1 Introduction to Computer Vision

1.1 What is Computer Vision?

Computer Vision is a field of artificial intelligence that enables computers to interpret and make decisions based on visual data, such as images and videos. It seeks to automate tasks that the human visual system can perform.

1.2 How is Computer Vision used Today?

Computer Vision is applied in various industries, such as:

- **Healthcare**: Medical image analysis and diagnostics.
- **Automotive**: Autonomous vehicles and driver-assist systems.
- **Retail**: Inventory management and customer analytics.
- **Security**: Facial recognition and surveillance.
- **Agriculture**: Crop monitoring and precision farming.

1.3 Why is the .jpeg Format Popular?

The JPEG format is popular because:

- It uses lossy compression to significantly reduce file size while maintaining acceptable image quality.
- It is widely supported across devices and platforms.
- It is efficient for web use, enabling faster image loading.

1.4 Example of Lossless Compression Technique

PNG (Portable Network Graphics) is an example of lossless compression. It preserves the exact image data, ensuring no loss in quality.

1.5 Challenges of Computer Vision

- Understanding diverse environments and contexts.
- Handling occlusions and distortions in images.
- Computational resource requirements.
- Interpreting subjective aspects, such as emotion or intent in images.

2 Image Processing: Introduction and Fundamentals

2.1 Basic Relationship Between Pixels

- **Neighbor**: Pixels directly adjacent to a given pixel.
- **Adjacency**: Two pixels are adjacent if they share a common edge or vertex.

Comfort Fun Access

History of Computer Vision


Marvin Minsky, MIT
Turing award, 1969

"In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer:

connect a camera to a computer and get the machine to describe what it sees."

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

Half a century later
we're still working on it

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

History of Computer Vision

Style transfer - In CV, "style transfer" ~~means~~ refers to a technique that takes the content of one image & applies the artistic style of another image to it, generating a new image that retains the original content, while adopting the visual characteristics.

-It allows us to "paint" an img in the style of a diff. image.


Marvin Minsky, MIT
Turing award, 1969

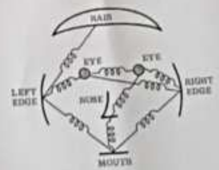

Gerald Sussman, MIT
AI Researcher since 1964

1960's: interpretation of synthetic worlds


Larry Roberts
"Father of Computer Vision"

Larry Roberts PhD Thesis, MIT, 1963, → Massachusetts Institute of Technology
Machine Perception of Three-Dimensional Solids

1970's: some progress on interpreting selected images



The representation and matching of pictorial structures
Fischler and Elschlager, 1973

what is png in the...

In the context of lossless compression, "PNG" stands for "Portable Network Graphics," which is a file format that utilizes lossless compression techniques, meaning that when a PNG image is compressed, no data is lost and the original image can be fully reconstructed upon decompression; making it ideal for preserving image quality in applications like logos, graphics, and detailed images where pixel precision is crucial.

Key points about PNG and lossless compression:

No data loss:

Unlike lossy compression formats like JPEG, PNG ensures that no image information is discarded during compression, resulting in perfect image fidelity.

Wide color support:

PNG can handle a wide range of colors, including transparency, making it suitable for various graphic design needs.

Compression algorithm:

PNG uses a compression algorithm called "Deflate" which effectively reduces file size while maintaining image quality.

Lossless Compression: A Complete Guide -...
Is PNG lossless? Yes, PNG - which

PNG - Wi
PNG (Portab
(PNG, officia

3D imaging - It refers to the process of capturing & analyzing data from 3-dimen-tional space, allowing for the creation of detailed, volumetric rep. of objs & scenes. Unlike 2D imaging, that offers flat, linear views limited to length & width, 3D imaging adds depth, providing a comprehensive rep. in 3-dim (x, y, z): length, width, height.

Features / key points

i) It provides depth with length & width.
ii) to capture a 3d info, multiple imgs of an obj are taken from diff. angles, that create a 3d model all together.
iii) Application - medical imaging (MRI, CT etc), product design etc.

- MRI → magnetic resonance imaging
- CT → computed tomography

1.4 Example of Lossless Compression Technique

PNG (Portable Network Graphics) is an example of lossless compression. It preserves the exact image data, ensuring no loss in quality.

1.5 Challenges of Computer Vision

- Understanding diverse environments and contexts.
- Handling occlusions and distortions in images.
- Computational resource requirements.
- Interpreting subjective aspects, such as emotion or intent in images.

2 Image Processing: Introduction and Fundamentals

2.1 Basic Relationship Between Pixels

- **Neighbors**: Pixels directly adjacent to a given pixel.
- **Adjacency**: Two pixels are adjacent if they share a common edge or vertex.

2.2 Path, Foreground, and Background

- **Path**: A sequence of adjacent pixels.
- **Foreground**: The primary objects of interest in an image.
- **Background**: The surrounding area of the foreground.
- **Boundary**: The edge separating the foreground from the background.

2.3 Distance Measures

- **Euclidean Distance**: Straight-line distance between two points. D_e
- **City Block Distance**: Distance measured along grid lines (Manhattan distance). D_4
- **Chessboard Distance**: Maximum of horizontal and vertical distances. D_8

3 Digital Image: Definition and Representation

A digital image is a numerical representation of a 2D visual pattern. It consists of pixels, each represented by a numeric value corresponding to intensity or color.

(i) add mat
(ii) add diagram
(iii) add that elm (often mat)

4 Electromagnetic Spectrum

Images are formed by capturing specific ranges of the electromagnetic spectrum, including visible light, infrared, and ultraviolet.

9:31

challenges of com...

Computer vision has several challenges, including:

- **Data quality:** The quality of the data used to train computer vision systems is important to ensure accurate results. Data can be inaccurate, missing, or unbalanced, which can lead to bias.
- **Data leakage:** Including inappropriate information in the training data can compromise the model's performance.
- **Lighting conditions:** Changes in lighting can make it difficult for the system to recognize objects. A model trained on high-resolution images might not recognize objects in low-light conditions.
- **Occlusion:** Objects partially blocking the view can impact a model's performance.
- **Labeling data:** Labeling data can be time-consuming.
- **Scaling:** Computer vision systems need to be able to scale.
- **Privacy and ethics:** Computer vision raises privacy and ethical issues.
- **Cost:** Computer vision can be expensive.
- **Experienced professionals:** There may be a lack of experienced professionals in the field.
- **Regular monitoring:** Computer vision systems may need to be monitored regularly.

5 Common Problems with Computer Vision...

Top Comp Opportur

Introduction

Digital Image

is a two-dimensional function

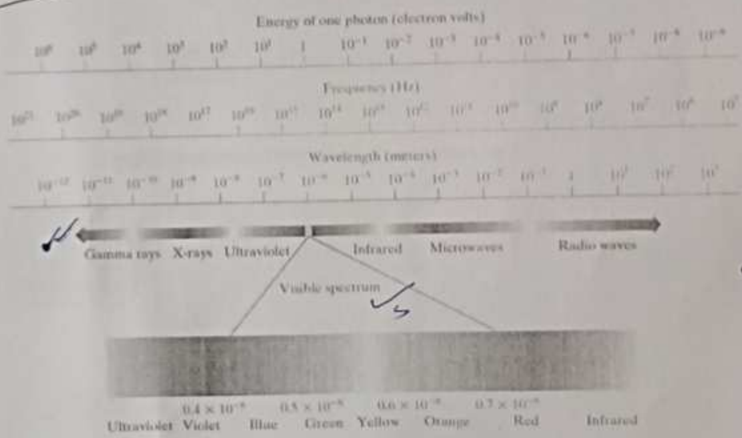
$$f(x, y)$$

where x and y are spatial coordinates, and

amplitude of f is called intensity or gray level at the point (x, y)

Light and EM Spectrum

wide range spectrum that consists uv rays, visible lights, x-rays, infrared etc



expanded

Light and EM Spectrum

eye can ~~from~~ in

The colors that humans perceive in an object are determined by the nature of the light reflected from the object.

e.g. green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths

wavelengths

..Light and EM Spectrum

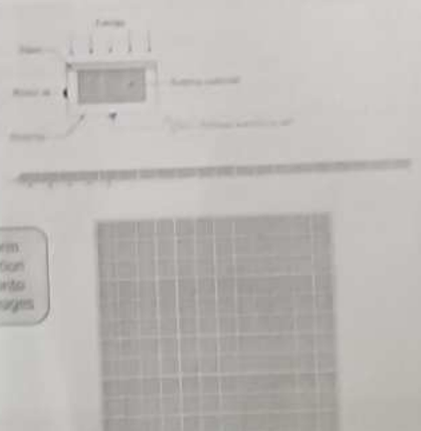
img that has only 2 colors b&w & it's combinations

gray scale imgs | b/w imgs

- Monochromatic light (void of color
  - **Intensity** is the only attribute, from black to white.
  - Monochromatic images are referred to as **gray-scale** images)
- Chromatic light bands: 0.43 to 0.79 μm
  - The quality of a chromatic light source:
  - **Radiance**: total amount of energy
  - **Luminance (lm)**: the amount of energy an observer perceives from a light source
  - **Brightness** (a subjective descriptor of light perception that is impossible to measure.) It embodies the achromatic notion of intensity and one of the key factors in describing color sensation.

the amount of energy reflected by a src

Image Acquisition



a
b
c
FIGURE 2.12
(a) Single imaging sensor
(b) Line sensor
(c) Array sensor

Transform illumination energy into digital images

Image Acquisition Using a Single Sensor

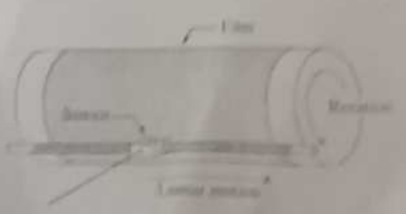


FIGURE 2.13
Combining a single sensor with motion to generate a 2-D image.

3 Digital Image: Definition and Representation

A digital image is a numerical representation of a 2D visual pattern. It consists of pixels, each represented by a numeric value corresponding to intensity or color.

4 Electromagnetic Spectrum

Images are formed by capturing specific ranges of the electromagnetic spectrum, including visible light, infrared, and ultraviolet.

5 Image Acquisition Techniques

5.1 Single Sensor

Uses a single photosensitive element, often with a rotating mirror or a moving sensor.

5.2 Sensor Strips

Linear arrays of sensors used in scanners.

5.3 2D Array of Sensors

Used in cameras to capture images in one shot. CMOS = Complementary metal oxide semiconductor

6 Sample Image Formation Model

The formation of an image involves:

- **Illumination**: Light falling on an object.
- **Reflectance**: Light reflected by the object.

The image intensity is a function of both illumination and reflectance.

7 Image Sampling and Quantization

Sampling converts a continuous signal (img) into discrete ~~values~~ pixels by selecting intervals, while **Quantization** maps these ~~values~~ intensity vals of pixels to finite levels.

8 Smoothing Spatial Filters

- **Linear Filters**: Use weighted averages for smoothing.
- **Average Filter Masks**: Assign equal weights to neighboring pixels. X

9 Order Statistic (Non-linear) Filters

Filters such as the **Median Filter** reorder pixel values to remove noise without blurring edges.

