HA NOI UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Computer Vision

Chapter 2: Image formation, acquisition and digitization

Content

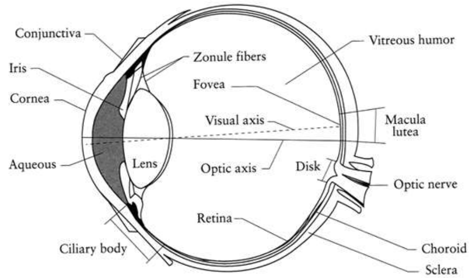
- Image formation
  - Human vision
  - Image formation
- Acquisition and digitization: Digital camera
  - Imaging sensor
  - 2D signal and sampling
- Color:
  - Primary color, additive/ subtractive color, color spaces
- Digital image representation and formats

Image formation

Image formation studies the forward process of producing images and videos.

- Image formation encompasses the radiometric and geometric processes by which 2D images of 3D objects are formed. To produce a real image, the nature of the visual sensors (i.e. CCD and CMOS cameras), should be studied.
- Imaging process is a mapping of an object to an image plane.
- With digital images, the image formation process also includes analog to digital conversion, sampling
- **Human color vision (Perception) :** In the case of computer vision the light incident on the sensor comprises the image. In the case of visual perception, the human eye has a color dependent response to light which is the spectral sensitivity of human vision.

The Eye



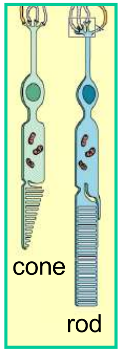
- The human eye is a camera
  - **Iris** - colored annulus with radial muscles
  - **Pupil** - the hole (aperture) whose size is controlled by the iris
  - What's the sensor?
    - photoreceptor cells (rods and cones) in the **retina**

Slide by Steve Seitz

Two types of light-sensitive receptors

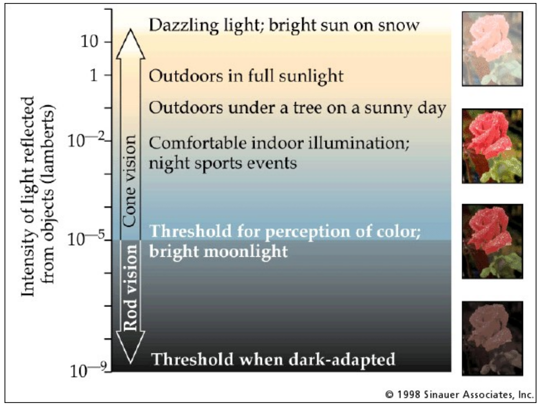
Cones
- cone-shaped
- less sensitive
- operate in high light
- color vision

Rods
- rod-shaped
- highly sensitive
- operate at night
- gray-scale vision



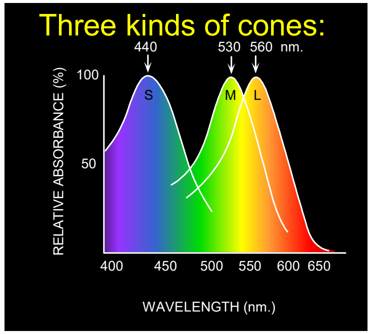
James Hays

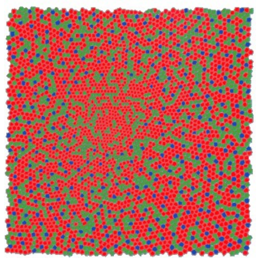
Rod / Cone sensitivity



SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Physiology of Color Vision

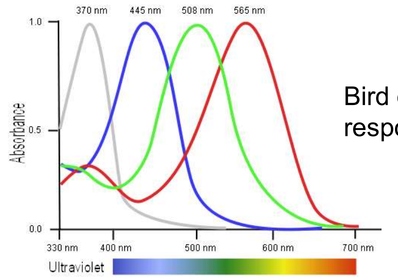




© Stephen E. Palmer, 2002

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Tetrachromatism



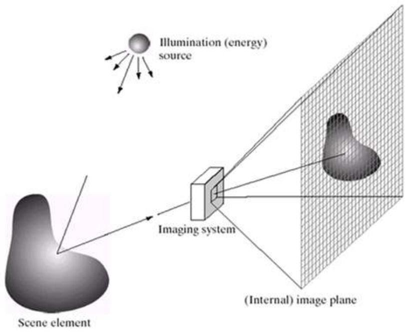
Bird cone responses

- Most birds, and many other animals, have cones for ultraviolet light.
- Some humans seem to have four cones (12% of females).

