# Digital Image Fundamentals

*Chang-Su Kim*

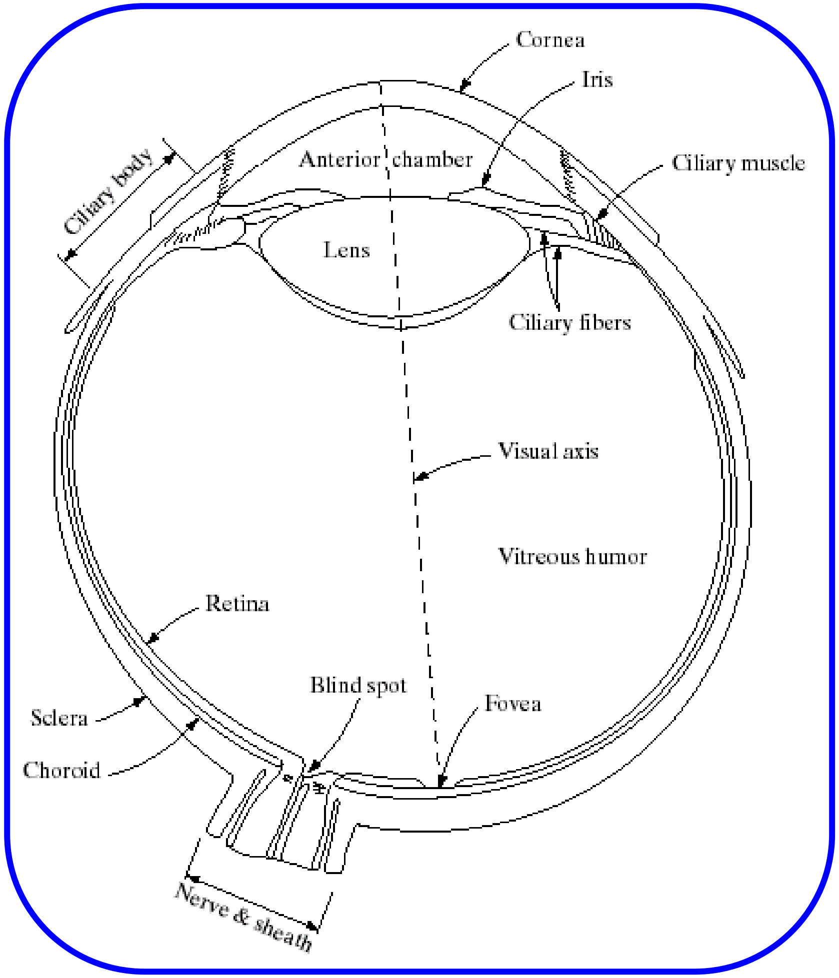
# Contents

- Human Visual System
- Basic Relationships between Pixels
- Image Fidelity Criterion

# Human Visual System

- Understanding of HVS
  - Measure of image fidelity or quality
  - Design and evaluation of DIP systems

- Human eye
  - Retina: film
  - Fovea: center of retina
  - Cones and rods: discrete light receptors on retina



# Cones and Rods

| Cones | Rods |
|---|---|
| Bright-light vision (photopic vision) | Dim-light vision (scotopic vision) |
| Sensitive to high level of illumination | Sensitive to low level of illumination |
| Concentrated around fovea (region of interest) | Widely distributed (overall picture) |
| Sensitive to colors | Insensitive to colors |
| 6-7 millions | 75-150 millions |

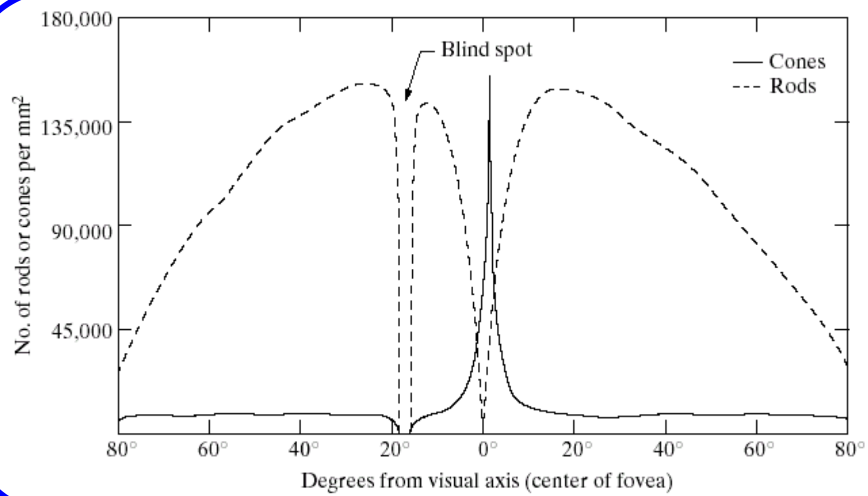
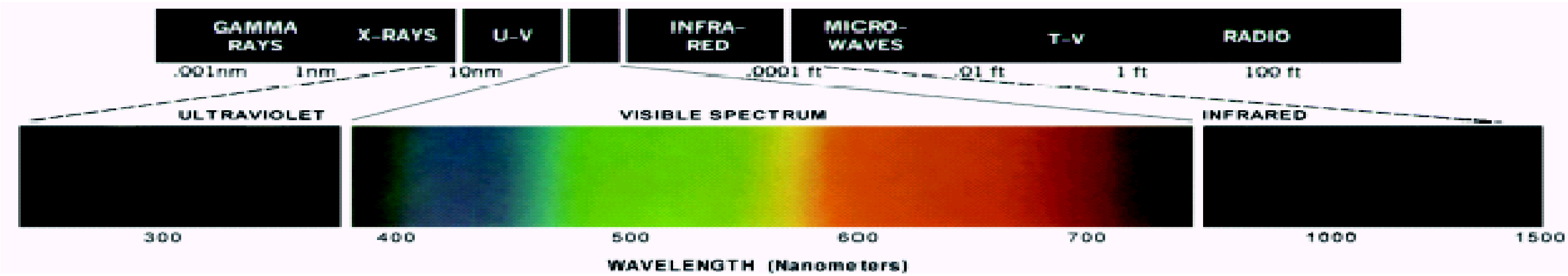