10 Representation of Objects

Objects are represented in terms of boundaries, regions, and key points within

Point detection) this are specific, salient point of interest in a img that represent human body joints, facial landmarks etc

Image Acquisition Using a Single Sensor

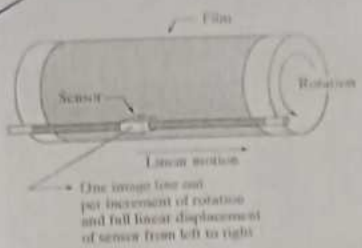


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Image Acquisition Using Sensor Strips

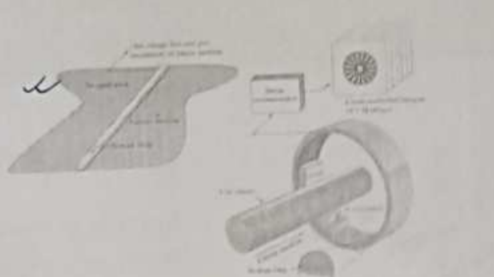


FIGURE 2.14

Image Acquisition Process

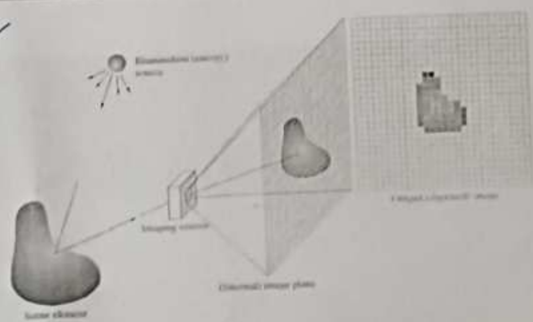


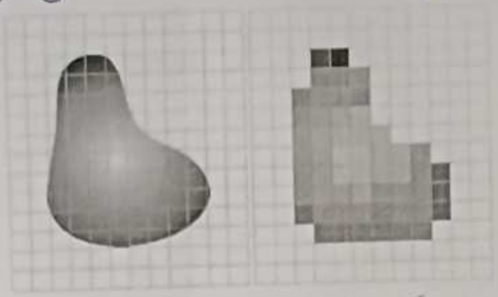
FIGURE 2.15

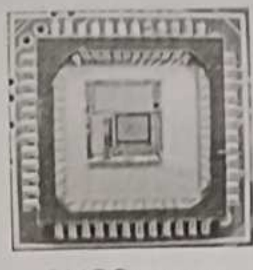
Sensor Array

10-11/30

Sensor Array

→ Add this to digital image's diagram





CMOS sensor

FIGURE 2.17 [illegible] = digitized image or digital image

A Simple Image Formation Model

$$f(x,y) = i(x,y) \cdot r(x,y)$$

$f(x,y)$: intensity at the point (x,y)

$i(x,y)$: illumination at the point (x,y)

(the amount of source illumination incident on the scene) (the amt of source (light) falling on the scene)

$r(x,y)$: reflectance/transmissivity at the point (x,y)

(the amount of illumination reflected/transmitted by the object)

where $0 < i(x,y) < \infty$ and $0 < r(x,y) < 1$

Some Typical Ranges of illumination

- Illumination
  - Lumen — A unit of light flow or luminous flux
  - Lumen per square meter (lm/m^2) — The metric unit of measure to illuminance of a surface
  - On a clear day, the sun may produce in excess of 90,000 lm/m^2 of illumination on the surface of the Earth
  - On a cloudy day, the sun may produce less than 10,000 lm/m^2 of illumination on the surface of the Earth
  - On a clear evening, the moon yields about 0.1 lm/m^2 of illumination
  - The typical illumination level in a commercial office is about 1000 lm/m^2

[illegible]

i) Image Acquisition using Single Sensor: Figure shows the components of Single (Light) Sensor eg: Photodiode, the use of filter in front of a sensor improves selectivity [illegible]

In order to generate a 2-D image using single sensor there has to be relative displacement in both x and y directions between the sensor & area to be imaged. 11

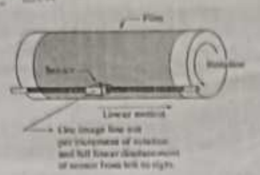
The figure shows the arrangement [illegible] as shown in below figure.

The film negative (image to be scanned) is mounted on a drum whose rotation makes one dimensional displacement.

The single sensor is mounted on a lead screw which provides displacement in another direction.

Such arrangement provides image capturing but is slow. This is a slow and inexpensive method of obtaining high-resolution images.

[illegible]



ii) Image Acquisition using Sensor Strip: This is consists of an in-line sensor arrangement of in the form of a sensor strip as shown in fig (b).

The strip provides imaging in one direction and its motion

in perpendicular direction provides imaging in other direction. 12

Figure below shows the arrangement of sensor strips which is mostly used in flat bed scanners.

Sensing devices with [illegible] in-line sensors are [illegible] used.

It gives one line of image at a time and the [illegible] motion makes 2-D scanning.

Sensor strips mounted in a ring configuration are frequently used in medical applications which gives cross sectional images of 3-D objects as shown below.

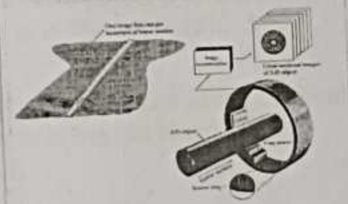


Fig: Image Acquisition using line sensor

iii) Image Acquisition using Sensor Arrays: Fig (c) shows individual sensors are arranged in 2-D array. 13

This is also the predominant arrangement found in digital cameras, electromagnetic & ultrasonic sensing devices.

A typical sensor for these cameras arrangement is CCD array that contains a packaged rugged array of [illegible] elements & more (CCD: Charged Coupled Devices).

The principal manner in which array sensors are used is shown below.

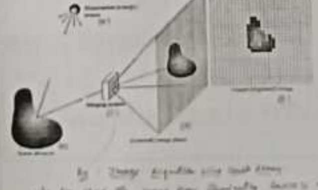


fig: Image Acquisition using sensor array

→ The fig shows the energy from illuminating source is reflected by scene element.

→ The imaging system collects the incoming energy from scene element and focus it on to an image plane.

→ The sensor array (which is coincident with the focal plane) produces output proportional to the integral of light received at each sensor.

→ The digitizer [illegible] output [illegible] to digital form results in a digital image (shown in e).

• Image Sampling and Quantization: The o/p of most sensors is a continuous voltage waveform whose amplitude and spatial behaviour are related to the physical phenomenon being sensed. 14

To create a digital image, we need to convert the continuous sensed data into digital form.

This involves two processes: i) Sampling

ii) Quantization

→ Digitizing the coordinate values is called Sampling.

→ Digitizing the amplitude (intensity) values is called quantization.

Basic Relationships Between Pixels

- **Neighbors** of a pixel p at coordinates (x,y)

  - **4-neighbors of p**, denoted by $\mathbf{N_4(p)}$:
    (x-1, y), (x+1, y), (x,y-1), and (x, y+1).

  - **4 diagonal neighbors of p**, denoted by $\mathbf{N_D(p)}$:
    (x-1, y-1), (x+1, y+1), (x+1,y-1), and (x-1, y+1).

  - **8 neighbors of p**, denoted $\mathbf{N_8(p)}$
    $N_8(p) = N_4(p) \cup N_D(p)$

Basic Relationships Between Pixels

- **Adjacency**
  Let V be the set of intensity values

  - **4-adjacency**: Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.

  - **8-adjacency**: Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.

Basic Relationships Between Pixels

- **Adjacency**
  Let V be the set of intensity values

  - **m-adjacency**: Two pixels p and q with values from V are m-adjacent if

    (i) q is in the set $N_4(p)$, or

    (ii) q is in the set $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

a large degree by no. of samples and discrete gray levels used in sampling & quantization.

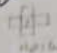
* Some basic Relationship between Pixels:

Neighbors of a Pixel:

• A pixel 'p' at coordinate (x,y) has four horizontal and vertical neighbors whose coordinates are given by (x+1,y), (x-1,y), (x,y+1), (x,y-1). This set of pixels called the 4-neighbors of 'p' and is denoted by $N_4(P)$.

• Each pixel is a unit distance from (x,y).

• Some neighbors of p lie outside the digital image if (x,y) is on the border of the image.



$N_4(P)$ set

• The four diagonal neighbors of 'p' have coordinates (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1) and are denoted by $N_D(P)$. 17

• These points, together with the 4-neighbors are called the 8-neighbors of 'P', and is denoted by $N_8(P)$.

$N_D(P)$ $N_8(P)$ Pixel coordinates

• Some of the points in $N_D(P)$ & $N_8(P)$ fall outside the image if (x,y) is on the border of image.

Adjacency, Connectivity, Regions and Boundaries:

• Connectivity b/w pixels is a fundamental concept that simplifies the definitions of such as regions and boundaries.

• Two neighbor pixels satisfy a specified criterion of similarity i.e. their gray levels (equal) then they are said to be "Connected".

• Let 'V' be the set of gray-level values used to define adjacency.

• In a binary image, V = {1} if we are referring to adjacency of pixels with value 1.

• Same idea to the grayscale image, but here 'V' contains more elements in b/w 0 – 255.

Three types of adjacency:

a) 4-adjacency: 2 pixels 'P' and 'q' with values from 'V' are 4-adjacent if 'q' is in the set $N_4(P)$.

18/46

(b) 8-adjacency: 2 pixels 'P' and 'q' with values from 'V' are 8-adjacent if 'q' is in the set $N_8(P)$.

(c) m-adjacency (mixed): Two pixels 'P' & 'q' are with values from 'V' are m-adjacent if 18

i) 'q' is in $N_4(P)$, or

ii) 'q' is in $N_D(P)$ & the set $N_4(P) \cap N_4(q)$ has no pixel whose values are from 'V'.

• Mixed adjacency is a modification of 8-adjacency which is introduced to eliminate ambiguities that often arise with 8-adjacency.

• Consider the following pixel arrangement to show the ambiguity of 8-adjacency and its elimination is m-adjacency:

0	1	1		0	1	1		0	1	1
0	1	0		0	1	0		0	1	0
0	0	1		0	0	1		0	0	1

Binary Image Pixel arrangement — 8-adjacency — m-adjacency

• Let 'R' be a subset of pixels in an image.

• If 'R' is a connected set then 'R' is said ...

18/46

• If one or more neighbors in the set 'R' are not connected, then it is said to be "boundary" of the region 'R'.

Distance Measures: For pixels P, q & z, with coordinates (x,y), (s,t), (v,w), respectively 'D' is a distance function or metric if

a) $D(P,q) \geq 0$ $(D(P,q) = 0$ iff $P = q)$,

b) $D(P,q) = D(q,P)$ &

c) $D(P,z) \leq D(P,q) + D(q,z)$

Cont

• The "Euclidean Distance" between P & q.

$$D_e(P,q) = \left[(x-s)^2 + (y-t)^2\right]^{1/2}$$ 19

• D_4 distance (city block distance) between P & q.

$$D_4(P,q) = |x-s| + |y-t|$$

Basic Relationships Between Pixels

- **Path**
  - A (digital) path (or curve) from pixel p with coordinates (x_0, y_0) to pixel q with coordinates (x_n, y_n) is a sequence of distinct pixels with coordinates

    4 adjacent

    $$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

    where (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$.
  - Here n is the *length* of the path.
  - If $(x_0, y_0) = (x_n, y_n)$, the path is ***closed*** path.
  - We can define 4-, 8-, and m-paths based on the type of adjacency used.

