Digital Image Processing (EC 61501)

Instructors:

Prof. P.K. Biswas and Dr. Saumik Bhattacharya

Lectures:

Monday: 10:00 – 11:00
Wednesday: 8:00-10:00
Thursday: 10:00-11:00

E-mail: pkb@ece.iitkgp.ac.in, saumik@ece.iitkgp.ac.in

Office: A-206, E&ECE Department

Logistics

- Duration
  - 8 Weeks
- Assessment
  - Continuous
- Main textbook
  - "Fundamentals of digital image-processing", Anil. K. Jain, Prentice-Hall
  - "Digital Image Processing", Rafael C. Gonzalez & Richard E. Woods, Addison-Wesley

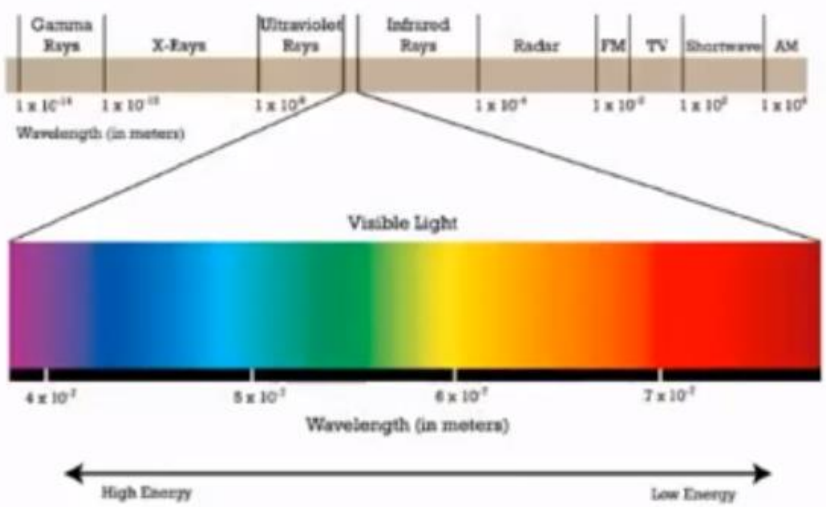
  - NPTEL lectures
  - Other sources: Lectures of CMU, RICE, Utah, UCSD etc.

Objectives

- To know about generation and processing of digital images
- Acts as a foundation of computer vision and pattern recognition
- Theory+ MATLAB demonstrations
- Coding based assignments/ projects

What is an image?

- We are mostly familiar with images that we can see with our eyes. These are the images we take with our cameras. They form a very small part of the spectrum!



The electromagnetic spectrum

What is an image?

- Images can be captured from other parts of the spectrum. For example:
  - Gamma images
  - X ray images
  - Ultra violet images
  - Optical-microscopy (light-microscopy) images
  - Infrared images
  - Satellite images
  - Thermal images

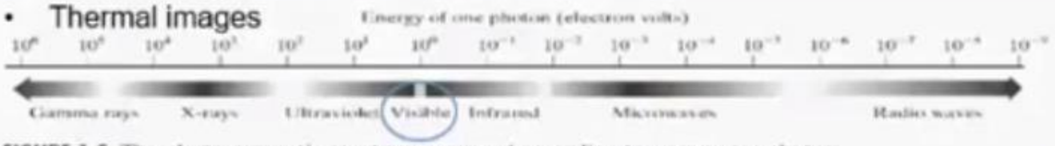


FIGURE 1.5 The electromagnetic spectrum arranged according to energy per photon.

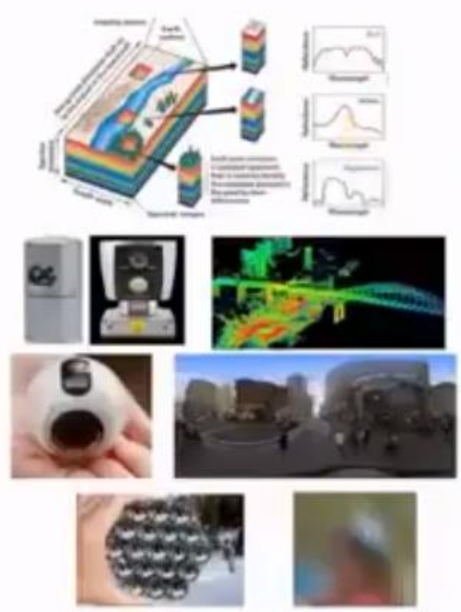
Other Acquisition Platforms

Hyperspectral Image Sensor
- $I(x,y)$ in R^D, $D \approx 48$

3D/Depth Sensor: LiDAR, Kinect
- $I(x,y)$ in R

Panoramic Cameras
- $I(\alpha,\beta)$, α,β in $[0,2\pi]$, $[0,\pi]$

Light-field cameras
- Lenslet system



Topics in DIP

- Image Processing topics include:
  - Image modelling
  - Image restoration, enhancement, reconstruction
  - Image compression
  - Analysis, detection, recognition, understanding

Image Processing vs Computer Vision

- Image Processing refers to a lower level (modelling/signal analysis/noise removal) of processing of an image signal compared to Computer Vision.
- For example:
  - Finding good mathematical features to describe a human silhouette is image processing.
  - Using the features which describe a human silhouette as part of a system that detects humans in an image is computer vision.
- Computer Vision refers to any type of science that attempts to make a digital computer carry human vision tasks. It is a bigger and less well-defined area compared to Image Processing. The two areas overlap.