**FIGURE 2.2** Distribution of rods and cones in the retina.

# Red, Green, and Blue Cones



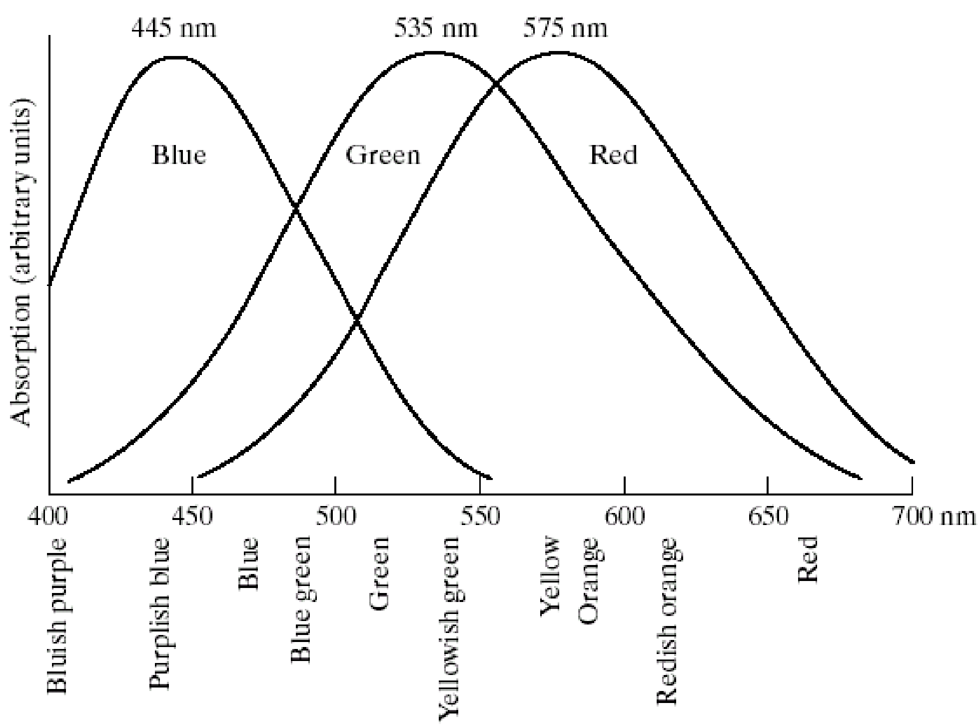




**FIGURE 6.3** Absorption of light by the red, green, and blue cones in the human eye as a function of wavelength.

# Color Blindness

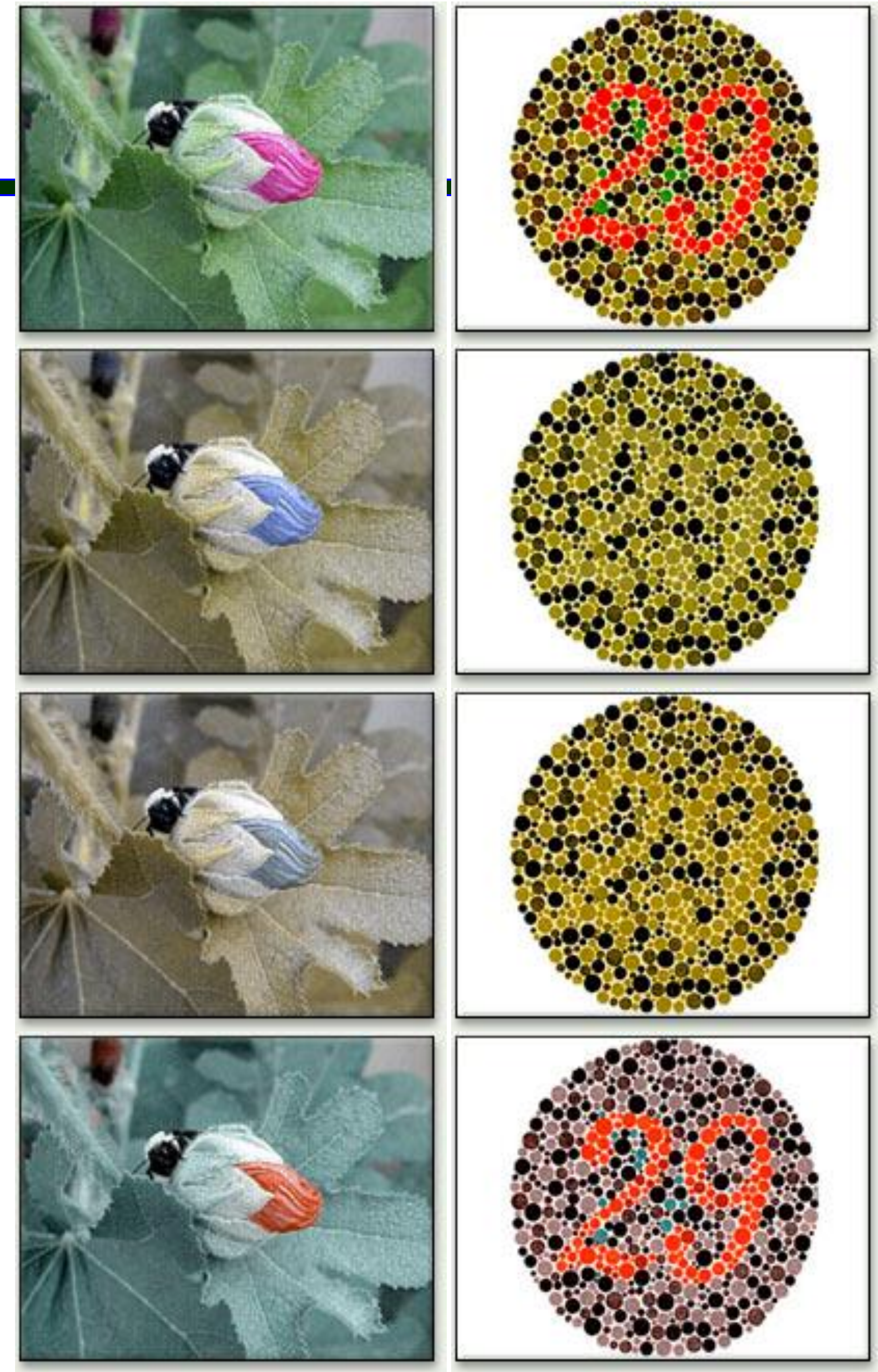
- Color deficiency
  - Not color blindness
- This happens when
  - One or more cones are missing
  - Their peak sensitivities are different from normal ones
- R-cone and G-cone deficiencies are more common than B-cone deficiency

normal

R-cone

G-cone

B-cone



# Radiance, Luminance and Brightness

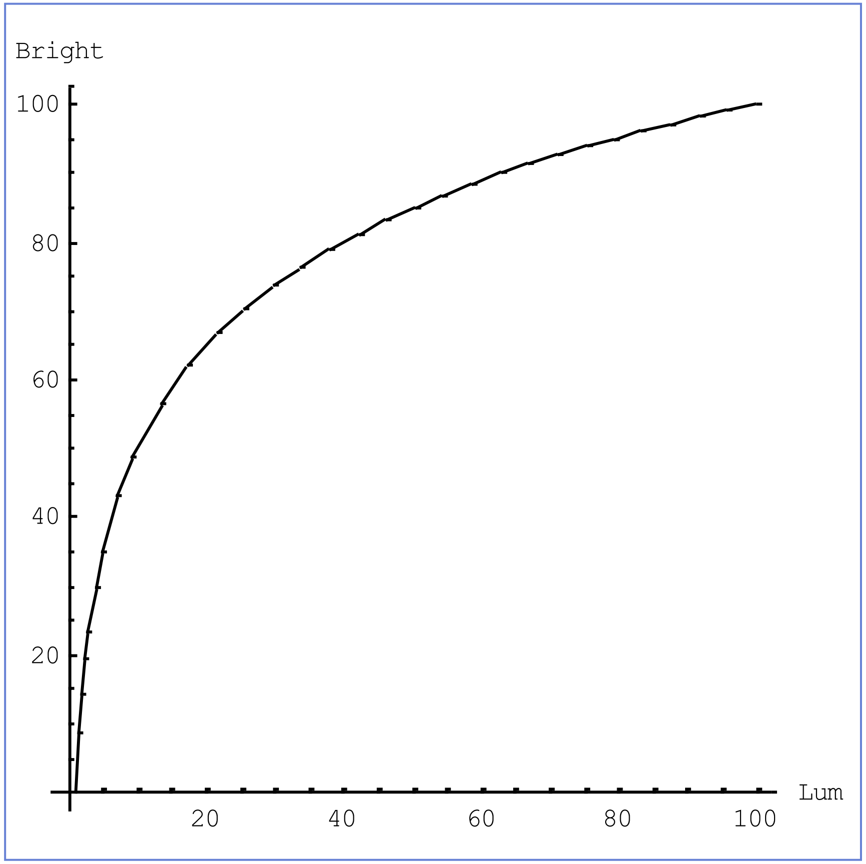
- Radiance
  - The total amount of energy that flows from a light source
- Luminance
  - The amount of energy an observer perceives from a light source
- Brightness
  - Subjective description of luminance

