



Multimedia Technology & Applications

IMAGE TECHNOLOGY

Part A: Graphics & Image Representation

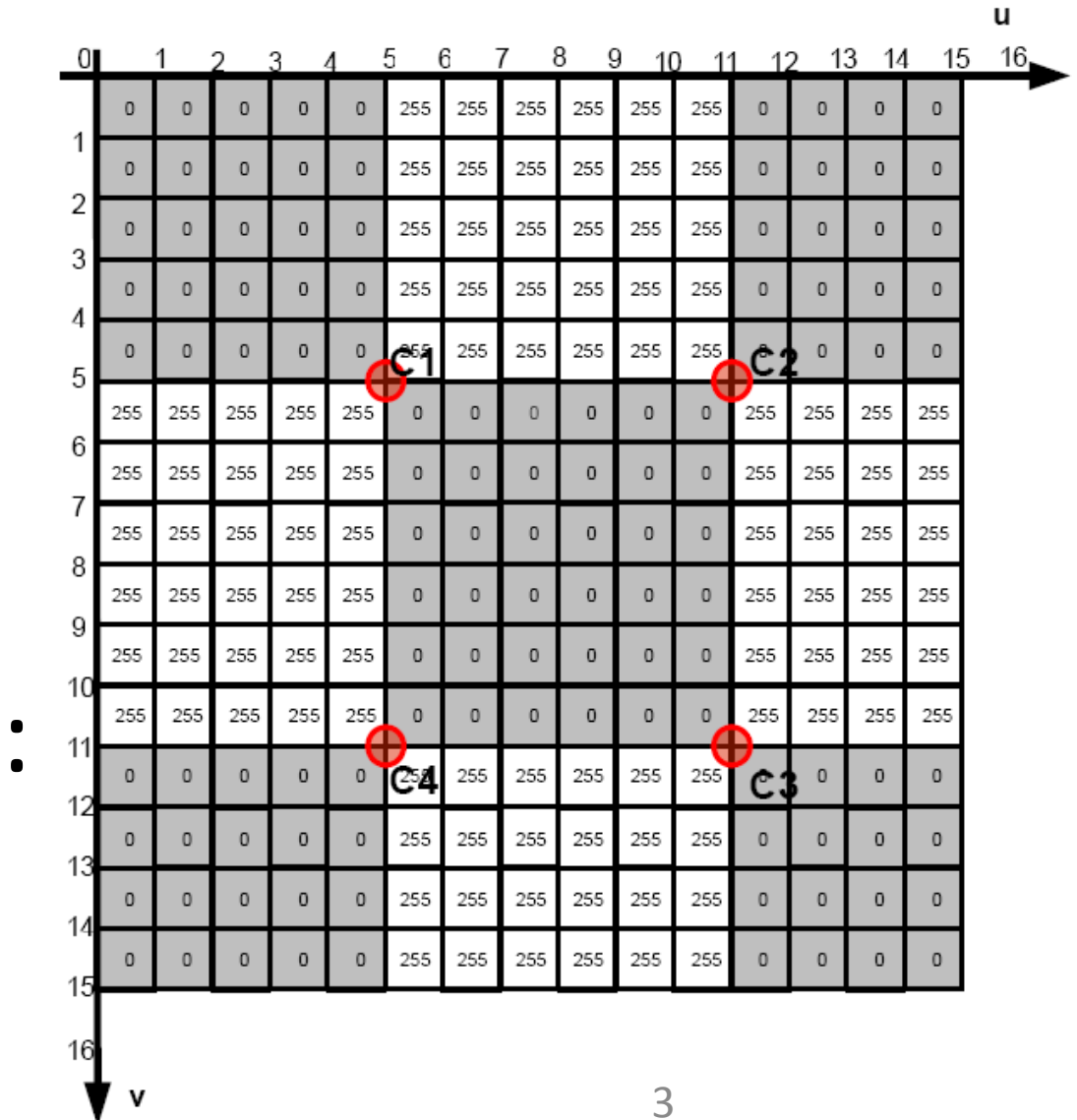
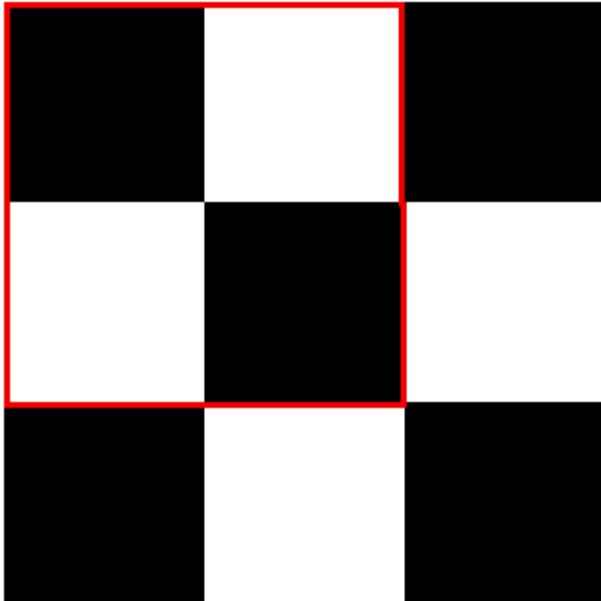
- Types of Images
- Image Format, Pixel, Resolution and Image Size, Sampling

Part B: Color Representation & Models

Part C: Image Processing & Analysis

- Dithering
- Histogram for Image Processing
- Image Compression Standard

Image: A two dimensional array



Types of Images

- Bitmap
- Vector graphics
- Procedural modelling

Bitmap Images

Computer displays

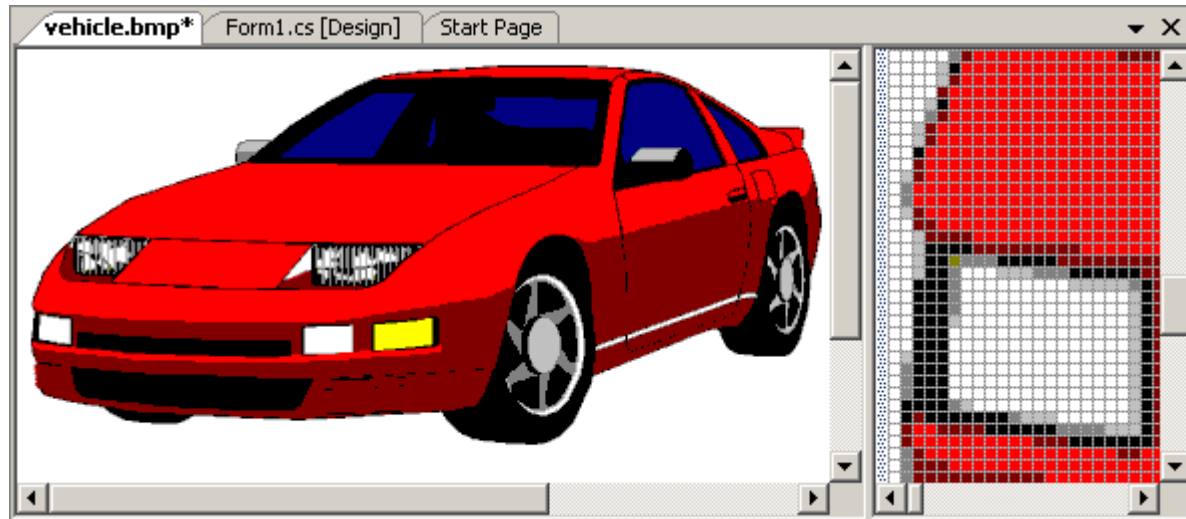
- Most modern computers have **bitmapped displays**, where each on-screen pixel directly corresponds to a small number of bits in memory.
- The screen is refreshed simply by scanning through pixels and coloring them according to each set of bits.

Bitmap Images

- Two-dimensional array of pixel (*picture element*)
- Also called as *pixmap*s or *raster graphics*
- Each pixel is a number representing the color at position (r, c) .
- Commonly created by digital cameras, scanners, paint programs (Corel Paint Shop Pro) and image processing programs (Adobe Photoshop).
- Bitmaps are appropriate for photographic images with continuously varying colors.

Bitmap Sample

- Created by software



Bitmap Sample

- Captured by camera



Bitmap Images

- 1-Bit images



Each pixel is stored as a single bit (0 or 1), so also referred to as **binary image**.

Monochrome 1 –bit Lena image

- 8-Bit Gray-level images

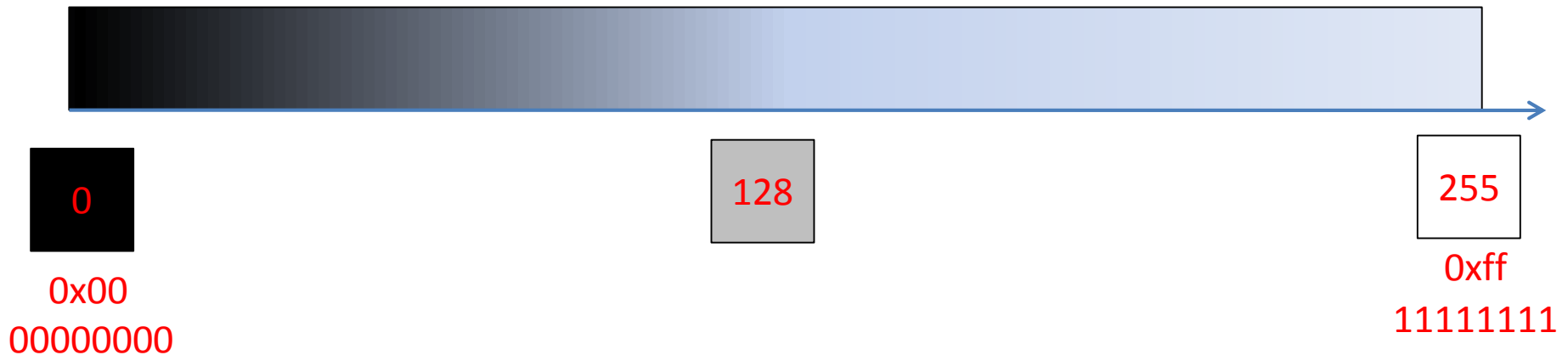


Each pixel has a gray-value between 0 and 255, represented by a single byte.

Each pixel is usually stored as a byte (a value between 0 to 255), so a 640x480 grayscale image requires 300 kB of storage ($640 \times 480 = 307,200$).

