COMP7055
Computer Vision

Chapter 1

Introduction + Image Acquisition

Teaching Team

- Course Instructor
  - Prof. Xiaoqing GUO
  - Email: xiaoqingguo@hkbu.edu.hk
- Teaching Assistant
  - Mr. Zhenshun LIU
  - Email: cszsliu@comp.hkbu.edu.hk

Lectures and Lab Hours

- Lecture -- Thursday 18:30 – 20:20
- Lab -- Thursday: 20:30 – 21:20 (TA will assist)
  - Venue: FSC801C, FSC801D

Schedule

Week	Date	Time	Lecture/Lab	Venue	Remarks
1	15 Jan	18:30 – 20:20	Lecture (Introduction + Image Acquisition)	FSC801C,D	Conventional CVPR algorithms
	15 Jan	20:30 – 21:20	No lab	FSC801C,D	
2	22 Jan	18:30 – 20:20	Lecture (Image enhancement)	FSC801C,D	
	22 Jan	20:30 – 21:20	Lab 1	FSC801C,D	
3	29 Jan	18:30 – 20:20	Lecture (Feature extractor)	FSC801C,D	
	29 Jan	20:30 – 21:20	Lab 2	FSC801C,D	
4	5 Feb	18:30 – 20:20	Lecture (Image classification and segmentation)	FSC801C,D	
	5 Feb	20:30 – 21:20	Lab 3	FSC801C,D	

Schedule

Week	Date	Time	Lecture/Lab	Venue	Remarks
5	12 Feb	18:30 – 20:20	Lecture (Deep learning for CVPR) + Course project briefing	FSC801C,D	Deep learning for CVPR
	12 Feb	20:30 – 21:20	Quiz (Conventional CVPR)	FSC801C,D	
6	19 Feb	18:30 – 20:20	University Holidays: Chinese New Year		
	19 Feb	20:30 – 21:20			
7	26 Feb	18:30 – 20:20	Lecture (Segmentation)	FSC801C,D	
	26 Feb	20:30 – 21:20	Lab 4	FSC801C,D	
8	5 Mar	18:30 – 20:20	Lecture (Object detection)	FSC801C,D	
	5 Mar	20:30 – 21:20	Lab 5	FSC801C,D	
9	12 Mar	18:30 – 20:20	Lecture (Temporal processing)	FSC801C,D	
	12 Mar	20:30 – 21:20	Lab 6	FSC801C,D	
10	19 Mar	18:30 – 20:20	Lecture (Data generation)	FSC801C,D	
	19 Mar	20:30 – 21:20	Lab 7	FSC801C,D	
11	26 Mar	18:30 – 20:20	Lecture (CVPR review) + Quiz (DL-based CVPR)	FSC801C,D	
	26 Mar	20:30 – 21:20	Project consultation (optional)	FSC801C,D	

Schedule

Week	Date	Time	Lecture/Lab	Venue	Remarks
12	2 April	18:30 – 20:20	University Holidays: Easter		
	2 April	20:30 – 21:20			
13	9 April	18:30 – 20:20			
	9 April	20:30 – 21:20			
14	16 April	18:30 – 20:20	Group Project Presentation	FSC801C,D	By students
	16 April	20:30 – 21:20	Group Project Presentation	FSC801C,D	
15	23 April	18:30 – 20:20	Group Project Presentation	FSC801C,D	
	23 April	20:30 – 21:20	Group Project Presentation	FSC801C,D	

Course Learning Outcomes

- Explain basic theories and techniques in computer vision and pattern recognition
- Identify various approaches of computer vision and pattern recognition, and design the components of computer vision and pattern recognition systems
- Describe and discuss the basic functions and methods for image processing
- Design and develop a computer vision and pattern recognition application prototype

Evaluation

- Continuous Assessment 50%
  - *Assignment* (2 in-class quizzes) *5%*
  - *Lab exercises* (7 labs) *10%*
  - *Group project* (proposal + presentation + report + code) *35%*

- Final Exam 50%

In order to pass the course, students have to obtain (i) ***at least 30%*** *in the final examination and (ii)* ***at least 35%*** *in the total score.*

Quiz and Final Exam

- Quiz -- 1 hour per quiz
  - Closed-book exam
  - LLM prohibited
  - Mobile phones prohibited
  - Only one A4 sheet of notes allowed
  - Regular calculators permitted

- Final exam -- 2 hour
  - No book, no LLM, no phones, no notes. Only regular calculators permitted

Lab Exercise

- Submit the code to Moodle in 48h after each course.
- Late submission is not accepted without previous arrangement with the instructor.
- If approved, late submission receives 10% per day penalty.

Group Project

- 2-3 students form a group
- Focused on one task (will specify the options later)
- Related to our course (segmentation, detection, classification, ...)
- 10 mins presentation per person + 5 mins Q&A per group (present your responsible part(s))
- One proposal and report per-group (specify your role)
- Code should be submitted as well

Note: Pay attention to workload and contribution balance.

Group Project Timeline

- **12 Feb 2026 (Week 5)**
  - ✓ Project briefing
- **19 Feb 2026 (Week 6)**
  - ✓ Submit group member list
- **20 Mar 2026 (Week 10, 20%)**
  - ✓ Submit a project proposal and plan, which includes but not limited to, your preliminary studies, rationale of your proposed methodology and your initial design **via Moodle.** A good proposal shows you understand what you're doing.
- **16/23 April 2026 (Week 14-15, 40%)**
  - ✓ presentation and demonstration, and Q&A
- **30 April 2026 (40%)**
  - ✓ Submit final report, with your ppt and source codes **via Moodle**

***Late submission Policy: 10% deduction per day**

References

Reference Books:

- R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, 2nd Edition, John Wiley & Sons, Inc, 2000.
- L.G. Shapiro and G. C. Stockman, Computer Vision, Prentice Hall, 2001.
- Richard Szeliski, Computer Vision: Algorithms and Applications, Springer, 2010
- Klette, Reinhard, Concise computer vision - An introduction into theory and algorithm, Springer, 2014.
- Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
- Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J. Dive into deep learning 2019

Hands-on:

- Joseph Howse , Joe Minichino, Learning OpenCV 4 Computer Vision with Python 3: Get to grips with tools, techniques, and algorithms for computer vision and machine learning, 3rd Edition, 2020
- Alberto Fernandez Villan, Mastering OpenCV 4 with Python: A practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7, 2019

Research:

- IEEE Transactions on Pattern Analysis and Machine Intelligence
- IEEE Transactions on Image Processing
- IEEE Transactions on Information Forensics and Security
- International Journal on Computer Vision
- Journal of Pattern Recognition
- Conferences: CVPR, ICCV, ECCV

FAQ (1)

1. Does this course require math
   - *Yes*
   - *Able to read and understand the physical meaning of math*
   - *Traditional CVPR algorithms need some calculations*

2. Does this course require programming? If so, what language(s) will be used?
   - *Yes, Python*
   - *IDE: VSCode (you can use other IDE, e.g.pytram, Jupyter, Spyder)*
   - *API: OpenCV, sklearn (conventional CVPR) and PyTorch (deep learning)*
   - *Will schedule lab sessions for OpenCV and PyTorch*

FAQ (2)

3. Who should take this course?

- ➢ *Those who are interested in computer vision and pattern recognition as well as its applications*
- ➢ *Those who would like to work in AI and related areas*
- ➢ *Those who would like to do research in computer vision, pattern recognition and machine learning*

4. Will this course teach computer vision and pattern recognition in-depth?

- ➢ *No. This course gives a broad overview of computer vision and pattern recognition as well as its applications*
- ➢ *Students who are interested in this area could contact me for more materials.*

Course Contents

- Introduction to computer vision and pattern recognition
  - Computer vision applications
  - Image acquisition
  - Color models/spaces

Course Contents

- Introduction to computer vision and pattern recognition
  - Computer vision applications
  - Image acquisition
  - Color models/spaces

What is Computer Vision?

- "**Computer vision** is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do." [From Wikipedia]
- Teach machine/computer how to see and understand images

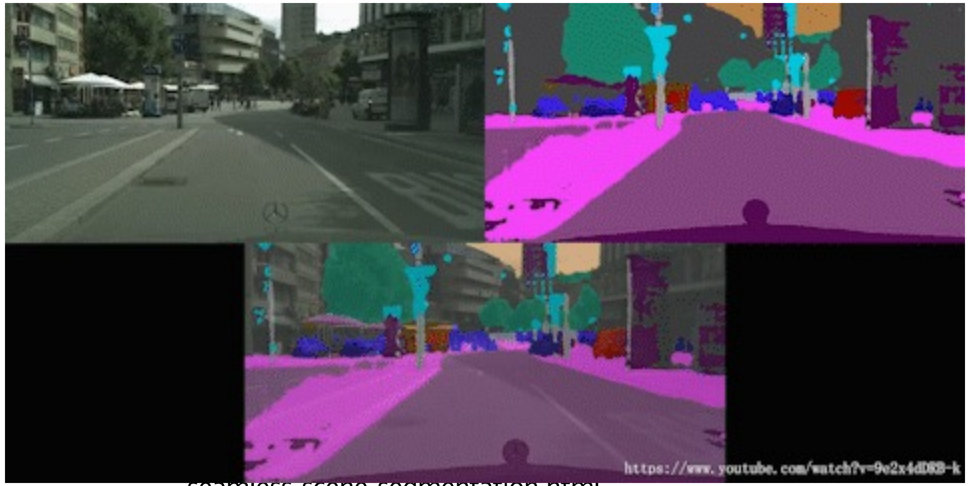
What is Pattern Recognition?

- "The field of **pattern recognition** is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories" [From Wikipedia]
- Teach machine/computer how to recognize a pattern

Computer Vision and Pattern Recognition Applications: Autonomous Driving

- Scene segmentation
- Road sign recognition
- Car recognition
- Pedestrian recognition

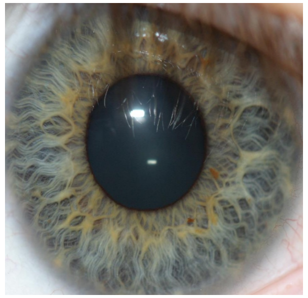
https://www.greencarcongress.com/2014/11/20141114-toshibaimage.html



https://www.youtube.com/watch?v=9e2x4dDRB-k -seamless-scene-segmentation.html

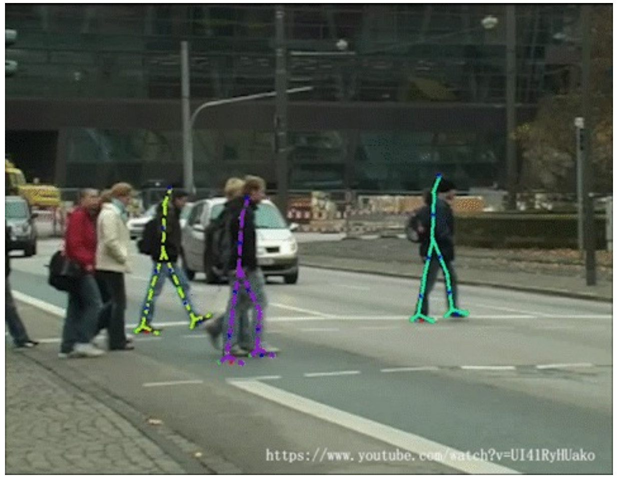
Computer Vision and Pattern Recognition Applications: Biometrics Recognition

- Face
- Fingerprint
- Iris
- Gait
  - …

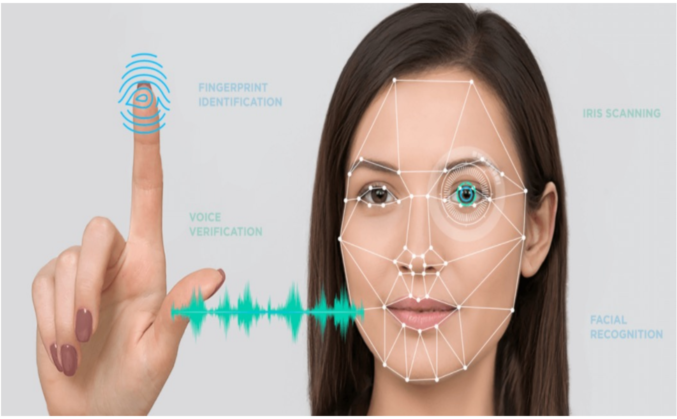


https://en.wikipedia.org/wiki/Iris_recognition

Border Control



https://www.youtube.com/watch?v=UI41RyHUako



https://mobidev.biz/blog/multimodal-biometrics-verification-system-ai-machine-learning

