Winter 2019

DMET 901 – Computer Vision

Image Representation and Properties (1)

Seif Eldawlatly

Image Representation

- Static Image: A two variable function $f(x, y)$, where (x, y) are coordinates in a plane

- Dynamic image: A three variable function $f(x, y, t)$, where t is time (video frame)

- $f(x, y)$ represents the brightness (intensity) at location (x, y)

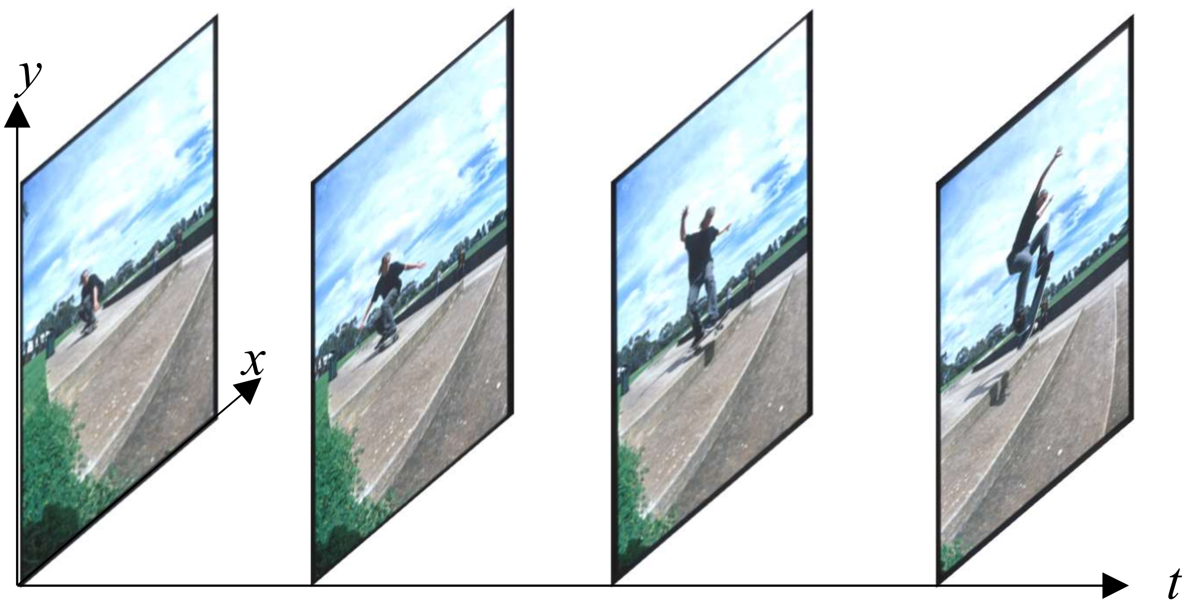


Image Representation

- Computerized image processing uses digital image functions represented by matrices

- The domain of the image function R

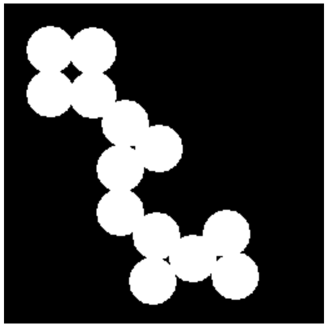
$$R = \{(x,y),\ 1 \le x \le x_m,\ 1 \le y \le y_n\}$$

  where x_m and y_n represent the maximal image coordinates.

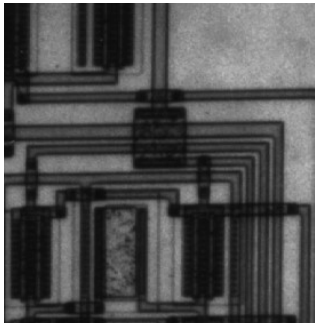
- In monochromatic Images, the lowest value (0) corresponds to black and the highest (1 or 255) to white

Image Formats

- Common Formats
  - Binary
    - 1 bit per pixel (0 or 1)



  - Intensity (Gray Scale)
    - 8 bits per pixel (0→ 255) or (0→ 1)



  - RGB (Red-Green-Blue)
    - 24 bits per pixel (8x3) (each 8 bits : 0 →255)

Image Formats

- RGB format: (Red-Green-Blue)

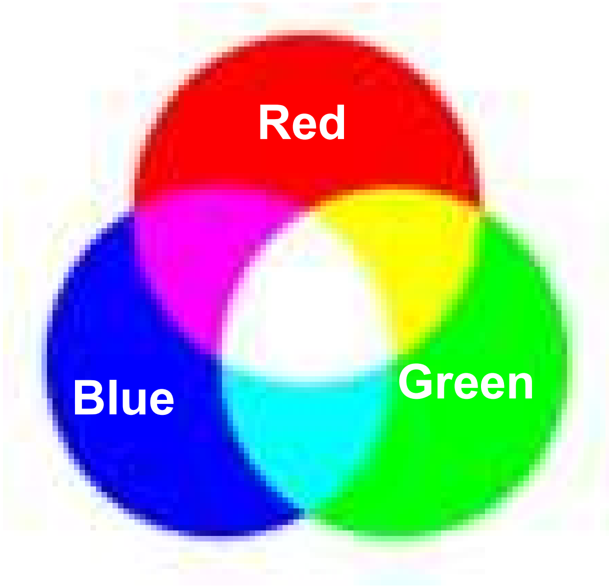
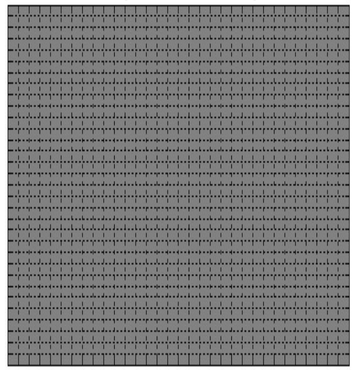


