in longer wavelengths. However, the white paper remains relatively similarly white in both these lighting conditions. This is because, the S cone system becomes less sensitive when adapting to day light and the L cones become less sensitive under fluorescent light. Thus chromatic adaptation is a closely related phenomenon which controls the spectral sensitivity of the three cones independently. This is illustrated in Figure 5.16 where the height of these curves can vary independently. While these curves represent the adaptation of the cones, they can also occur in the single and double opponent cells. Figure 5.17 provides a visual demonstration of this. Note that here the eye adapts to cyan and yellow both of which can only be produced by the opponent cells. This can be thought of to be similar to the white balance control of a camera.

However, note that all these theories put together cannot explain several things. Take the example of the image in Figure 5.19. If using only trichromatic or opponent theory, we would describe this image as the walls of a building that is white in color in some regions and gray in others. But instead, we perceive this as a building with white walls and dark shingles. The gray wall indicates a shadow and we are able to perceive it. Not only so, we are able to factor out the shadow and find out the true color of the building. In fact, even if we consider adaptation theory, some of the following cannot be explained.

- First, we know that adaptation takes some time, sometimes minutes. In that case, when we switch of the reading lamp, we would expect our eye to adapt to the reduced illumination. And the time when the eye is getting adapted we would see gray paper instead of the white paper. That does not happen.

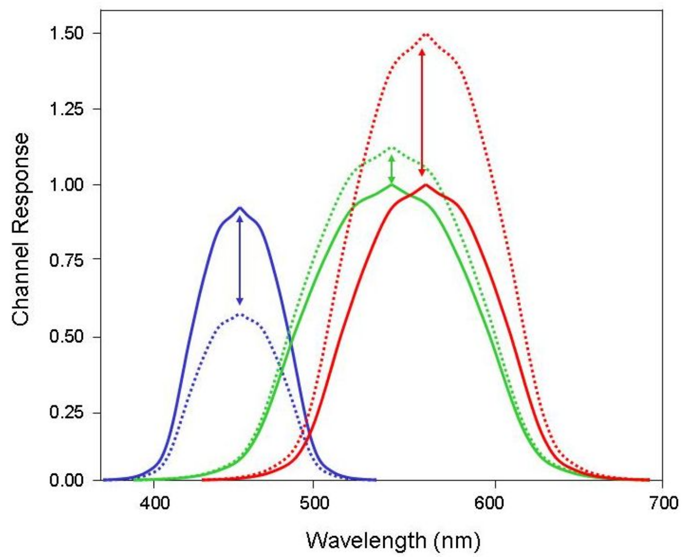

Figure 5.16: Illustration of chromatic adaptation as the independent sensitivity regulation of the three cone spectral responses.

- Second, even if we assume instantaneous adaptation, we should not be able to have a sense of the change in illumination after the eye adapts. Yet when the reading lamp is switched off, we know that this light is less than what was there before. But the paper is still white.

- Third, adaptation can only explain a situation where there is a change in illumination. Lightness constancy happens also with same illumination. For example, in Figure 5.18, we perceive part of the table in the shadow of the cup. We perceive the reflectance of that area to be same as that of the rest of the table. However, when a line is used to explicitly define the boundary of this shadow, we perceive that the part of the table has a different reflectance. No illumination change happened to cause this. This cannot be explained by adaptation theories.

5.5 Higher Visual Mechanisms

There are several important cognitive visual mechanisms that impact color appearance, like memory, object recognition, continuity perception and so on. Though not all of it is yet clearly understood, a large body of work have started the initial inroads to understand these processes. In this section, we are going to explore some of these and see some early models developed to explain such phenomena.

We know that the light that reaches the eye from any point on the surface is the product of its illumination and reflectance (Equation 5.1). However, we do not know the components of this product, but only the product itself. Also, there is no inherent difference in the representation of the reflectance and the illumination that can be exploited. Thus recovering the product from this is again an under-determined inverse problem. However, human beings can solve this with amazing efficiency and accuracy most of the time.