e.g.) Infrared heater

- High radiance – it is a heater !
- Low luminance – human eyes are sensitive only to visible spectrum

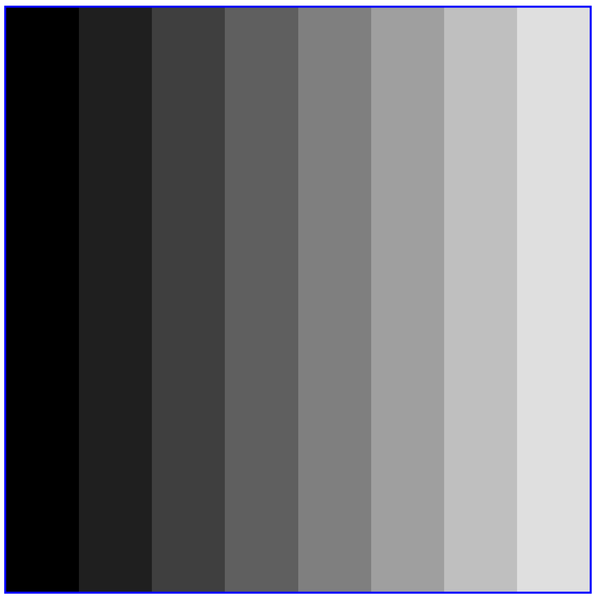
# Radiance, Luminance and Brightness

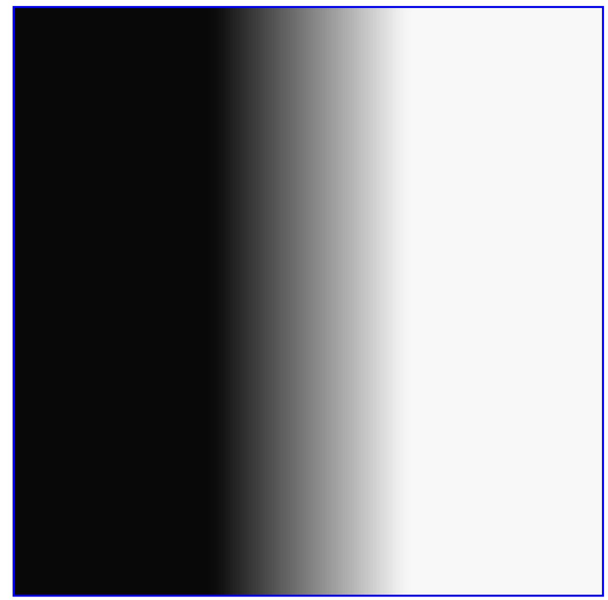
- Brightness (*b*) vs. Luminance (*l*)
  - Log relation (approximate)
  - $b = 50 \log_{10} l$ $(1 <= l <= 100)$
    - $l = 10$, $b=50$
    - $l = 100$, $b = 100$
    - To be two times brighter, luminance should be squared

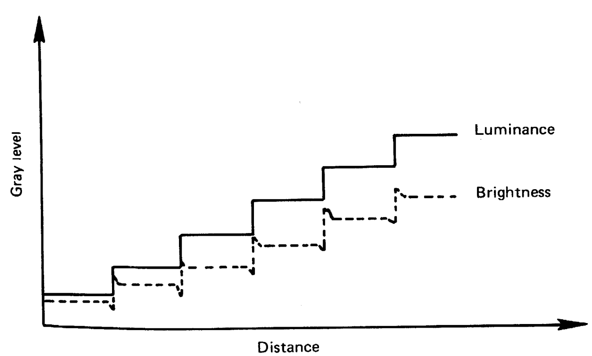


# Mach Bands

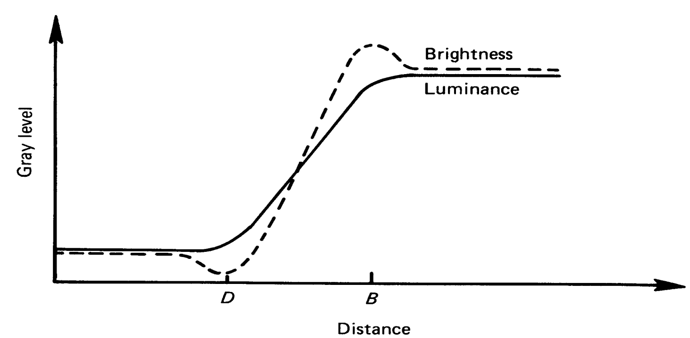
- Perceived brightness depends on **surroundings** as well as luminance








Luminance versus brightness.



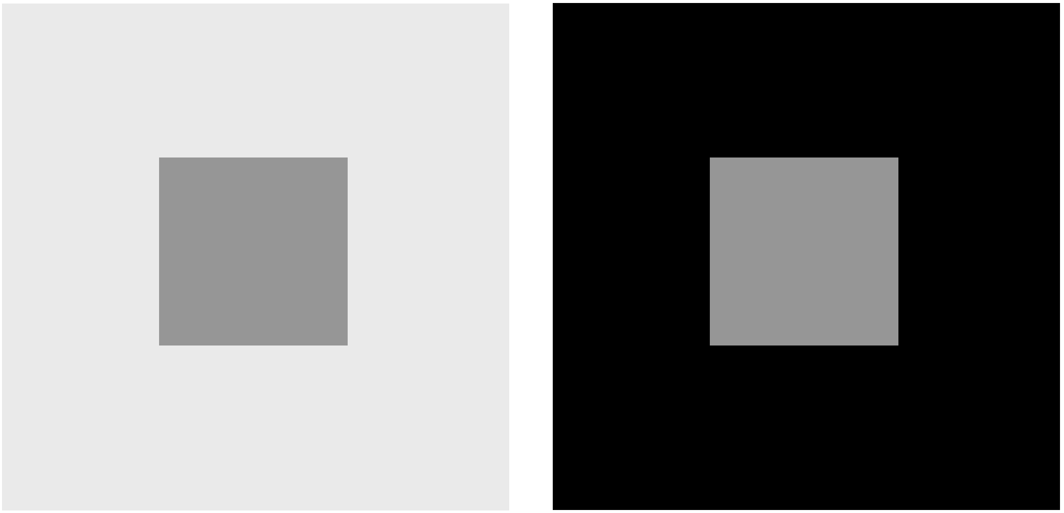

Mach band effect.