Grayscale image of Lena

Black-and-White (Grey scale)



Black-and-White (Grey scale)

8-bit color images

- Many systems can make use of 8 bits of color information (the so-called “256 colors”) in producing a screen image.
- Such image files use the concept of a **lookup table** to store color information.
 - Basically, the image stores not color, but instead just a set of bytes, each of which is actually an index into a table with 3-byte values that specify the color for a pixel with that lookup table index.

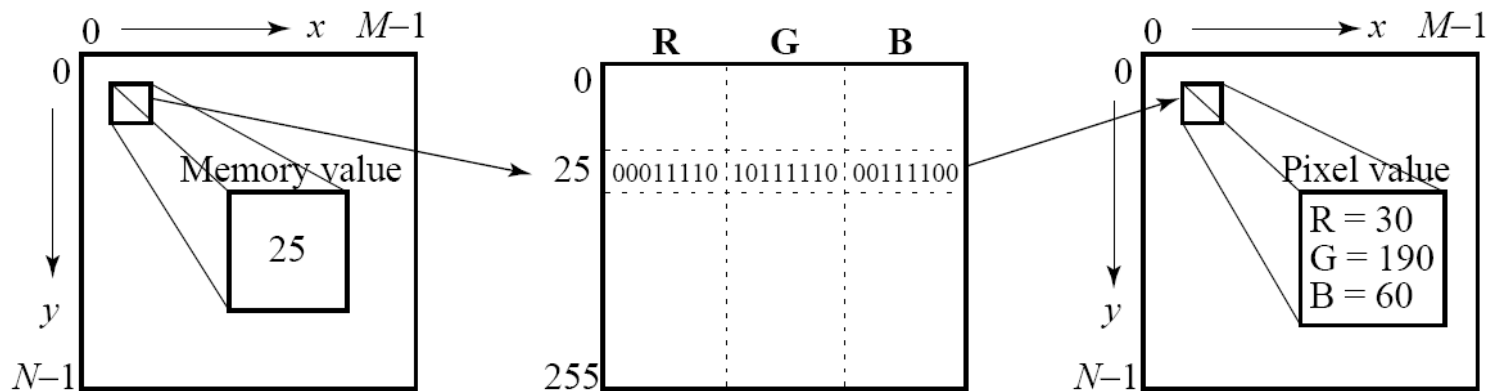


Example of 8-bit color image.

- Note the great savings in space for 8-bit images, over 24-bit ones: a 640x480 8-bit color image only requires 300 kB of storage, compared to 921.6 kB for a color image (again, without any compression applied).

COLOR LOOK-UP TABLES (LUTS)

- Look-up table is used to represent colour.
- Store only the index of the colour LUT for each pixel.
- Look up the table to find the colour (RGB) for the index.



Color LUT for 8-bit color images.

The idea used in 8-bit color images is to store only the index, or code value, for each pixel. E.g., if a pixel stores the value 25, the meaning is to go to row 25 in a color look-up table (LUT).

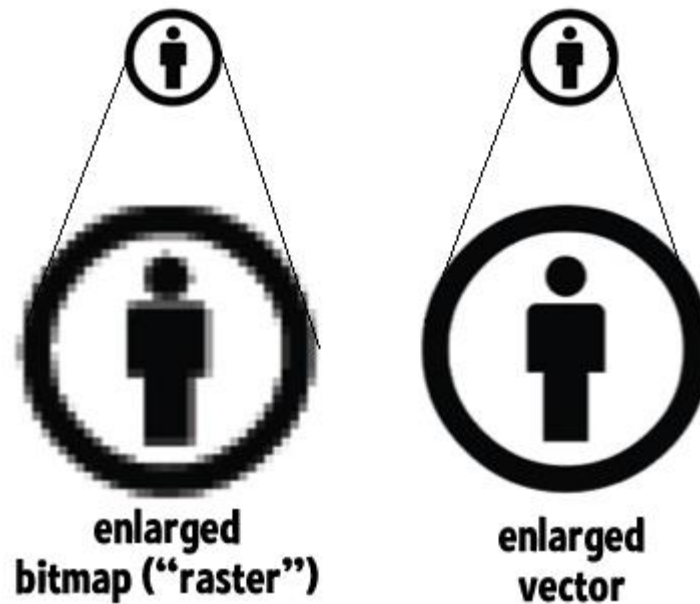
Vector Graphics

- Created in programs such as Adobe Illustrator and Corel Draw
- Use object specifications and mathematical equations to describe shapes to which colors are applied.
- Vector graphic images are appropriate for cleanly delineated shapes and colors, like cartoon or poster pictures.

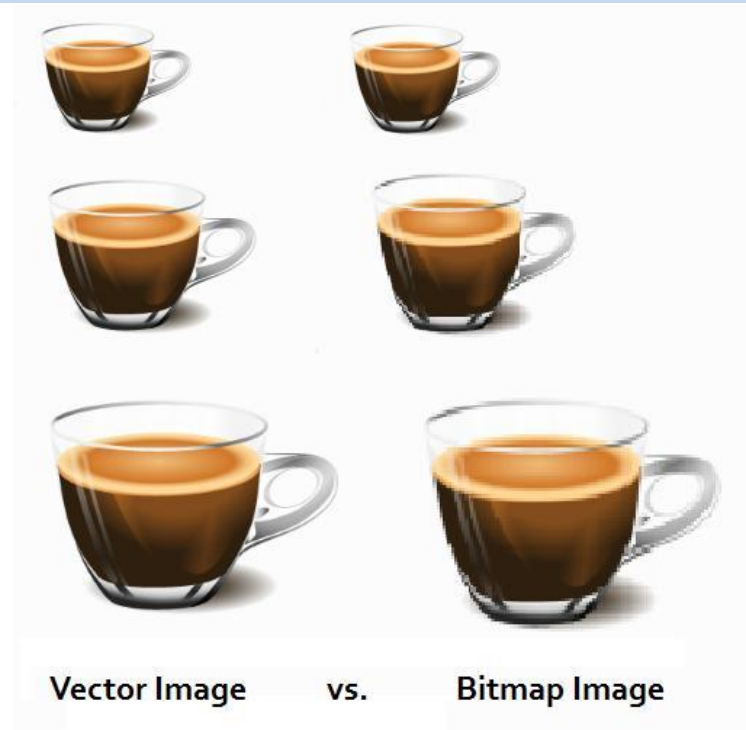
Vector formats

- As opposed to the raster image formats - (where the data describes the characteristics of each individual pixel), vector image formats contain a geometric description which can be rendered smoothly at any desired display size.
- At some point, all vector graphics must be rasterized in order to be displayed on digital monitors. Vector images may also be displayed with analog CRT technology such as that used in some electronic test equipment, medical monitors, radar displays, laser shows and early video games.
- Plotters are printers that use vector data rather than pixel data to draw graphics.





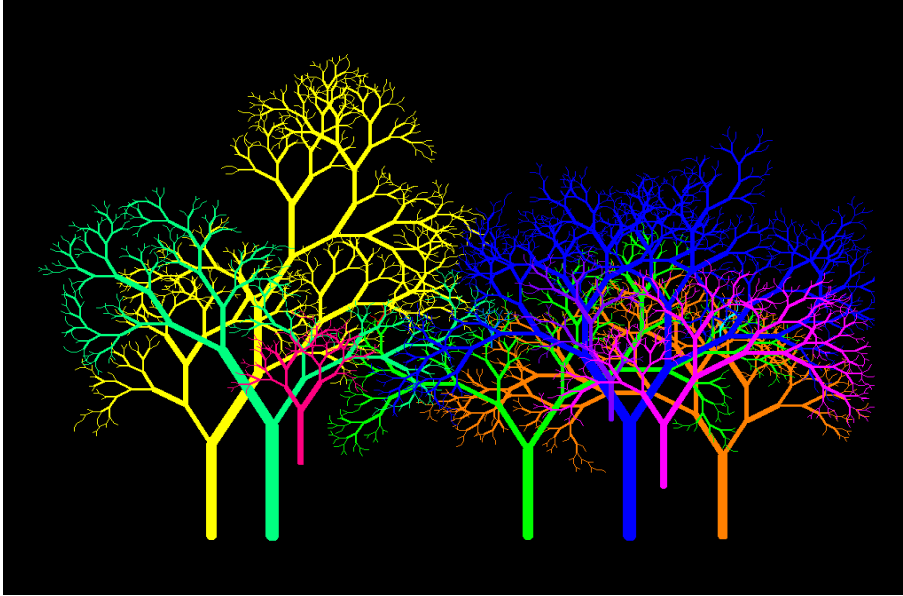
Comparison of Bitmap and Vector



Comparison of Bitmap and Vector

Procedural Modeling

- Also known as **algorithmic art** because of its aesthetic appeal.
- Procedurally modeled images: generate complex patterns, shapes, and colors in non-intuitive ways.
- Computer program uses some combination of mathematics, logic, control structures, and recursion to determine the color of pixels and content of the overall picture.
- Examples: Fractals and Fibonacci spirals.



A fractal is a graphical image characterized by a recursively repeating structure. For fractals, this self-embedded structure can be repeatedly infinitely.

Digital Image File Types

File Suffix	Our Abbreviation	File Type	Characteristics
Bitmap Images			
<i>.bmp</i>	BMP	Windows bitmap	1 to 24 bit color depth, 32-bit if alpha channel is used. Can use lossless RLE or no compression. RGB or indexed color.
<i>.gif</i>	GIF	Graphics Interchange Format	Used on the web. Allows 256 RGB colors. Can be used for simple animations. Uses LZW compression. Originally proprietary to CompuServe.
<i>.jpeg</i> or <i>.jpg</i>	JPEG	Joint Photographic Experts Group	For continuous tone pictures. Lossy compression. Level of compression can be specified.
<i>.png</i>	PNG	Portable Network Graphics	Designed as an alternative to <i>.gif</i> files. Compressed with lossless method. 1 to 64 bit color with transparency channel.
<i>.psd</i>	PSD	Adobe Photoshop	Supports a variety of color models and bit depths. Saves image layers created in photographic editing.
<i>.psp</i>	PSP	Corel Paint Shop Pro	Similar to <i>.psd</i> .
<i>.raw</i>		Photoshop	Uncompressed raw file. Could be black and white, grayscale, or RGB color.
<i>.tif</i> or <i>.tiff</i>	TIFF	Tagged Image File Format	Often used for traditional print graphics. Can be compressed with lossy or lossless methods, including RLE, JPEG, and LZW. Comes in many varieties.