Image Digitization

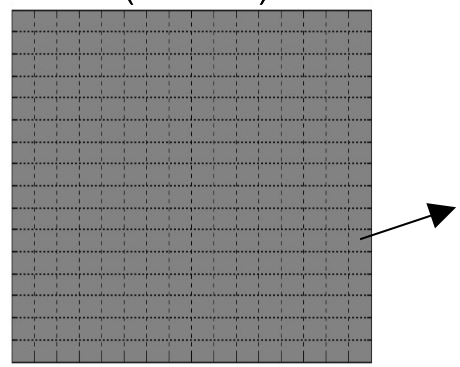
- Image digitization consists of two stages
  - Sampling: Converting the continuous function $f(x, y)$ to an M x N matrix

Original Scene

Sampling Matrix (32 x 32)



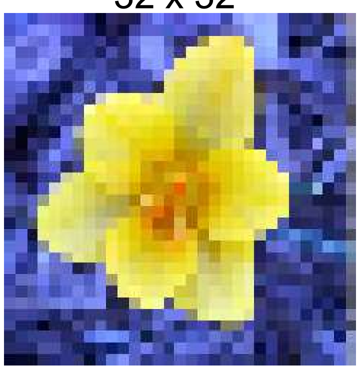
Sampling Matrix (16 x 16)



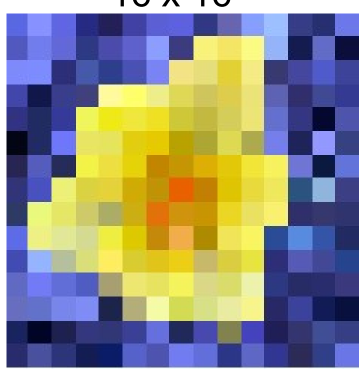
Pixel (for 2D)
Voxel (for 3D)

  - Quantization: Assigning to each continuous sample an integer value

32 x 32

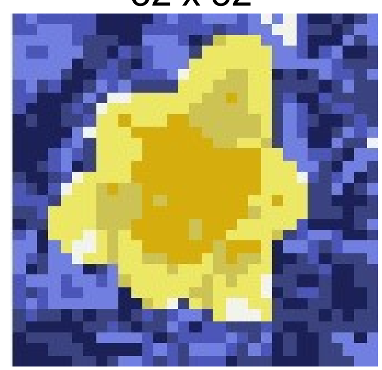


16 x 16



Each pixel is represented by 3 x 8 bits
→ Number of colors = 2^{24} = 16,777,216

32 x 32



Each pixel is represented by 3 x 3 bits
→ Number of colors = 2^{9} = 512

Zooming and Shrinking

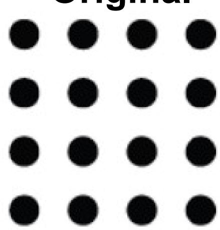
- Zooming versus Sampling
  - Sampling is done during capturing the image
  - Zooming (shrinking) is applied to the captured digital image

- Zooming

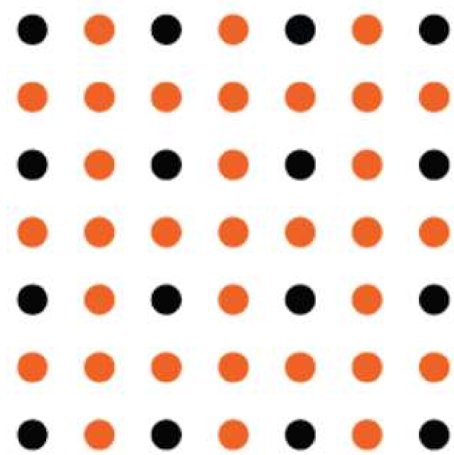
  1- Nearest-neighbor Interpolation:

  To double the size of an image, replace the brightness of each pixel (i, j) with four pixels (i, j), $(i, j+1)$, $(i+1, j)$, $(i+1, j+1)$ of the same brightness

Original



Zoom in (x2)



● : Original pixel

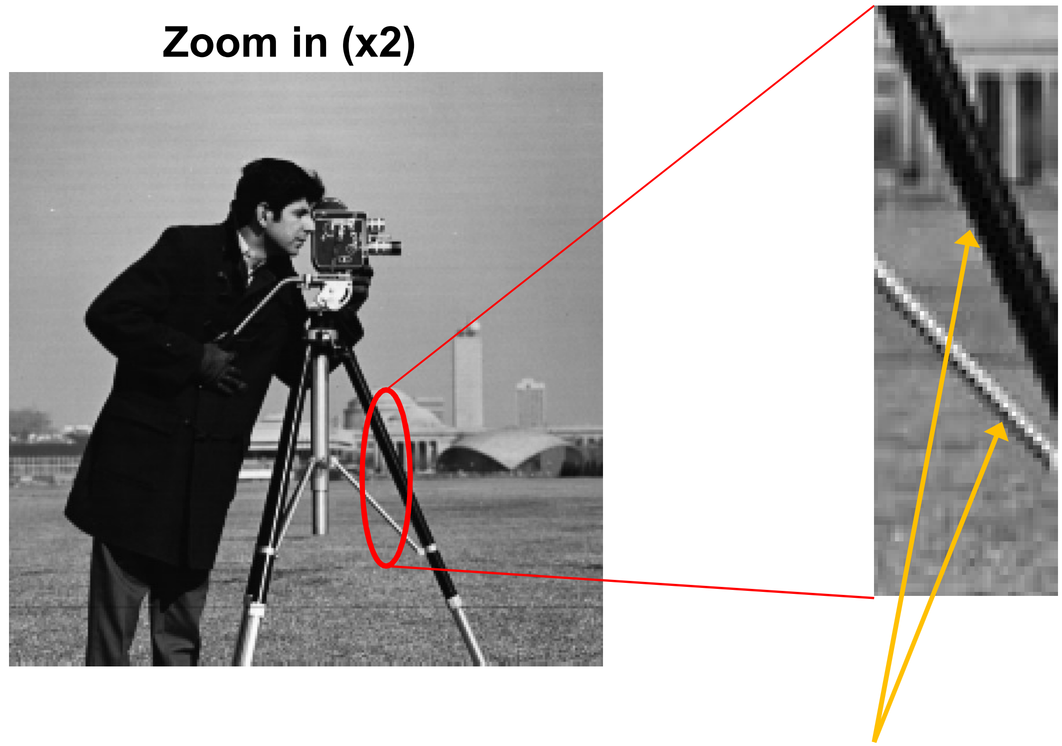
● : New pixel

Zooming and Shrinking

- Problem with nearest-neighbor interpolation

Original

Zoom in (x2)