FaceID in iPhone X

Announced on 12 September 2017

Computer Vision and Pattern Recognition Applications: Human behavior understanding

https://www.youtube.com/watch?v=Uz
hJiO9oDKg

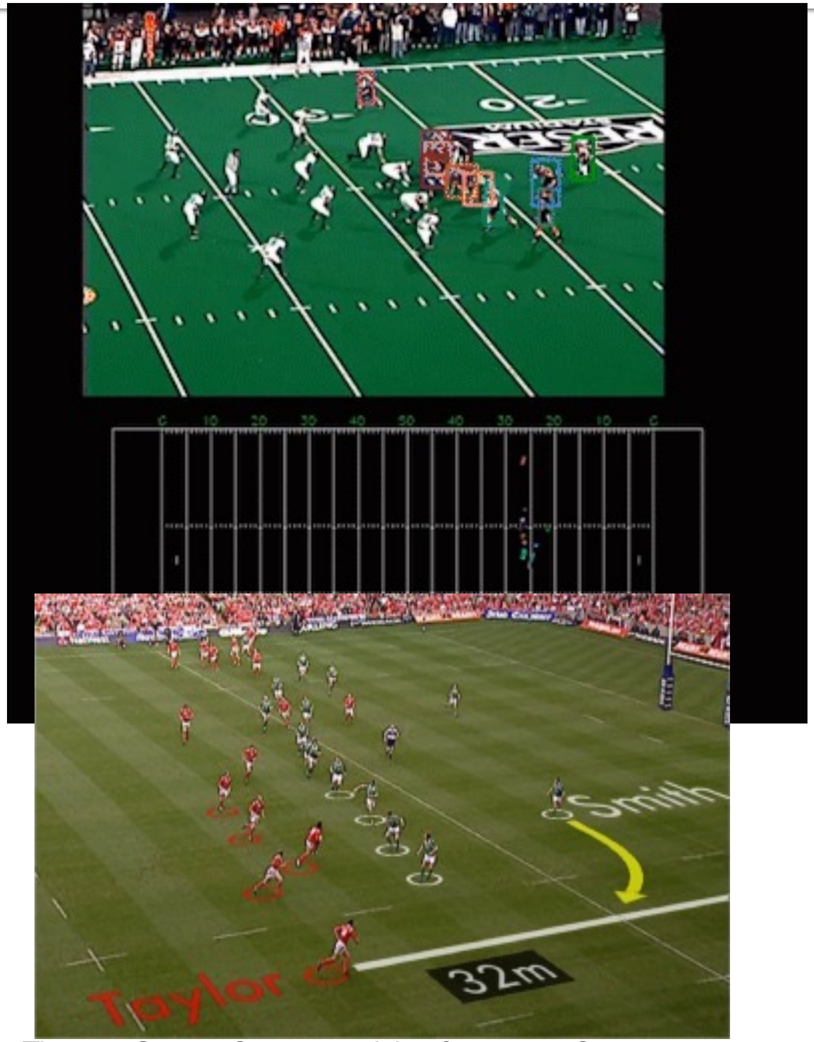
X-box Kinect

https://www.youtube.com/watch?v=m-_t_aIIQwg

https://www.youtube.com/watch?v=pW6nZXeWlGM

Computer Vision and Pattern Recognition Applications: Sports Analysis

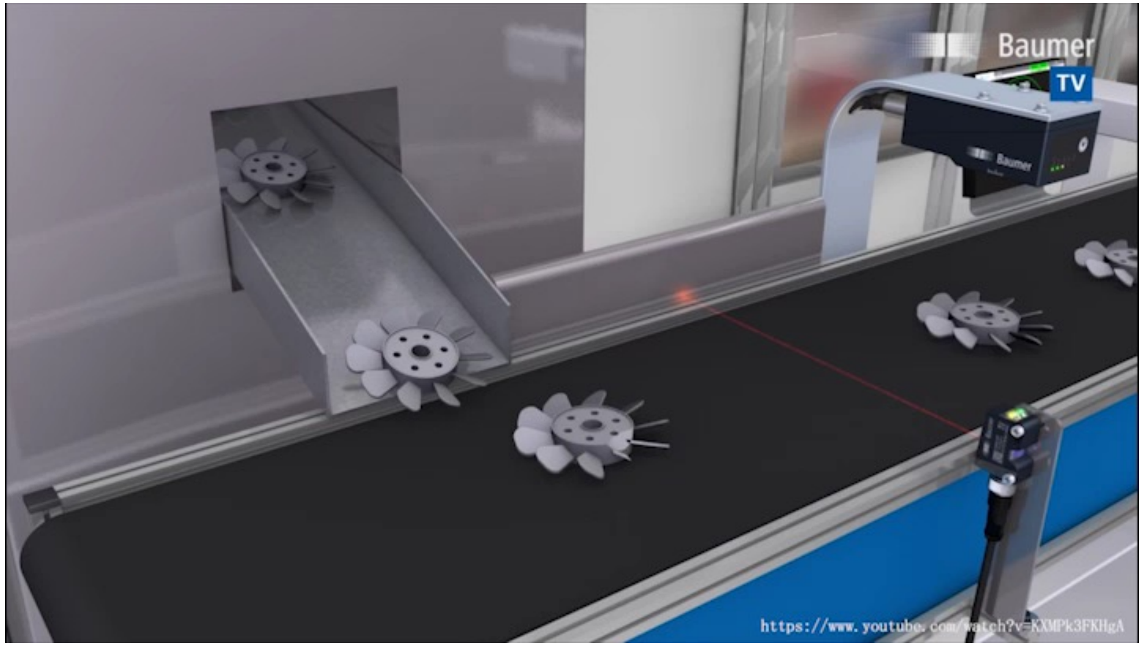
https://www.youtube.com/watch?v=7IVUyhMkyms



Thomas G et al. Computer vision for sports: Current applications and research topics, CVIU 2017

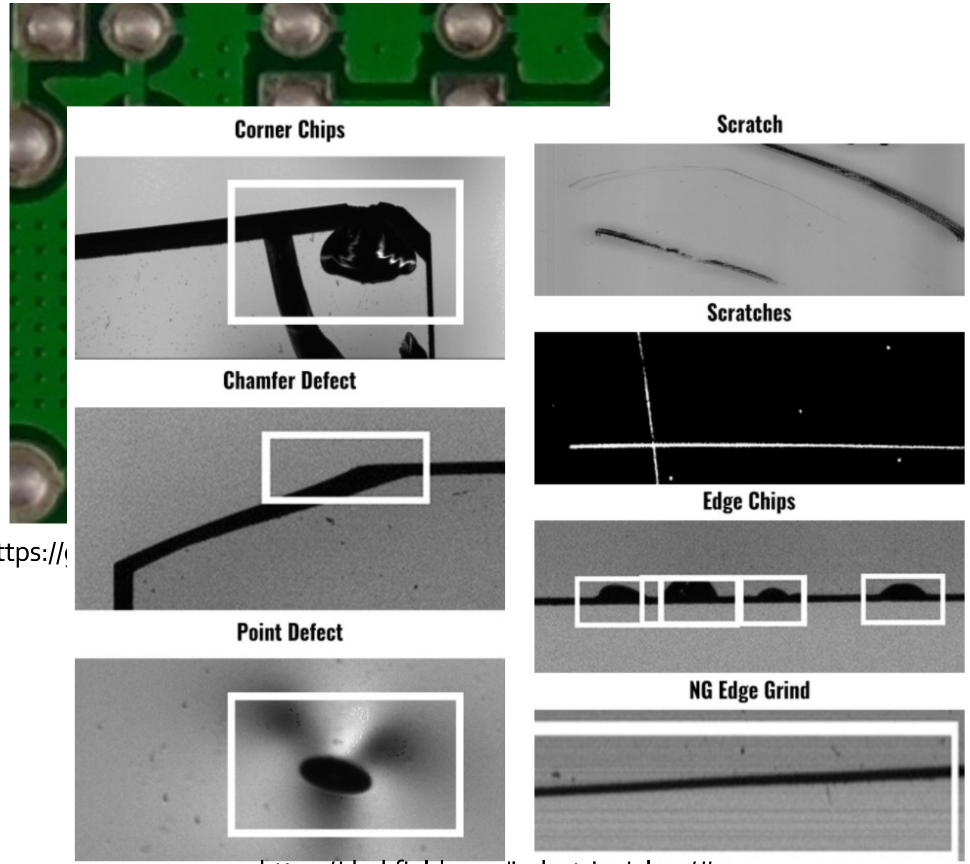
Computer Vision and Pattern Recognition Applications: Defect Recognition

- Defect recognition



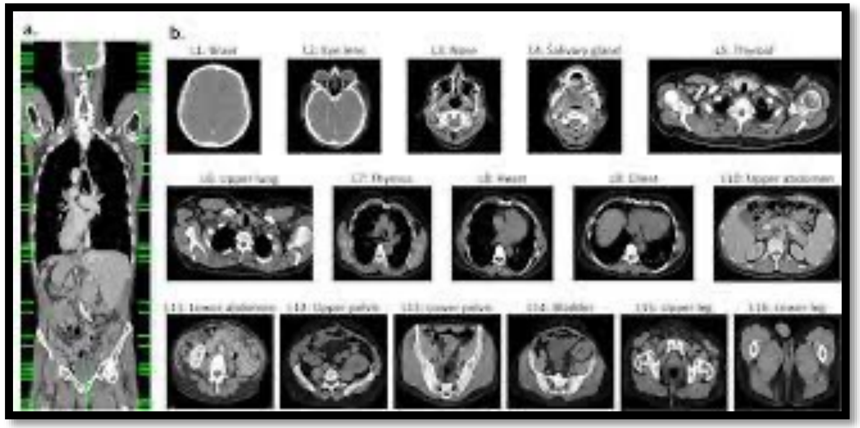
https://www.youtube.com/watch?v=KXMPk3FKHgA

