MITMECHE

Introduction to Computer Vision

Department of Mechanical Engineering
Massachusetts Institute of Technology

Computer Vision

- Image data representation
- Image statistics
- Color models
- Thresholding
- Image restoration
- Depth sensing
- Camera model
- Feature extraction
- Image topic classification
- Time series visual feedback
- April tags for robotic application
- Numerous application in robotics, autonomous driving

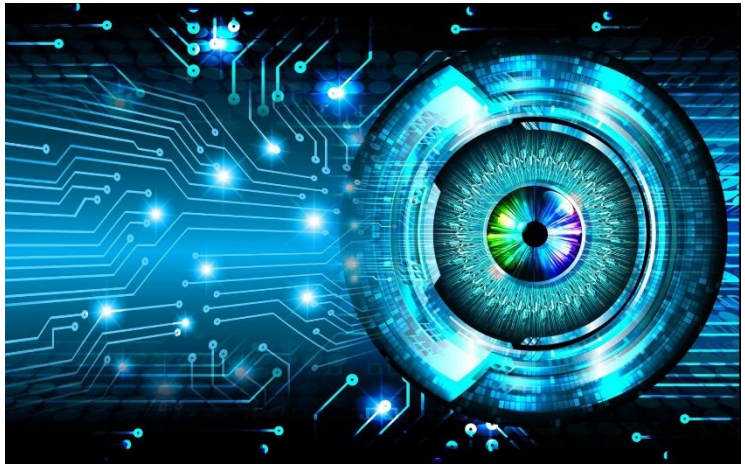
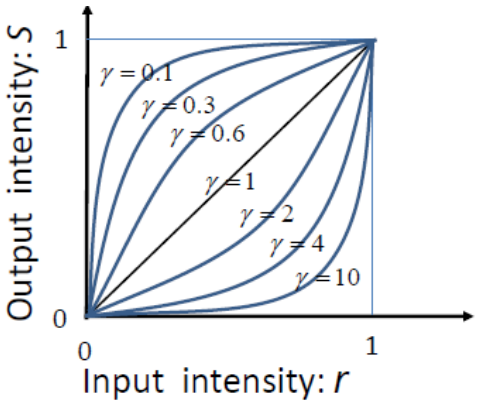


Image Representation and Transformation

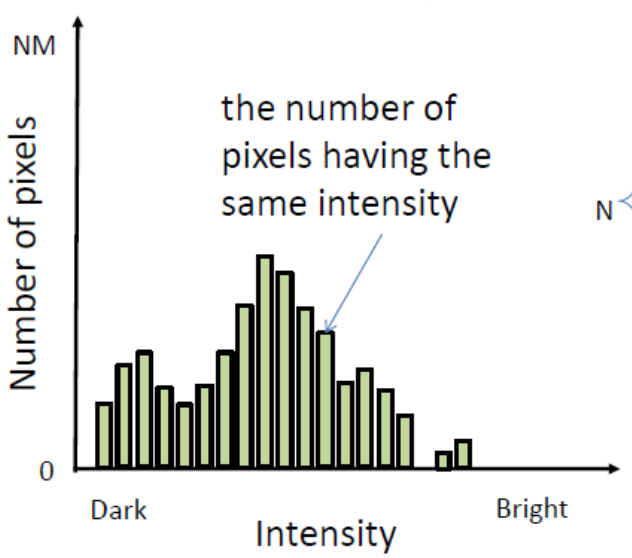
- Multi-dimensional matrix with color channel and spatial position`
- Bitmap file: large size, simple structure, uncompressed, lossless
- Joint Photographic Experts Group: lossy compression 24 bit color
- Graphics Interchange Format: LZW Lossless compression
- Portable Network Graphics: Open source, lossless compression
- Prefer to work with simple 2D * 3 color channel matrix data format
- Intensity Gamma transformation for contrast



