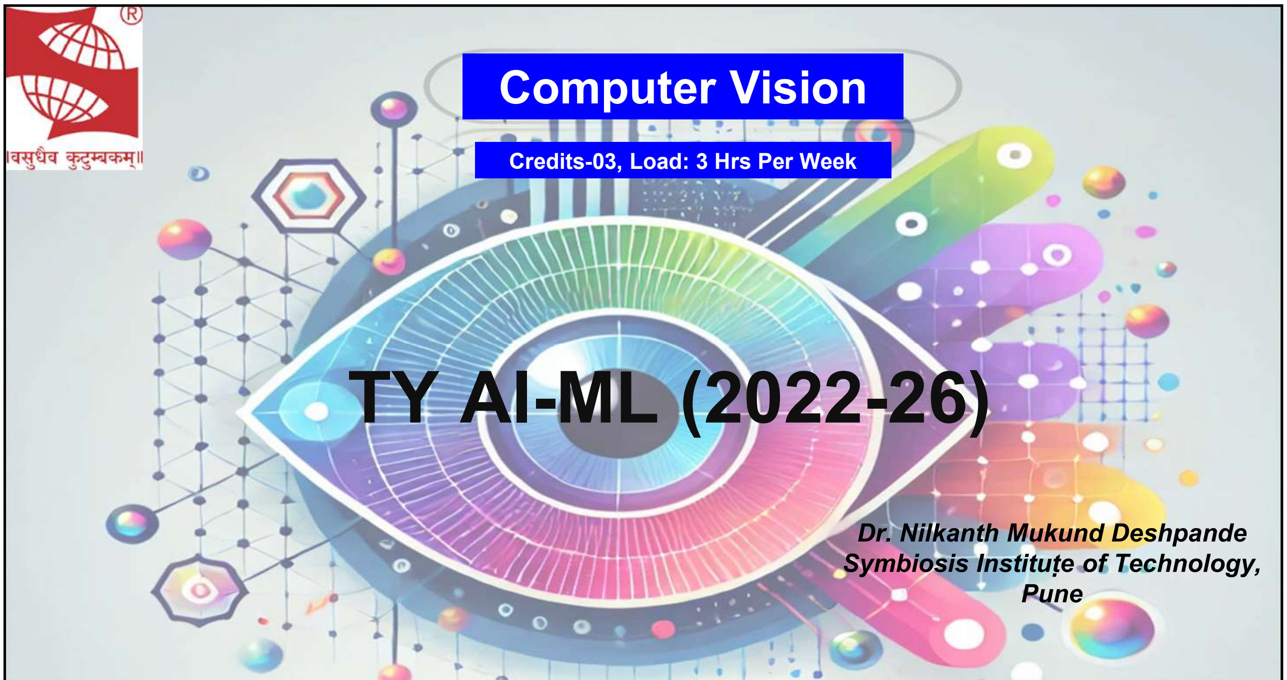


Computer Vision

Credits-03, Load: 3 Hrs Per Week

TY AI-ML (2022-26)

Dr. Nilkanth Mukund Deshpande
Symbiosis Institute of Technology, Pune

Syllabus

Course Name : Computer Vision
Course Code : TE7484
Faculty : Engineering
Course Credit : 3
Course Level : 3
Sub-Committee (Specialization) : Computer Science

Learning Objectives

Understand basics of computer vision as well as its mission of making computers see and interpret the world as humans do, by learning core concepts of the field and receiving an introduction to human vision capabilities.

Learn both image and video recognition, including image classification and annotation, object recognition and image search, various object detection techniques, motion estimation, and artificial intelligence.