James Hays

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

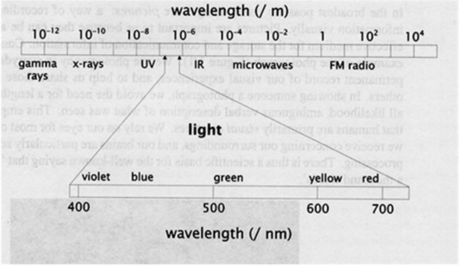
Image formation



Adapted from S. Seitz

What is light?

- Light: The visible portion of the electromagnetic (EM) spectrum.
- Light occurs between wavelengths of approximately 400 and 700 nanometers.



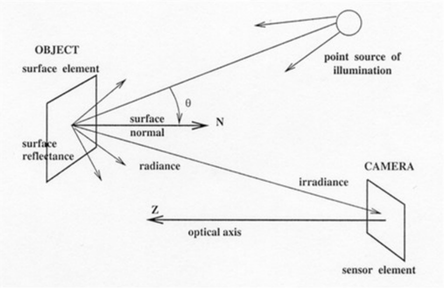
Photometric image formation

- Illumination source: Sun, light …
- Photometric measurement:
  - Perceptual brightness of visible electromagnetic energy of light.
- Optical system (lenses):
  - An object (scene) may be illuminated by the light from an emitting source.
  - The light incident on the object is reflected in a manner dependent on the surface properties of the object
  - An illuminated object will scatter light toward a lens and the lens will collect and focus the light to create the image
- Imaging sensor: CCD (**c**harge-**c**oupled **d**evice) or CMOS sensors cameras provide the 2D sensed signal.
- Digital camera: 2D sensed signal is pass to *analog-to-digital converter (sampling),* it create the digital image

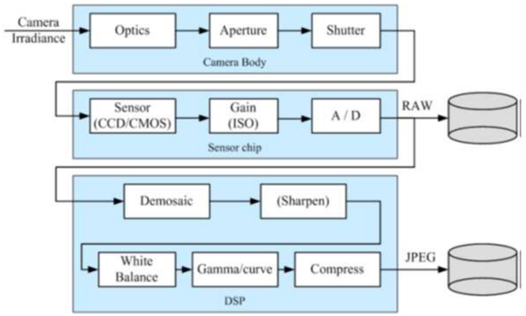
Photometric image formation

- Modeling the image formation process: 3D geometric features in the world are projected into 2D features in an image.
- A simplified model of photometric image formation is illustrated.

  - The scene is illuminated by a single source.
  - The scene reflects radiation towards the camera.
  - The camera senses it via CCD/ CMOS



Acquisition and digitization: Digital camera



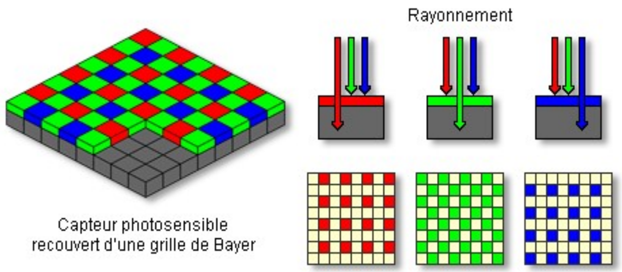
Digital camera: Image sensing and processing pipeline

Adapted from S. Seitz

Digital camera

- **Image acquisition**:
  - Optical system, aperture (capture), shutter
  - Imaging sensor: CCD/ CMOS sensor camera consists of a array of photodiodes. Each cell in the is light-sensitive diode that converts photons to electrons.
  - 2D sensed signal of image, video
- **Digitization** (ADC): Sampling and Quantization
  - Sampling the 2D sensed signal create the samples or pixels
  - Quantizing the sample values as the integer values of pixels
- **Processing** (DSP- Digital Signal Processing):
  - Cameras perform a variety of digital signal processing operations to *enhance* the image before *compressing* and *storing* the pixel values in standard format file.