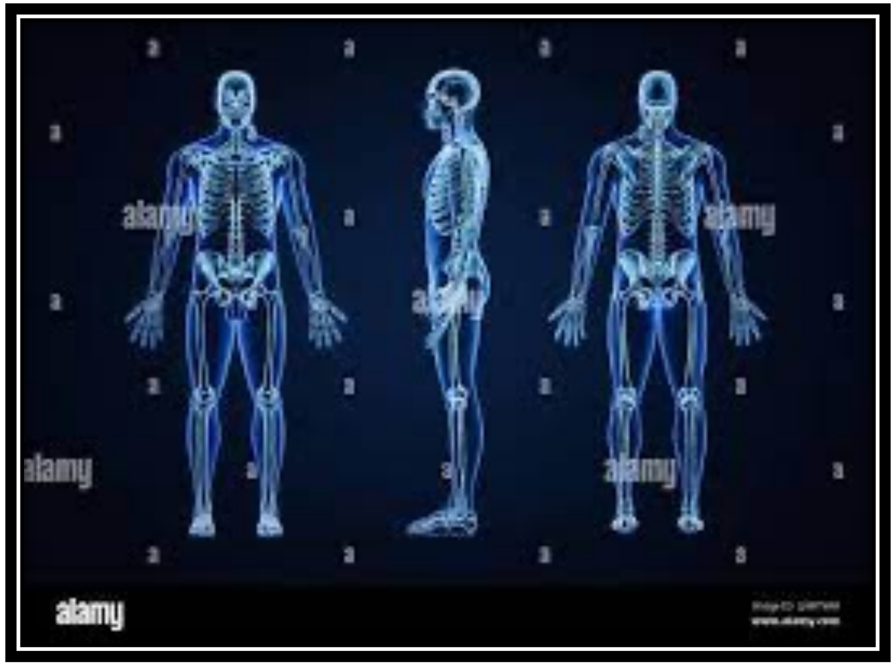
https://



https://darkfield.com/industries/glass/#

Computer Vision and Pattern Recognition Applications: Medical Imaging

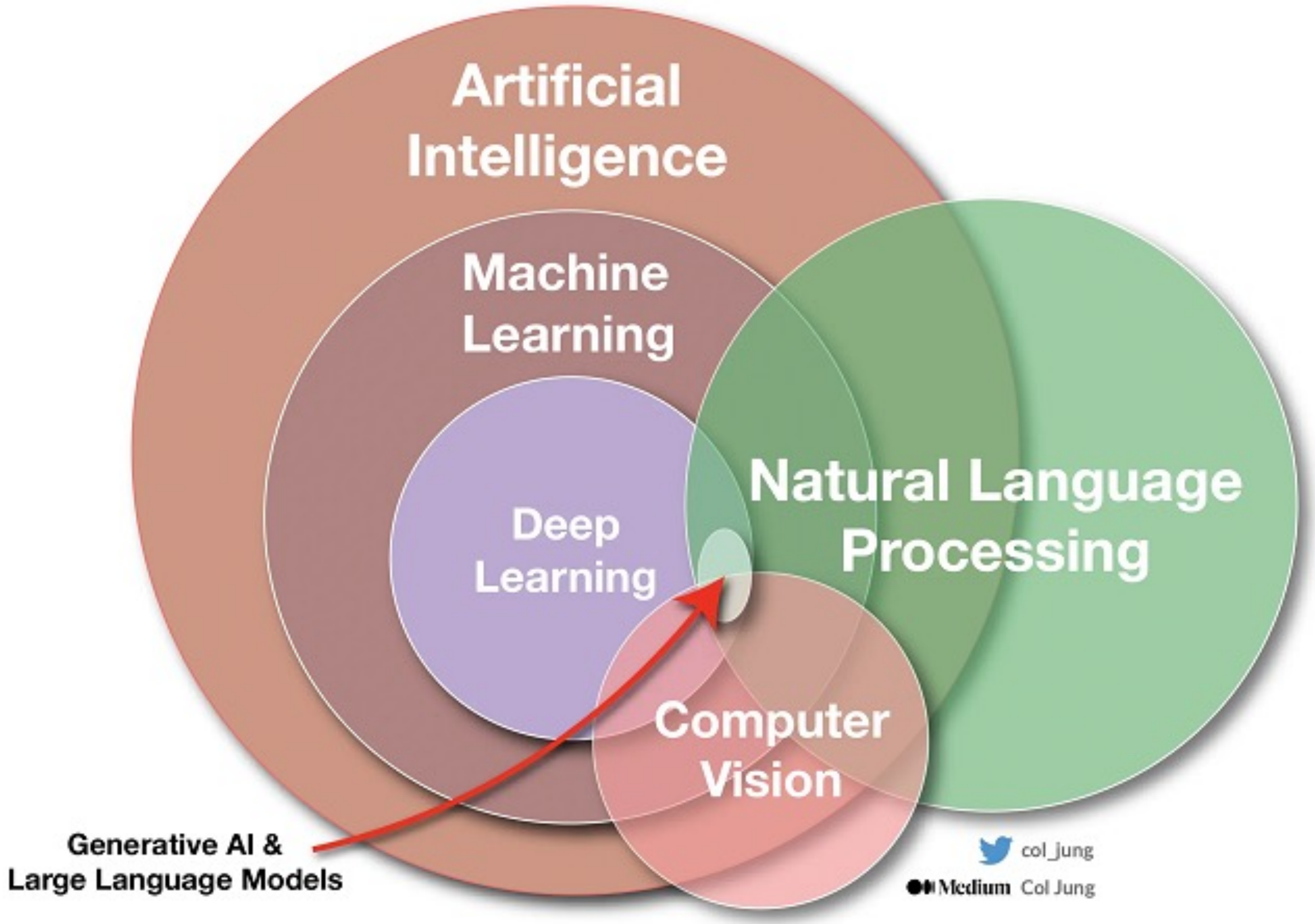




Video: https://www.youtube.com/watch?v=9fAcjfnWyso?

AI 2.0: Generative AI and Foundation Models

- Multimodality Computer Vision Applications
  - ✓ Interactive chat/conversation (LLM)
  - ✓ Prompt engineering
  - ✓ Generation ability: Gen Arts



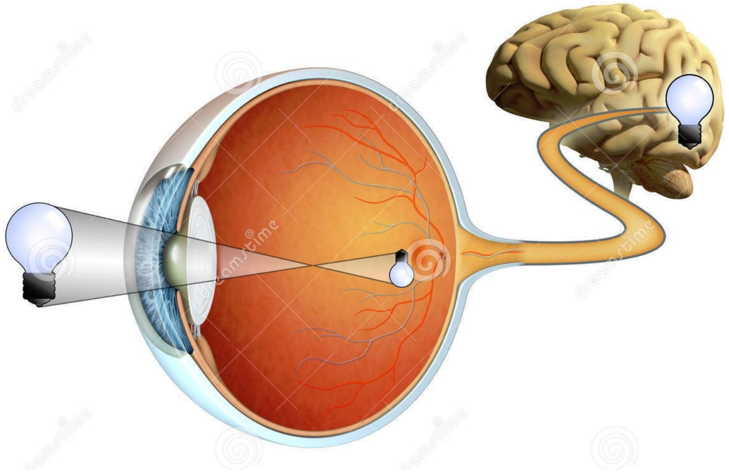
Source: https://docs.thinkwisesoftware.com/docs/sf/llm

Course Contents

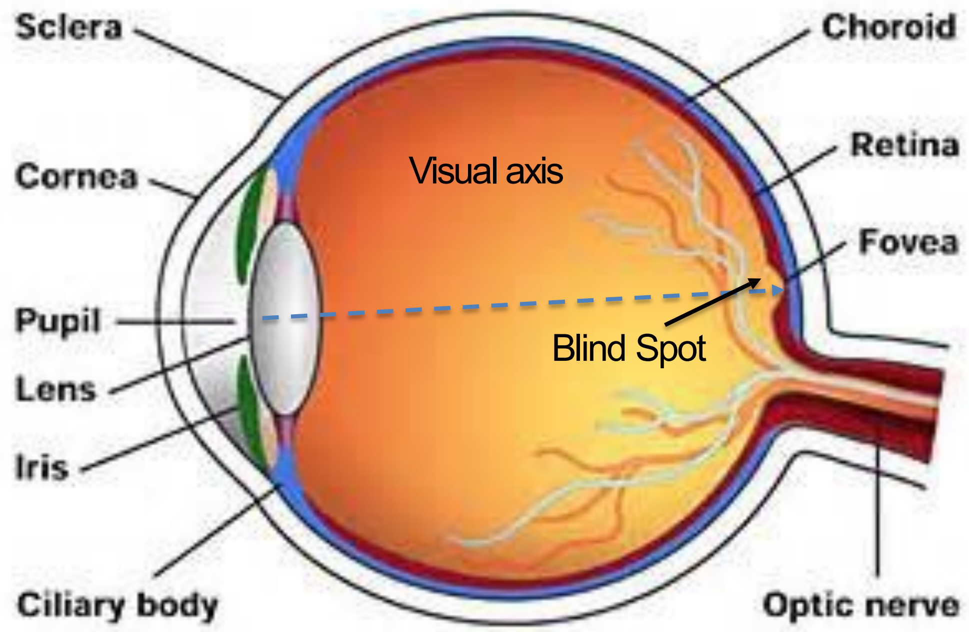
- Introduction to computer vision and pattern recognition
  - Computer vision applications
  - Image acquisition
  - Color models/spaces

Human Visual System

- Knowledge of how images form in the eye can help us with processing digital images
  - Perceptual based Image Processing
    - Focus on perceptually significant information
    - Discard perceptually insignificant information
- Issues:
  - **Biological**
  - **Psychophysical**



Structure of Human Eye



- Light passes through **cornea**, **iris**, **lens** and form image on **retina**.
- Retina contains 2 types of photoreceptors: **rods** and **cones**
- **Fovea** is about 1.5 mm in width, contains about 337,000 cones
- Focal length about 17 mm

Rods and Cones in the Retina

The retina is covered with light receptors called **cones** (6-7 million) and **rods** (75-150 million)

- **Cones** are concentrated around the **fovea**, detect color at bright light – photopic vision
- **Rods** spread at back of eye, general vision – scotopic vision.

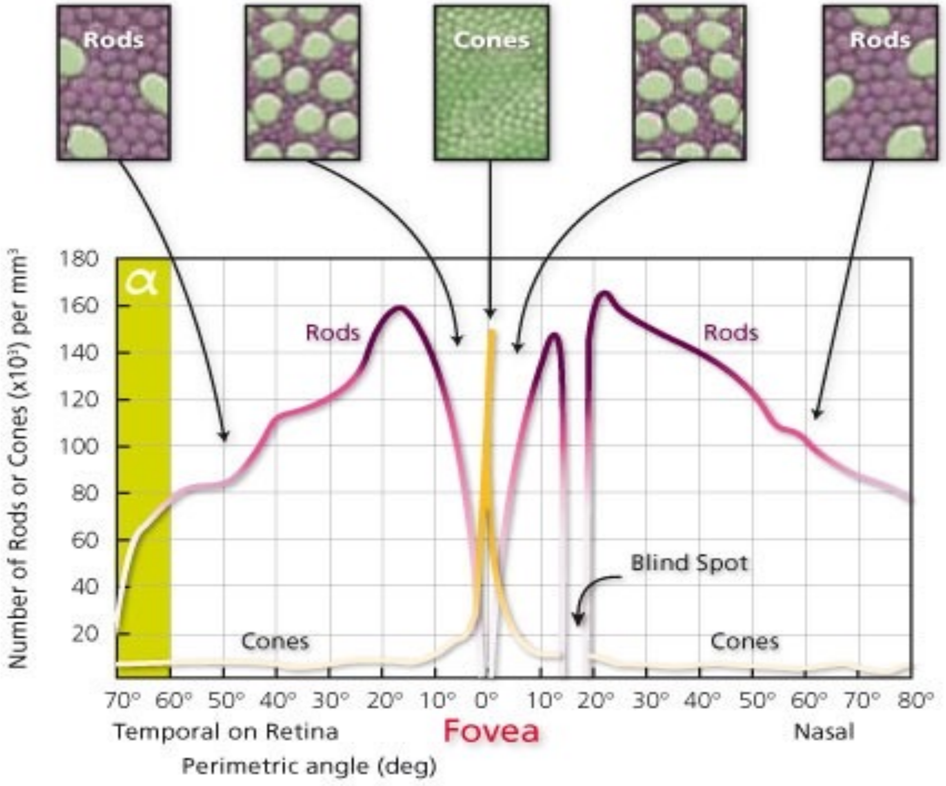


Image Formation in the Eye

- Muscles within the eye can be used to change the shape of the lens allowing us focus on objects that are near or far away
- An image is focused onto the retina causing rods and cones to become excited which ultimately send signals to the brain