Gamma Transformation $s = cr^{\gamma}$

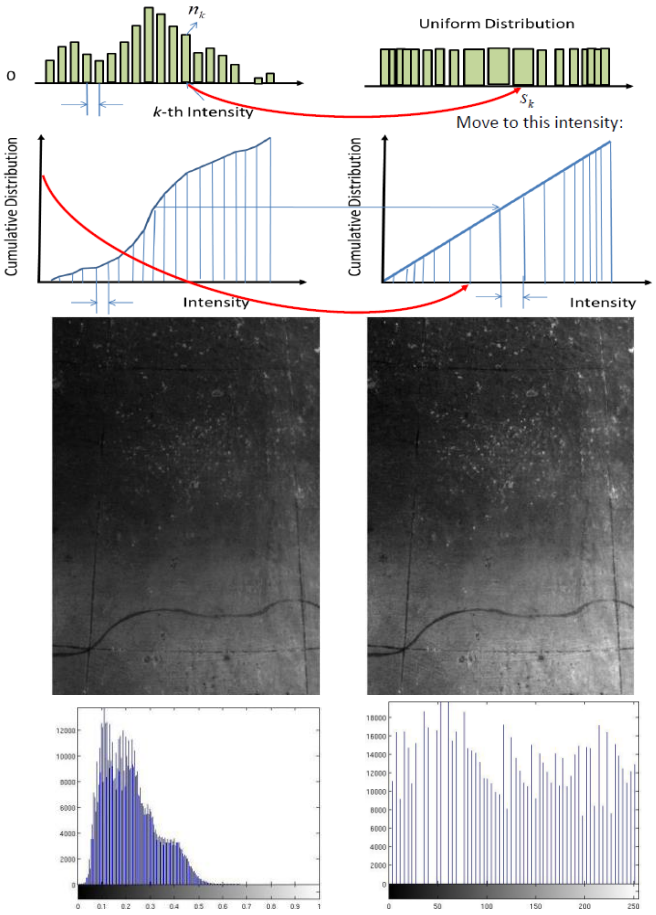
Histogram Processing

The histogram of a digital image indicates how the intensity of pixels distributes among the whole pixels.

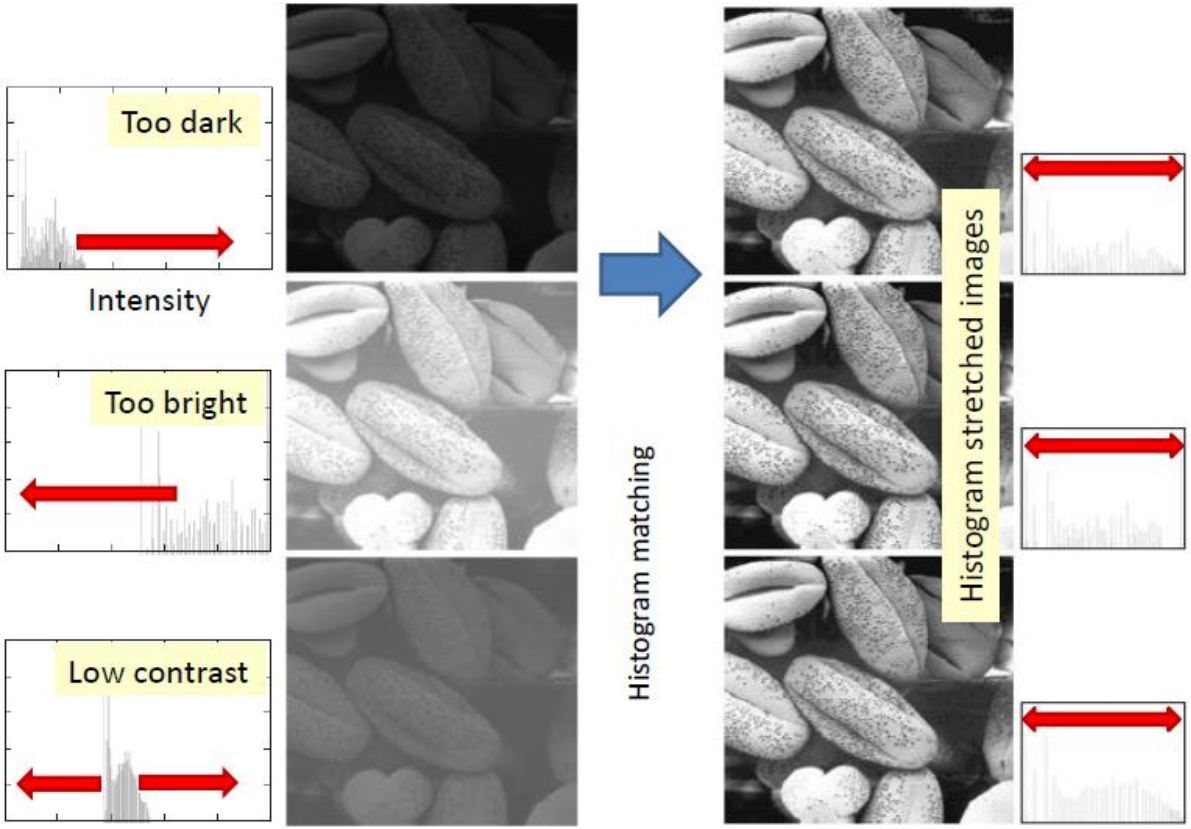


Histograms can be computed quite easily for digital images.

