UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA

Lecture notes for
Multimedia Data Processing

Color and its Representation

Last updated on: 26/03/2020

The Light

- Light is an electromagnetic radiation which is characterized by the energy transported. This energy depends on the wavelengths and the number of photons in the radiation.
- The fundamental radiometric quantities that describe an electromagnetic wave are:

Radiant energy : Q_e

Radiant flux : $P_e = \dfrac{dQ_e}{dt}$

- From these quantities are then derived the two quantities capable of describing the point and extended light sources respectively:

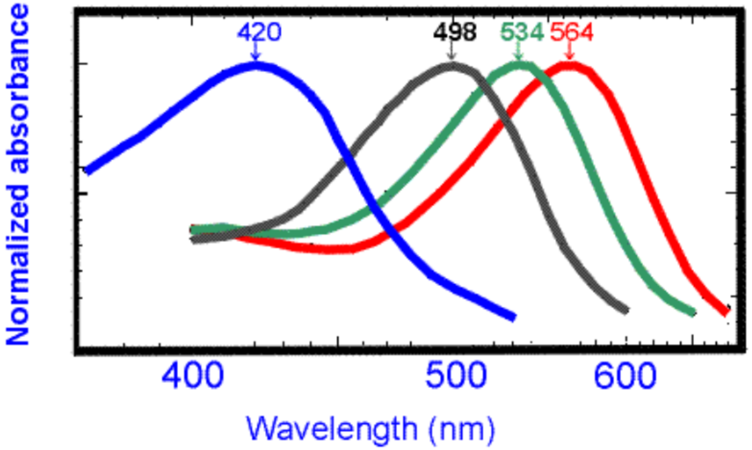
Radiant intensity: $I_e = \dfrac{dP_e}{d\Omega}$

Radiance: $L_e = \dfrac{d^2 P_e}{dA_p d\Omega}$

- These quantities (total) are the combined result of the properties relating to each wavelength (spectral quantities).

The Human Vision

- Our eye perceives the color (hue) based on the wavelength of light that produces the stimulus on the retina from red (longer wavelength), orange, yellow, green, blue, violet.
- The visible spectrum ranges from 380 to 780 nm
- The retina has about 120 million rods that are sensitive to brightness and achromatic, and about 8 million cones sensitive to different wavelengths placed in the fovea.

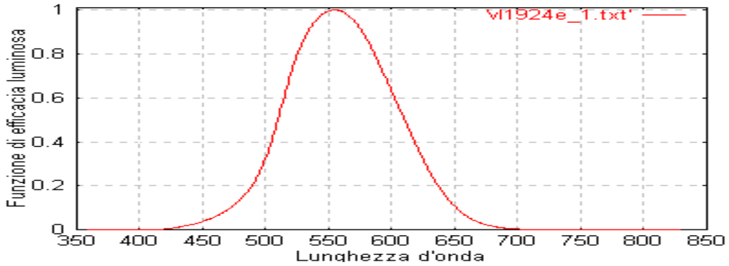


After Bowmaker & Dartnall, 1980

- Color perception occurs through the combination of three primary stimuli (trichrome): the signal is filtered at the retina level and divided into three colors blue red and green (green-yellow).

Photometry

- The radiometric quantities are not suitable for quantifying the light emission in terms of the response of the human visual system that is characterized by an uneven sensitivity to different wavelengths.
- The sensitivity of the eye was standardized in 1924 by the *Commision Internationale de l'Eclairage* (CIE) through the definition of the **Standard Observer** and its function of luminous efficacy:



- By weighting the radiometric measurements with the luminous efficacy function, the photometric quantities are obtained:

Luminous Intensity: $I_v = K_m \int_{380}^{780} I_{e,\lambda} V(\lambda) d\lambda$ m.u.: cd

Luminance: $L_v = K_m \int_{380}^{780} L_{e,\lambda} V(\lambda) d\lambda$ m.u.: $\frac{cd}{m^2}$

Quantify the Color

- Our eye (differently from the ear) is not able to perform a spectral analysis, but reports a sensation resulting from the combination of all visible wavelengths.
- It is also impossible to communicate the sensation corresponding to a certain stimulus. What we can communicate is that two different stimuli produce the same sensation.
- We can define three types of colors:
  - **spectral colors** (produced by radiation containing a narrow band of wavelengths), **magenta** or **tyrian purple/porpora** (radiation containing a combination of the extremes of the visible spectrum, i.e. blue-purple and red-yellow), **white** (continuous spectrum radiation)
- The theory of Young-Helmholtz (1801) hypothesizes that in order to reproduce the sensation associated to a specific color, three primary radiations can be suitably combined (trichrome systems)
- The first experimental confirmation obtained from experiments on the eye occurred in 1964.