Vector Graphics			
<i>.ai</i>	AI	Adobe Illustrator	Proprietary vector format.
<i>.swf</i>	SWf	Shockwave Flash	Proprietary vector format; can contain stills, animations, video, and sound.
<i>.cdr</i>	CDR	Corel Draw	Proprietary vector format.
<i>.dxf</i>	DXF	AutoCAD ASCII Drawing Interchange Format	ASCII text stores vector data.
Metafiles			
<i>.cgm</i>	CGM	Computer Graphics Metafile	ANSI, ISO standard.
<i>.emf, .wmf</i>	EMF, WMF	Enhanced metafile and Windows metafile	Windows platform.
<i>.eps</i>	EPS	Encapsulated Postscript	Used for output to Postscript device.
<i>.pdf</i>	PDF	Portable Document Format	An open standard working toward ISO standardization. Windows, MAC, Unix, Linux
<i>.pict</i>	PICT	Picture	Macintosh. Can use RLE or JPEG compression. Grayscale, RGB, CMYK, or indexed color.
<i>.wmf</i>	WMF	Windows metafile	16-bit format. Can be binary or text. Not portable to other platforms.

JPEG/JFIF

- [JPEG](#) (Joint Photographic Experts Group) is a [lossy](#) compression method; JPEG-compressed images are usually stored in the [JFIF](#) (JPEG File Interchange Format) file format.
- Nearly every digital camera can save images in the JPEG/JFIF format, which supports eight-bit grayscale images and 24-bit color images (eight bits each for red, green, and blue).
- JPEG applies lossy compression to images, which can result in a significant reduction of the file size.
- Applications can determine the degree of compression to apply, and the amount of compression affects the visual quality of the result.

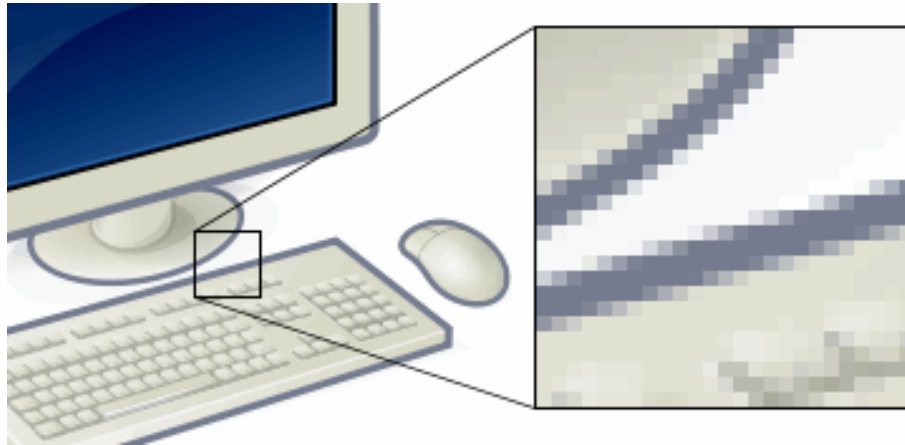
GIF

- **GIF** ([Graphics Interchange Format](#)) is in normal use limited to an 8-bit palette, or 256 colors (while 24-bit color depth is technically possible).
- GIF is most suitable for storing graphics with few colors, such as simple diagrams, shapes, logos, and cartoon style images, as it uses lossless compression, which is more effective when large areas have a single color, and less effective for photographic images.
- Due to GIF's simplicity and age, it achieved almost universal software support. Due to its animation capabilities, it is still widely used to provide image animation effects, despite its low compression ratio compared to modern video formats.

Pixel

- When you display a bitmap image on your computer, the **logical pixel**—that is, the number representing a color and stored for a given position in the image file—is mapped to a **physical pixel** on the computer screen.
- For an image file, pixel dimensions is defined as the number of pixels horizontally (i.e., width, w) and vertically (i.e., height, h) denoted $w \times h$.
- For example, your digital camera might take digital images with pixel dimensions of 1600×1200 .
- Similarly, your computer screen has a fixed maximum pixel dimensions—e.g., 1024×768 or 1400×1050 .

Pixel



- Each pixel is a sample of an original image.
- More samples typically provide more accurate representations of the original.
- The intensity of each pixel is variable. In color imaging systems, a color is typically represented by three or four component intensities such as red, green, and blue, or cyan, magenta, yellow, and black.

Resolution

- Resolution is defined as the number of pixels in an image file per unit of spatial measure.
- For example, resolution can be measured in **pixels per inch** (ppi).
- It is assumed that the same number of pixels are used in the horizontal and vertical directions, so a 200 ppi image file will print out using 200 pixels to determine the colors for each inch in both the horizontal and vertical directions.

- Resolution of a printer is a matter of how many dots of color it can print over an area. A common measurement is ***dots per inch*** (DPI).
- For example, an inkjet printer might be able to print a maximum of 1440 DPI. The printer and its software map the pixels in an image file to the dots of color printed.
- There may be more or fewer pixels per inch than dots printed. Consider your printer's resolution when you create an image to be printed.

Image Size

- Image size is defined as the physical dimensions of an image when it is printed out or displayed on a computer, e.g., in inches or centimeters.
- By this definition, image size is a function of the pixel dimensions and resolution, as follows:

For an image with resolution r and pixel dimensions $w \times h$ where w is the width and h is the height, the printed image size will be $a \times b$ as given by

$$\begin{aligned}a &= w/r \\ \text{and} \\ b &= h/r\end{aligned}$$

For an image with pixel dimensions $w \times h$ where w is the width and h is the height displayed on a computer display with resolution r at 100% magnification, the displayed image size will be $a \times b$ as given by

$$\begin{aligned}a &= w/r \\ \text{and} \\ b &= h/r\end{aligned}$$

(It is assumed that the display resolution is given in pixels per inch or centimeters per inch.)

When display on a computer

Resampling

- Changing the number of pixels in an image is called ***resampling***.
- Increase the pixel dimensions by ***upsampling*** or decrease the dimensions by ***downsampling***.
- Resampling always involves some kind of interpolation, averaging, or estimation, and thus it cannot improve the quality of an image.
- The additional pixels created by upsampling are just “estimates” of what the original pixel values would have been if the original image captured at higher pixel dimensions.
- Pixel values from downsampling are just averages of the information originally captured.

Resampling

- In the case of digital imaging, if the sampling rate is too low, then the image takes on a shape or pattern different what was actually being photographed—blockiness, blurriness, jagged edges, or moiré patterns (patterns that are created from two other patterns overlapping each other at an angle).
- Will get back in bitmap later (color representation and frequency transform).

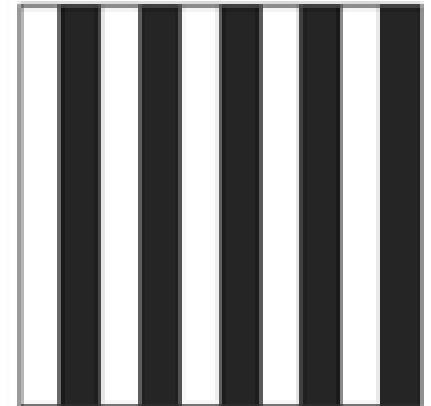
Aliasing

- Imagine that this picture has been divided into sampling areas so that only 15 samples are taken across a row.
- If the color changes even one time within one of the sample areas, then the two colors in that area cannot both be represented by the sample.
- Mathematically speaking, the spatial frequencies of the original scene will be aliased to lower frequencies in the digital photograph. Visually, the reconstructed image looks **blocky** and the **edges of objects are jagged**.

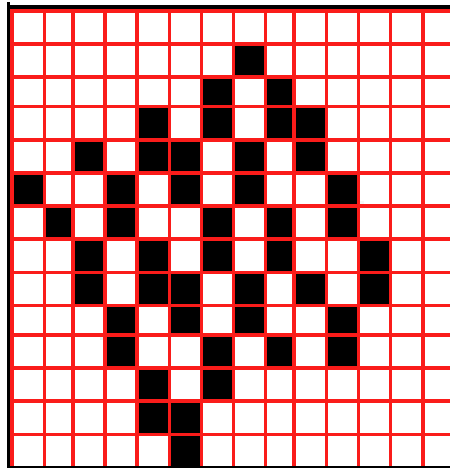
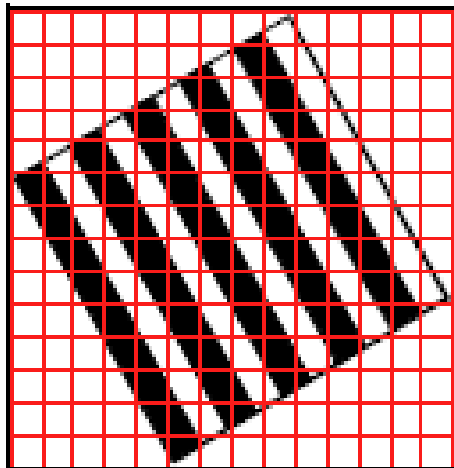


Moiré Effect

- Occur when there is a pattern in the image being photographed, and the sampling rate for the digital image is not high enough to capture the frequency of the pattern.
- a different pattern will result in the reconstructed image
- In the figure, the color changes at a perfectly regular rate, with a pattern that repeats five times in the horizontal direction. What would happen if we sampled this image five times, at regularly spaced intervals?
- Depending on where the sampling started, the resulting image would be either all black or all white.