We learn from a histogram the darkness, brightness, and contrast of an image.

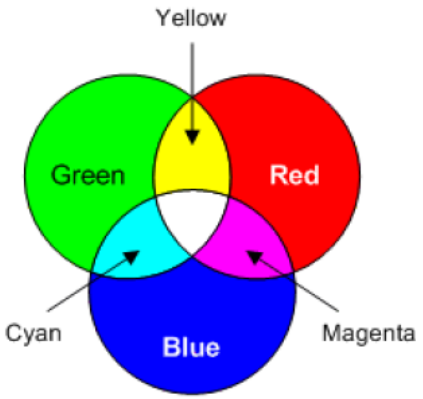


Histogram Processing



Color Model

Mixture of Light

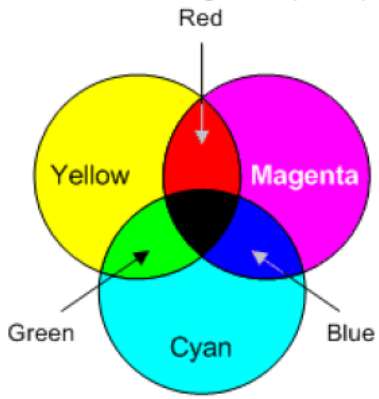


Red, Green, and Blue (RGB) – Primary colors of light (Additive Primaries)

Color Monitors and Video Cameras

Magenta, Yellow, and Cyan – Secondary colors of light (Subtractive primaries)

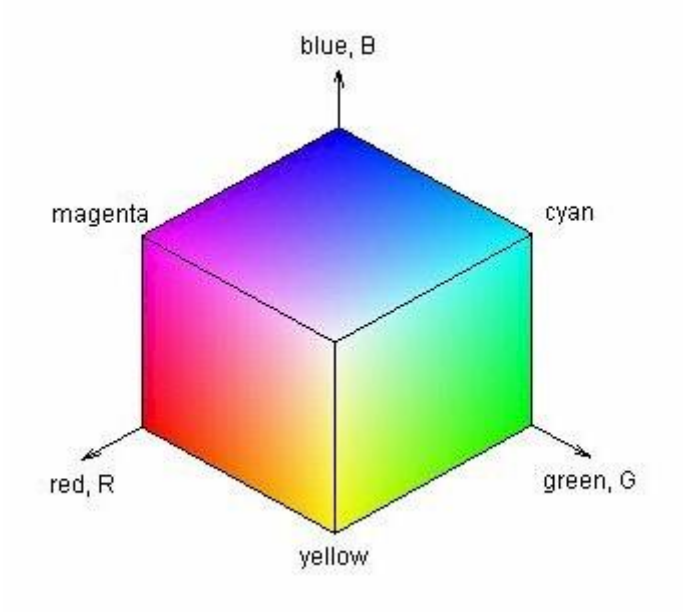
Mixture of Pigment (Paint)



Magenta, Yellow, and Cyan – Primary colors of pigment

Printers

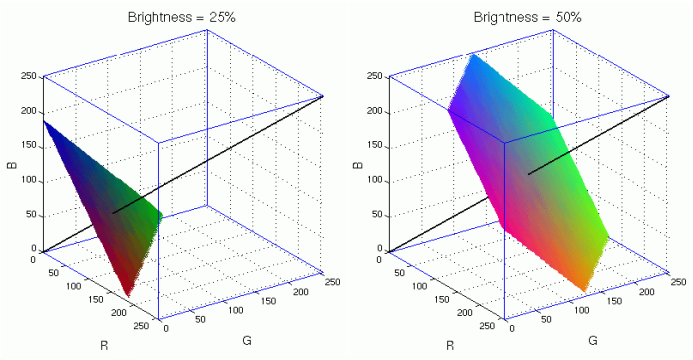
Red, Green, and Blue
Secondary colors of pigment

