Lecture 2
Human visual system, perception & pixel

Dr. Xiran Cai

Email: caixr@shanghaitech.edu.cn

Office: 3-438 SIST

Tel: 20684431

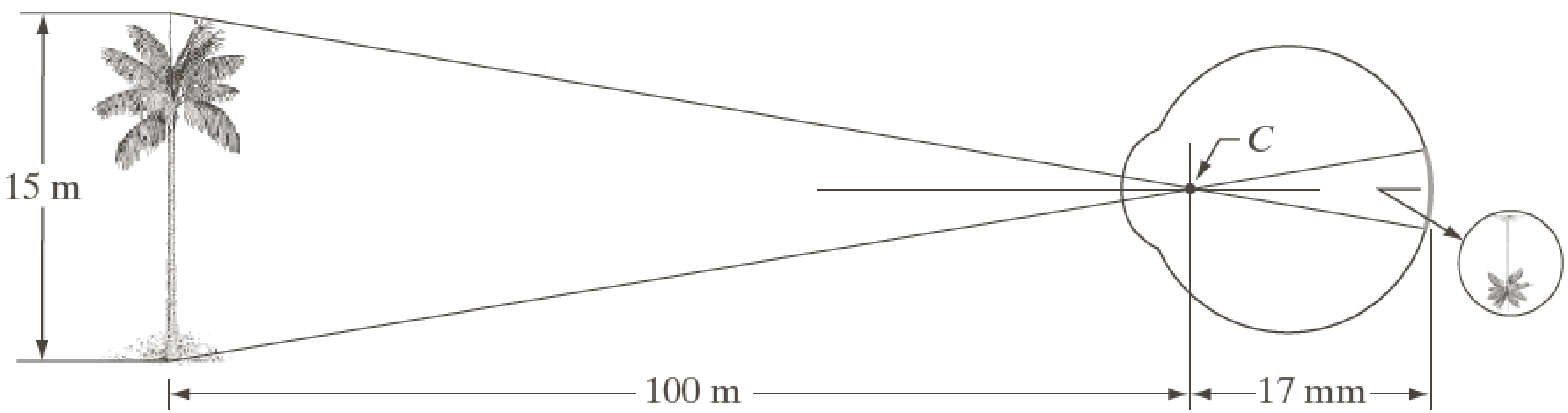
ShanghaiTech University

Outline

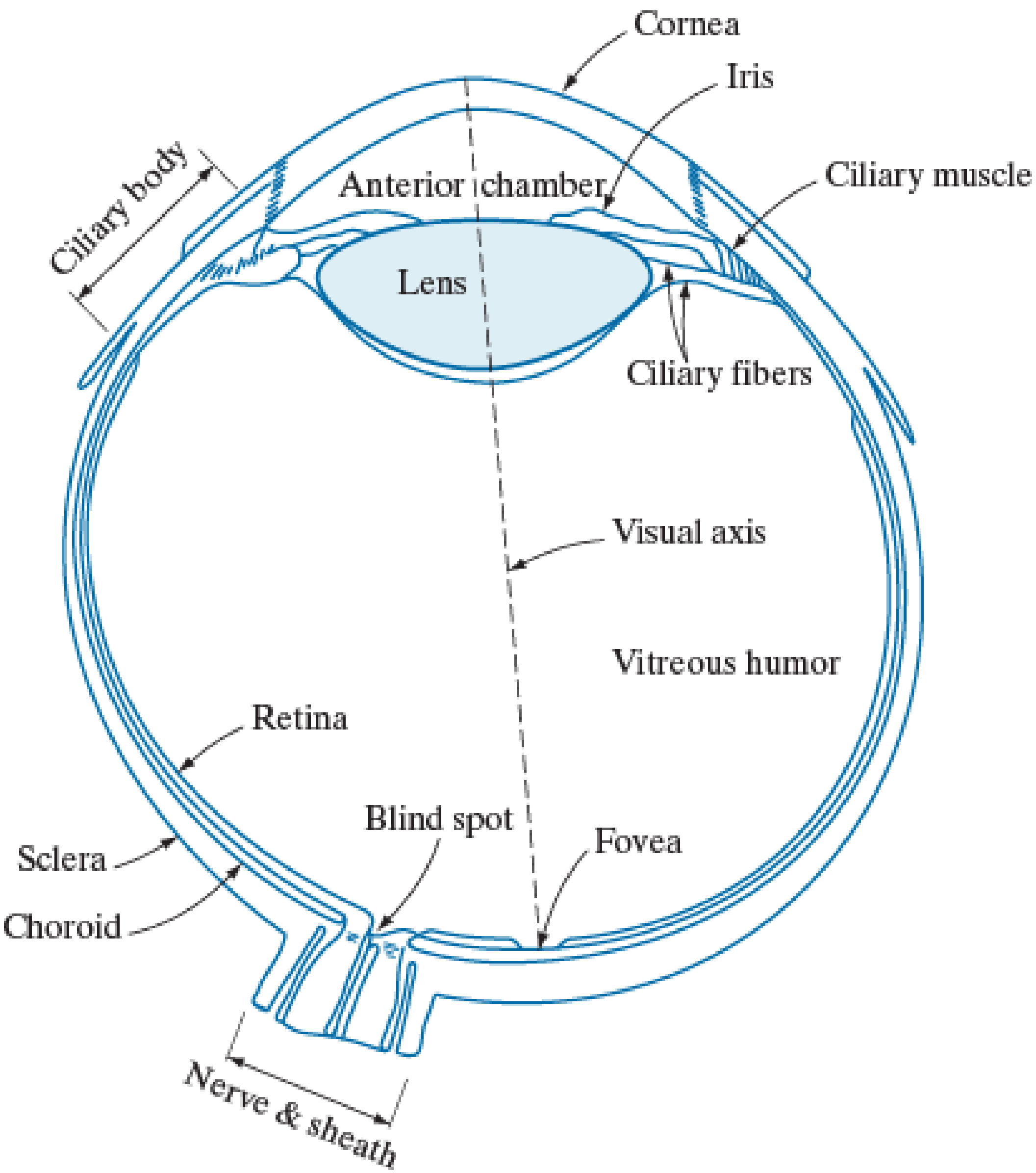
- Anatomical structure of eye ball
- Image formation and brightness adaptation in the eyes
- Visual illusions
- Pixel
- Basic operation of DIP