RGB and XYZ

- The first numerical realization of the tristimulus space made on experimental observations is by Maxwell in 1860.
- The reference monochromatic radiations, considered as orthogonal triad, were:

$$\lambda_b = 456.9nm, \lambda_g = 525.1nm, \lambda_r = 630.2nm$$

- The RGB space was formalized and standardized by the CIE in 1931, choosing as a reference:

$$\lambda_b = 435.8nm, \lambda_g = 546.1nm, \lambda_r = 700.0nm$$

- In the definition, an average of the data from multiple observers was taken.
- It was observed that the RGB coordinates must also be negative to reproduce the stimuli corresponding to all the components of the spectrum. *There was no highlighting of the stimulus luminance component*.
- Thus was born the XYZ system in which the Y represents the luminance component and the colorimetric functions are all positive.
- This space is the reference for all subsequent studies on colorimetry.

CIEXYZ Color Space

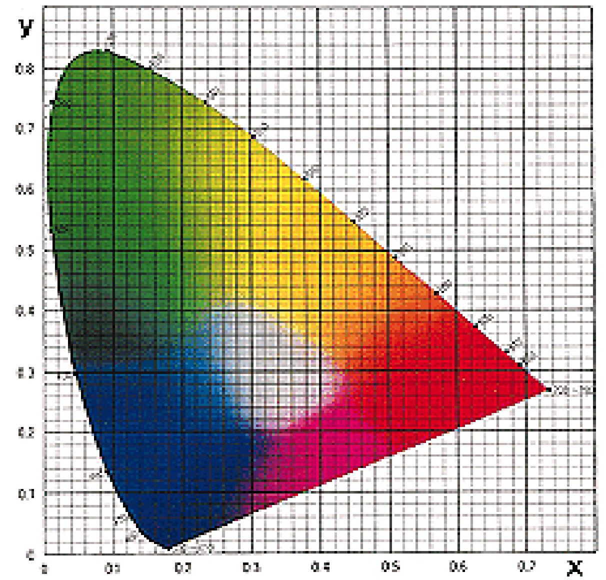
- The conversion from CIERGB to CIEXYZ is defined by the following linear transformation:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4900 & 0.3100 & 0.2000 \\ 0.1770 & 0.8124 & 0.0106 \\ 0 & 0.0100 & 0.9900 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- In the XYZ space the chromaticity is defined by the intersection of the tristimulus vectors with the X + Y + Z = 1 plane, therefore:

$$x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z}, z = \frac{Z}{X+Y+Z}$$

- For the obvious consideration that z = 1-x-y, it is customary to represent the projection of this figure on the Z = 0 plane, obtaining a graph in x, y:



CIELAB Color Space

- Today, the CIEXYZ space is considered the absolute reference for all color definitions, but it has one major flaw: it describes the physical nature of the color and not its perception by the observer.
- A series of color spaces are then defined in 1976, among which **CIELAB** has become the de-facto standard in industrial colorimetric practice.
- The first coordinate of this space can be obtained from the following formulas:

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \qquad \left(\frac{Y}{Y_n}\right) > 0.008856$$

$$L^* = 903.3\left(\frac{Y}{Y_n}\right) \qquad \left(\frac{Y}{Y_n}\right) \leq 0.008856$$

- This is called *CIE 1976 Lightness* and reproduces the perception of the brightness of the Standard Observer, referring to an illuminant whose characteristics are (Xn, Yn = 100, Zn).
- This scale is uniform with an excellent approximation and it is defined in the range [0,100].

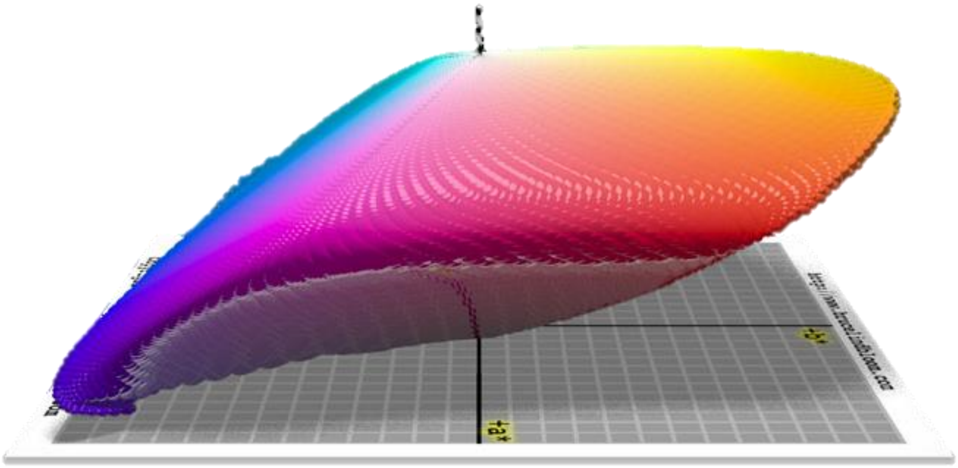
CIELAB Color Space

- The other two coordinates are a * and b * defined as:

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right]$$
$$b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right]$$

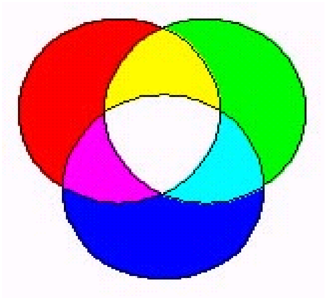
- where $f(x) = \begin{cases} x^{\frac{1}{3}} & x > 0.008856 \\ 7.787x + \frac{16}{116} & x \le 0.008856 \end{cases}$

- On each plane with constant lightness hue angle and chroma (saturation) can be defined. The difference between colors can be viewed as the Euclidean distance between the color vectors.

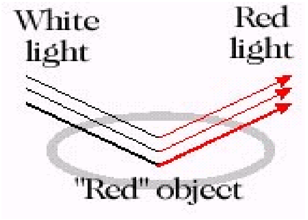


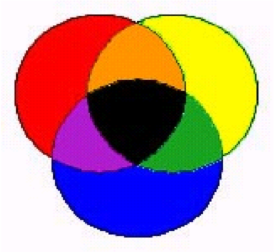
Color Reproduction

- A color can be reproduced with three fundamental techniques capable of reproducing the same stimulus:
- **additive synthesis:** mixture of lights of different chromaticity



- **subtractive synthesis:** mixture or overlap of media that absorb different spectral components differently



