Digital Image Processing

Lecture #11

Ming-Sui (Amy) Lee

Course Information

- **Following Schedule**

04/01	Lecture 5	05/20	Lecture 10
04/08	Lecture 6	05/27	Lecture 11
04/15	Lecture 7	06/03	Lecture 12
04/22	Midterm	06/10	Demo
04/29	Lecture 8	06/17	Demo
05/06	Proposal	06/24	Final Package Due
05/13	Lecture 9		

Announcement

- Final demo video
  - Due: 11:59 a.m. (noon) on Jun. 9, 2021
    - 10~12 minutes for each team
    - Video format: mpg, avi, mp4 or wmv
    - Provide a valid link on NTU COOL for us to download
    - DIP_Teamxx_FinalDemo.mpg/avi/mp4/wmv
  - Remember to include
    - Paper title / Motivation / Problem definition /
    - Algorithm / Experimental results
    - Reference

Color Image Processing

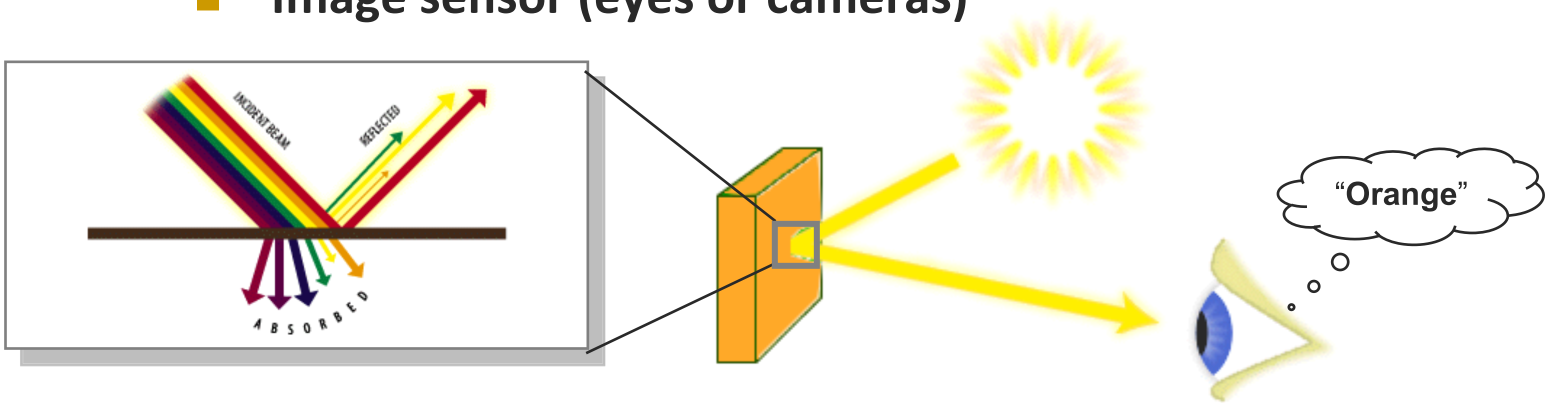
Color Image Processing

Introduction

- **Color Vision**
  - What is color
  - Visible spectrum
  - Spectral colors
  - Color perception
  - Color systems
  - Tri-stimulus colors
  - CIE primaries
  - Color mixing
- **Color models**
  - Overview
  - RGB
  - CMY/CMYK
  - YIQ/YUV
  - HSI/HSV
  - L*a*b
  - Transform
  - Comparison
- **Color image processing**
  - Individual vs. vector
  - Tonal correction
  - Color correction
  - Hist. equalization
  - Sharpening
  - Color quantization
  - Pseudo-coloring
  - Re-coloring
  - Multi-spectral images

Color Image Processing

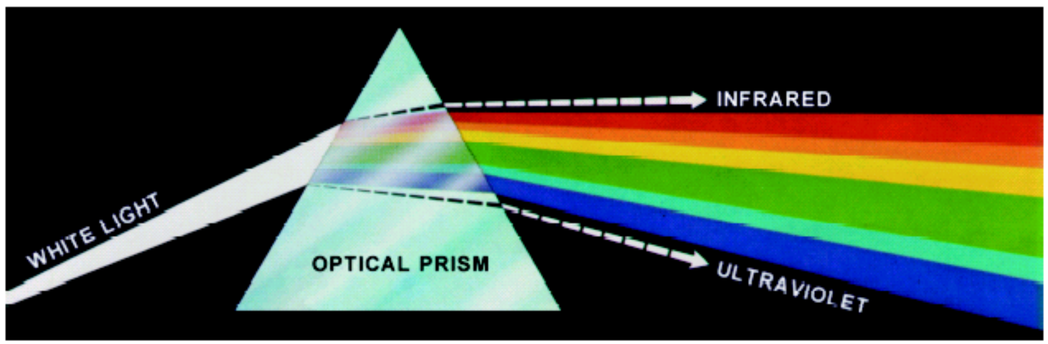
- Human perception, not directly measurable
  - Related to the light spectrum of a stimulus
  - Depends on
    - Light source
    - Reflectance (Reflecting objects)
    - Image sensor (eyes or cameras)



- Some or all of the light may be absorbed
- Dominate wavelength reflected by objects determines the color tone

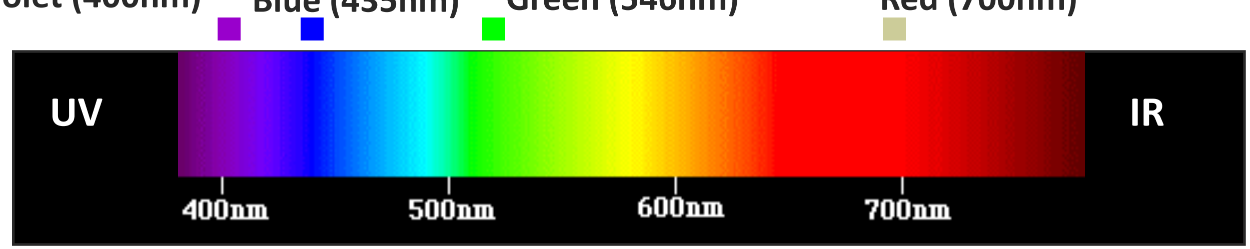
Color Image Processing

- Newton's experiment



- Visible spectrum

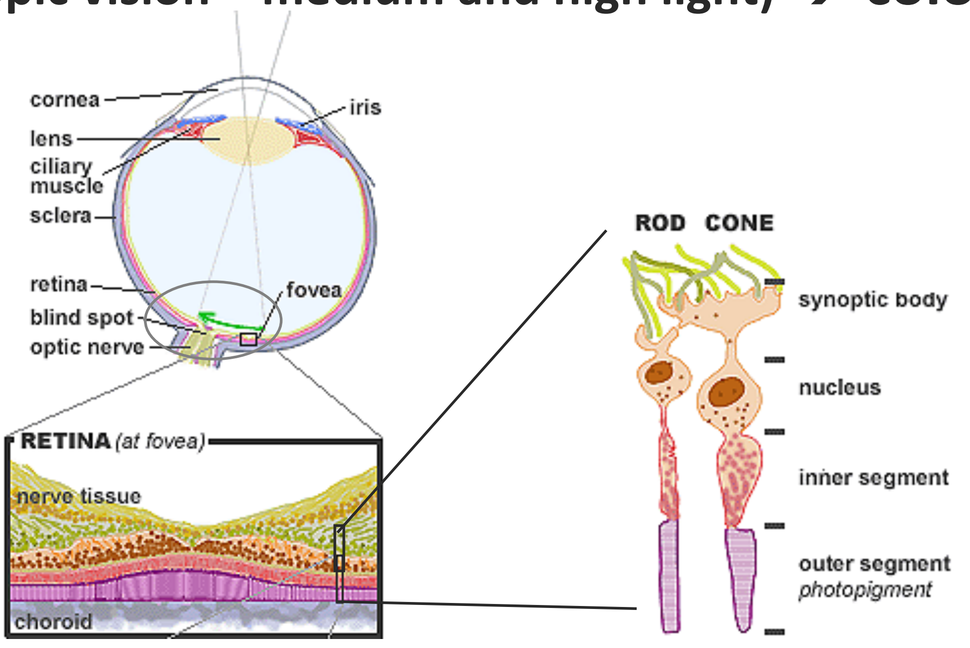
Violet (400nm) Blue (435nm) Green (546nm) Red (700nm)



Shorter wavelength wavelength (in nanometers) Longer wavelength

Color Image Processing

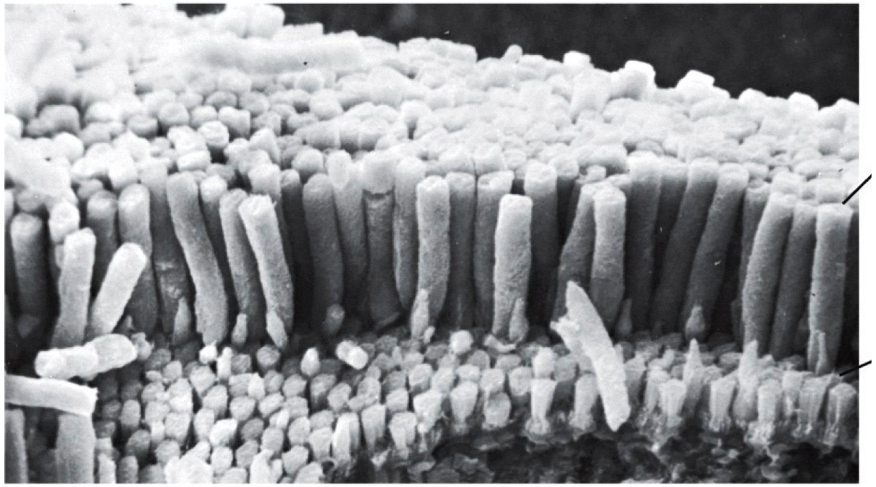
- Retina
  - a light-sensitive layer at the back of the eye
- Photosensitive cells
  - Rods (scotopic vision – low light) → luminance
  - Cones (photopic vision – medium and high light) → color



Rods: 75~150 millions
Cones: 6-7 millions

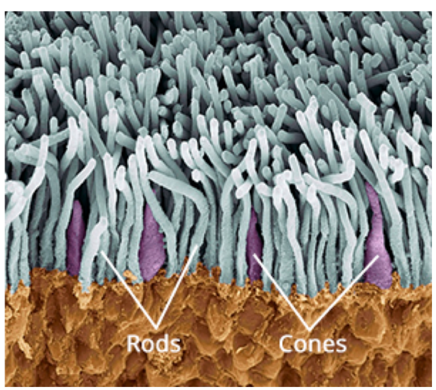
Color Image Processing

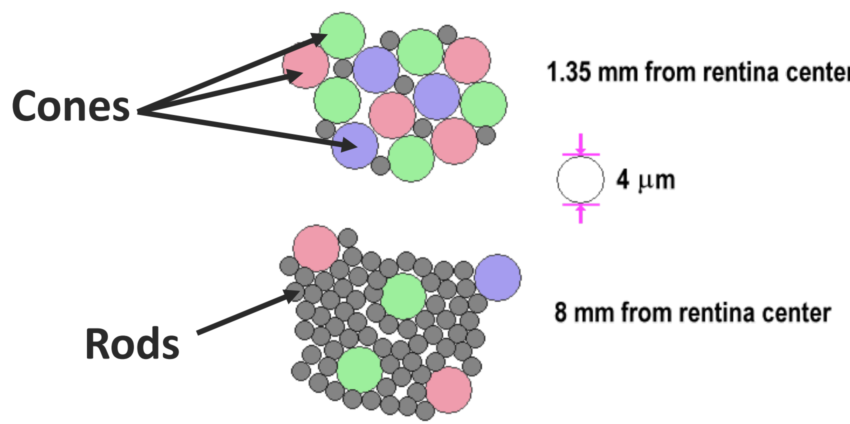
- Rods 
  - Low illumination levels (scotopic vision)
  - Provide our night vision ability
  - A thousand times more sensitive to light than cones
  - Much slower to respond to light than cones
  - Distribute primarily in the periphery of the visual field



Color Image Processing

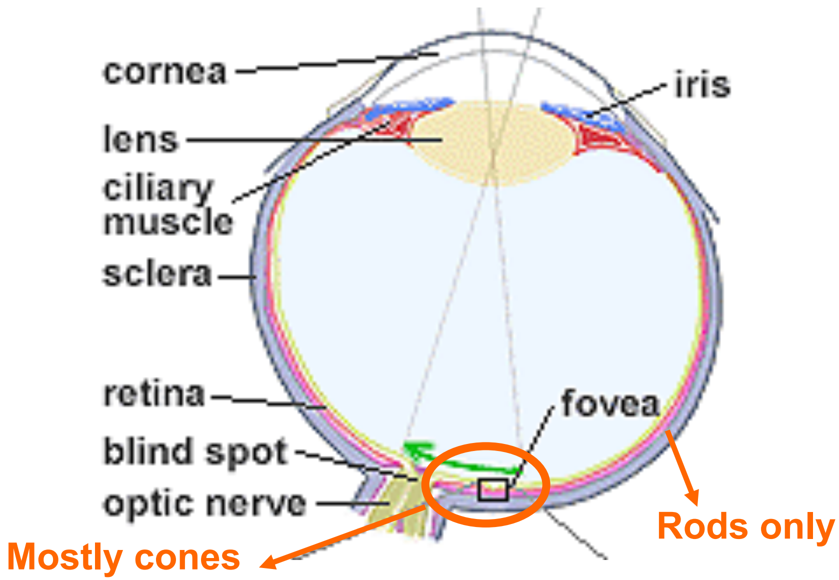
- **Cones** 
  - **High illumination levels (Photopic vision)**
  - **Less sensitive than rods**
  - **Not evenly distributed - density decreases with distance from fovea**