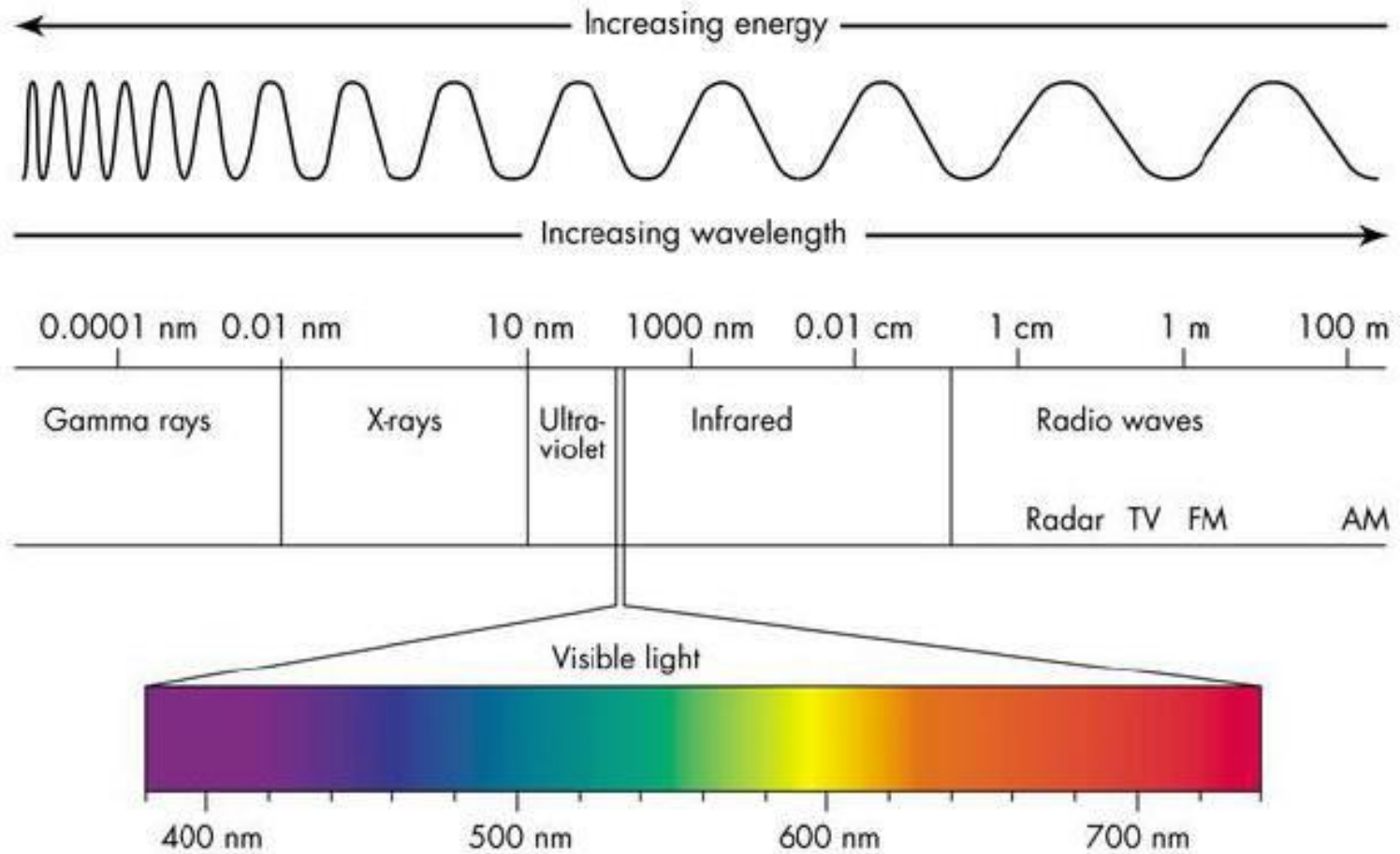


- Moiré effects can result when the original pattern is more complex and the pattern is tilted at an angle with respect to the sampling.
- Imagine the image that would result from tilting the original striped picture and then sampling in the horizontal and vertical directions, as shown in Figure.
- The red grid shows the sampling blocks. Assume that if more than half a sampling block is filled with black from the original striped image, then that block becomes black. Otherwise, it is white.

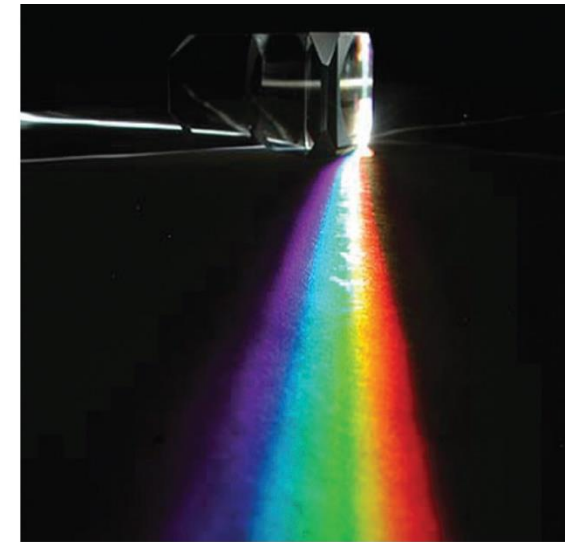
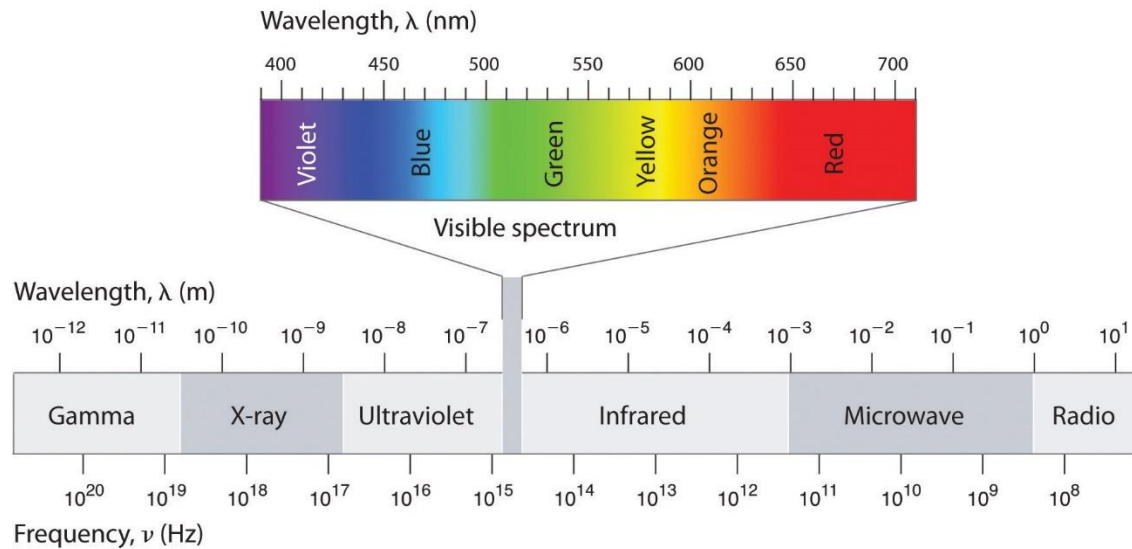


Color Representation & Models

- Color can be described as the characteristic of a light that emitted from a source or reflected from a surface.
- Color is both a **physical** and **psychological** phenomenon - composed of electromagnetic waves.
- For humans, the wavelengths of visible colors fall between approximately 370 nm and 780 nm.
- These waves fall upon the color receptors of the eyes, the human brain translates the **interaction between the waves and the eyes** as color perception.



Electromagnetic Spectrum and Visible Light



(b)

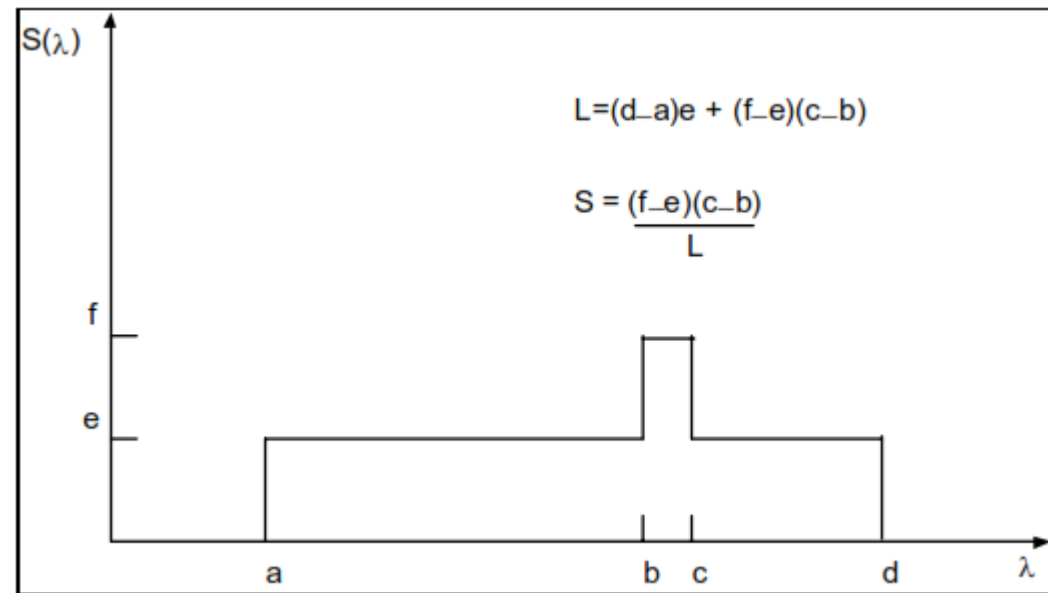
Electromagnetic Spectrum and Visible Light

Spectral Density Function

- the colors we see around us are almost always produced by a combination of wavelengths.
- A spectrograph breaks up a color into its component wavelengths, producing a spectral density function $P(\lambda)$.
- A spectral density function shows the contributions of the wavelengths λ to a given perceived color as λ varies across the visible spectrum.

- Spectral density functions are one mathematical way to represent colors, but not a very convenient way for computers.
- One problem is that two colors that are perceived to be identical may, on analysis, produce different spectral density curves.
- If we want to use a spectral density curve to tell a computer to present a particular shade of green, which “green” spectral density curve is the best one to use?

- a simpler spectral density graph - each color in the spectrum can be characterized by a unique graph that has a simple shape
- However, the dimensions of hue, saturation, and brightness do not correspond very well to the way computer monitors are engineered.



Spectral density graph showing hue, saturation, and lightness

The graph for each color gives the color's *dominant wavelength, equivalent to the hue; its saturation (i.e., color purity); and its luminance.*

Power & Energy of a light

Two colored lights can be of the same wavelength but of different power. A light's *power, or energy per unit time, is a physical property not defined by human perception*. Power is related to brightness in that if two lights have the same wavelength but the first has greater power, then the first will appear brighter than the second.

Brightness

subjective perception and has no precise mathematical definition

Luminance

mathematical definition that relates a light's wavelength and power to how bright it is perceived to be.

Interestingly, lights of equal power but different wavelengths do not appear equally bright. The brightest wavelengths are about 550nm.

RGB Color System

- The most suitable for display.
- Referring to three primary colors, which are red (R), green (G) and blue (B).
- 24-bit RGB computer monitor with each primary color has 8-bit data or 256 discrete values.
- Simplicity in producing color as it only involves process of **additive** color mixture.

- Red, green, and blue are good choices as primary colors because the cones of the eyes—the colors receptors—are especially sensitive to these hues.

$$C = rR + gG + bB$$

- where r , g , and b indicate the relative amounts of red, green, and blue energy respectively. R , G , and B are constant values based on the wavelengths chosen for the red, green and blue components.
- The values r , g , and b are referred to as the values of the RGB **color components** (also called **color channels** in application programs).