Diagonal lines appear in a stairway-like shape

Zooming and Shrinking

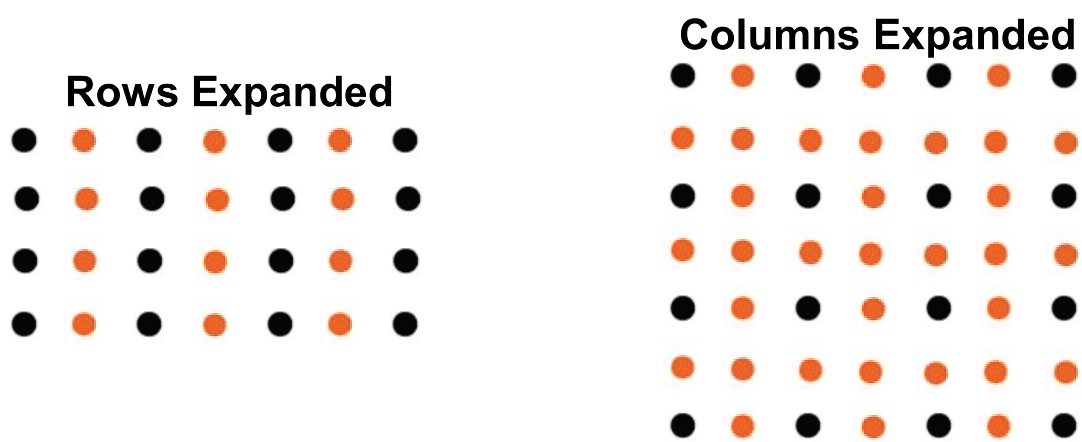
- Zooming

  2- Bilinear Interpolation

  a- Expand the rows as follows:

The brightness of each new pixel is the average of the brightness of the left and right original pixels

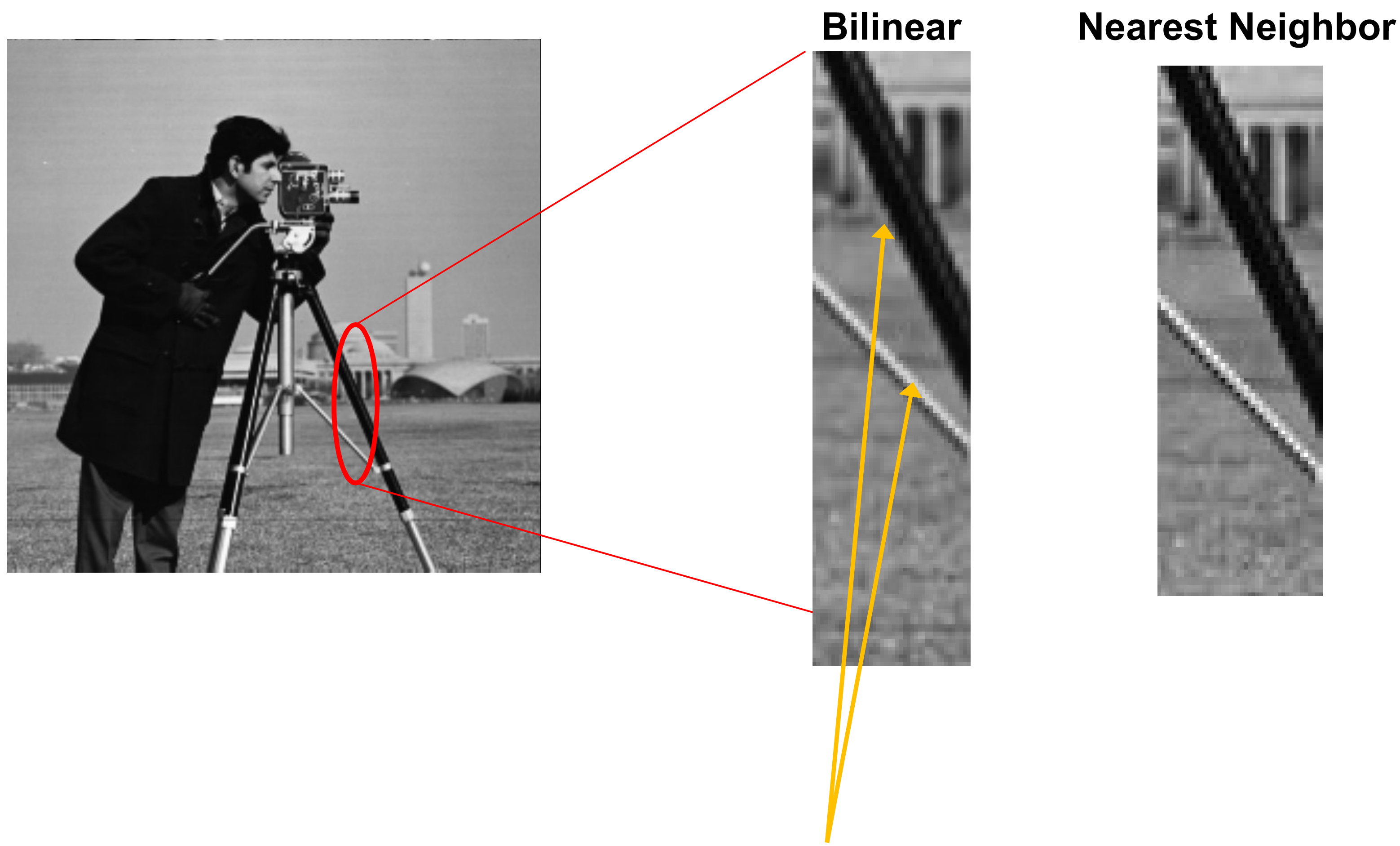
  b- Expand the columns as follows:



The brightness of each new pixel is the average of the brightness of the above and below pixels