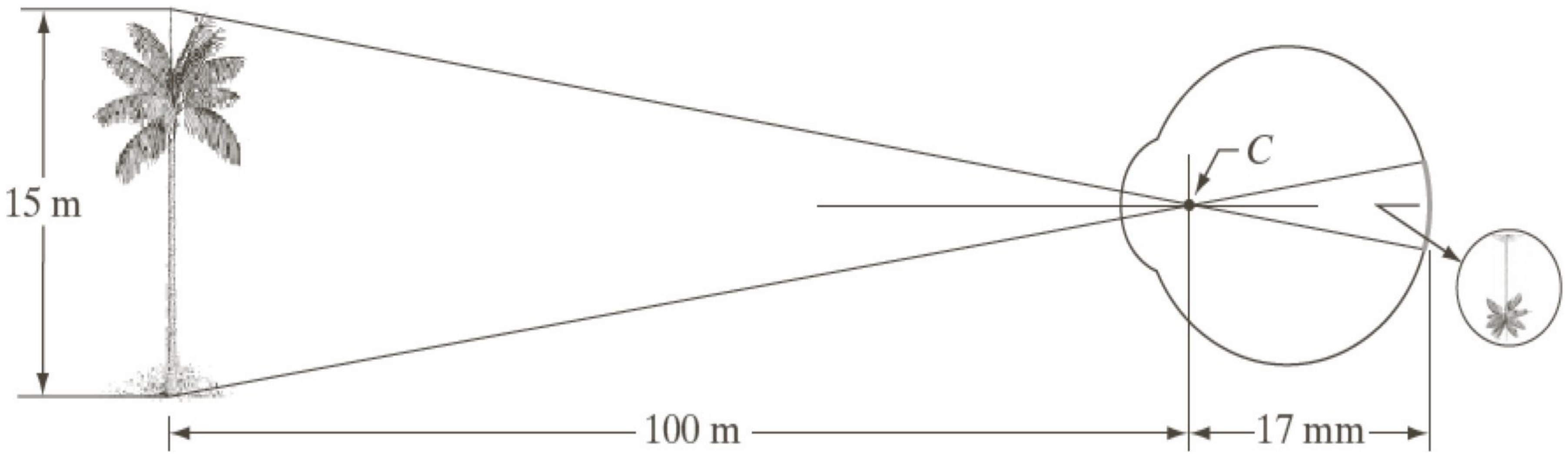
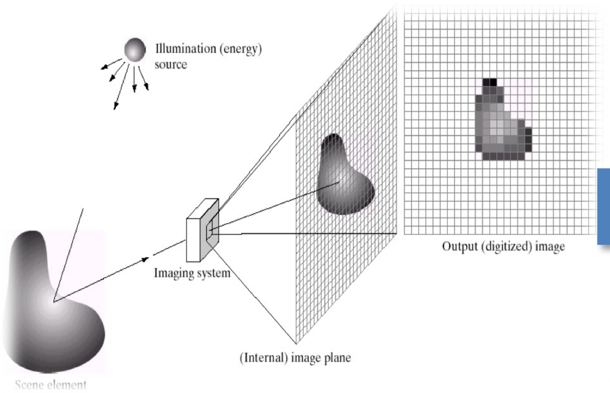


Image Acquisition



Object/Scene Element

Camera,
Scanner,
X-ray Machine,
etc

Computer, LCD/LED TV,
Printer, Film, etc.…

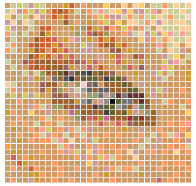
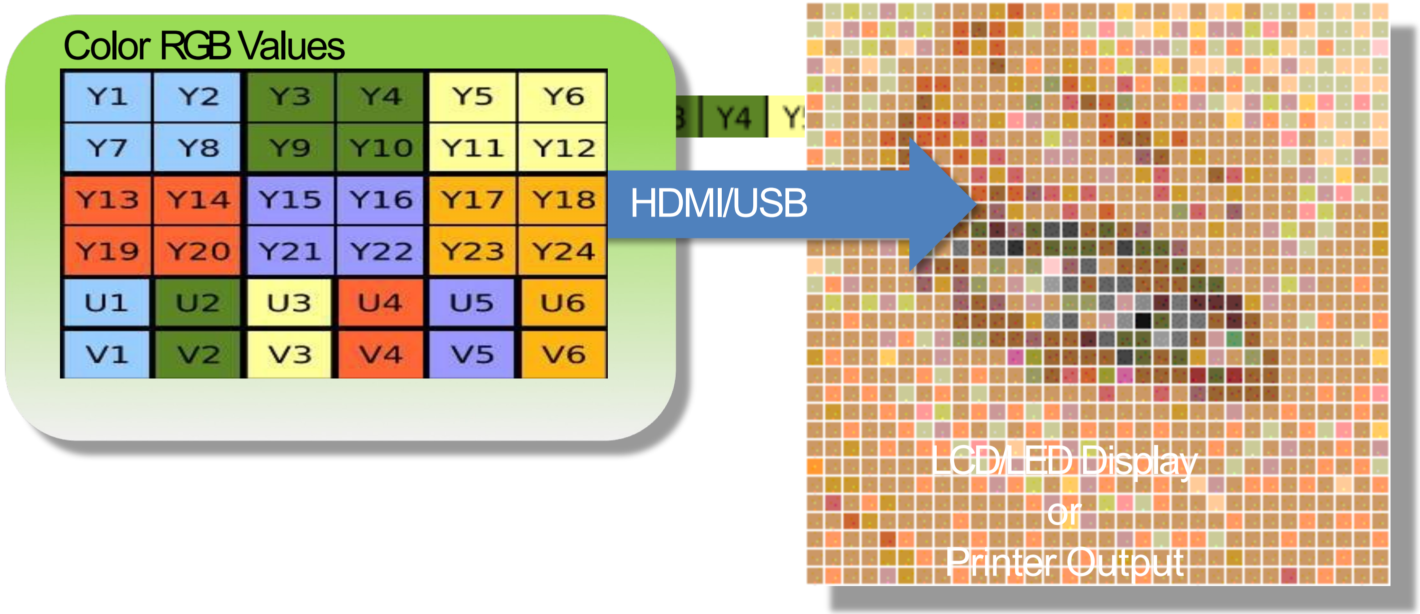
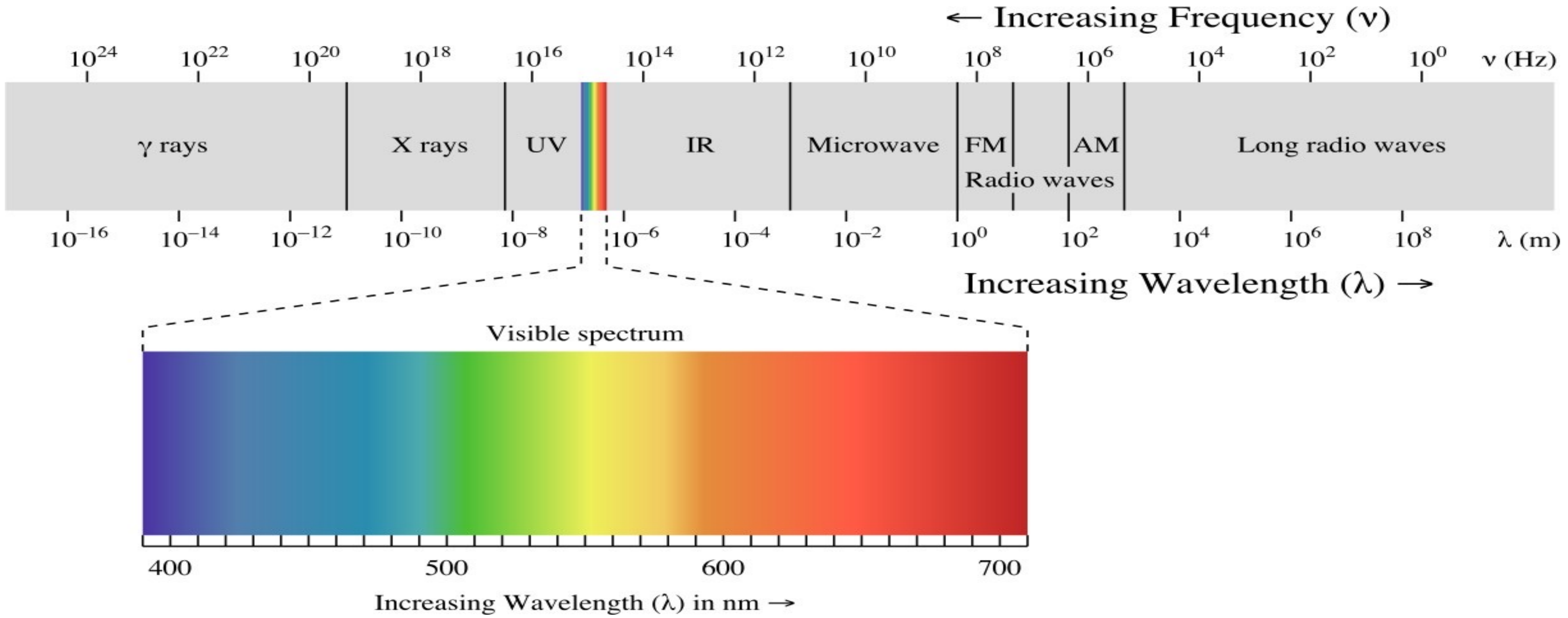


Image Representation



Sources of Image Data

- Normally images are formed from Electromagnetic (EM) radiation



https://en.wikipedia.org/wiki/Electromagnetic_spectrum

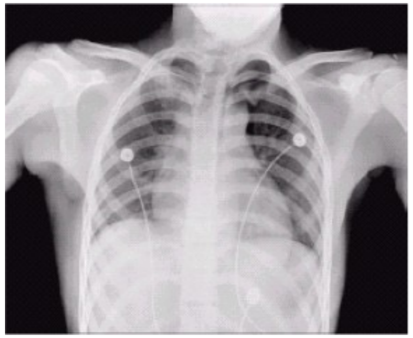
How to Visualize Light

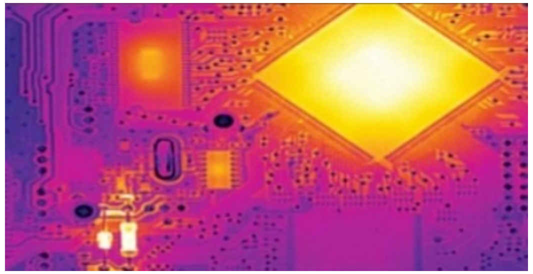
Three major sources

- **Reflected light**
  - Photographic images, scanned images, …
- **Absorbed light**
  - X-ray image, …
- **Emitted light**
  - Infrared image, …