RGB Color space

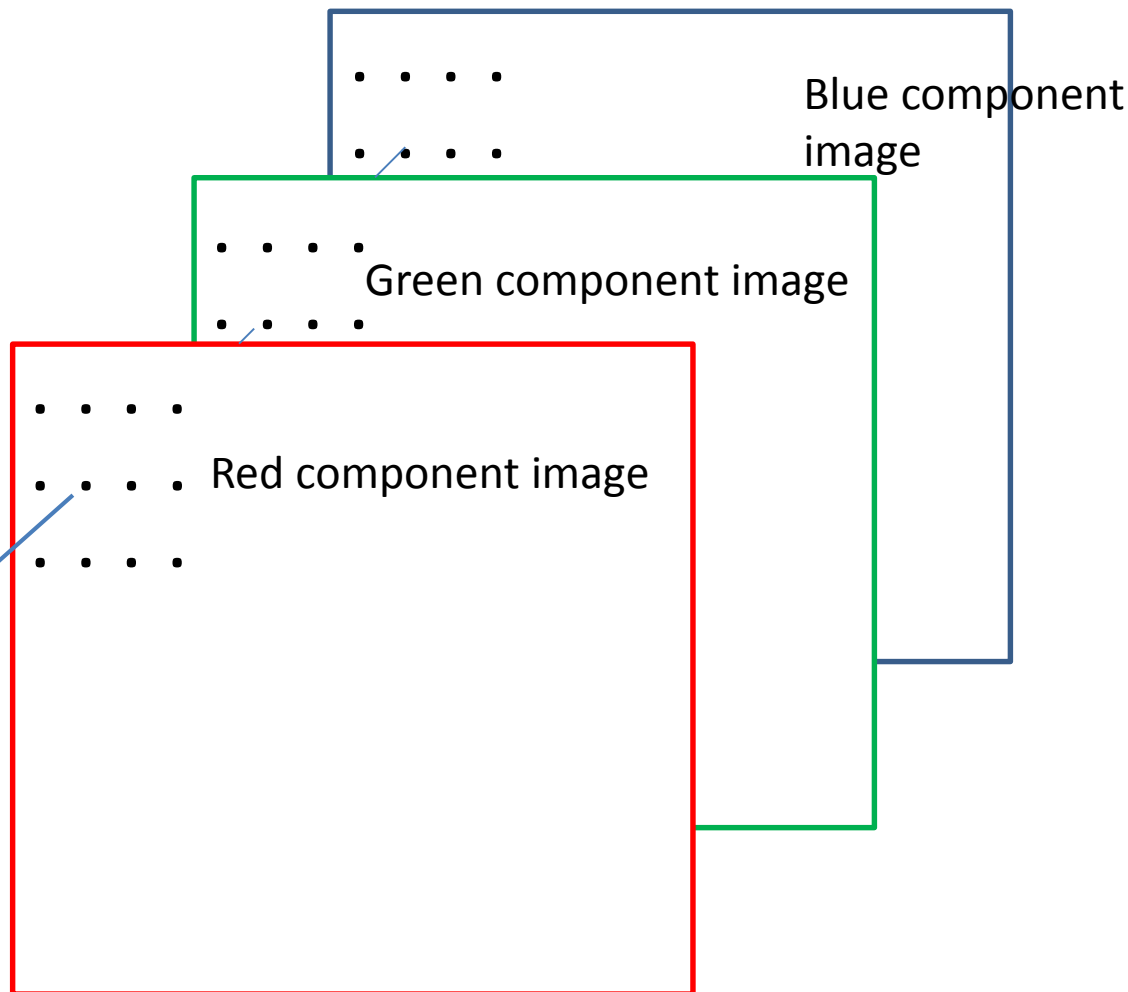
- An additional channel, the alpha channel, is used to specific the opacity value of a color pixel.

8	8	8	8
Alpha	Red	Green	Blue
Bit 31-24	Bit 23-16	Bit 15-8	Bit 7-0

RGB Pixel & Memory layout

Three color components of a color pixel.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



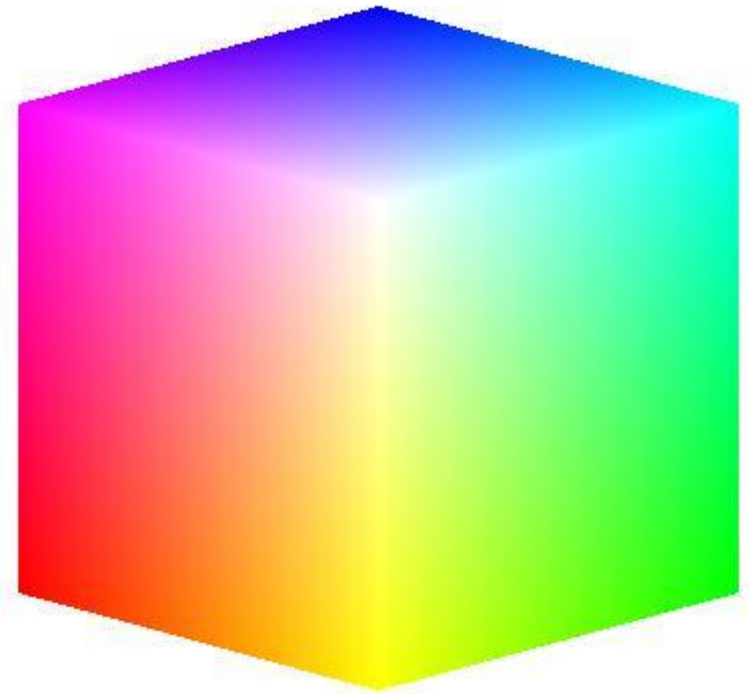
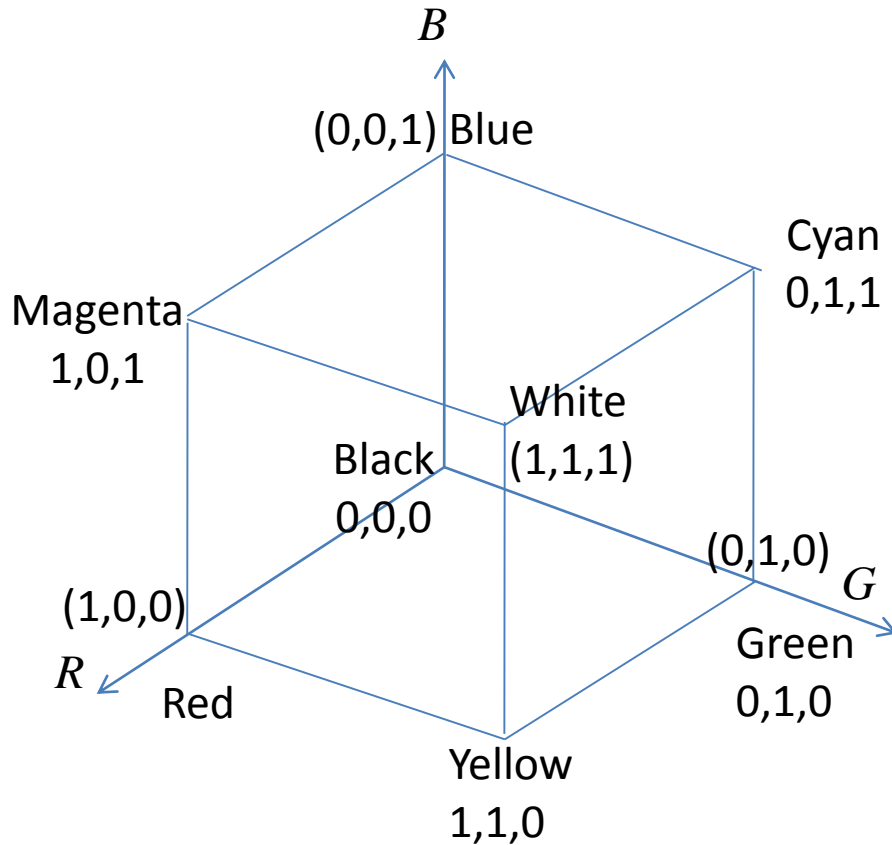
Memory storage representation:



Can also be stored in this way:



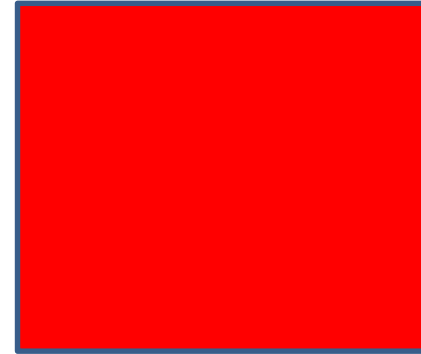
RGB color cube



For example, the light orange described as $(1, 0.65, 0.15)$ above would become $(255, 166, 38)$ in an RGB mode with maximum values of 255.

Creating a RGB image

```
comp1 = zeros(100,100);  
comp2 = zeros(100,100);  
comp3 = zeros(100,100);  
comp1(1:100,1:100)=255;  
imshow(uint8(cat(3,comp1,comp2,comp3)))  
figure  
imshow(uint8(cat(3,comp2,comp1,comp3)))
```



Creating a RGB image

```
comp1 = zeros(100,100);  
comp2 = zeros(100,100);  
comp3 = zeros(100,100);  
comp1(1:100,1:100)=255;  
comp2(1:100,1:100)=166;  
comp3(1:100,1:100)=38;
```

```
figure  
imshow(uint8(cat(3,comp1,comp2,comp3)))
```



Conversion

- Let an RGB color pixel be given by (R, G, B), where R, G, and B are the red, green, and blue color components, respectively. Then the corresponding **grayscale** value is given by (L, L, L), where

$$L = 0.30R + 0.59G + 0.11B$$

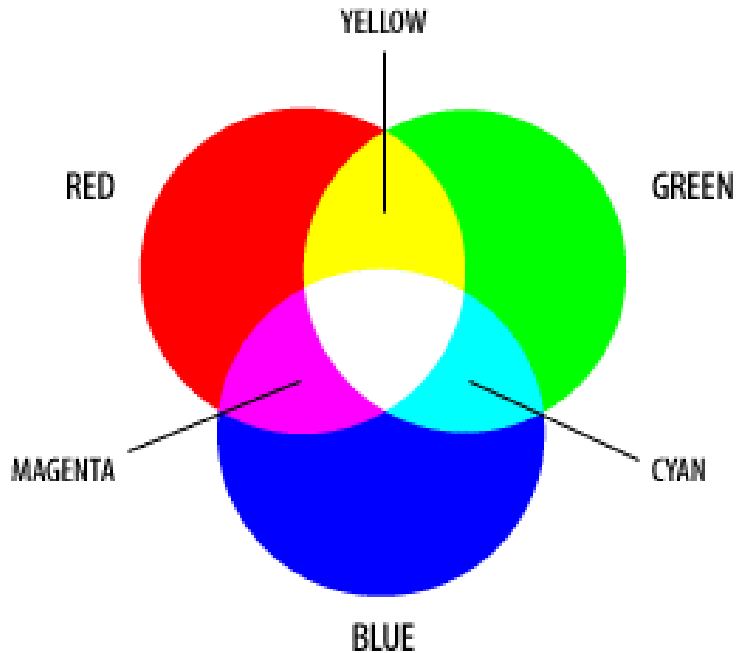
- Since all three color components are equal in a gray pixel, only one of the three values needs to be stored. Thus a 24-bit RGB pixel can be stored as an 8-bit grayscale pixel.