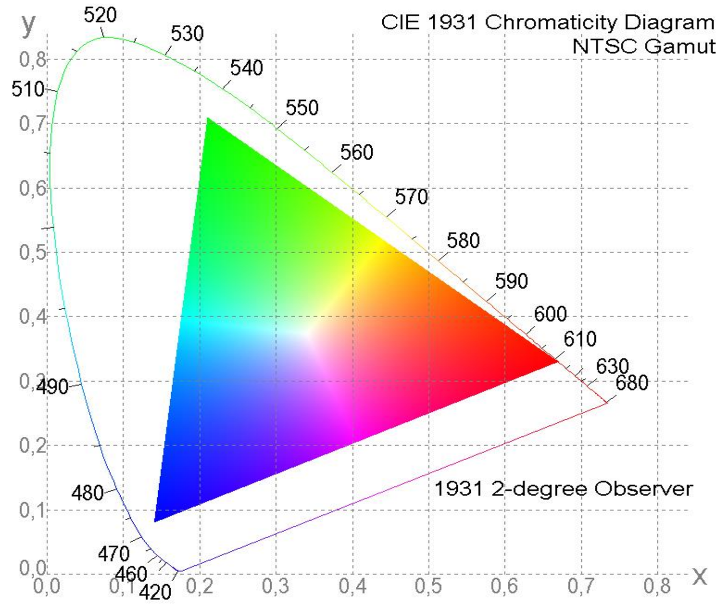


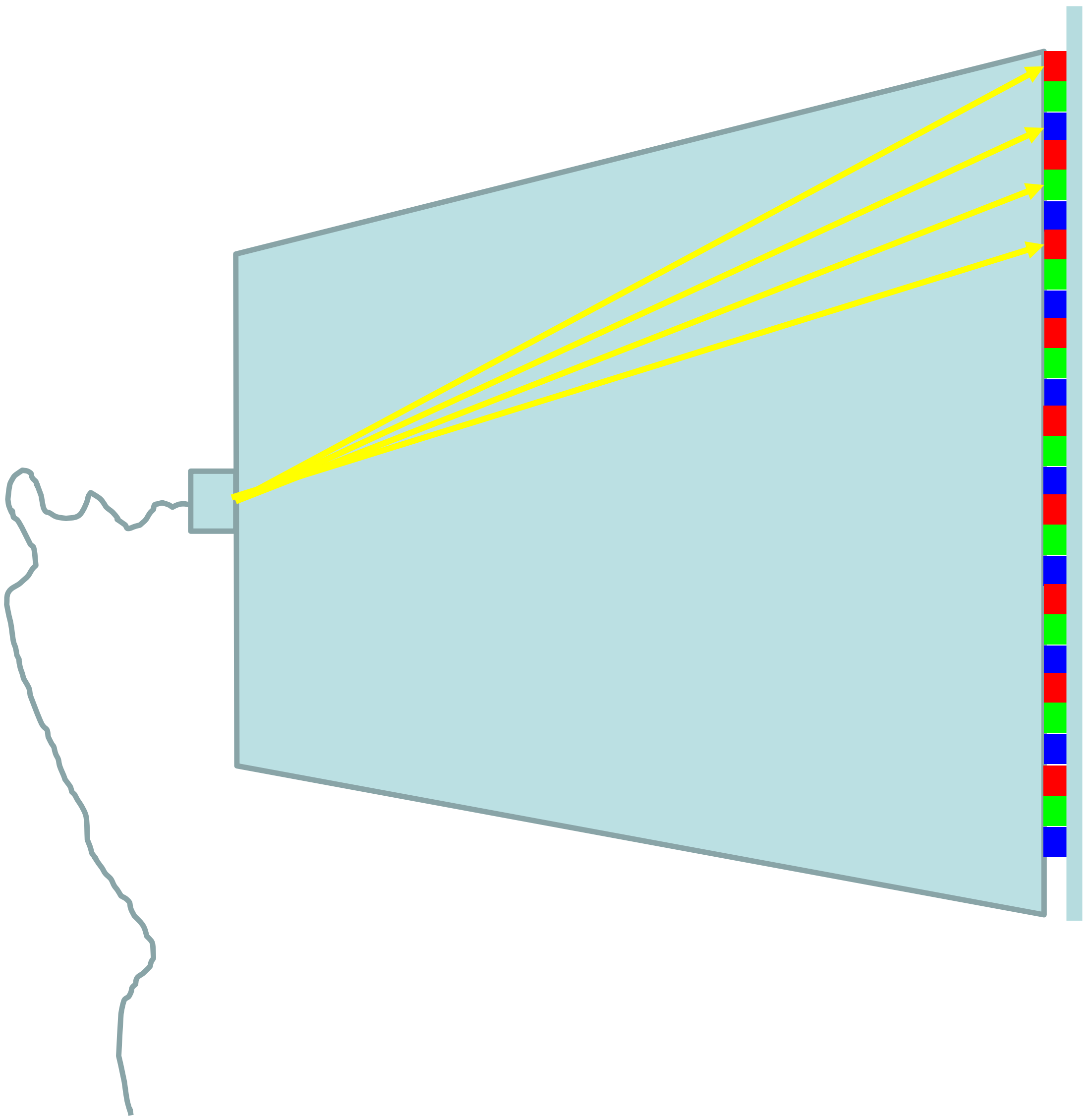
- **spatial integration:** differently colored and close points that overlap in the human eye, thus producing a result similar to additive synthesis (this process occurs in monitors and in screening inks printing)

CRT Monitor

- The television system is conceived starting from the cathode ray tube structure (CRT) which synthesizes the color through spatial integration of RGB phosphors arranged in mosaic on the screen.
- The color space is therefore built on the basis of the chromaticity of the lights emitted by the three types of phosphors.
- An RGB space is thus obtained in which only the colors enclosed in a cube can be displayed, the vertices of the cube represent the colorimetric coordinates of the phosphors.

The intersection of this space with the R + G + B = 1 plane produces a triangle that can be represented in the CIEXYZ space and it is called *gamut* of the monitor.





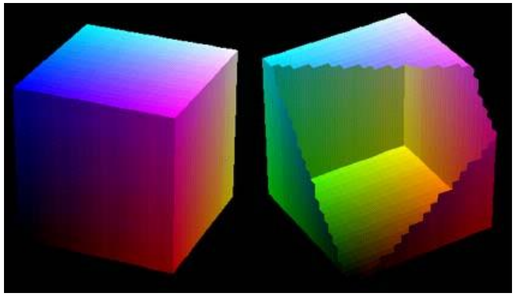
Gamma Correction

- The relationship between the luminance emitted by the CRT monitor and the value of the signal that drives the electron beam is not linear but exponential:

$$\frac{L_v}{L_{v,\max}} = \left(\frac{E}{E_{\max}}\right)^{\gamma}$$

- Since the electronic circuit necessary for linearization was very expensive, the signal was corrected with the so-called gamma correction before being transmitted over the air. Thus, The gamma correction refers to the exponent and not to the colors.
- Starting from the NTSC standard, the choice was $\gamma = 2.2$

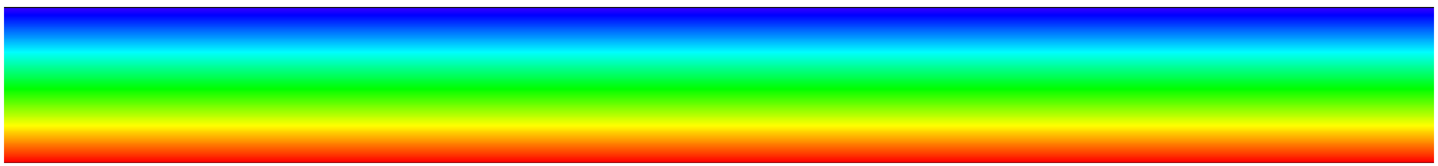
Realistic values for gamma would be between 2.35 and 2.55, but the television standards are designed for situations in which the screen is viewed in environments with soft light and therefore leave a margin to obtain an increase in contrast.



