MIT-ADT UNIVERSITY
PUNE, INDIA

Course Code – Foundations of Artificial Intelligence

Class -L.Y. (SEM-I), AIA

Unit - I Introduction to Digital Image Processing

Dr. Nilima Kulkarni- Course Coordinator
Prof. Jyoti Manoorkar

AY 2025-2026 SEM-I

Introduction

"One picture is worth more than ten thousand words"

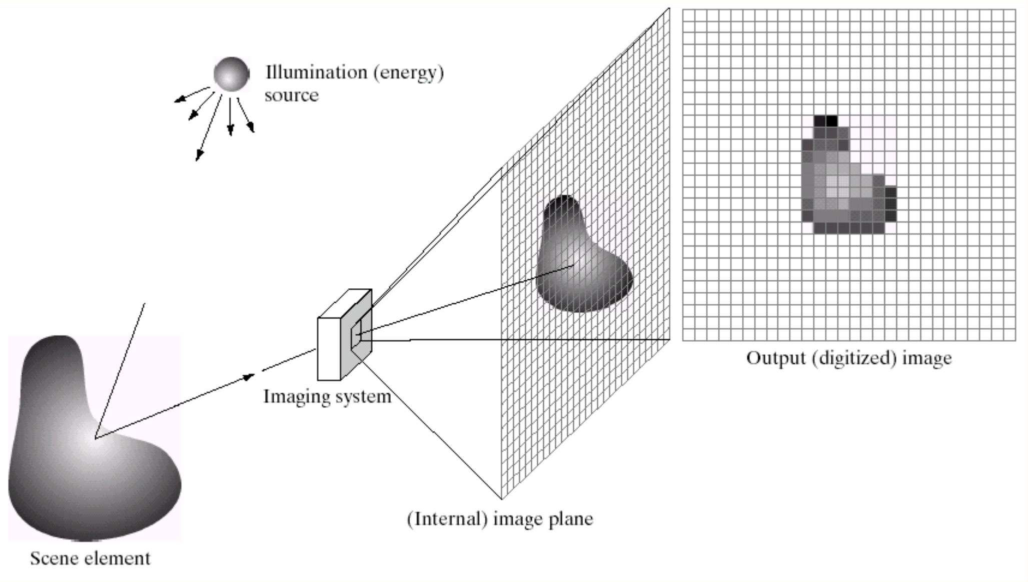
Anonymous

Contents

- Introduction to Image processing
- Fundamental steps in image processing
- Components of image processing system
- Pixels
- Coordinate conventions
- Imaging Geometry
- Spatial Domain
- Frequency Domain
- Sampling and quantization
- Basic relationship between pixels
- Applications of Image Processing

What is a Digital Image?

A **digital image** is a representation of a two-dimensional image as a finite set of digital values, called picture elements or pixels

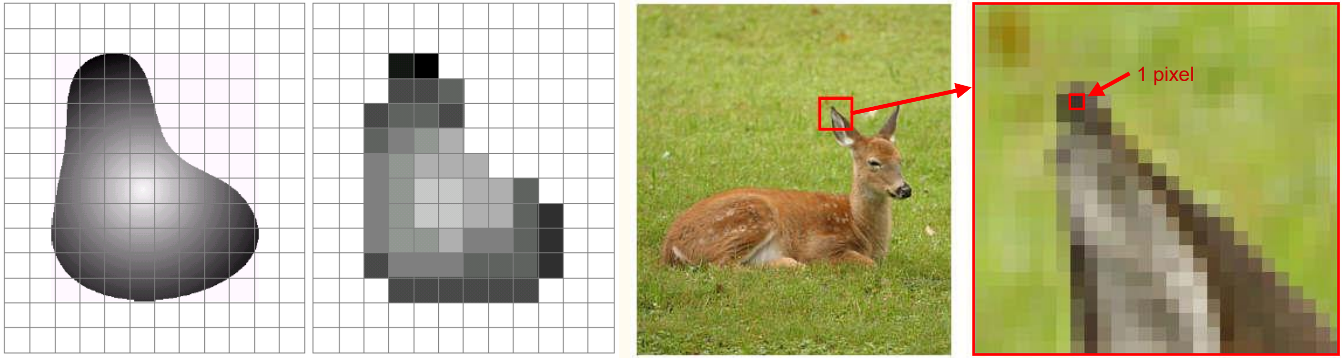


Images taken from Gonzalez & Woods, Digital Image Processing (2002)

What is a Digital Image? (cont…)

Pixel values typically represent gray levels, colours, heights, opacities etc

Remember *digitization* implies that a digital image is an *approximation* of a real scene



Images taken from Gonzalez & Woods, Digital Image Processing (2002)

What is a Digital Image? (cont...)

Common image formats include:

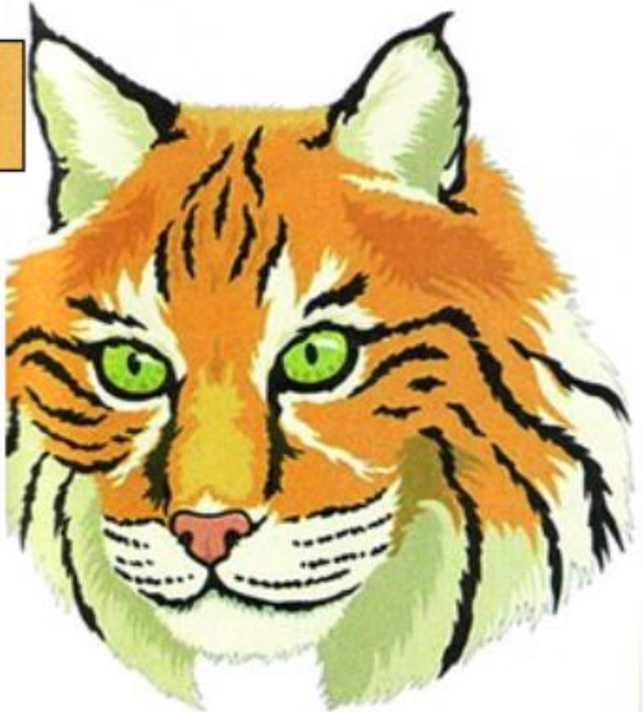
- 1 sample per point (B&W or Grayscale)
- 3 samples per point (Red, Green, and Blue)

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

- An **image** is the graphical and visual representation of some **information** that can be **displayed** on a computer screen or **printed** out

- Images come in a variety of forms:
  - Photographs
  - Drawings
  - Paintings
  - Television and motion pictures
  - Semantics
  - Maps etc.

- Images show us the prominent features of the objects that they represent.



These images are composed quite differently, each is an effective representation of its subject

What is Digital Image Processing?

Explanation:
Digital Image Processing (DIP) involves using **computer algorithms to perform operations on images** to improve their quality, extract information, or transform them for further analysis.

Key Points:

- Input: Digital image (array of pixel values)
- Processing: Using mathematical techniques
- Output: Enhanced or modified image

Example:

- **Instagram Filters:** When you adjust brightness or apply a filter, the app is performing DIP.

What is Digital Image Processing?