Zooming and Shrinking

- Bilinear interpolation zooming

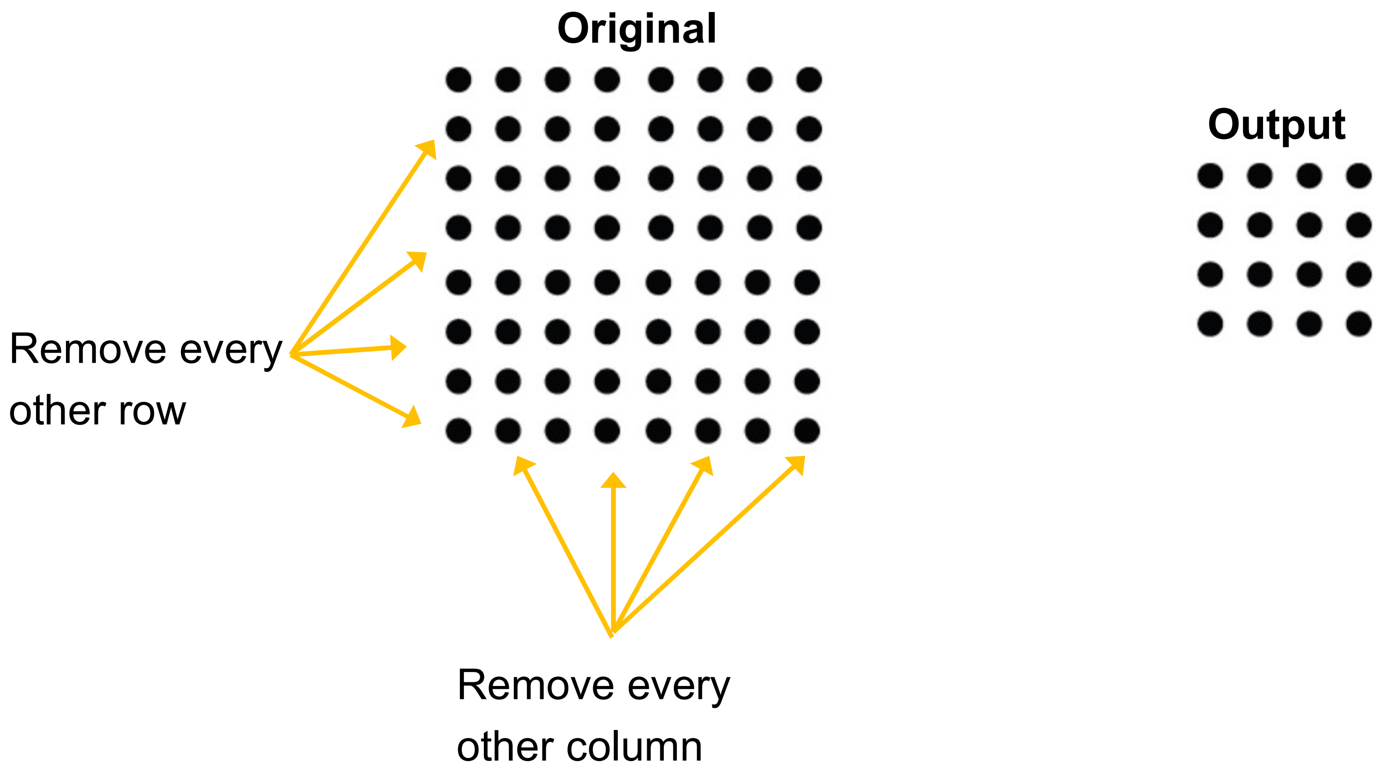


Bilinear interpolation has less stairway-effect

Zooming and Shrinking

- Shrinking

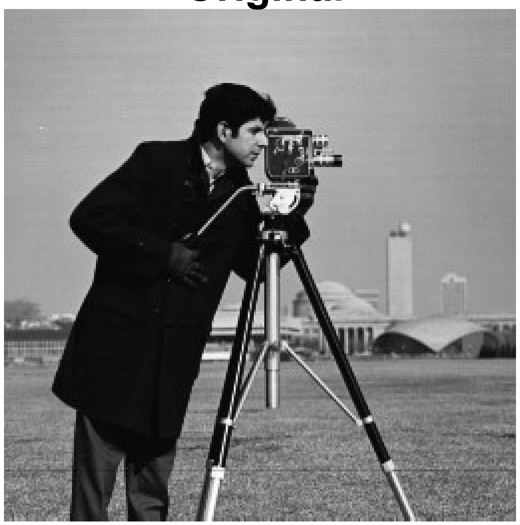
  To reduce the size of an image by half, remove every other row and every other column