Graphical representation of human eye

- Objects captured as focused images on the image plane at retinas
- Perspective projection based on pinhole model geometry,
- Image size depends on distance of object
- In practice, optical devices with lens



Human Eye



Fovea: 中央凹
Retina: 视网膜
Iris: 虹膜
Pupil: the opening of the iris (2-8 mm) 瞳孔

Lens: 晶状体
Blind spot: 盲点

Receptors:
Rod: 杆状体
(scotopic or dim-light vision)
Cone: 锥状体
(photopic or bright-light vision)

Human Visual Perception

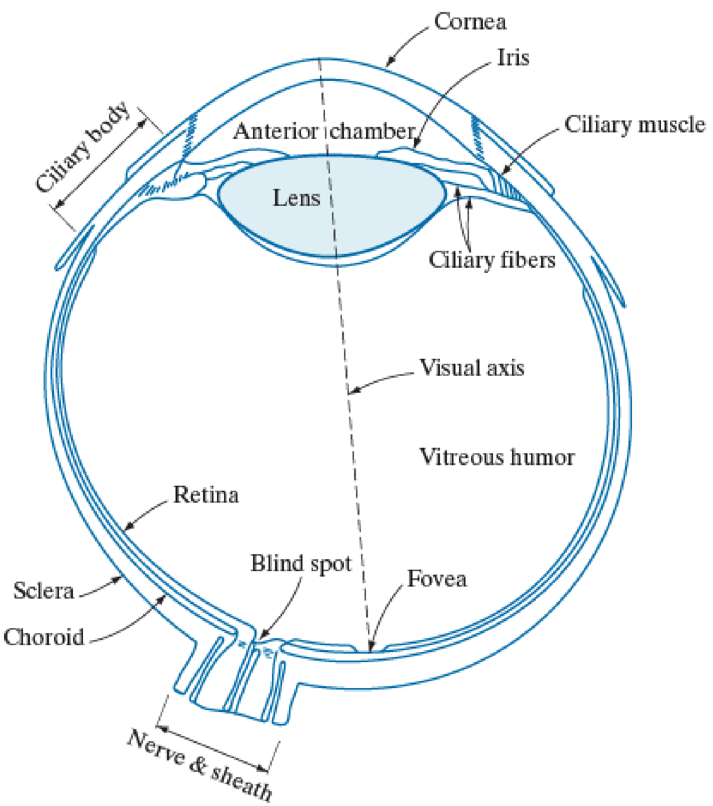


Figure 2.1: Sagittal Eye ball anatomy

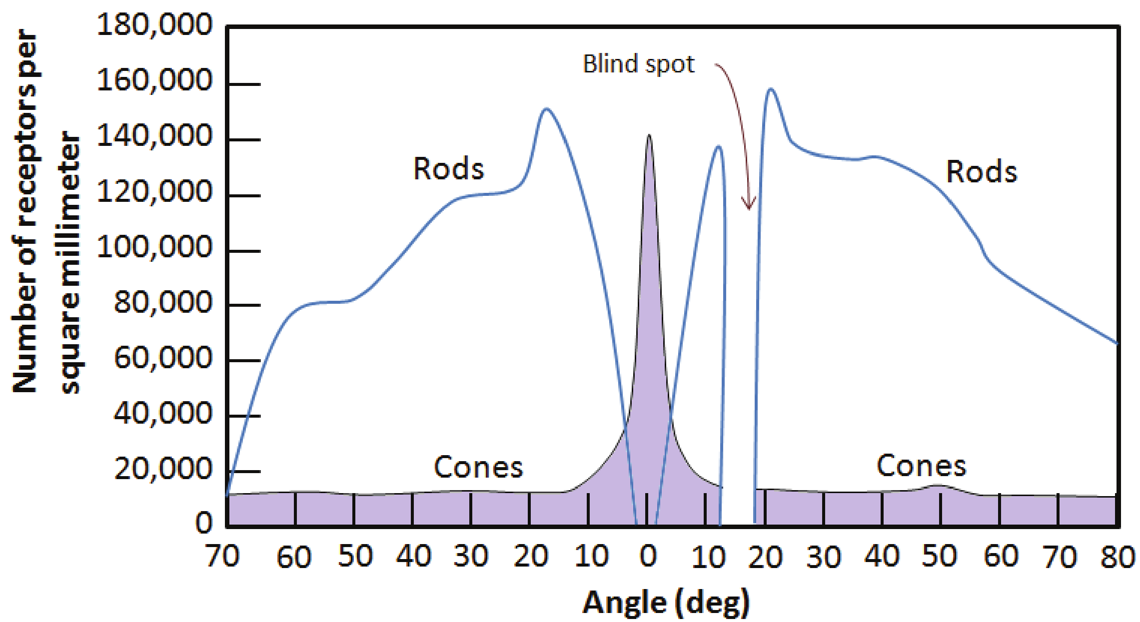