Helmholtz's Unconscious Reference Theory

The first theory to explain the separation of illumination from refelctance was by Helmholtz called *unconscious reference theory*. He assumed that the human mind somehow knows the illumination $I(\lambda)$. We know the color

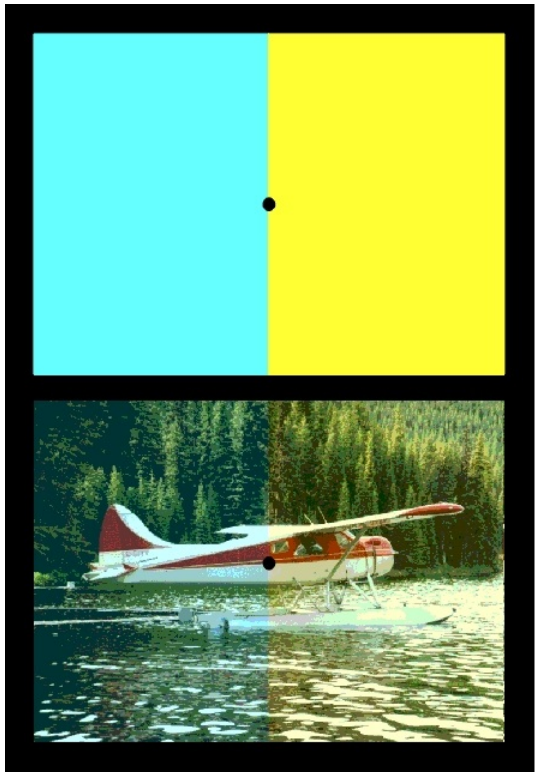

Figure 5.17: A demonstration of chromatic adaptation. Fixate on the black spot between the blue and yellow regions for about 30 secs. Then shift your gaze to the white spot in the center of the airplane image. Note that you do not notice the hue-shift between the left and right half of the image following this adaptation.

stimuli $C(\lambda)$ from the retinal image. And finding the reflectance of the image is basically a simple division. However, he left out all the details of how do we really know $I(\lambda)$ which is critical part of the puzzle.

Hering's Relational Theory

Hering made significant progress in the direction of finding a computational model to solve this under-determined inverse problem. He proposed that this information can be directly extracted from the images. The key is, due to change in illumination, the absolute amount of light reflected from surfaces change. But the relative amount of light reflected across different surfaces remain constant. The black ink surly reflects almost 100 times more photons in the sunlight than indoors, but so does the white paper. Hence there relative reflectance remains same. He proposed that we perceive the relative contrast of the elements in the scene. Wallach proved his claim by showing that luminance ratios across objects is important. To do this, Wallach used two different scenes from two pairs of projectors. In one pair, luminance A was projected in the center by one projector and luminance B are projected around it in an annular ring from the second projector. In the second pair, luminance C was projected in the annular ring surround. The subjects were asked to match the light projected at the center of this ring from the second projector so that it matches A. Note that C is different than B. What the subjects did was adjust the gray in the center of the second pair to D such that $\frac{C}{D} = \frac{B}{A}$. This showed that contrast is important for human vision.

The next question is how do we calculate such luminance ratios? Wallach further showed that the ratio at the edge is important. For example, Figure 5.20 the luminance profile of one horizontal line across the image is shown. Note that the grays in the annular ring and center of the black region are both the same. But we perceive the contrast that exists at the edges only. So this told, that edges are very important for our perception.

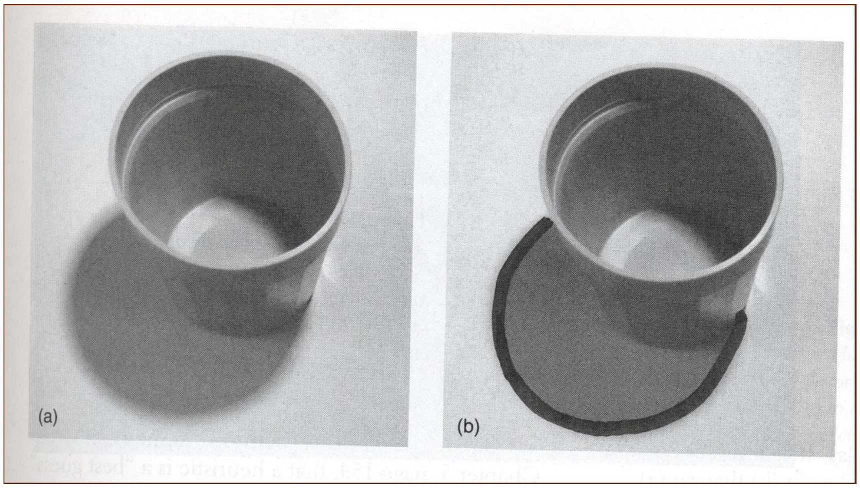

Figure 5.18: Lightness constancy with no change in illumination cannot be explained by adaptation theories.