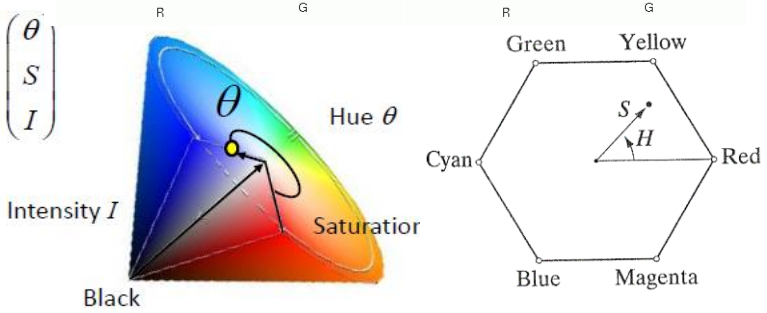


RGB color cube

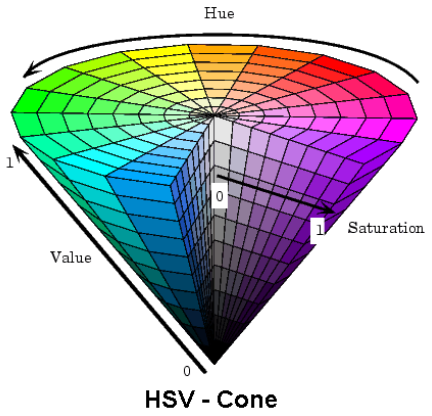
Color Model HIS, HSV

Hue, saturation, intensity





Hue, saturation, value



HSV - Cone

RGB to HSI Conversion

$$H = \begin{cases} \theta: & B \le G \\ 2\pi - \theta: & B > G \end{cases} \qquad \theta = \cos^{-1}\left(\frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right)$$

$$I = \frac{1}{3}(R+G+B)$$

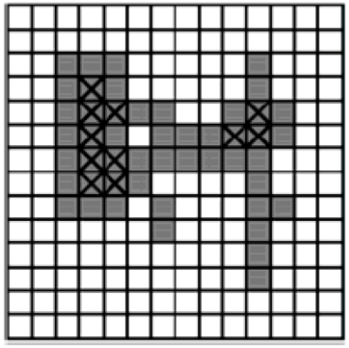
$$S = 1 - \frac{3}{(R+G+B)} \min(R,G,B)$$

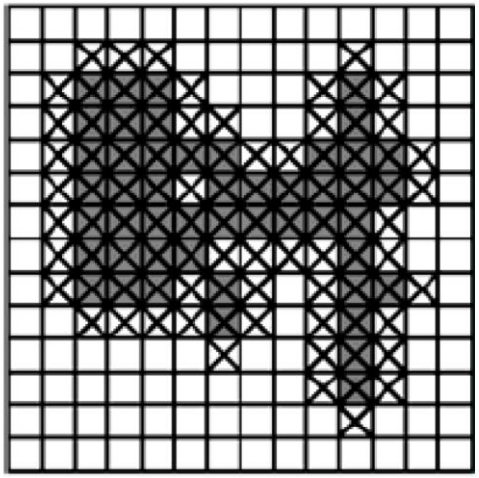
MITMECHE

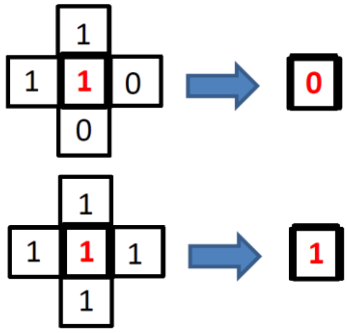
AsiaUS Business Connection

Workshop TIM THE BEAVER

Erosion and Dilation

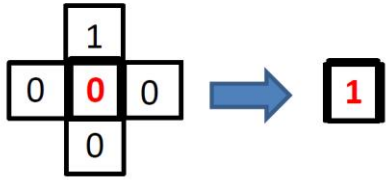




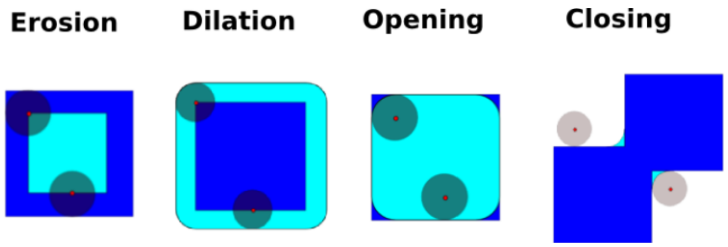


Erosion

(B)

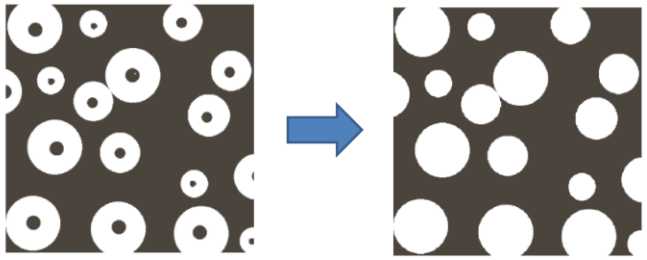


Dilation



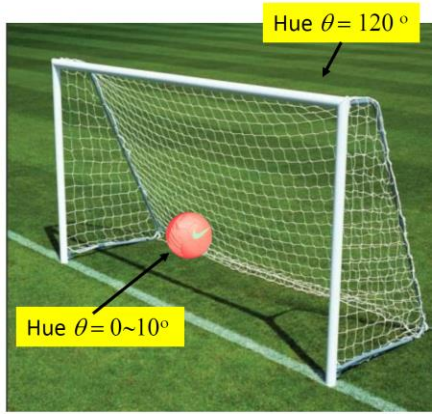
Opening = Erosion first, followed by dilation