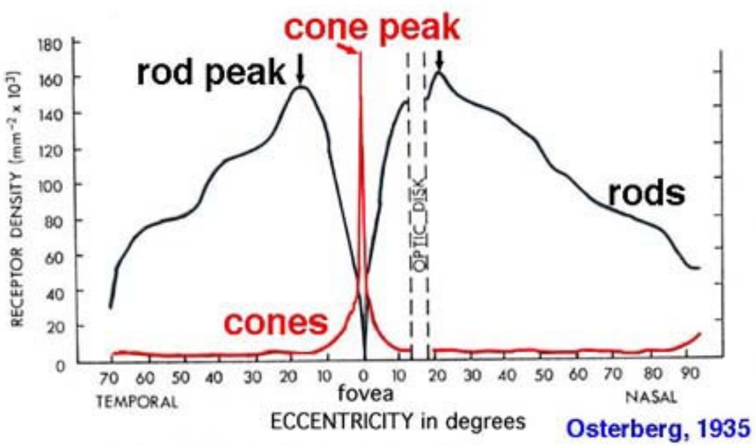




//Note// **Bayer filter**
use twice as many green elements as red or blue to mimic the human eye's greater resolving power with green light

Color Image Processing

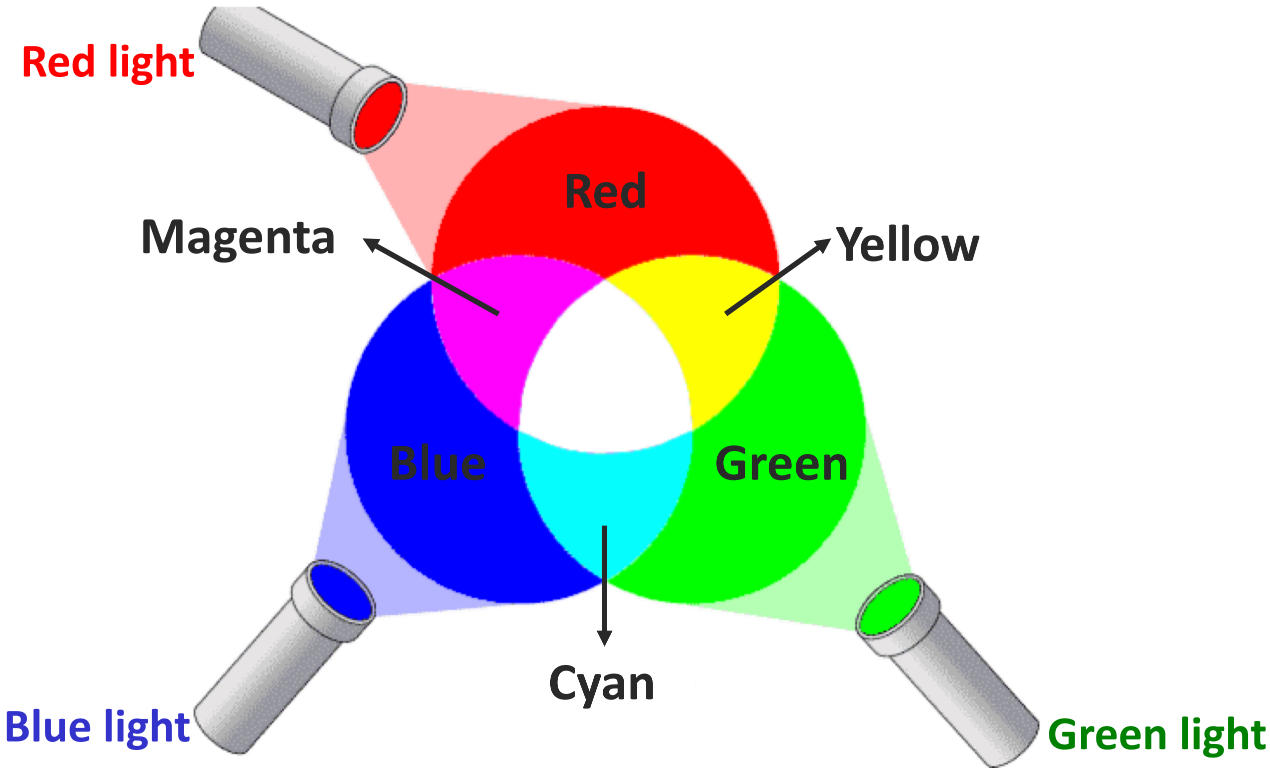




- Rod and cone densities
  - The peak number of cones occurs in the fovea
  - The rods peak about 20 degrees from the center
  - Rods: 75~150 (120) millions
  - Cones: 6~7 (7) millions

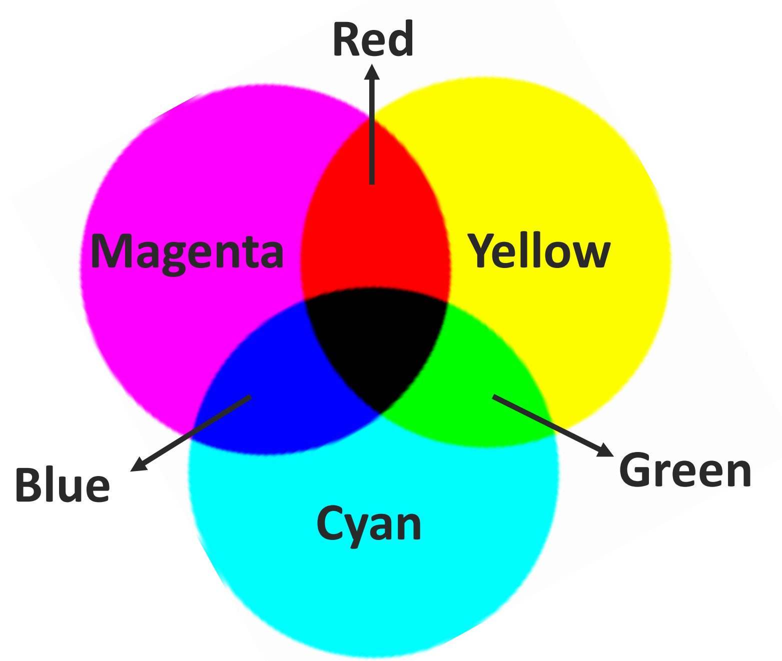
Color Image Processing

- **Additive Color Matching**
  - Primary colors can be combined to generate different composite colors



Color Image Processing

- **Subtractive Color Matching**
  - Composite color is the difference between two added colors
  - Used for painting and printing



Chromaticity

- Chromaticity (色度)

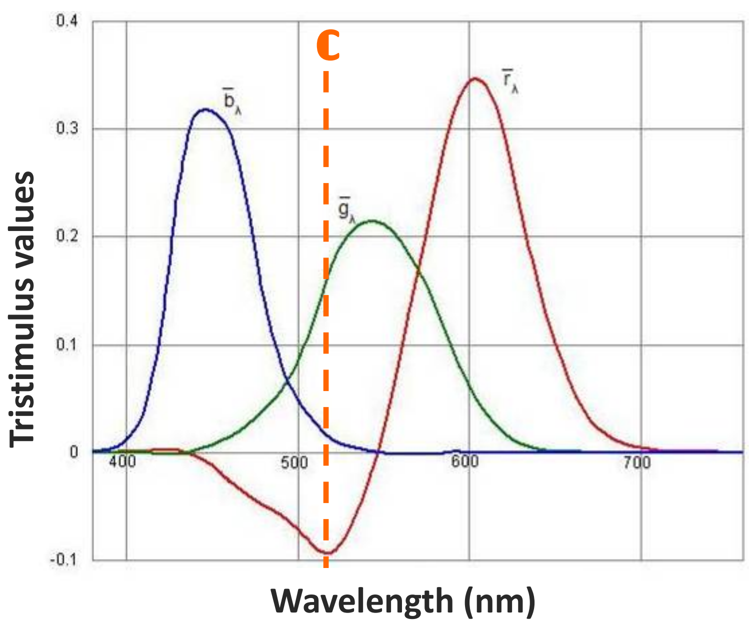
  = Hue (色調) + Saturation (飽和度)
  - Hue: dominant wavelength and color
  - Saturation: relative purity or the amount of white light mixed with a hue
- Color = Brightness + Chromaticity
- Tristimulus values

  the amount of R, G, and B needed to form any particular color : X, Y, Z
  - Trichromatic coefficients

$$x = X/(X+Y+Z) \qquad y = Y/(X+Y+Z) \qquad z = Z/(X+Y+Z)$$

Color Image Processing

- **Tristimulus**
  - **The primary sources recommended by CIE**
    - **CIE: International Commission on Illuminations 1931**
    - **700nm(red), 546.1nm(green) and 435.8nm(blue)**
    - **Be able to match all the wavelengths of the visible spectrum**



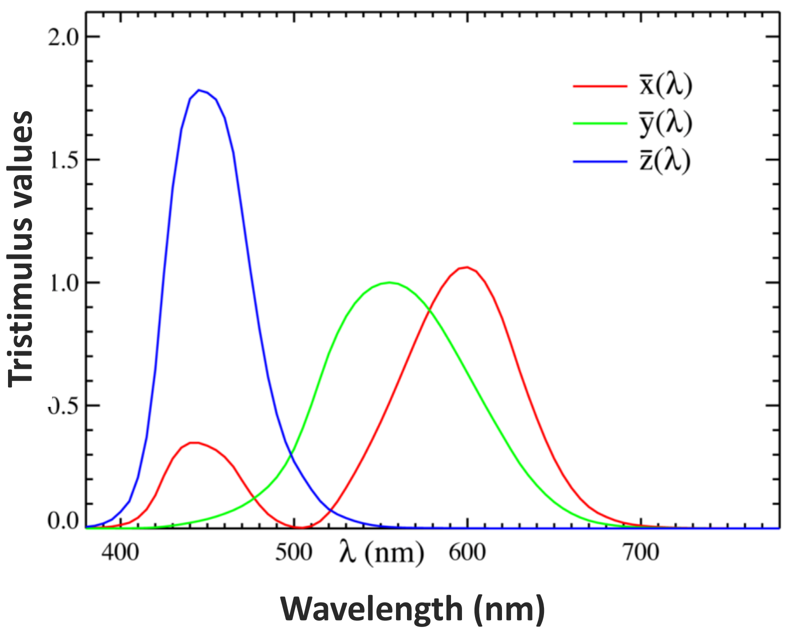
α : **The amount of the k-th primary needed to produce a color C**

β : **reference white color**

α / β : **the tristimulus value of color C**

Color Image Processing

- CIE XYZ System
  - Three "tristimulus" of a color: *X, Y* and *Z*
  - Three color matching functions: $\bar{x}$ $\bar{y}$ $\bar{z}$
  - Make all tristimulus values all positive



$$X = \int_\lambda I_\lambda \bar{x}(\lambda) d\lambda$$

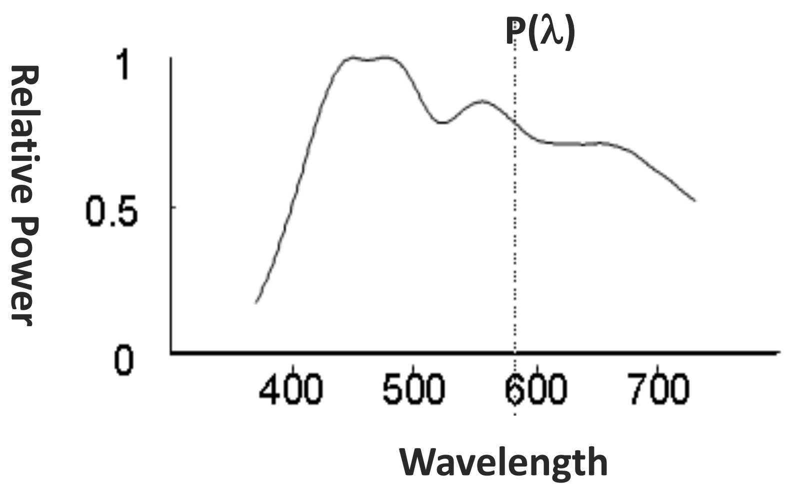
$$Y = \int_\lambda I_\lambda \bar{y}(\lambda) d\lambda$$

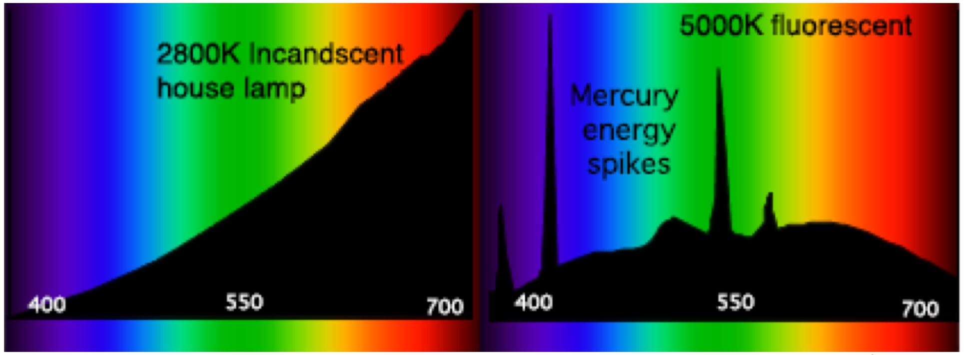
$$Z = \int_\lambda I_\lambda \bar{z}(\lambda) d\lambda$$

