Digital Image Processing

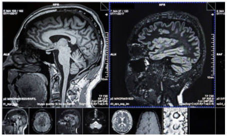
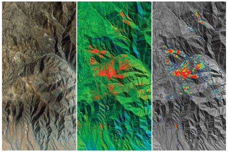
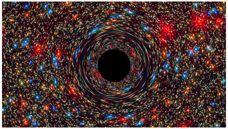
Noel Jeffrey Pinton

University of the Philippines - Cebu / Department of Computer Science

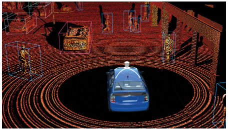
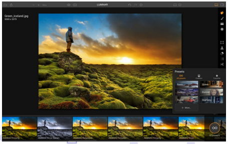
CMSC 178IP

Outline

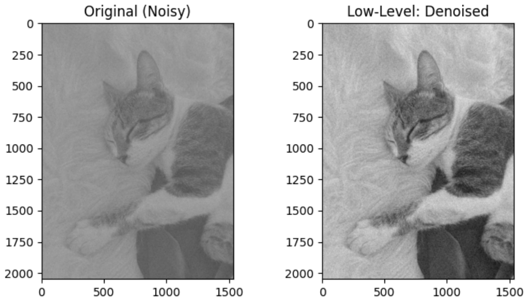
What is Digital Image Processing?

- Digital Image Processing (DIP) refers to the manipulation of image data using a digital computer.
- It involves converting an image into a numerical representation and applying algorithms to extract, enhance, or analyze information.
- Applications span medicine (CT, MRI), remote sensing, astronomy, robotics, industrial inspection, and multimedia.

Low-Level Processing

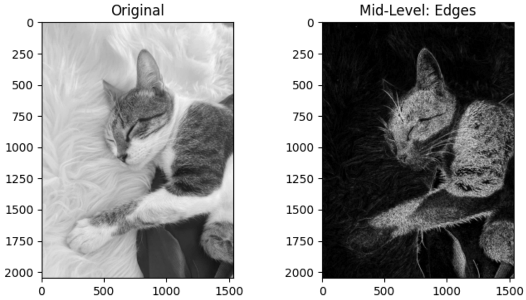


- Deals with basic operations on pixels.
- Examples: image denoising, smoothing, sharpening, contrast adjustment.
- Mathematical model:

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

where $*$ denotes convolution.

Mid-Level Processing



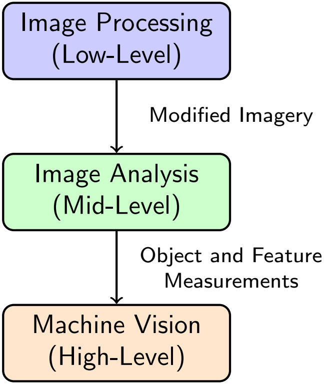
- Extracts meaningful structures from an image.
- Example: **Edge detection** using Sobel operator.
- Sobel convolution kernels:

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * f(x, y), \quad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * f(x, y)$$

High-Level Processing

- Interprets objects and scenes.
- Involves recognition, classification, and semantic understanding.
- Example: object classification using machine learning.

Three Layers of Image Understanding (Diagram)



- **Low-Level (Image Processing)** Denoising, contrast enhancement in CT scans.
- **Mid-Level (Image Analysis)** Tumor segmentation, edge detection of structures.
- **High-Level (Machine Vision)** Tumor classification, object recognition in scenes.

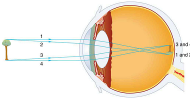
What are Image Sensors?

- **Definition:** An image sensor is a device that converts light or other forms of electromagnetic radiation into a digital signal (an image).
- **Importance:** They are the interface between the physical world and digital image processing systems.
- **Examples:** Optical cameras (CCD, CMOS), Radar sensors, Medical scanners (MRI, CT), Infrared sensors, and the Human eye.

Digital Camera

PET/CT Scanner

Radar

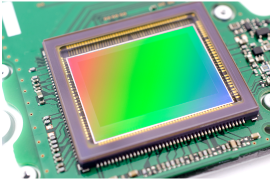


Human Eye

Optical Image Sensors

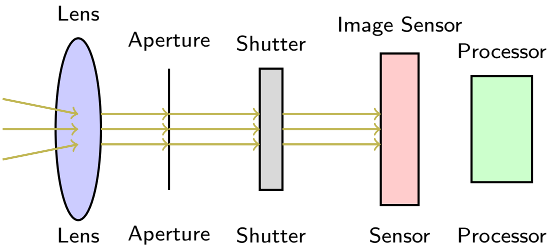
- Use visible light to capture images.
- Two main types: CCD (Charge-Coupled Device) and CMOS (Complementary Metal-Oxide-Semiconductor).
- Widely used in smartphones, digital cameras, surveillance, and scientific imaging.

