Week 8

Video Basics

Video Analysis

- Object Detection/Recognition
- Motion Detection/Recognition
- Activity Detection/Recognition

Video is a sequence of Images displayed at certain rate to stimulate motion!

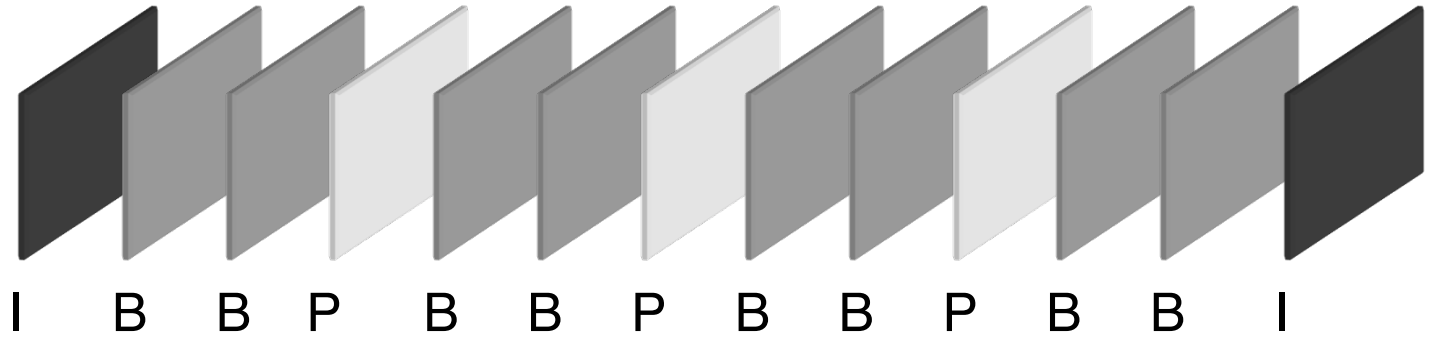
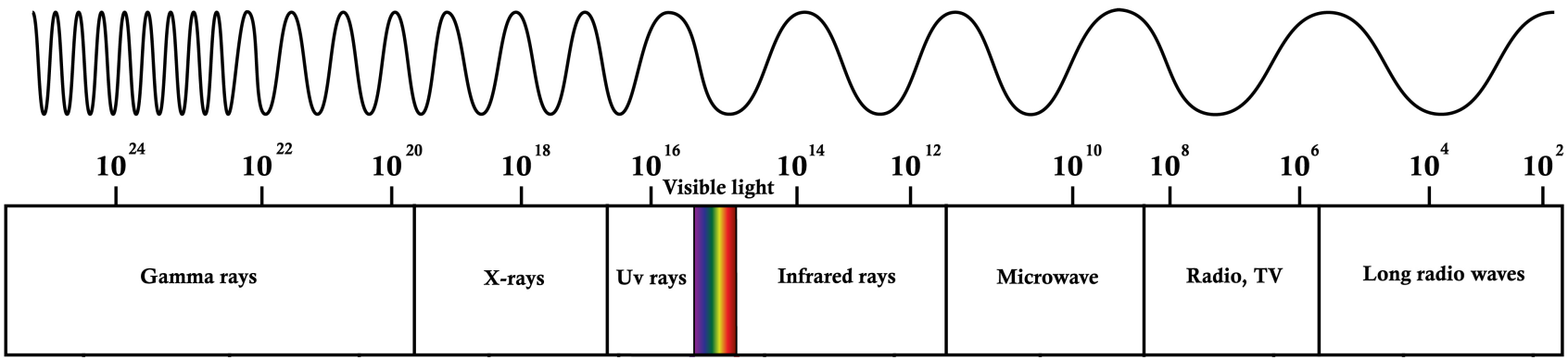
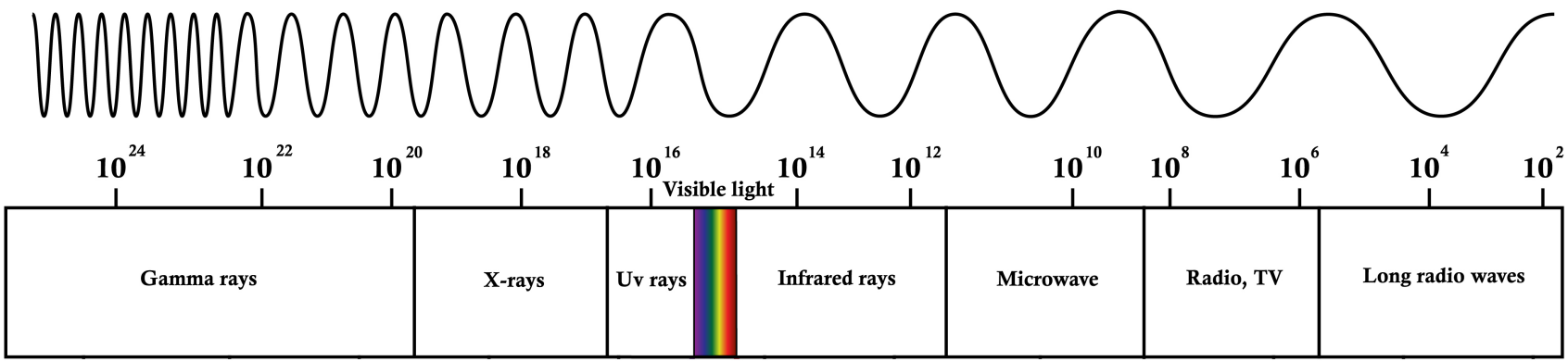