# Simultaneous Contrast

- Perceived brightness depends on **surroundings** as well as luminance

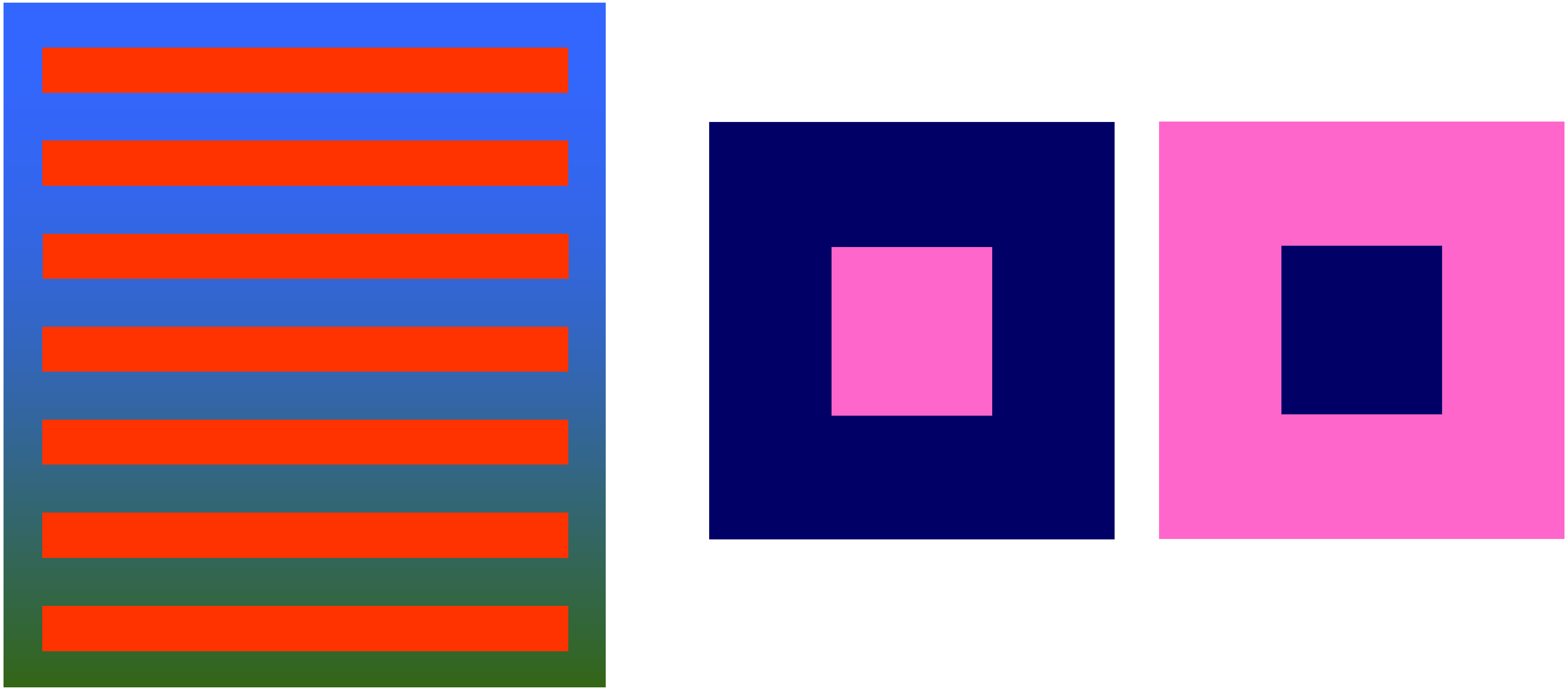
# Simultaneous Contrast

- Perceived brightness depends on **surroundings** as well as luminance



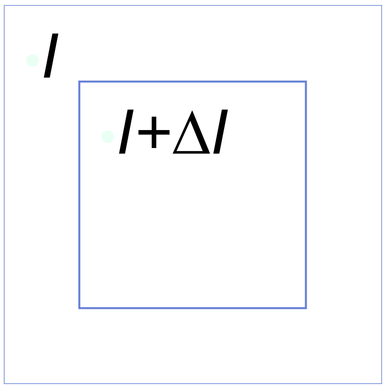
# Simultaneous Contrast

- Perceived color also depends on **surroundings**



# Brightness Discrimination

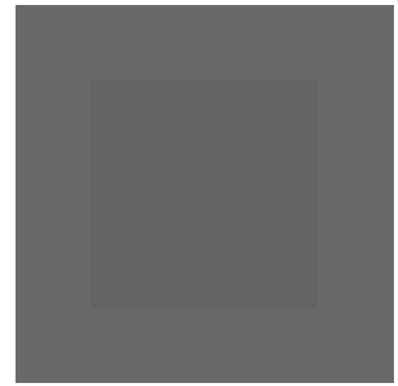
- Experiments




$\Delta I_c$ : JND (just noticeable difference)

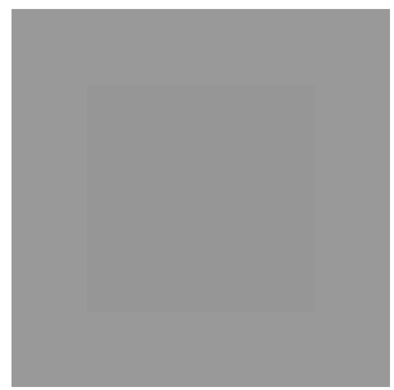
(100,101)

(100,102)

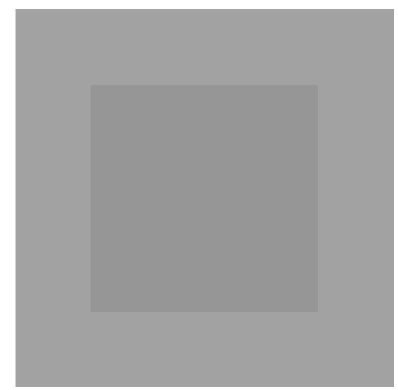


(100,105)

(150,151)



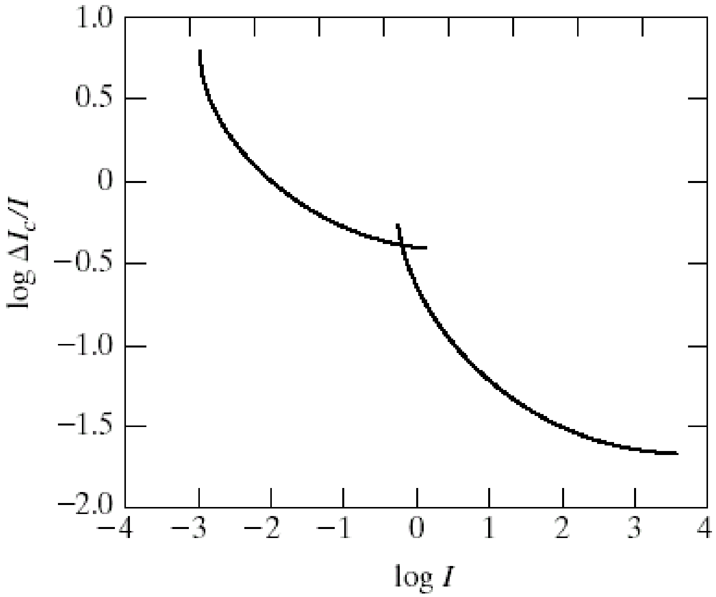
(150,153)