Photo of natural scene

X-Ray of Chest

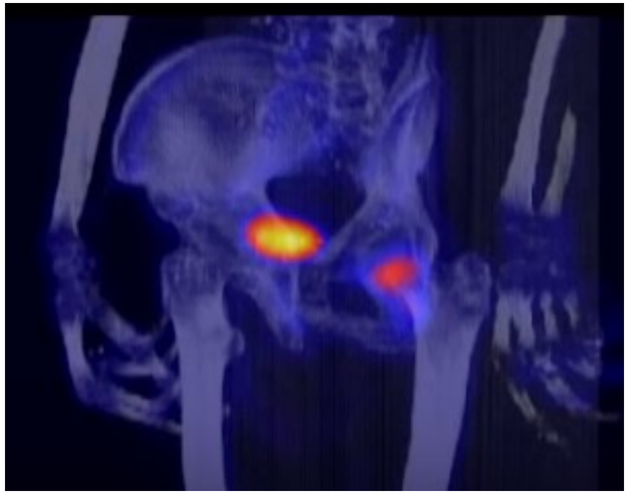


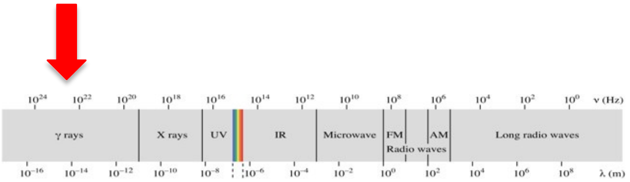


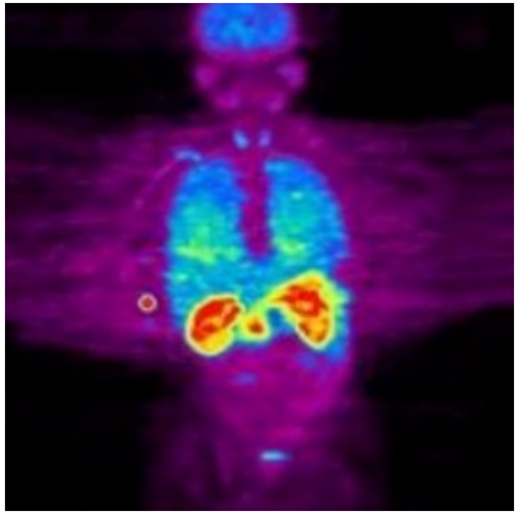
Thermal/IR view of a Chip

Gamma Ray Imaging

1. Inject patient with isotope
2. Positron emission tomography
3. Gas cloud in constellation Cygnus
4. Radioactive valve



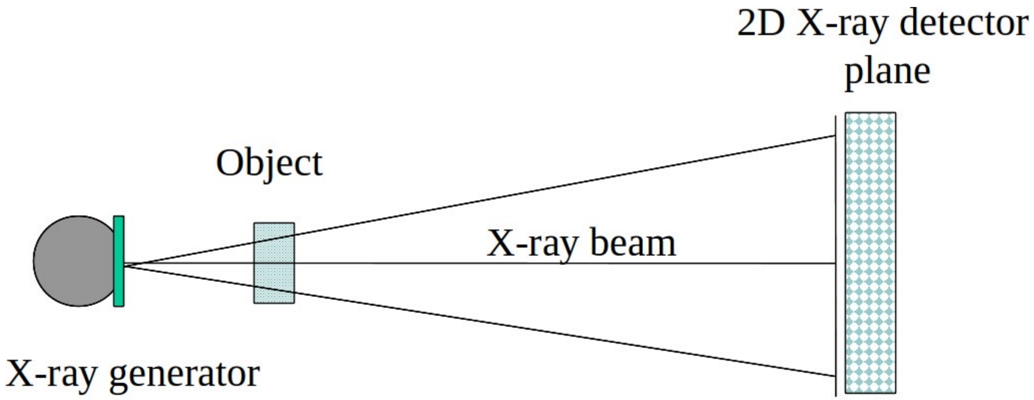




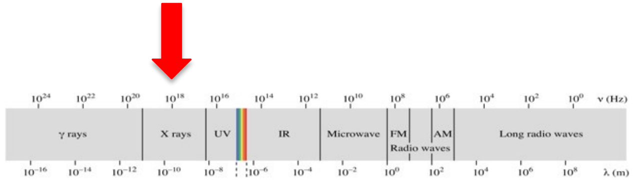
Gamma ray exposed images (emitted light)

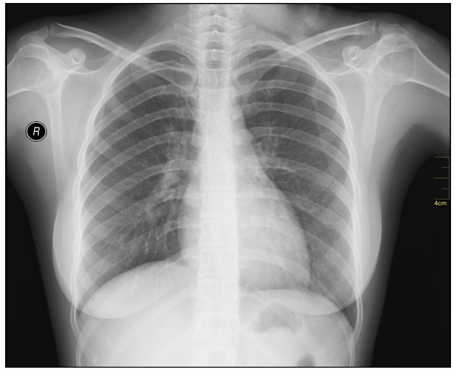
YouTube Example:
https://www.youtube.com/watch?v=ghzQj_E3A-g

X-Ray Imaging



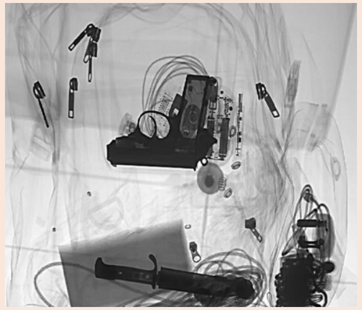
Object is Magnified





A projection of **absorbed** X-Rays as they pass through a patient.

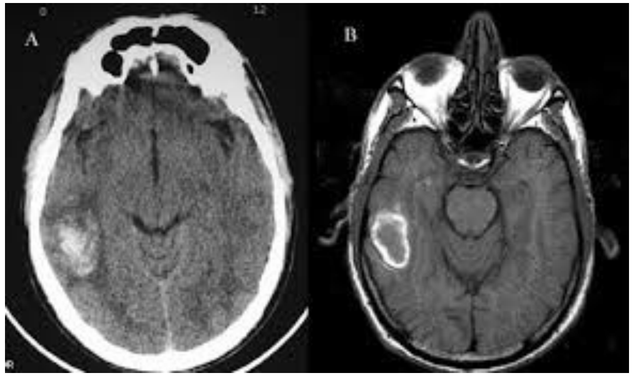
Other examples



Airport Inspection

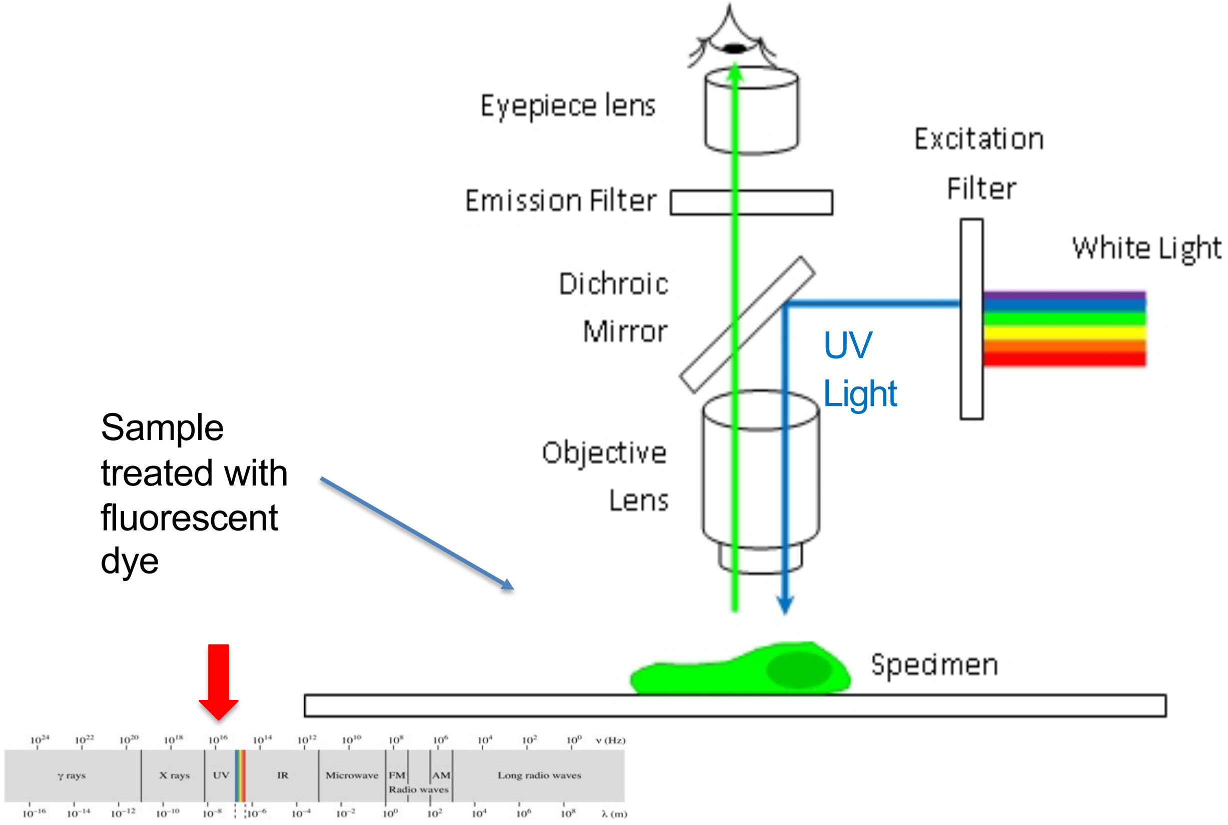
Computer Tomography (CT)

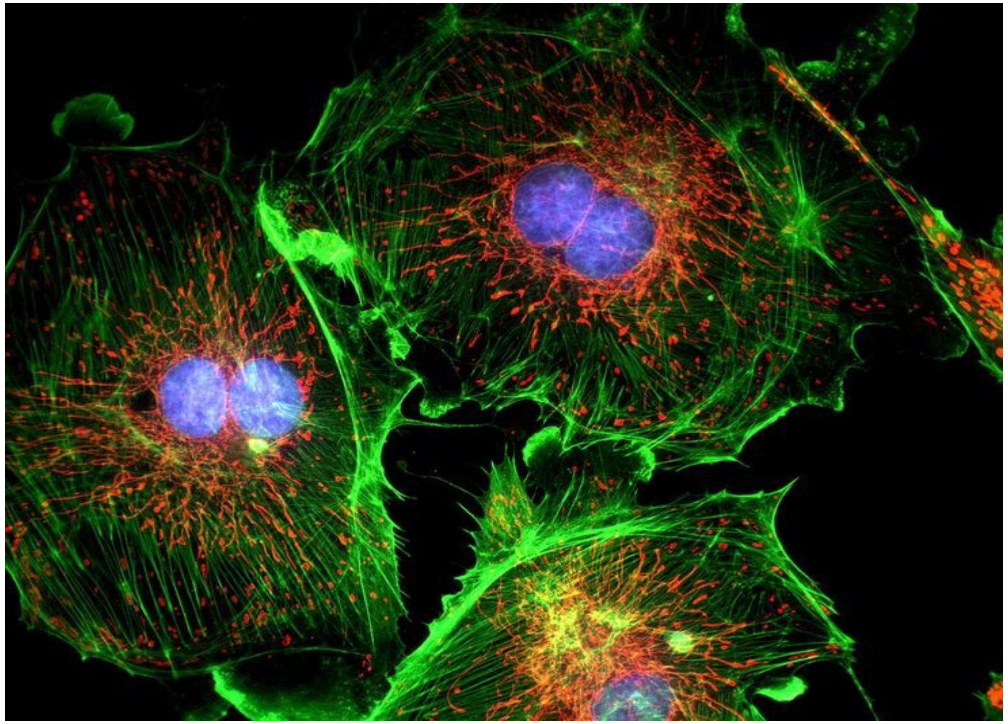
- Another important use of X-Rays is Computer Tomography (CT) scanning. It can create 3D image of object from many projections.



Ultraviolet Imaging

- Fluorescence Microscopy

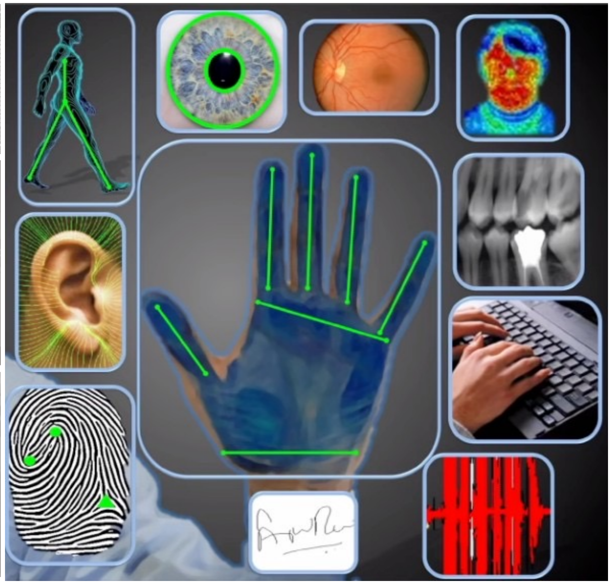


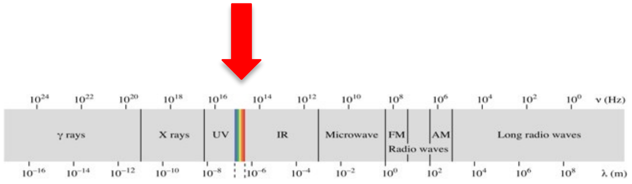


YouTube Example:
https://www.youtube.com/watch?v=PCJ13LjncMc

Visible-Band Imaging

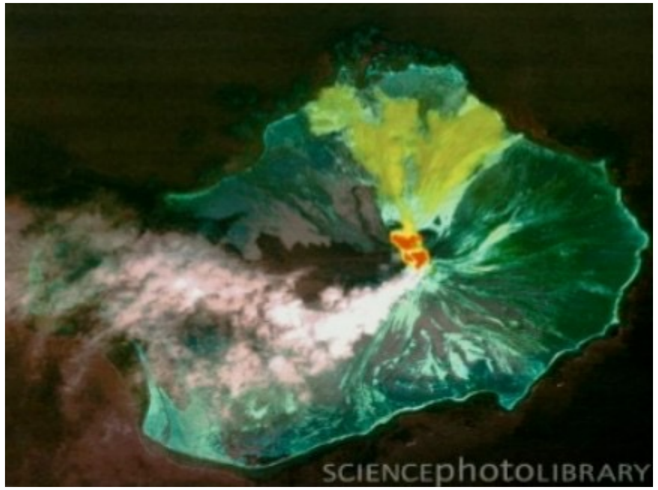
- Smartphones
- Consumer Digital Imaging
- Light Microscope
- Remote Sensing (Satellite)
- Manufacturing/Inspection
- Spacecraft Imaging
- License Plate Recognition
- Biometrics