Image acquisition is the process of capturing EM energy radiated/passed by objects!



- Gamma-Ray Imaging
- X-Ray Imaging
- Infrared Imaging
- Visible Spectrum Imaging

Image source: https://s-media-cache-ak0.pinimg.com/originals/9a/19/9e/9a199eb6c2b084694f3f0a36b71b9269.jpg



Visible light is an electromagnetic wave in the 400 nm - 700 nm range!

Image source: https://s-media-cache-ak0.pinimg.com/originals/9a/19/9e/9a199eb6c2b084694f3f0a36b71b9269.jpg

Image Sensing

- Sensors are used to measure reflected energy
- Light is converted to voltage
- ADC is used to convert analog to digital voltage

Single sensor?

Move the sensor over the surface and measure the reflected energy!

Sensor Strip

Move the strip over the surface!

Example?

Sensor Array

- A array of sensors captures the light
- No mechanical movement
- CMOS and CCD are popular sensors

Sampling and Quantization

Digitizing the coordinate values is called sampling!

Digitizing the amplitude values is called quantization!

Image Terminology

- Pixels -- picture elements in digital images
- Image Resolution -- number of pixels in a digital image
  - –width x height (e.g., 640X480)
- Bits/pixel – also contributes to the quality of the image

Image Representation

24-bit color Image

Size=786 kb

512*512 pixels

8-bit Gray Image

Size=263 kb

512*512 pixels

1-bit BW Image

Size=33 kb

512*512 pixels

Which image is bigger size?

Size=786 kb

Size=263 kb

8-bit color-map Image

Size=263 kb

512*512 pixels

Color Look-up Tables

- The idea is to store only the code value for each pixel.
- If a pixel stores the value 25, the meaning is to go to row 25 in a color look-up table (LUT).

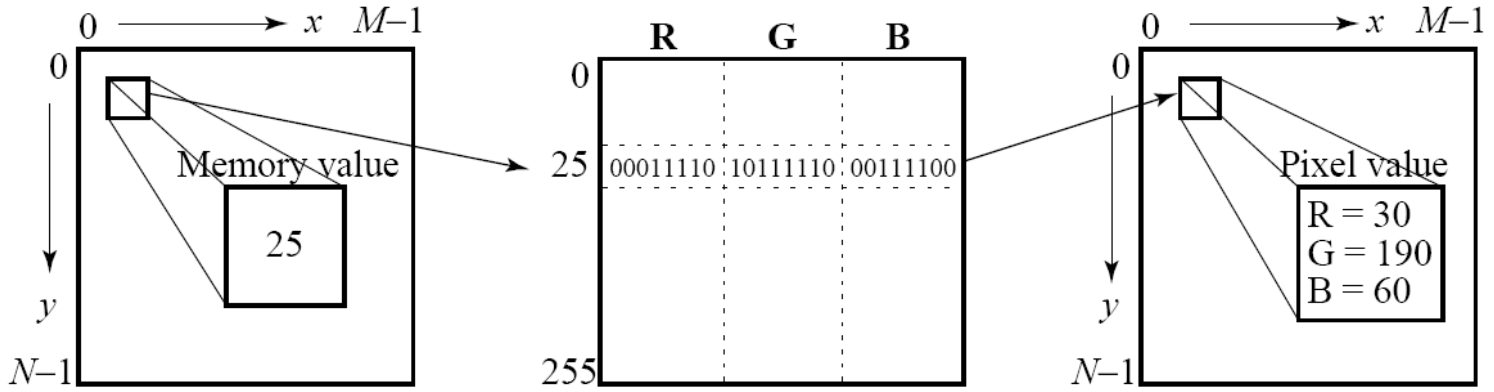


Image File Formats

- Some formats are restricted to particular hardware / operating system platforms.
- There are applications that convert formats from one system to another.
- Most image formats incorporate compression, lossless or lossy.