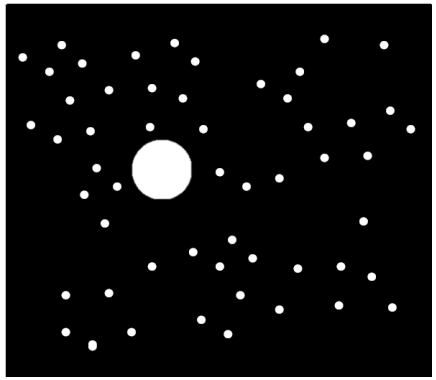
Closing = Dilation first, followed by erosion



Closing fills small holes.

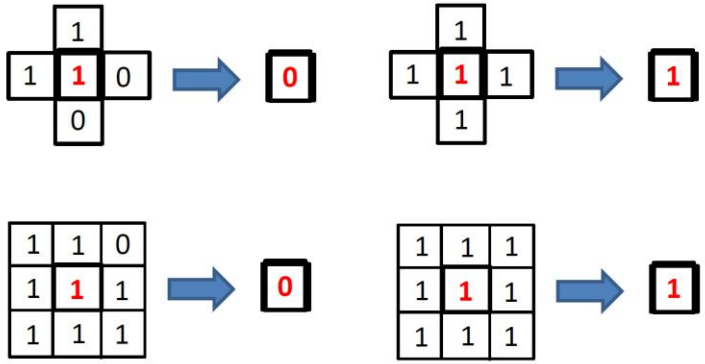
Color Thresholding and De-noising





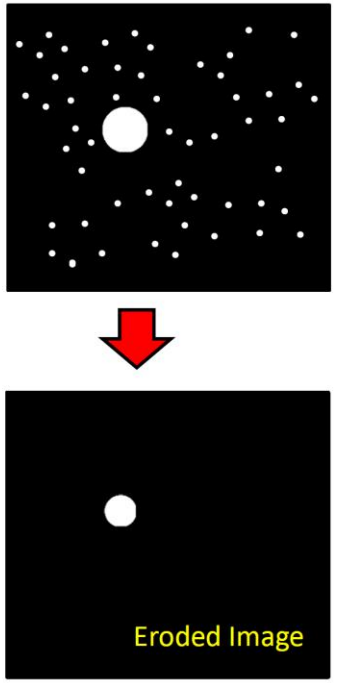
Simple rules for eliminating noise

- Spatial Filtering: Erosion, Opening;
  Unless all the adjacent 4 or 8 pixels are 1s, replace the middle pixel by 0.



Erosion can eliminate isolated dots. The segmented ball gets smaller.