Sensor array : an example



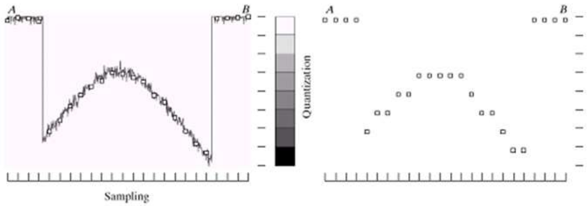
Real scene -> digital Image

Scene + Light → Sensor array → **Electric charge (Continuous signal)** → Analog/Digital conversion → **Number (discrete signal)**

Digitization = **Sampling (lấy mẫu)**
+ **Quantization (Lượng tử hóa)**

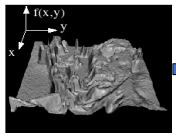
Sampling and quantization

- **Sample** the 2D space on a regular grid
- **Quantize** each sample (round to nearest integer)



Sampling and quantization

- **Sample** the 2D space on a regular grid
- **Quantize** each sample (round to nearest integer)



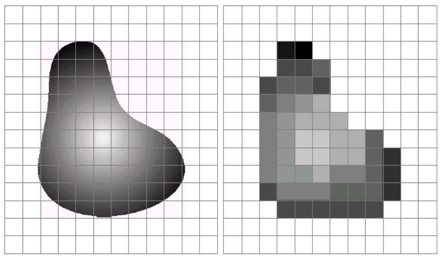
j →

i ↓

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

2D

Digital image



a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Spatial resolution (sampling)

200 X 278 | 50 X 70 | 12 X 18

Gray-level resolution (Quantization)


8 bits


4 bits


2 bits

Color spaces

- Color spaces; different types of color modes
- Color represented by vector of components
  - ❖ Red, Green, Blue (**RGB**)
  - ❖ Hue, Saturation, Value (**HSV**)
  - ❖ Luminance, chrominance (**YUV,** LUV)
  - ❖ **XYZ**
- Color convert: RGB – YUV

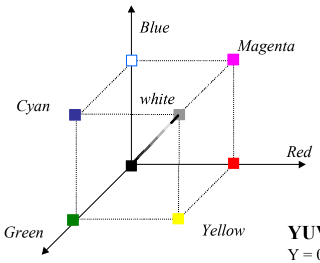
Y = 0.299R + 0.587G + 0.114B

U = 0.493 (B - Y) ; V = 0.877 (R - Y)

$$\begin{bmatrix} Y \\ C_R \\ C_B \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ 0.439 & -0.368 & -0.071 \\ -0.148 & -0.291 & 0.439 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

Color coordinate system

Màu sắc



RGB

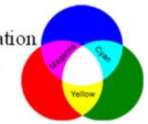
Black	(0, 0, 0)
Red	(255, 0, 0)
Green	(0, 255, 0)
Yellow	(255, 255, 0)
Blue	(0, 0, 255)
Magenta	(255, 0, 255)
Cyan	(0, 255, 255)
White	(255, 255, 255)

YUV (Luminance)
Y = 0.299R + 0.587G + 0.114B
U = R - Y
V = B - Y

Y Cb Cr (JPEG)
Cb = U/2 + 0.5
Cr = V/1.6 + 0.5

Color: Additive/Subtractive primary color

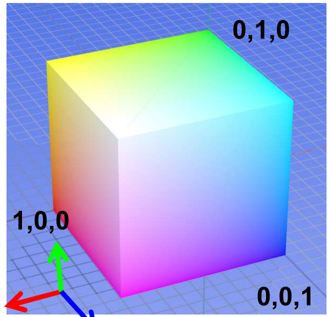
Colors: combination of RGB



- **Primary color**: Red, Green, Blue (RGB)
- **Additive colors**:
  - – Combination of RGB can be mixed to produce Cyan, Magenta, Yellow (CMY) &White.
  - – ***Additive color reproduction system***:
    - Combination of RGB to reproduce a colored light.
- **Subtractive colors** CMY can be mixed to produce RBG & black
  - – ***Subtractive color reproduction system***: A white light sequentially passes through cyan, magenta, yellow filters to reproduce a colored light.