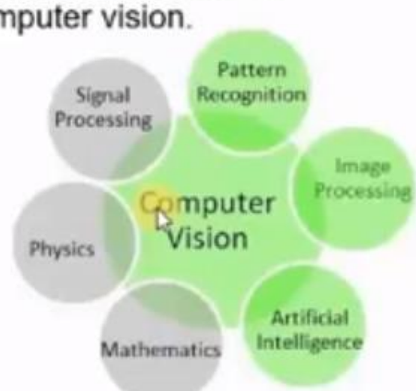
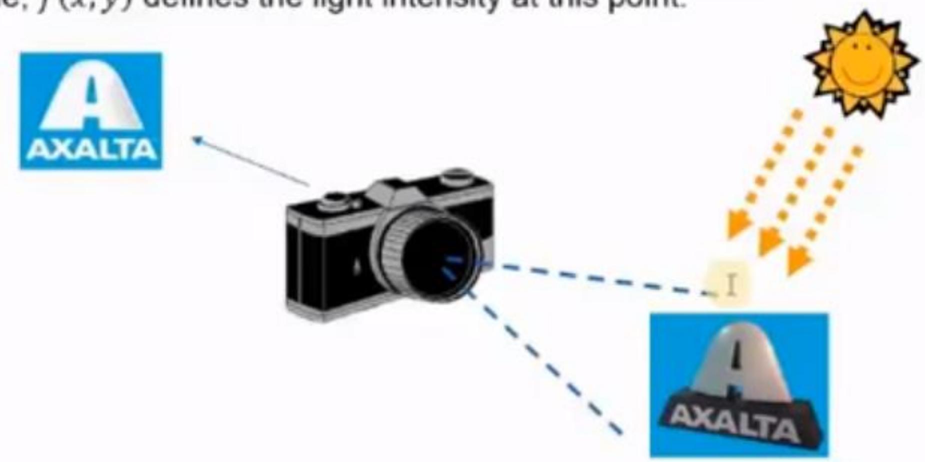


Image acquisition

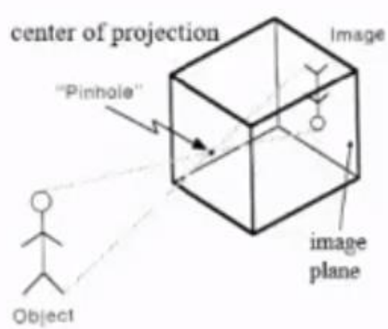
- An image is a projection of a 3D scene into a 2D projection plane.
- An image can be defined as a function of two variables (x, y) as $f(x, y)\colon R^2 \to R$, where for each position (x, y) in the projection plane, $f(x, y)$ defines the light intensity at this point.



Camera Geometry

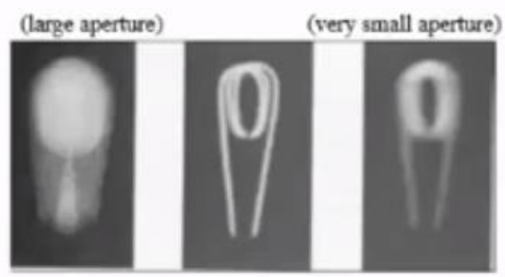
- The simplest device to form an image of a 3D scene on a 2D surface is the "pinhole" camera.
- Rays of light pass through a "pinhole" and form an inverted image of the object on the image plane.

perspective projection



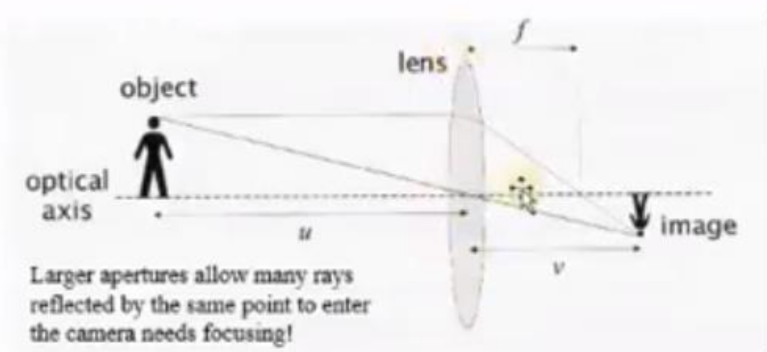
Diffraction and pinhole optics

- If we narrow the pinhole, only a small amount of light is let in.
  - When light passes through a small aperture, it does not travel in a straight line.
  - It is scattered in many direction (diffraction - a quantum effect).
- If we use a wide pinhole, light from the source spreads across the image (i.e., not properly focused), making it blurry.



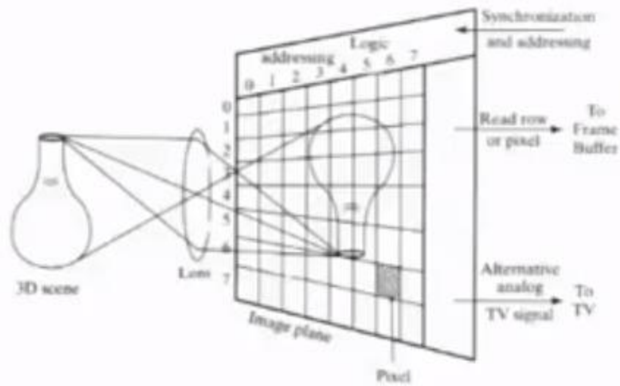
Camera Optics

- In practice, lens is used to duplicate the pinhole geometry without resorting to undesirably small apertures.
- Lens are placed in the aperture to focus the bundle of rays from each scene point onto the corresponding point in the image plane – this leads to sharp images!



CCD cameras

- An array of tiny solid state cells convert light energy into electrical charge.
  - Manufactured on chips typically measuring about 1cm x 1cm



Modern Camera

- The output of a CCD array is a continuous electric signal (*video signal*) which is generated by scanning the photo-sensors in a given order (e.g., line by line) and reading out their voltages.

- The video signal is sent to an electronic device called **frame grabber**.

- The frame grabber digitizes the signal into a 2D, rectangular array *N* x *M* of integer values, stored in the **frame buffer**.