Popular Formats

- 8-bit GIF: one of the most important formats because of its historical connection to the WWW and HTML markup language as the first image type recognized by net browsers.
- JPEG: currently the most important common file format.
- PNG: most popular lossless image format.
- TIFF: flexible file format due to the addition of tags.
- EXIF: allows the addition of image metadata.
- PS and PDF: vector based language, popular in publishing and academia

Microsoft Formats

- Vectored: WMF
- Non-vectored: BMP

Color Models

- RGB
- YCbCr
- YUV
- YIQ
- CMY
- HSV

CMY - Cyan, Magenta, Yellow

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

RGB to CMY

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

CMY to RGB

RGB to YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

RGB to YIQ

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595879 & -0.274133 & -0.321746 \\ 0.211205 & -0.523083 & 0.311878 \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

RGB to YCbCr

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

Video Analysis

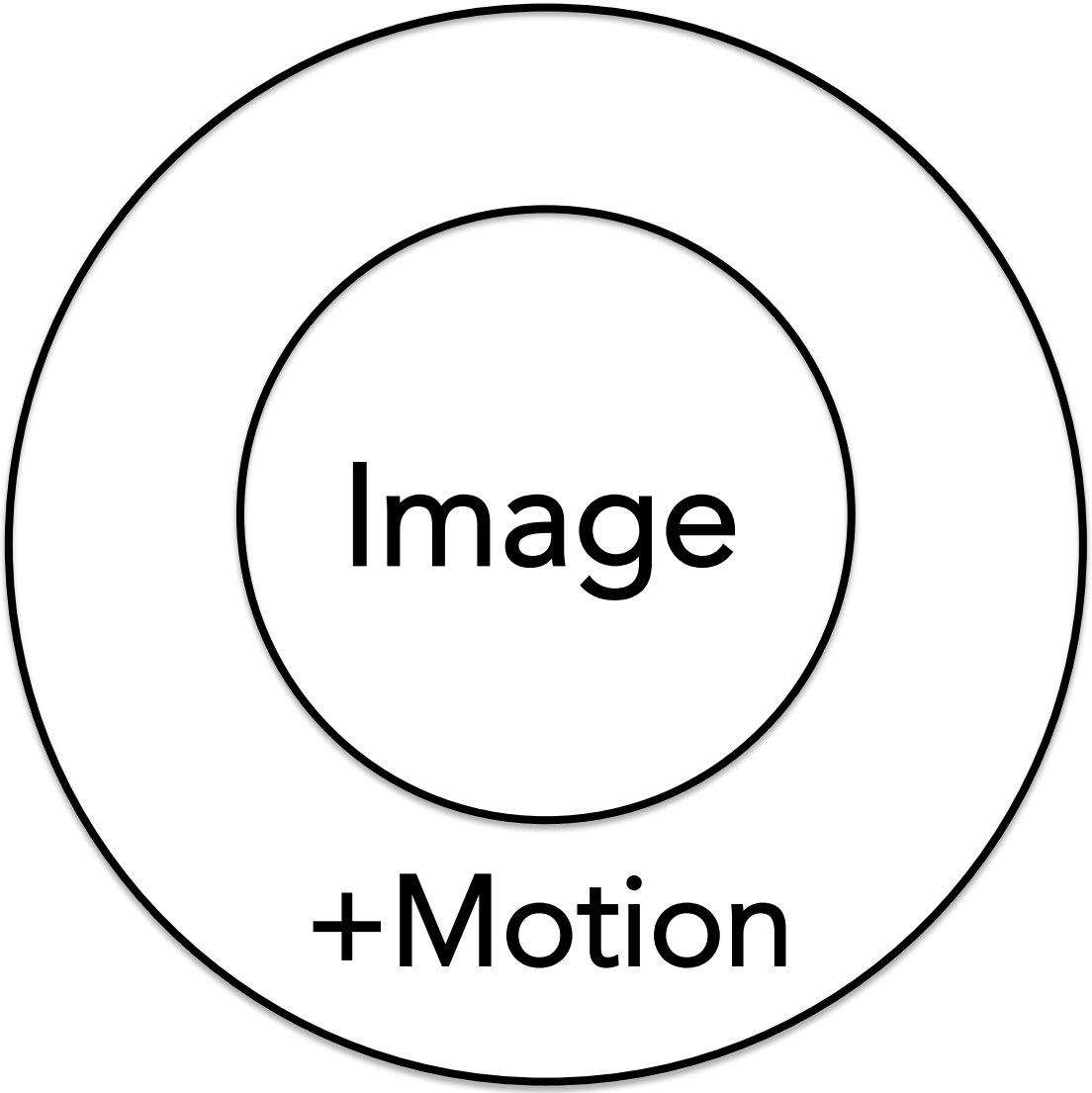


Image Analysis

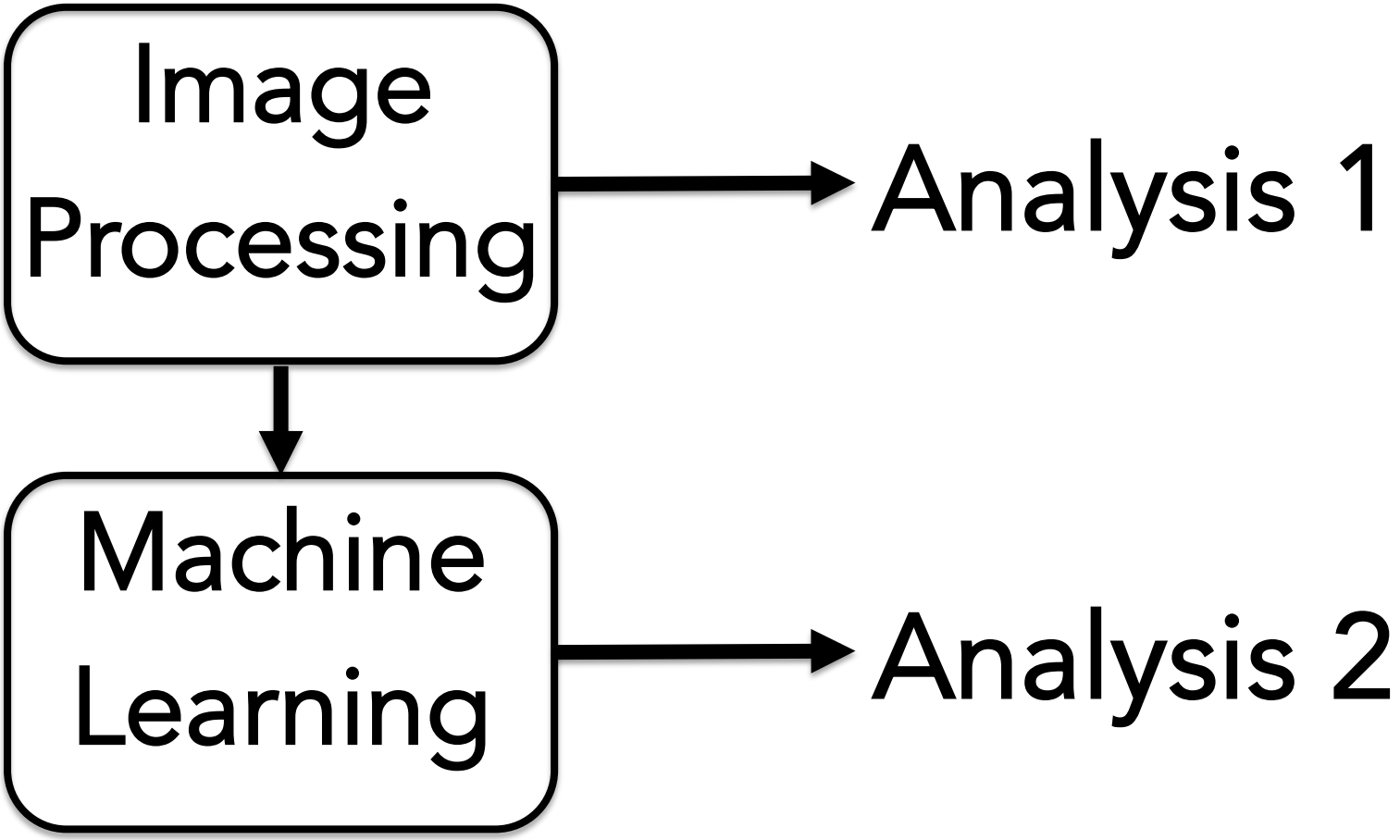
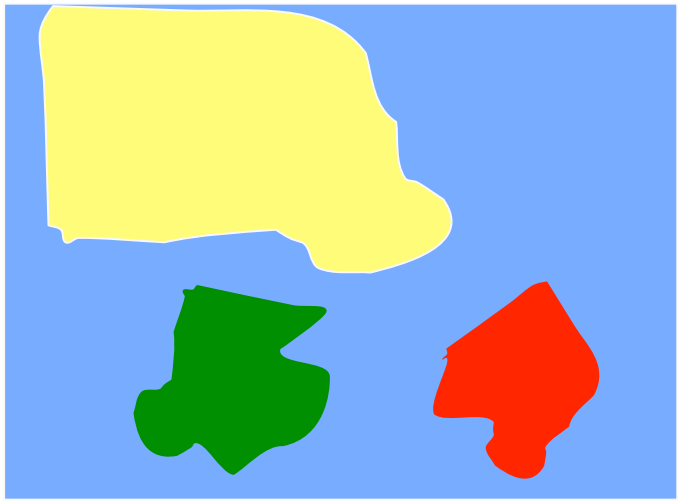


Image Segmentation

Partitioning an image into its constituent regions!