Examples: Adjacency and Path

$V = \{1, 2\}$

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Examples: Adjacency and Path

$V = \{1, 2\}$

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}$$

8-adjacent

Examples: Adjacency and Path

$V = \{1, 2\}$

0	1	1
0	2	0
0	0	1

0	1	1
0	2	0
0	0	1

8-adjacent

0	1	1
0	2	0
0	0	1

m-adjacent

Examples: Adjacency and Path

$V = \{1, 2\}$

0	1	1
0	2	0
0	0	1

0	1	1
0	2	0
0	0	1

8-adjacent

0	1	1
0	2	0
0	0	1

m-adjacent

The 8-path from (1,3) to (3,3):
(i) (1,3), (1,2), (2,2), (3,3)
(ii) (1,3), (2,2), (3,3)

The m-path from (1,3) to (3,3):
(1,3), (1,2), (2,2), (3,3)

Basic Relationships Between Pixels

- **Connected in S**

  Let S represent a subset of pixels in an image. Two pixels p with coordinates (x_0, y_0) and q with coordinates (x_n, y_n) are said to be **connected in S** if there exists a path

  $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$

  where $\forall i, 0 \le i \le n, (x_i, y_i) \in S$

Basic Relationships Between Pixels

Let S represent a subset of pixels in an image

- For every pixel p in S, the set of pixels in S that are connected to p is called a ***connected component*** of S.
- If S has only one connected component, then S is called ***Connected Set.***
- We call R a **region** of the image if R is a connected set
- Two regions, R_i and R_j are said to be ***adjacent*** if their union forms a connected set.
- Regions that are ~~not to be~~ adjacent are said to be ***disjoint.***

Basic Relationships Between Pixels

- Boundary (or border)
  - The ***boundary*** of the region R is the set of pixels in the region that have one or more neighbors that are not in R.
  - If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns of the image.
- Foreground and background
  - An image contains K disjoint regions, R_k, $k = 1, 2, \ldots, K$. Let R_u denote the union of all the K regions, and let $(R_u)^c$ denote its complement.

    All the points in R_u is called **foreground**;

    All the points in $(R_u)^c$ is called **background**.

Question 1

- **In the following arrangement of pixels, are the two regions (of 1s) adjacent? (if 8-adjacency is used)**

1	1	1
1	0	1
0	1	0
0	0	1
1	1	1
1	1	1

Region 1

Region 2

Question 2

- In the following arrangement of pixels, the two regions (of 1s) are disjoint (if 4-adjacency is used)

1	1	1
1	0	1
0	1	0
0	0	1
1	1	1
1	1	1

Region 1

Region 2

- In the following arrangement of pixels, the two regions (of 1s) are disjoint (if 4-adjacency is used)

1	1	1
1	0	1
0	1	0
0	0	1
1	1	1
1	1	1

foreground

background

Question 3

- In the following arrangement of pixels, the circled point is part of the boundary of the 1-valued pixels if 8-adjacency is used, true or false?

0	0	0	0	0
0	1	1	0	0
0	1	1	0	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0

Question 4

• ... R is said to be Region.

• If one or more neighbors in the set 'R' are not connected, then it is said to be "boundary" of the Region 'R'.

Distance Measures:- For pixels P, q & z, with coordinates (x, y), (s, t), (v, w), respectively, 'D' is a distance function or metric if

(a) $D(P,q) \geq 0$ $\quad (D(P,q) = 0$ iff $P = q)$,

(b) $D(P,q) = D(q,P)$ &

(c) $D(P,z) \leq D(P,q) + D(q,z)$.

Cont...

• The "Euclidean Distance" between P & q,

$$D_e(P,q) = \left[(x-s)^2 + (y-t)^2\right]^{1/2}$$

19

• D_4 distance (city-block distance) between P & q,

$$D_4(P,q) = |x-s| + |y-t|.$$

$$\begin{matrix} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{matrix}$$

Eg:- $D_4(P,q) \leq 2$

• D_8 distance (chess board Distance) b/w P & q,

$$D_8(P,q) = \max(|x-s|, |y-t|)$$

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

Eg:- $D_8(P,q) \leq 2$

——— o ———

(10)

* **Introduction to the mathematical tools used in DIP**

20

I. Array versus Matrix Operations:-

• Consider two 2x2 images: $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ & $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$.

• Array Product : $\begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$

• Matrix Product : $\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix}$

• Array operation involving one or more images is carried out on a pixel-by-pixel basis.

II. Linear versus Non-Linear Operations

Distance Measures

The following are the different Distance measures:

a. Euclidean Distance :
$D_e(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}$

b. City Block Distance:
$D_4(p, q) = |x-s| + |y-t|$

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

c. Chess Board Distance:
$D_8(p, q) = \max(|x-s|, |y-t|)$

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

5

Question 5

- In the following arrangement of pixels, what's the value of the chessboard distance between the circled two points?

	1	2	3	4	5
1	0	0	0	0	0
2	0	0	1	1	0
3	0	1	1	0	0
4	0	1	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0

$\Rightarrow 2$

(can move diagonaly)

22-23/41

5

Question 6

- In the following arrangement of pixels, what's the value of the city-block distance between the circled two points?

0	0	0	0	0
0	0	1	1	0
0	1	1	0	0
0	1	0	0	0
0	0	0	0	0
0	0	0	0	0

(can't move diagonaly)

4

Example: Addition of Noisy Images for Noise Reduction

Noiseless image: f(x,y)

Noise: n(x,y) (at every pair of coordinates (x,y), the noise is uncorrelated and has zero average value)

Corrupted image: g(x,y)

or noisy g(x,y) = f(x,y) + n(x,y)

Reducing the noise by adding a set of noisy images, $\{g_i(x,y)\}$

$$\overline{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

Example: Addition of Noisy Images for Noise Reduction

$$\overline{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

$$E\{\overline{g}(x,y)\} = E\left\{\frac{1}{K}\sum_{i=1}^{K} g_i(x,y)\right\}$$

$$= E\left\{\frac{1}{K}\sum_{i=1}^{K}[f(x,y)+n_i(x,y)]\right\}$$

$$= f(x,y) + E\left\{\frac{1}{K}\sum_{i=1}^{K} n_i(x,y)\right\}$$

$$= f(x,y)$$

Example: Addition of Noisy Images for Noise Reduction

- In astronomy, imaging under very low light levels frequently causes sensor noise to render single images virtually useless for analysis.

- In astronomical observations, similar sensors are used for noise reduction by observing the same scene over long periods of time. Image averaging is then used to reduce the noise.

FIGURE 2.26 (a) Image of Galaxy Pair NGC 3314 [illegible] by additive Gaussian noise. (b)-(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)

An Example of Image Subtraction: Mask Mode Radiography

Mask h(x,y): an X-ray image of a region of a patient's body

Live images f(x,y): X-ray images after injection of the contrast medium

Enhanced detail g(x,y)

$$g(x,y) = f(x,y) - h(x,y)$$

The procedure gives a movie showing how the contrast medium propagates through the various arteries in the area being observed.

a b
c d

FIGURE 2.28 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

An Example of Image Multiplication

a b c

FIGURE 2.29 Shading correction. (a) Shaded SEM image of a tungsten filament and support, magnified approximately 130 times. (b) The shading pattern. (c) Product of (a) by the reciprocal of (b). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

mathematical operation in DIP

Linear v/s non linear operation

$$H[f(x,y)] = g(x,y)$$

$$H[a_i f_i(x,y) + a_j f_j(x,y)]$$

$$= H[a_i f_i(x,y)] + H[a_j f_j(x,y)] \rightarrow \text{Additivity}$$

$$= a_i H[f_i(x,y)] + a_j H[f_j(x,y)] \rightarrow \text{Homogeneity}$$

$$= a_i g_i(x,y) + a_j g_j(x,y)$$

H is said to be linear operator, if it meets above qualification.
H is said to be non-linear operator, if it doesn't meet the above qualification.

Set and Logical Operations

- Let A be the elements of a gray-scale image gray scale img
  The elements of A are triplets of the form (x, y, z), where x and y are spatial coordinates and z denotes the intensity at the point (x, y).

$$A = \{(x, y, z) \mid z = f(x, y)\}$$

- The complement of A is denoted A^c

$$A^c = \{(x, y, K - z) \mid (x, y, z) \in A\}$$

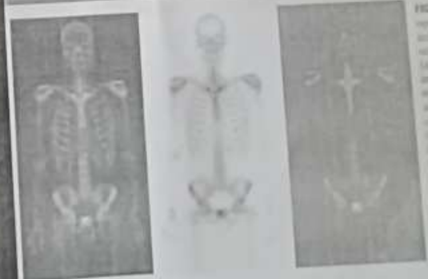
$K = 2^k - 1$; k is the number of intensity bits used to represent z

Set and Logical Operations

- The union of two gray-scale images (sets) A and B is defined as the set

$$A \cup B = \{\max_z(a, b) \mid a \in A, b \in B\}$$

Set and Logical Operations



Spatial Operations

- Single-pixel operations

Alter the values of an image's pixels based on the intensity.

$$s = T(z)$$

e.g.,

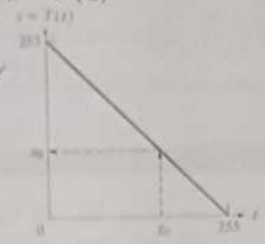
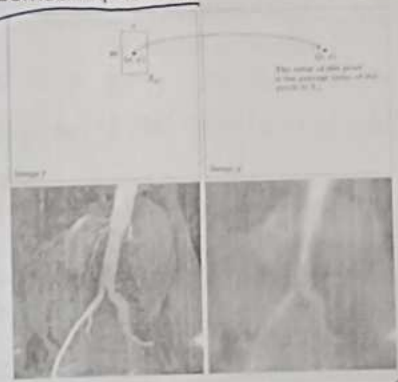


FIGURE 2.34 Intensity transformation function used to obtain the negative of an 8-bit image. The dashed arrows show transformation of an arbitrary input intensity value z_0 into its corresponding output value s_0.

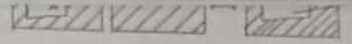
Spatial Operations

- Neighborhood operations



Here the pixel vals of Smoothed img is determined by the avg pixel vals of ordi -nal ~~shaped~~ Sharpened img

Thank You



Eg:- Logic Operations on Image:

—— o ——

(mine)

Need for Image Transform:-

"Transform" is basically a mathematical tool to represent a signal. The need for transforms is as follows:

i) Mathematical Convenience: (Every action in time domain will have an impact in frequency domain.

Convolution in Time Domain $\Longleftrightarrow$ Multiplication in frequency Domain)

↓ Complex ↓ Simple

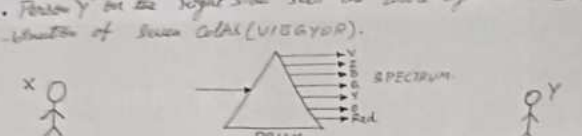
ii) To Extract more Information:- (It allow us to extract more relevant information.)

- To illustrate this, consider the example of Prism experiment. ①

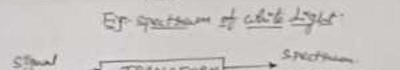
Cont...

For,
i) extracting data
ii) Analyzing the data for predictions & detection
iii) noise reduction
iv) Calculating various metrics from img
v) Doing various mathematical operations etc.
vi) Detection (edge, etc)
vii) sharpening (smoothing

- Person 'X' on the left side sees the light as white light;
- Person 'Y' on the right side sees the white light as a combination of seven colors (VIBGYOR). 24



Eg: spectrum of white light.



Eg: Concept of Transform.

- Here, 'Y' is getting more information than 'X', by using prism.
- Similarly the T/f tool allows us to extract more information.

iii) Other needs:-

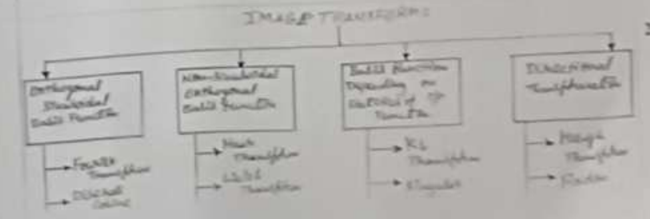
- The transformation (may isolate critical components of the image pattern, so that they are directly accessible to analysis)
- The transformation (may place the image data in a more compact form so that they can be stored & transmitted efficiently.)