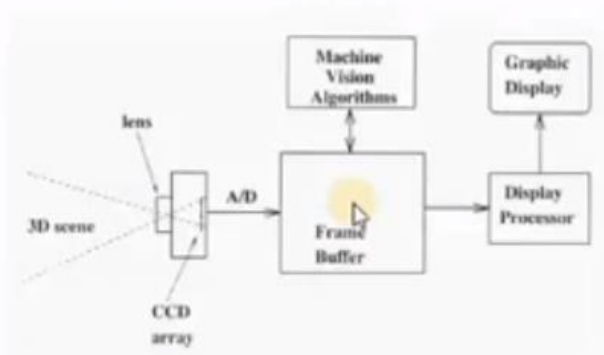
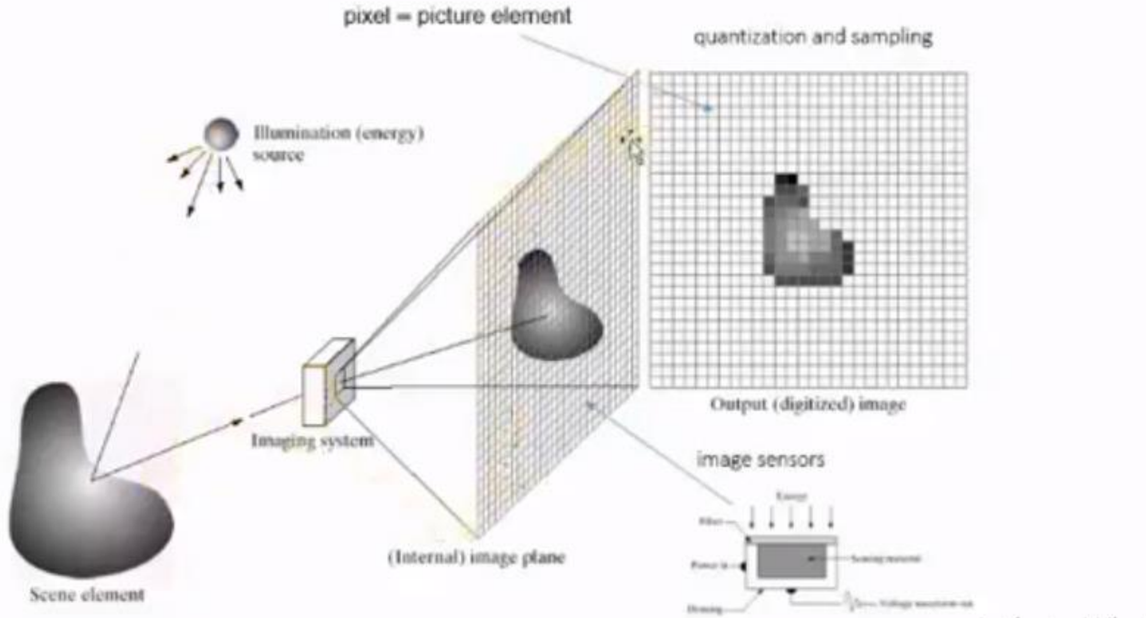
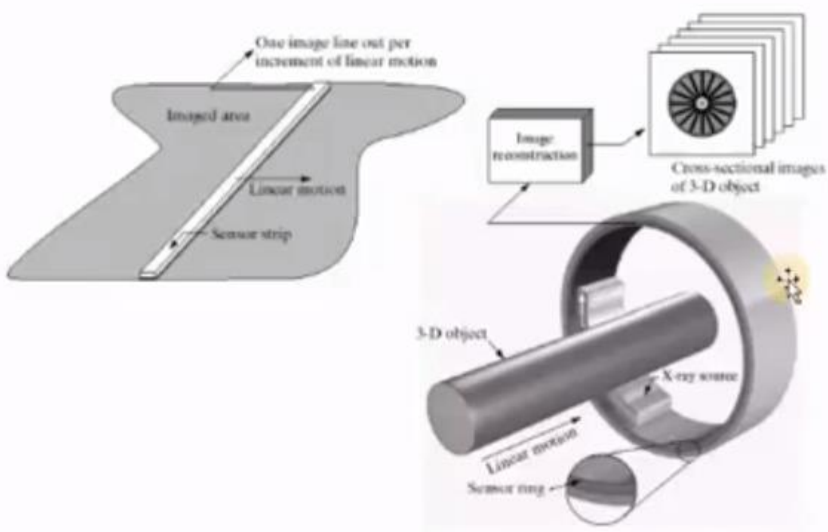


Image acquisition



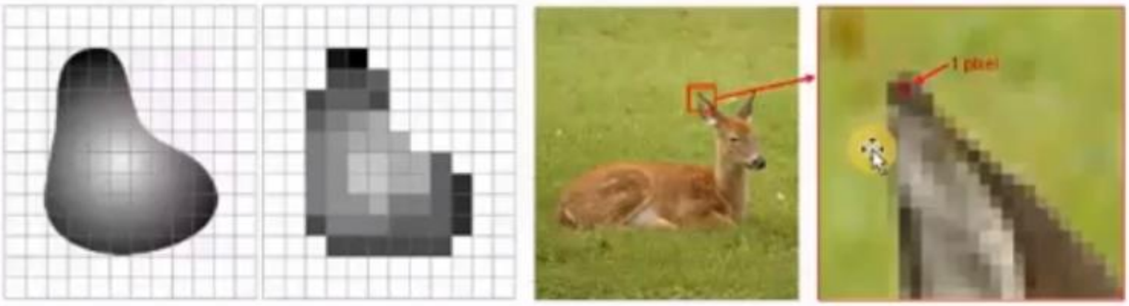
Sensor Strips



a b
FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

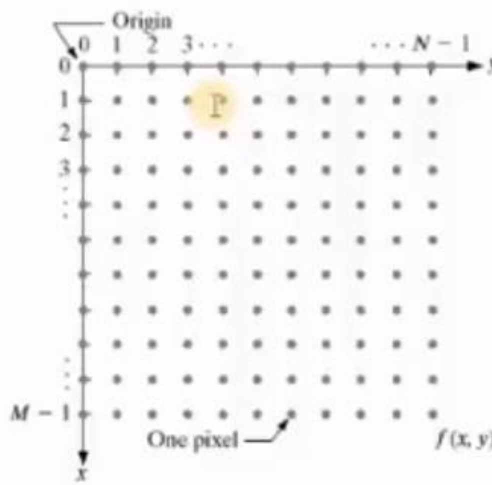
Pixels

- Pixel values typically represent gray levels, colors, distance from camera, etc.
- **Remember** *digitization* implies that a digital image is an *approximation* of a real scene



Image

- An image is a function defined on a 2D coordinate $f(x,y)$.
- The value of $f(x,y)$ is the intensity.
- 3 such functions can be defined for a color image, each represents one color component
- A digital image can be represented as a matrix.