Spectral power distribution

Color Image Processing

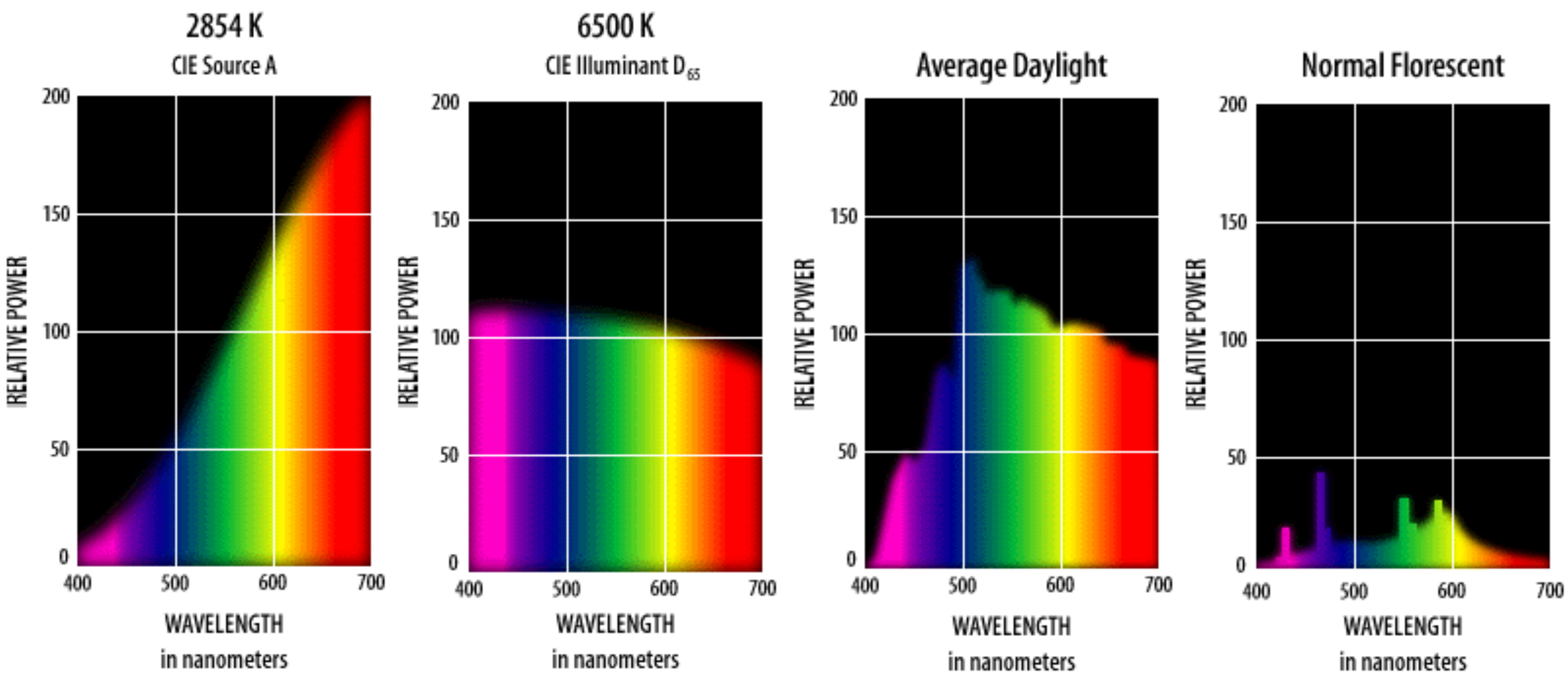
- Spectral Power Distribution (SPD)
  - Light may be precisely characterized by giving the power of the light at each wavelength in the visible spectrum
  - SPD is a function $P(\lambda)$ that defines the power of the light at each wavelength





Color Image Processing

- Spectral Power Distribution (SPD)
  - Examples



The color temperature of a lamp (bulb) describes how the light appears when the human eye looks directly at the illuminated bulb. Color temperature is measured by a unit called the kelvin (K), a scale that starts at absolute zero (-273 degrees C).

Color Image Processing

- ## Spectral Power Distribution (SPD)

The color appearance change due to different illumination may be quantified in terms of the SPD of the light

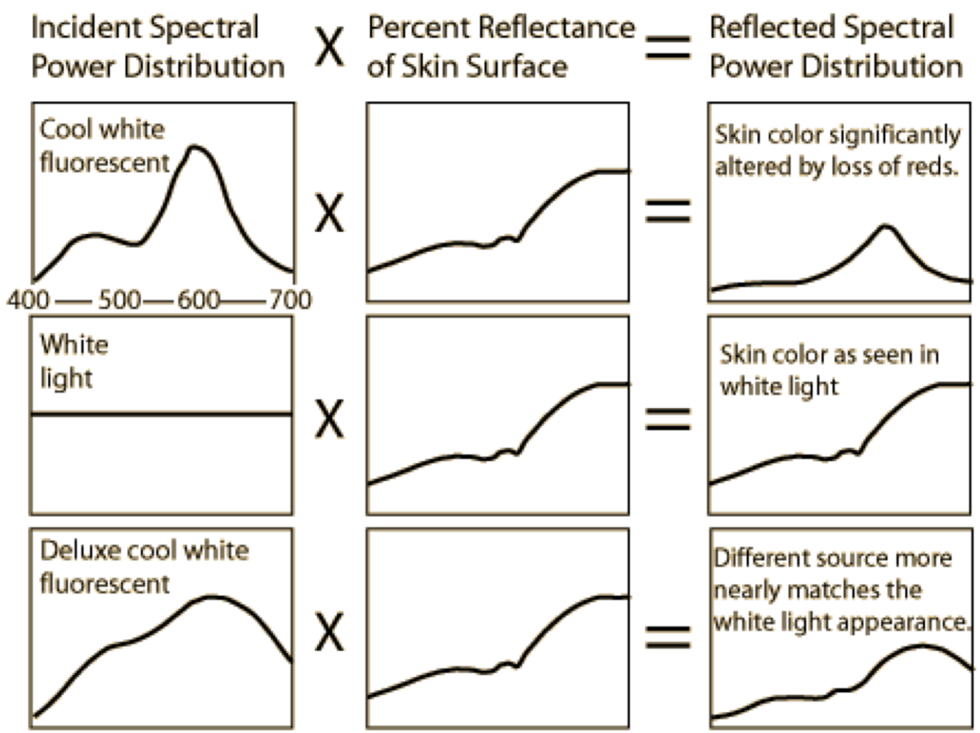
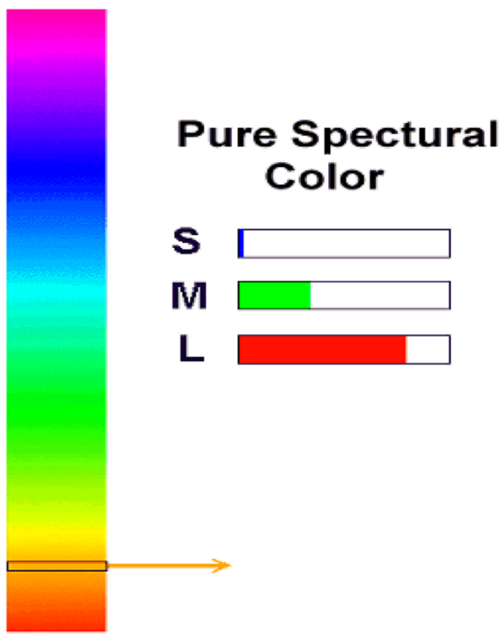
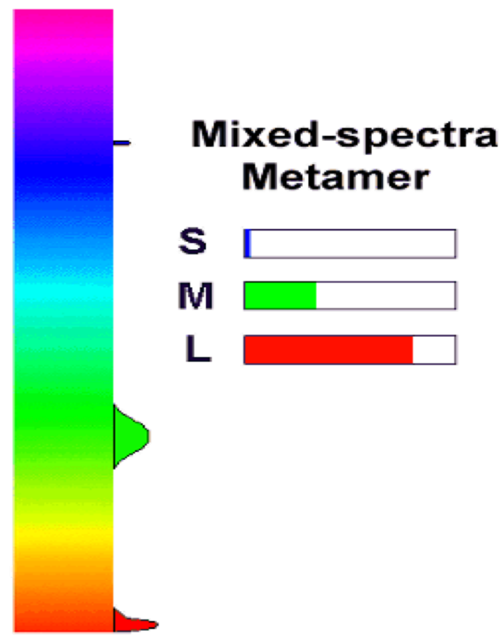


*Image source: *HyperPhysics*

Color Image Processing

- **Metamer**
  - Two colors that appear the same visually might have different SPD's (with different spectral composition)





奇美廣告
@台大誠品

https://www.youtube.com/watch?v=I9HG5ewTrJ8

Color Image Processing

- **Chromaticity Diagram**
  - Any color can be defined by its tristimulus values (X, Y, Z) or chromaticity coordinates (x,y,z)
  - Mathematical conversions between x,y,z and X,Y,Z

$$x = \frac{X}{X+Y+Z}; \qquad y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$

  - x,y,z represent the proportions of the X primary, Y primary and Z primary respectively in a given color mixture

Color Image Processing

- **C.I.E. Chromaticity Diagram (1931)**
  - Commission Internationale d'Eclairage (international commission on illumination)
  - Develop light measurement standard (most widely used standard today)
  - Achromatic light

    "white" or uncolored light
  - Chromatic light

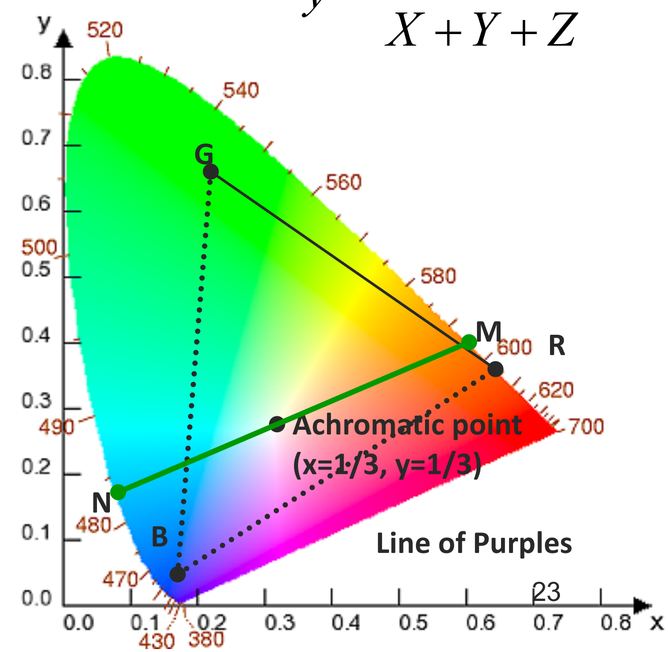
    colored light
  - Monochromatic light

    contained only one wavelength

    (laser light)

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

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



Color Image Processing

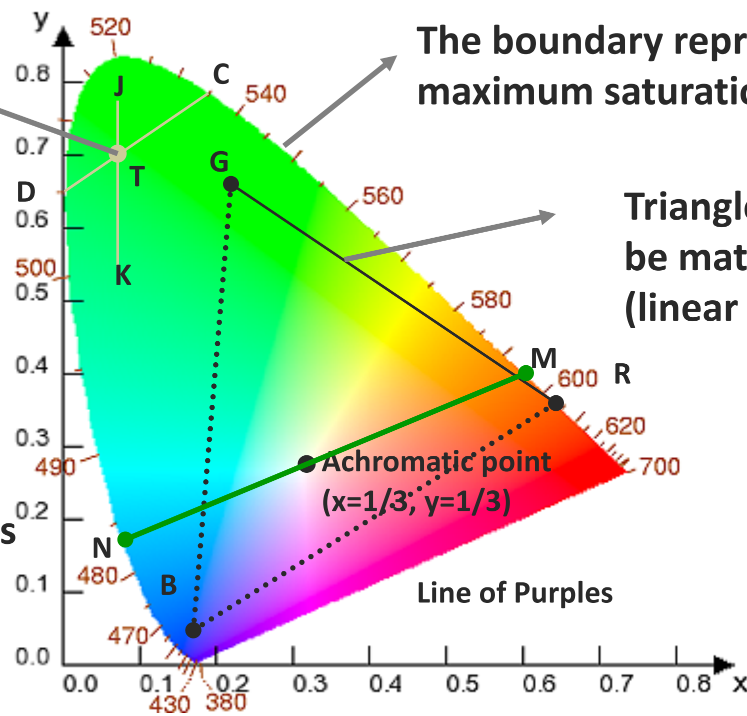
- ## C.I.E. Chromaticity Diagram

The pairs (J,K) and (C,D) can produce the same color T if combined properly

The boundary represents maximum saturation for the spectral colors

Triangle of colors that can be matched by R, G, B (linear combination)

No three colors in the diagram can span the whole color space



The connecting line pass through achromatic point form a complementary color pair, e.g. (M,N)

All colors that can be humanly perceived can be plotted within this space

CIE 1931 *xy* Chromaticity Diagram

Color Image Processing

- Color Reproduction
  - CIE chromaticity diagram is used to show color reproduction of various color imaging methods
  - Gamut
    - The range of colors accessible to a given process
    - Color coverage examples:

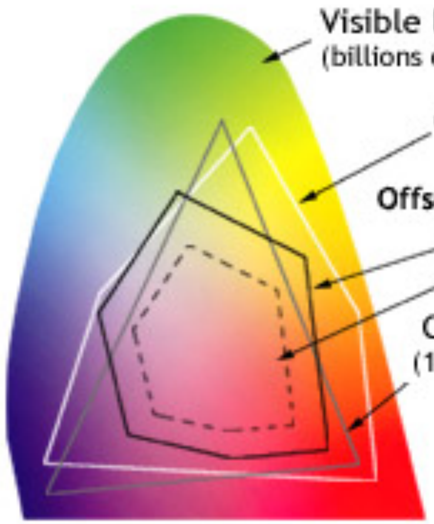
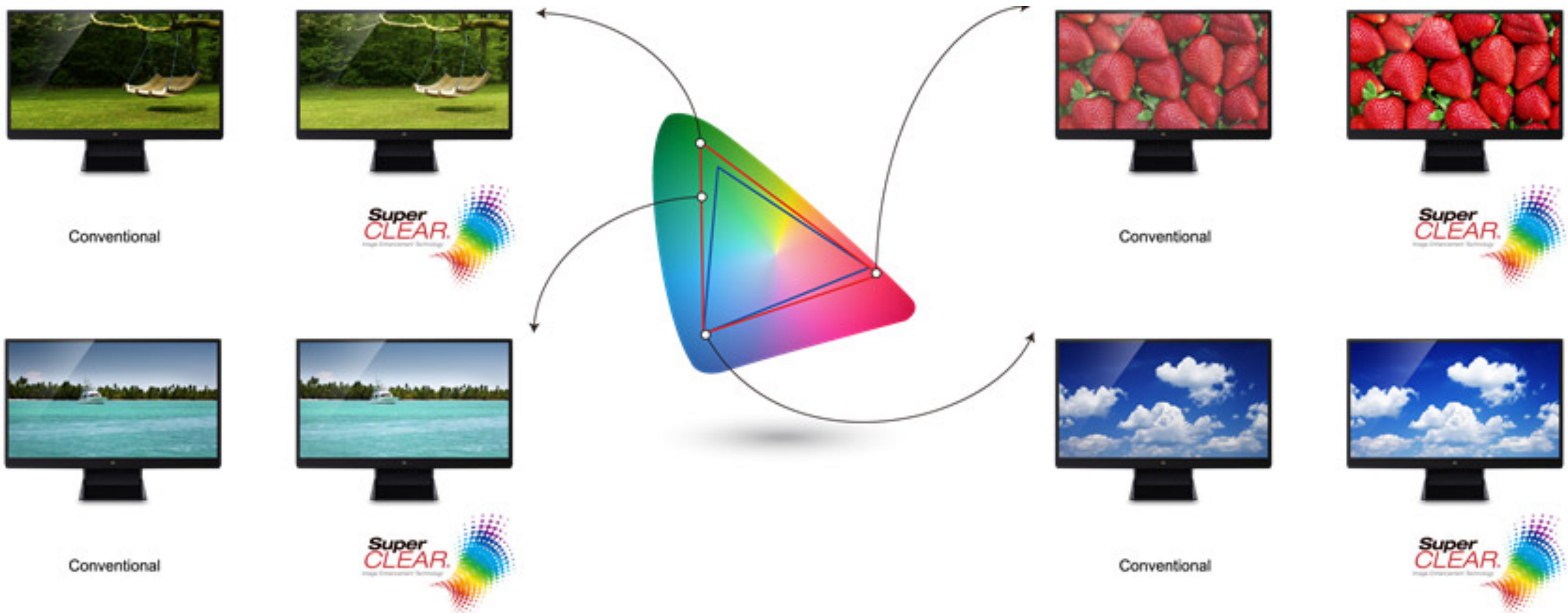