—— o ——

> Classification of Image Transforms:- Image transforms can be classified on the nature of the "basis function".

i) Transforms with orthogonal basis function.
ii) Transforms with non-sinusoidal orthogonal basis function.
iii) Transforms whose basis functions depend on the statistics of the input data.
iv) Directional Transformation.

Cont...



25

Some Typical Ranges of Reflectance

- Reflectance
  - 0.01 for black velvet
  - 0.65 for stainless steel
  - 0.80 for flat-white wall paint
  - 0.90 for silver-plated metal
  - 0.93 for snow

Image Sampling and Quantization

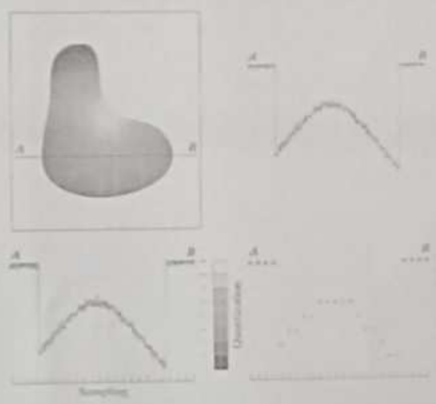
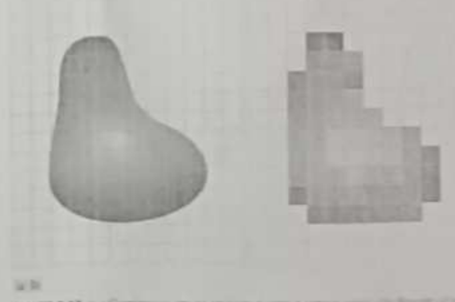


Image Sampling and Quantization



digital form known as a digital image (Raw is x').

* Image Sampling and Quantization:- The o/p of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. 14

To create a digital image, we need to convert the continuous sensed data into digital form.

This involves two processes: i) Sampling
ii) Quantization.

- Digitizing the coordinate values is called Sampling.
- Digitizing the amplitude (intensity) values is called quantization.

• Let us consider a continuous image, f(x, y) where x, y coordinates and amplitudes are continuous as shown in fig(a).

• To convert the image into digital form, we have to sample the function in both the amplitude & coordinates.

• As per the definitions of Sampling & quantization, we perform the operations on the image in fig(a).

• The 1-D function in fig(b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in fig(a).

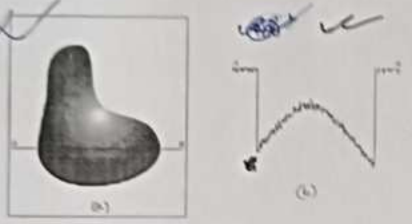


Fig:- Generation of Digital Image:

• To sample this 1 function, we take equally spaced samples along line AB as shown in fig(c). 15

• The location of each sample is given by a tick mark in bottom part of the figure.

• The samples are shown as small white squares superimposed on the function. The set of these discrete locations gives sampled function.

• The values of the samples are still a continuous gray level values.

• In order to form a digital function, the gray-level values must be converted (quantized) into discrete quantities.

• The right side of fig(c) shows eight discrete gray-levels ranging from black to white.

• These gray-levels are assigned to each corresponding sample to obtain the quantized value.

• The digital samples obtained in Sampling and quantization is represented in fig(d).

• Starting at the top of the image and carrying out this procedure line by line produces a digital image.

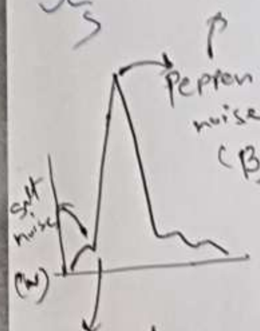
Fig:- Generation of Digital Image:

• When a sensing array is used for image acquisition, there is no motion & the no. of sensors in array establishes the limit of Sampling in both directions.

• The figure(2.1) shows continuous image projected onto the plane ⑤

P.T.O.

sharp / abnormal drop of the axis towards



sharp / abnormal raise of the axis towards of array sensor and fig(2.4) shows the image after sampling and quantization.

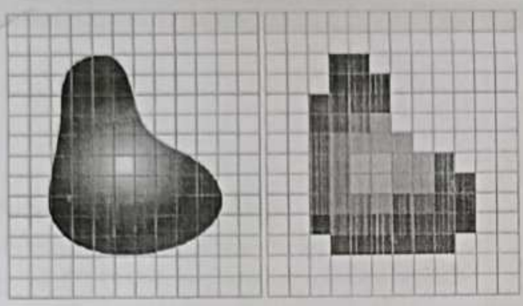


Fig 2:- (a) Continuous Image (b) Digital Image.

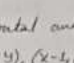
- Clearly, the quality of a digital image is determined to a large degree by no. of samples and discrete gray levels used in sampling & quantization.

---

* Some basic Relationships between Pixels:-

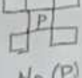
Neighbors of a Pixel:-

- A pixel 'p' at coordinates (x,y) has four horizontal and vertical neighbors whose coordinates are given by $(x+1,y)$, $(x-1,y)$, $(x,y+1)$, $(x,y-1)$. This set of pixels called the 4 neighbors of 'p' and is denoted by $N_4(P)$.
- Each pixel is a unit distance from (x,y).
- Some Neighbors of p lie outside the digital image if (x,y) is on the border of the image.

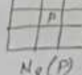


$N_4(P)$ Set.

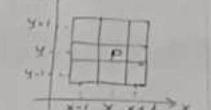
- The four diagonal neighbors of 'p' have coordinates $(x+1,y+1)$, $(x+1,y-1)$, $(x-1,y+1)$, $(x-1,y-1)$ and are denoted by $N_D(P)$.
- These points, together with the 4-neighbors, are called the 8-neighbors of 'p', and is denoted by $N_8(P)$.



$N_D(P)$



$N_8(P)$



Pixel Coordinates.

- Some of the points in $N_D(P)$ & $N_8(P)$ fall outside the image if (x,y) is on the border of image.

Adjacency, Connectivity, Regions and Boundaries:

- Connectivity b/w pixels is a fundamental concept the simplifies the definitions of such as regions and boundaries.
- Two neighbor pixels satisfy a specified criterion of similarity in their gray levels (equal) then they are said to be "Connected."

8 Smoothing Spatial Filters

- **Linear Filters**: Use weighted averages for smoothing.
- **Average Filter Masks**: Assign equal weights to neighboring pixels.

9 Order Statistic (Non-linear) Filters

Filters such as the **Median Filter** reorder pixel values to remove noise without blurring edges.

MASK == Filters

10 Representation of Objects

Objects are represented in terms of boundaries, regions, and key points within an image.

11 Median Filtering for Noise Reduction and Sharpening

Median Filtering reduces noise while preserving edges, useful for tasks such as sharpening. → use for salt-pepper noise reduction

12 Laplace Operator

A second-order derivative operator used for edge detection by identifying areas of rapid intensity change.

13 Unsharp Masking and High-Boost Filtering

These techniques enhance edges by subtracting a smoothed version of the image from the original.

14 Spatial Correlation and Convolution

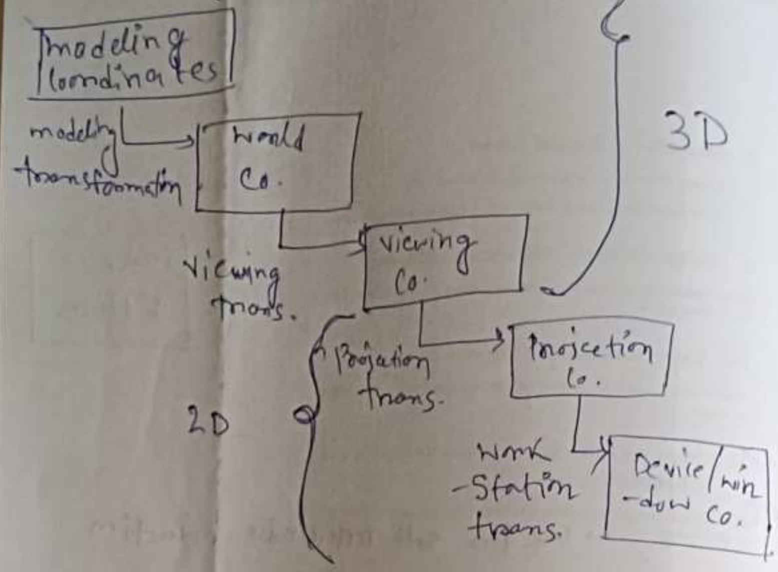
- **Spatial Correlation**: Measures similarity between patterns.

- **Convolution**: Applies a filter to an image, combining values in a local neighborhood.

15 Image Segmentation (Concept)

Image segmentation partitions an image into regions or objects based on properties like color, intensity, and texture.

3D viewing Process



Spatial Domain vs. Transform Domain

- Spatial domain is a domain that (image plane itself, directly process the intensity values of the image plane)
- Transform domain is a domain that (process the transform coefficients, not directly process the intensity values of the image plane)

5

Spatial Domain Process

$$g(x, y) = T[f(x, y)]$$

$f(x, y)$: input image

$g(x, y)$: output image

T : an operator on f, defined over a neighborhood of point (x, y)

5

Spatial Domain Process

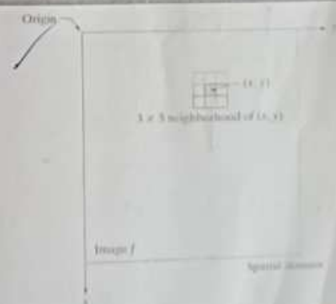


FIGURE 3.1 A 3 × 3 neighborhood about a point (x, y) in an image in the spatial domain. The neighborhood is moved from pixel to pixel in the image to generate an output image.

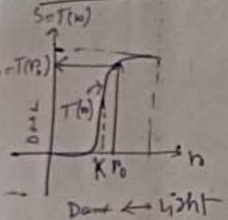
• Spatial domain Process

Intensity transformation function

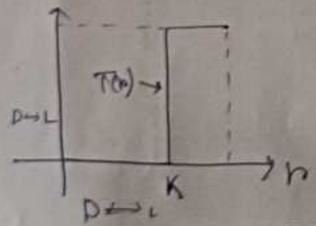
$S = T(r)$

- i) used to get the img
- ii) used to calculate the pixel val. of a smoothened img from the pixel vals of it's sharpened img
- iii) can control brightness & darkness
- iv) can invert the pixel vals.

a) Contrast stretching func



b) Thresholding func



~~(transformation domain)~~ Process

- Image negatives → $\boxed{S = L - 1 - r}$
- Log transformation → $\boxed{S = C\log(1+r)}$
- Power-Law (Gamma) Transfo. → $\boxed{S = Cr^{\gamma}}$ [C = 1, for all cases]

Range (0 to L-1)

Eg – CRT (cathode Ray Tube), have an intensity-to-voltage response that is a power function with exponent varying from approx 1.8 to 2.5

$\boxed{S = r^{1/2.5}}$

10:30

← 2b Spatial Tra...