(150,162)

# Weber's Ratio

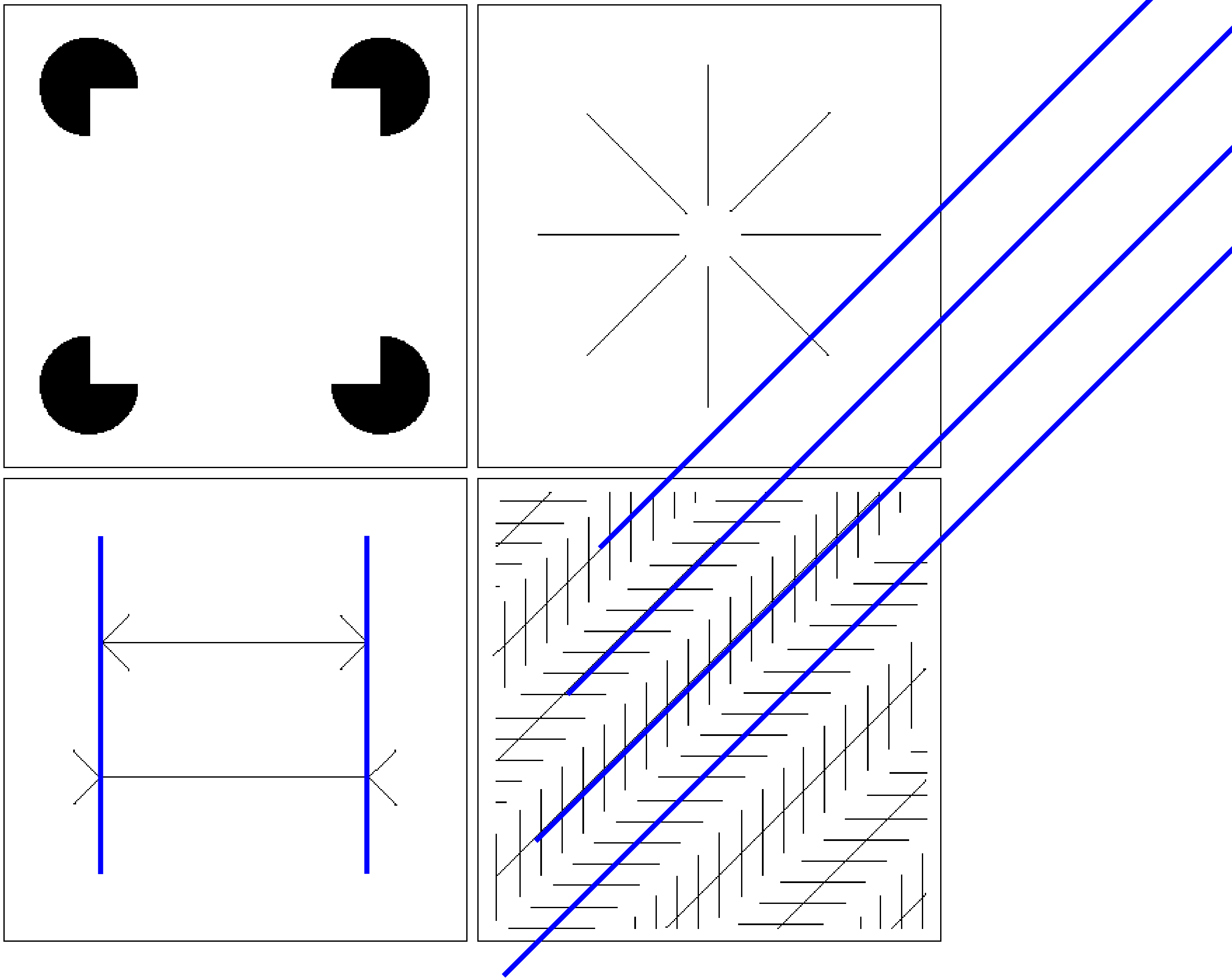
- $I$ and $I+\Delta I_c$
  - Their differences can be just noticeable
  - $\Delta I_c$ depends on $I$
- Weber's law
  - $$\frac{\Delta I_c}{I} = \text{constant}$$
  - More recent result