Digital Camera



CMOS Image Sensor

Parts of a Typical Optical Camera

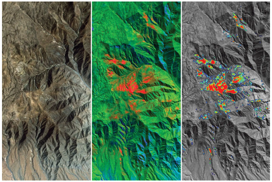


- Light enters through the **lens**, passes the **aperture** and **shutter**.
- The **sensor** converts photons into electrical signals.
- The **processor** handles image formation, storage, and compression.

Radar Image Sensors

- Use radio waves to capture surface features.
- Useful in remote sensing, weather monitoring, and military applications.
- Capable of imaging through clouds, smoke, and even vegetation.

Radar

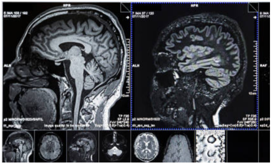


Remote Sensing

Medical Image Sensors

- Specialized sensors capture non-visible signals: X-rays, gamma rays, magnetic resonance.
- Used in CT (X-ray), MRI (magnetic fields), Ultrasound (sound waves).
- Critical for diagnostics, treatment planning, and research.

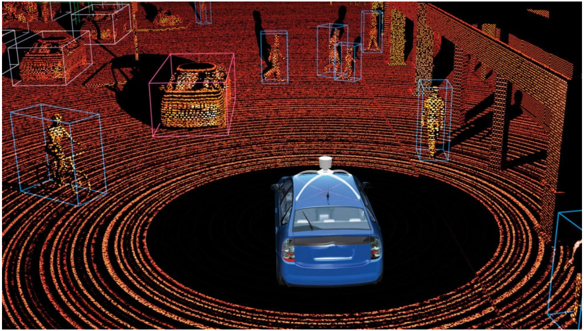
Computed Tomography Scanner



CT Scan Image

Other Types of Image Sensors

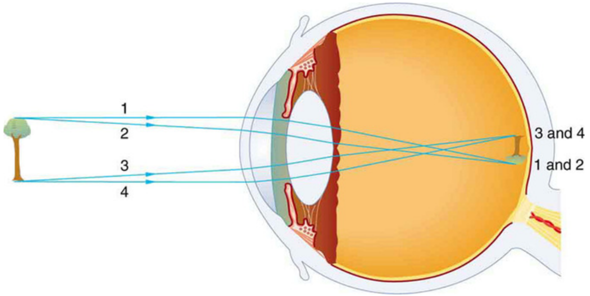
- Infrared (night vision, thermal imaging).
- Hyperspectral sensors (satellite and agricultural monitoring).
- LIDAR (3D mapping and autonomous vehicles).



LIDAR for Self-Driving Cars

The Human Eye as a Sensor

- The human eye is a biological sensor that converts light into neural signals.
- Rods: sensitive to light intensity (low-light vision).
- Cones: sensitive to color (RGB perception).
- Inspiration for artificial imaging systems.



Human Eye

What is an Image?

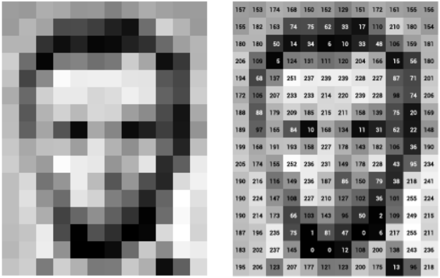
- An **image** is a two-dimensional function:

$$f(x, y) : \mathbb{Z}^2 \to \{0, 1, \ldots, L-1\}$$

  where (x, y) are spatial coordinates and $f(x, y)$ is the intensity at that point.

- Each element of the image is a **pixel** (picture element).
- **Intensity values:**
  - $0 \to$ Black (minimum intensity).
  - $L-1$ (maximum value) $\to$ White (maximum intensity).
  - For 8-bit images, $L = 256$, so intensity $\in [0, 255]$.
- Images can be represented in different forms: binary, grayscale, color (RGB), multispectral, hyperspectral.

Mathematical Representation



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

- A grayscale image is represented as a 2D function:

$$f(x, y) : \mathbb{Z}^2 \to \mathbb{R},$$

  where $f(x, y)$ gives the intensity at pixel coordinates (x, y).
- A color image is represented as a vector function:

$$f(x, y) = \big(f_R(x, y), f_G(x, y), f_B(x, y)\big).$$

Binary Image Representation

- A **binary image** has only two possible intensity values:

$$f(x, y) \in \{0, 1\}$$

- Typically:
  - $0 \rightarrow$ Black (background).
  - $1 \rightarrow$ White (foreground / object).
- Used in:
  - Document scanning (text vs. background).
  - Masks in image processing.
  - Morphological operations (erosion, dilation).
  - Object detection (silhouettes, shapes).
- Can be obtained by **thresholding** a grayscale image:

$$f_b(x, y) = \begin{cases} 1, & f(x, y) \geq T \\ 0, & f(x, y) < T \end{cases}$$