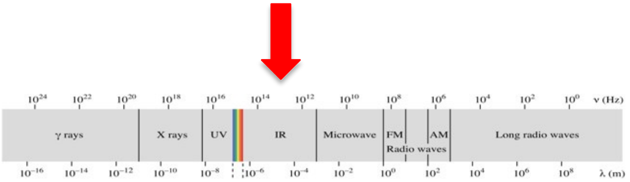




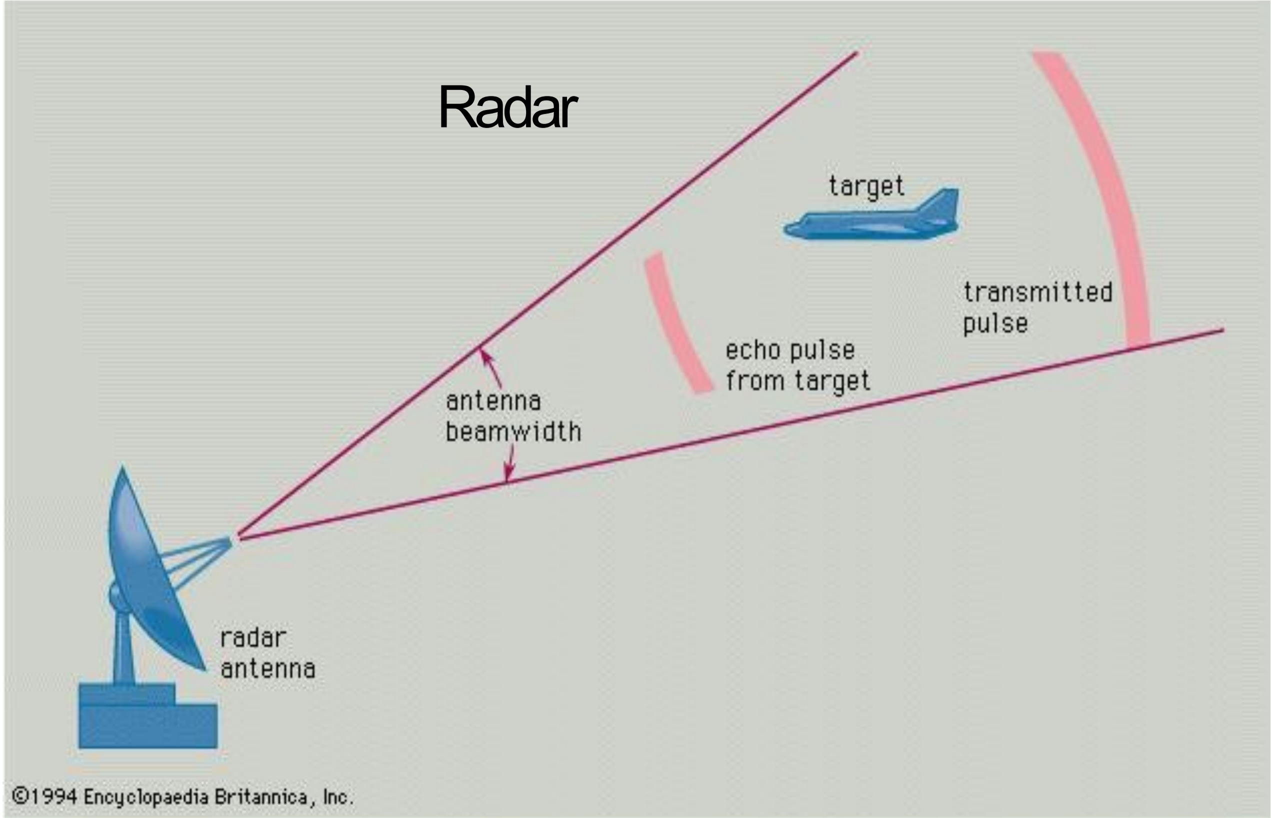
Infra-Red Imaging

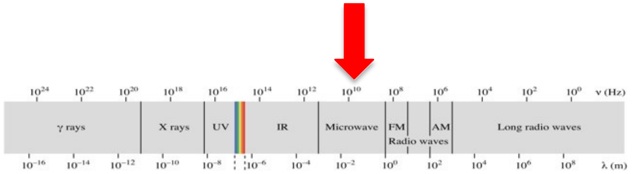
- Night vision system used by soldiers
- IR satellite view of Augustine volcano



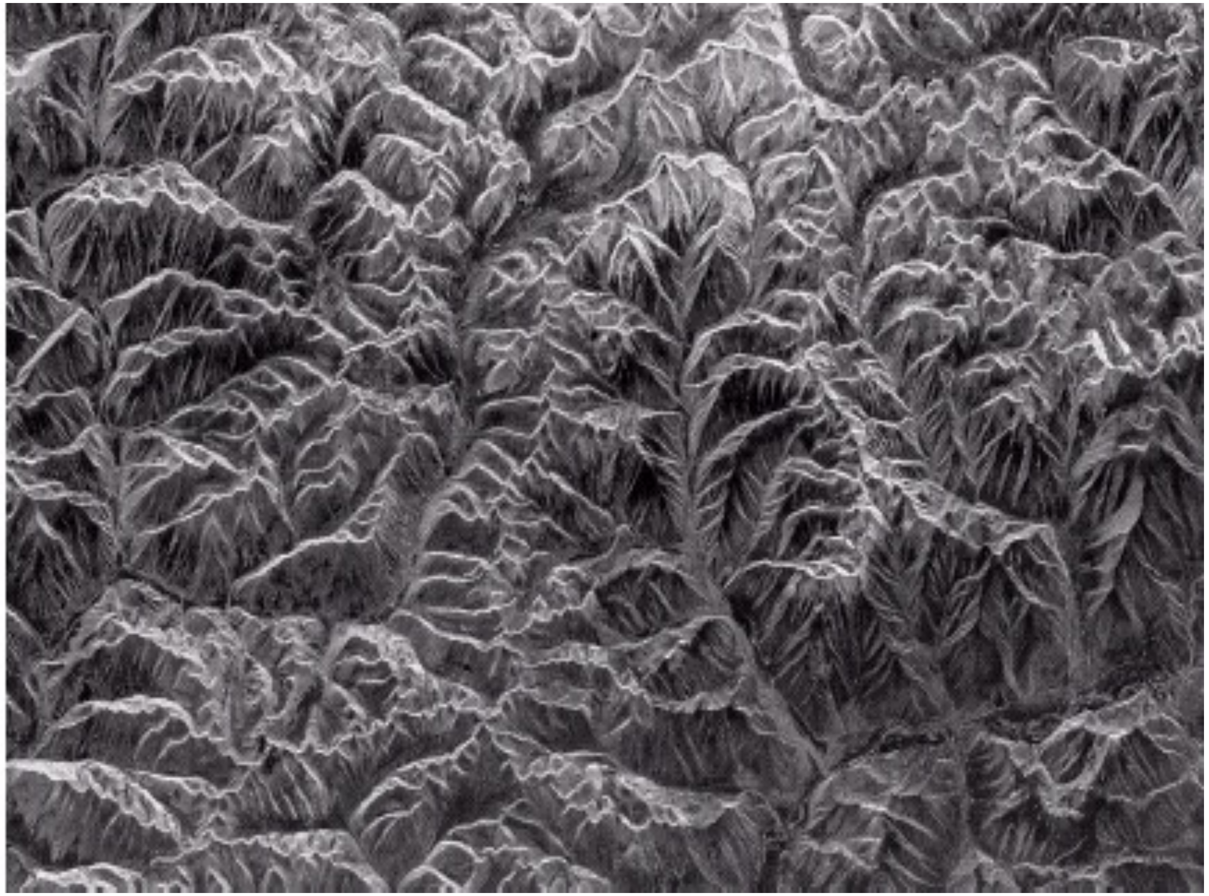


Radar Imaging



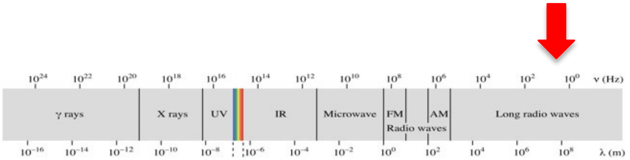


Spacebome radar image of mountains in southeast Tibet

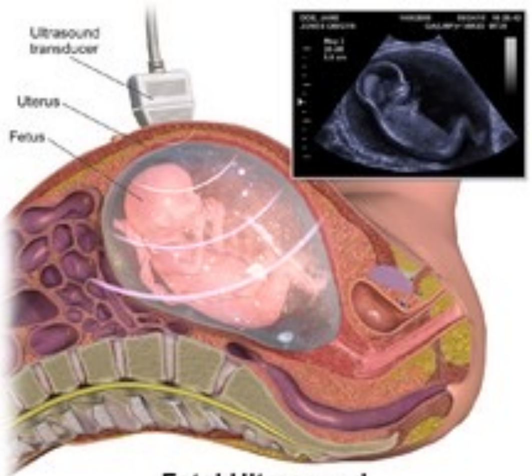


Magnetic Resonance Imaging

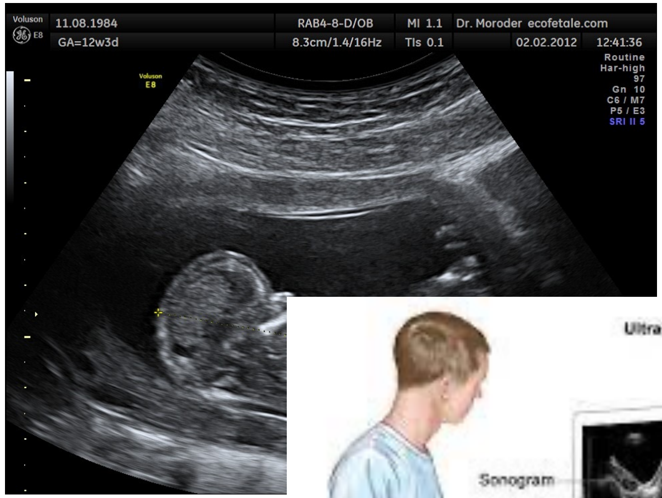
- MRI (Magnetic Resonance Imaging)