RGB Color Space at The Computer

- The representation of color in electronic machines is typically designed according to the video cards that will have to drive a CRT monitor.

- A typical representation is the so-called *truecolor* which uses three 8-bits values for each pixel to control the three signals brought to the monitor.



- If not otherwise specified, when working with digital images on the computer you are working with values acquired and corrected with a gamma factor.

- On the other hand, when generating synthetic images in which human perception is relevant, it is important to know the transformations made by the video card, which can introduce other corrections.

Transmissive Color Spaces

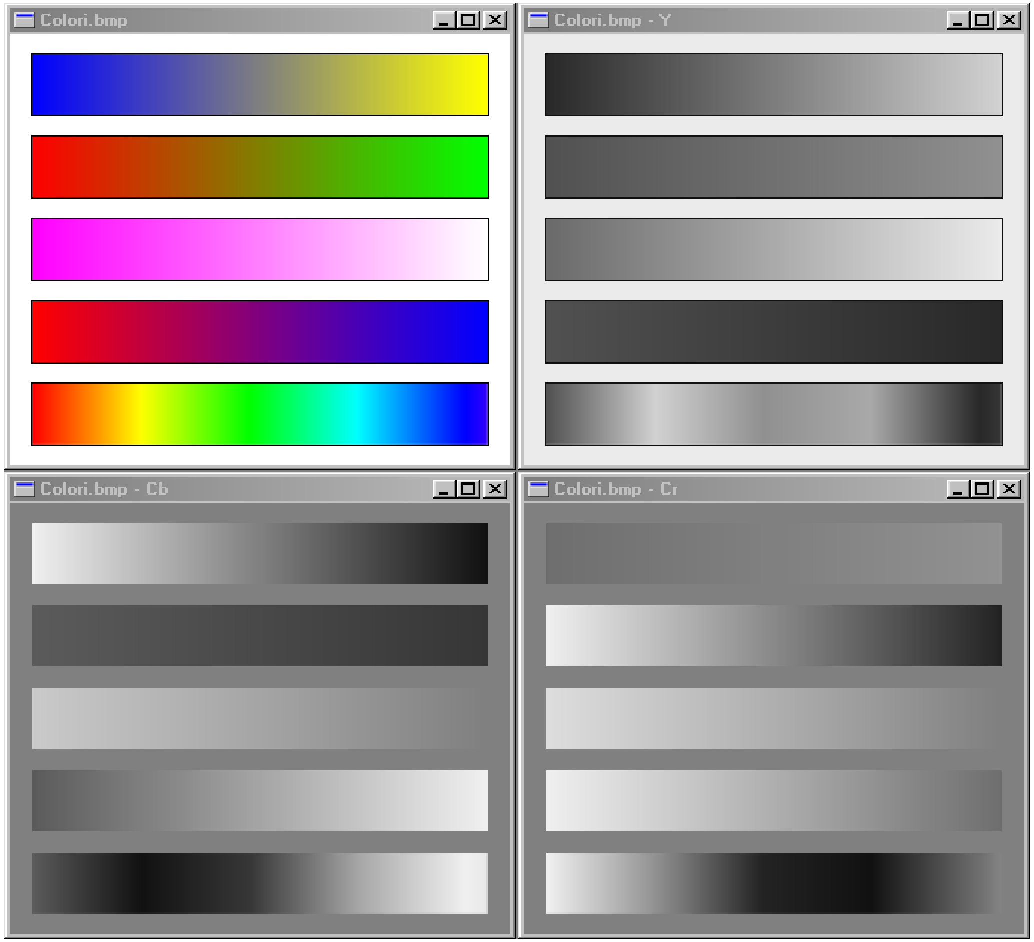
- When it was created, the video signal had to be visible also to black and white televisions, so it was thought to separate the luminance component (Y) from two other, called chrominance.
- A series of standards was born, among which we remember:
  - **YIQ** (NTSC)
  - **YUV** (PAL)
  - **YCC** (Kodak PhotoCD)
  - **YC_BC_R** (Digital Video, JPEG, MPEG)
- While the YIQ and YUV standards are important in the world of analog signals, the YCBCR standard has assumed a dominant role in digital.
- The conversion formulas in the case of RGB values [0..255] are:

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

- These conversion formulas are taken from the JPEG File Interchange Format, Version 1.02

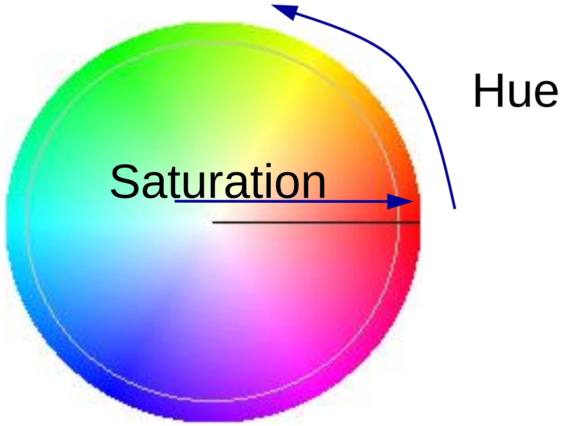
YC_BC_R Example

- In this example the effect of a YCBCR decomposition are shown:



Ordered Space for Computer Graphics

- In addition to the RGB standard, these spaces are used to introduce numerical color specification.
- This type of transformation is useful if you are processing images for an interface with the human operator.
- They are not devised to be precise and are obtained as a transformation from an indefinite (in terms of coordinates and range) RGB space.
- **Representation with luminance and chrominance (A.H. Munsell)**
- All characterized by two basic concepts:
  - H = Hue “Tinta”
  - S = Saturation “Purezza”



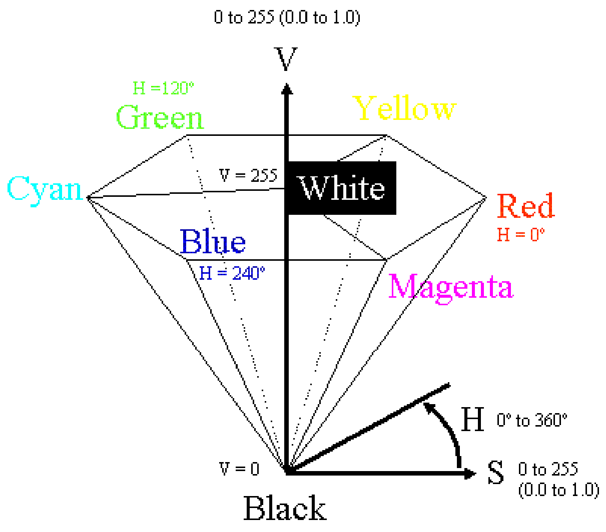
HSV Color Space

- **HSV** (hue, saturation, value): H is an angle between 0 and 360 degrees, S and V are values between 0 and 1. This is a $0 \le R,G,B \le 1$ color space transformation, where for each point are defined:

$$Max = \max(R,G,B)$$
$$Min = \min(R,G,B)$$

- The coordinates are then given by the following formulas:

$$V = Max$$
$$S = \frac{(Max - Min)}{Max}$$
$$H = \frac{\pi}{3}\begin{cases} \dfrac{G-B}{(Max-Min)} & Max = R \\ 2 + \dfrac{B-R}{(Max-Min)} & Max = G \\ 4 + \dfrac{R-G}{(Max-Min)} & Max = B \end{cases}$$



HLS Color Space

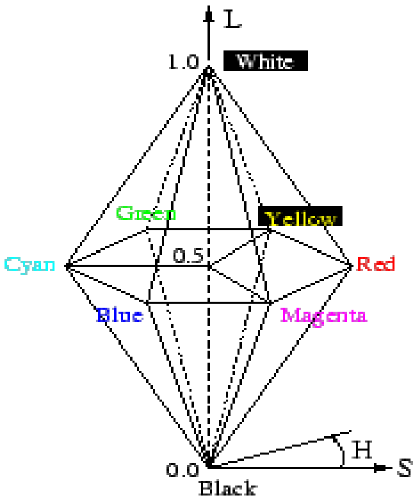
- Introduced by TEKTRONIX as its "standard color"
- It is nothing more than a reworked version of the HSV system where "value" is replaced by "lightness" :

$$L = \frac{Max + Min}{2}$$

- Only the S definition changed:

$$S = \begin{cases} \dfrac{Max - Min}{Max + Min} & L \le 0.5 \\ \dfrac{Max - Min}{2 - (Max + Min)} & L > 0.5 \end{cases}$$

  - The pyramid structure of the HSV is doubled as shown on the right.
  - This, for example, is the system used in the standard dialog box for selecting Windows colors



Chromatic Quantization

- Like spatial resolution, color resolution is also a compromise between cost and performance.
- It is not always necessary to have all the colors of the *truecolor* images (24 bits per pixel, 8 for each channel).
- It is therefore possible to **save storage space** and **reduce the computational load** by reducing the number of colors.
- To reduce the number of color levels:
  - we can use **fewer bits** per band (e.g. 5-5-5 or 5-6-5 in the case of 15 and 16-bit *hicolor* respectively)
  - we can a **look-up table** (color map): a finite number of colors (e.g. 256) stored on a table is chosen and the pixel value is a pointer to that table which contains RGB triad.
- The list of available colors is called *palette*

- "In ancient times" the CGA card had a 4-color palette, the 16-colors EGA and the "revolution" occurred with the VGA card that allowed a 256-colors palette!