CMYK Color System

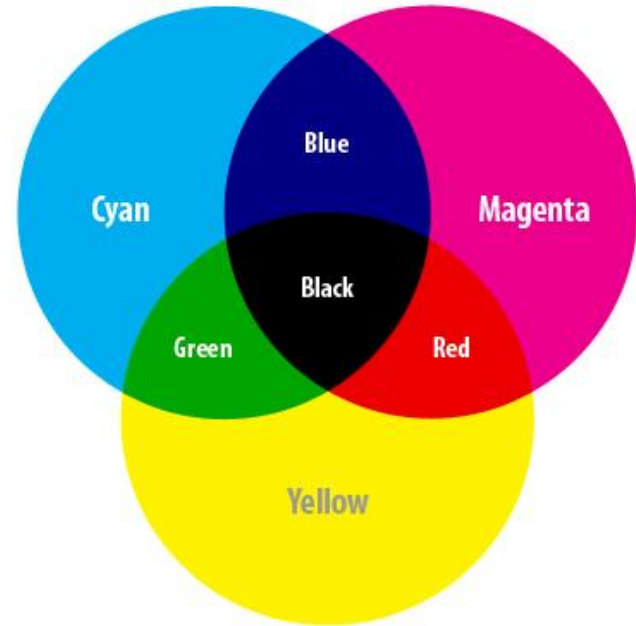
- Defined based on **subtractive** mixture.
- Designed for printing purpose.
- **C** is referring to cyan, **M** for magenta and **Y** for yellow. **K** for black is added to influence the printing brightness. Besides, text is typically printed in black.
- To convert the RGB color system to CMYK color system, first the RGB is convert to CMY and then a suitable **K** value is determine based on the CMY values.

- **RGB is a *additive, projected light* color system.** All colors begin with black "darkness", to which different color "lights" are added to produce visible colors. RGB "maxes" at white, which is the equivalent of having all "lights" on at full brightness (red, green, blue).
- **CMYK is a *subtractive, reflected light* color system.** All colors start with white "paper", to which different color "inks" are added to absorb (subtract) light that is reflected. In theory, CMY are all you need to create black (applying all 3 colors at 100%). Alas, that usually results in a muddy, brownish black, so the addition of K (black) is added to the printing process. It also makes it easier to print black text (since you don't have to register 3 separate colors).

RGB and CMYK Model



(a) RGB



(b) CMYK

Additive and subtractive color system, (a) RGB, (b) CMYK

CMY – RGB - CMY

- For a pixel represented in RGB color, the red, green, and blue color components are, respectively, R , G , and B . Then the equivalent C , M , and Y color components are given by

$$C = 1 - R$$

$$M = 1 - G$$

$$Y = 1 - B$$

- Similarly, RGB values can be computed from CMY values with

$$R = 1 - C$$

$$G = 1 - M$$

$$B = 1 - Y$$

(The values can be given in the range of [0 255] or normalized to [0 1].)

CMYK

For a pixel represented in the CMY color model, the cyan, magenta, and yellow color components are, respectively, C , M , and Y . Let K be the minimum of C , M , and Y . Then the equivalent color components in the CMYK model,

C_{new} , M_{new} , Y_{new} , and K are given by

$$K = \min(C, M, Y)$$

$$C_{new} = C - K$$

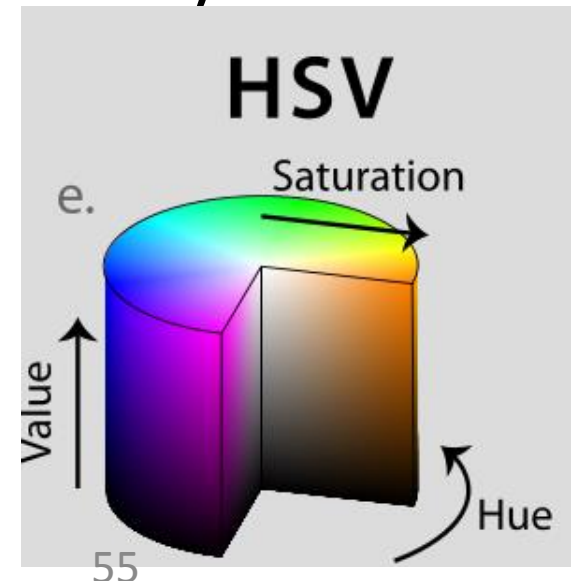
$$M_{new} = M - K$$

$$Y_{new} = Y - K$$

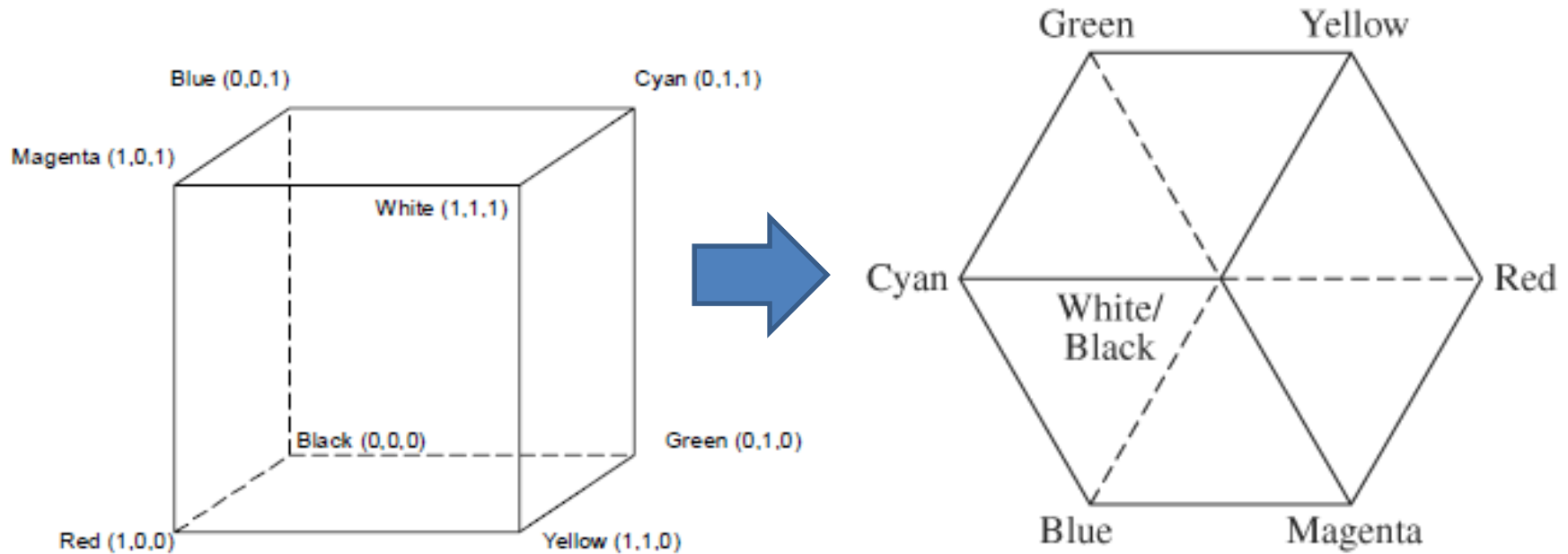
HSV color space

- Hue, Saturation and Value (HSV) system considerably closer than the RGB system to the way in which humans experience and describe color sensation.
- Hue -> tint (the essential color)
- Saturation -> shade (the purity of the color)
- Value -> tone (luminance)

H and S represent chromaticity while V both describing luminance.



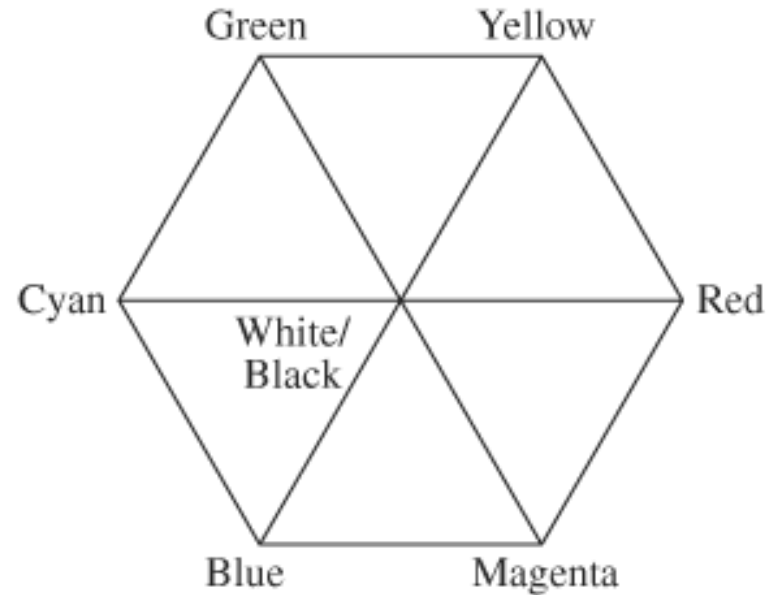
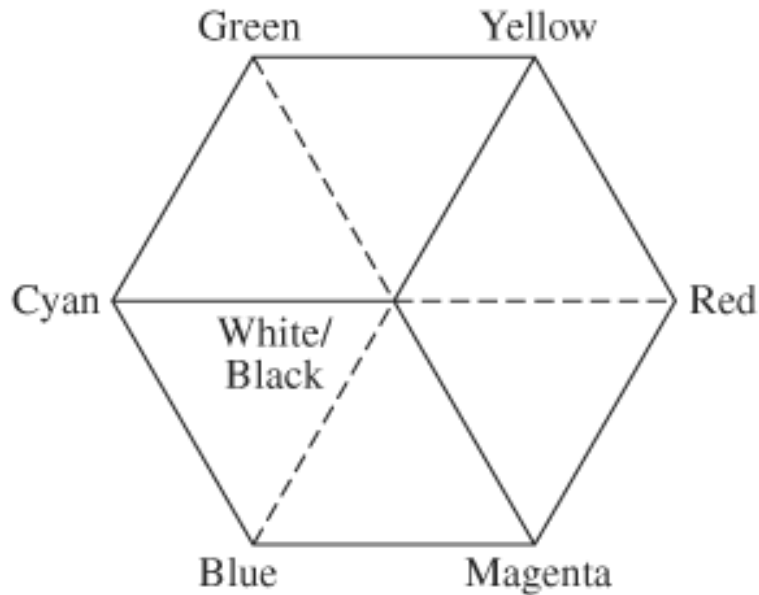
RGB to HSV