Figure 5.19: Surface Based Stage: Shadows does not affect our perception.

Retinex Theory

However, note that we not only have a feel for the reflectance of adjacent objects that share a edge, but also of objects which are separated further apart in the image and do not share an edge. In that case, the question was, how do we really integrate the perceived contrast at the edges to create a perception of relative reflectance across different objects of the scene. Land and McCann provided a reasonable answer by proposing the *retinex theory* in 1971.

The basic assumption of this theory is that the illumination is continuous. This means, that all the images in the scene are due to drastic change in reflectance and not in illumination. Let us consider a scene in Figure 5.20 that has areas of three different reflectance coded by the different grays. Let the reflectance of area 1, 2 and 3 be R_1, R_2 and R_3 respectively. Note that 1 and 3 are non-adjacent areas that do not share an edge. Let the luminance in the direction x (vertical) be denoted by $I(x)$. The assumption is $I(x)$ is continuous. Therefore, the contrast ratio at edge of area 1 and 2 is given by

$$C_{1\rightarrow 2} = \frac{R_1 \times I(x)}{R_2 \times I(x)} = \frac{R_1}{R_2}$$

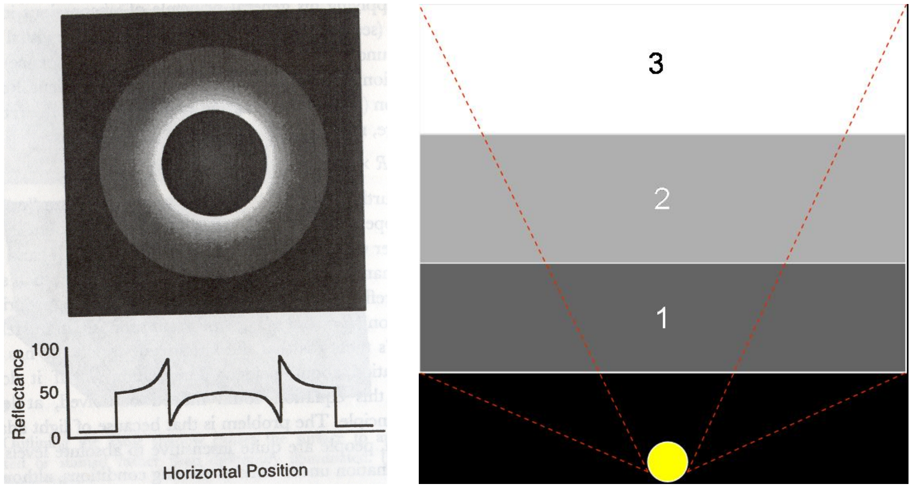

Figure 5.20: Left: We perceive contrast ratio at the edges. Right: The integration of ratio at the edges by the retinex theory.

Similarly

$$C_{2\to 3} = \frac{R_2}{R_3}$$

Therefore, to find the relative reflectance of area 1 and 3, we need to find the ratio of R_1 and R_3. From the above equations,

$$\frac{R_1}{R_3} = C_{1\to 2} \times C_{2\to 3}$$

Thus, retinex theory shows that if the illumination is continuous, then the reflectance is integrated across the image by multiplying the contrasts along edges. Note that the same result is obtained irrespective of whether we started from $3 \to 2 \to 1$ or $1 \to 2 \to 3$.

Scaling Problem: Thus, the retinex theory tells us how we get an idea or perceive that relative reflectance across different objects. However, when we look at objects, we perceive them as having deterministic luminance in the scale of white to black. So how do we solve the scaling problem to come to this absolute representation from the relative representation.