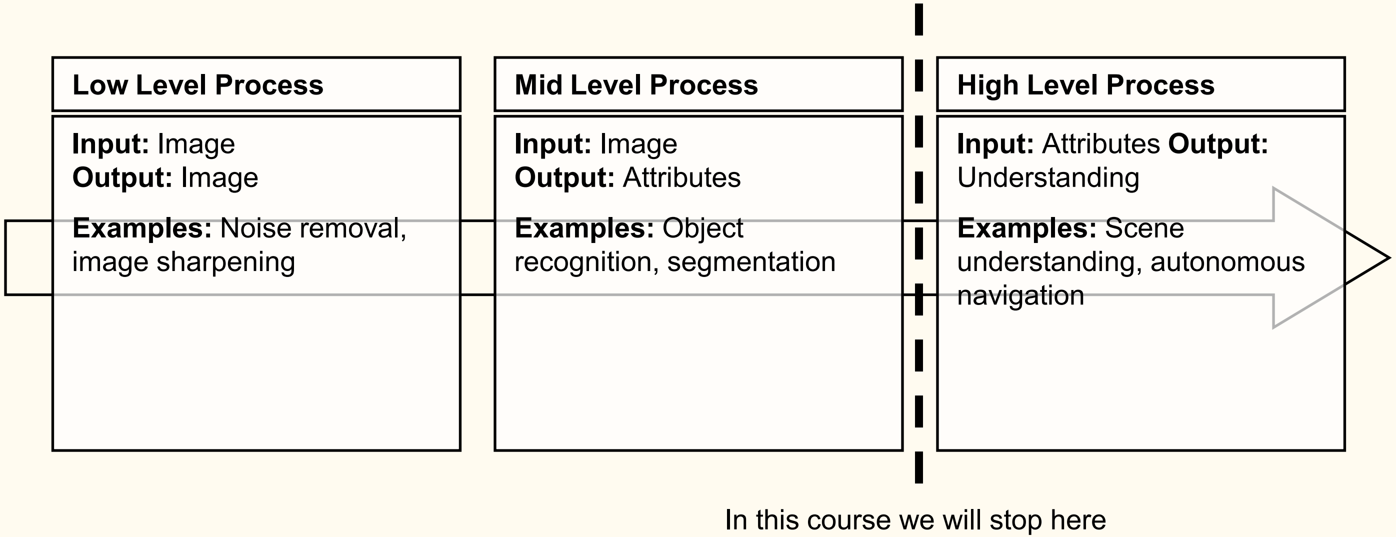
Digital image processing focuses on two major tasks

- Improvement of pictorial information for human interpretation
- Processing of image data for storage, transmission and representation for autonomous machine perception

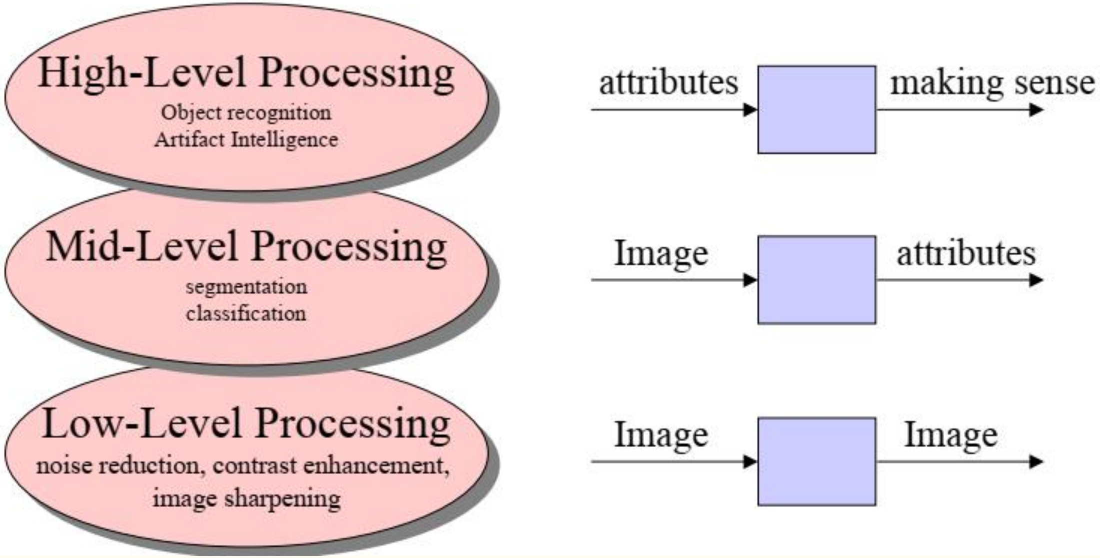
Some argument about where image processing ends and fields such as image analysis and computer vision start

What is DIP? (cont...)

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



- ***Image processing to computer vision***

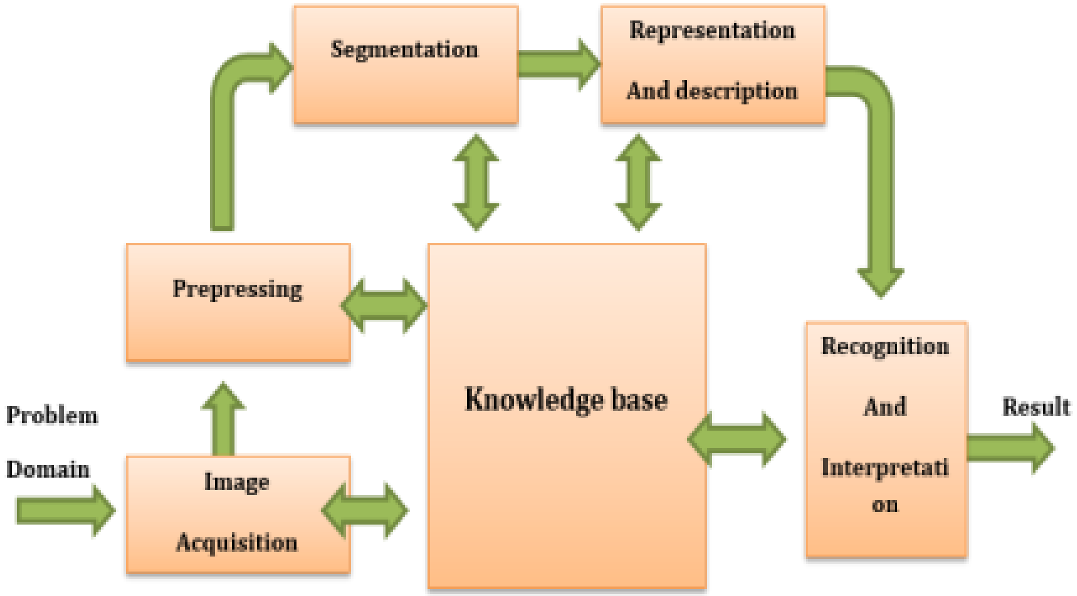


Fundamental Steps in Image Processing

Description:
Digital image processing typically involves a **series of well-defined steps**. Each step addre

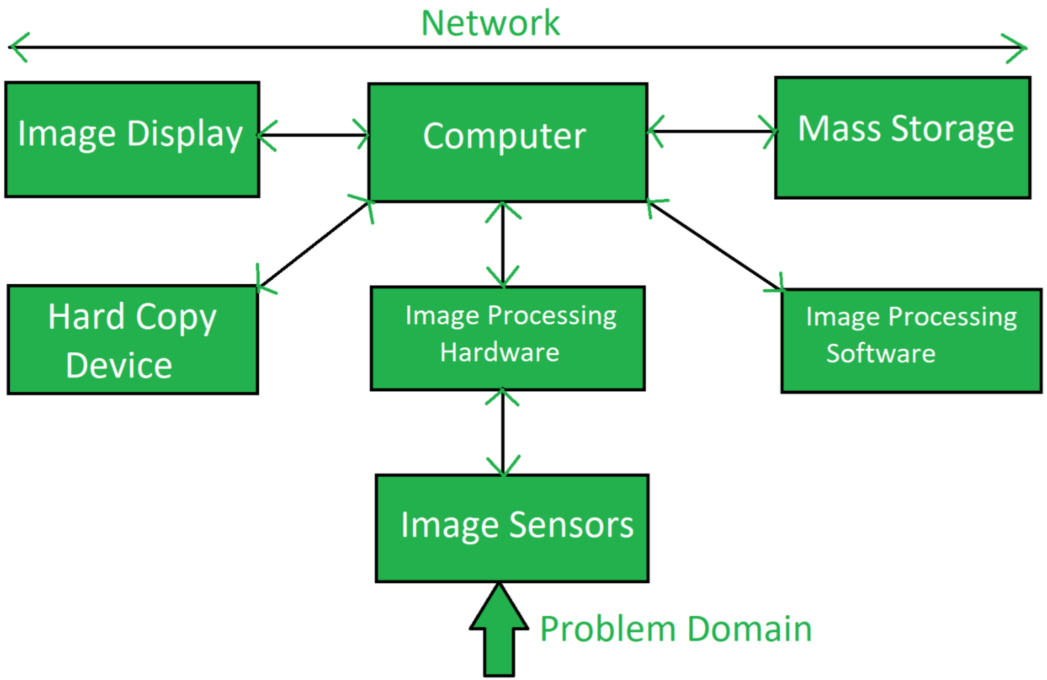
Steps with Explanation and Example:

1. **Image Acquisition**
   - Capturing an image using a sensor (camera/scanner).
   - *Example:* Taking a photo with your smartphone.
2. **Preprocessing**
   - Improving image quality by removing noise or correcting illumination.
   - *Example:* Smoothing a blurry photo.
3. **Segmentation**
   - Dividing the image into regions of interest.
   - *Example:* Identifying faces in a group photo.
4. **Representation and Description**
   - Converting segmented regions into a form suitable for analysis.
   - *Example:* Calculating area and shape of an object.
5. **Recognition**
   - Assigning labels to objects.
   - *Example:* Recognizing handwritten digits.
6. **Knowledge Base**
   - Using prior knowledge to guide processing.
   - *Example:* Knowing a car's typical shape helps in detection.



Components of Image Processing System

Explanation:
An image processing system has **hardware and software components** working together.



Components with Description:

- **Image Sensors**
  - Devices capturing images (e.g., CCD camera, flatbed scanner).
- **Digitizer**
  - Converts analog signals to digital.
- **Processing Hardware**
  - Computers or specialized processors that run algorithms.
- **Software**
  - Programs implementing image enhancement, segmentation, etc.
- **Mass Storage**
  - Databases storing images and results.
- **Display**
  - Monitors or projectors to visualize results.
- **Hardcopy Devices**
  - Printers for output.
- **Networking**
  - Sharing images across systems.

Real-Life Example: A medical imaging workstation includes an MRI scanner (sensor), digitizer, image processing software, and storage server.

What is a Pixel?

Explanation:
A **pixel (picture element)** is the **smallest unit** of a digital image, representing intensity or color.

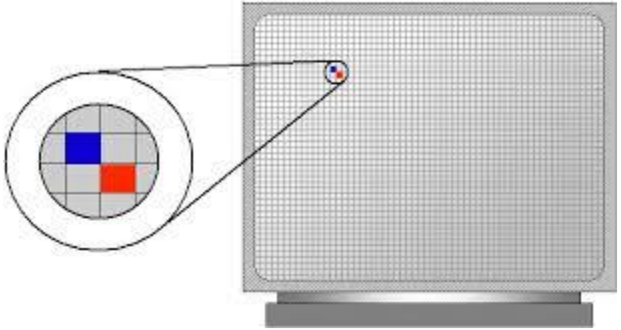
Details:

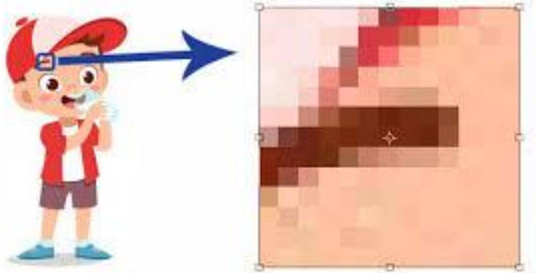
- Each pixel has coordinates (x, y).
- A grayscale image pixel contains one value (0–255).
- A color image pixel has three values (R, G, B).

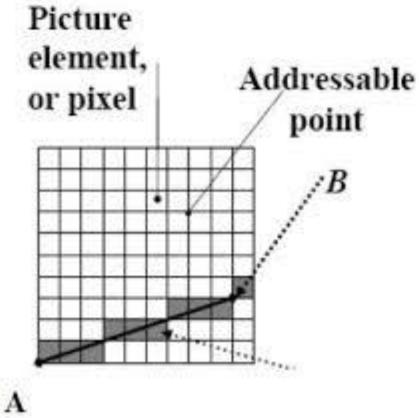
Example:

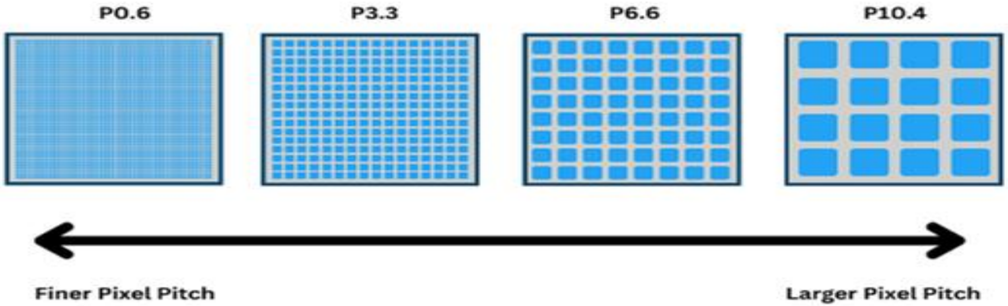
- A 1024×768 image has **786,432 pixels**.
- In an 8-bit grayscale image, 0=black, 255=white.

- A **pixel** is the smallest element of resolution on a computer screen (Screen Resolution)
- A pixel is the basic unit of a digital images. Digital image is a picture that may be stored in, displayed on, processed by a computer.
- As mentioned, **bitmap** is composed of a matrix elements called **pixels**
- Each pixel can be in a specific **colour** and each pixel consists of two or more colors.
- The range of these colours is known as the **colour depth** .
- The color depth determined “How much data in bits used to determined the number of colors”.
- Colour depth is measured in **bits per pixel**
  - Remember: a **bit** (**binary digit**) is either **1** or **0** and that there are eight bits in a **byte**



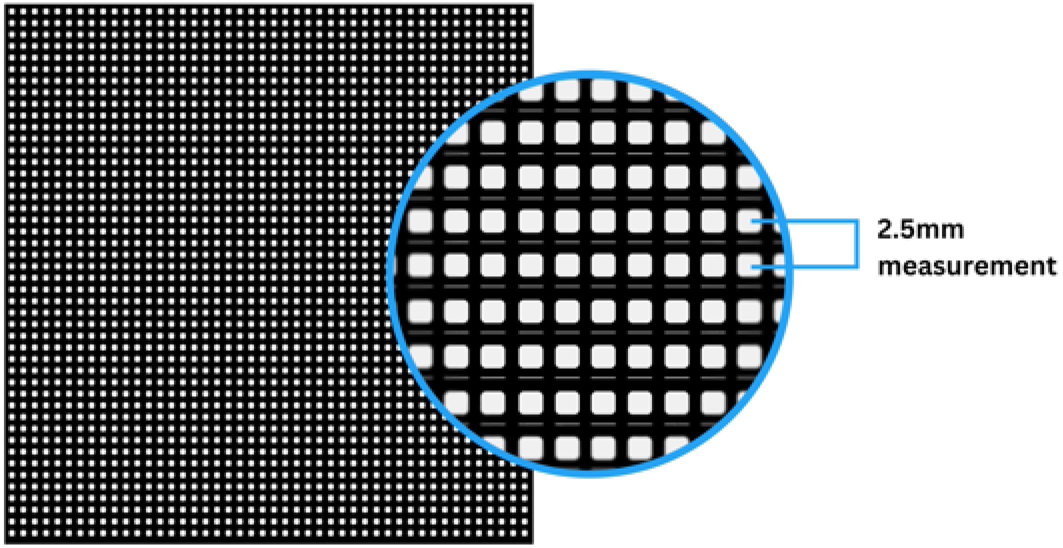






Pixel pitch describes the density of the pixels (LED clusters) on an LED display and correlates with resolution. Sometimes referred to as pitch or dot pitch, the pixel pitch is the distance in millimeters from the center of a pixel to the center of the adjacent pixel. Since pixel pitch indicates the amount of space between two pixels, a smaller pixel pitch means there is less empty space between pixels. This equates to higher pixel density and improved screen resolution.