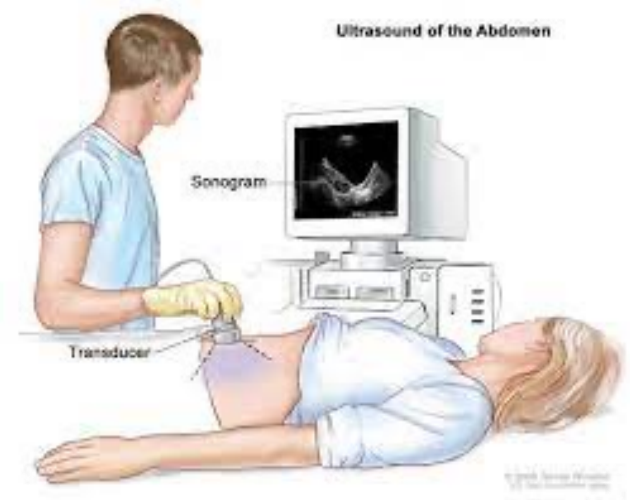


Ultrasound Image



Fetal Ultrasound

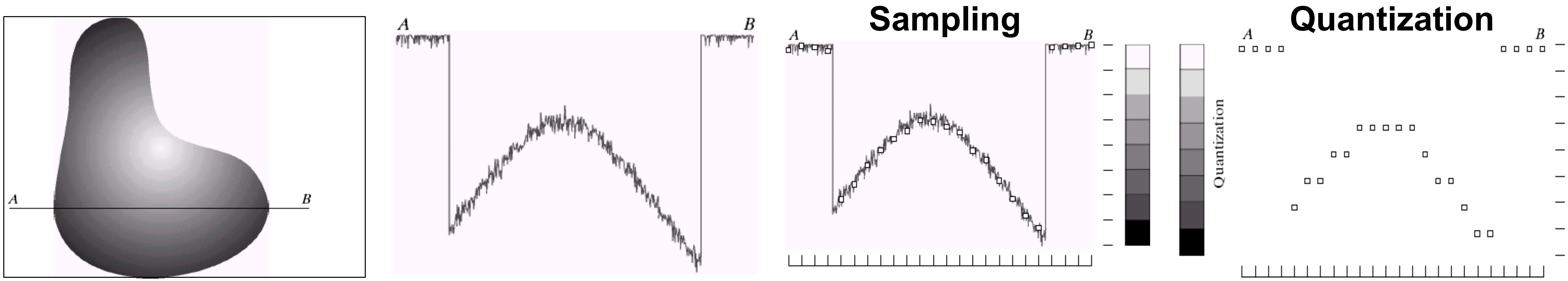




Computer Generated Images

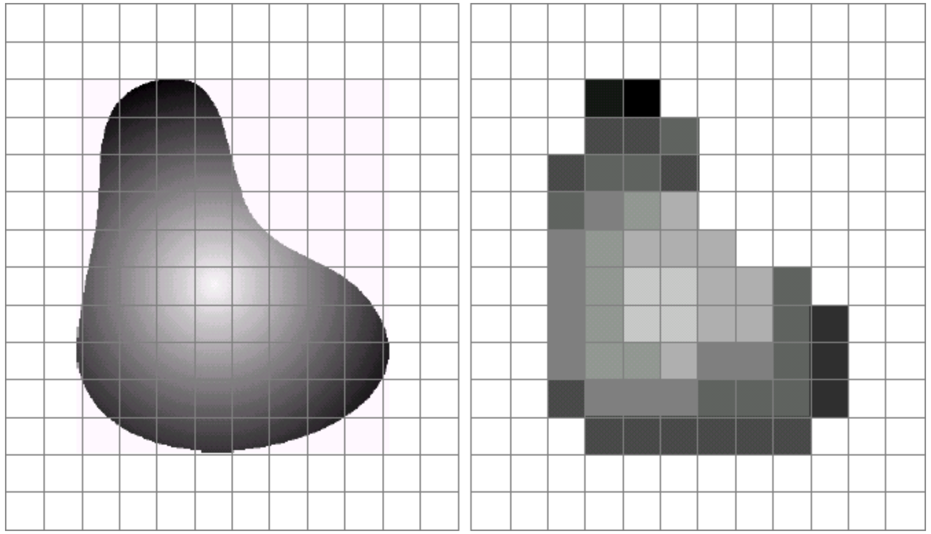
Image Sampling and Quantization

- A digital sensor can only measure a limited number of samples at a discrete set of energy levels. It cannot capture everything continuously.
- **Sampling is about space.** We divide the image into a grid of pixels. More samples means higher resolution.
- **Quantization is about intensity.** Quantization is the process of converting a continuous analogue signal into a digital representation of this signal



2-D Image Digitization Process

- Remember that a digital image is always only an approximation of a real world scene



What is a Pixel?

- Pixel = Picture Element
- The smallest discrete component of an image or picture on a screen is known as a pixel.
- The greater the number of pixels, the greater is the resolution.
- Each pixel is a sample of an original image, where more samples typically provide more-accurate representations of the original.

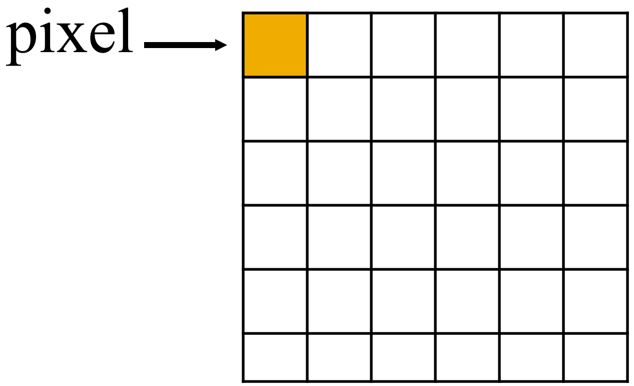


Image Resolutions: Spatial and Gray-level

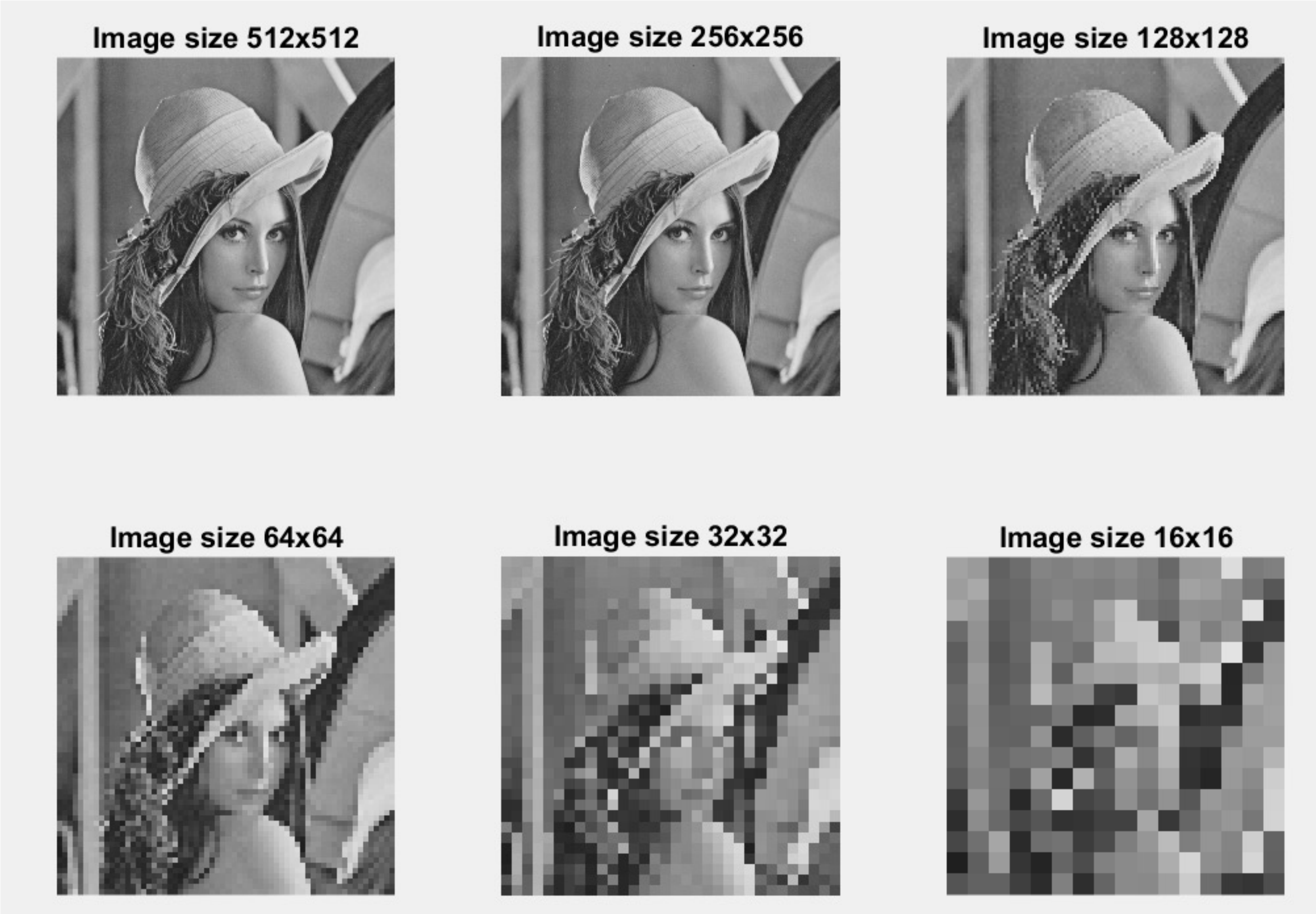
- Spatial Resolution:
  - Number of lines per millimeter, or Dot Per Inch (DPI).
  - Often say spatial resolution of MxN pixels.
  - The larger, the better.
  - Define how fine the sampling is.
- Gray-level Resolution:
  - Smallest discernible change in gray level.
  - Often say gray-level resolution of L.
  - 8 bits -> $2^8 – 1 = 255$

Spatial Resolution

- The spatial resolution of an image is determined by how sampling was carried out
- Image size M*N
- Spatial resolution simply refers to the smallest discernable detail in an image
  - Vision specialists will often talk about pixel size
  - Graphic designers will talk about dots per inch (DPI)

Spatial Resolution Examples

Spatial Resolution Examples



Intensity/Gray Level Resolution

- Intensity level resolution refers to the number of intensity levels used to represent the image
  - The more intensity levels used, the finer the level of detail discernable in an image
  - Intensity level resolution is usually given in terms of the number of bits used to store each intensity level

Number of Bits	Number of Intensity Levels	Examples
1	2	0, 1
2	4	00, 01, 10, 11
4	16	0000, 0101, 1111
8	256	00110011, 01010101
16	65,536	1010101010101010

Intensity Level Resolution Examples

256 grey levels (8 bits per pixel)

128 grey levels (7 bpp)

64 grey levels (6 bpp)

32 grey levels (5 bpp)

16 grey levels (4 bpp)

8 grey levels (3 bpp)

4 grey levels (2 bpp)

2 grey levels (1 bpp)

Resolution: How Much is Enough?

- The big question with resolution is always how much is enough?
- This all depends on what is in the image and what you would like to do with it
- Key questions include
  - Does the image look aesthetically pleasing?
  - Can you see what you need to see within the image?

Resolution Selection Example

- The picture on the right is fine for counting the number of cars, but not for reading the number plate

Intensity Level and Resolution Selection

Low Detail
(Face)

Medium Detail
(Cameraman)

High Detail
(Crowd)

Pixel and Resolution

- Image size vs resolution and bit(s)/pixel
  - Image size increases when either resolution or bits/pixel increases
  - Image size = *resolution x bits/pixel*

1-bit Lena image size

= 4000 x 3000 x 1 bit

≅ 12M bits = 1.5M Bytes

8-bits Lena image size

= 4000 x 3000 x 8 bits

≅ 96 M bits = 12M Bytes

Pixel and Resolution

- What about color images?
  - Encode color information for each pixel
  - For display, usually we use Red, Green and Blue to represent each pixel
  - If we use 8 bits for each color channel, the image size
    = 4000x3000 (image resolution)
    x 8 (bits) x 3 (color channels)
    ≅ 288Mbits = 36M Bytes
    => compression is required

Course Contents

- Introduction to computer vision and pattern recognition
  - Computer vision applications
  - Image acquisition
  - Color models/spaces

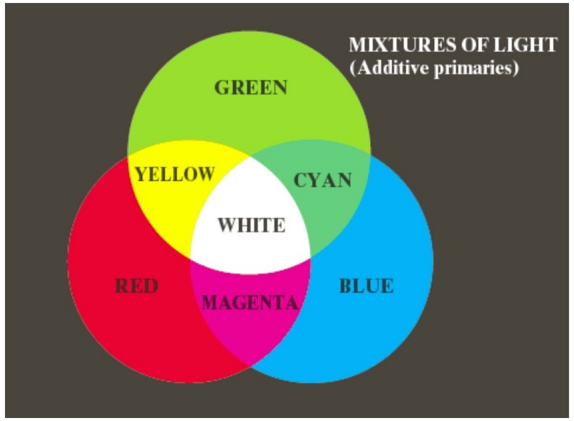
Color

Color Models/Spaces

- Color Model:
  - An abstract mathematical system to represent color.
  - A 3 dimensional abstractions.
  - Defines three primary colors along three dimensions
  - Is typically limited in the range of colors they can represent and hence often can’t represent all colors in the visible spectrum