The retinex theory also proposed an *anchoring heuristic* for this purpose. It showed that we perceive the region of highest luminance as white. Black is always considered as no light. An experiment was done where a pair of projector was used to project a circle of dark gray surrounded by a circle of light gray. The subjects were then asked to identify the colors. The field of view of their vision was not restricted in any way so that they can see all the other lights and objects on the scene. In the second stage, the same scene was projected, but now the subject's field of view was restricted so that they cannot see anything else other than this projected scene, they perceived that lighter gray as white and the other one as a darker gray close to the black. This shows how we solve the scaling problem.

However, note that the biggest shortcoming of the retinex theory is in assumption that the illumination is continuous which is almost never true in real world. Shadows are caused due to discontinuous illumination. This in turn emphasizes the importance of detecting the reflectance and the illumination edges. The stimuli is a multiplication of illumination and reflectance. Hence, edges in the illumination and reflectance can both cause an edge in the stimuli. Identifying them are critical for judging the reflectance of a surface.

Thus, the whole problem if to find two maps, *illumination maps* and *reflectance maps*, the product of which gives the retinal image. Of course, the problem is under constrained. But is there enough reasonable heuristic assumptions that we can make, so that we can actually solve this problem from images only?

How do the humans achieve it?

However, as we mentioned before, humans achieve this with relative ease and in no time. There has been evidence that the humans use the following heuristic to achieve this task. Note that some of these heuristics may need information that is not present in the retinal image like depth. Nevertheless, it is important to know what are such heuristics that we use.

- *Fuzziness:* Note that only point light sources can cause sharp edges. Extended light sources, as is common in everyday world, usually have a penumbra region which causes the illumination edge to be fuzzy. However, a reflectance edge can be very sharp. This is exactly what changes our perception of the shadow in Figure 5.18.

- *Planarity:* Human system also has information about the depth at every pixel of the retinal image. As we will study in the subsequent chapters, images from two eyes are used to extract this depth information. Notice that in most regular scene, we have a simple light source like a bulb or sunlight that should produce continuous illumination for planar scene geometry. Most of the time, the discontinuities are due to non-coplanar geometry which can be detected easily by the depth information. So, edges around non-coplanar geometry are usually classified as illumination edges.

- *Magnitude of Ratios:* Also, the ratio of reflectance of a good white to a good black object may be at most $10 : 1$. Thus the contrast ratio along an reflectance edge is hardly more than $10 : 1$. However, illumination edges can have a contrast ratio as large as $10,000 : 1$ (sunlight and shadow). This magnitude also plays a role in detecting the two types of edges.

- *Chrominance:* Finally, additional information of the chrominance helps. For example, illumination edges would almost always show a change in only luminance and not in hue and saturation. However, the same is not true for reflectance edges.

Even in the presence of color, following are a few heuristics that had been found to help constrain the problem more.

- *Consistency of Illumination:* This assumes that the hue and saturation of the illumination is same at all regions of the scene.

- *Restricted Illumination:* Realistically illumination can only be of a few types. For example, it is very rare, if not impossible, for a scene to be lit by a monochromatic light.

- *Restricted Reflectance:* Realistically, objects reflectance can also be of some particular categories. For example, an object that only reflects one wavelength if difficult to find.

Based on these there have been some heuristic on the property of illumination and reflectance edges. Figure 5.21 shows the kind of spectrum we can expect on either side of an edge. If the edge is an illumination edge, crossing over of spectrums in unlikely. Just the opposite is the case for reflectance. Using this a theory was proposed. If the response of any two wavelengths of the reflectance or the illumination is considered across the edge. And these two response are joined by straight line as shown in Figure 5.21. There may be four cases as to how the lines will be arranged with respect to each other. In two of them, they will cross caused *spectral cross points*. In the third, they will not cross, but will have opposite slopes. All these cases are likely to happen only in case of a reflectance edge. The fourth and only case where the lines are parallel is due to a illumination edge.