Image source: www.worgx.com

Color Image Processing

- ## Color Reproduction



http://ap.viewsonic.com/me/products/lcd/superclear/

Color Image Processing

http://genk.vn/dien-thoai/anh-tai-dai-chien-camera-phone-2013-20131122233531025.chn

Color Spaces

Color Spaces

- RGB
  - used in CRT monitors
- YIQ, YUV
  - formerly used in NTSC (National Television System Committee) television broadcasts, employed mainly in North and Central America, and Japan
- YCbCr
  - used in image and video compression: JPEG and MPEG
- CMYK
  - used in the printing process
- HSI, HSV
  - used by artists as it is more intuitive to think about a color in terms of hue and saturation
- Lab
  - Commonly used for astronomical image

Color Spaces

- **Linear Transformation of RGB color spaces**
  - **RGB: used in CRT monitors**
  - **YIQ: used in NTSC TV,** $T = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & 0.322 \\ 0.211 & -0.523 & 0.312 \end{pmatrix}$
  - **YUV: used in PAL, SECAM,** $T = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & 0.312 \end{pmatrix}$
  - **YCbCr: used for JPEG, MPEG,** $T = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.4187 & -0.0813 \\ -0.1687 & -0.3313 & 0.5 \end{pmatrix}$

Color Spaces

- **Other transformations**
  - **CMY: cyan, magenta, yellow**
    - **Complementary to red, green and blue, respectively**
    - **Useful in color printers and copiers**
    - **Used for subtractive synthesis from white in color printers**

      (C,M,Y)=(1,1,1)-(R,G,B)
  - **CMYK: like CMY, uses black (K)**
    - **Given (C,M,Y)**

      K=min(C,M,Y); C=C-K; M=M-K; Y=Y-K

Image source: www.worgx.com

Color Spaces

- **RGB to CIEXYZ**

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- **CIEXYZ to CIELAB**

$$\mathbf{L}^* = \mathbf{116}\, f\left(\frac{Y}{Y_N}\right)\text{-}\mathbf{16}\,;\ \ \mathbf{a}^* = \mathbf{500}\left[f\left(\frac{X}{X_N}\right) - f\left(\frac{Y}{Y_N}\right)\right];\ \ \mathbf{b}^* = \mathbf{200}\left[f\left(\frac{Y}{Y_N}\right) - f\left(\frac{Z}{Z_N}\right)\right]$$

where $f(t) = \begin{cases} (t)^{\frac{1}{3}}, & if\ t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29}, & otherwise \end{cases}$

and X_N, Y_N, Z_N **are the CIEXYZ tristimulus values of the reference white point. Under Illuminant D65, the values are:**

$X_N = 95.047, Y_N = 100.000, Z_N = 108.883$

Color Spaces

- **CIELAB to CIEXYZ**

$$X = X_N f^{-1}\left(\frac{1}{116}(L^* + 16) + \frac{1}{500}a^*\right)$$

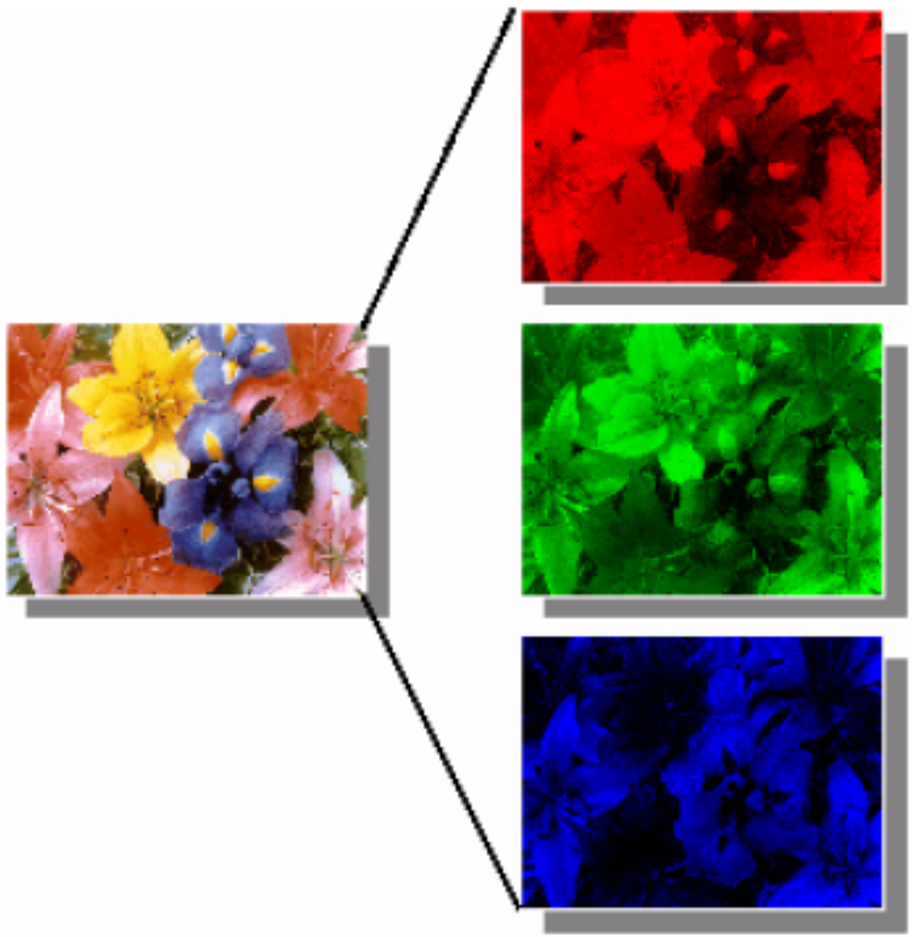
$$Y = Y_N f^{-1}\left(\frac{1}{116}(L^* + 16)\right)$$

$$Z = Z_N f^{-1}\left(\frac{1}{116}(L^* + 16) - \frac{1}{200}b^*\right)$$

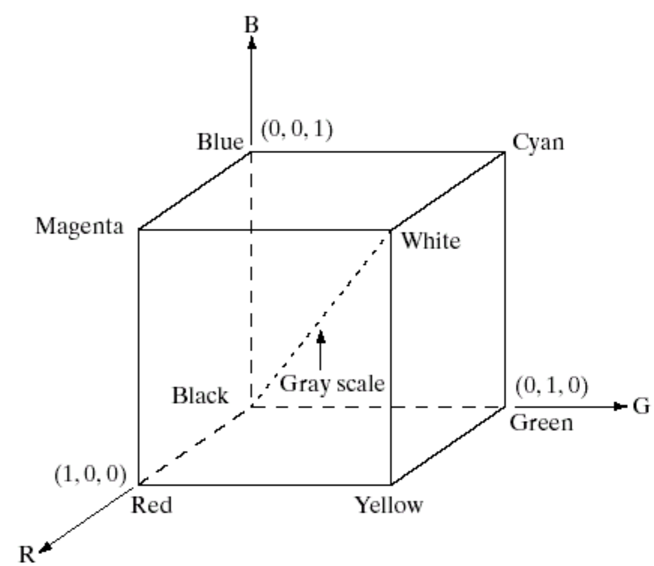
$$f^{-1}(t) = \begin{cases} t^3, & if\ t > (\frac{6}{29}) \\ 3\left(\frac{6}{29}\right)^2 (t - \frac{4}{29}), & otherwise \end{cases}$$

Color Spaces

- Color space examples - RGB

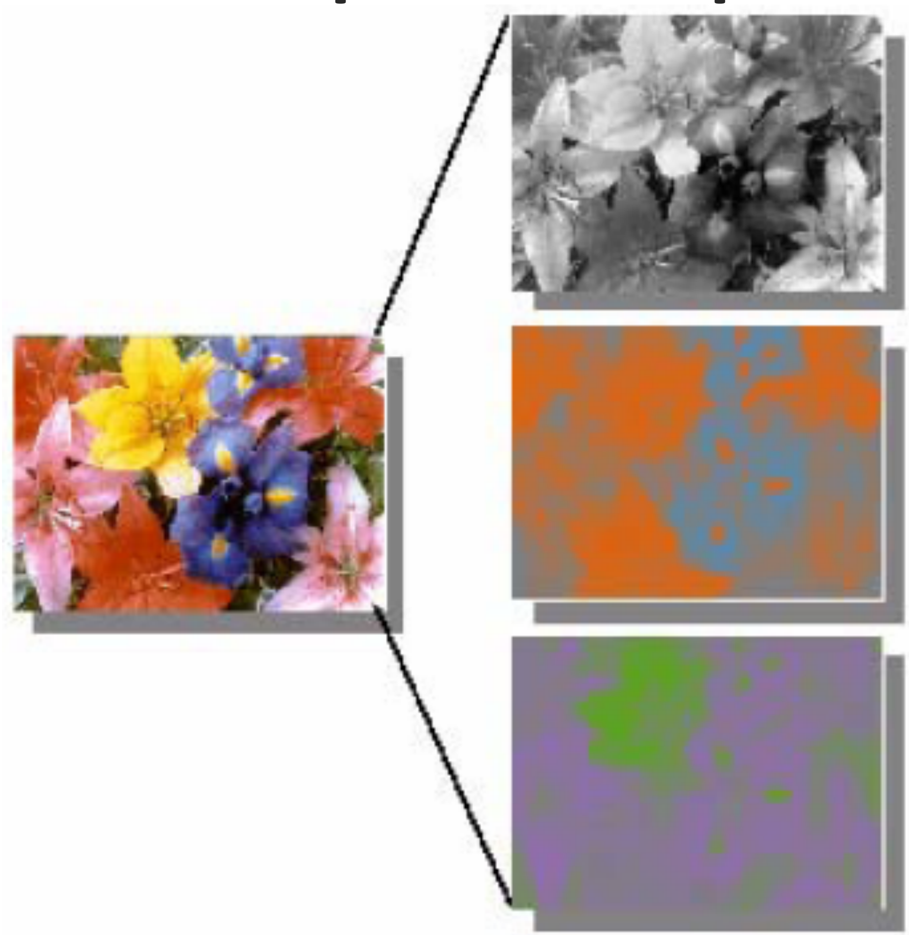


- Cartesian coordinate system (cube)
- Each color appears in its primary spectral components of R, G, and B



Color Spaces

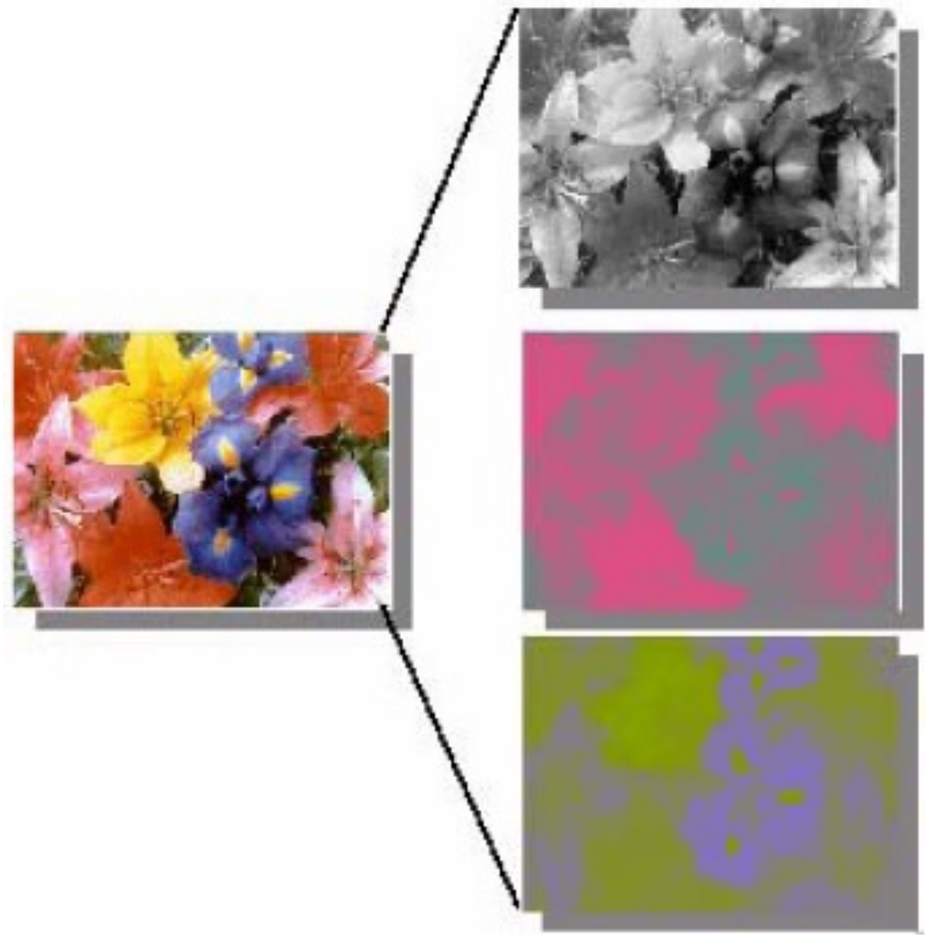
- **Color space examples - YIQ**



$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & 0.322 \\ 0.211 & -0.523 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Color Spaces