Example

bitmap image with 256 colors

Let's see the contents of the rectangle at the coordinates
x=100..104, y=100..104

190	190	191	188	186
190	188	188	**188**	188
182	182	187	183	186
167	170	170	174	182
158	158	158	170	173

RGB map pointers values

Palette:

...			
186	1.0000	0.6118	0.3216
187	1.0000	0.5490	0.3529
188	**1.0000**	**0.5686**	**0.4000**
189	1.0000	0.6353	0.3255
190	1.0000	0.6118	0.4510
191	1.0000	0.6471	0.4196
...			

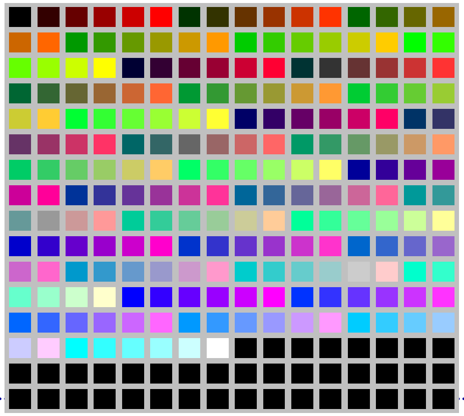
RGB [0.0,1.0] values

Palette Selection

- The choice of the palette plays a fundamental role in the quantization of colors.
- It is possible to **use a standard palette** that contains a subset of the possible colors, chosen by dividing the RGB cube into a standard number of steps for each channel.
- This operation can be done by choosing a step equal to:

$$N = \left\lfloor \sqrt[3]{Elem} \right\rfloor$$

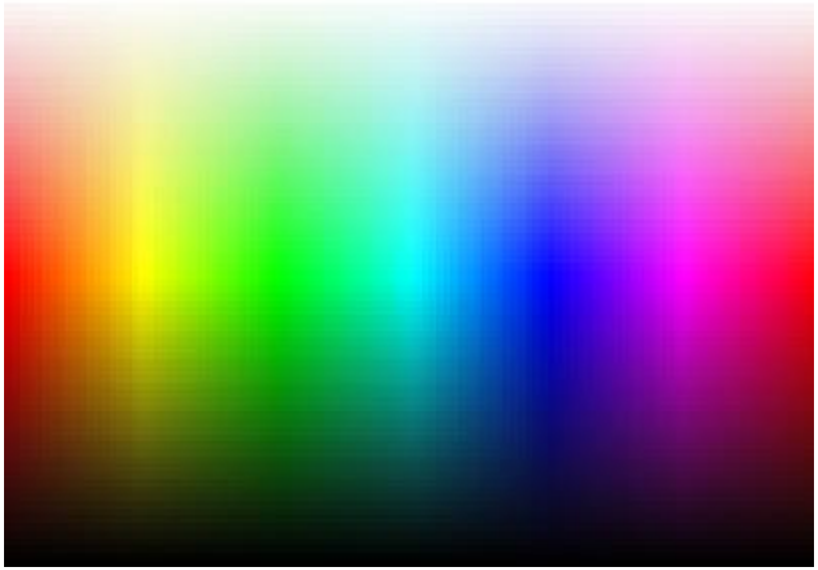
- where *Elem* is the number of elements to insert in the final palette. It is easy to verify how in the case of 256 elements the number of steps should be 6, leading to a palette of 216 equally spaced elements, while the remaining can be chosen at will (for example, shades of pink to better represent the skin).



Optimized Palette

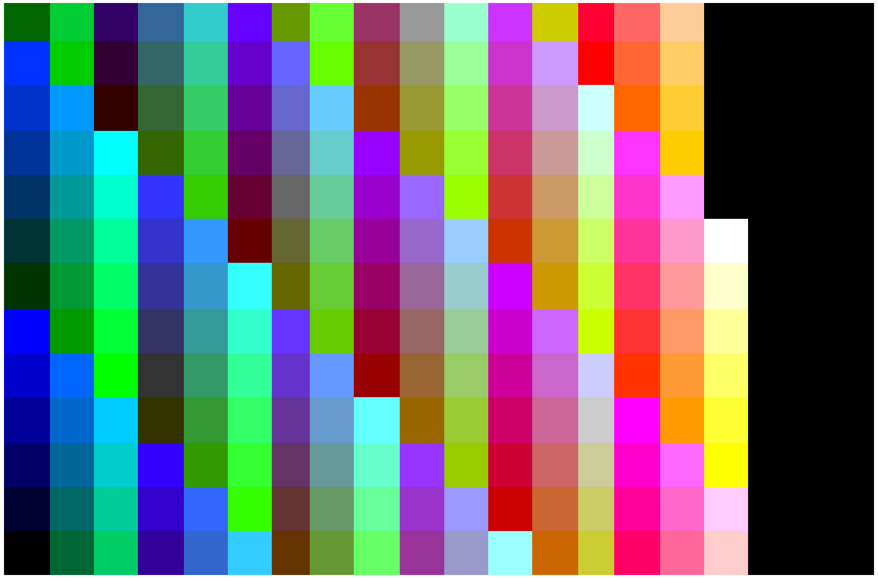
- Choosing a palette that best represents the color space of the image optimize the colors reduction.
- In 1980 Paul Heckbert proposed the Median Cut Algorithm, a technique for obtaining an optimized palette:
  - the **three-dimensional histogram** of the colors contained in the image is calculated;
  - the minimum parallelepiped oriented like the axes that contains all the colors must then be found;
  - a cut on the longer direction of the *box* must be performed. The cut should be in the point that leaves (approximately) the same number of elements in the two halves (median cut).
  - the cutting process is repeated for the box with most elements (or the largest as volume) until as many boxes as the required colors have been obtained. The representative colors will be given by the box’s colors average.
- A further point to focus on is the technique to assign the palette to the image, since this implies a concept of similarity between colors.
- Usually graphics programs use the **Euclidean distance in the RGB space** as a metric, even if this is not perceptually so meaningful.