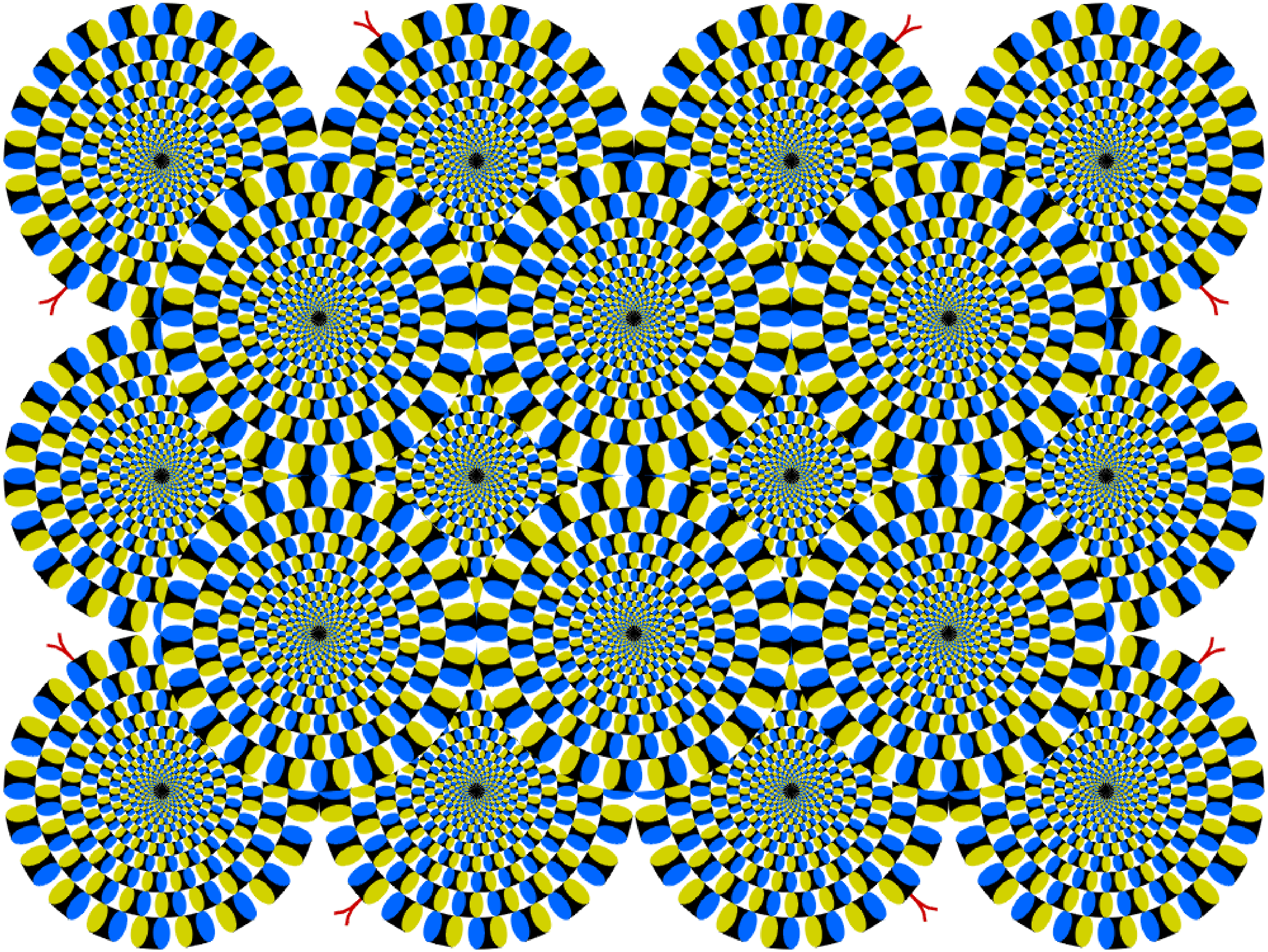
# Optical Illusions

a b
c d

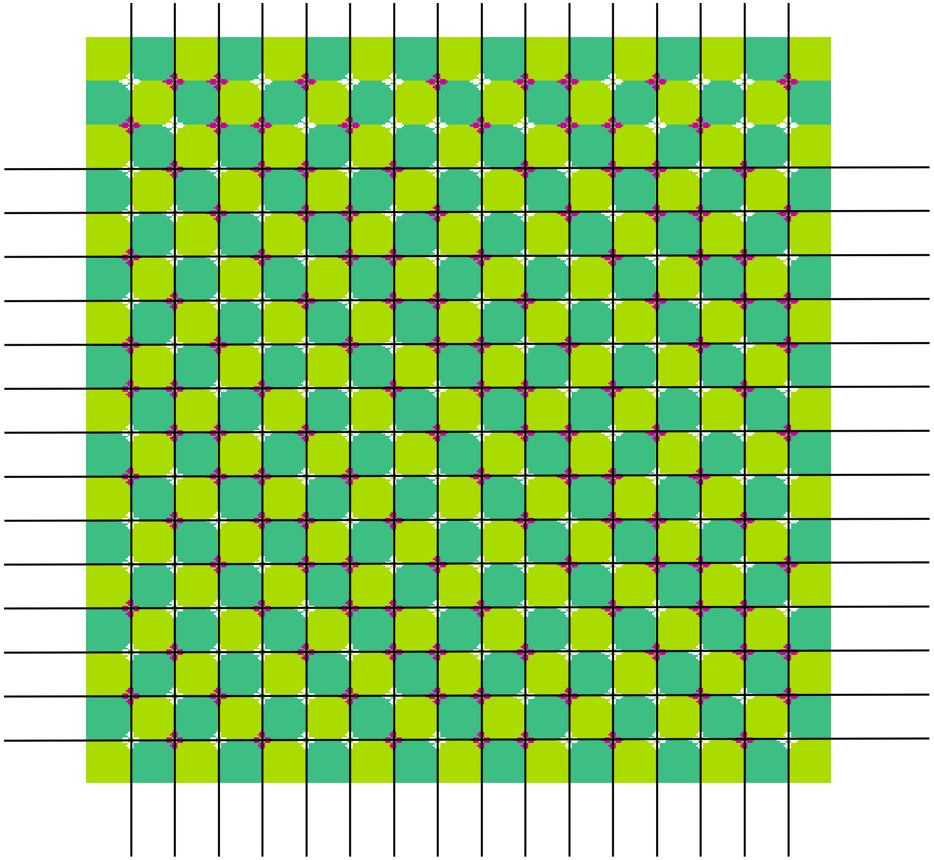
**FIGURE 2.9** Some well-known optical illusions.



# More Illusion Examples



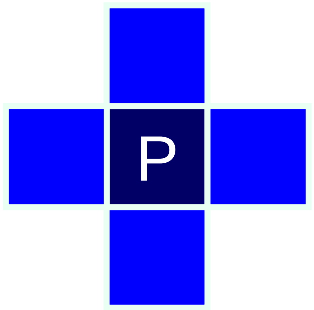
# More Illusion Examples

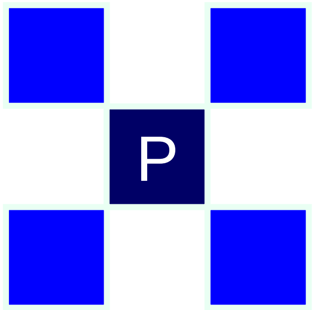


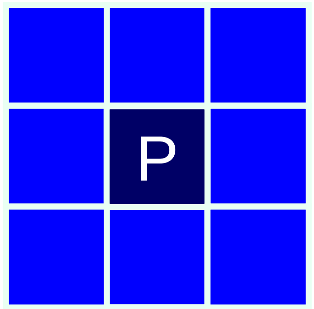
We do not understand HVS fully.

# Basic Relationships between Pixels

- Neighborhood
  - 4-neighbors of p = (x,y): $N_4(p)$
    - (x+1,y), (x-1,y), (x,y-1), (x, y+1)
  - Four diagonal neighbors of p: $N_D(p)$
    - (x+1,y+1), (x+1,y-1)
    - (x-1,y+1), (x-1,y-1)
  - 8-neighbors of p: $N_8(p) = N_4(p) \cup N_D(p)$
- Note that neighborhood depends on pixel coordinates only (not on pixel values)



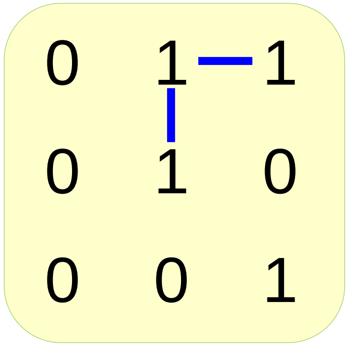



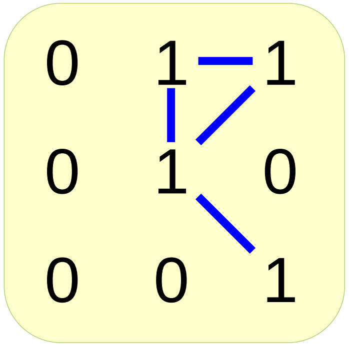



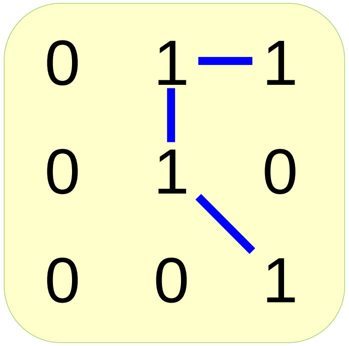
# Basic Relationships between Pixels

- Let V be a set of similar gray values
  - e.g. V={1} in binary images
- Adjacency
  - p and q are 4-adjacent if
    - $f(p), f(q) \in V$, and
    - $q \in N_4(p)$ (equivalent to $p \in N_4(q)$)
  - p and q are 8-adjacent if
    - $f(p), f(q) \in V$, and
    - $q \in N_8(p)$
  - p and q are m-adjacent if
    - $f(p), f(q) \in V$, and
    - (i) $q \in N_4(p)$, or