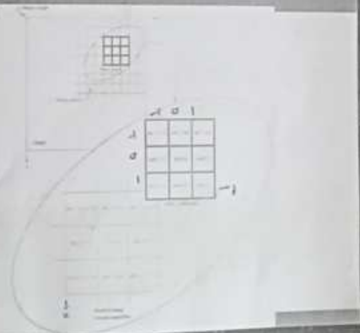
Spatial Filtering

A spatial filter consists of (a) a neighborhood, and (b) a **predefined operation**

Eg – Linear spatial filtering of an image of size MxN with a filter of size mxn is given by the expression

$$g(x,y) = \sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)\,f(x+s,y+t)$$

Spatial Filtering



Spatial Correlation

The correlation of a filter $w(x,y)$ of size $m \times n$ with an image $f(x,y)$, denoted as $w(x,y) ☆ f(x,y)$

$$w(x,y) ☆ f(x,y) = \sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)\,f(x+s,y+t)$$

- **use of log transformation**

i) Reduces skewness
ii) Stabilites variance
iii) Simplifies relationships
iv) Improves model fit
v) Compresses data range
vi) Handles exponential growth
vii) Enhances interpretability

- **------ gamma --**

i) Enhances img brightness
ii) Applies non-linear mapping
iii) Improves contrast
iv) Controlled by γ (gamma) values
v) Used in display systems.

- **--- image negative - -**

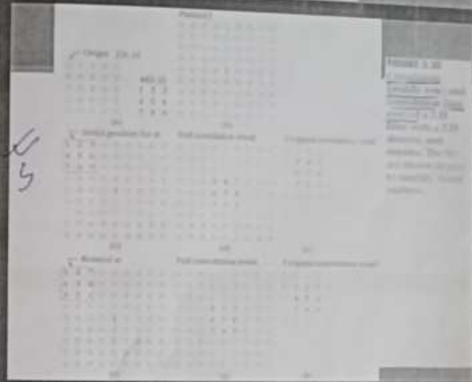
i) Inverts pixel values
ii) converts bright areas to dark & vice versa.
iii) useful for medical imaging (eg x-rays)

- **piece wise linear transformation**
- **Contrast stretching** - Expands the range of intensity levels in an image, so that it spans the full intensity range of the recording medium or display device.
- **Intensity-level slicing** - Highlighting a specific range of intensities in an image (often ROI of interest).

Spatial Convolution

The convolution of a filter $w(x, y)$ of size $m \times n$ with an image $f(x, y)$, denoted as $w(x, y) \star f(x, y)$

$$w(x, y) \star f(x, y) = \sum \sum w(s, t) f(x-s, y-t)$$



Smoothing Spatial Filters

Smoothing filters are used for blurring and for noise reduction

Blurring is used in removal of small details and bridging of small gaps in lines or curves

Smoothing spatial filters include linear filters and nonlinear filters.

Spatial Smoothing Linear Filters

• ~~Exm~~ Use

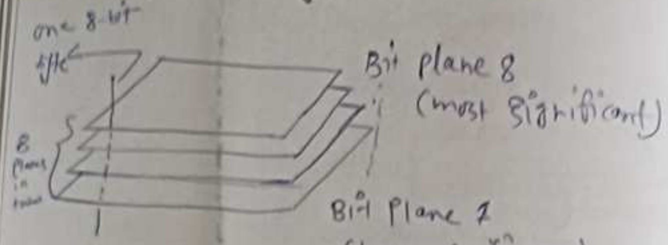
i) Divides the intensity range into segments

ii) Applies diff. linear transformations to each segment.

iii) Enhances specific intensity ranges.

iv) Controll the brightness of an image

• Bit-Plane Slicing



Bit-Plane rep. of an 8-bit img

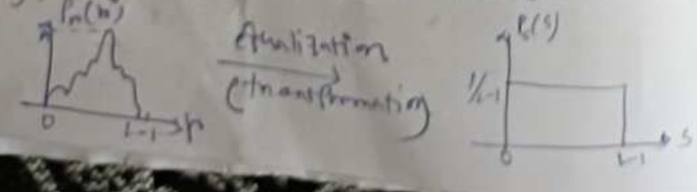
• Each bit plane is a binary image.

• Histogram Processing

Histogram $h(r_k) = n_k$
r_k = k^{th} intensity val.
n_k = No. of pixels in the img with intensity r_k
Normalized histogram $p(r_k) = \frac{n_k}{MN}$
n_k = the no. of pixels in the img of size $M \times N$ with intensity r_k.

• Histogram Equalization

The intensity levels in an img may be viewed as random variables in the interval $[0, L-1]$.
Let $P_r(r)$ & $P_s(s)$ denote the probability density function (PDF) of random variables r & s.



Smoothing Spatial Filters

Smoothing filters are used for blurring and for noise reduction

Blurring is used in removal of small details and bridging of small gaps in lines or curves

Smoothing spatial filters include linear filters and nonlinear filters.

Spatial Smoothing Linear Filters → use weighted avg for smoothing

The general implementation for filtering an M × N image with a weighted averaging filter of size m × n is given

$$g(x,y) = \frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t) f(x+s, y+t)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)}$$

where $m = 2a+1$, $n = 2b+1$.

eg

Two Smoothing Averaging Filter Masks

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Benifits of histogram equalization

i) Improves contrast
ii) Improves image quality.
iii) Improves image suitability
iv) Improves visibility
v) Improves image properties.

10:40

← 2b Spatial Tra...

Order-statistic (Nonlinear) Filters → It's a type of nonlinear filter that reorders pixel vals to remove noise without blurring. It remove noise while preserving edges, so it's very useful for sharpening.

- Nonlinear
- Based on ordering (ranking) the pixels contained in the filter mask
- Replacing the value of the center pixel with the value determined by the ranking result

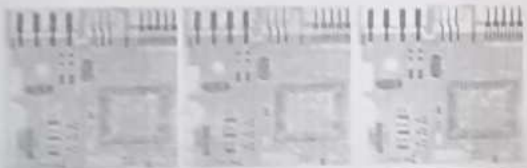
E.g., median filter, max filter, min filter

these → Img sharpening, non say noise reduction

effects

i) Edge preservation
ii) noise reduction with detail retention
iii) Robustness to outliers
iv) Finer details preservation

Example: Use of Median Filtering for Noise Reduction



a b c

FIGURE 3.35 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3 × 3 averaging mask. (c) Noise reduction with a 3 × 3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

→ median filter

Sharpening Spatial Filters

- Foundation
- Laplacian Operator
- Unsharp Masking and Highboost Filtering
- Using First-Order Derivatives for Nonlinear Image Sharpening — The Gradient

Sharpening Spatial Filters

- Foundation
- Laplacian Operator
- Unsharp Masking and Highboost Filtering
- Using First-Order Derivatives for Nonlinear Image Sharpening — The Gradient

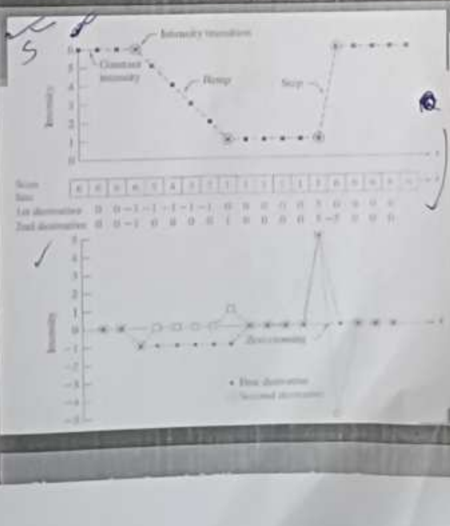
Sharpening Spatial Filters: Foundation

- The first-order derivative of a one-dimensional function f(x) is the difference

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

- The second-order derivative of f(x) as the difference

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$$



a
b
c

FIGURE 3.36 Illustration of the first and second derivatives of a 1-D digital function representing a section of a horizontal intensity profile from an image. In (a) and (c) data points are joined by dashed lines as a visualization aid.

10:41

what is median ,ma...

In computer vision, a "median filter" replaces a pixel value with the median value of its neighboring pixels within a specified window, effectively removing noise while preserving edges, while a "max filter" replaces a pixel with the highest value in its neighborhood, highlighting bright areas, and a "min filter" replaces a pixel with the lowest value, effectively darkening the image and highlighting dark areas; all three are considered "non-linear" filters based on order statistics.

intensity
white > Black
(brightest) (darkest)
255 0

Key points about each filter:

Median filter:

- Most commonly used for removing "salt and pepper" noise, which appears as random bright and dark pixels.

11

Home Search Saved Notifications

Detection of isolated points

• The Laplacian

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$= f(x+1,y) + f(x-1,y) + f(y,y+1) + f(x,y-1) - 4f(x,y)$$

$$g(x,y) = \begin{cases} 1, & \text{if } |R(x,y)| \geq T \quad \left[R = \sum_{k=1}^{9} w_k z_k\right] \\ 0 & \text{otherwise} \end{cases}$$

Sharpening Spatial Filters: Laplace Operator

The second-order isotropic derivative operator is the Laplacian for a function (image) f(x,y)

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial^2 f}{\partial x^2} = f(x+1,y) + f(x-1,y) - 2f(x,y)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x,y+1) + f(x,y-1) - 2f(x,y)$$

$$\nabla^2 f = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y)$$

Sharpening Spatial Filters: Laplace Operator

0	1	0	1	1	1
1	-4	1	1	-8	1
0	1	0	1	1	1
0	-1	0	-1	-1	-1
-1	4	-1	-1	8	-1
0	-1	0	-1	-1	-1

a b
c d

FIGURE 3.37 (a) Filter mask used to implement Eq. (3.6-6). (b) Mask used to implement an extension of this equation that includes the diagonal terms. (c) and (d) Two other implementations of the Laplacian found frequently in practice.

Sharpening Spatial Filters: Laplace Operator

by using Laplace operator

Image sharpening in the way of using the Laplacian:

$$g(x,y) = f(x,y) + c\left[\nabla^2 f(x,y)\right]$$

where,

f(x, y) is input image,

g(x, y) is sharpenend images, (o/p img)

c = -1 if $\nabla^2 f(x,y)$ corresponding to Fig. 3.37(a) or (b)

and c = 1 if either of the other two filters is used

c = -1 to 1

10:48

what is laplace ord...

Show more

বাংলায় | In English

Introduction. The Laplacian operator is a second-order differential operator in n-dimensional Euclidean space, denoted as ∇^2. It is the divergence of the gradient of a function. 16 Dec 2023

(Prrv)

https://medium.com › aimonks › e...

Exploring the Laplacian Operator: A Key Tool in Computer Vision for ...

About featured snippets | Feedback

People also ask

What is the Laplace operator in computer vision?

Home | Search | Saved | Notifications

Unsharp Masking and Highboost Filtering

- Unsharp masking
  - Sharpen images consists of subtracting an unsharp (smoothed) version of an image from the original image
  - e.g., printing and publishing industry (prai)

- Steps
  1. Blur the original image
  2. Subtract the blurred image from the original
  3. Add the mask to the original

Unsharp Masking and Highboost Filtering

Let $\bar{f}(x, y)$ denote the blurred image, unsharp masking is

$$g_{mask}(x, y) = f(x, y) - \bar{f}(x, y)$$

Then add a weighted portion of the mask back to the original

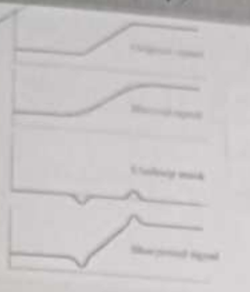
$$g(x, y) = f(x, y) + k * g_{mask}(x, y) \qquad k \geq 0$$

when $k > 1$, the process is referred to as highboost filtering.

highly sharpened img

Unsharp Masking: Demo

Unsharp Masking: Demo (i7)



a
b
c
d

FIGURE 3.39 1-D illustration of the mechanics of unsharp masking. (a) Original signal. (b) Blurred signal with original shown dashed for reference. (c) Unsharp mask. (d) Sharpened signal, obtained by adding (c) to (a).