Study various computer vision applications

Apply deep learning techniques to implement various computer vision concepts on different platforms

Learn various object detection techniques

Syllabus

Sr. No.	Topic	Actual Teaching Hours
1	Basics of Digital ImagingImage Acquisition, Sampling, Quantization, Difference in Monochrome and Multichrome imaging, concept of color spaces, point processing techniques, mask processing methods, image filtering, shape in images, edge detection, gradient	9
2	Image Representation and Region AnalysisShape Descriptors-contour based, region based, Boundary based; Thresholding based segmentation, Watershed based Segmentation, Gray level Co-occurrence Matrix-energy, entropy, maximum probability, contrast, correlati	9
3	Computer Vision ApplicationsImage Fusion and Clustering-K-means, Vector Quantization, Hierarchical Clustering, Partitioned Clustering, Image Inpainting, Multisensor image fusion, character recognition, face recognition, Trademark databases, Medical Imagi	9

Syllabus

4	Deep Learning in Computer Vision Inception architectures, ResNet, GoogLeNet DenseNet, atrous DenseNet and beyond, Fine-grained image recognition, Detection and classification of facial attributes, Content-based image retrieval, Computing semantic image e	9
5	Object DetectionANN, Region-based convolutional neural network, From R-CNN to Fast R-CNN, Faster R-CNN, Region-based fully-convolutional network, Single shot detectors, Speed vs. accuracy tradeoff, Fun with pedestrian detectors, Object detection problem,	9
	Total	45

Books Recommended

Book	Author	Publisher
Computer and Robot Vision	Robert Haralick and Linda Shapiro	Addison Wesley, Vol I & II.
Computer Vision	David A. Forsyth, Jean Ponce	A Modern Approach, PHI
Digital Image processing	Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins	Pearson Education , Fourth Impression, 2008, ISBN: 978-81-7758-898-9.

Course Outcomes

CO1: Apply fundamental digital imaging concepts, including image acquisition, sampling, quantization, color spaces, and basic processing techniques to enhance and analyze images. *(Bloom's Level 3: Apply)*

CO2: Implement segmentation techniques and region analysis using shape descriptors, thresholding, and co-occurrence matrix features to extract meaningful information from images. *(Bloom's Level 3: Apply)*

CO3: Develop and execute clustering, image fusion, and recognition methods for real-world applications, including character recognition, face recognition, and medical imaging. *(Bloom's Level 4: Analyze)*

CO4: Design and evaluate deep learning models such as Inception, ResNet, and DenseNet for advanced tasks like fine-grained image recognition, content-based retrieval, and semantic analysis. *(Bloom's Level 5: Evaluate)*

CO5: Create and optimize object detection systems using advanced algorithms like R-CNN variants and single-shot detectors, addressing the trade-offs between speed and accuracy. *(Bloom's Level 5: Create)*

Evaluation Plan

Sr. No.	Component	CO	Max marks	Weight	Tentative date
1.	Case study based Technical Poster	CO1	07	24%	10 Feb.- 2025
2.	Unit Test	CO2, CO3	15	50%	17-22 March-2025
3.	Kaggle Competition/ Hakathon on Computer vision	CO4, CO5	08	26%	21 April-2025

The Origins of Digital Image Processing

- *An image- defined as a two-dimensional function, where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point.*
- *When x, y, and the intensity values of f are all finite, discrete quantities, we call the image a digital image.*
- *The field of digital image processing refers to processing digital images by means of a digital computer.*
- **Vision** is the most advanced of our senses, so images play the most important role in human perception.
- Unlike humans, who are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines cover almost the entire EM spectrum, ranging from gamma to radio waves.
- They can operate on images generated by sources that humans are not accustomed to associating with images.
- These include ultrasound, electron microscopy, and computer-generated images.
- Thus, digital image processing encompasses a wide and varied field of applications

The Origins of Digital Image Processing

Examples of different images

- **Digital photographs**
A familiar example of a digital image, digital cameras use a sensor to capture light and convert it into digital information.

- **Satellite images**
Acquired using electronic sensors, satellite images are a type of digital image.

- **Sonographs**
These images are created by measuring how soundwaves bounce off surfaces and converting them into bits for a computer to interpret.

- **Scanned documents**
Digital images can be created by scanning documents such as manuscripts, printed texts, and artwork.

- **Gigapixel images**
These are an example of a digital image mosaic, which is a combination of non-overlapping images.