- Turn RGB cube around, tilt it, so that you looking straight into origin (white/black)

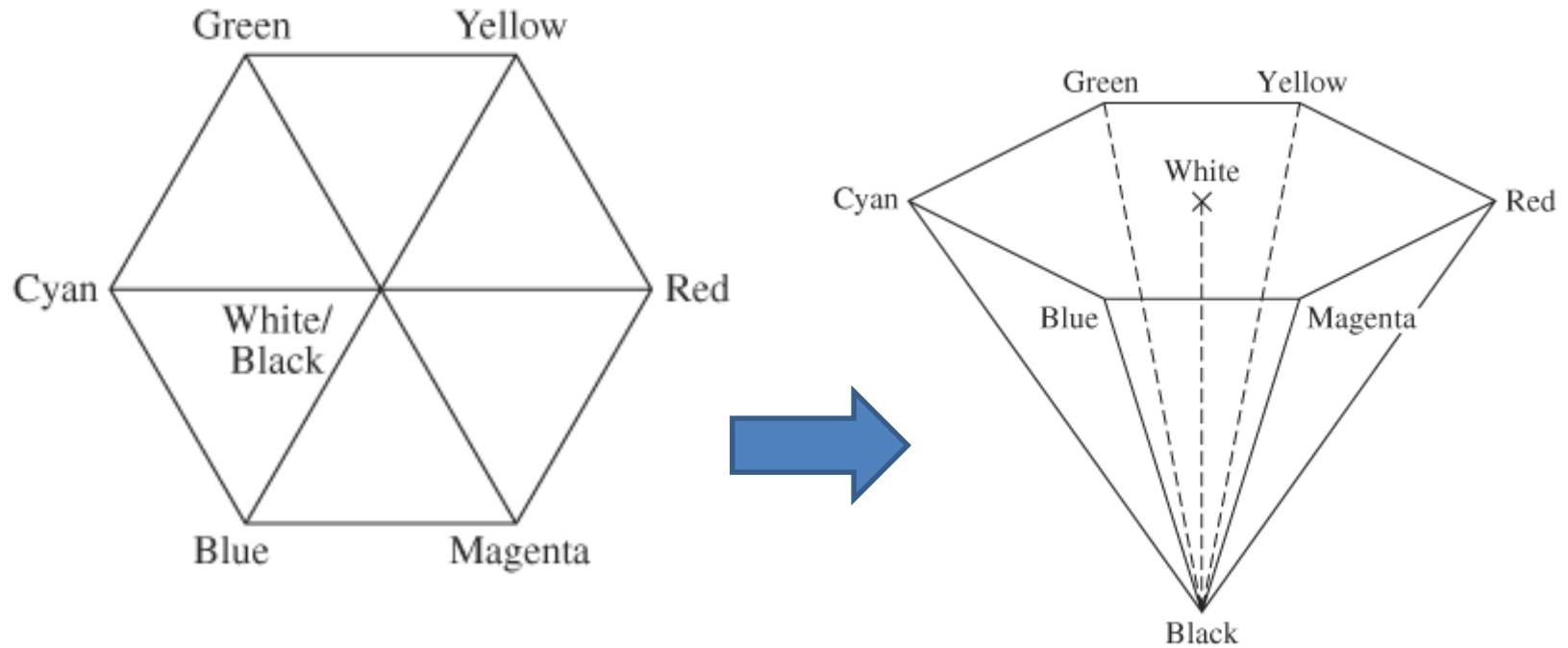
- RGB color cube viewed from the top

RGB to HSV



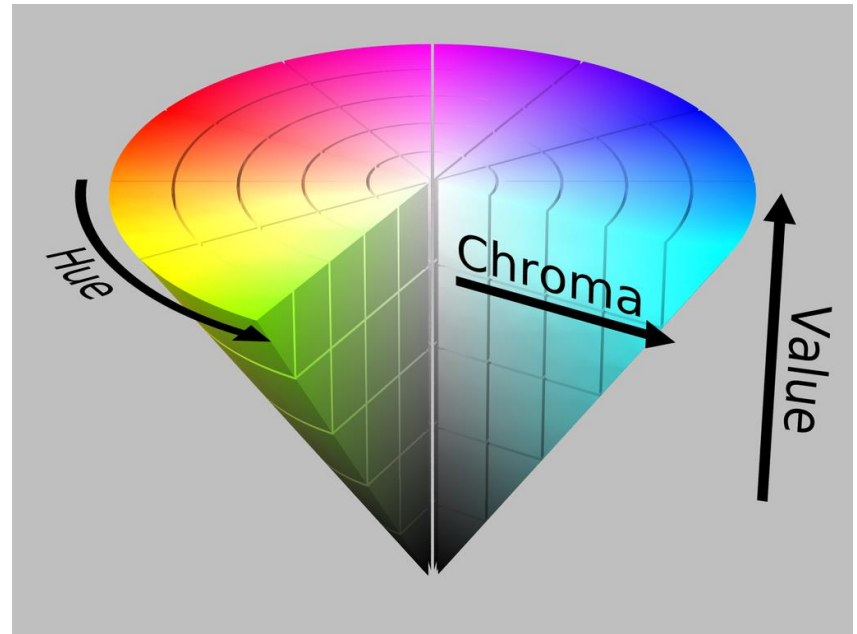
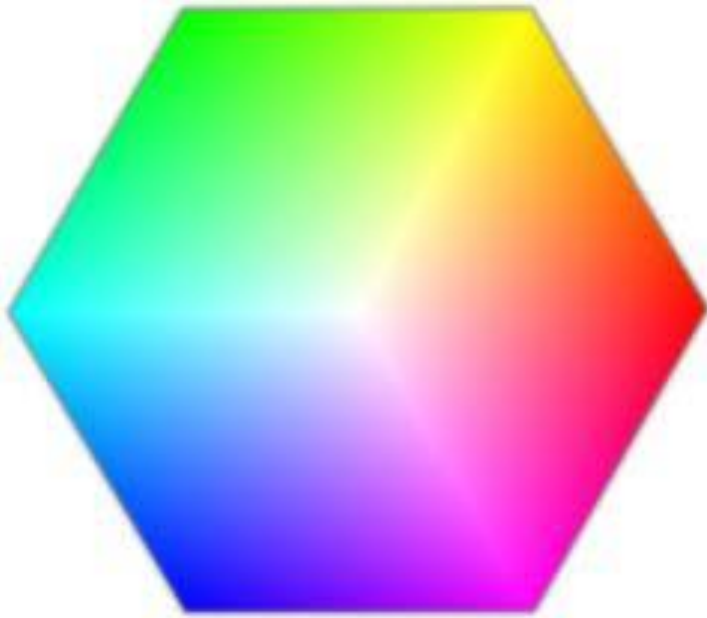
- RGB color cube collapsed to 2D

HSV Color Space



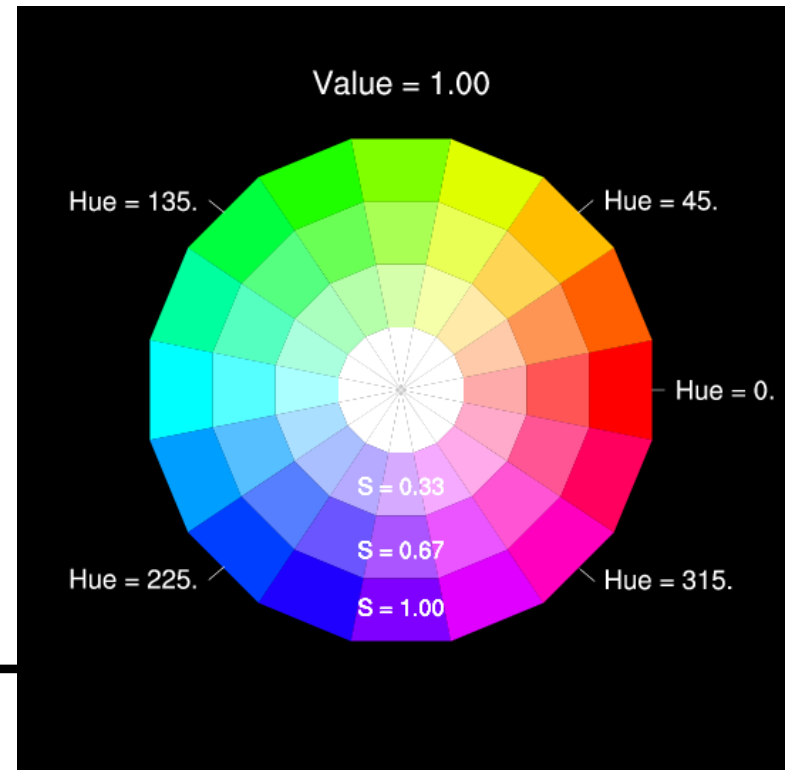
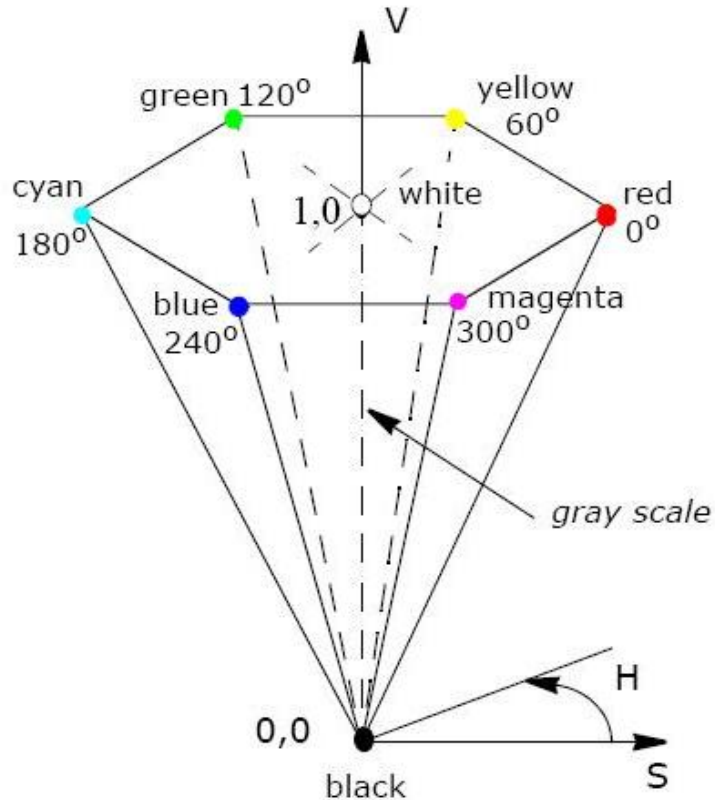
- Now, expand into 3D by pulling down on the center point.