Storage

- Common image formats include:
  - 1 sample per point (grayscale)
  - 3 samples per point (Red, Green, and Blue)
  - Video (above information **plus** time)

- For most of this course we will focus on grey-scale images

What is DIP?

- The continuum from image processing to computer vision can be broken up into low-, mid- and high-level processes

Low Level Process	Mid Level Process	High Level Process
Input: Image **Output:** Image	**Input:** Image **Output:** Attributes	**Input:** Attributes **Output:** Understanding
Examples: Noise removal, image sharpening	**Examples:** Object recognition, segmentation	**Examples:** Scene understanding, autonomous navigation

In this course we will stop here

Image as a function

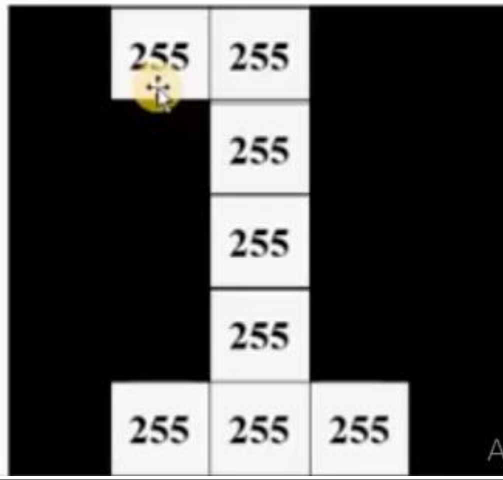
- The rectangular grid presented in previous slides implies that digital images are two-dimensional (2D) signals $f(x, y)$.
- In video the concept of time is present as well, since we acquire a sequence of frames and not a single image frame. Therefore, a video signal could be described as a three-dimensional (3D) signal $f(x, y, t)$.
- Four-dimensional (4D) image signals $f(x, y, z, t)$ also exist. It is a term used to describe the study of three-dimensional (3D) specimens as they change over time. Examples are CT and MRI scans.

Sampling of Captured Image

- Sampling of an image is basically sampling of a 2D signal.
- The continuous image coordinates (x, y) are replaced with a set of discrete values.
- That means we only observe the image signal at certain locations.
- In the example below two identical images are sampled at different rates.
- Obviously the higher the sampling rate the better the quality of the image.
- After a specific sampling rate the human eye is not able to perceive an improved image.
- For the image below sampling which yields a digital image of size 256×256 is efficient so that the human eye perceives the image as an analogue one with good quality.

Quantization

- Quantisation of an image is basically discretization of the image signal $f(x,y)$ (discretization of the image amplitude).
- After sampling and quantization both pixel coordinates (x,y) and image values $f(x,y)$ are represented with binary numbers.
- Below you see an image quantised in two levels (binary).
- For images of the so called gray level type where $f(x,y)$ is a scalar and represents all shades of the gray colour, ranging from the absolute black (0) to the absolute white (255), we normally use 256 gray levels for $f(x,y)$.
- The value of $f(x,y)$ is called the intensity of the image.



Gray level/ Bit-depth

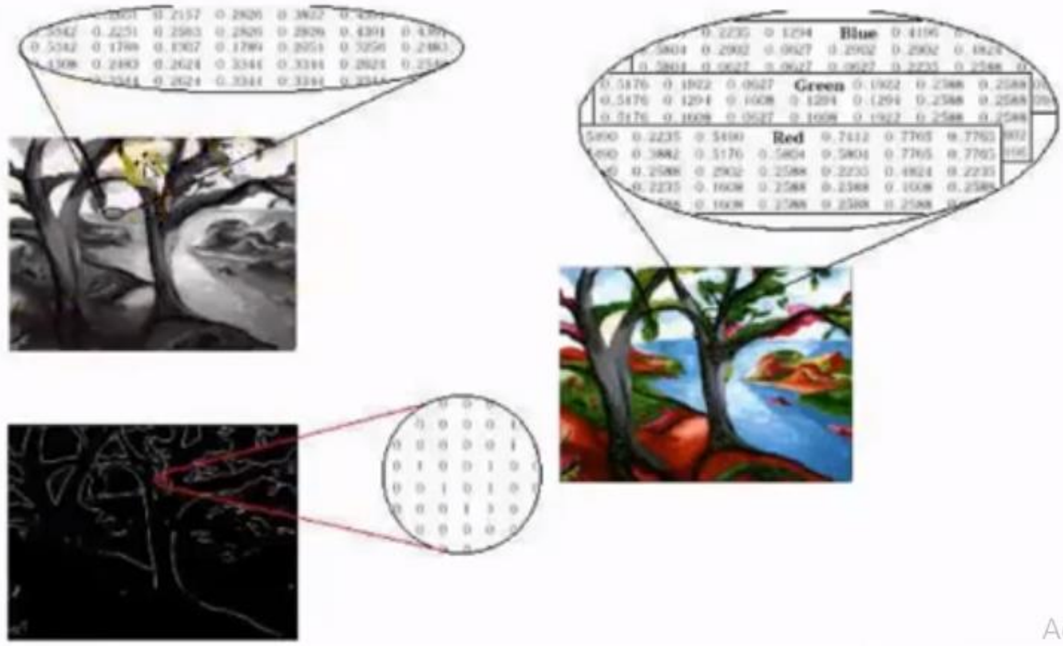
- Obviously the more quantisation levels we assign to digital images the better their quality.
- The term quality gives rooms for discussion. When can we say that an image is of good quality?
- Image processing scientists use various mathematical metrics to asses the quality of in image. We will come across with some of them later.