- **Virtual-reality photography**
This type of photography provides an interactive viewing experience. Digital images are made up of pixels, which are tiny squares of color that a computer arranges to create an image. The most common digital image formats are GIF, PNG, and JPEG.

Computer Vision

- Computer vision's goal is to <u>use computers to emulate human vision</u>, including learning and being able to make inferences and take actions based on visual inputs.

- This area itself is a branch of artificial intelligence (AI) whose objective is to emulate human intelligence.

- The field of AI is in its earliest stages of infancy in terms of development, with progress having been much slower than originally anticipated.

- The area of image analysis (also called image understanding) is in between image processing and computer vision.

The Origins of Digital Image Processing

- One of the first applications of digital images was in the newspaper industry, when pictures were first sent by submarine cable between London and New York.
- Introduction of the Bartlane cable picture transmission system in the early 1920s reduced the time required to transport a picture across the Atlantic from more than a week to less than three hours.
- Specialized printing equipment coded pictures for cable transmission and then reconstructed them at the receiving end.

A digital picture produced in 1921 from a coded tape by a telegraph printer with special type faces. (McFarlane.)

A digital picture made in 1922 from a tape punched after the signals had crossed the Atlantic twice. (McFarlane.)

Unretouched cable picture of Generals Pershing and Foch, transmitted in 1929 from London to New York by 15-tone equipment. (McFarlane.)

The Origins of Digital Image Processing

- The first computers for meaningful image processing tasks appeared in the early 1960s.
- The birth of what we call digital image processing today.
- Work on using computer techniques for improving images from a space probe began at the Jet Propulsion Laboratory (Pasadena, California) in 1964.
- Here, a computer processed pictures of the moon transmitted by Ranger 7 to correct various types of image distortion inherent in the on-board television camera.
- The imaging lessons learned with *Ranger 7* served as the basis for improved methods used to enhance and restore images.

The first picture of the moon by a U.S. spacecraft. *Ranger 7* took this image on July 31, 1964 at 9:09 A.M. EDT, about 17 minutes before impacting the lunar surface. (Courtesy of NASA.)

Fundamental Steps in Digital Image Processing

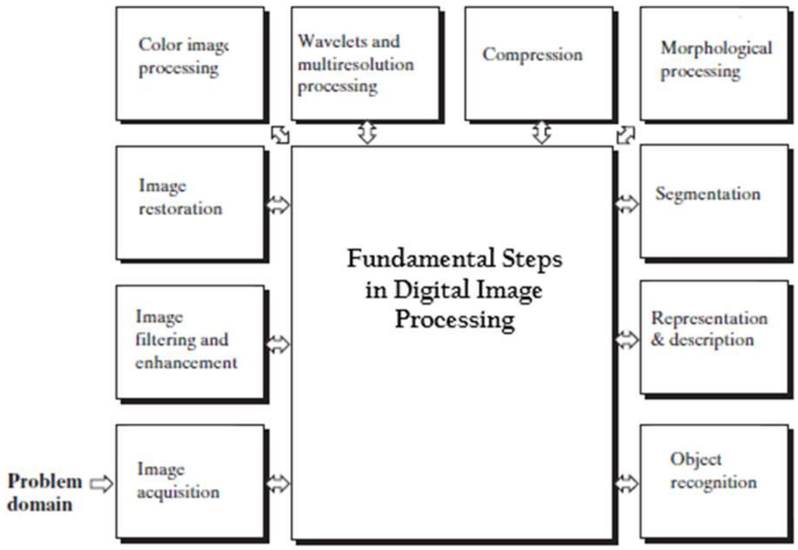


Image Acquisition

Elements of Visual Perception

Structure of the Human Eye

- The eye is nearly a sphere, with an average diameter of approximately 20 mm.
- Three membranes enclose the eye: the *cornea* and *sclera* outer cover, the choroid, and the *retina*.
- The cornea is a tough, transparent tissue covering the eye's anterior surface.
- Continuous with the cornea, the sclera is an opaque membrane enclosing the optic globe's remainder.