Palettes Comparison



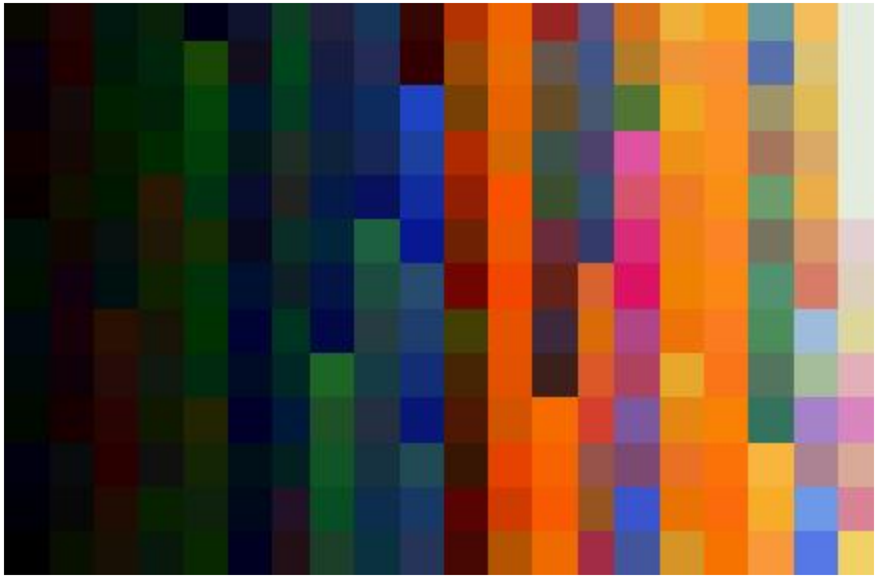
- Original 24-bits image with 21828 colors.
- (It is not a synthetic image, it is a *strelitzia* on a dark background)

Palettes Comparison



- Reduced 256-color image with standard palette.

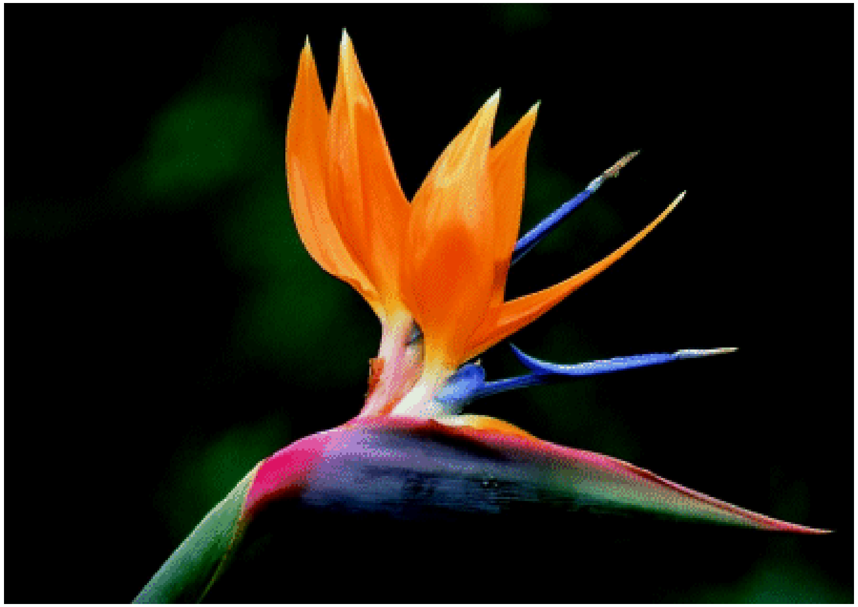
Palettes Comparison



- 256-color reduced image with Median Cut.

Original Image (21828 Colors)

256-Colors Image



128-Colors Image

16-Colors Image


8-Colors Image