2.5 Pixel Pitch



1 bit per pixel = 2 colours (monochrome)

2 bits per pixel = 4 colours

4 bits per pixel = 16 colours

8 bits per pixel = 256 colours

Generally good enough for colour images

16 bits per pixel = 65536 colours

Better quality for photograph-like images, also known as *high colour*

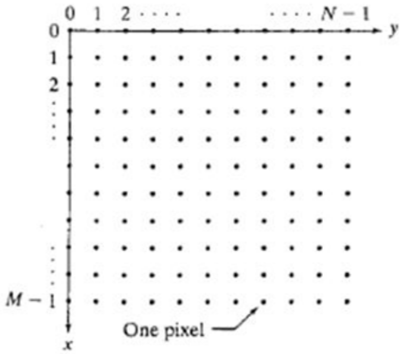
24 bits per pixel = >16 million possible colours

Used to recreate photo realistic images, also known as *true colour*

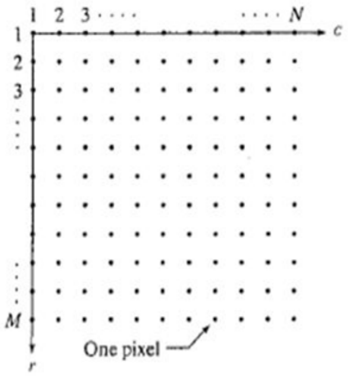
Coordinate Conventions

- Images use a **2D coordinate system** to identify pixel positions.
    - The **origin (0,0)** is usually at the **top-left corner**.
    - X increases to the right.
    - Y increases downward.

Digital Image Representation Coordinate Conventions



(A)



(B)

The result of sampling and quantization is a matrix of real numbers.

Assume that an image f(x,y) is sampled so that the resulting image has M rows and N columns. We say that the image is of size M x N.

The image origin is defined to be at (x, y)=(0, 0). The next coordinate value along the first row of the image are (x, y)= (0, 1).

Note that x ranges from 0 to M-1, and y ranges from 0 to N-1 in integer increments

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & f(0,2) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & f(1,2) & \dots & f(1,N-1) \\ \vdots & \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & f(M-1,2) & \dots & f(M-1,N-1) \end{bmatrix}$$

Origin from (0,0) or (1,1) both way possible

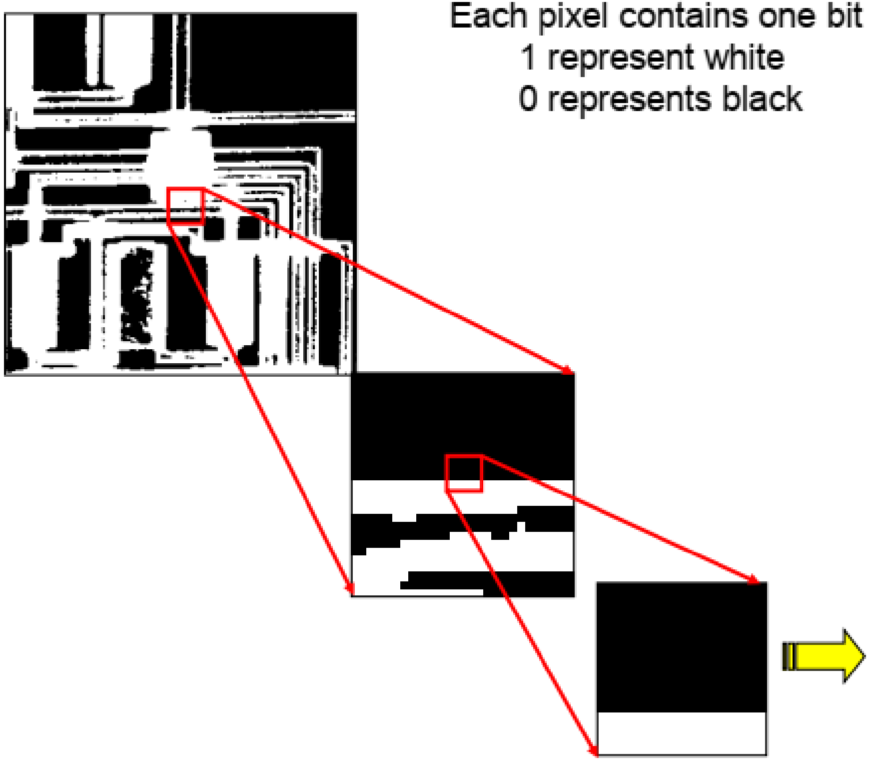
Image Types : Binary Image

Binary image or black and white image

Each pixel contains one bit :

1 represent white

0 represents black



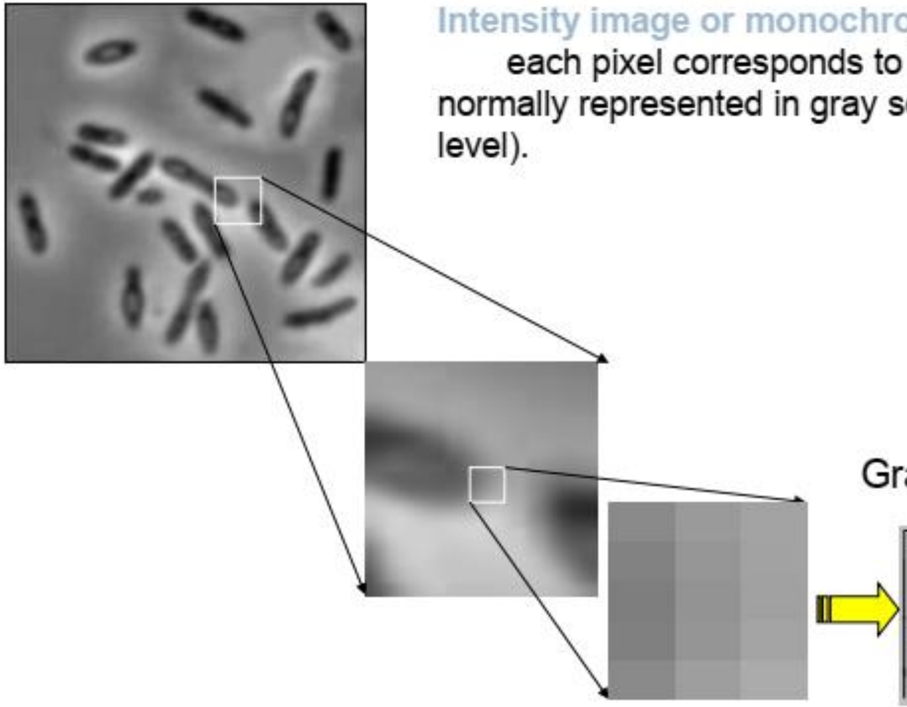
Binary data

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Digital Image Types : Intensity Image

Intensity image or monochrome image

each pixel corresponds to light intensity normally represented in gray scale (gray level).