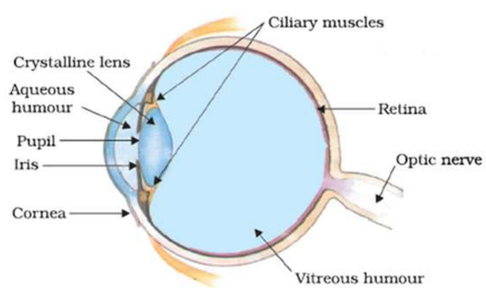
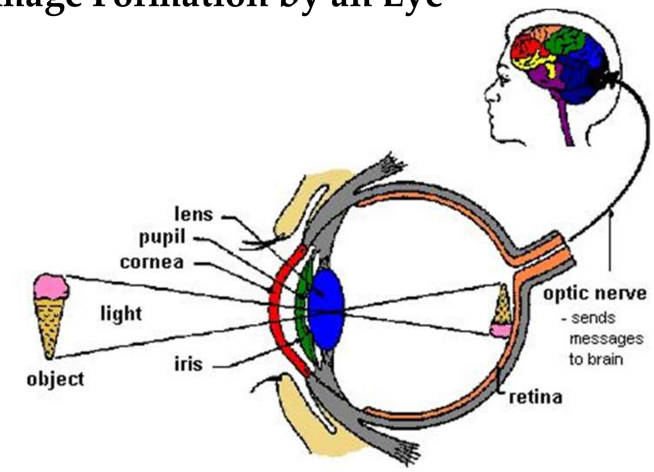


Image Acquisition

Image Formation by an Eye



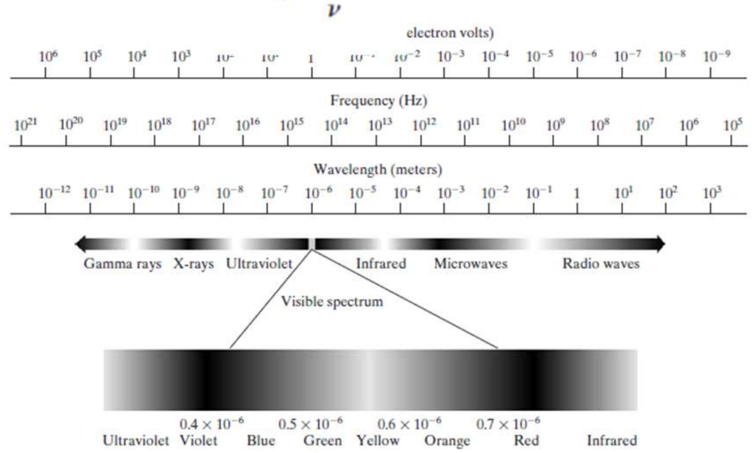
https://www.youtube.com/watch?v=YcedXDN6a88

Image Acquisition

Light and the Electromagnetic Spectrum

The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy. Wavelength and frequency are related by the expression where c is the speed of light

$$\lambda = \frac{c}{\nu}$$



The electromagnetic spectrum

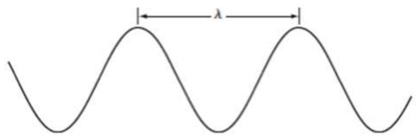
Image Acquisition

Light and the Electromagnetic Spectrum

- Electromagnetic waves can be visualized as propagating sinusoidal waves with wavelength, or they can be thought of as a stream of massless particles, each traveling in a wavelike pattern and moving at the speed of light.
- Each massless particle contains a certain amount (or bundle) of energy. Each bundle of energy is called a *photon*.

$$E = h\nu$$

where h is Planck's constant, e-energy, v-frequency



The electromagnetic spectrum

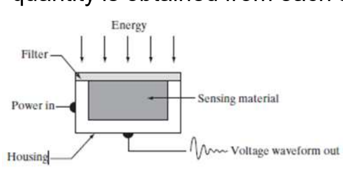
- **Energy is proportional to frequency**, so the higher-frequency (shorter wavelength) electromagnetic phenomena carry more energy per photon.
- Thus, radio waves have photons with low energies, microwaves have more energy than radio waves, infrared still more, then visible, ultraviolet, X-rays, and finally gamma rays, the most energetic of all.
- This is the reason why gamma rays are so dangerous to living organisms.