Color spaces: RGB

Default color space



R = 1
(G=0,B=0)

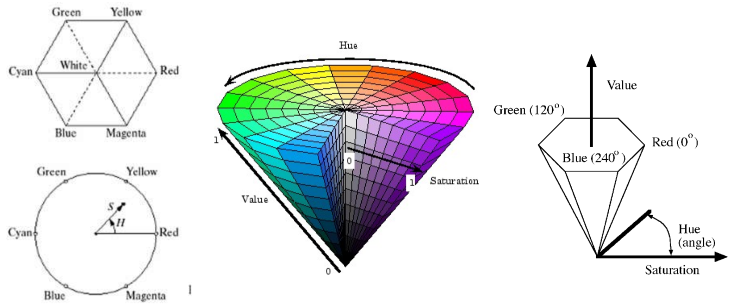
G = 1
(R=0,B=0)

B = 1
(R=0,G=0)

Any color = r*R + g*G + b*B
- Strongly correlated channels
- Non-perceptual

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

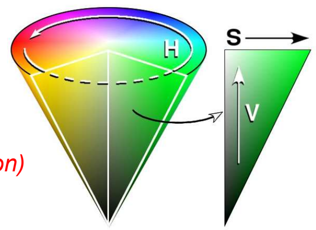
Nonlinear color spaces: HSV



- Perceptually meaningful dimensions:
  - Hue, Saturation (chroma)
  - Value (Intensity)

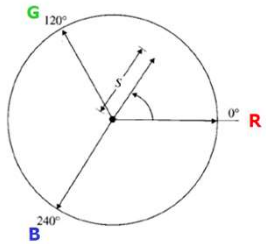
HSV (Hue – Saturation- Value)

- The Hue-Saturation-Value (HSV) color space is use for segmentation and recognition
  - Non-linear conversion
  - Visual representation of colors
- We identify for a pixel:
  - The pixel *intensity (value)*
  - The pixel *color (hue + saturation)*
- RGB does not have this seperation



HSV (Hue – Saturation- Value)

- **Hue (H)** is coded as an angle between 0 and 360
- **Saturation (S)** is coded as a radius between 0 and 1
  - S = 0 : gray
  - S = 1 : pure color
- **Value (V)** = MAX (Red, Green, Blue)



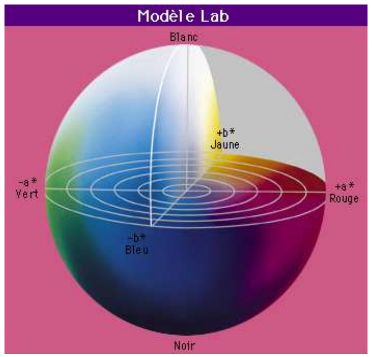
HSV (Hue – Saturation- Value)

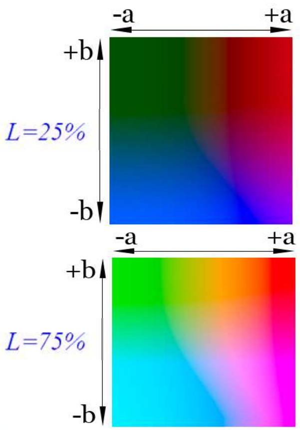
- If we know the color of the object we are looking for, can model it using a **hue interval**
- Take care, because it is an angle (periodic value)
  - Hue < 60° means nothing
    - *Is 350° smaller or bigger than 60°?*
  - Define an interval: 350° < Hue < 60° (for example)
- This interval is valid if **S**aturation > *threshold* (otherwise gray level)
- This is independant of **V**alue , which is more sensible to light conditions

Lab color space

- The **Lab** system (sometimes **L*a*b***) is based on a study from human vision
  - independant from all technologies
  - presenting colors as seen by the human eyes
- Colors are defined using 3 values
  - **L** is the luminance, going from 0% (black) to 100% (white)
  - a* represents an axis going from green (negative value, -127) to red (positive value, +127)
  - b* represents an axis going from blue (negative value, -127) to yellow (positive value,+127)

Lab color space





Color space vs. illumination conditions

- collected 10 images of the cube under varying illuminatio conditions

- separately cropped every color to get 6 datasets for the 6 different colors

Changes in color due to varying Illumination conditions

- Compute the density plot: Check the distribution of a particular color say, blue or yellow in different color spaces. The density plot or the 2D Histogram gives an idea about the variations in values for a given color