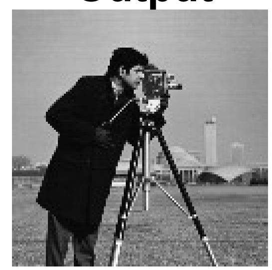
Zooming and Shrinking

- Shrinking example

Original

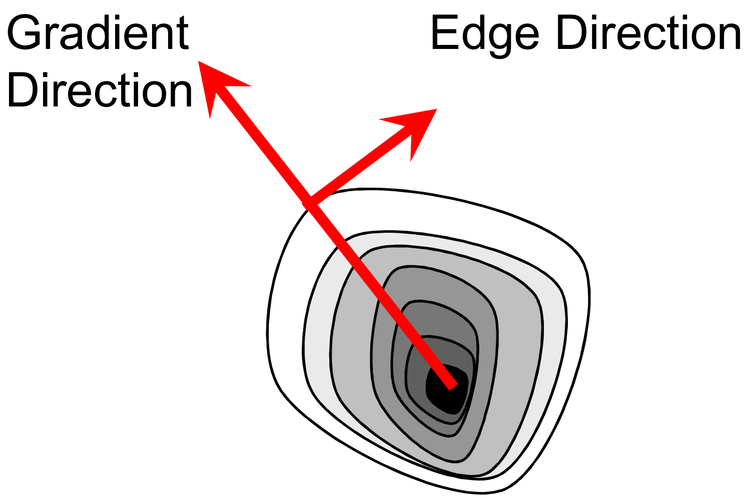


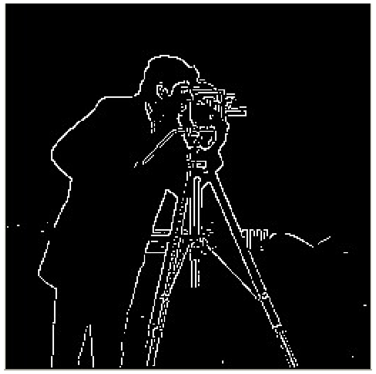
Output



Digital Image Properties: Edge

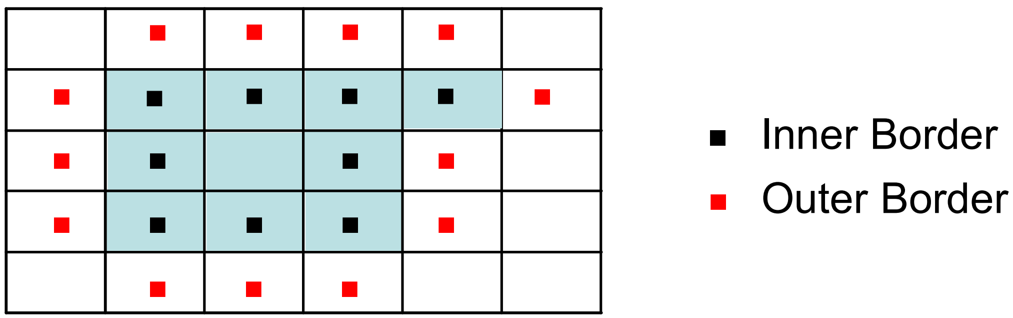
- Edges are pixels where brightness changes abruptly
- An edge is a vector variable with magnitude and direction
- The direction of the edge is perpendicular to the direction of the gradient
- The gradient gives the direction of maximal growth of the image function from black to white





Digital Image Properties: Border

- The border of a region R is the set of pixels within the region that have one or more neighbors outside



- The border is a global concept related to a region, while the edge represents a local property (at the pixel level)

Visual Perception of Images

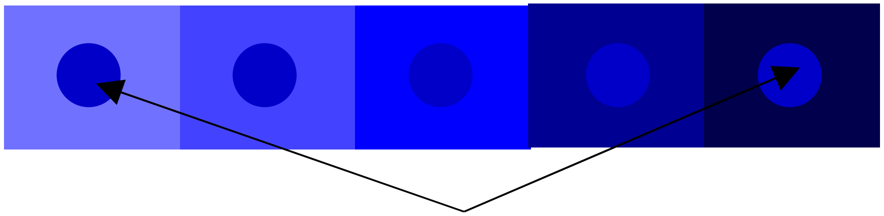
- Principles of human image perception are necessary to design image processing algorithms

- *Contrast*
  - Ratio between average brightness of an object and the background

Low Contrast

Higher Contrast

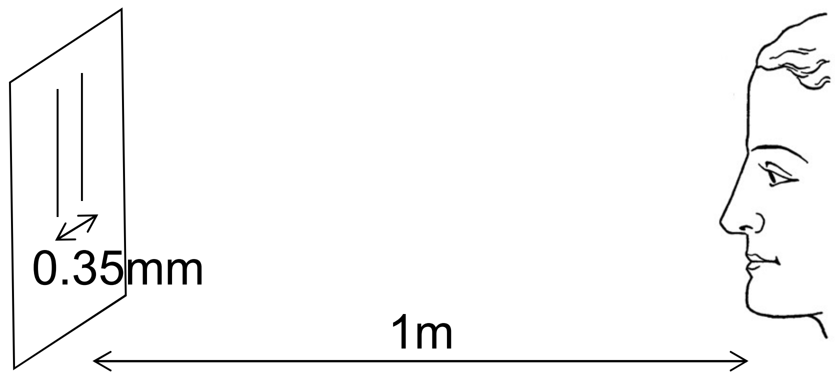
  - For the same perception, higher brightness requires higher contrast
  - Conditional contrast effect



Same brightness, different perception

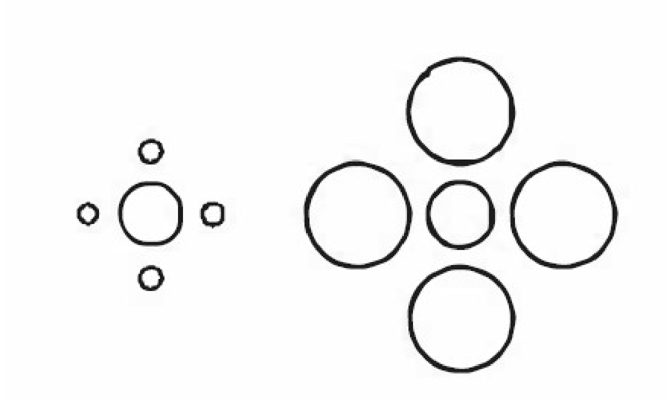
Visual Perception of Images

- *Acuity*
  - The ability to detect details in images
  - Acuity decreases with increasing distance
  - Resolution in an image is bounded by the resolution ability of the human eye
  - Maximum theoretical resolution of human eye is 0.35 mm line pair at a distance of 1m