Image Acquisition

Transforming Illumination energy into Digital Images:

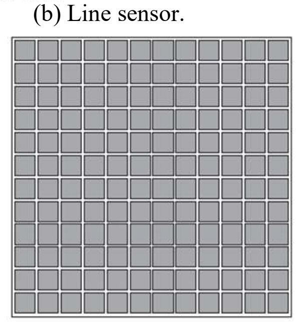
Acquisition using single sensor

- Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected.
- The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response.

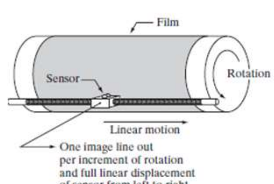


(a) Single imaging sensor.

(b) Line sensor.



(c) Array sensor.

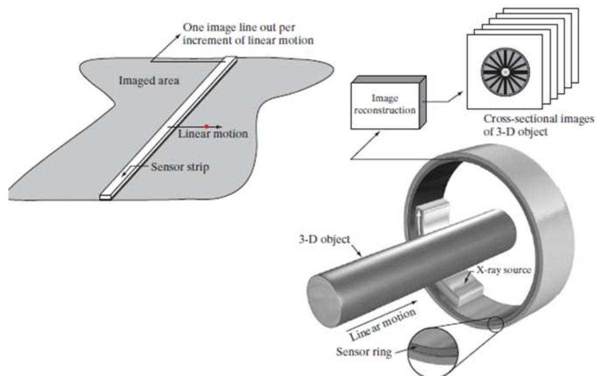


Combining a single sensor with motion to generate a 2-D image.

Image Acquisition

Transforming Illumination energy into Digital Images:
Acquisition using sensor strips

- The strip provides imaging elements in one direction.
- Motion perpendicular to the strip provides imaging in the other direction

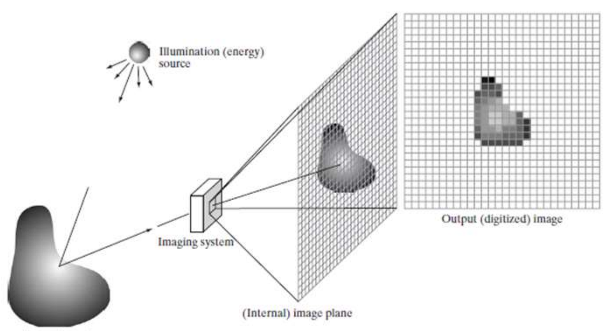


(a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

Image Acquisition

Simple Image Formation Model

- The strip provides imaging elements in one direction.
- Motion perpendicular to the strip provides imaging in the other direction



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

Image Acquisition

Simple Image Formation Model

Image denoted as a 2-dimensional function of the form

f(x,y) ……………….. (1)

When an image is generated from a physical process, its **intensity values are proportional to energy** radiated by a physical source
(e.g., electromagnetic waves)
Hence,
0<f(x,y)< *Infinity…....(2)*

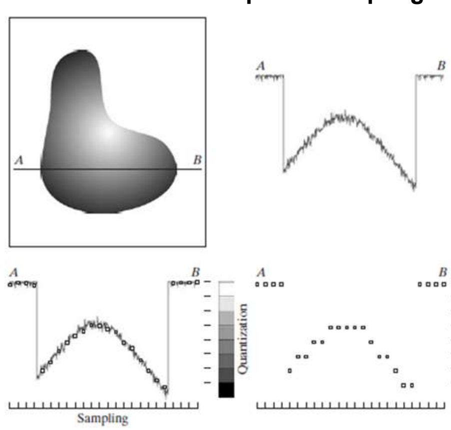
F(x,y) has two components,

1. The amount of source **illumination incident** on the scene being viewed ***(illumination-i(x,y))***
2. The amount of **illumination reflected** by the objects in the scene ***(reflectance-r(x,y).***