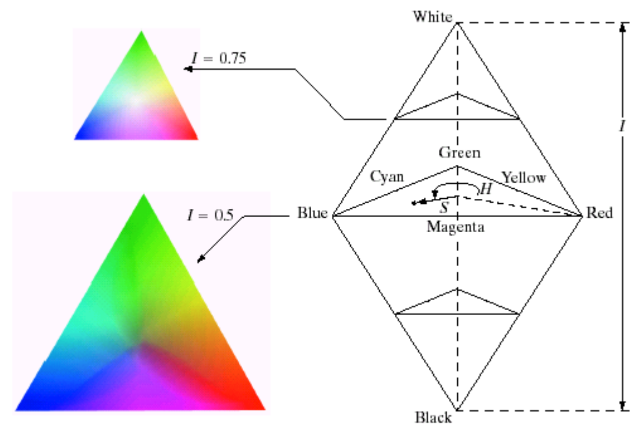
- **Color space examples - YUV**



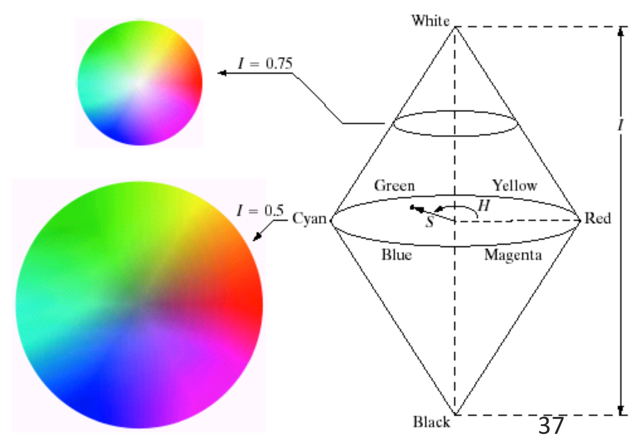
$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Color Spaces

- ## HSI Color Model
  - **The HSI space is represented by a vertical intensity axis, the length (saturation) of a vector from the axis to a color point, and the angle (hue) this vector makes with the red axis (*i.e.* spectral colors)**
  - **The power of HSI color model is to allow independent control over hue, saturation, and intensity**



triangular



circular

Color Spaces

- **RGB ←→ HSI Transformations**
  - **RGB→HSI**

$$H = \begin{cases} \theta & \text{if } B \le G \\ 360 - \theta & \text{if } B > G \end{cases} \qquad \theta = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \qquad I = \frac{1}{3}(R+G+B)$$

  - **HSI → RGB**

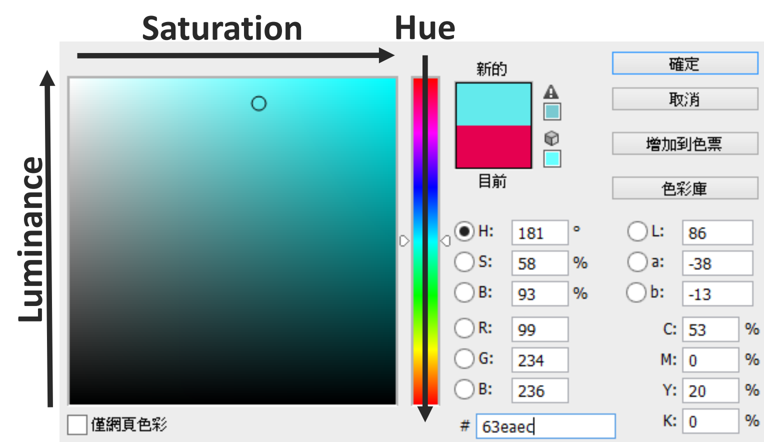
e.g. RG sector $(0^\circ < \mathrm{H} < 120^\circ)$

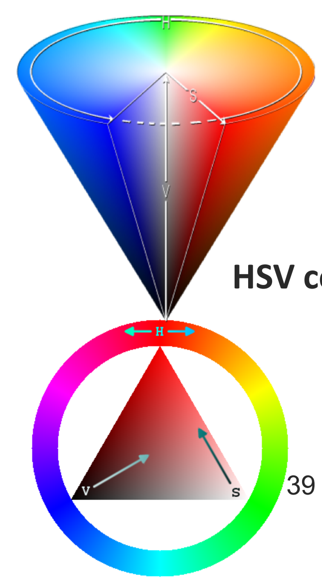
$$B = I(1-S) \qquad R = I\left[1 + \frac{S\cos H}{\cos(60^\circ - H)}\right]$$

$$G = 1 - (R+B)$$

Color Spaces

- **Attributes of color**
  - **Hue**
    - dominant wavelength of a color (*i.e.* spectral colors)
    - "tone" of a color
      e.g."Red" and "Green" are primarily describing hue)
  - **Saturation**
    - purity of a color (*i.e.* , how vivid a color appears)
    - fully saturated color – no mixture of white
  - **Brightness (Value)**
    - Luminance/lightness of a color





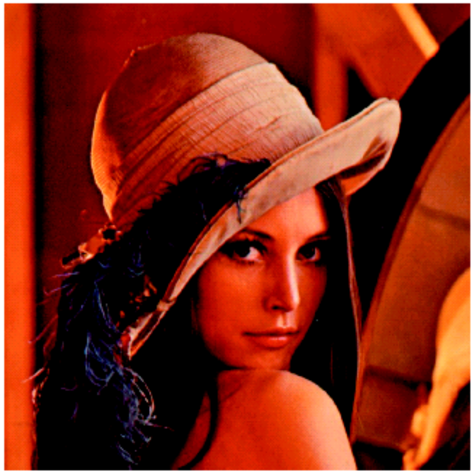
HSV cone

Color Spaces

- **Comparison of HSI & HSV**
  - **The difference between HSI and HSV (hue, saturation, value) lies in the computation of the brightness component (*I* or *V*), which determines the distribution and dynamic range of both brightness (*I* or *V*) and saturation (*S*)**

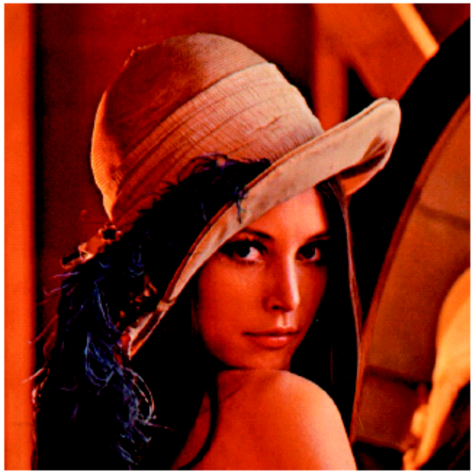
Color Spaces

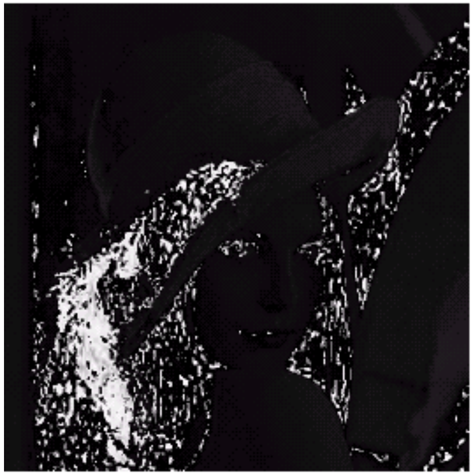
- **Example**
  - **RGB**



Color Spaces

- **Example**
  - **HSI**





Color Spaces

- **Example**

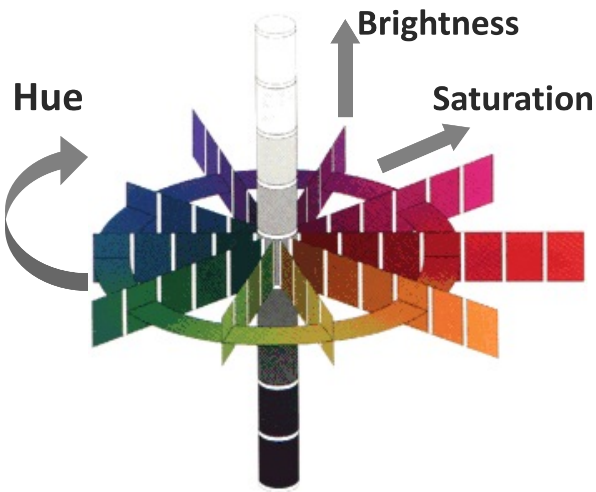
CMYK

RGB

HSI

Color Spaces

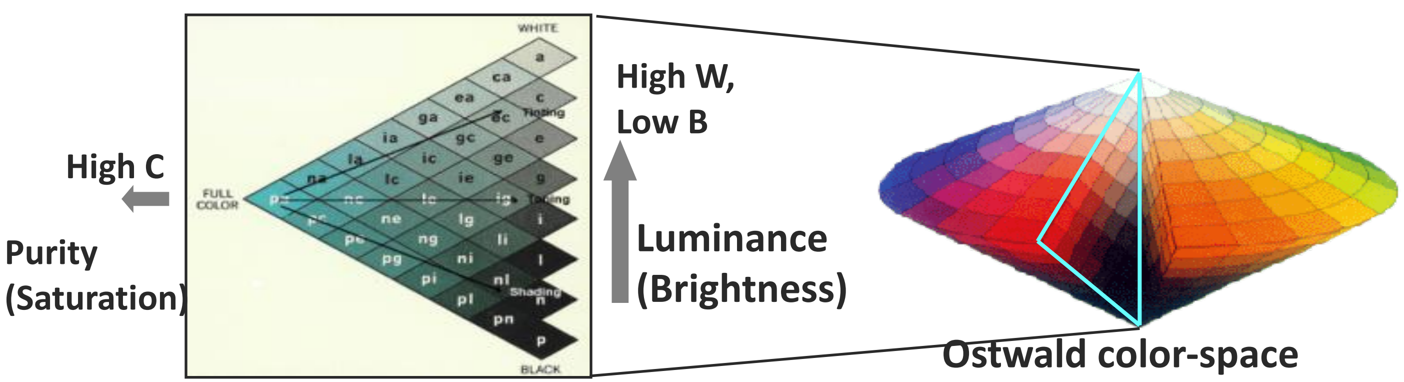
- **Munsell color system**
  - Hue: 100 equally spaced hues around the circle (10 hues, each subdivided into 10 subdivisions)
  - Saturation: 0 (gray) to 10-18 (full color), depending on the hue
  - Brightness: values from 0 for black and 10 for white



Munsell color tree

Color Spaces

- **Ostwald color system**
  - Natural color system
    - Dominant Wavelength (*Hue*)
    - Purity (*Saturation*)
    - Luminance (*Brightness*)
  - The Ostwald color space is represented by values C,W, and B to represent the percentages of the circle
    - e.g., (C,W,B)=(35,15,50) represents 35% full color, 15% white, and 50% black



Examples of Color Image Processing

Color Image Processing

- **Full Color Image Processing**
  - Full-color and interpretations of its various color space components
- **Method 1**
  - Process each component image individually and form a composite processed color image from the individually processed components
- **Method 2**
  - Work with color pixels directly

Color Image Processing

- Full Color Image Processing
  - Example

http://www.learnopencv.com/image-alignment-ecc-in-opencv-c-python/

Color Image Processing

- **Full Color Image Processing**
  - **Example**

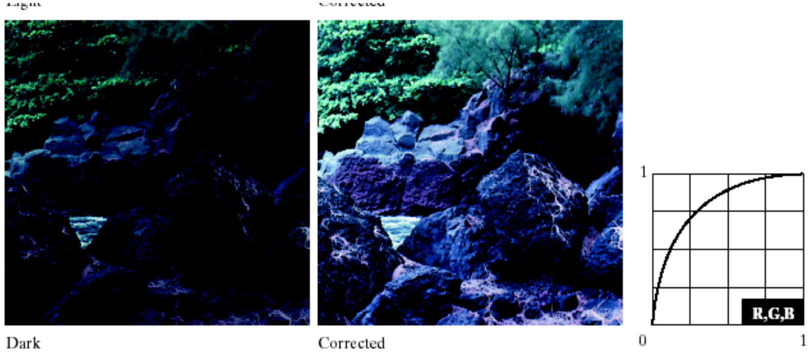
http://www.aventdudomainepublic.org/prokoudine
https://mfleisig.wordpress.com/2011/03/31/rgb-3-color-mixing-from-black-and-white/

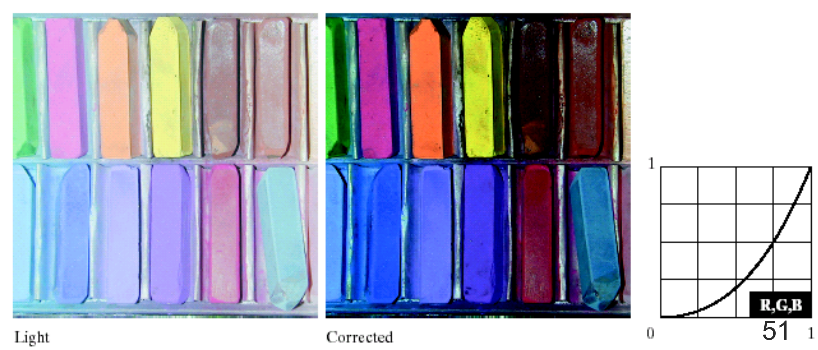
Color Image Processing

- Color Image Tone Correction
  - Tonal correction to provide a proper key (tone) of an image (just like to correct the brightness of a gray-level image)
  - Hue of color is not changed
  - For RGB and CMYK -- map all color components with the same transformation function
  - For HSI – only the intensity component is modified

Color Image Processing

- **Color Image Tone Correction**
  - **Example**





Color Image Processing

- **Color Image Smoothing**
  - **Example**

    RGB channels, Intensity component and Difference

Color Image Processing

- Color Image Sharpening
  - Example:

    RGB channels, Intensity component and Difference

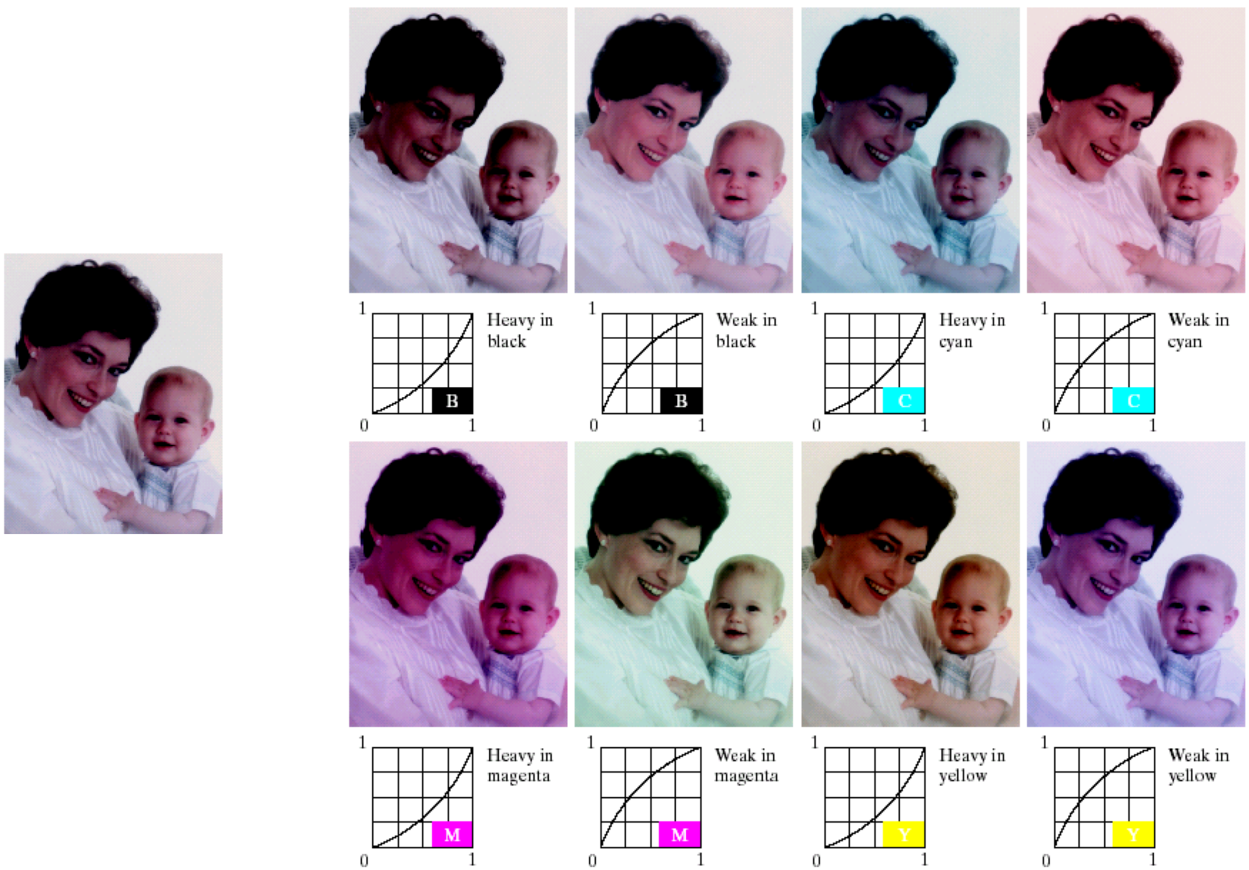
Color Image Processing

- White Balancing
  - Tells the camera sensor what temperature of light the camera is taking a picture of
  - Remove unrealistic color casts to get accurate colors in pictures

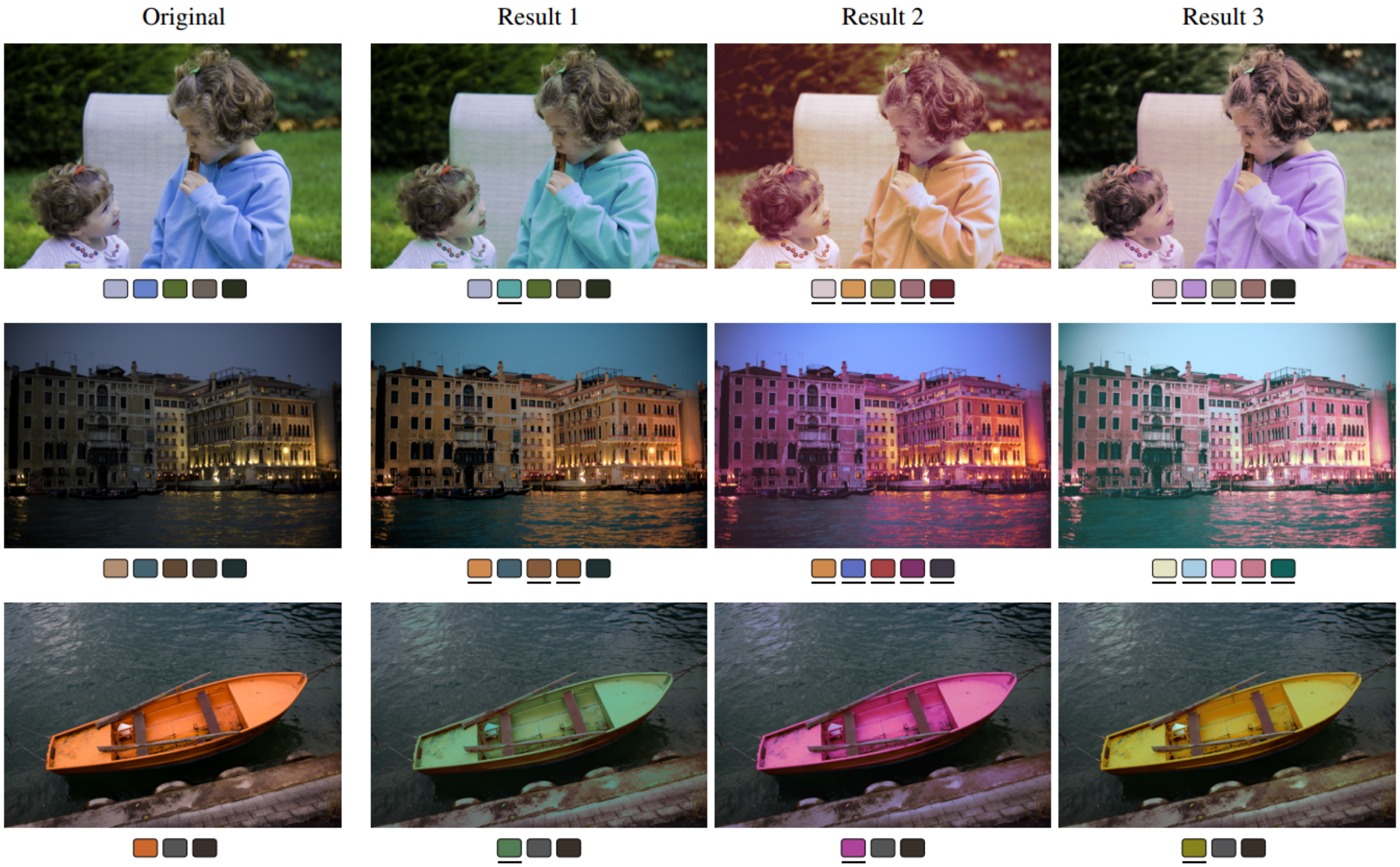
⇨

Color Image Processing

- ## Color Balancing Correction



Color Image Processing



Chang, Huiwen, et al. "Palette-based photo recoloring." ACM Transactions on Graphics (TOG) 34.4 (2015): 139.

Color Image Processing

- Image Inpainting

Color Image Processing

- **Dehazing**

http://johanneskopf.de/publications/deep_photo/

Color Image Processing

-  

Color Image Processing

- **Mosaicking**

Color Image Processing

- **Super Resolution**





Color Image Processing

- **High-Dynamic-Range Imaging**

Image Compression

Image Compression

- ## Need for Compression

Multimedia image data	Size of image	Bits/pixel or Bits/sample	Uncompressed size (B for bytes)	Bandwidth (b for bits)	Transmission time	
					56K modem	780Kb DSL
Grayscale image	512x512	8bpp	262KB	2.1Mb per image	42 sec.	3 sec.
Color image	512x512	24bpp	786KB	6.29Mb per image	110 sec.	7.9 sec.
Medical image	2048x 1680	12bpp	5.16MB	41.3Mb per image	12 min.	51.4 sec.
SHD image	2048x 2048	24bpp	12.58MB	100Mb per image	29 min.	2 min.

Image Compression

- Some Image Compression Formats
  - RIFF – Resource Interchange File Format
  - GIF – Graphics Interchange File Format
  - PNG – Portable Network Graphics
  - JPEG – Joint Photographic Expert Groups
  - JPEG 2000

Image Compression

- JPEG Background
  - Make no assumptions about the type of image
    - applicable to practically any kind of continuous-tone digital source image
  - With tractable computational complexity
  - Support flexibility by allowing the following modes of operation
    - Lossless and Lossy Encoding
    - Sequential Encoding
    - Progressive Encoding
    - Hierarchical Encoding

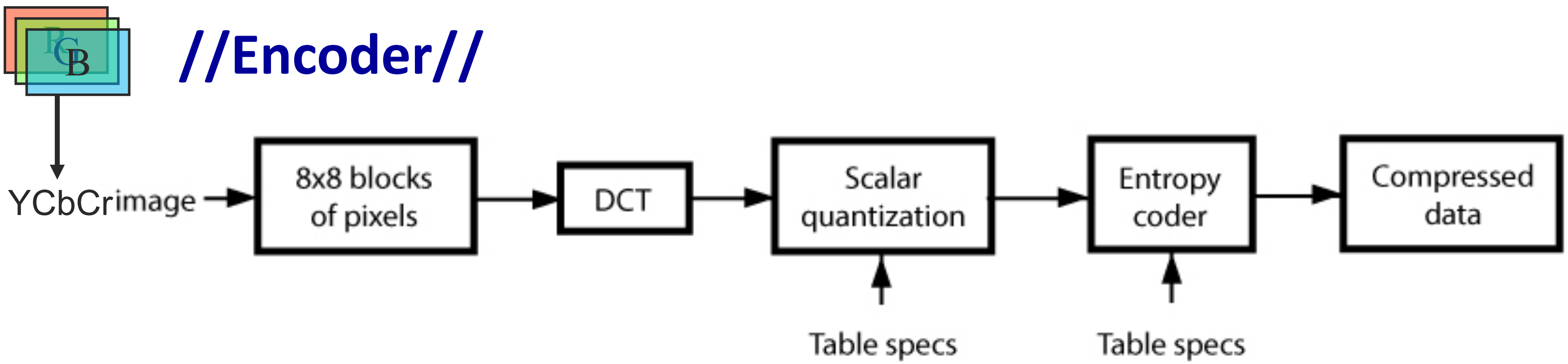
Image Compression

- JPEG Image Coding
  - The standard algorithm for compression of still images
    - Started in June 1987
    - Finalized and accepted in 1991
  - Reasonably low computational complexity
  - Capable of producing compressed images of high quality
  - Provide both ‘lossless’ and ‘lossy’ compression of arbitrarily sized graylevel and color images

Image Compression

- **JPEG Lossy Codec Scheme**

//Encoder//



//Decoder//

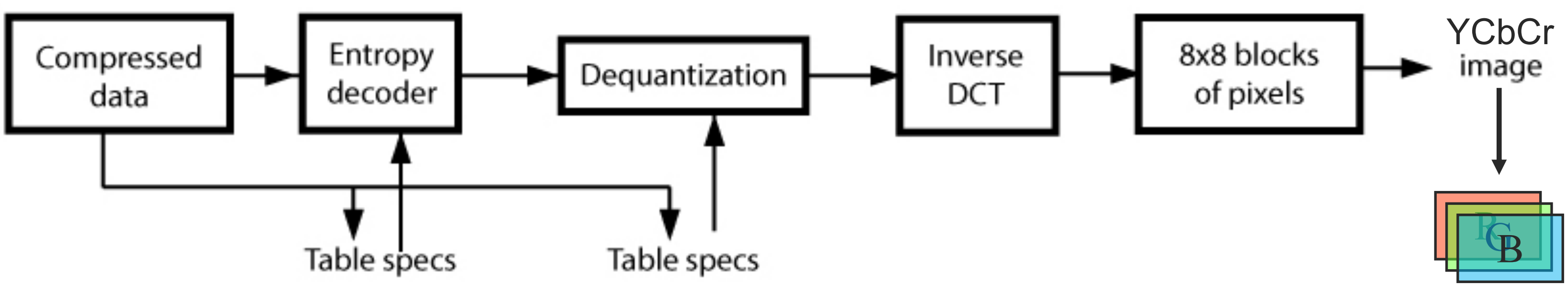


Image Compression

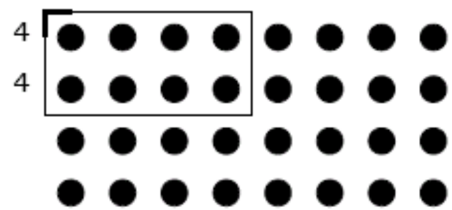
- **JPEG Algorithmic Overview**
  - Convert the RGB color channels of the input image into YCbCr components
  - $$Y = 0.299*R + 0.587*G + 0.114*B$$
    $$Cb = -0.299*R - 0.587*G + 0.886*B$$
    $$Cr = 0.701*R - 0.587*G - 0.114*B$$
    - each of them is encoded independently
  - Each channel is divided into a series of 8x8 pixel blocks followed by being processed in a raster scan sequence from left to right, top to bottom
  - Each 8x8 block of pixels is analyzed using DCT (transform coding) and DCT coefficients (1 DC and 63 AC's) are quantized
  - After DCT and quantization, coefficients are entropy coded

Image Compression

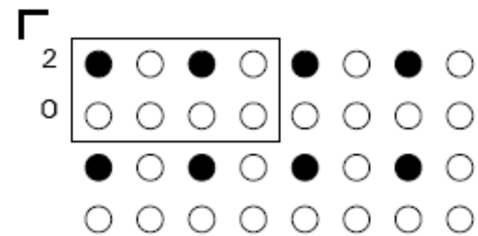
- **Why not RGB?**
  - Similarly to the processing in the human visual system, transformed color spaces such as YIQ, YUV, YCbCr represent color with luminance (Y) and chrominance (the other 2 channels)
    - We can subsample the chrominance channels (e.g. 4:2:2, 4:2:0 subsampling scheme)
    - We can quantize the chrominance channels more coarsely
    - The chrominance channels are rather uncorrelated with the luminance channel, which yields better compression