256×256 256 levels

256×256 32 levels

256×256 2 levels

Gray level images/Colour images/Binary images



Key Stages in DIP

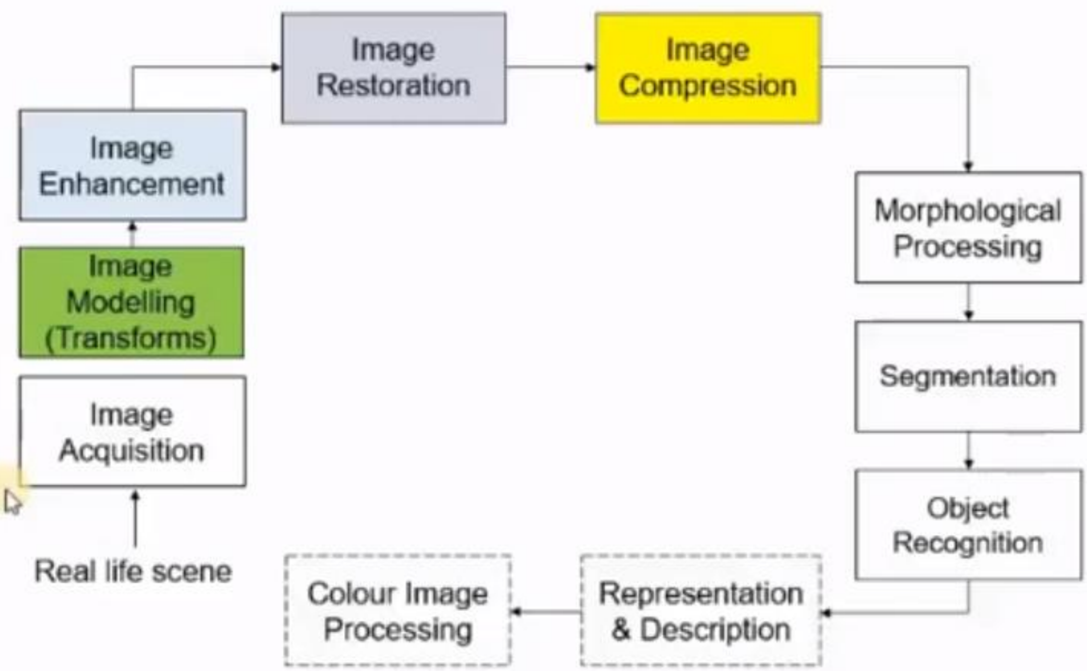