Thresholding

If above threshold than region 1 otherwise region 2

Generally Histogram of intensities is used to choose the Threshold

- Basic Global Thresholding
- Otsu's Method
- Multiple Threshold
- Variable Thresholding

Region Growing

Start with one pixel and grow it by adding adjacent pixels until adjacent pixels are dissimilar!

Which pixels are “similar”?

Region Growing Results on Lena

Usually statistical tests are used to decide which pixels to add to the region!

E.g., mean and variance of pixels currently in the region!

If similarity is too low, start a new region!

Clustering-Based

1. Start with initial k centroids
2. Partition the pixels into K clusters
3. Calculate distortion in each cluster (e.g. least-squares error)
4. Re-calculate cluster centroids and repeat 2-3 until distortion is below threshold

Clustering-Based

- Use K-means to find the clusters
- Use predefined K or hierarchical partitioning
- In hierarchical clustering, keep creating new clusters until distortion is minimized

Some Clustering Methods

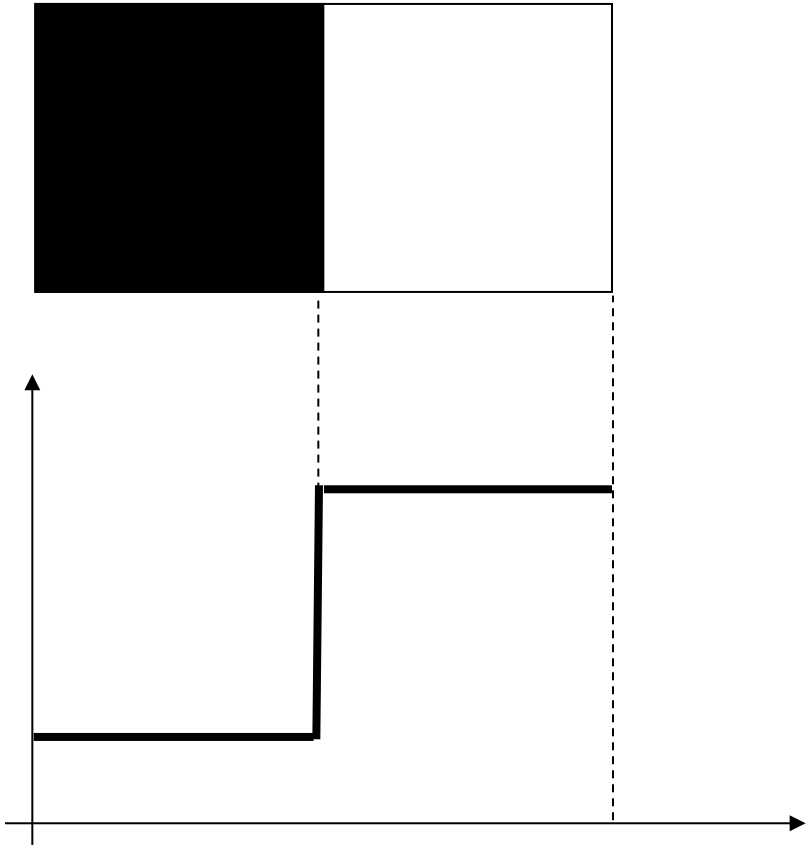
- K-means Clustering and Variants
- Isodata Clustering
- Histogram-Based Clustering
- Graph-Theoretic Clustering
- Self Organizing Maps