Source: Vikas Gupta, Learn OpenCV

Color space vs. illumination conditions

- Similar illumination: very compact

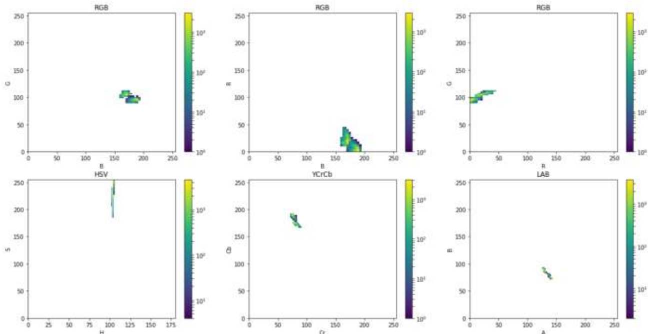


Fig.: Density Plot showing the variation of values in color channels for 2 similar bright images of **blue color**

Source: Vikas Gupta, Learn OpenCV

Color space vs. illumination conditions

- Similar illumination: very compact

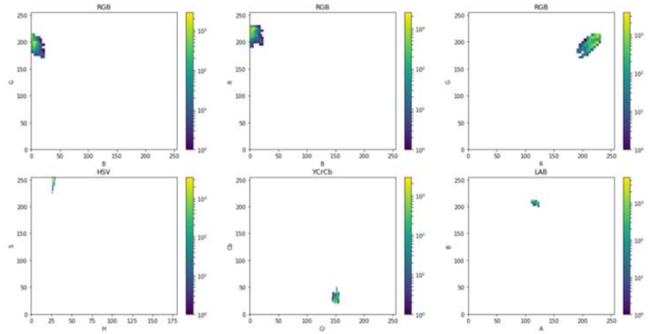


Fig.: Density Plot showing the variation of values in color channels for 2 similar bright images of **yellow color**

Source: Vikas Gupta, Learn OpenCV

Color space vs. illumination conditions

- Different illumination:

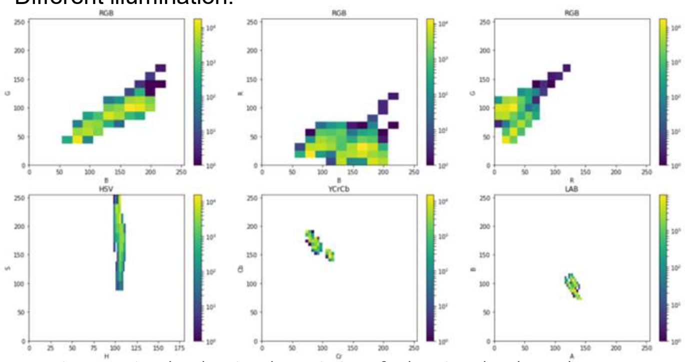


Fig.: Density Plot showing the variation of values in color channels under varying illumination for the **blue color**

Source: Vikas Gupta, Learn OpenCV

Color space vs. illumination conditions

- Different illumination:

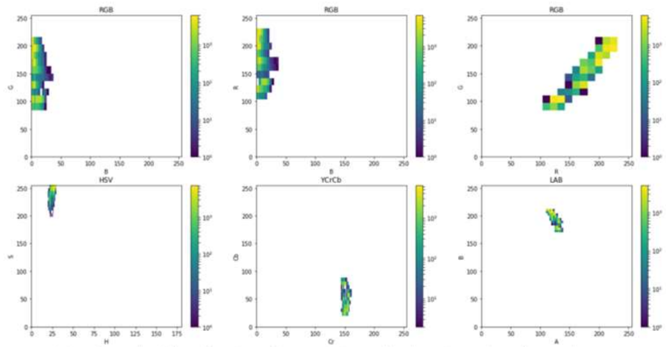


Fig.: Density Plot showing the variation of values in color channels under varying illumination for the **yellow color**

Source: Vikas Gupta, Learn OpenCV

Color space vs illumination conditions

- Different illumination:
  - RGB space: the variation in the value of channels is very hight
  - HSV: compact in **H**. Only H contains information about the absolute color ➔ a choix
  - YCrCb, LAB: compact in CrCb and in **AB**
    - Higher level of compactness is in LAB
  - Convert to other color spaces (OpenCV):
    - cvtColor(bgr, ycb, COLOR_BGR2YCrCb);
    - cvtColor(bgr, hsv, COLOR_BGR2HSV);
    - cvtColor(bgr, lab, COLOR_BGR2Lab);