Unsharp Masking and Highboost Filtering: Example

a
b
c
d
e

FIGURE 3.40 (a) Original image. (b) Result of blurring with a Gaussian filter. (c) Unsharp mask. (d) Result of using unsharp masking. (e) Result of using highboost filtering.

~ edge detection

Image Sharpening based on First-Order Derivatives (Sobel operator)

For function $f(x, y)$, the gradient of f at coordinates (x, y) is defined as (gradient)

5

$$\nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$$

The *magnitude* of vector ∇f, denoted as $M(x, y)$ (mag. of grad.)

Gradient Image $\quad M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^{\,2} + g_y^{\,2}}$

Prewitt mask for detecting diagonal edges

0	1	1
-1	0	1
-1	-1	0

-1	-1	0
-1	0	1
0	1	1

Sobel

0	1	2
-1	0	1
-2	-1	0

2	-1	0
-1	0	1
0	1	2

Image Sharpening based on First-Order Derivatives

The *magnitude* of vector ∇f, denoted as $M(x, y)$

$$M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2}$$

$$M(x, y) \approx |g_x| + |g_y|$$

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

$$M(x, y) \approx |z_8 - z_5| + |z_6 - z_5|$$

Image Sharpening based on First-Order Derivatives

Roberts Cross-gradient Operators

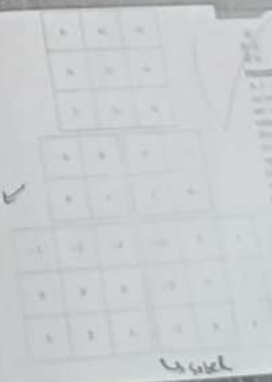
$$M(x, y) \approx |z_9 - z_5| + |z_8 - z_6|$$

Sobel Operators

$$M(x, y) \approx |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Image Sharpening based on First-Order Derivatives



Roberts

Sobel

Prewitt

Spatial resolution

i) Smallest discernible detail in a image

ii) measured in dpi (~~distance~~ dots per inch), dots(pixels) per unit distance(dpud), line per unit distance (lpud).

iii) It focuses on the no. of pixels in an img.

iv) makes img sharper & clearer.

v) Eg- A 4k img shows more detail than an HD image

Intensity resolution

i) Smallest ~~discernable~~ discernible change in intensity.

ii) measured in bits: 8 bits, 12 bits, 16 bits etc.

iii) Focuses on the no. of brightness/ colors levels per pixel.

iv) makes img smoother with more colors / brightness options.

v) Eg - A 16-bit img has more shades than an 8-bit img.

15 Image Segmentation (Concept)

Image segmentation partitions an image into regions or objects based on properties like color, intensity, and texture.

16 Addition of Two Images

Combining pixel values of two images to create a composite image. ~~Contrast~~ ~~transfer~~ ~~betw~~ ~~body~~

17 Spatial and Intensity Resolution

diff? {

- Spatial Resolution: The smallest discernible detail in an image.
- Intensity Resolution: The smallest discernible change in intensity.

18 Connected Component Analysis

(Identifies and labels connected regions in a binary image) f.f.o

19 Light-Sensitive Receptors

- Rods: Sensitive to low light levels.
- Cones: Detect color and function in bright light.

20 Edge Detection Using Prewitt and Sobel Operators

- Prewitt and Sobel: First-order derivative operators for edge detection.
- Marr-Hildreth: A second-order method using Laplacian of Gaussian. - for edge detection

derivative operators

[redacted]

21 Line and Point Detection

(Algorithms are designed to identify specific structures such as lines and points in an image)

22 Image Classification

22.1 Problem and Challenges

- Variability in illumination, perspective, and noise.

22.2 Nearest Neighbor Classifier

- L1 Distance: Sum of absolute differences. (ass)

22.3 Hyperparameters & model

Values set before the learning process to optimize performance.

22.4 Kth Nearest Neighbor

Considers the majority class among k closest points.

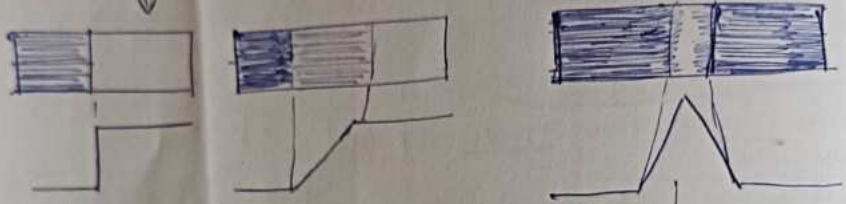
22.5 Linear Classification (Parametric Approach)

Models the decision boundary as a linear function.

Edge detection

Edges are pixels where the brightness function changes abruptly.

Edge models



L → R models (ideal rep.) of a Step, Ramp, & Roof edge & their correspon-ding intensity profile.

R = entire spatial region occupied by an img

Here, R is partitioned into n sub-regions $R_1, R_2 \ldots R_n$

such that,

a) $\bigcup_{i=1}^{n} R_i = R$

b) R_i is a connected set, $i = 1, 2, \ldots n$

c) $R_i \cap R_j = \phi$

d) $Q(R_i) = \text{TRUE}, \; i = 1, 2, \ldots n$

e) $Q(R_i \cup R_j) = \text{FALSE}$ for any adjacent regions R_i & R_j

10:58

what is image seg...

Image segmentation is used in many fields, including:

- **Medical imaging:** To detect and label pixels that represent tumors in a patient's brain or other organs
- **Agriculture:** To use image segmentation in agricultural applications
- **Satellite images:** To use image segmentation in satellite images
- **Autonomous vehicles:** To use image segmentation in AI for autonomous vehicles

Some techniques used in image segmentation include:

- **Watershed segmentation:** Treats the image as a topographic surface, where pixel values represent elevation. This technique is useful for separating overlapping

Home Search Saved Notifications

11:00

how to do Addition...

Key points about image addition:

Pixel-wise operation:

The addition happens at the level of individual pixels, meaning the value of each pixel in the resulting image is the sum of the values of the corresponding pixels in the original images.

The addition operation is based on indivi-dual pixel addition

Same dimensions:

Both input images must have the same dimensions (width and height) for the addition to work correctly.

Color space compatibility:

The images should be in the same color space (e.g., RGB, grayscale) to ensure meaningful addition.

Potential overflow:

If the pixel values add up to exceed the maximum value allowed by the data type (e.g., 255 for 8-bit grayscale), the result might be clipped to the maximum value, leading to potential information loss.

Advanced Techniques for Edge Detection

- The Marr-Hildreth edge detector / operator

$$G(x,y)=e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad \sigma: \text{space constant.}$$

$e^{-\frac{x^2+y^2}{2\sigma^2}}$

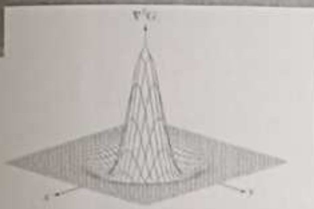
Laplacian of Gaussian (LoG)

$$\nabla^2 G(x,y)=\frac{\partial^2 G(x,y)}{\partial x^2}+\frac{\partial^2 G(x,y)}{\partial y^2}$$

$$=\frac{\partial}{\partial x}\left[\frac{-x}{\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\right]+\frac{\partial}{\partial y}\left[\frac{-y}{\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\right]$$

$$=\left[\frac{x^2}{\sigma^4}-\frac{1}{\sigma^2}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}+\left[\frac{y^2}{\sigma^4}-\frac{1}{\sigma^2}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$=\left[\frac{x^2+y^2-\sigma^2}{\sigma^4}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}$$



0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

a b
c d

FIGURE 10.21 (a) Three-dimensional plot of the *negative* of the LoG. (b) Negative of the LoG displayed as an image. (c) Cross section of (a) showing zero crossings. (d) 5 × 5 mask approximation to the shape in (a). The negative of this mask would be used in practice.

Marr-Hildreth Algorithm

1. Filter the input image with an nxn Gaussian lowpass filter. n is the smallest odd integer greater than or equal to 6σ $n \geq 6\sigma$

2. Compute the Laplacian of the image resulting from Step 1.

3. Find the zero crossing of the image from Step 2.

$$g(x,y)=\nabla^2\left[G(x,y)\star f(x,y)\right]$$

11:19

connected compo...

Key points about connected component analysis:

Binary image input:
This technique primarily operates on binary images, where each pixel is either considered "on" (foreground) or "off" (background).

Labeling connected regions:
The core function of CCA is to assign unique labels to each separate connected region within the image, allowing you to differentiate between different objects.

Connectivity definition:
When determining if pixels are connected, you can use different connectivity rules like 4-connectivity (only checking adjacent pixels horizontally and vertically) or 8-connectivity (including diagonals).

Applications of connected component analysis:

Home Search Saved Notifications

11:19

connected compo...

Applications of connected component analysis:

Object counting:

Counting the number of distinct objects in an image by identifying and counting the labeled connected components.

Character recognition:

Analyzing individual characters on a license plate by isolating each character as a connected component.

Medical image analysis:

Identifying and analyzing individual cells or tissue regions in a medical scan

Defect detection:

Finding flaws or cracks in an image by identifying small connected regions that deviate from the expected pattern

Home Search Saved Notifications

12:52

Trending Now DSA Web Tech Foundational C

Prewitt Operator

→ 1st order derivative operator for edge detection

The Prewitt operator was developed by Judith M. S. Prewitt. Prewitt operator is used for edge detection in an image. Prewitt operator detects both types of edges, these are:

- Horizontal edges or along the x-axis,
- Vertical Edges or along the y-axis.

Wherever there is a sudden change in pixel intensities, an edge is detected by the mask. Since the edge is defined as the change in pixel intensities, it can be calculated by using differentiation. Prewitt mask is a first-order derivative mask. In the graph representation of Prewitt-mask's result, the edge is represented by the local maxima or local minima.

Both the first and second derivative masks follow these three properties:

- More weight means more edge detection.
- The opposite sign should be present in the mask. (+ and

Open in App

- The opposite sign should be present in the mask. (+ and -)
- The Sum of the mask values must be equal to zero.

Prewitt operator provides us two masks one for detecting edges in the horizontal direction and another for detecting edges in a vertical direction.

Prewitt Operator [X-axis] = [-1 0 1; -1 0 1; -1 0 1]

Prewitt Operator [Y-axis] = [-1 -1 -1; 0 0 0; 1 1 1]

Steps:

- Read the image.
- Convert into grayscale if it is colored.
- Convert into the double format.
- Define the mask or filter.
- Detect the edges along X-axis.
- Detect the edges along Y-axis.
- Combine the edges detected along the X and Y axes.
- Display all the images.