Image Compression

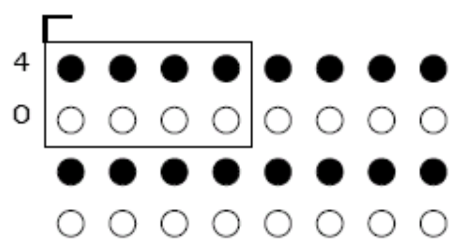
- **Subsampling**
  - **4:4:4**



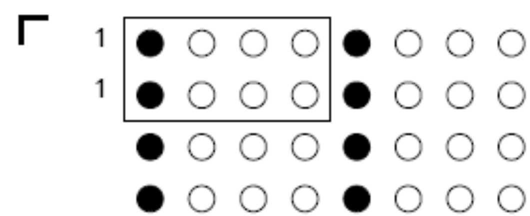
4:2:0



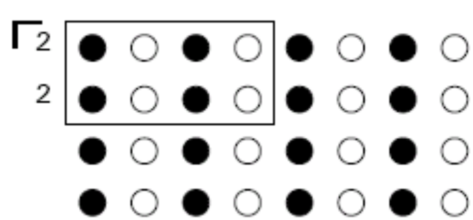
  - **4:4:0**



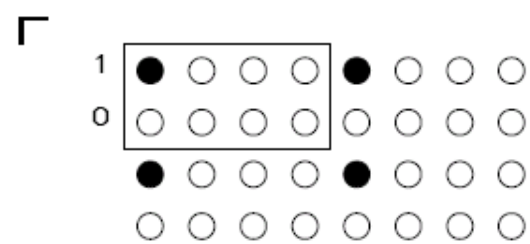
4:1:1



  - **4:2:2**



4:1:0



DCT and IDCT

- One-dimensional example

Spatial Domain | **Frequency Domain**

Input function

190	184	186	182	167	123	63	38

DCT

400.56	143.81	-67.76	-16.33	7.42	-4.73	5.49	0.05

Use all 8 DCT values for decoding

400.56	143.81	-67.76	-16.33	7.42	-4.73	5.49	0.05

Input function computed from DCT using IDCT

190	184	186	182	167	123	63	38

Error

0	0	0	0	0	0	0	0

DCT and IDCT

- One-dimensional example

Spatial Domain | **Frequency Domain**

Input function

190	184	186	182	167	123	63	38

DCT

400.56	143.81	-67.76	-16.33	7.42	-4.73	5.49	0.05

Zeros at last 4 DCT values

400.56	143.81	-67.76	-16.33	0.00	0.00	0.00	0.00

Input function computed from DCT using IDCT

187.6	186.8	186.5	182.4	163.4	122.6	70.5	33.0

Error

2.40	-2.80	-0.50	-0.40	3.60	0.40	-7.50	5.0

DCT and IDCT

- One-dimensional example

Spatial Domain | **Frequency Domain**

Input function

190	184	186	182	167	123	63	38

DCT

400.56	143.81	-67.76	-16.33	7.42	-4.73	5.49	0.05

Zeros at last 7 DCT values

400.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Input function computed from DCT using IDCT

141.6	141.6	141.6	141.6	141.6	141.6	141.6	141.6

→ average

Error

48.4	42.4	44.4	40.4	25.40	-18.6	-78.6	-103.6

DCT and IDCT (cont'd)

- Comments
  - By zeroing out the last four values
    - the maximum error is 7 out of 255, which is 2.7 %
  - By compressing the data 50 %
    - an error of 2.7 % is introduced
    - acceptable quality

DCT and IDCT

- **2D Discrete Cosine Transform (DCT)**
  - **Transform**

$$F(u,v) = \frac{2C(u)C(v)}{n} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(j,k) cos\left[\frac{(2j+1)u\pi}{2n}\right] cos\left[\frac{(2k+1)v\pi}{2n}\right]$$

  - **Inverse DCT**

$$f(j,k) = \sum_{v=0}^{n-1} \sum_{u=0}^{n-1} C(u)C(v)F(u,v) cos\left[\frac{(2j+1)u\pi}{2n}\right] cos\left[\frac{(2k+1)v\pi}{2n}\right]$$

$$C(w) = \begin{cases} 1/\sqrt{2}, & if\ w = 0 \\ 1, & if\ w = 1, 2, \dots, n-1 \end{cases}$$

DCT and IDCT (cont'd)

- **2D Discrete Cosine Transform (DCT)**

$$F(u,v) = \frac{2C(u)C(v)}{n} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(j,k) cos\left[\frac{(2j+1)u\pi}{2n}\right] cos\left[\frac{(2k+1)v\pi}{2n}\right]$$

  - **Transform of an 8x8 block with u=1, v=0:**

$$F_{10} = \frac{c_1 c_0}{4}\left[\cos\frac{\pi}{16}\left(\sum_{i=0}^{7} f(0,j) - \sum_{i=0}^{7} f(7,j)\right) + \cos\frac{3\pi}{16}\left(\sum_{i=0}^{7} f(1,j) - \sum_{i=0}^{7} f(6,j)\right) + \right.$$

$$\left. \cos\frac{5\pi}{16}\left(\sum_{i=0}^{7} f(2,j) - \sum_{i=0}^{7} f(5,j)\right) + \cos\frac{7\pi}{16}\left(\sum_{i=0}^{7} f(3,j) - \sum_{i=0}^{7} f(4,j)\right)\right]$$

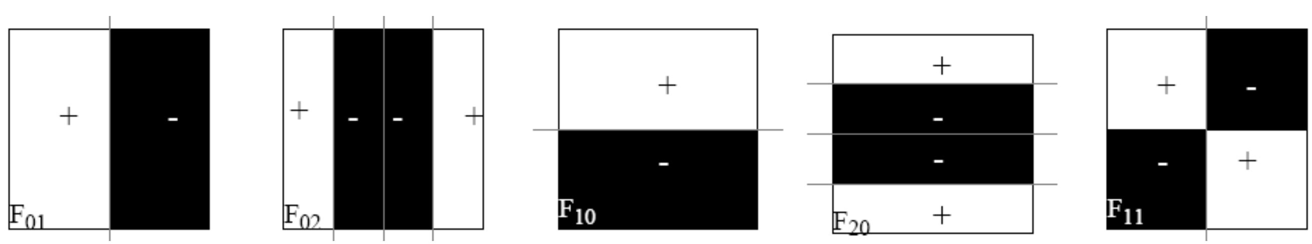


Image Compression

- **8x8 DCT Basis Functions**

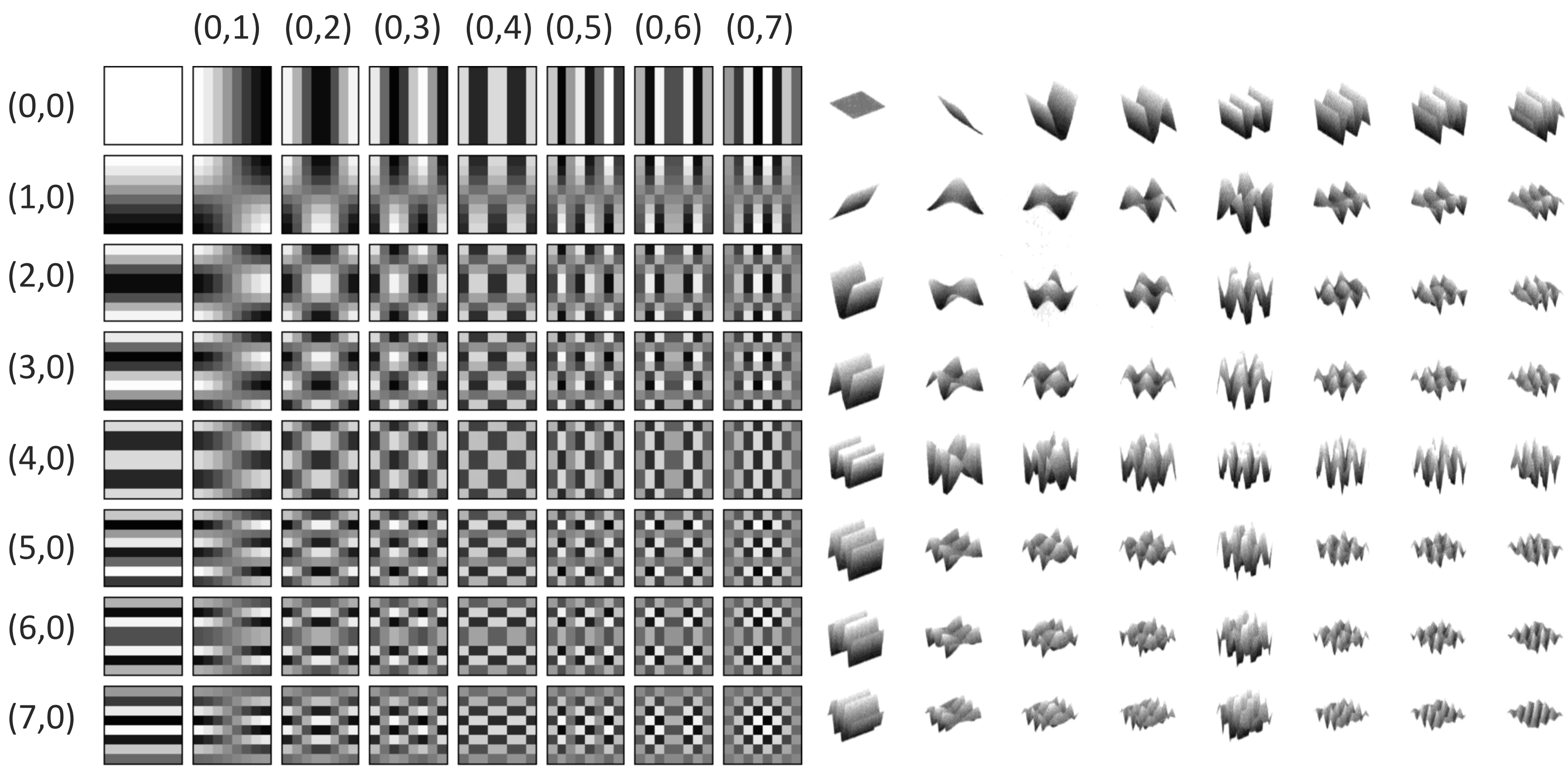


Image Compression

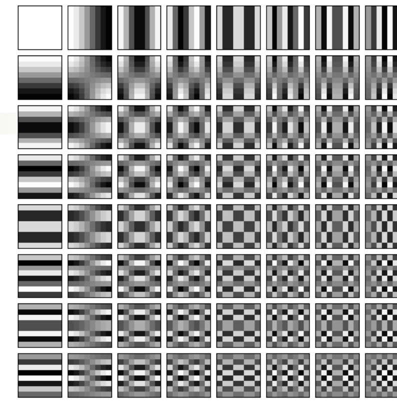
- The Meaning of DCT Coefficients
  - Represent the spatial frequency content within an 8x8 image block
  - The (0,0) coefficient is called DC coefficient
    - The average of the 64 image pixel values in the block
  - The rest of 63 coefficients are called AC coefficients
  - Move to the right of the block

    → the energy in higher horizontal frequencies
  - Move down the block

    → the energy in higher vertical frequencies

Image Compression

- **Examples of 8x8 DCT Coefficients**



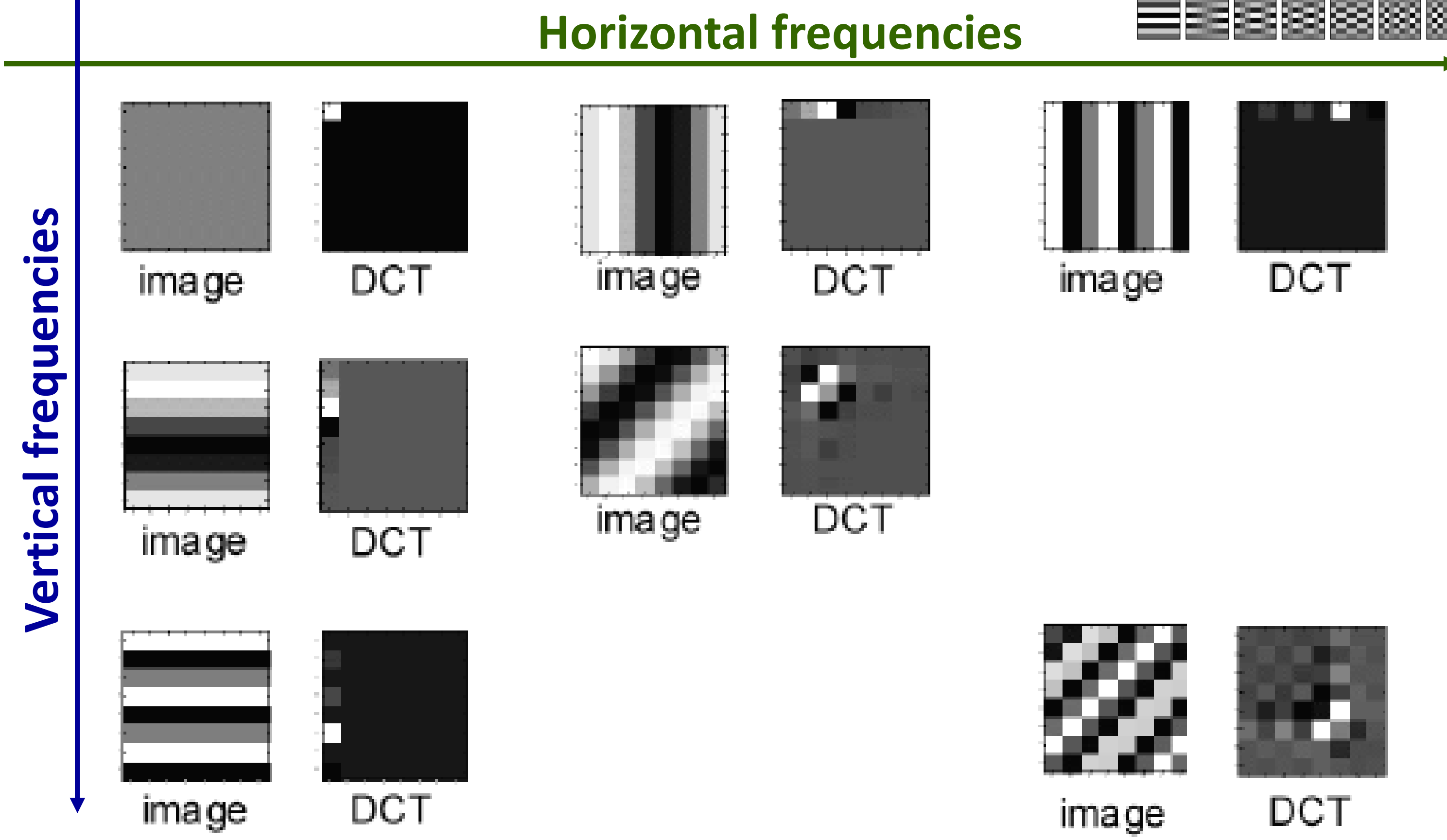


Image Compression

- DCT Coefficient Quantization
  - Each DCT coefficient is quantized independently
  - More quantization levels are required for
    - Low frequency coefficients
    - Luminance channel