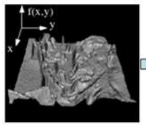
Image representation
Continuous Images as functions

- Monochromatic Image: A continuous brightness function of a number of variables f, from R^2 to R:
  - *f*(*x, y*) gives the intensity at position (x, y)
  - Realistically, we expect the image only to be defined over a rectangle, with a finite range
- A color image include 3 brightness functions of 3 color pasted together (3 color component signals). We can write this as a "vector-valued" function:

$$f(x,y)=\begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

Digital images representation

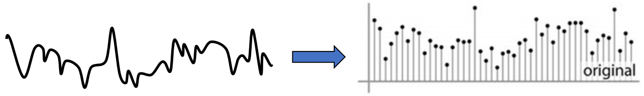
- Sample the 2D space on a regular grid is pixel
- Quantize each sample (round to nearest integer)
- Image data is represented as a matrix of integer values.



j →, i ↓

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

2D



Analog signal

Digital signal: Data

Adapted from S. Seitz

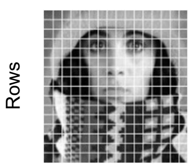
Definition: Digital images

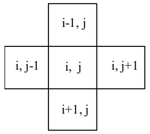
- Digital image functions *f* represented as matrices X(i,j).
- **Image data** is represented by a rectangular array of integers
- An integer represents the brightness or darkness of the monochromatic image at that point (pixel). Limited brightness integer values (8 bit) = gray levels = values 0 to 255
- **Definition**: **Digital image is a matrix X(i,j)** of pixels, N:number of rows, M: number of columns, Q: integer brightness values (levels) of pixels

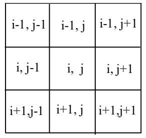
$$X(i,j)=\begin{matrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{matrix}$$

Digital gray image

- Example: Matrix X(i,j) of pixels of a gray level image
- Image data: 2D array X(i,j) of integer brightness value uint8 of pixels at coordinates (i,j).

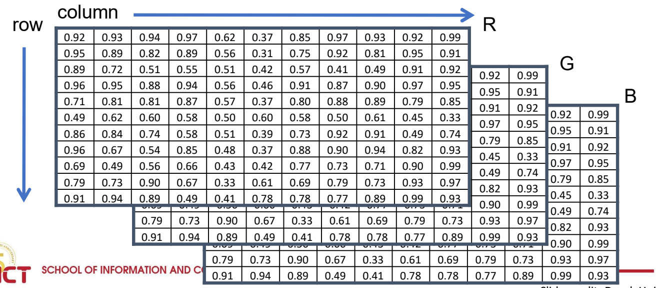






RGB color images in Matlab

- Images represented as a matrix X(i,j)
- Suppose we have a NxM RGB image called "Im"
  - Im(1,1,1) = top-left pixel value in R-channel
  - Im(y, x, b) = y pixels down, x pixels to right in the b^{th} channel
  - Im(N, M, 3) = bottom-right pixel in B-channel
- imread(filename) returns a uint8 image (values 0 to 255)
  - Convert to double format (values 0 to 1) with im2double



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

Slide credit: Derek Hoiem

Digital image format

- ❖ **Parameters for digital image formats**:
  - **Digital image resolution:** (height x width) in pixels
  - **Quantization** (bits per pixel):
    - Gray level image: 8 bits/ pixel
    - RGB color image: 24 bits/ pixel
    - Binary image: 1 bit/ pixel
- ❖ **Digital Image Storage**: file stored in two parts: Header; Data
- ❖ **Common image file formats:**
  - GIF (Graphic Interchange Format) -
  - PNG (Portable Network Graphics)
  - JPEG (Joint Photographic Experts Group)
  - TIFF (Tagged Image File Format)
  - PGM (Portable Gray Map)
  - FITS (Flexible Image Transport System)

Digital video format

- Parameters for digital video formats
  - Digital image resolution (height x width) in pixels
  - Quantization (bits per pixel)
  - Frame rate (frames per second)
- Standard video file formats
  - AVI, M-JPEG,
  - H26X (ITU_T:H.261, H.263, H.263, H264)
  - MPEG-1, MPEG-2, MPEG-4 Part 10 / H264 AVC,mp4…


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY
Thank you for your attention!
soict.hust.edu.vn/
fb.com/groups/soict