Edge Based Segmentation

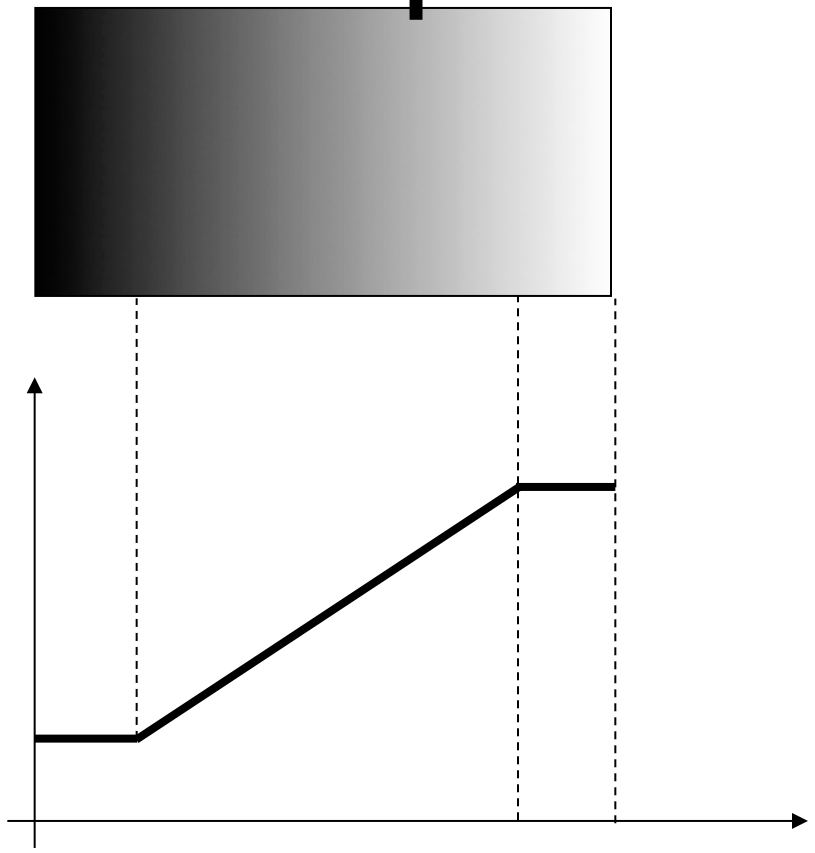
Edge Detection

81	82	26	24
82	33	25	25
81	82	26	24

Ideal



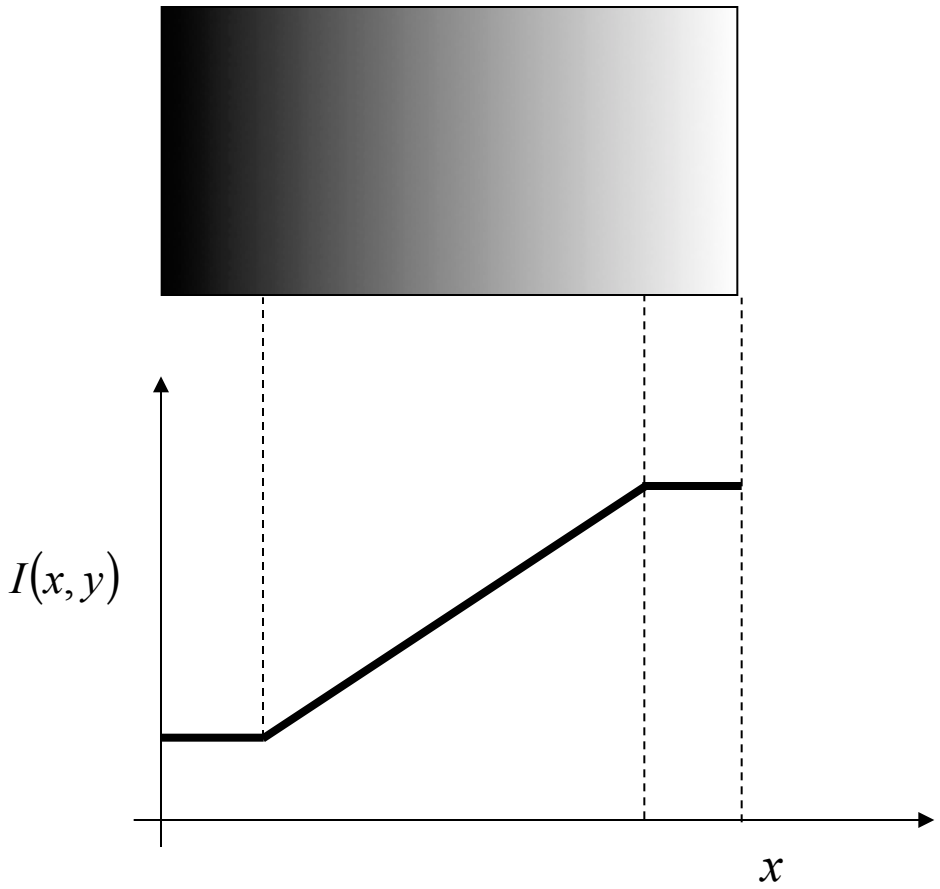
Ramp

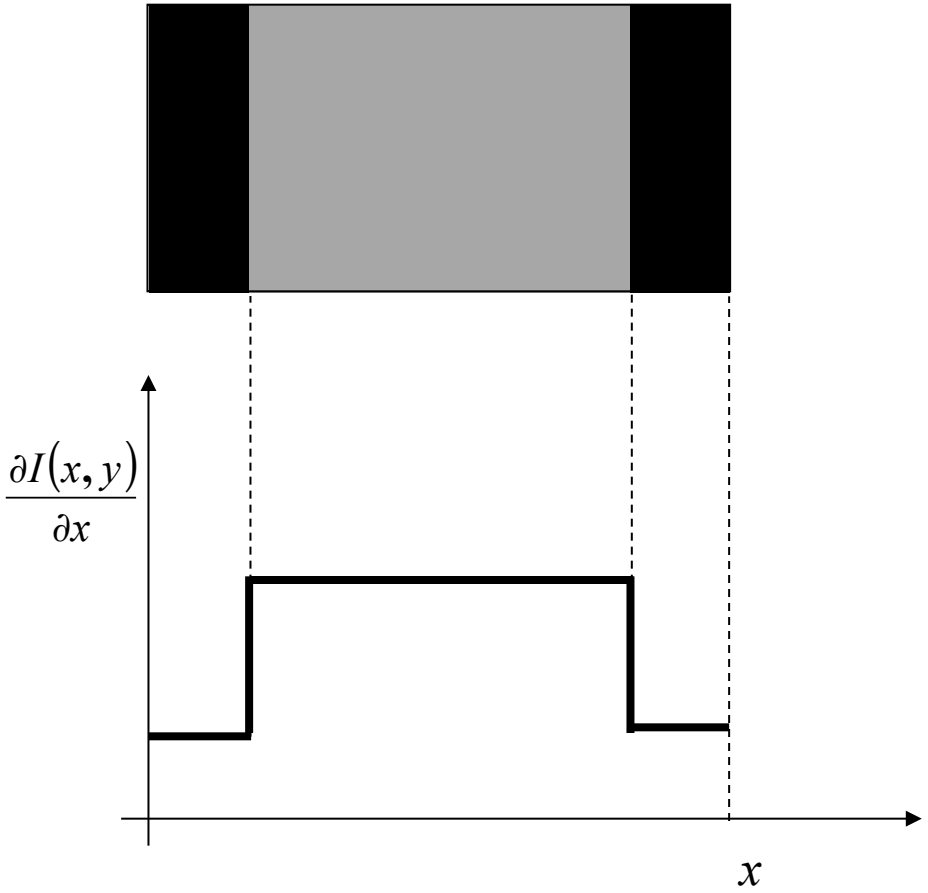


Use derivatives to detect slow (ramp) change!