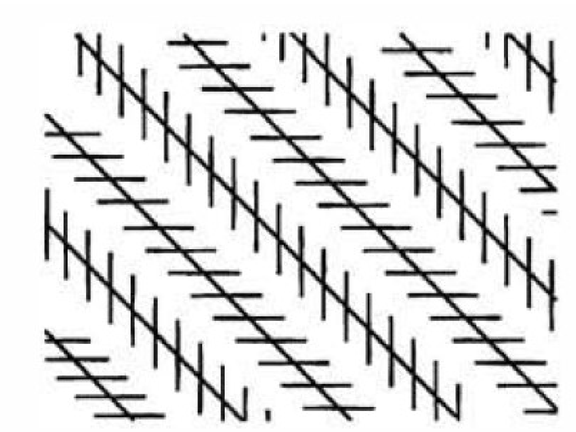


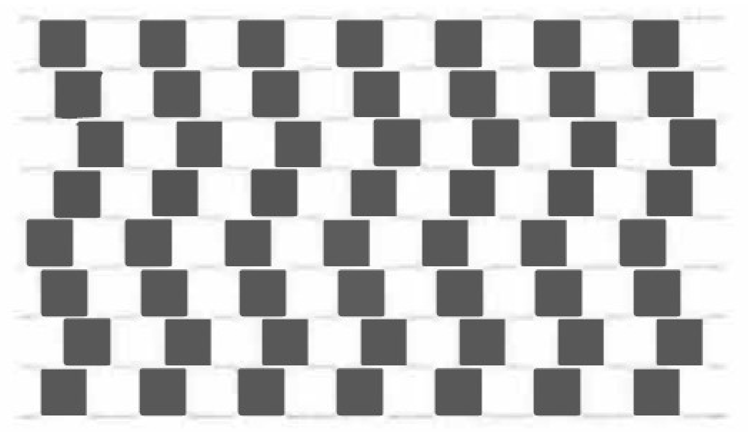
Visual Perception of Images

- Visual Illusions
  - Borders carry a lot of information for humans



Two circles with same diameter appear to have different sizes



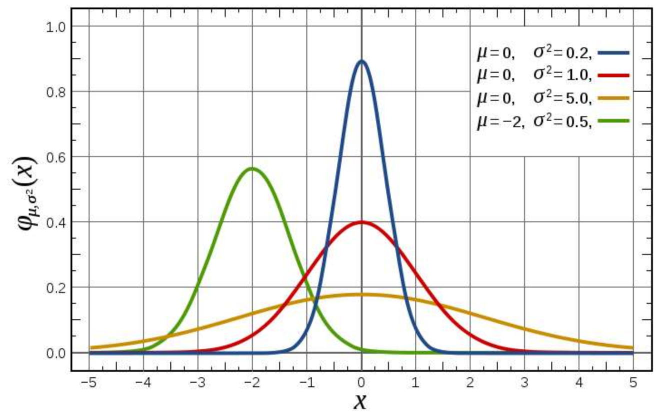


Perception of a dominant shape can be fooled by nearby shapes

Noise in Images

- Noise can occur during image capture, transmission or processing
- Gaussian Noise

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right)$$



- Types of Noise

  1. Additive Noise

$$f(x,y) = g(x,y) + v(x,y)$$

where g: original image, v: noise

Original Image

+ Gaussian Noise (μ = 0, σ = 20)

+ Gaussian Noise (μ = 0, σ = 60)

Noise in Images

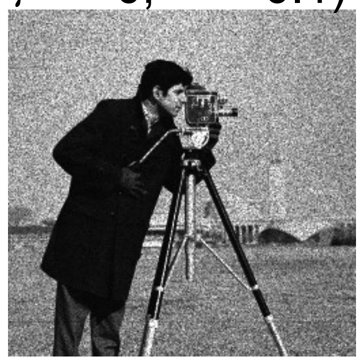
2. Multiplicative Noise

$$f(x, y) = g(x, y)(1 + v(x, y))$$

Noise is a function of the original brightness.

Original Image

x Gaussian Noise
(μ = 0, σ = 0.1)



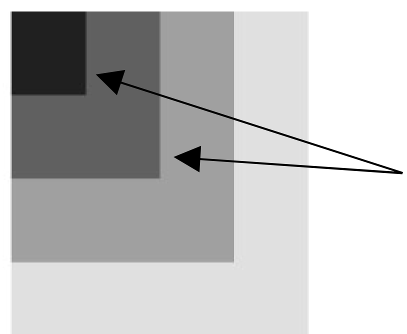
x Gaussian Noise
(μ = 0, σ = 0.2)

3. Quantization Noise

Occurs when insufficient quantization levels are used

256 Colors

4 Colors



False Contours

Noise in Images

4. Impulse Noise (Salt and Pepper)

Randomly changing the brightness of some pixels to either black or white

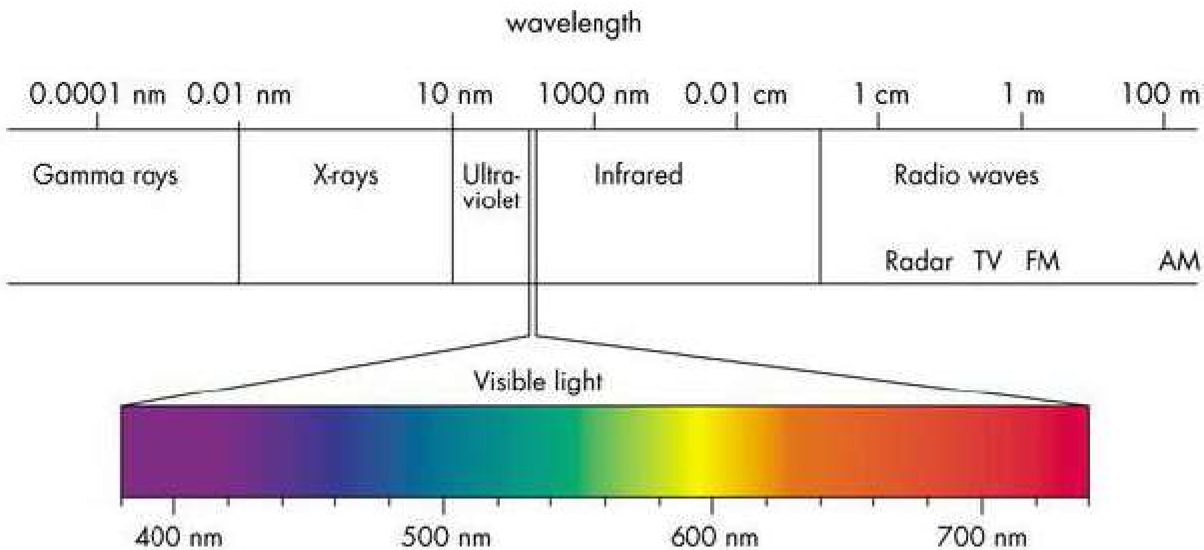
Original Image

With salt and pepper

- To eliminate noise:
  - No information about noise properties → Local pre-processing
  - Noise parameters are known → Image restoration techniques