Image Acquisition

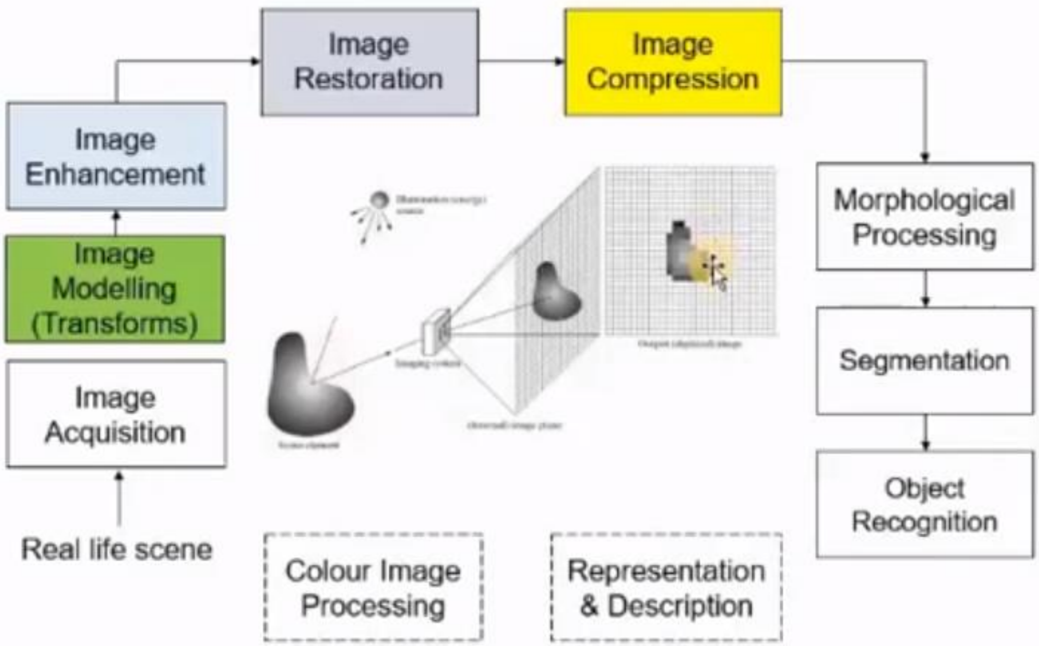


Image Modelling

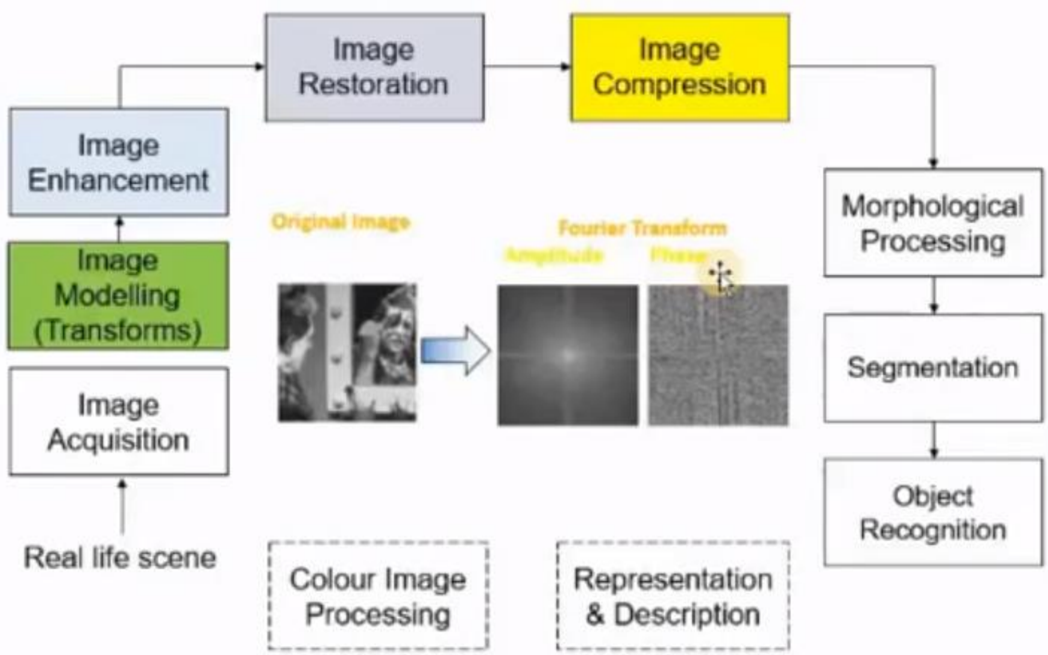


Image Enhancement

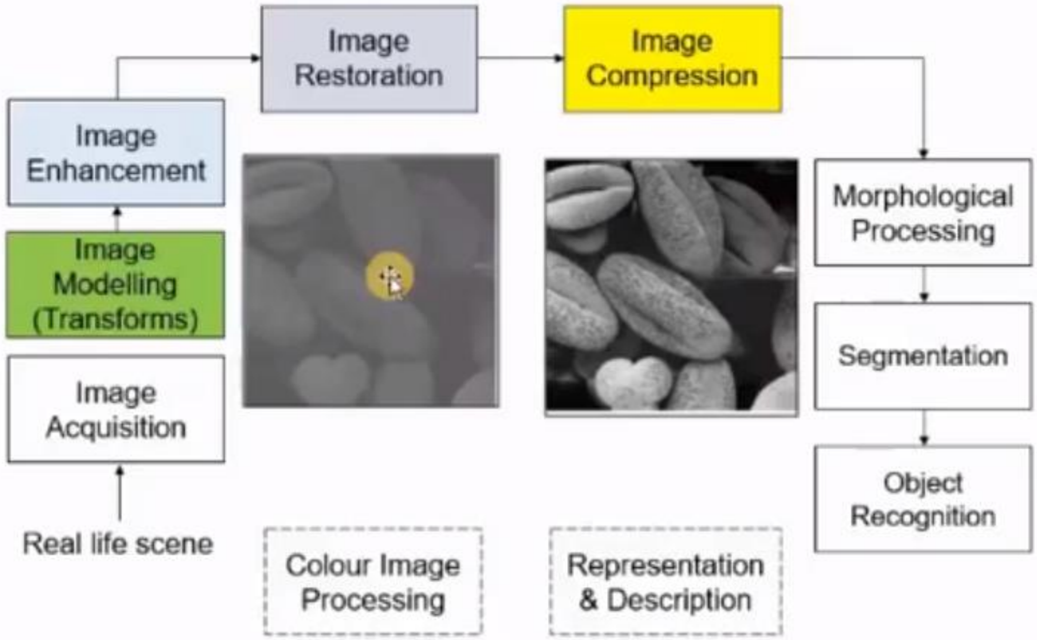