Imtool() is the inbuilt function in Matlab. It is used to display the image. It takes 2 parameters, the

Basic Edge Detection by Using First-Order Derivative

$$\nabla f \equiv grad(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$$

The magnitude of ∇f

$$M(x,y) = \text{mag}(\nabla f) = \sqrt{g_x^{\,2} + g_y^{\,2}}$$

The direction of ∇f

$$\alpha(x,y) = \arctan(g_y / g_x)$$

The direction of the edge

$$\phi = \alpha - 90^\circ$$

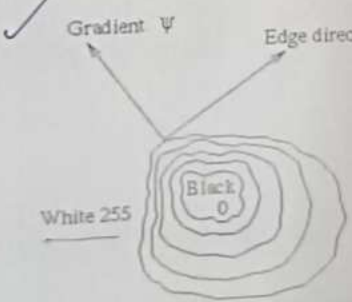
Basic Edge Detection by Using First-Order Derivative

Edge normal: $\nabla f \equiv grad(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$

Edge unit normal: $\nabla f / \text{mag}(\nabla f)$

In practice,sometimes the magnitude is approximated by

$$\text{mag}(\nabla f) = \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right| \quad \text{or} \quad \text{mag}(\nabla f) = \max\left(\left|\frac{\partial f}{\partial x}\right|, \left|\frac{\partial f}{\partial y}\right|\right)$$

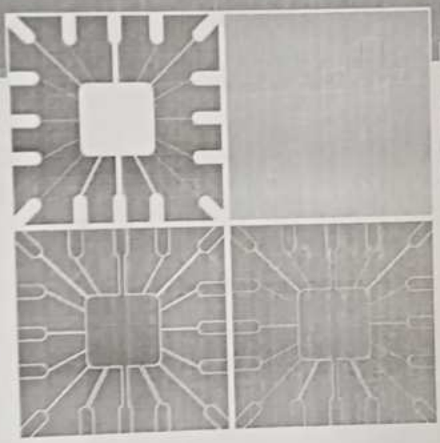


Line Detection

- Second derivatives result in a stronger response and produce thinner lines than first derivatives

- Double-line effect of the second derivative must be handled properly

- For this various algos are used.



a b
c d

FIGURE 10.5
(a) Original image. (b) Laplacian image; the magnified section shows the positive/negative double-line effect characteristic of the Laplacian. (c) Absolute value of the Laplacian. (d) Positive values of the Laplacian.

Detecting Line in Specified Directions

- Let R_1, R_2, R_3, and R_4 denote the responses of the masks in Fig. 10.6. If, at a given point in the image, $|R_k|>|R_j|$, for all $j \neq k$, that point is said to be more likely associated with a line in the direction of mask k.

what is point detec...

In computer vision, "point detection" refers to the process of identifying specific, salient points of interest within an image, often called "key-points" or "interest points," which are locations that stand out due to significant changes in intensity or texture, allowing for further analysis of the image like object tracking, pose estimation, or facial recognition; essentially, it's about pinpointing key locations within an image that provide meaningful information about the scene or object depicted.

From this Point detection (Laplacian) mask is used

Key points about point detection:

What it detects:

Corners, junctions, or other distinct features in an image that are relatively invariant to changes in viewpoint, scale, or lighting.

Applications:

- **Object tracking:** Identifying key points on an object to track its movement across frames.
- **Human pose estimation:** Locating joints on a human body.
- **Facial recognition:** Identifying key facial landmarks like eyes, nose, and mouth.
- **Augmented reality:** Precisely placing virtual objects in the real world.

Common algorithms:

- **Harris Corner Detector:** A classic algorithm that uses the local image gradient to identify corners.
- **Scale Invariant Feature Transform (SIFT):** Detects keypoints across different scales and rotations.
- **Oriented FAST and Rotated BRIEF (ORB):** A faster alternative to SIFT with good performance.

A Beginners Guide to Computer Vision (Part...

23 Apr 2020 — Here comes the concept of Interest Points. Look...

I/P img → img preprocessing → Analysis → o/p & visualization

- Image classification?

Image classification pipeline

Image data → Artificial feature extraction → weight training → Result o/p

- Image classification – It's a tool that labels an image based on its visual content.

- With N examples, how fast are training & prediction? j.i.c TC (Nearest neighbour classification)
  - Train $O(1)$
  - Predict $O(N)$

  But this is bad, we want reverse

- For setting hyper parameters –

  a) choose hyper parameters
  b) split data into train & test,
  choose hyperparameters } → Bad

  a) "
  b) "
  c) split data into train, val & test,
  choose hyper parameters } → Better

1:14

what is image class...

Image classification is a computer vision task that involves assigning a label to an image based on its contents.

How it works

Image classification analyzes an image at the pixel level to determine the most appropriate label. It uses training data of already labeled images to assign a probability to each class.

Why it's important

Image classification is a fundamental task in computer vision and machine learning. It provides valuable data and insights that can help inform decisions.

How it's performed

Image classification is typically performed using classification networks such as CNNs.

Types of classification

There are two types of image classification: supervised and unsupervised, depending on the interaction between the analyst and what are problems...

Some challenges and problems in image classification in computer vision include:

Ethical concerns
Biased training data can lead to biased predictions that reinforce social biases.

Poor lighting
Changes in brightness, shadows, and dark spots can make it difficult for algorithms to recognize objects.

Deformation
In the real world, objects can deform, making it difficult for algorithms to identify them.

Scalability
As complexity increases, it's important to ensure scalability.

Accuracy
Inaccurate or mislabeled data can lead to model errors.

Data quantity
There may be a large number of images.

Data dimensionality
The data may have high dimensionality.

Labeled data
There may be a lack of labeled data.

Variations
There may be variations in lighting, scale, and orientation.

Seeing Through the Machine's Eyes: Top...
Poor Lighting, Changes in...

Common Image Cl...

First classifier: Nearest Neighbor

→ img is classified as the class of the nearest neighbour image, where the distance of unknown img from labelled img is calculated by L1 distance

```
def train(images, labels):
    # Machine learning!
    return model
```

→ Memorize all data and labels

```
def predict(model, test_images):
    # Use model to predict labels
    return test_labels
```

→ Predict the label of the most similar training image

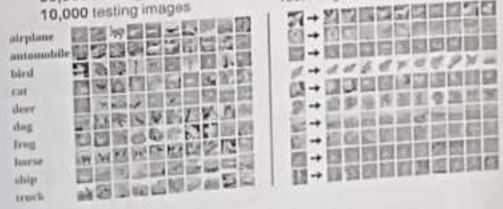
Example Dataset: CIFAR10

~~10 classes~~
50,000 training images
10,000 testing images

Example Dataset: CIFAR10

10 classes
50,000 training images
10,000 testing images

Test images and nearest neighbors



Distance Metric to compare images

L1 distance: $d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$

test image

56	32	10	18
90	23	128	133
24	26	178	200
2	0	255	220

−

training image

10	20	24	17
8	10	89	100
12	16	178	170
4	32	233	112

=

pixel-wise absolute value differences

46	12	14	1
82	13	39	33
12	10	0	30
2	32	22	108

$\xrightarrow{\text{add}}$ 456

what is nearest nei...

Applications in computer vision:

Image classification:

Identifying the category of an image (e.g., cat, dog, car) based on its similarity to labeled training images.

Object detection:

Locating and classifying objects within an image by comparing small image patches to known object templates.

Limitations:

- **Computational cost:** Can be computationally expensive for large datasets as it requires calculating distances to all training images for each new image.
- **Sensitivity to outliers:** Can be sensitive to noise or outliers in the training data.
- **Curse of dimensionality:** Performance can degrade significantly when dealing with

• Inefficient for complex dataset or training process

Hyperparameters

What is the best value of **k** to use?
What is the best **distance** to use?

These are **hyperparameters**: choices about the algorithm that we set rather than learn

Very problem-dependent.
Must try them all out and see what works best.

Setting Hyperparameters

Idea #1: Choose hyperparameters that work best on the data

Your Dataset

Setting Hyperparameters

Idea #1: Choose hyperparameters that work best on the data

BAD: $K = 1$ always works perfectly on training data

Your Dataset

Setting Hyperparameters

Idea #1: Choose hyperparameters that work best on the data

BAD: $K = 1$ always works perfectly on training data

Your Dataset

Idea #2: Split data into **train** and **test**, choose hyperparameters that work best on test data

train | test

• Setting hyper parameters

Cross validation → It means split data into folds, try each fold as validation & avg the results.

f1	f2	f3	f4	f5	f6	f7	test

↳ 7-fold c.v.

---

what is hyperpara...

In computer vision, a hyperparameter is a configuration variable that is manually set before training a machine learning model. to get the best performance

→ K=7 i.e. 7-fold cross validation works better than other c.v.

Definition

A hyperparameter is a parameter whose value is set before the machine learning process begins.

Purpose

Hyperparameters define the model's configuration and settings, and guide the learning process. to get the best performance

Examples

Examples of hyperparameters in computer vision include the learning rate, batch size, number of layers, filter sizes, pooling strategies, dropout rates, and activation functions.

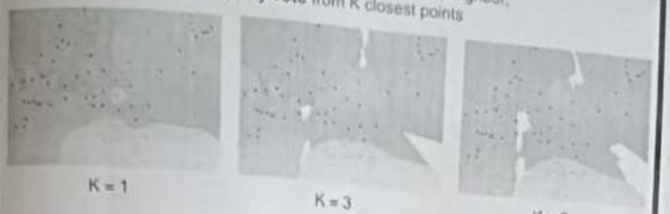
Importance

Choosing the right hyperparameters is important because they directly

Search Saved Notifications

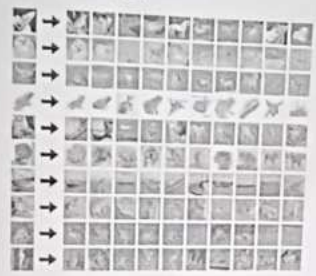
K-Nearest Neighbors

Instead of copying label from nearest neighbor, take **majority vote** from K closest points

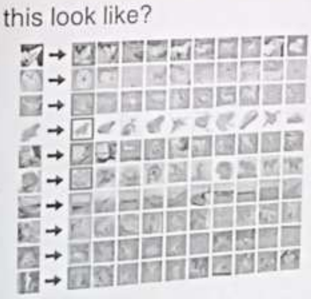


K = 1 K = 3 K = 5

What does this look like?



What does this look like?



K-Nearest Neighbors: Distance Metric

L1 (Manhattan) distance

$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$

sum of abs distance of difference

L2 (Euclidean) distance

$$d_2(I_1, I_2) = \sqrt{\sum_p (I_1^p - I_2^p)^2}$$

K-Nearest Neighbors: Distance Metric

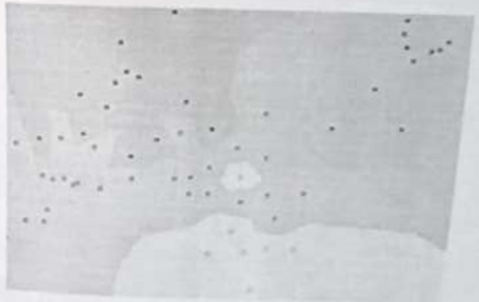
L1 (Manhattan) distance

$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$

K = 1

L2 (Euclidean) distance

$$d_2(I_1, I_2) = \sqrt{\sum_p \left(I_1^p - I_2^p\right)^2}$$



K = 1

1:26

Q what is k nearest n...

In computer vision, "K-Nearest Neighbor" (KNN) is a machine learning algorithm used for image classification, where a new image is categorized based on the class labels of the "K" most similar images (its nearest neighbors) from a pre-labeled training dataset, essentially deciding the class of a new image by comparing it to known images that are closest to it in feature space.

Supervised

Euclidean distance metric is mainly used here.

Key points about KNN in computer vision:

How it works:

- **Feature extraction:** First, each image in the training set is converted into a feature vector (a ... of numerical values ...enting ima...

1:26

what is k nearest n...

"K" value:

The value of "K" is a hyperparameter that needs to be tuned, determining how many neighbors to consider when making a classification decision.