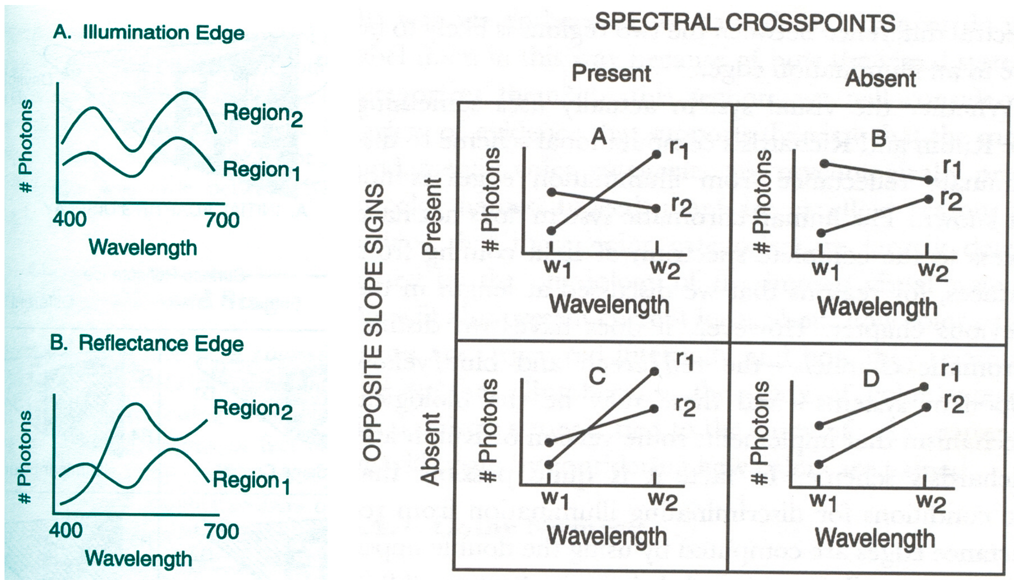

Figure 5.21: Left: The spectrum of colors on both sides of an illumination edge and a reflectance edge. Right: Spectral cross points and opposite slopes can only happen for reflectance edge.

5.6 Category Based Color Perception

Note that we have an amazing capability of naming and grouping colors. In fact, we categorize colors *perceptually* much more than we do many other physical quantities. For example, Figure 5.22 shows a rainbow that is nothing but a continuous change of wavelengths of light over space. Notice that it has all the different wavelengths occupying equal regions spatially. Compare this with the bottom element of the figure, where we have showed a continuous change in luminance from white to black, with each level occupying equal region spatially. Note that the luminance spectrum looks very continuous. However, the rainbow gives the sensation of seven colors arranged in stripes with fuzzy boundaries. This shows that probably in our perception, all different colors are not categorized similarly. Probably, some colors are more important than others. In this section we will see if such a categorization exists and if so, how does it help in our color perception.

One important question in this direction is, does the categorization of color have a perceptual basis, or is it purely due to culture. Did this naming develop just because we had to communicate? Or, our biological make up has some influence on the naming? For a long time, a theory that was prevalent was that this naming is completely due to cultural idiosyncracies and does not have any perceptual basis. This theory of *cultural relativism* was propounded by Sapir-Whorf. But the very fact, that all humans across all cultures and countries would probably see the rainbow with the same seven colors inspite of the presence of all different wavelengths, raises questions about the validity of this theory.

This theory was refuted strongly by Berlin and Kay's work from UC-Berkeley in the early 70s. Berlin and Kay first identified 11 basic colors in English based on the following rules.

- A basic color should be *monoleximic*. So, colors like off-white, whitish-brown etc were ruled out.

- Basic color should have a *primary chromatic reference*, i.e it should be identified with a wavelength. This eliminated colors like gold and silver. Orange was also a contender for elimination. However, other then fruit, it has a strong association with the wavelength of light and hence was retained.