Grayscale Image

Binary Image (Threshold=128)

Grayscale (Monochrome) Image Representation

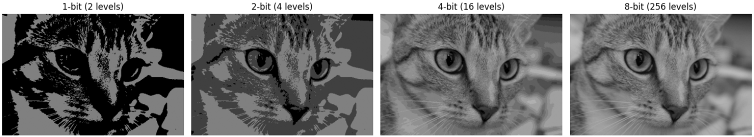
- A **grayscale image** (also called **monochrome**) uses shades of gray to represent intensity.
- Each pixel stores an integer value in range:

$$f(x, y) \in [0, L-1], \quad L = 2^k$$

  where k is the number of bits per pixel (bpp).
- Typical case: $k = 8 \Rightarrow L = 256$ gray levels.
  - $0 \rightarrow$ Black
  - $255 \rightarrow$ White
  - Values in between $\rightarrow$ shades of gray
- Applications:
  - Medical imaging (X-rays, CT, MRI)
  - Remote sensing
  - Industrial inspection

Grayscale Image of a Cat (Chelsea)

Grayscale Bit-Depth Representation



- A **grayscale image** uses k bits per pixel.
- The number of gray levels is:

$$L = 2^k$$

- Examples:
  - $k = 1 \;\Rightarrow\;$ 2 levels (black and white).
  - $k = 2 \;\Rightarrow\;$ 4 levels.
  - $k = 4 \;\Rightarrow\;$ 16 levels.
  - $k = 8 \;\Rightarrow\;$ 256 levels (standard).
- Increasing k provides smoother tonal transitions and better detail.
- Lower bit-depths cause **posterization**, where smooth gradients appear as sharp steps.

RGB Color Image Representation

- An **RGB color image** uses three channels per pixel:

$$f(x,y) = \big(f_R(x,y), f_G(x,y), f_B(x,y)\big)$$

where f_R, f_G, f_B correspond to red, green, and blue intensity values.

- Each channel typically has $k = 8$ bits:

$$f_c(x,y) \in [0,255], \quad c \in \{R,G,B\}$$

giving $256^3 \approx 16.7$ million possible colors.

Additive Color Mixing:

- $(255,0,0) \to$ Pure Red
- $(0,255,0) \to$ Pure Green
- $(0,0,255) \to$ Pure Blue
- $(255,255,255) \to$ White
- $(0,0,0) \to$ Black

Applications:

- Photography and video
- Computer graphics
- Display and visualization

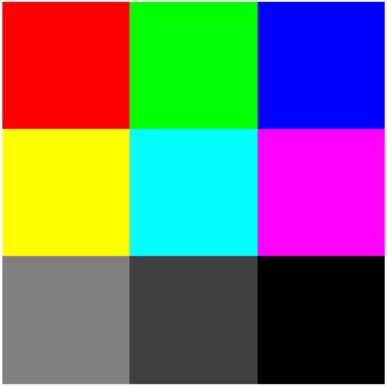
K-bit Color Image Representation (Composite)

- Each pixel stores 3 channels: Red, Green, Blue.
- Bit-depth k controls the range of intensity values per channel:

$$f_c(x, y) \in \{0, 1, \ldots, 2^k - 1\}, \quad c \in \{R, G, B\}$$

- Higher k $\longrightarrow$ smoother gradients and more realistic images.
- Lower k $\longrightarrow$ fewer possible colors, producing banding or posterization.

How RGB Values are Stored (Example)



3x3 color patch image generated from RGB values.

$$\begin{bmatrix} (255,0,0) & (0,255,0) & (0,0,255) \\ (255,255,0) & (0,255,255) & (255,0,255) \\ (128,128,128) & (64,64,64) & (0,0,0) \end{bmatrix}$$

- Each entry is a pixel stored as (R, G, B).
- Values range from 0 to 255 for 8-bit images.
- The 3 channels together define the final pixel color.

Explanation: How RGB is Stored

- A digital image is represented as a 3D array (tensor):

$$f(x,y) = \big(f_R(x,y), f_G(x,y), f_B(x,y)\big)$$

  where each $f_c(x,y)$ is an integer in $[0,255]$ (for $k=8$ bits).
- For an image of size $M \times N$, storage requires:

$$M \times N \times 3 \quad \text{values}$$

  (the factor 3 corresponds to R, G, and B channels).
- Example: a 1920×1080 HD image:

$$1920 \times 1080 \times 3 \approx 6.2\,\text{million values}$$

- Each value usually occupies 1 byte (8 bits), so total memory:

$$1920 \times 1080 \times 3\,\text{bytes} \approx 6.2\,\text{MB}$$

Thank You!

Thank you for your attention!

Questions or Comments?

CMSC 178IP - Digital Image Processing