Gray scale values

$$\begin{bmatrix} 10 & 10 & 16 & 28 \\ 9 & 6 & 26 & 37 \\ 15 & 25 & 13 & 22 \\ 32 & 15 & 87 & 39 \end{bmatrix}$$

A ***binary image*** is represented by pixels consisting of 1 bit each, which can represent two tones (typically black and white), using the values 0 for black and 1 for white or vice versa.

Eg:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

A ***grayscale image*** is composed of pixels represented by multiple bits of information, typically ranging from 2 to 8 bits or more.

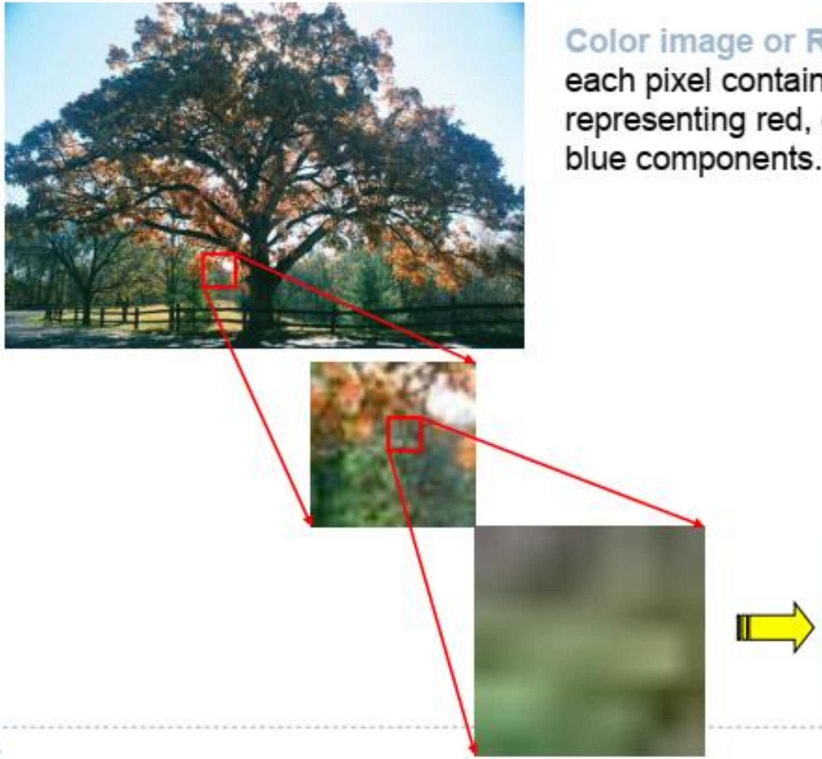
Eg:

$$\begin{bmatrix} 10 & 10 & 16 & 28 \\ 9 & 6 & 26 & 37 \\ 15 & 25 & 13 & 22 \\ 32 & 15 & 87 & 39 \end{bmatrix}$$

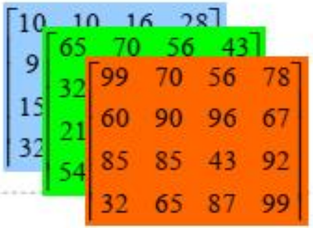
Example: In a 2-bit image, there are four possible combinations: 00, 01, 10, and 11. If "00" represents black, and "11" represents white, then "01" equals dark gray and "10" equals light gray. The bit depth is two, but the number of tones that can be represented is 2^2 or 4. At 8 bits, 256 (2^8) different tones can be assigned to each pixel.

Digital Image Types : RGB Image

Color image or RGB image:
each pixel contains a vector representing red, green and blue components.



RGB components



A ***color image*** is typically represented by a bit depth ranging from 8 to 24 or higher. With a 24-bit image, the bits are often divided into three groupings: 8 for red, 8 for green, and 8 for blue. Combinations of those bits are used to represent other colors. A 24-bit image offers 16.7 million (2^{24}) color values. Increasingly scanners are capturing 10 bits or more per color channel and often outputting 8 bits to compensate for "noise" in the scanner and to present an image that more closely mimics human perception.

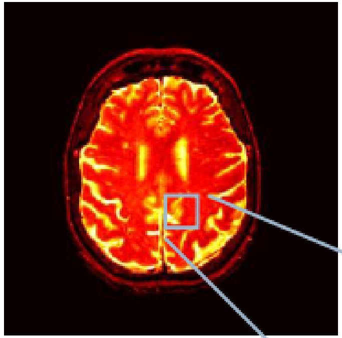
Eg:

$$R=\begin{bmatrix} 99 & 70 & 56 & 78 \\ 60 & 90 & 96 & 67 \\ 85 & 85 & 43 & 92 \\ 32 & 65 & 87 & 99 \end{bmatrix} \quad G=\begin{bmatrix} 65 & 70 & 56 & 43 \\ 32 & 54 & 96 & 67 \\ 21 & 54 & 47 & 42 \\ 54 & 65 & 65 & 39 \end{bmatrix} \quad B=\begin{bmatrix} 10 & 10 & 16 & 28 \\ 9 & 6 & 26 & 37 \\ 15 & 25 & 13 & 22 \\ 32 & 15 & 87 & 39 \end{bmatrix}$$

Image Types : Index Image

Index image

Each pixel contains index number pointing to a color in a color table



$$\begin{bmatrix} 1 & 4 & 9 \\ 6 & 4 & 7 \\ 6 & 5 & 2 \end{bmatrix}$$

Index value

Color Table

Index No.	Red component	Green component	Blue component
1	0.1	0.5	0.3
2	1.0	0.0	0.0
3	0.0	1.0	0.0
4	0.5	0.5	0.5
5	0.2	0.8	0.9
...	...	...	...

Imaging Geometry

Explanation:

Imaging geometry describes **how a 3D scene is projected onto a 2D image plane**.

Concepts:

- **Perspective Projection:**
  - Distant objects appear smaller.
- **Orthographic Projection:**
  - Parallel projection without depth distortion.
- **Focal Length:**
  - Distance between lens and sensor, affecting field of view.

Real-Time Example:

- When you take a photo of a building, the top looks narrower due to perspective.

Imaging geometry, in digital image fundamentals, is the study of the spatial relationships between the object being imaged and the image plane.

It involves understanding how a 3D scene is projected onto a 2D image, considering factors like camera position, orientation, and the properties of the lens.

This understanding is crucial for various applications, including image processing, computer vision, and medical imaging.

Spatial and Frequency Domains

Why domain transformation required..

Many times, image processing tasks are best performed in a domain other than the spatial domain.

Moreover, it is easy to detect some features in a particular domain, i.e., a new information can be obtained in other domains.

Difference between spatial domain and frequency domain

In spatial domain, we deal with images as it is. The value of the pixels of the image change with respect to scene.
Whereas in frequency domain, we deal with the rate at which the pixel values are changing in spatial domain.
For simplicity, Let’s put it this way.

Spatial domain

In simple spatial domain, we directly deal with the image matrix.