    Because human visual system is more sensitive to the above two components
  - The quantized DC coefficient of a block is encoded as **the difference from the DC term of the previous block** in the processing order (since usually there's strong correlation between DC's of adjacent blocks)

//Note// the quantization table is not standardized

Quantization

- Quantization table for Y component

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

- Quantization table for Cb & Cr components

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Image Compression

- **Example of Quantization**

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Source image sample

235.6	-1.0	-12.1	-5.2	2.1	-1.7	-2.7	1.3
-22.6	-17.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

DCT coefficients

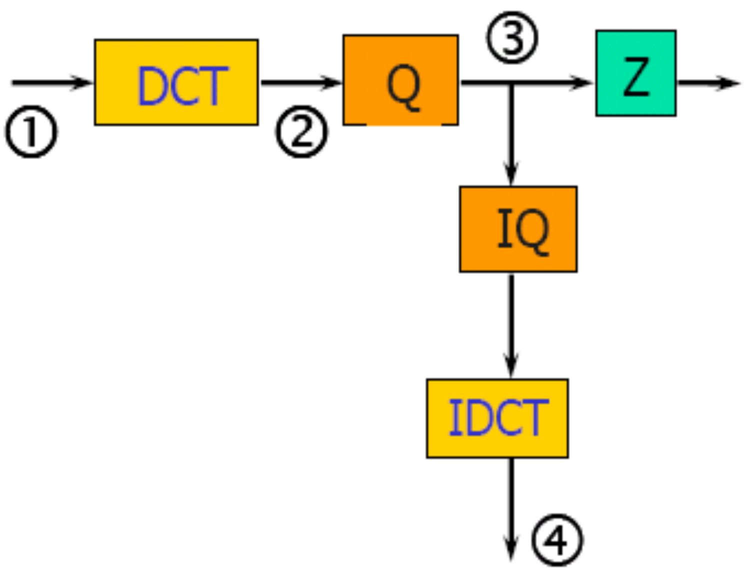
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Quantization table

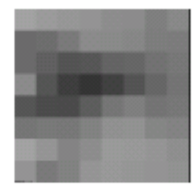
15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantized coefficients

Effects of DCT and Quantization



①



DCT

②

994	-34	20	38	22	0	5	-1
-49	17	8	-17	-4	-1	4	-2
149	35	-37	-15	1	3	-1	-4
58	8	-32	-3	-9	-11	2	-4
-27	5	26	6	9	7	8	0
9	41	53	8	0	1	-1	2
0	2	-1	-6	-2	-3	0	1
3	-14	-12	2	-1	-7	-5	-1

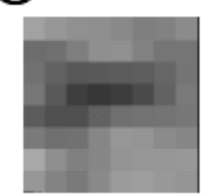
QUANTIZATION
(Step size = 30)

③

33	-1	1	1	1	0	0	0
-2	1	0	-1	0	0	0	0
5	1	-1	0	0	0	0	0
2	0	-1	0	0	0	0	0
-1	0	1	0	0	0	0	0
0	1	2	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

IQ, IDCT

④



Mean square error
=38.75

Image Compression

- **Entropy Coding of DCT Coefficients**
  - **The quantized DCT coefficients are scanned in zigzag order**

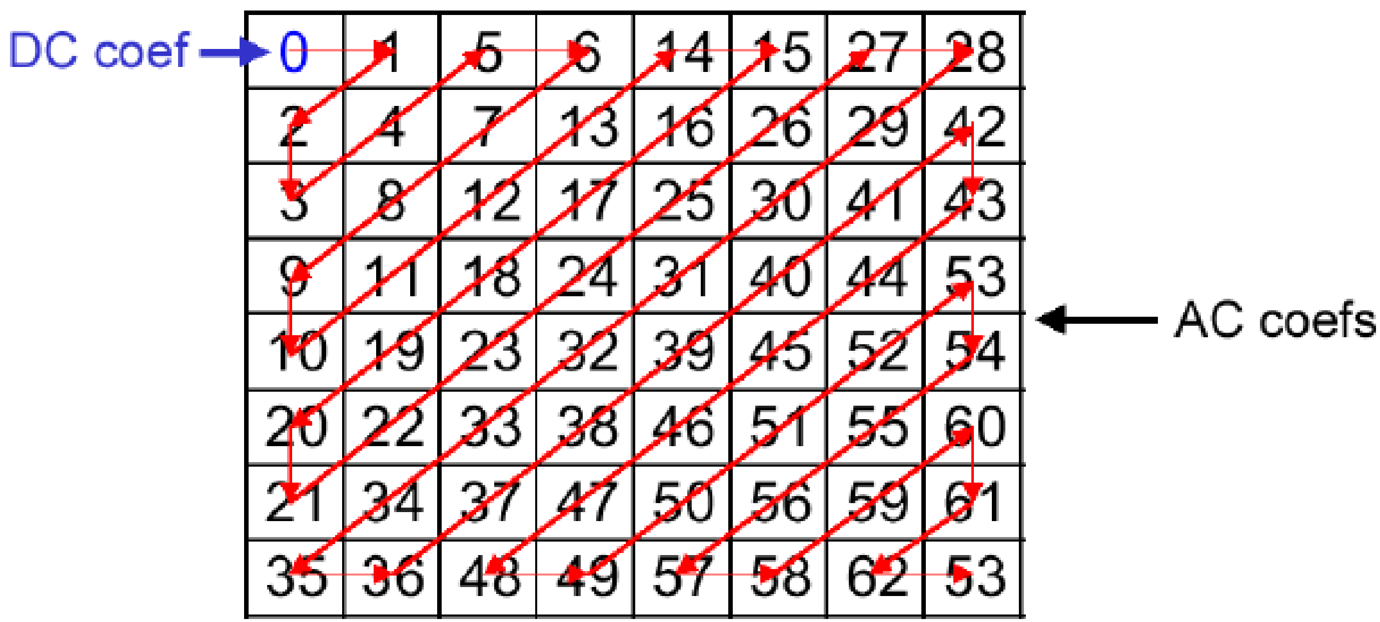


Image Compression

- **Entropy Coding of DCT Coefficients**
  - **Zigzag ordering places low frequency coefficients before high frequency coefficients**
    - **Low frequency → more likely to be non-zero**
    - **High frequency → more likely to be null after quantization**

→ the sequence of quantized coefficients is often a long sequence of null values

168	0	-1	0	0	0	1	-1
-1	0	0	0	1	0	0	0
0	0	1	1	0	0	0	1
0	0	0	0	0	0	0	-1
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	-1	1	0	0	0	0

Zigzag →

168,0,-1,0,0,-1,0,0,0,0,1,
0,1,0,0,0,1,1,0,0,0,0,0,0,
0,0,0,1,,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,0,0,0,0,-1,1,0,
0,0,-1,0,0,0,0,0,0,0,0,0,0

Differential Pulse Coded Modulation

- DC coefficients frequently occupied a significant percentage of bitrate, special treatment is worthwhile
- DPCM (Differential Pulse Coded Modulation)
  - use the previous QDC to predict the current QDC

$$X_{DPCM} = X_i - X_{i-1}$$

eg:

DC=138	DC=140
DC=136	DC=138

$\xrightarrow{\text{DPCM}}$ 138 → 2→ -4 → 2

Entropy Coding

- **AC coefficients are represented in an intermediate run-length representation, which encodes the runs of zero preceding any non-null coefficient**

  → Symbol 1: (runlength)

  → Symbol 2: (amplitude)

eg:

168,0,-1,0,0,-1,0,0,0,0,1,
0,1,0,0,0,1,1,0,0,0,0,0,0,
0,0,0,1,-1,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,0,0,0,0,-1,1,0,
0,0,-1,0,0,0,0,0,0,0,0,0,0

→

(1,-1) (2,-1) (4,1) (1,1) (3,1)
(0,1) (9,1) (0,-1) (14,1) (4,-1)
(0,1) (3,-1) EOB

Entropy Coding

- Entropy Coding of DCT Coefficients
  - For both DC and AC coefficients,
    - Each symbol 1 is Huffman encoded
      - Variable length prefix code
    - Each symbol 2 is encoded with Variable Length Integer (VLI) code
      - Different from a Huffman code
      - Symbol length is known in advance

An example

(a)8×8影像

159	153	158	152	140	138	132	132
164	162	162	157	151	142	134	132
167	168	161	160	158	145	139	134
164	168	161	166	162	152	149	141
171	166	168	167	163	162	147	151
173	164	169	170	166	166	162	161
175	169	172	176	174	172	174	166
173	172	175	173	180	181	177	172

DCT

(b)減128後再DCT

260	49	-16	5	2	4	0	1
-79	36	-2	-7	1	-3	-1	-2
0	-8	3	-2	-2	1	5	1
-8	-4	5	-4	1	7	6	-2
-2	-6	-1	0	-4	-1	0	-1
-3	-2	-1	-1	1	2	-5	-1
-4	-1	1	0	0	-2	2	0
-1	1	1	1	-1	1	0	0

量化

(c)量化後系數

16	4	-2	0	0	0	0	0
-7	3	0	0	0	0	0	0
0	-1	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Zig.Zag Scan 和 Model

長度：-, 0, 0, 1, 0, 2, 0, 54

振幅：16, 4, -7, 3, -2, -1, -1

DC

編碼：1010110100100100000110

(*)1111010111110000001010

(b)量化表

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(e)量化後係數

16	4	-2	0	0	0	0	0
-7	3	0	0	0	0	0	0
0	-1	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

* DC值需和前一個8×8 block相減後才編碼，此例DC編碼為1010110，size為5， Amplitude為-9。

反量化

(f)DCT係數

256	44	-20	0	0	0	0	0
-84	36	0	0	0	0	0	0
0	-13	0	0	0	0	0	0
-14	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

反DCT

(g)還原後影像

153	153	152	149	144	137	131	126
158	158	157	154	149	142	135	131
165	164	163	160	155	148	141	137
167	167	167	164	159	153	147	143
167	168	168	167	163	158	153	149
166	168	169	170	168	165	161	159
168	170	173	175	175	184	171	170
169	172	176	179	181	180	179	177

Image Compression

- DCT induces some artifacts
  - Blur
    - truncation of high frequency coefficients
  - Graininess
    - coarse quantization of some coefficients
  - Blocking artifacts
    - each block is processed independently with different quantization strategies
- How to solve it?
  - overlapping block
  - post processing

Image Compression

- Example

$$C_r = \frac{Number\ of\ bits\ needed\ for\ the\ original\ signal}{Number\ of\ bits\ needed\ for\ the\ compressed\ signal}$$

786488 bytes

23116 bytes (Cr=34.0)

Image Compression

- Example

786488 bytes

11738 bytes (Cr=67.0)

Image Compression

- **Example**

786488 bytes

49746 bytes (Cr=15.80)

Image Compression

- **Example**

786488 bytes

15730 bytes (Cr=50.0)

Image Compression

- **Example**

786488 bytes

26614 bytes (Cr=29.55)

Image Compression

- Example

786488 bytes

7865 bytes (Cr=100.0)

JPEG Performance

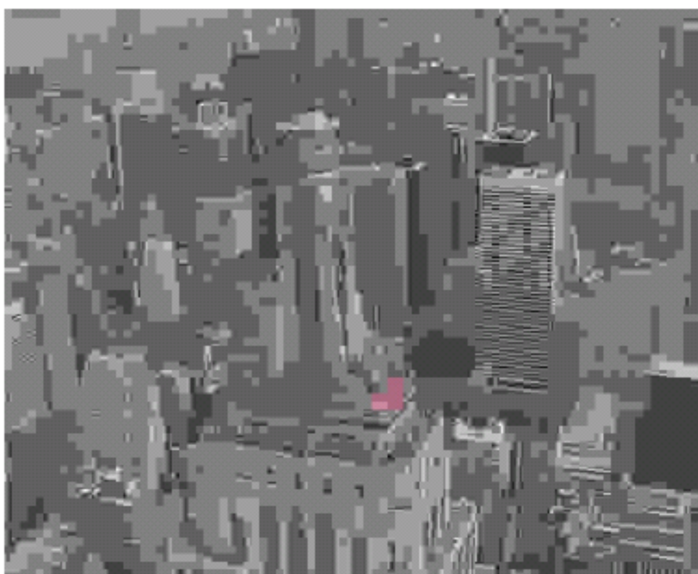
- Guideline for compression and picture quality
- Highly correlated to image content

Bits/Pixel	Quality	Compression Ratio
>=2	Indistinguishable	8-to-1
1.5	Excellent	10.7-to-1
0.75	Very Good	21.4-to-1
0.5	Good	32-to-1
0.25	Fair	64-to-1

JPEG 2000

- **Created by JPEG in 2000**
  - **wavelet-based method**
  - **compression ratio around 100:1**

JPEG Bpp: 0.0185 **JPEG 2000**



PSNR: 22.63 dB

PSNR: 26.00 dB

JPEG 2000

- **Another example**

Bpp: 0.075

JPEG

PSNR: 29.67 dB

JPEG 2000

PSNR: 30.86 dB