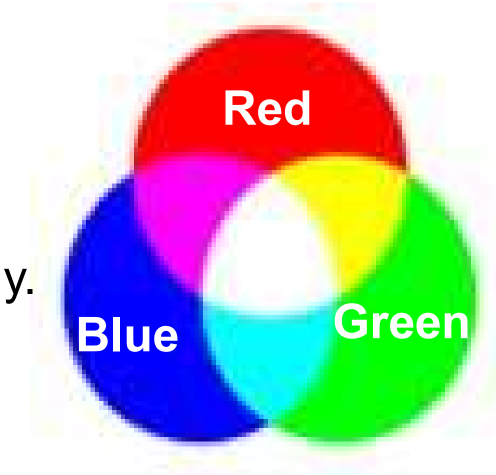
Color Images

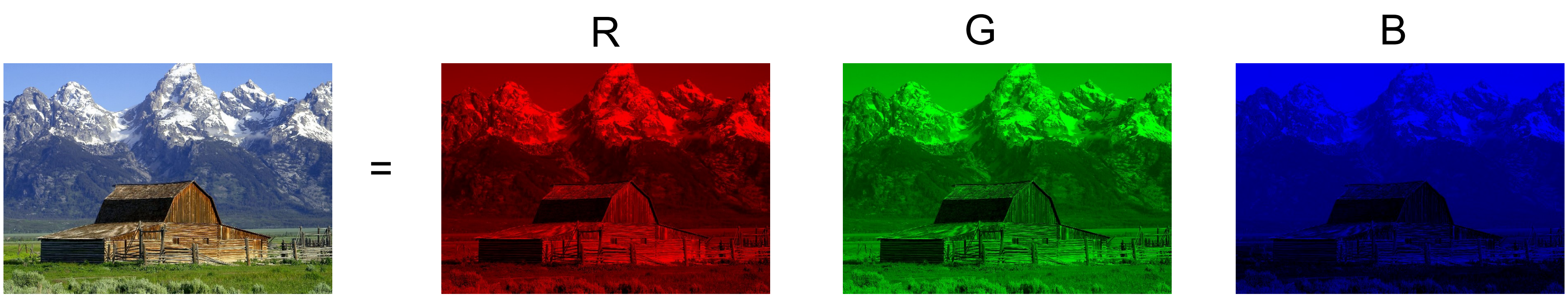
- Visible light represents a very narrow section of the electromagnetic spectrum



Color Spaces

- RGB Color Space (Red-Green-Blue)
  - Additive Mixing
  - The color of each pixel is represented by a vector (r, g, b), where r, g, and b represent the intensity of the primary colors Red, Green, and Blue, respectively. If r, g, and b are each represented by 8-bits, the number of possible colors is then 2^{24}.





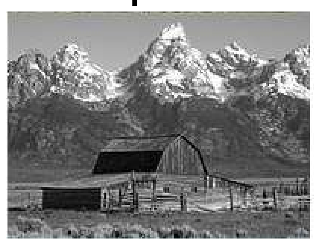
Color Spaces

- YIQ Color Space
  - Additive Mixing

  - The Y component represents intensity, while the I (in-phase) and Q (quadrature) components represent color

=

Y

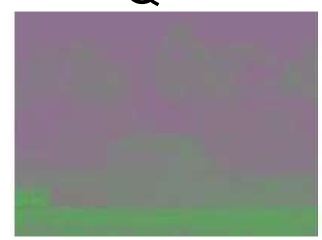


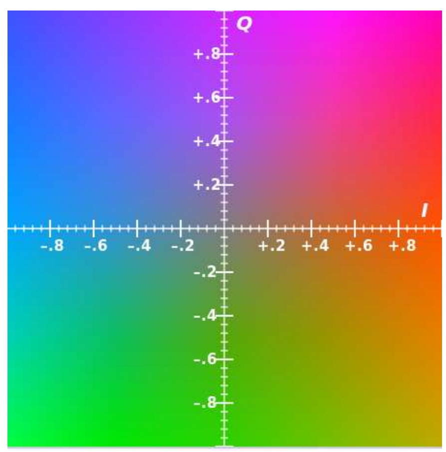
+

I

+

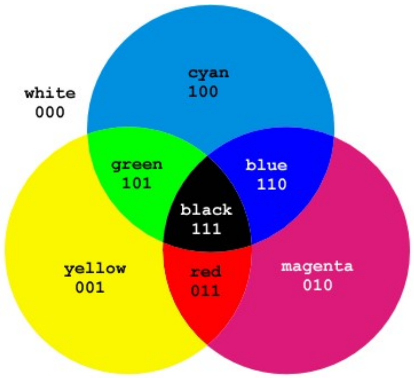
Q



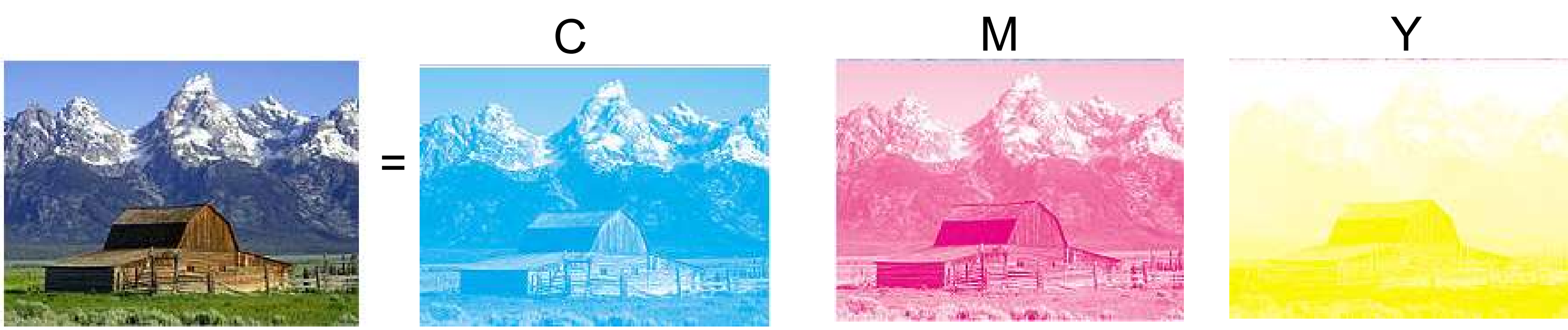


Color Spaces

- CMYK Color Space (Cyan-Magenta-Yellow-Black)
  - Subtractive Mixing
  - Used in printers. It describes the color of the ink that needs to be used so that light reflecting from the white paper and passing through the ink produces the desired color