Advantages:

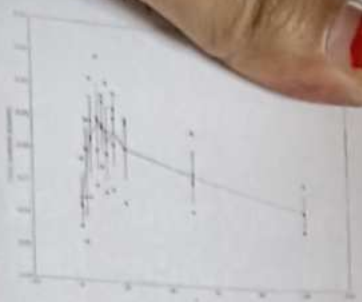
- Simple to implement
- No complex training process needed, as it only stores the training data
- Can be effective for simple classification tasks

Disadvantages:

- Can be computationally expensive for large datasets
- Sensitive to the choice of distance metric
- May not perform well on complex image classification tasks with high dimensionality → large dim (curse of dim)

Home Search Saved Notifications

• sensitive to outliers



Example of
-fold cross-validation
he value of **k**.

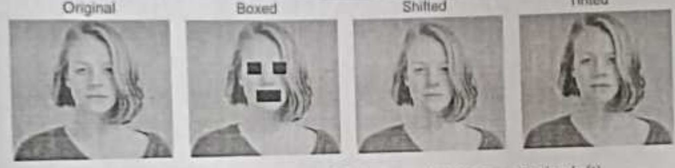
Each point: single
outcome.

The line goes
through the mean, bars
indicated standard
deviation

(Seems that k ~= 7 works best
for this data)

k-Nearest Neighbor on images **never used.**

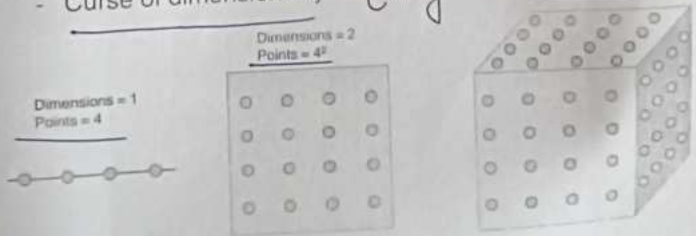
- Very slow at test time (computationally expensive)
- Distance metrics on pixels are not informative (sensitive on chosing distance metric)



(all 3 images have same L2 distance to the one on the left)

k-Nearest Neighbor on images **never used.**

- Curse of dimensionality (may not perform well for complex images)



38-40/53

K-Nearest Neighbors: Summary

In **Image classification** we start with a **training set** of images and labels, and must predict labels on the **test set**

The **K-Nearest Neighbors** classifier predicts labels based on nearest training examples

Distance metric and K are **hyperparameters**

Choose hyperparameters using the **validation set**; only run on the test set once at the very end!

11

1:28 3.00 KB/S 20

what is linear clad...

Example algorithms using linear classification:

- **Perceptron:** A basic linear classifier that updates weights based on misclassified data
- **Logistic Regression:** A probabilistic linear classifier that outputs the probability of belonging to a certain class
- **Linear Discriminant Analysis (LDA):** A dimensionality reduction technique that aims to find the linear combination of features that best separates classes

will not perform well for non linear data

Limitations of linear classification:

Non-linear data:

If the data cannot be easily separated by a straight line (non-linearly separable), a linear classifier will not perform well.

(or) less efficient for non linear data

Complex patterns:

For complex image recognition tasks where features are not linearly related, more advanced non-linear classifiers are often needed.

as non linear data isn't used so it's hard for it to identify complex patterns

Linear Classifier - an overview | ScienceDire...

Linear vs. Classifica 28 Jun 2024

(or)

As non-linearity isn't introduced, the model is unable to identify complex pattern

- **Hyperplane** – (Perceptron theory) → (Conversion from normal to hyperplane) (form ML)

It's a decision boundary that divides the i/p space into 2 or more regions, each corresponding to a diff. class or o/p label.

<u>In general we know</u> (from perceptron theory)

$$y_k = \cancel{\phi(x)} = \phi\left(\sum_{j=1}^{n} w_{kj} \cdot x_j + b\right) \quad [\phi(x) = \text{A.F.}]$$

where, $X = \sum_{j=1}^{n} w_{kj} * x_j + b$ is a hyperplane function.

[w = weights, x = input, b = bias] [y_k = o/p]

In perceptron learning we reduce the hyperplane func. into hyperplane eqn, that is

$$X = 0$$

$$\Rightarrow \sum_{j=1}^{n} w_{kj} \cdot x_j + b = 0$$

If we are solving 2 class problem or binary class problem, this eqn is reduced into 2 eqns,

$$\sum_{j=1}^{n} w_{kj} * x_j + b \geq 0 \quad \text{— (i)}$$

$$\sum_{j=1}^{n} w_{kj} * x_j + b < 0 \quad \text{— (ii)}$$

} hyperplane or 4D

That's how we convert from normal to hyperplane.

---

what is parametric...

In computer vision, a "parametric approach" in linear classification refers to a method where the classification model is defined by a fixed set of parameters (like weights and biases) that are learned from training data, allowing for efficient prediction by simply applying a linear function to the input image features, unlike non-parametric methods which might require storing the entire training data for comparison; essentially, it means the model assumes a specific form (like a linear equation) to make predictions, with the key focus being on optimizing those parameters to best fit the data.

→ $f(x, W) = Wx$ [W = weights or parameters] $= Wx + b$ add (bias) (me)

Key points about parametric linear classification:

Fixed model structure:

Home Search Saved Notifications

Representing Digital Images

TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,360,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Spatial and Intensity Resolution

- Spatial resolution
  - A measure of the smallest discernible detail in an image
  - stated with *line pairs per unit distance, dots (pixels) per unit distance, dots per inch (dpi)*
- Intensity resolution
  - The smallest discernible change in intensity level
  - stated with *8 bits, 12 bits, 16 bits, etc.*

Spatial and Intensity Resolution

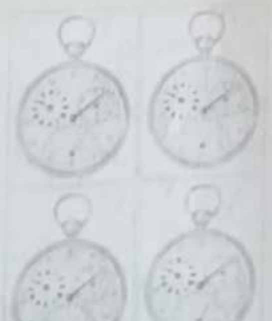


FIGURE 2.20

Representing Digital Images

The representation of an M×N numerical array in MATLAB

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

Representing Digital Images

- Discrete intensity interval [0, L-1], $L=2^k$

L → no. of discrete intensity levels used

The number b of bits required to store a M × N digitized image

$b = M \times N \times k$

k = no. of bits used to rep. each pixel in an img

Representing Digital Images

TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
64	4,096	8,192	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
128	16,384	32,768	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
256	65,536	131,072	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
512	262,144	524,288	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1024	1,048,576	2,097,152	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2048	4,194,304	8,388,608	12,582,912	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4096	16,777,216	33,554,432	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
8192	67,108,864	134,217,728	201,326,592	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

eg

N = = M
ify ∴ N = 32,
L = 2, ∴ k = 1
∴ b = M×N×k
= 32×32×1 = 1024 → no. of bits rel. to store a 32×32 img with 2 intensity levels
ify b = 32×32×2 = 2028
(L=4, 2^2) → 4 intensity levels
like that

The human eye is a → camera.

Iris – Colored annulus with radial muscles.

pupil – the hole (aperture) whose size is controlled by iris.

Film – photoreceptor cells (rods & cones) in retina.

↓

layer of tissue at the back of the eyeball.

Two types of light-sensitive receptors

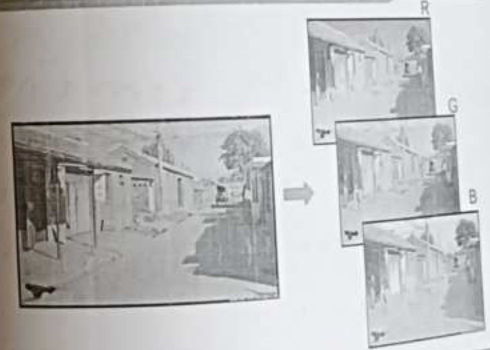
Cones

- cone-shaped
- less sensitive
- operate in high light color vision
- Detects Colourful imgs
- Fast to respond
- sensitive to high lights (brightness)

Rods

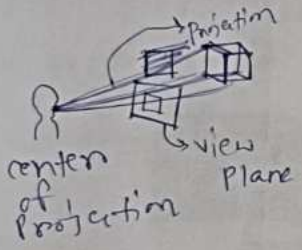
- rod-shaped
- highly sensitive
- operate at night
- gray-scale vision (b/w)
- slower to respond
- ~~a frequent~~ sensitive to low light (darkness)

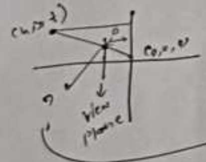
Color Image



Thank You

Projection - In general, it ~~mean~~ refers to transform point in n-space to m-space (m<n). In CV, it refers to map viewing coordinates to 2D screen coordinates.



- 

What are the coordinates of the point resulting from projection of (x, y, z) onto the view plane? (From 3D → 2D coordinates)

$$? = \left(\frac{xD}{z}, \frac{yD}{z}\right)$$ formula

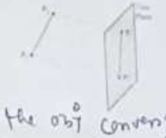
[D = Distance from view plane]

Taxonomy of Projections

- Planar geometric projection
  - Parallel
    - Orthographic
      - Top
      - Front
      - Side
      - Axonometric
    - Oblique
  - Perspective
    - One-point
    - Two-point
    - Three-point

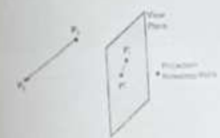
Parallel & Perspective

Parallel Projection → Lines from the obj's vertices intersect the viewing plane to form a parallel projection



Perspective Projection → Lines from the obj converge at a single point, called the center of projection.

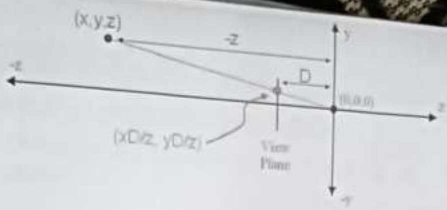
(Pro from diff - diffi)



15-16/24

Parallel Projection

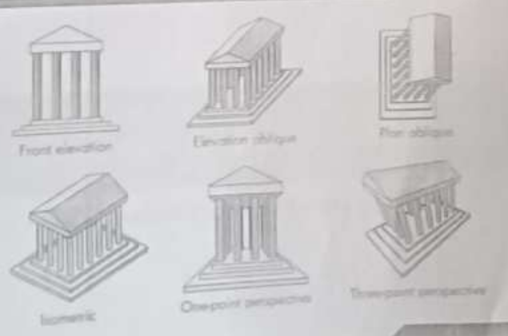
- Center of projection is at infinity
  - Direction of projection (DOP) same for all points



Perspective vs. Parallel

- Perspective projection
  - \+ Size varies inversely with distance – looks realistic   size $\propto \frac{1}{dist}$
  - – Distance and angles are not(in general) preserved
  - – Parallel lines do not (in general) remain parallel   – more realistic looking
- Parallel projection
  - \+ Good for exact measurements   size $\propto$ dist
  - \+ Parallel lines remain parallel
  - – Angles are not (in general) preserved
  - – Less realistic looking

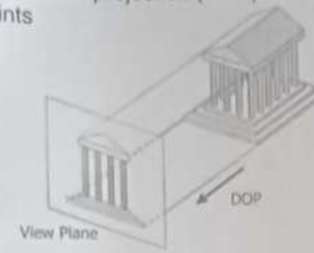
Classical Viewing



(Thank You) 22-23/24

Parallel Projection

- Center of projection is at infinity
  - Direction of projection (DOP) same for all points



Orthographic & Oblique

- Orthographic parallel projection
  - the projection is perpendicular to the view plane

- Oblique parallel projection
  - The projectors are inclined with respect to the view plane

Orthographic Projections

- DOP perpendicular to view plane