- A basic color should be *general purpose* i.e widely used for different kinds of objects. This ruled out colors

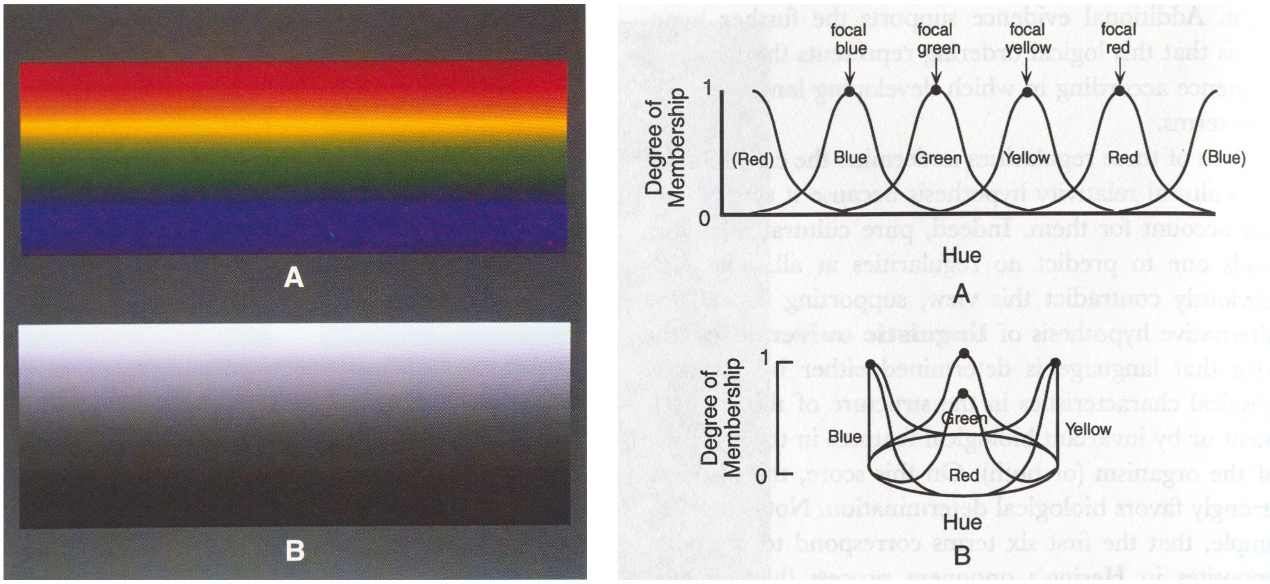

Figure 5.22: Left: Compare the rainbow which is a continuous spectrum of wavelength with the continuous spectrum of luminance. While the luminance looks like a continuous field, the rainbow looks more like discrete colors whose boundaries are fuzzy. Right: Rosch's focal and boundary colors.

like blonde (used for hair) and roan (used for horses).

- A basic color should be *frequently used*. This ruled out the colors like mauve, taupe and burgundy which are not used much regularly, but mostly by color professionals to distinguish the subtlety of different color.

Based on these four rules Berlin and Kay identified 11 basic colors as *red, green, blue, orange, yellow, pink, brown, white, black, gray* and *purple*.

Next they studied 20 different languages, 8 by travelling to different countries and interacting with them and the rest from literature to see how the basic colors in those languages relate to these 11 basic colors in English. Note that if the theory of cultural relativism hold, you will find almost no regularity in the basic colors developed out of different cultures. However, strangely enough they found that most cultures abide by these 11 colors. However, a few more colors can be added to the basic color list based on their study of other languages and make it a list of 16 basic colors. The five colors added were based on the frequency of use of these colors in most other languages. These are a term of *sky-blue*, terms referring to *warm colors* (red and yellow), *cool colors* (blue and green), *light warm* (white and red and yellow) and *dark cool* (black and blue and green). This completed Berlin and Kay's list of 16 basic color terms out of millions of color that we can see. Note that colors are not arranged linearly in this basic set. For example red or yellow belongs both under warm and light warm and blue and green belong both under cool and dark cool. In fact a hierarchy can be constructed with these basic colors based on the inclusion of one in another category. Berlin and Kay built such a hierarchy and investigated if this hierarchical perceptual organization is similar across the world. It turned out that it was and thus ruled out the cultural relativism theory. From then on, a lot of work has been done on the categorization of color.