Spatial Domain-

An image can be represented in the form of a 2D matrix where each element of the matrix represents pixel intensity. This state of 2D matrices that depict the intensity distribution of an image is called Spatial Domain. It can be represented as shown below-

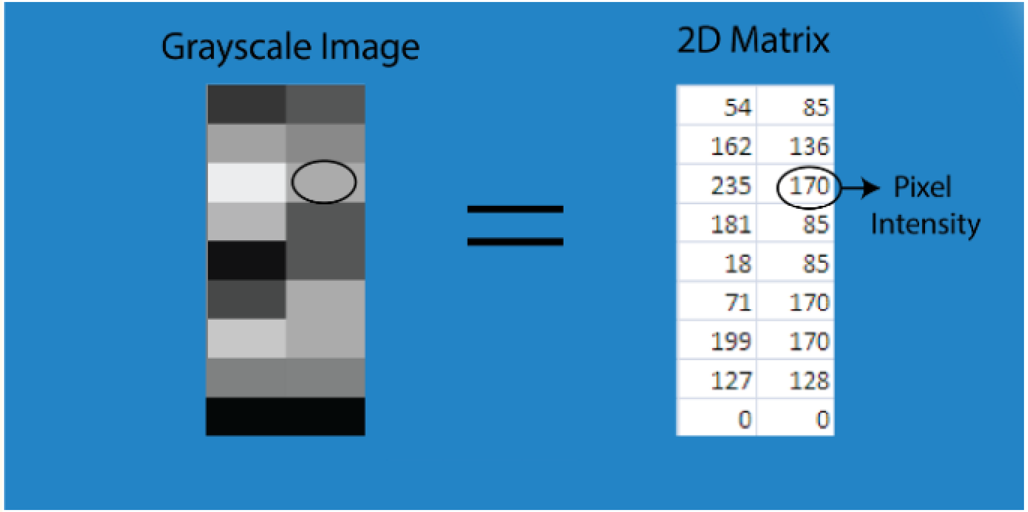
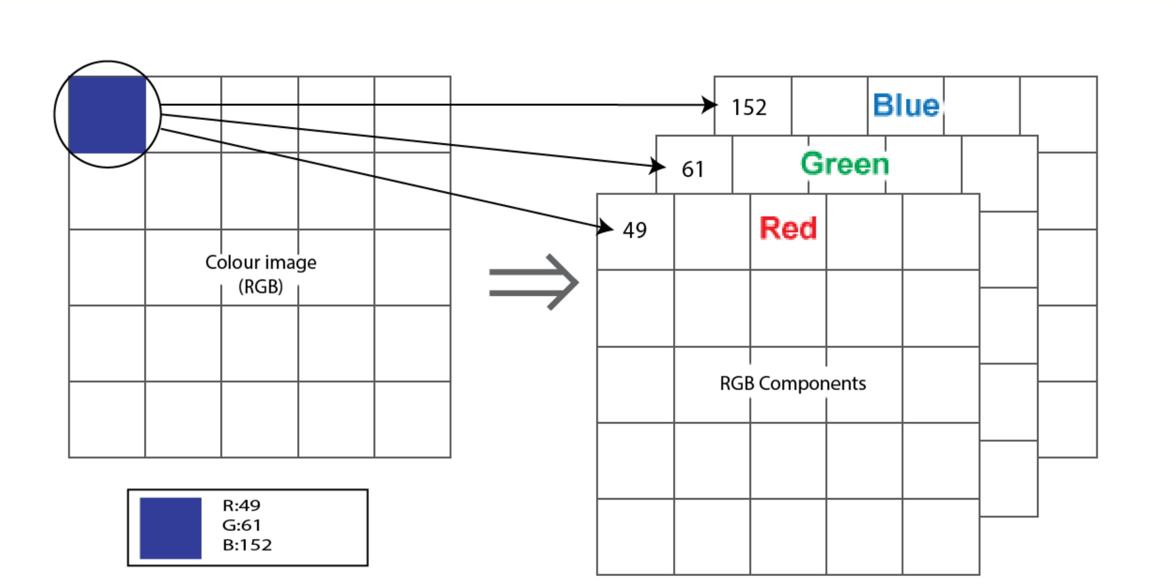


Illustration of Spatial Domain

For the RGB image, the spatial domain is represented as a 3D vector of 2D matrices. Each 2D matrix contains the intensities for a single color as shown below



Each pixel intensity is represented as $I(x,y)$ where x,y is the co-ordinate of the pixel in the 2D matrix. Different operations are carried out in this value.

For example- operation T(say, addition of 5 to all the pixel) is carried out in $I(x,y)$ which means that each pixel value is increased by 5.

This can be written as-

$I'(x,y) = T[I(x,y)]$

where, $I'(x,y)$ is the new intensity after adding 5 to $I(x,y)$.

Frequency Domain-

In frequency-domain methods are based on Fourier Transform of an image. Roughly, the term frequency in an image tells about the rate of change of pixel values.

Frequency Domain

Definition:

Processing images by **transforming them into frequency components** (how rapidly pixel intensities change).

Techniques:

- **Fourier Transform:**
  - Converts spatial representation to frequency representation.
- **Filtering:**
  - Emphasizing or reducing specific frequencies.
- ex>, JPEG compression removes high frequencies to reduce file size.

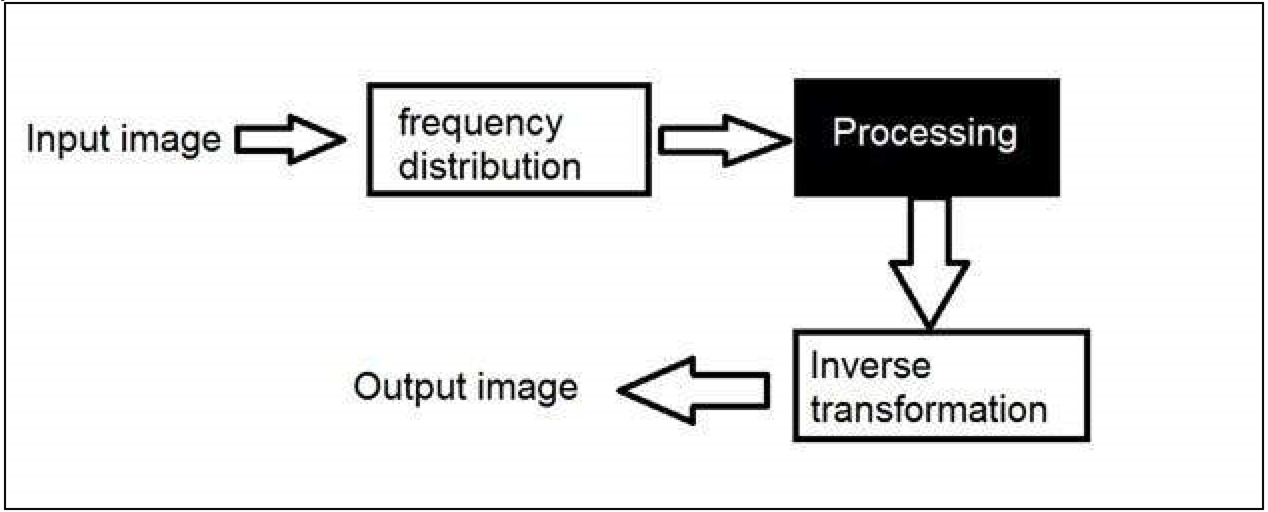
In spatial domain, we deal with images as it is. The value of the pixels of the image change with respect to scene.

Whereas in frequency domain, we deal with the rate at which the pixel values are changing in spatial domain.

Frequency Domain

We first transform the image to its frequency distribution. Then our black box system perform what ever processing it has to performed, and the output of the black box in this case is not an image, but a transformation. After performing inverse transformation, it is converted into an image which is then viewed in spatial domain.

It can be pictorially viewed as



Sampling and Quantization

Sampling and Quantization

Explanation:
These are **two key steps in digitizing analog images.**

Sampling:

- Selecting pixel locations (resolution).
- Example: 300 dpi scanner samples more densely than 72 dpi.