HSV



Source: http://en.wikipedia.org/wiki/HSL_and_HSV

- **Hue:** Position (degree)
- **Saturation:** Color's distance from central axis
- **Value (axis):** from black (0)-tip, to 1 (white)-surface



Hue represented by a position on the circle in degree, from 0 to 360, with red set at 0.

60

For example, what is the color of $(58^\circ, 0.88, 0.93)$?

* RGB to HSV/HLS is a non-linear transformation



↑ Lightness increases in this direction. Along the axis, black is at the bottom, white at the top

Saturation is 100% (pure colors) on the perimeter of the middle hexagon.

Hue is represented by the angle around the vertical axis.

Algorithm RGB to HSV

/* Input: r , g , and b , each real numbers in the range $[0...1]$.
Output: h , a real number in the range of $[0...360)$, except if $s = 0$,
in which case h is undefined. s and v are real numbers in the range
of $[0...1]$.*/

```
{  
    max = maximum(r,g,b)  
    min = minimum(r,g,b)  
    v = max  
    if max  $\neq$  0 then s = (max - min)/max  
    else s = 0  
    if s == 0 then h = undefined  
    else {  
        diff = max - min  
        if r == max then h = (g - b) / diff  
        else if g == max then h = 2 + (b - r) / diff  
        else if b == max then h = 4 + (r - g) / diff  
        h = h * 60  
        if h < 0 then h = h + 360  
    }  
}
```



YUV & YIQ

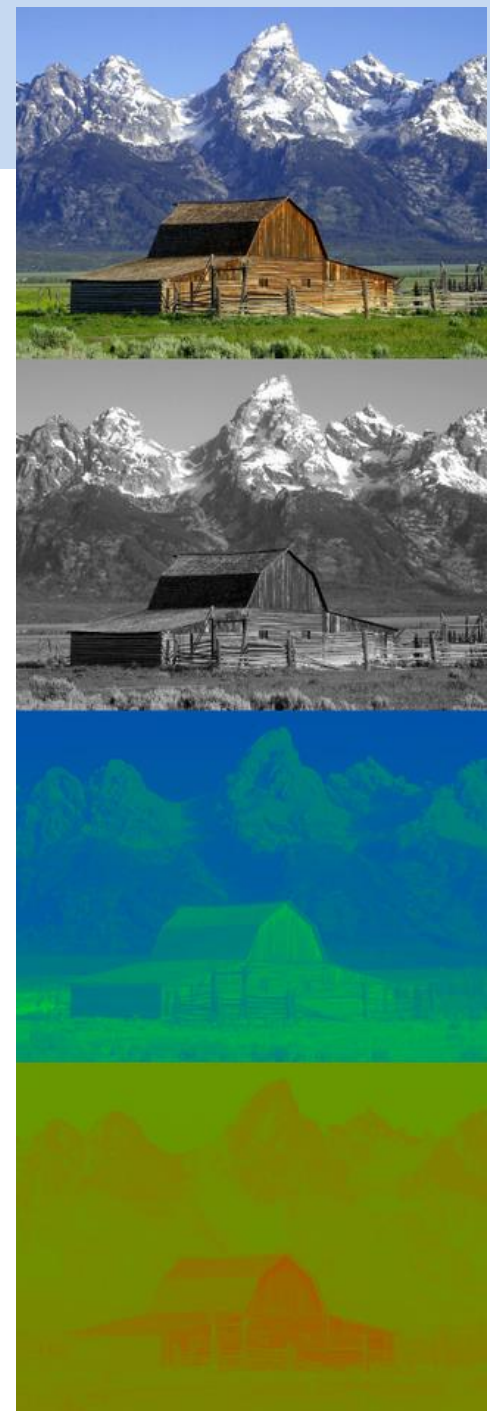
- Applied to color TV broadcasting
- Y – **luminance** component
- UV& IQ – **chromaticity** components: these are variable that changed by the brightness, color, contrast controls on a TV.
- Designed in order to compress RGB color while at the same time conserving the color bandwidth that suit human visual system.
- Not suitable for digital image processing.

RGB – YUV (PAL)

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

An image along with its Y', U, and V components respectively.

Source: <http://en.wikipedia.org/wiki/YUV>



RGB – YIQ (NTSC)

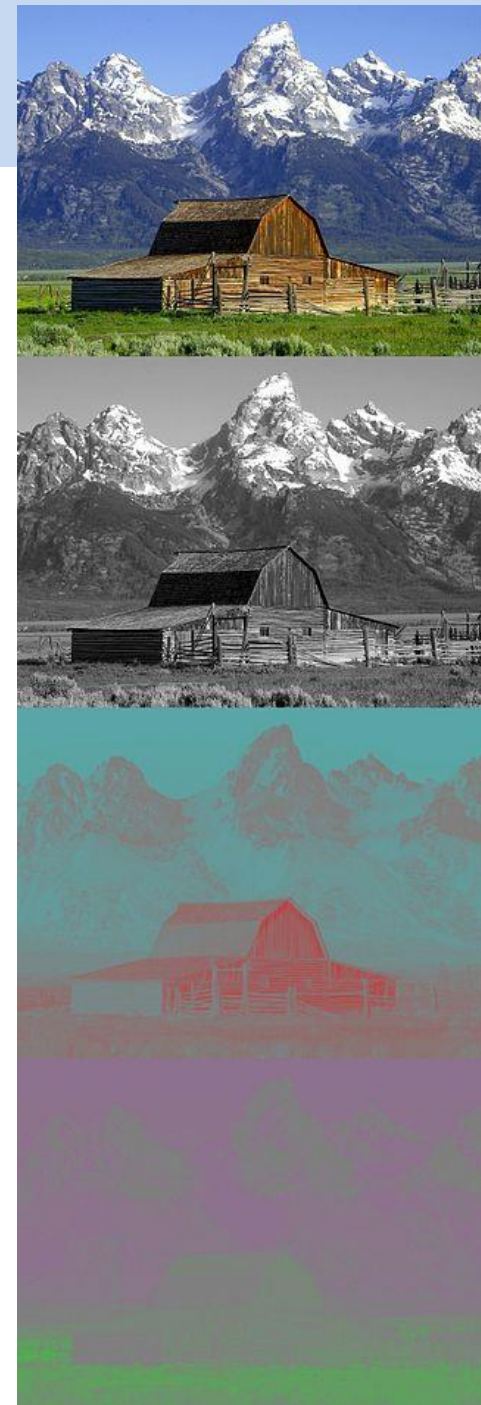
For a pixel represented in RGB color, let the red, green, and blue color components be, respectively, R , G , and B . Then the equivalent Y , I , and Q color components in the YIQ color model are given by

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

(Note that the values in the transformation matrix depend upon the particular choice of primaries for the RGB model.)

An image along with its Y , I , and Q components.

Source: <http://en.wikipedia.org/wiki/YIQ>

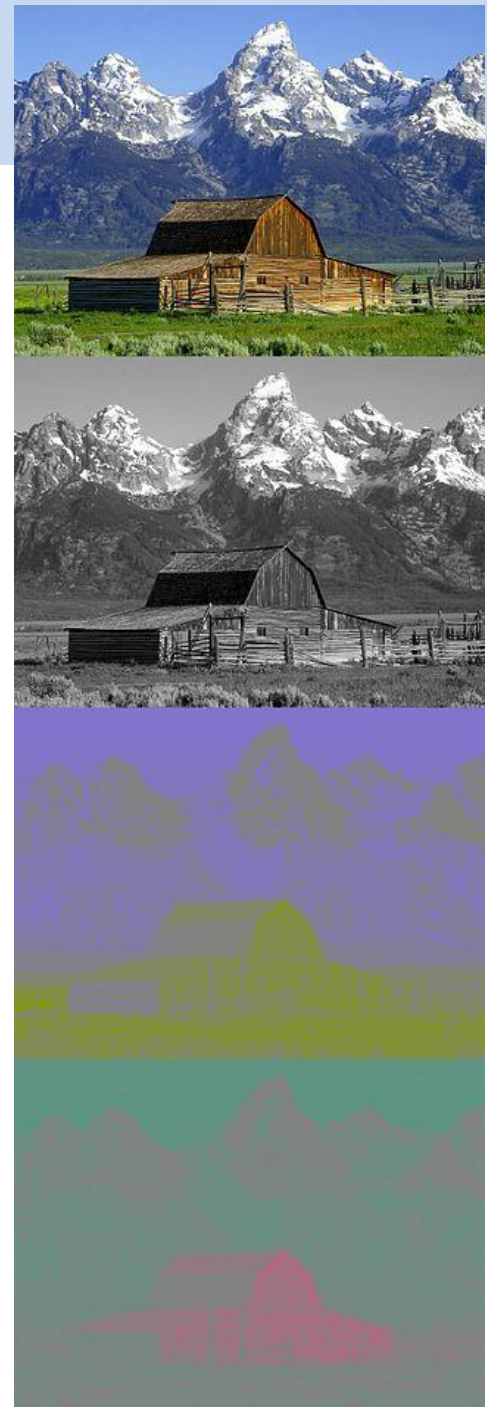


YCbCr Color Model

- Digital video – ITU-R BT. 601-4
- Used in JPEG image compression and MPEG video compression
- Closely related to YUV

A color image and its Y , C_B and C_R components. The Y image is essentially a greyscale copy of the main image.

Source: <http://en.wikipedia.org/wiki/YCbCr>



YCbCr Color Model

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

In practice, the resultant signals range from 16 to 219 for Y. The values from 0 to 15 are called *footroom*, while the values from 236 to 255 are called *headroom*.

C_B and C_R have a range of ± 112 and offset of +128. If R, G, B are floats in $[0 \dots 1]$, we obtain Y, C_B and C_R in $[0 \dots 255]$ via the transform

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$