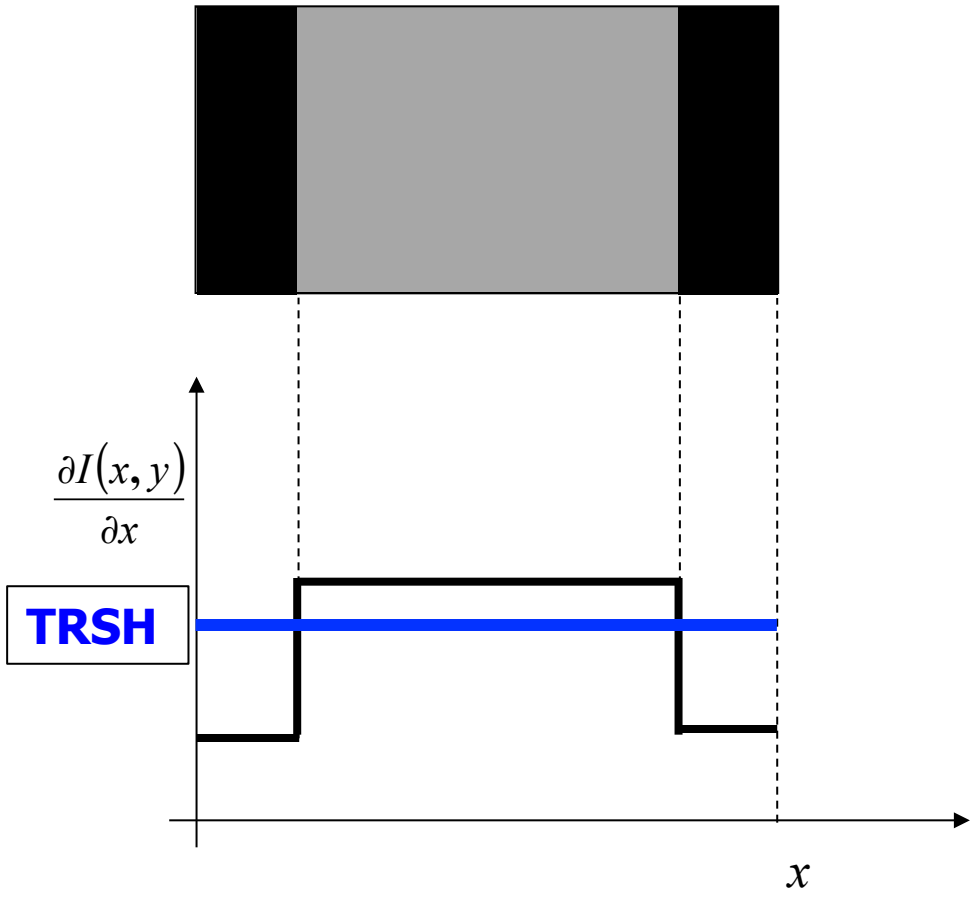
Original

First Derivative





$$\frac{\partial I(x,y)}{\partial x} > TRSH \Rightarrow Edge \quad Detected$$



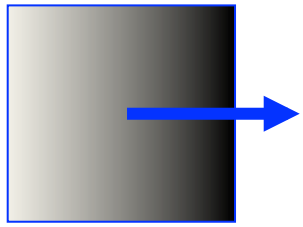
First Derivative is also Called Gradient!