Quantization:

- Assigning numerical values to pixel intensities.
- Example: 8-bit quantization gives 256 gray levels.
- Ex. Saving a photo in low resolution reduces quality due to coarse sampling and quantization.

In Digital Image Processing, signals captured from the physical world need to be translated into digital form by “Digitization” Process. In order to become suitable for digital processing, an image function f(x,y) must be digitized both spatially and in amplitude. This digitization process involves two main processes called

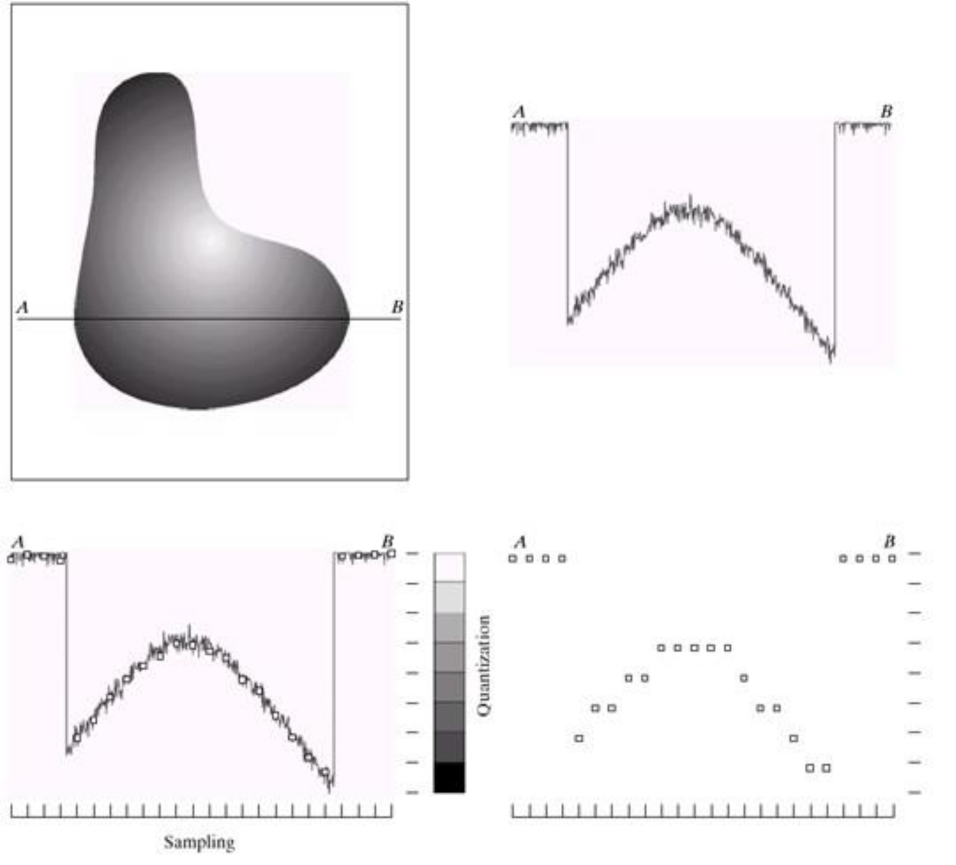
1. **Sampling:** Digitizing the co-ordinate value is called sampling.
2. **Quantization:** Digitizing the amplitude value is called quantization

Typically, a frame grabber or digitizer is used to sample and quantize the analogue video signal.

SAMPLING AND QUANTIZATION

Discretizing coordinate values is called Sampling

Discretizing the amplitude values is called Quantization



The *quality of a digital image* is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization.

However image content is also an important consideration in choosing these parameters

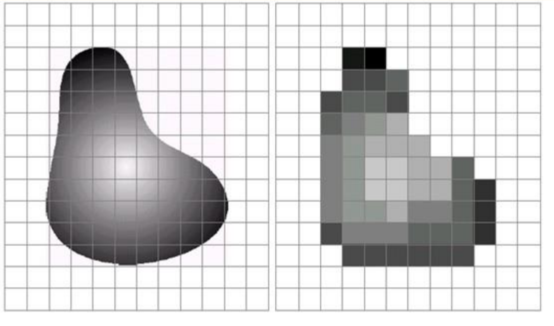
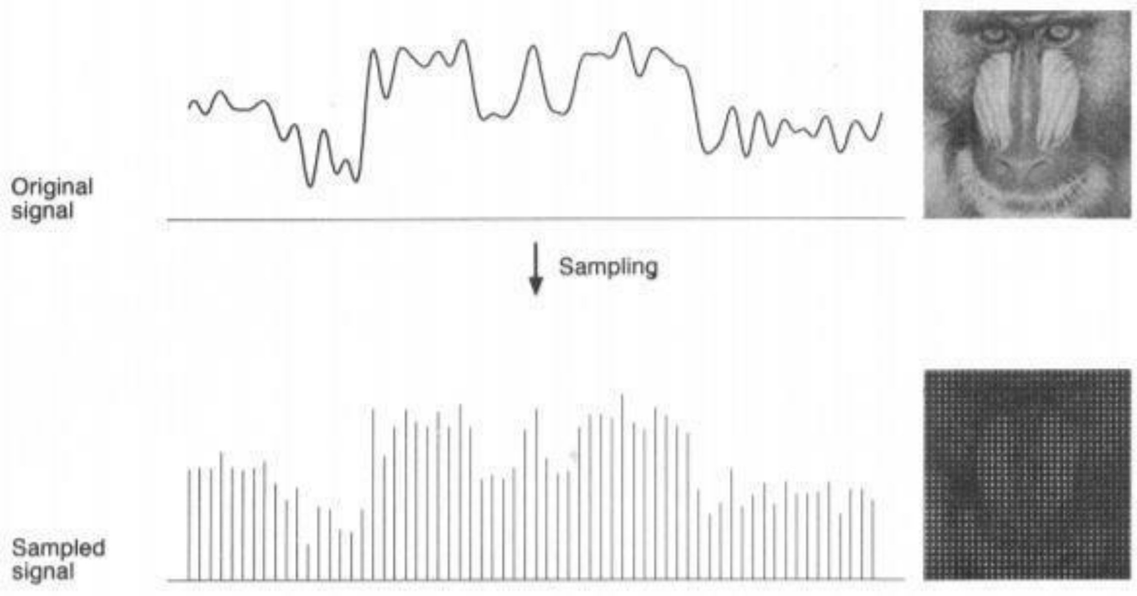


Image after sampling and Quantization

Sampling has a relationship with image pixels. The total number of pixels in an image can be calculated as Pixels = total no of rows * total no of columns.

For example, let's say we have total of 36 pixels, that means we have a square image of 6X 6. As we know in sampling, that more samples eventually result in more pixels. So it means that of our continuous signal, we have taken 36 samples on x axis. That refers to 36 pixels of this image. Also the number sample is directly equal to the number of sensors on CCD array.

Here is an example for image sampling and how it can be represented using a graph.



How Sampling Affects Image Resolution

The sampling rate (also called spatial resolution) dictates the level of detail captured in the digital image. An image sampled at a high rate contains more pixels and thus more detail, while low sampling rates lead to a coarse, pixelated look. For instance, high-definition images have a higher sampling rate than standard-definition images, which is why they appear clearer and more detailed.

Example: A 1024×1024 image has a high sampling rate with a total of 1,048,576 pixels, which provides a fine level of detail. A lower-resolution image, like 256×256, only contains 65,536 pixels, so it captures less detail and may appear pixelated when zoomed in.

Common Sampling Techniques

- Uniform Sampling: The most common technique where sampling points are evenly spaced, creating a grid-like pattern.
- Non-Uniform Sampling: Sampling points are spaced irregularly, sometimes used in specific applications like radar or medical imaging to capture particular details with more precision.