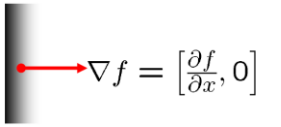
Difference Image

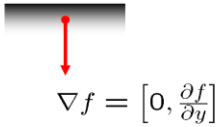


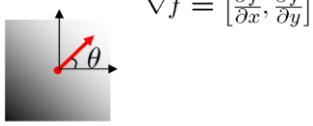
Kernel Based Edge Detection

- Originated from idea of gradient
- Kernel convolution

$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]$$







$$\theta = \tan^{-1}\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

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

Vertical

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

Horizontal

Prewitt Operator

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

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

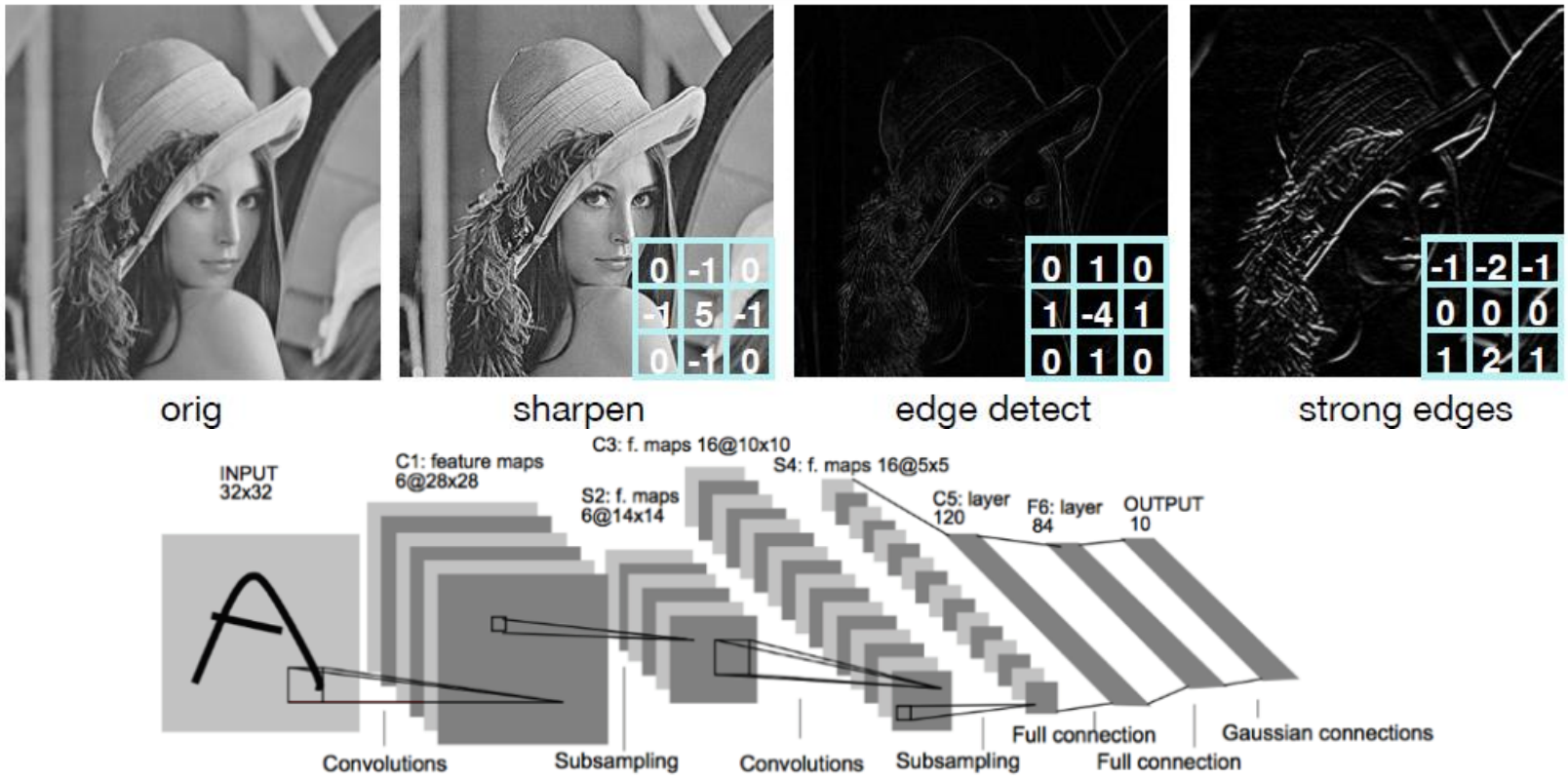
Sobel Operator

orig

edge detect

strong edges

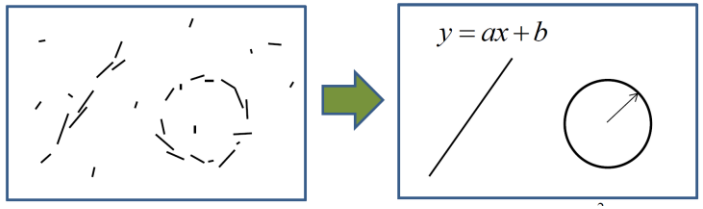
Convolutional Neural Network



CNN called LeNet by Yann LeCun (1998)

Hough Transformation

- Question: how to convert segmented edges to meaningful geometry?
- Utilize Hough transformation for parameterized primitive shapes

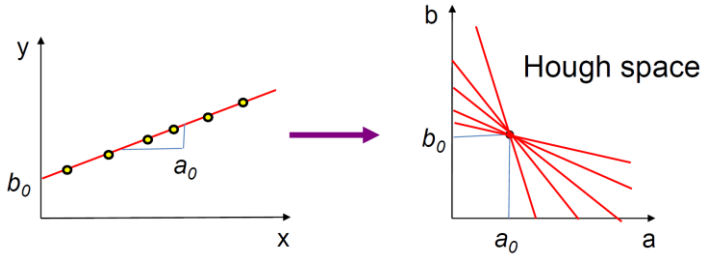


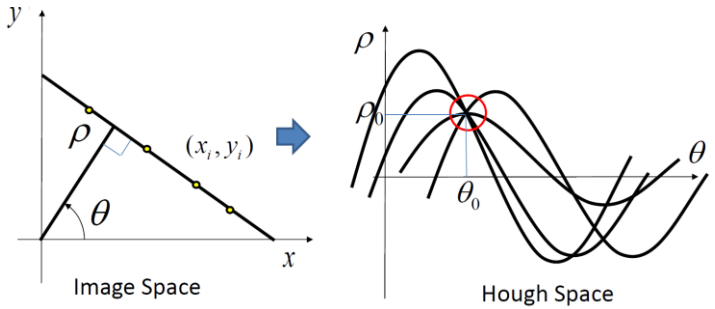
At each point of the (discrete) parameter space, find the number of lines that pass through it