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

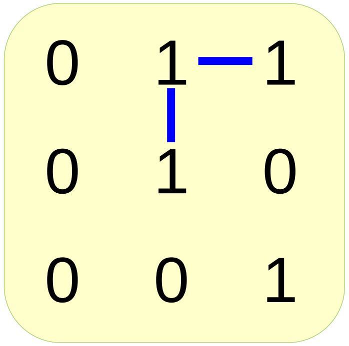


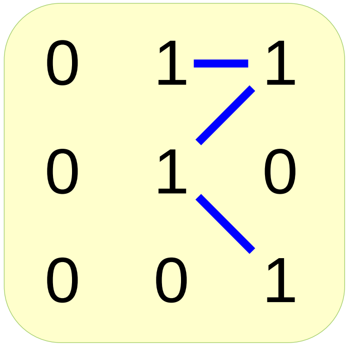



# Basic Relationships between Pixels

- Path or Curve (from $p_1$ to $p_n$)
  - There exists a sequence
    - $p_1, p_2, p_3, \ldots p_{n-1}, p_n$
    - s.t. each $p_i$ and $p_{i+1}$ are adjacent
  - Path length = n-1
  - Closed path if $p_1 = p_n$
  - 4-path, 8-path, m-path



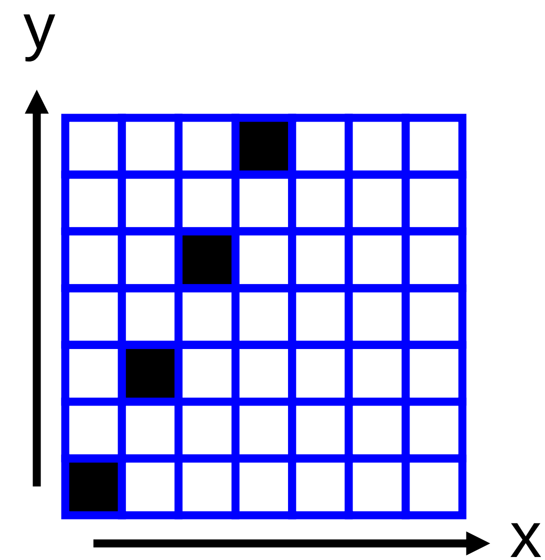

- Connected Set
  - A set of pixels S is connected if for each p, q in S, there exists a path between p and q
  - 4-connected set, 8-connected set, m-connected set

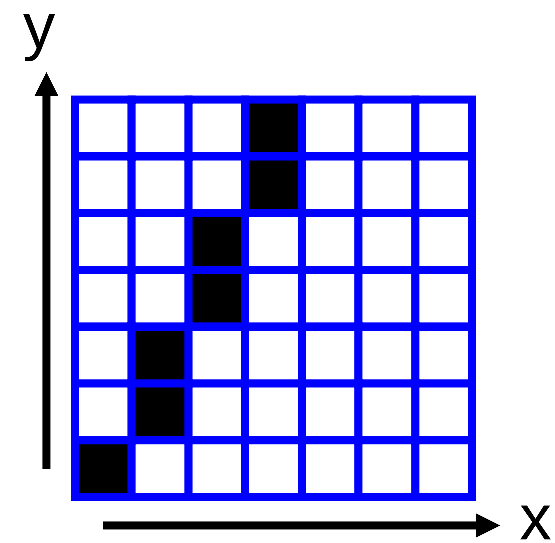
# Summary of Pixel Relationships

- Neighborhood
  - Coordinates concept
  - 4-neighborhood, 8-neighborhood
- Adjacency
  - Neighborhood + similar pixel values (same pixel value in binary case)
  - 4-adjacency, 8-adjacency, and m-adjacency
- Path
  - Pixel sequence in which each pair of consecutive elements are adjacent
  - Important in describing region boundaries
- Connected Set
  - Between every pair of pixels, there exists a path
  - Important in describing regions

# Scan Converting of Lines

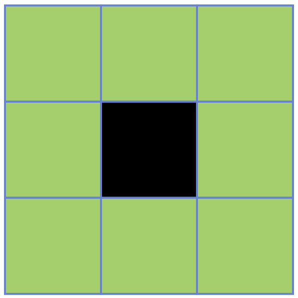
- Concept of path
- Draw the line y = f(x)=2x
  - for each x, plot (x, f(x))
    - (0,0), (1,2), (2,4), (3,6) …
  - Not a path
- Alternatively, x=g(y)=0.5y
  - for each y, plot (g(y), y)
    - (0,0), (0.5, 1), (1, 2), (1.5, 3), (2, 4), (2.5, 5), (3, 6)
    - (0, 0), (1, 1), (1, 2), (2, 3), (2, 4), (3, 5) (3, 6)
  - Be a path



# 2D Neighbors vs. 3D Neighbors

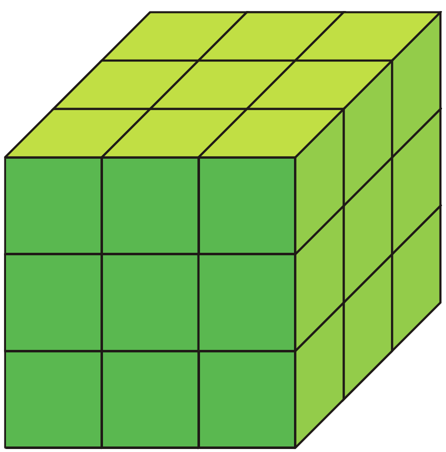
- 2D neighborhood (square pixel)



4-neighbors: share edge

8-neighbors: share edge or vertex

- 3D neighborhood (cube voxel)



6-neighbors: share face

18-neighbors: share face or edge

26-neighbors: share face, edge or vertex

# Distance between Pixel Coordinates

- D is a distance function (or metric or norm) if
  1. $D(p,q) \geq 0$ ($D(p,q)=0$ iff $p=q$)
  2. $D(p,q) = D(q,p)$
  3. $D(p,q) \leq D(p,z)+D(z,q)$

  where $p=(x,y)$, $q=(s,t)$, $z=(v,w)$ are pixel coordinates

- $D_n(p,q) = [\ |x-s|^n + |y-t|^n\ ]^{1/n}$
  - $D_2(p,q)= [\ (x-s)^2+(y-t)^2]^{1/2}$
    - Euclidian distance
    - Conditions 1 and 2 are obvious
    - Condition 3 is due to triangle inequality
  - $D_1(p,q) = |x-s| + |y-t|$
  - $D_\infty(p,q) = \max\{\ |x-s|, |y-t|\ \}$

| | | |
|---|---|---|
| 0 | p | 0 |
| 0 | 0 | 0 |
| 0 | 0 | q |

$D_2(p,q)=$ Sqrt(5)
$D_1(p,q)= 3$
$D_\infty(p,q)= 2$

# Distance between Pixel Coordinates

- It can be shown that $D_n(p,q)$ is a valid distance function for each positive number n
- For example, consider $D_1(p,q) = |x-s| + |y-t|$, where p = (x,y) and q = (s,t)
  - Condition 1
    - $D_1(p,q) \geq 0$ since it is the sum of absolute values
    - $D_1(p,q) = 0$ if and only if x=s and y=t (i.e. p=q)
  - Condition 2
    - $D_1(p,q) = |x-s| + |y-t| = |s-x| + |t-x| = D_1(q,p)$
  - Condition 3
    - Let z=(v,w) be a pixel coordinate
    - $|x-s| \leq |x-v| + |v-s|$ (Lemma in next slide)
    - Similarly, $|y-t| \leq |y-w| + |w-t|$
    - Therefore,

$$\begin{aligned} D_1(p,q) &= |x-s| + |y-t| \\ &\leq |x-v| + |v-s| + |y-w| + |w-t| \\ &= |x-v| + |y-w| + |v-s| + |w-t| \\ &= D_1(p,z) + D_1(z,q) \end{aligned}$$