Image Restoration

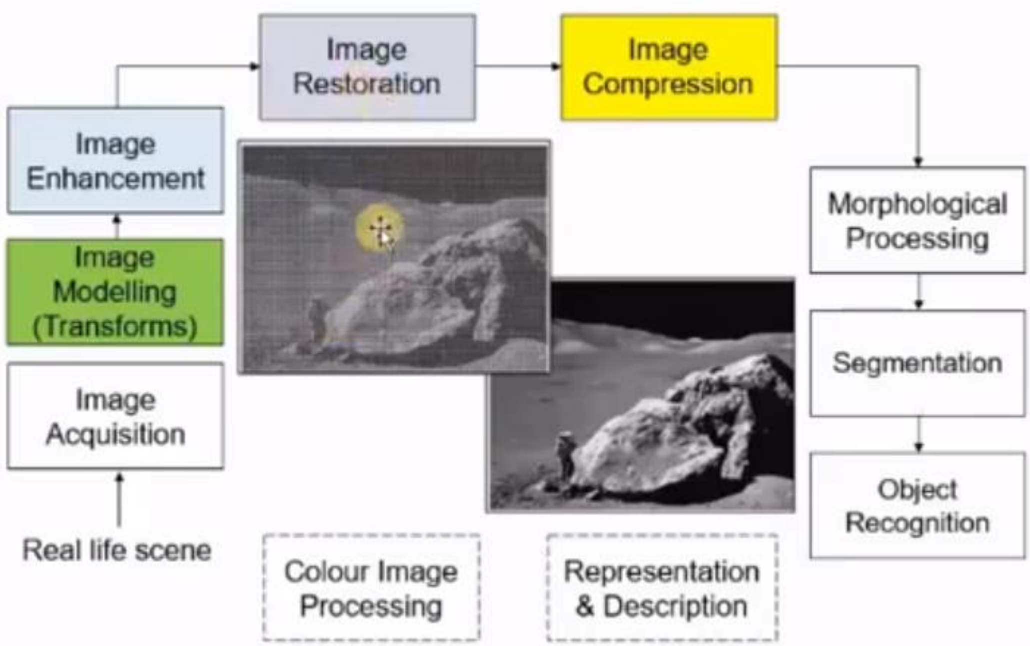
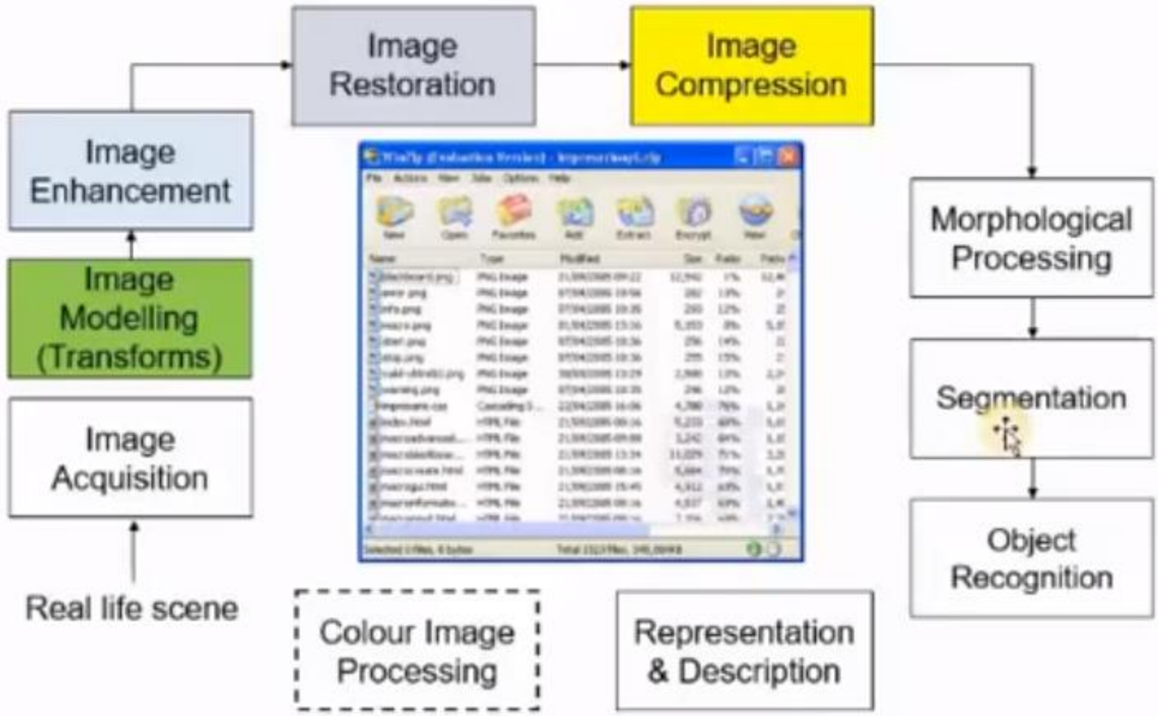
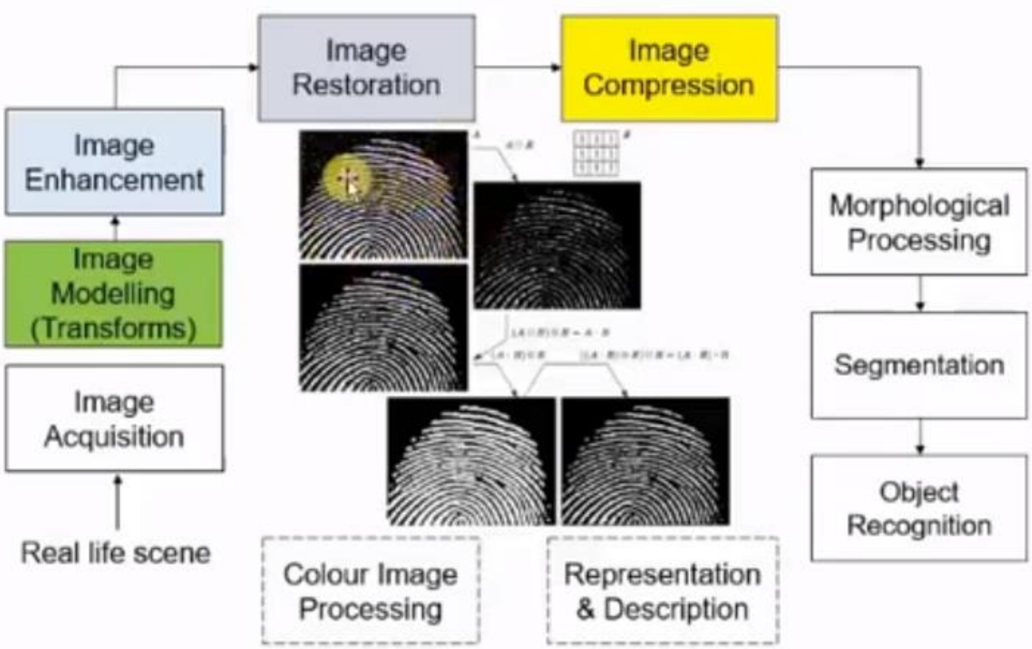