– Use an array of counters

The more lines, the more edges are collinear in the image space.

Peaks in the counter array (bright points) in the Hough space found by thresholding.





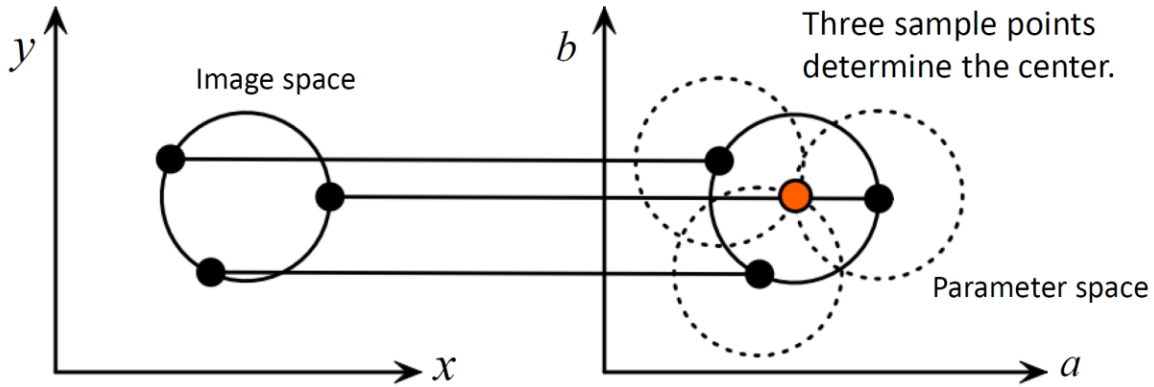
Hough Transformation for Circle

- Let R be a known radius, and (a, b) are unknown center coordinates.
- If a point at $(x_i \quad y_i)$ is on the circle, it must satisfy

$$x_i = a + R\cos\theta,\ y_i = b + R\sin\theta$$

- In the parameter space, a and b must be on a circle given by

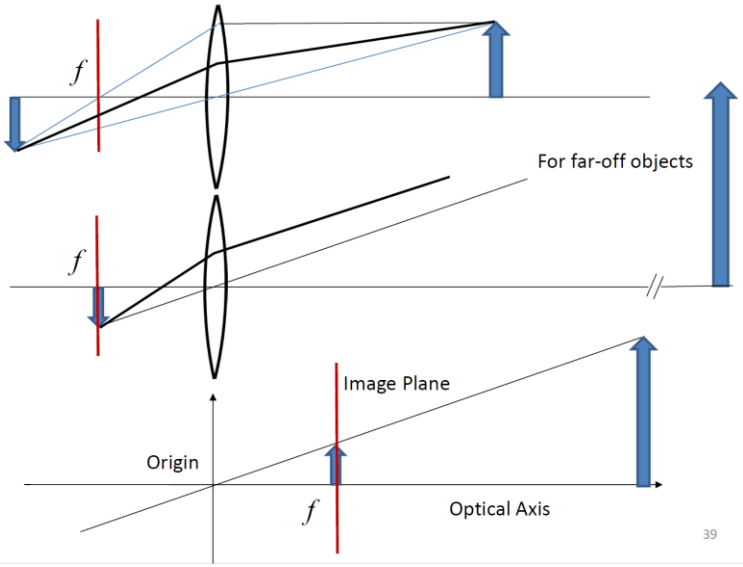
$$a = x_i - R\cos\theta,\ b = y_i - R\sin\theta$$



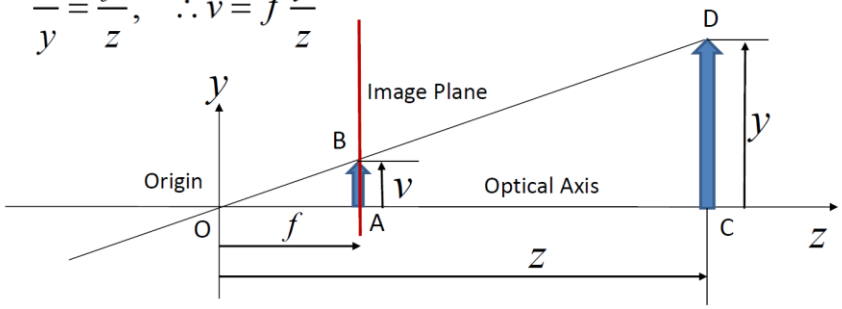
Each point in geometric space (left) generates a circle in parameter space (right). The circles in parameter space intersect at the (a, b) that is the center in geometric space.