What is Quantization in Digital Image Processing?

Quantization is the process of mapping the continuous range of pixel intensity values into discrete levels, converting real-world light and color values into digital values that a computer can process. This step is necessary because, unlike analog images, digital images can only store a limited number of intensity or color values for each pixel. Quantization reduces the infinite range of possible values into a finite set, allowing for easier storage and processing.

In grayscale images, quantization maps brightness levels to a specific number of shades, typically represented by bits (e.g., an 8-bit grayscale image has 256 levels of brightness). For color images, quantization divides the color space into discrete intervals across the RGB channels.

How Quantization Affects Image Quality

The bit depth of an image—measured in bits per pixel—determines the number of possible intensity values per pixel. Higher bit depths mean more intensity levels and smoother transitions between colors or shades of gray, while lower bit depths can cause banding or posterization, where transitions appear abrupt.

Example: A 1-bit image can only represent black and white, while an 8-bit image can represent 256 shades of gray. For color images, 24-bit color (8 bits for each RGB channel) provides over 16 million possible colors, resulting in rich and detailed color representation.

The Relationship Between Sampling and Quantization

In digital image processing, sampling and quantization work together to digitize an image accurately. Sampling defines spatial resolution (the amount of detail in the image), while quantization defines color or intensity resolution (the number of distinct colors or shades that can be represented). Both must be balanced to achieve an image that is both visually accurate and computationally manageable.

For instance, high spatial resolution with low quantization levels might yield a highly detailed image with poor color quality, whereas high quantization with low sampling can produce vibrant colors in a low-detail image.

Key Effects of Sampling and Quantization on Digital Images

1. Image Resolution and Clarity
   Higher sampling rates increase spatial resolution, resulting in clearer, more detailed images. High quantization levels improve color or grayscale detail, offering smoother transitions and richer colors.
2. File Size and Storage
   High sampling and quantization rates produce larger files. For example, a high-resolution image with 24-bit color depth will require significantly more storage than a lower-resolution, grayscale image of the same scene.
3. Processing Complexity
   More detailed images typically require more computational power for processing, so balancing sampling and quantization can help optimize performance in applications like real-time video processing or embedded systems.

Basic Relationships Between Pixels

Explanation:
Relationships describe **how pixels relate to their neighbors.**

Types:

- **Neighborhood:**
  - 4-neighbors (up, down, left, right)
  - 8-neighbors (including diagonals)

- **Adjacency:**
  - Pixels are adjacent if they share an edge or corner.

- **Connectivity:**
  - Path of adjacent pixels.

- **Distance Measures:**
  - Euclidean distanceCity-block distance

Example:

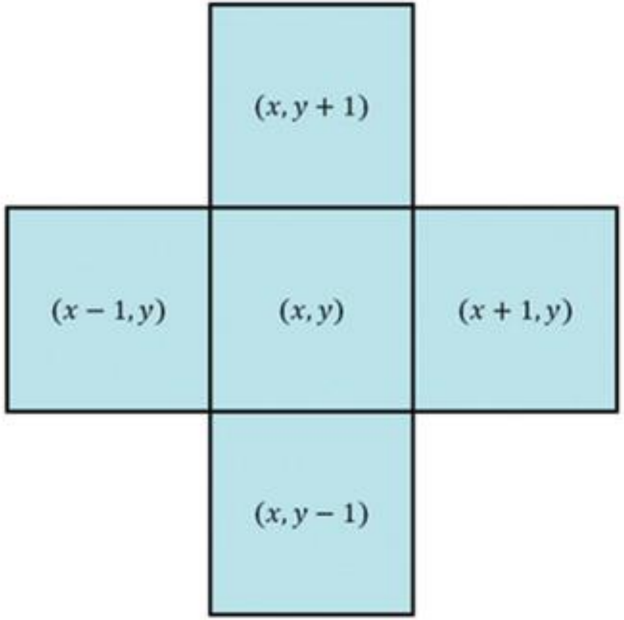
- To find connected regions in an image, you analyze adjacency and connectivity.

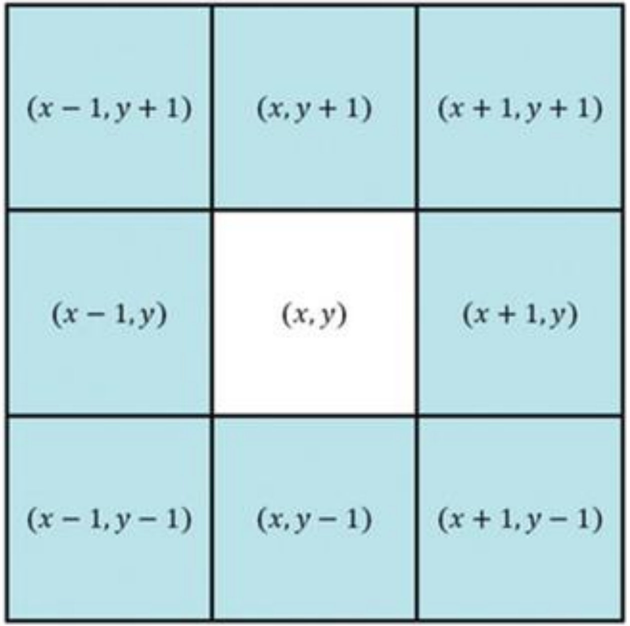
Real-Time Example:

- Flood-fill tool in paint programs uses connectivity to fill regions.

NEIGHBORS OF A PIXEL

- A pixel p at coordinates (x,y) has four horizontal and vertical neighbors having coordinates (x+1, y)(x, Y+1)(x-1, y)and (x, y-1).
- These are denoted by $N_4(p)$
- The four diagonal neighbors of p are (x+1, Y+1), (X-1,Y+1), (X+1, Y-1) and (X-1, Y-1)
- These are denoted by $N_D(p)$
- Together these two groups are called N8(p)
- For Boundary pixels number of Neighboring pixels will be less


$(x, y+1)$
$(x-1, y)$
(x, y)
$(x+1, y)$
$(x, y-1)$


$(x-1, y+1)$
$(x, y+1)$
$(x+1, y+1)$
$(x-1, y)$
(x, y)
$(x+1, y)$
$(x-1, y-1)$
$(x, y-1)$
$(x+1, y-1)$

Applications of Digital Image Processing

Overview:
DIP has **wide-ranging applications across domains.**

Key Applications:

- **Medical Imaging:**
  - MRI, CT scan enhancement and analysis.
- **Remote Sensing:**
  - Satellite imagery for agriculture and disaster management.
- **Industrial Inspection:**
  - Detecting defects in manufactured products.
- **Biometrics:**
  - Face and fingerprint recognition.
- **Document Processing:**
  - OCR (Optical Character Recognition).
- **Consumer Electronics:**
  - Digital cameras, smartphones.

Simple Real-Time Examples:

- **Face Unlock:** Your phone analyzes your face image.
- **Barcode Scanning:** Uses image processing to decode information.
- **Traffic Cameras:** Detect license plates automatically.

Thank You

https://cuitutorial.com/basic-relationships-between-pixels/#google_vignette

https://www.cse.iitm.ac.in/~sdas/courses/CV_DIP/PDF/NEIGH_CONN.pdf

https://images.search.yahoo.com/search/images?p=Basic+Relationships+Between+Pixels&fr=mcafee&type=E210US1274G0&imgurl=https%3A%2F%2Fimage1.slideserve.com%2F2965526%2Fbasic-relationships-between-pixels4-l.jpg#id=3&iurl=https%3A%2F%2Fslideplayer.com%2Fslide%2F13768544%2F85%2Fimages%2F29%2FBasic%2BRelationships%2BBetween%2BPixels.jpg&action=click