Aristotle defined an item to belong to an category when it followed a certain set of rules. Thus, every item of a category should follow strictly a set of rules. The question was, do the colors form such an categorization. As it turns out, they don't. The kind of categorization they form are more of *prototyping*. Basically, this means that a color may not satisfy all properties of a group, but some of more than groups and can thus belong *kind of* to many groups. This was experimented by identifying a few focal colors which were a subset of the basic colors. The rest of the colors were called the boundary colors. 329 color chips were used and the following two experiments were performed. First, the subjects were shown all the 329 chips and were asked to categorize each color to a class of focal color. The subjects found this to be very difficult job. Second, the subjects were again shown the same 329 chips and were asked to pick the 'best example' of the focal colors. Though this involved quite a bit of search, the subjects found it a much easier job. The difficulty was measured as proportional to the time taken for each color to perform these tasks. The second experiment was performed by Rosch on a old tribe at New Guinea called the Dani. It was found that these people have only two color terms: *mili* and *mola* for light-warm and dark-cool. Rosch's experiment was to teach them the other colors. It was found that they learned all the basic colors red, green , yellow and blue much faster than the others. The third experiment was again by Rosch, where he presented the subjects with a series of colors. The subject's task was to say if the color was an instance of the category name mentioned before. It was found that the subject responded very fast with the focal colors. In fact, the time the subject took to decide increased proportionally as the color moved away from the focal colors.

Based on this experiment, Rosch proposed a fuzzy set theory to define color categories. In traditional set theory, a item can either belong to a set (membership of 1) or not (membership of 0). There is no intermediate state. However, in a fuzzy set theory, a color can have partial belonging to a category (membership of a fraction between 0 and 1) and hence can belong to more than one set. Rosch define four color categories around four focal colors as shown in Figure 5.22. Each of these defines a fuzzy set of color, where the membership of a color in the set is defined by the curves shown in Figure 5.22. He identified four focal colors as red, green, blue and yellow. These colors have a membership of 1 in their categories. All other colors have a fractional membership in the adjacent categories. The better way to visualize is to have a circular representation. In that case, the purple can also be defined as colors having fuzzy memberships in the both red and blue category.

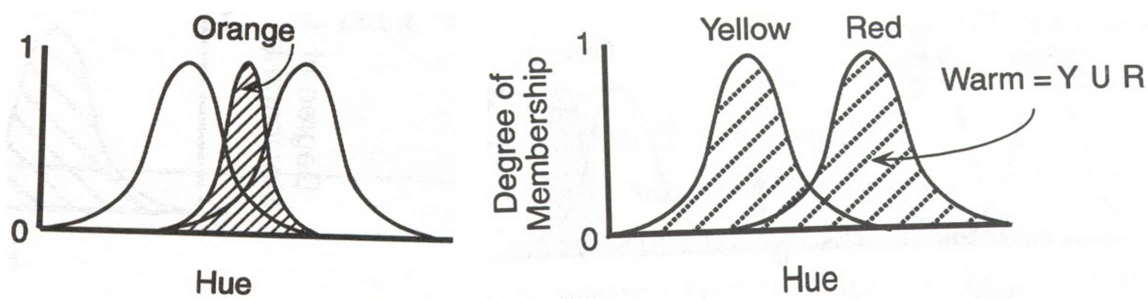

Figure 5.23: Left: Fuzzy AND of two colors creating a derived color. Right: Fuzzy OR of two colors creating a composite color.

A new theory was proposed by Kay and McDaniel in 1978. They proposed six focal colors: red, green, blue, yellow, white and black. Notice the relationship of these colors with the opponent colors in the opponent color theory. They proposed a class of derived colors which is a fuzzy AND of the focal colors. Ideally, a fuzzy AND should be the intersection of the fuzzy sets defined by the focal colors. However, Kay and McDaniel defined the fuzzy AND by a function that is 1 where the function for two categories meet. Then they fall off on both sides regularly until it reached 0 at the adjacent focal colors. This is illustrated in Figure 5.23 where the color orange have a membership function defined by the striped function. They defined five derived colors: orange (yellow AND red), gray (white AND black), purple (red AND blue), pink (white AND red), and brown (yellow AND black). Finally, they defined composite colors as a fuzzy OR of two or more focal colors. Fuzzy OR is defined as the maximum membership of the two colors at any point, or with a union of the sets as shown in Figure 5.23. They defined five composite colors: warm (red OR yellow), cool (blue OR green), light-warm (white OR red OR yellow) and dark-cool (black or blue or green). Thus they explained the sixteen basic colors proposed by Kay and Berlin. This concludes the category based color theories.