Hence,
f(x,y)= i(x,y) r(x,y)……..(3)
where* 0<i(x,y)< *Infinity, and* 0<r(x,y)< *1

Image Sampling and Quantization

Basic Concepts in Sampling and Quantization



Generating a digital image.
(a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

- The figure shows a continuous image *f* that is to be converted to digital form.
- An image may be continuous with respect to the *x*- and *y*-coordinates, and also in amplitude.
- To convert it to digital form, sample the function in both coordinates and in amplitude.
- Digitizing the coordinate values is called *sampling*.
- Digitizing the amplitude values is called *quantization*

Image Sampling and Quantization

Some Sampling techniques consequences

<u>Single Sensor for generating image:</u>

- When an image is generated by a ***<u>single sensing element combined with mechanical motion,</u>*** the output of the sensor is quantized as above.
- However, spatial sampling is accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data.
- Mechanical motion can be made very exact so, in principle, there is almost no limit as to how fine we can sample an image using this approach.
- In practice, limits on sampling accuracy are determined by other factors, such as the quality of the optical components of the system.

<u>Sensing strips for image acquisition:</u>

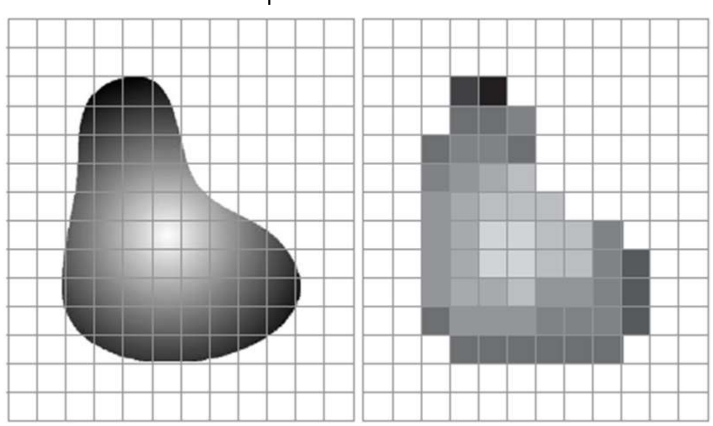
- When a sensing strip is used for image acquisition, the number of sensors in the strip establishes the sampling limitations in one image direction.
- Mechanical motion in the other direction can be controlled more accurately, but it makes little sense to try to achieve sampling density in one direction that exceeds the sampling limits established by the number of sensors in the other.
- Quantization of the sensor outputs completes the process of generating a digital image.

Image Sampling and Quantization

Some Sampling techniques consequences

<u>Image generated by sensing array:</u>

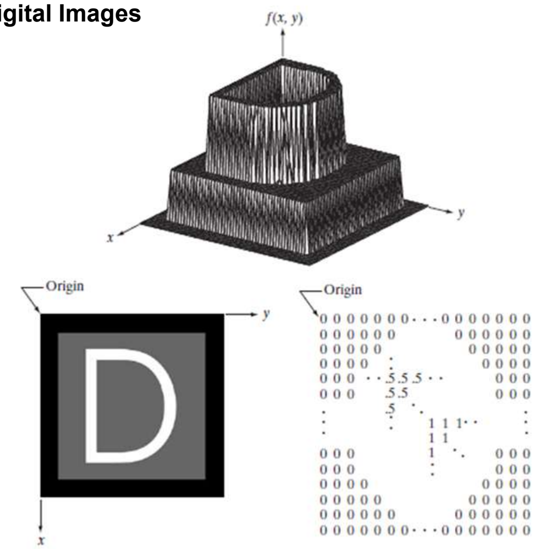
- When a ***<u>sensing array</u>*** is used for image acquisition, there is no motion and the number of sensors in the array establishes the limits of sampling in both directions. Quantization of the sensor outputs is as before.