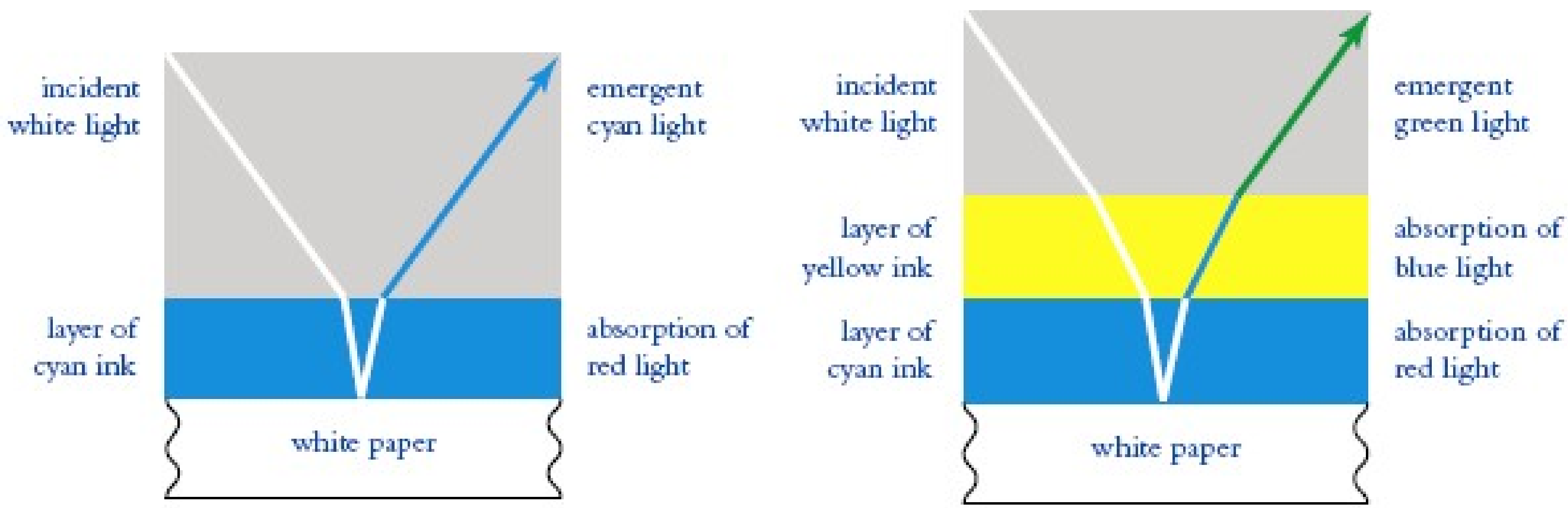


  - Example: Cyan is the complement of Red.
    When a white paper is painted with cyan, no red color is reflected



Color Spaces

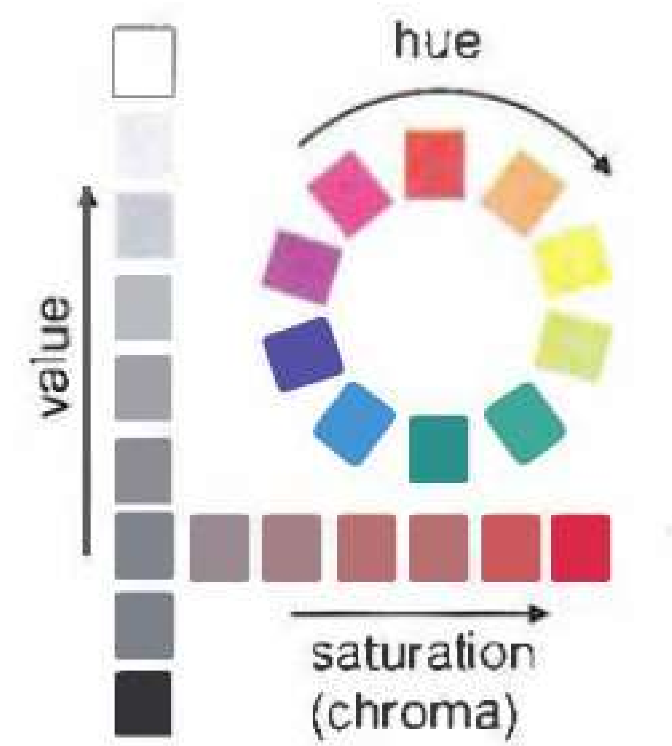
- In the printing process, thin layers of ink absorb some components of the incident light, so overlaying ink mixes colors subtractively
- Reflected light appears to be colored as ink reflects certain colors only



Coloured inks

Color Spaces

- HSV Color Space (Hue-Saturation-Value)
  - Additive Mixing
  - It’s closer to how painters mix colors
  - Hue corresponds to color
    Saturation corresponds to degree of color
    Value corresponds to brightness



Switching between Formats

- RGB to Gray Scale

$$I_{gray}(p) = \frac{I_R(p) + I_G(p) + I_B(p)}{3}$$

Better result could be obtained by weighting different components as follows

$$I_{gray}(p) = 0.3 I_R(p) + 0.59 I_G(p) + 0.11 I_B(p)$$

Switching between Formats

- Gray Scale to Binary

$$I_{bin}(p) = \begin{cases} 1 \text{ or } 255 & \text{if } I_{gray}(p) >= d \\ 0 & \text{Otherwise} \end{cases}$$