$$\overline{\nabla I} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial I(x,y)}{\partial x} \\ \dfrac{\partial I(x,y)}{\partial y} \end{bmatrix}$$

Magnitude $\nabla I = \left\| \overline{\nabla I} \right\| = \sqrt{\left[G_x^2 + G_y^2 \right]}$

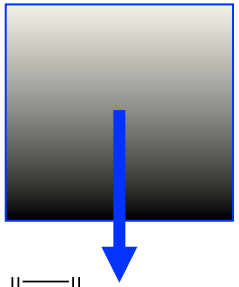
Direction $\theta(x, y) = \tan^{-1} \begin{pmatrix} G_x \\ G_y \end{pmatrix}$

Gradient Represents The Direction of Strongest Change



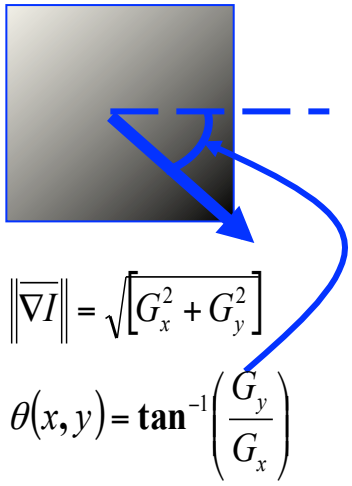
$$\left\|\overline{\nabla I}\right\| = G_x$$

$$\theta(x,y) = 0$$



$$\left\|\overline{\nabla I}\right\| = G_y$$

$$\theta(x,y) = -\frac{\pi}{2}$$

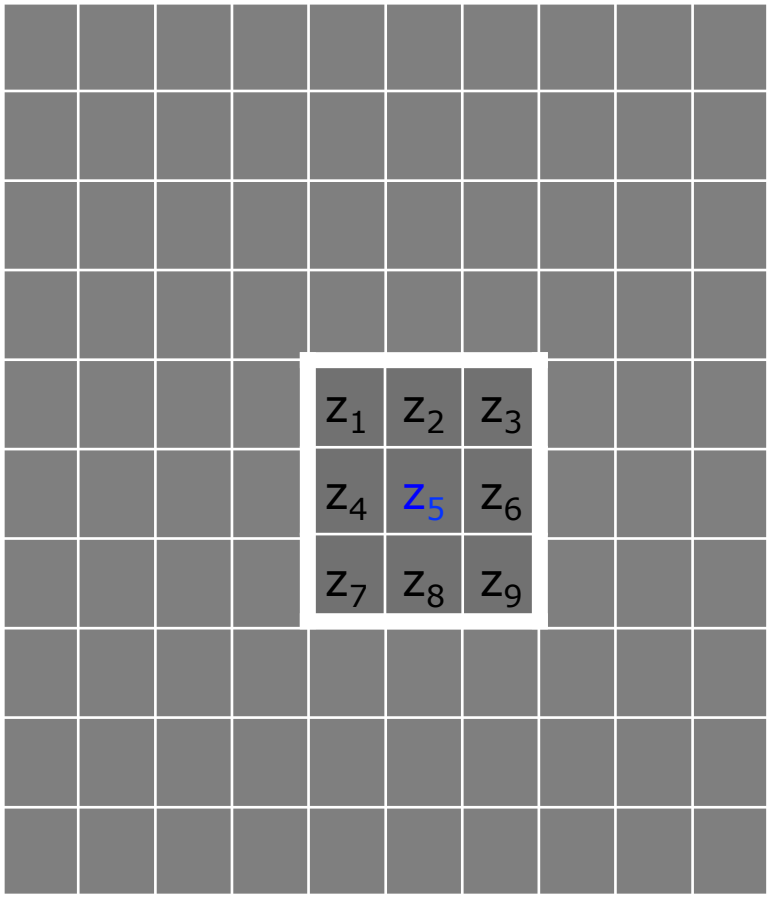


$$\left\|\overline{\nabla I}\right\| = \sqrt{\left[G_x^2 + G_y^2\right]}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

How do we calculate Gradient of pixel intensities?

Look at 3X3 neighborhood around a pixel



The Prewitt Edge Detector

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

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

$$G_x \approx (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

$$G_y \approx (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

The Sobel Edge Detector

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

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

$$G_x \approx (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_y \approx (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

Canny Edge Detector

- Smooth the image with Gaussian filter
- Determine gradient magnitude at each pixel
- Non-maximum suppression for thinning
- Uses two thresholds, th1>th2
  - Strong edge: Pixels with gradient magnitude > th1
  - Weak edge: Pixels with gradient magnitude < th1 and >th2, but connected to a strong edge

original image (Lena)

norm of the gradient

thinning

thresholding