(a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

Image Sampling and Quantization

Representing Digital Images



(a) Image plotted as a surface. (b) Image displayed as a visual intensity array.
(c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

Monochrome and Multichrome Images

Monochrome Imaging:

Definition:

- Monochrome images are represented using intensity values of a single channel.
- Mathematically, each pixel's intensity I(x, y) is a scalar value:

$$I(x, y) \in [0,255]$$

Where x and y are the spatial coordinates of the pixel, and the range [0, 255] represents grayscale intensity (for an 8-bit image).

Characteristics:

The image matrix I is two-dimensional:

$I=\{I(x, y)\}$, where I(x, y) represents the intensity at each pixel.

Advantages:

Data size is smaller due to the single channel of information.

Ideal for scenarios where color is unnecessary but detail and texture are crucial.

Monochrome and Multichrome Images

Multichrome Imaging:

Definition:
Multichrome images are represented using multiple color channels, typically Red (R), Green (G), and Blue (B).
The intensity at each pixel is represented as a vector:
I(x,y)=[R(x,y), G(x,y), B(x,y)]
where R(x,y),G(x,y),B(x,y) ∈ [0,255] for 8-bit color depth.

Characteristics:
The image matrix I is three-dimensional:
I={[R(x,y),G(x,y),B(x,y)]}, with each channel representing a layer of the image.

Advantages:
Multichrome imaging enables color reconstruction by combining R, G, and B values to represent realistic scenes.

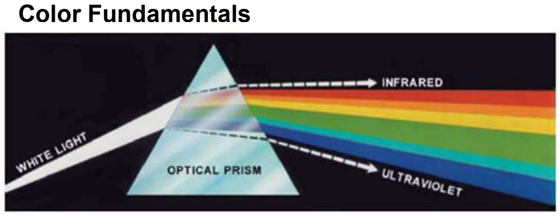
Dr. Nilkanth Deshpande | **Computer Vision** | **TY BTech (2022-26)**

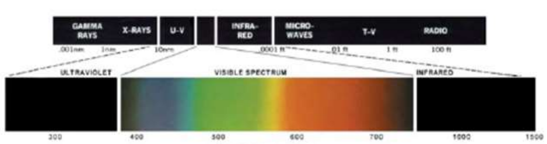
Concept of color spaces

Color Images processing (Broad Types):

a) **Full-color processing: The** images in question typically are acquired with a full-color sensor, such as a color TV camera or color scanner.
b) **Pseudo-color processing:** the problem is one of assigning a color to a particular monochrome intensity or range of intensities

Color Fundamentals





Wavelengths comprising the visible range of the electromagnetic spectrum.

- In 1666, Sir Isaac Newton discovered that when a beam of sunlight passes through a glass prism, the emerging beam of light is not white but consists of a continuous spectrum of colors ranging from violet at one end to red at the other.
- Basically, the colors that humans and some other animals perceive in an object are determined by the nature of the light reflected from the object

Dr. Nilkanth Deshpande | **Computer Vision** | **TY BTech (2022-26)**

Concept of color spaces

❑ Achromatic Light:

If the light is achromatic (void of color), its only attribute is its *intensity* or amount. Achromatic light is what viewers see on a black-and-white television set, and it has been an implicit component of our discussion of image processing thus far.

❑ Chromatic light:

It spans the electromagnetic spectrum from approximately 400 to 700 nm. Three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance, and brightness.

- ***Radiance*** is the total amount of energy that flows from the light source, usually measured in watts (W).
- ***Luminance***, measured in lumens (lm), gives a measure of the amount of energy an observer *perceives* from a light source.
- ***Brightness*** is a subjective descriptor that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation.

Concept of color spaces

➢ Most color models in use today are oriented either ***toward hardware*** (such as for color monitors and printers) or ***toward applications*** where color manipulation is a goal (such as in the creation of color graphics for animation).