# Distance between Pixel Coordinates

- Lemma: For any scalars x, v, s,

$$|x-s| \le |x-v| + |v-s|$$

- Proof)

1) $(x-v) \ge 0$ and $(v-s) \ge 0$:

$$|x-v|+|v-s| = x-v + v-s = x-s = |x-s|$$

2) $(x-v) < 0$ and $(v-s) < 0$:

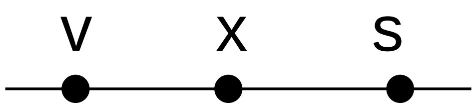
$$|x-v|+|v-s| = v-x + s-v = s-x = |s-x|$$

3) $(x-v) \ge 0$ and $(v-s) < 0$:

   a) $x \le s$:

   $v \le x \le s$

   Thus, $|x-s| < |v-s| \le |x-v| + |v-s|$

   b) $x>s$:

   $v<s<x$

   Thus, $|x-s| < |x-v| \le |x-v| + |v-s|$

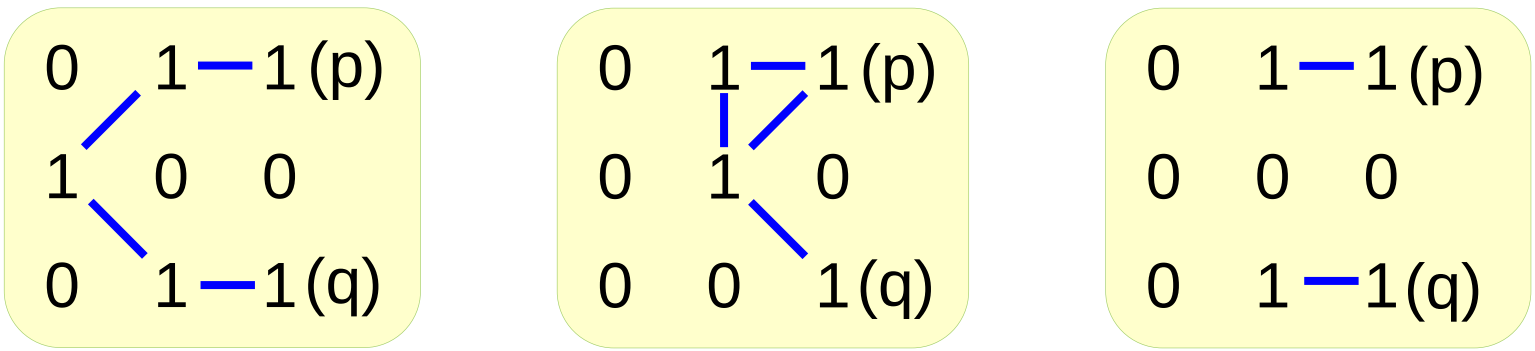
4) $(x-v) < 0$ and $(v-s) \ge 0$:

   Similar to Case 3)

# Distance between Pixel Coordinates

- f(p,q): the length of the shortest 8-path between p and q
  - If there is no 8-path, then f(p,q)=∞

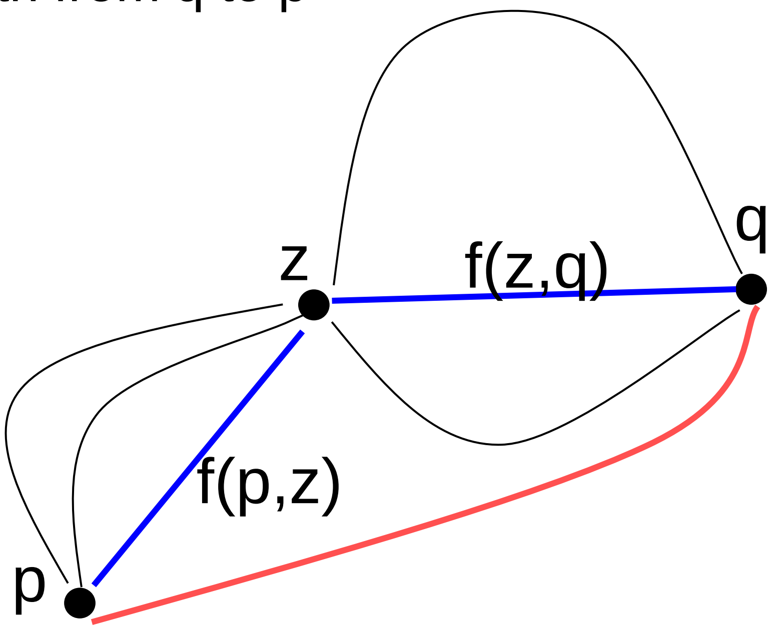
f(p,q) = 4 f(p,q) = 2 f(p,q) = ∞




- Is f(p,q) a valid distance function?
  - Yes, it is.

# Distance between Pixel Coordinates

- Conditon 1:
  - f(p,q) = the shortest path length ≥ 0
  - f(p,q) = 0  iff  p=q
- Condition 2:
  - f(p,q) = the shortest path length from p to q
    
    = the shortest path length from q to p
    
    = f(q,p)
- Condition 3:
  - Concatenation of two shortest paths
  - There exist a path from p to q, whose length is f(p,z) + f(z,q)
  - Therefore,
    
    f(p,q) ≤ f(p,z) + f(z,q)



# Image Fidelity Criteria

- i.e.) Image compression
  - f(x,y): original image of resolution M x N
  - g(x,y): reconstructed image of the same resolution
  - How similar g(x,y) is to f(x,y) ?

- MSE (Mean Square Error)

$$\mathrm{MSE} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(f(x,y)-g(x,y))^2$$

- PSNR (Peak Signal to Noise Ratio)

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\mathrm{MSE}} \quad \text{(dB)}$$

# Image Fidelity Criteria

- MAD (Mean Absolute Difference)

$$\mathrm{MAD} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}|f(x,y)-g(x,y)|$$

- Comparison
  - MAD is faster
  - MSE facilitates mathematical analysis
  - PSNR is intuitive
    - > 35 dB : almost the same as the original
    - < 25 dB : very poor quality
    - 28 – 32 dB : acceptable quality at very low bitrate applications