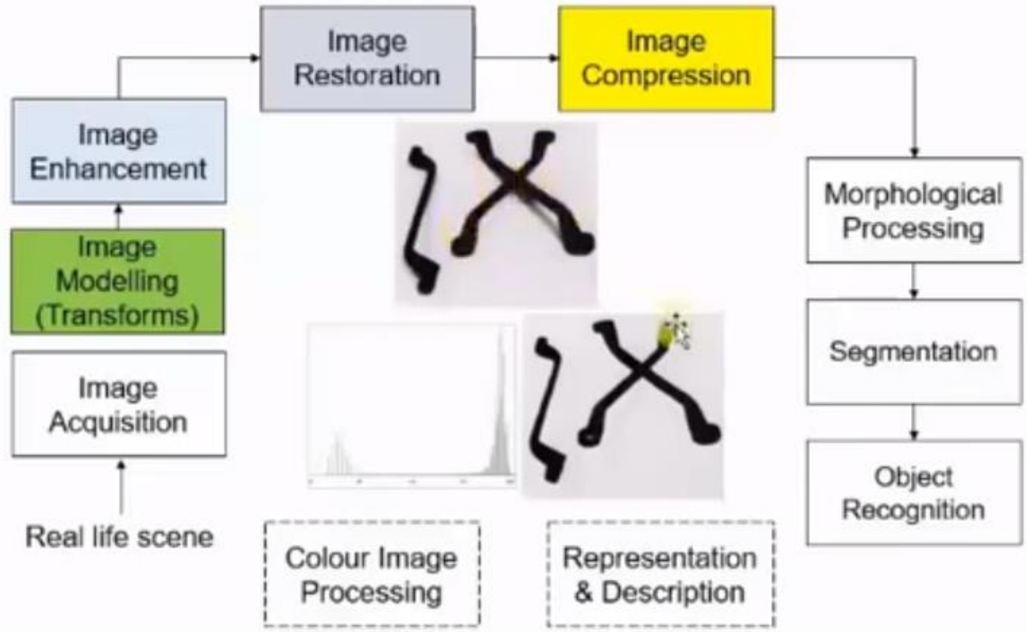
Image Compression



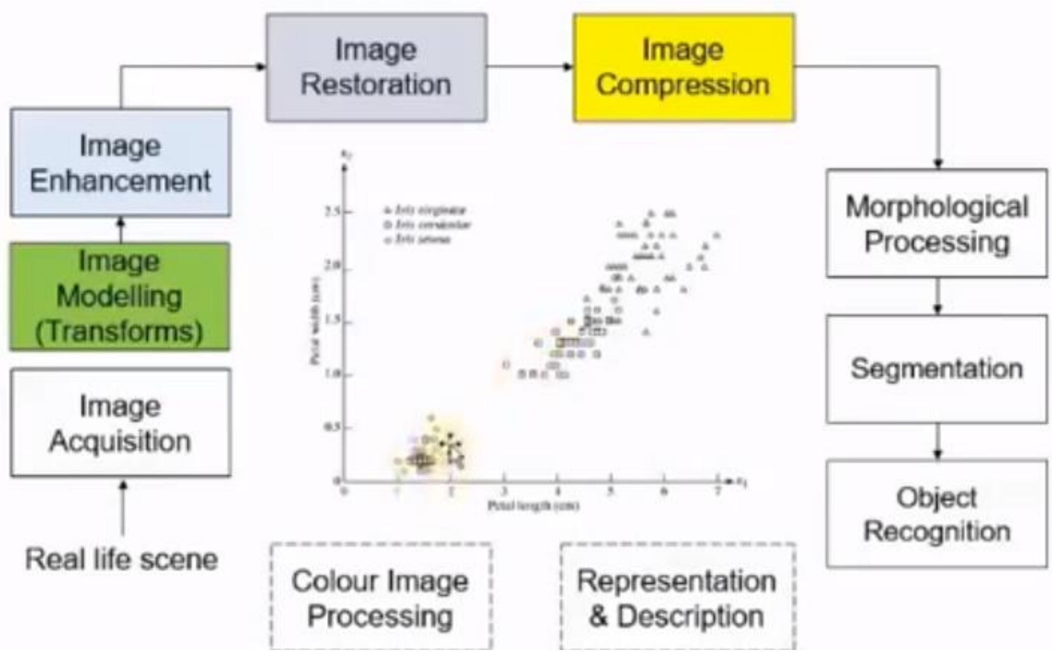
Morphological Processing



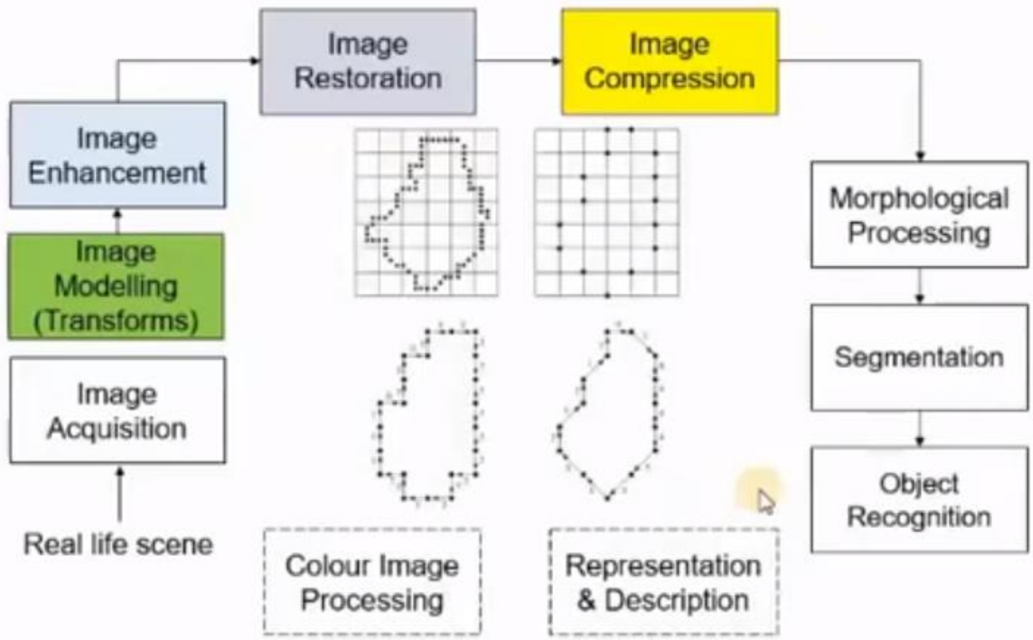
Segmentation



Object Recognition



Representation and Description



Colour Image Processing