In the case of Digital Image Processing, the hardware-oriented models most commonly used in practice are as:

1. **RGB (red, green, blue)** model for color monitors and a broad class of color video cameras;

2. **CMY** (cyan, magenta, yellow) and **CMYK** (cyan, magenta, yellow, black) models for color printing;

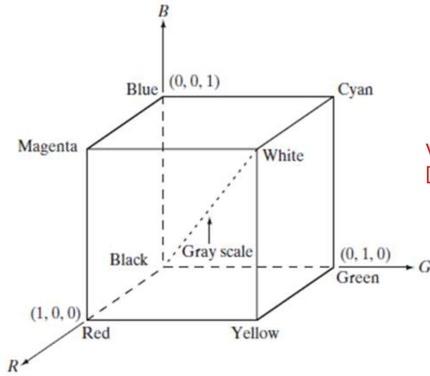
3. **HSI** (hue, saturation, intensity) model, which corresponds closely with the way humans describe and interpret color.

➢ The HSI model also has the advantage that it decouples the color and gray-scale information in an image, making it suitable for many of the gray-scale techniques developed.

Concept of color spaces

RGB (red, green, blue) model for color monitors and a broad class of color video cameras.

- ✓ Here each color appears in its primary spectral components of red, green, and blue.
- ✓ This model is based on a Cartesian coordinate system.

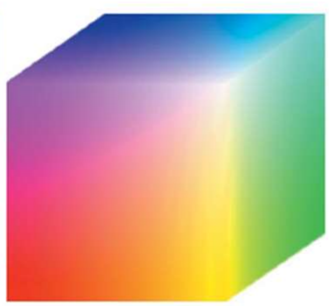


Values of *R*, *G*, and *B* are assumed to be in the range [0, 1].

Schematic of the RGB color cube. Points along the main diagonal have gray values, from black at the origin to white at point (1, 1, 1).

Concept of color spaces

Images represented in the RGB color model consist of three component images, one for each primary color. When fed into an RGB monitor, these three images combine on the screen to produce a composite color image



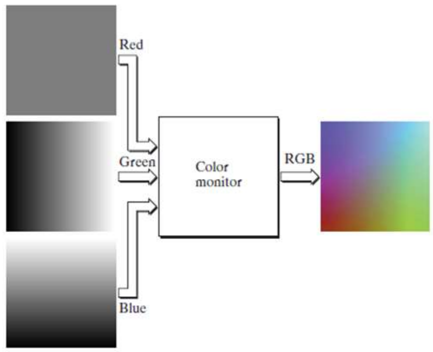
RGB 24-bit color cube.

- The number of bits used to represent each pixel in RGB space is called the *pixel depth*.
- Consider an RGB image in which each of the red, green, and blue images is an 8-bit image.
- Under these conditions each RGB *color* pixel [that is, a triplet of values (*R*, *G*, *B*)] is said to have a depth of 24 bits (3 image planes times the number of bits per plane).
- The term *full-color* image is used often to denote a 24-bit RGB color image.
- The total number of colors in a 24-bit RGB image shows the 24-bit RGB color cube corresponding to the diagram.

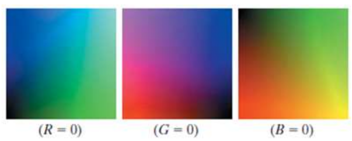
$(2^8)^3 = 16,777,216$.

Concept of color spaces

Acquiring **a color image**



(a) Generating the RGB image of the cross-sectional color plane (127, G, B).



(b) The three hidden surface planes in the color cube of previous slide

Concept of color spaces

The CMY and CMYK Color Models

- Cyan, magenta, and yellow are the secondary colors of light or, alternatively, the primary colors of pigments

For example, when a surface coated with cyan pigment is illuminated with white light, no red light is reflected from the surface.

That is, cyan subtracts red light from reflected white light, which itself is composed of equal amounts of red, green, and blue light.

- Most devices that deposit colored pigments on paper, such as color printers and copiers, require CMY data input or perform an RGB to CMY conversion internally. This conversion is performed using the simple operation.

- Similarly, pure magenta does not reflect green, and pure yellow does not reflect blue

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Where, color values normalized to the range [0, 1].

The above equation shows that RGB values can be obtained easily from a set of CMY values by subtracting the individual CMY values from 1.

Dr. Nilkanth Deshpande Computer Vision TY BTech (2022-26)