Camera Model

- For traditional, we know image size (v) and focus length (f)
- With depth sensing (e.g. Kinect, Lidar), we also have depth (z)
- Can compute y directly given these information



$$\frac{v}{y} = \frac{f}{z}, \quad \therefore v = f\frac{y}{z}$$

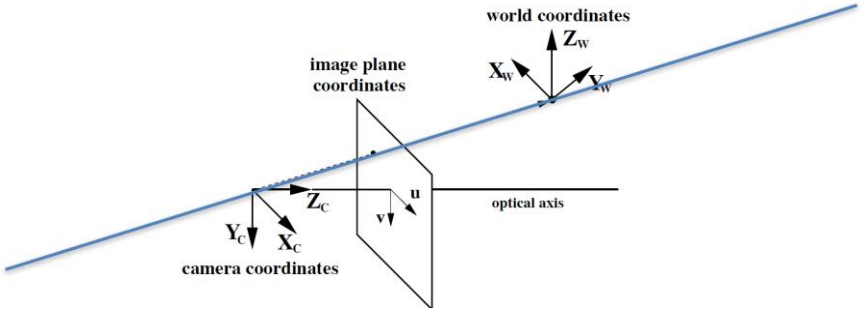


Kinect

640 x 480 pixels RGB Camera
640 x 480 pixels Infra-Red Depth Sensor
30 Hz USB 2

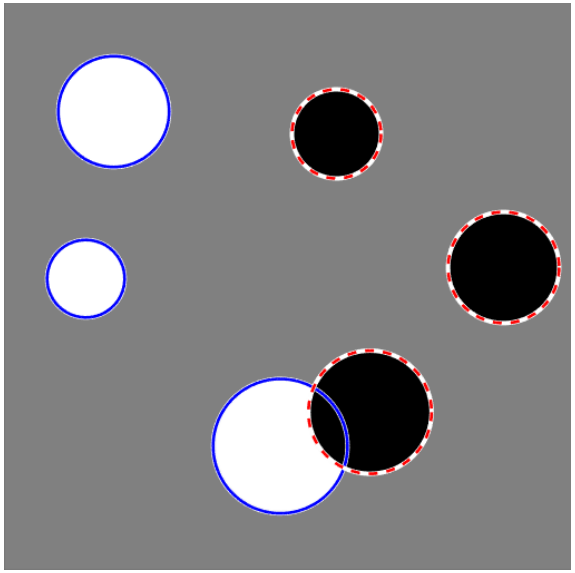
April Tags

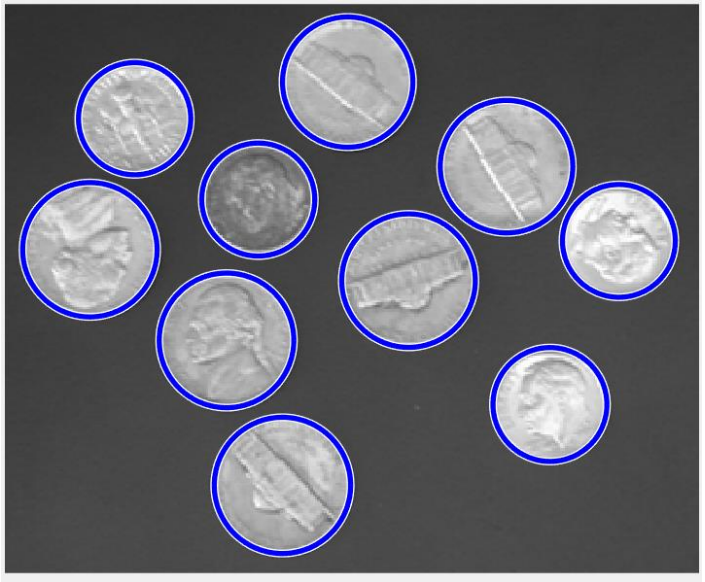
- Appearance similar to QR coe
- Widely used in mobile robot application
- 3D translation and rotation transformation matrix
- Determine orientation and position of camera to a plane



Hough Transformation Demo

- Matlab Hough transformation implementation





Thank You!