RGB Color Model

- **Assumes that light is used to generate colors**.
- Human perception is additive since black is the absence of light and white the presence of all wavelengths of light.
- Devices: computer monitor, LCD screens, projectors
- **RGB Color Model**
  - Red : 700 nm
  - Green : 546 nm
  - Blue : 436 nm



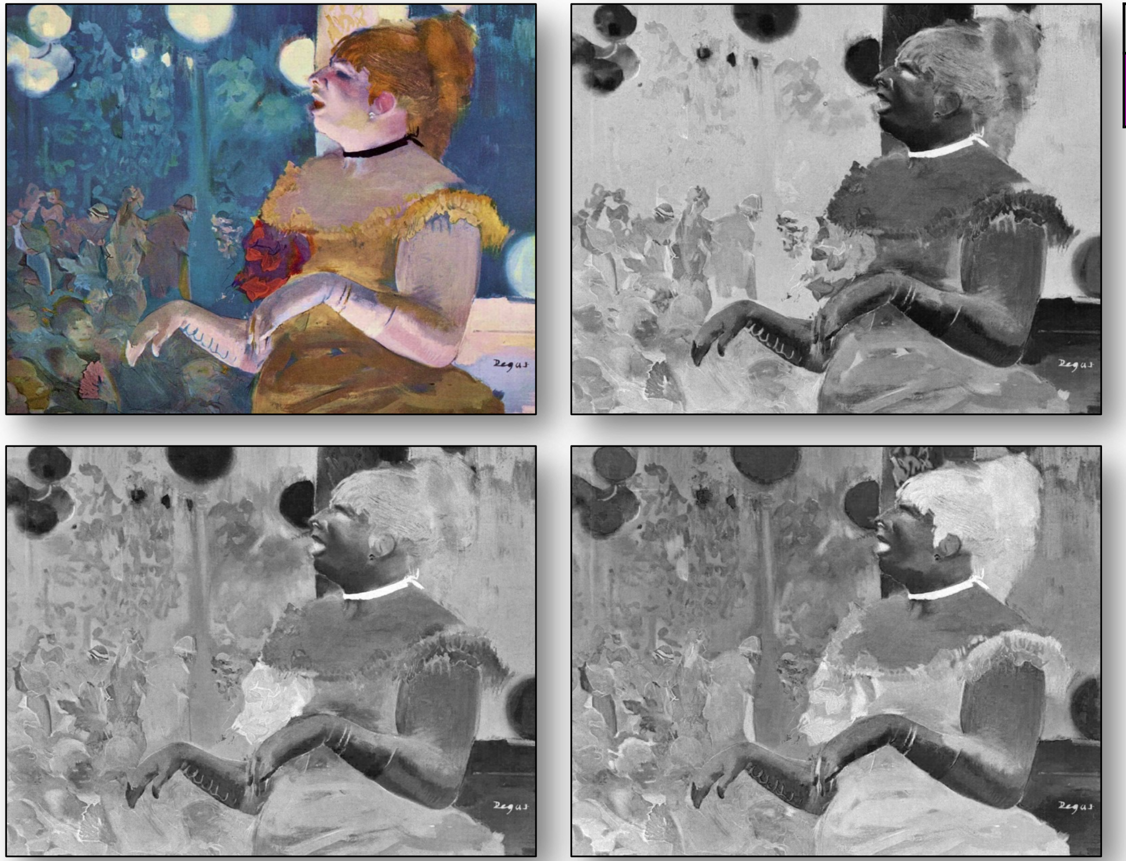
Different Color Models

- Different color models: RGB,CMY(CMYK), HSV, YCbCr
- RGB color model: eye, monitor, video
- CMY (CMYK) color model: printing
- HSV/HSB: close to Human Visual Perception
- YCbCr are commonly used for image/video compressions

RGB Channels

How different parts of the image are brighter in different channels depending on the color.

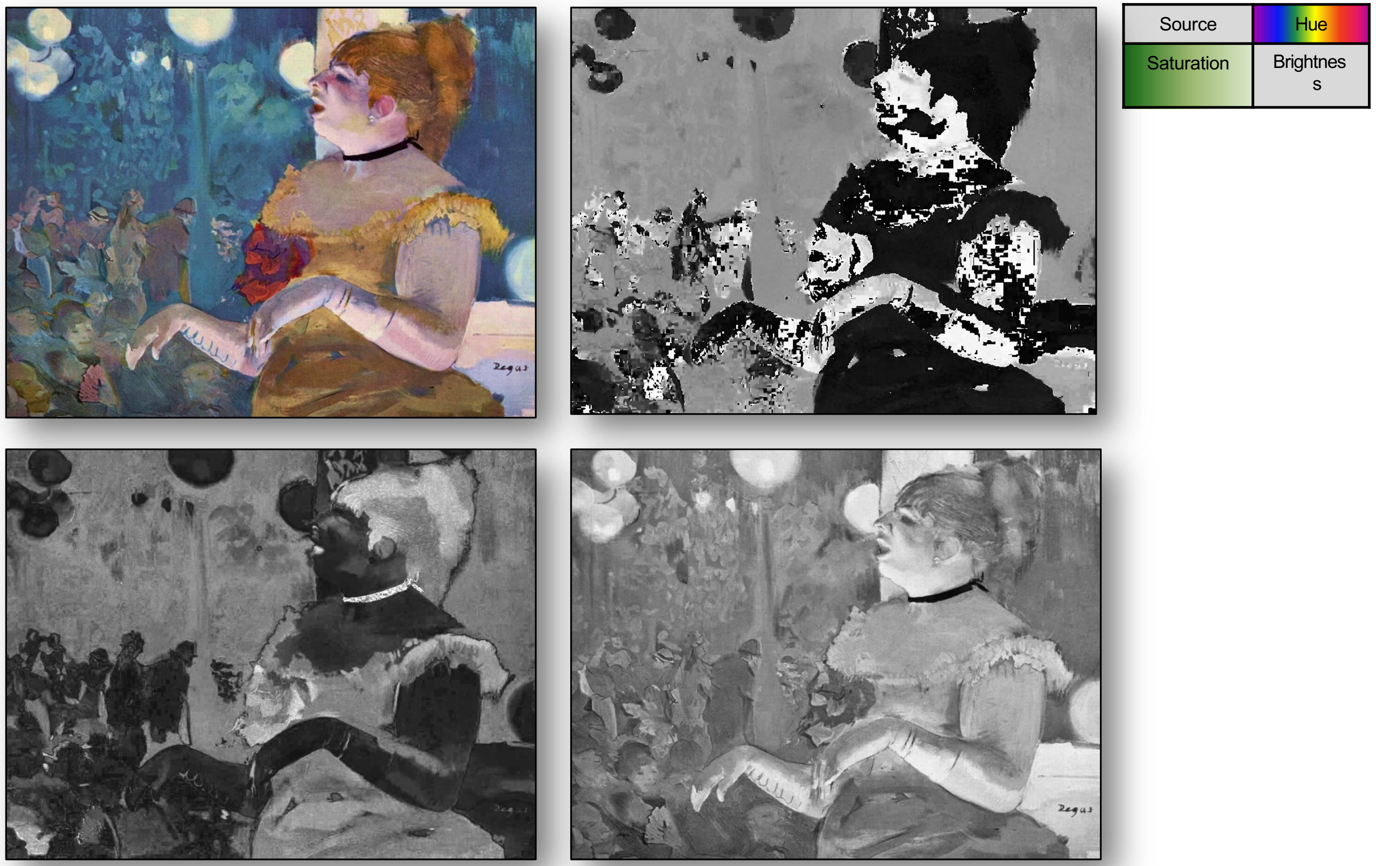
CMY Channels



Source	Cyan
Magenta	Yellow

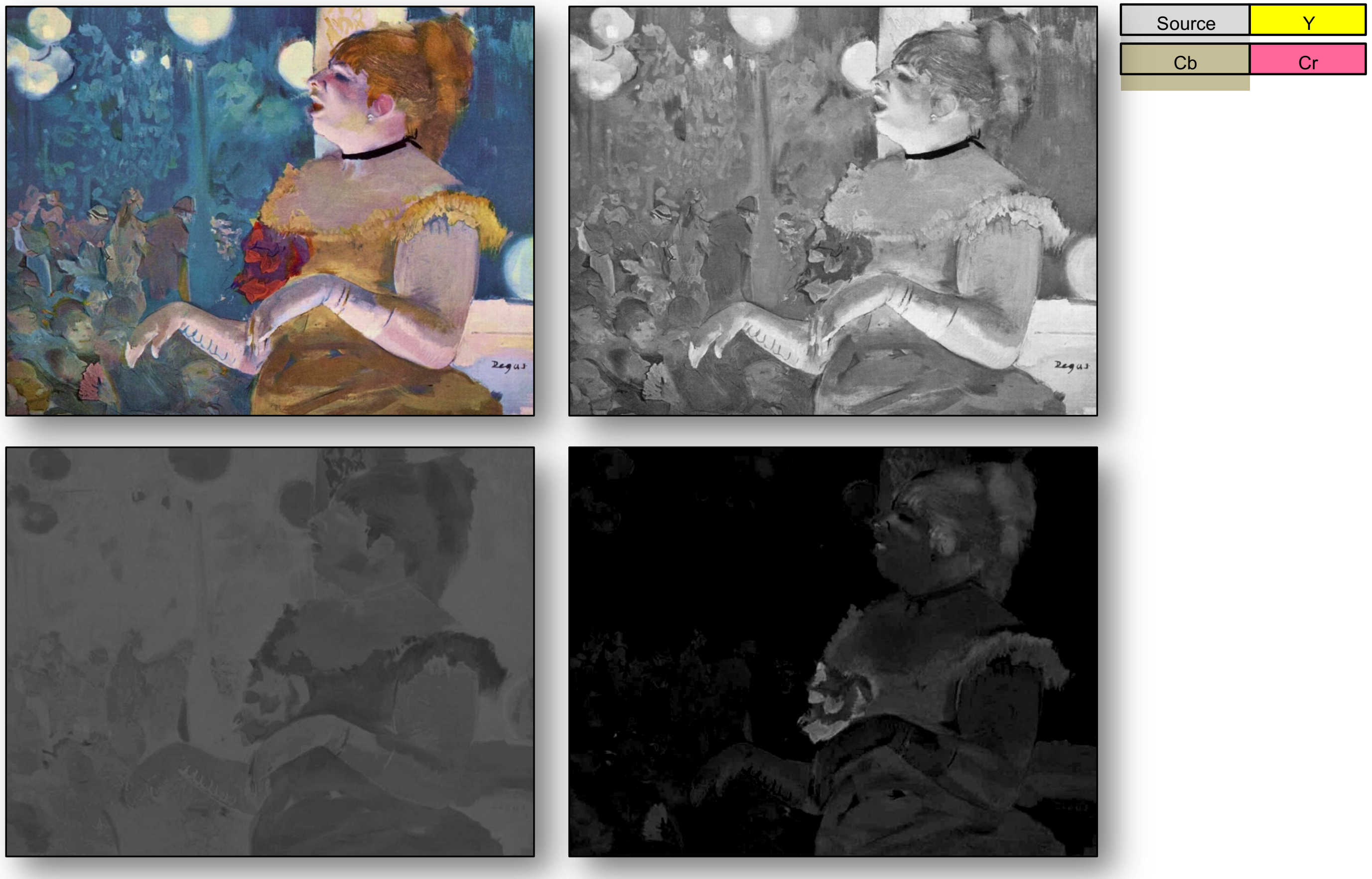
How different parts of the image are brighter in different channels depending on the color. But for printing colors.

HSV Channels



HSV channels - hue shows the color type, saturation shows the intensity of color, value shows brightness. Very different from RGB.

YCbCr Channels



YCbCr - Y is the luminance, the brightness. Cb and Cr are the color components. Notice the Y channel looks like a grayscale image - that's the brightness information separated out.

~ End ~