Figure 2.2: Distribution of rods and cones in the retina

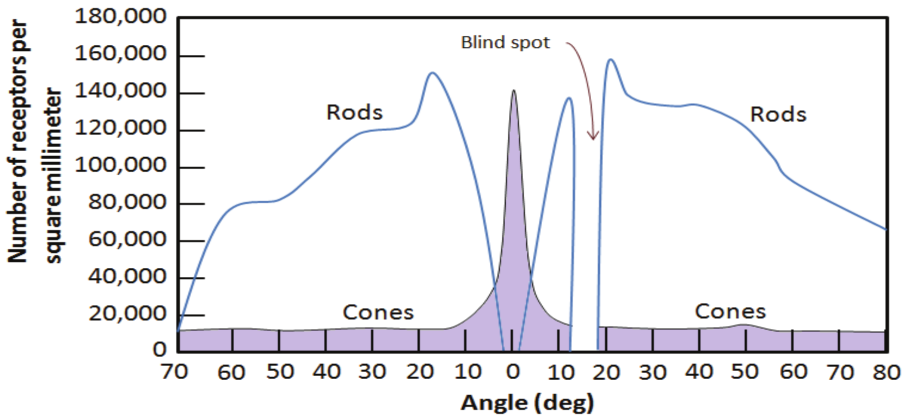
Human visual system perception elements

Rods:

- Sensitive to light intensity
- Night "scotopic" vision
- Achromatic
- Low acuity (many per nerve end)
- Peripheric vision
- Slow response
- 75-150 Million/Retina

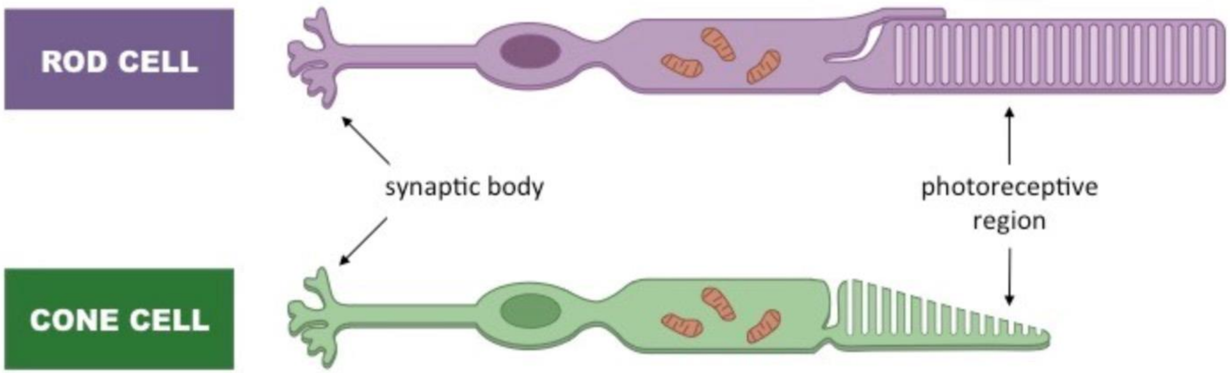
Cones:

- Only direct to direct light
- "Photopic" vision
- Chromatic (3 "colors")
- Concentrated in fovea (1 per nerve end)
- High visual acuity, spatial resolution.
- Fast response
- 6-7 Million/Retina



Rods and cones

Types of Photoreceptors (Rods and Cones)

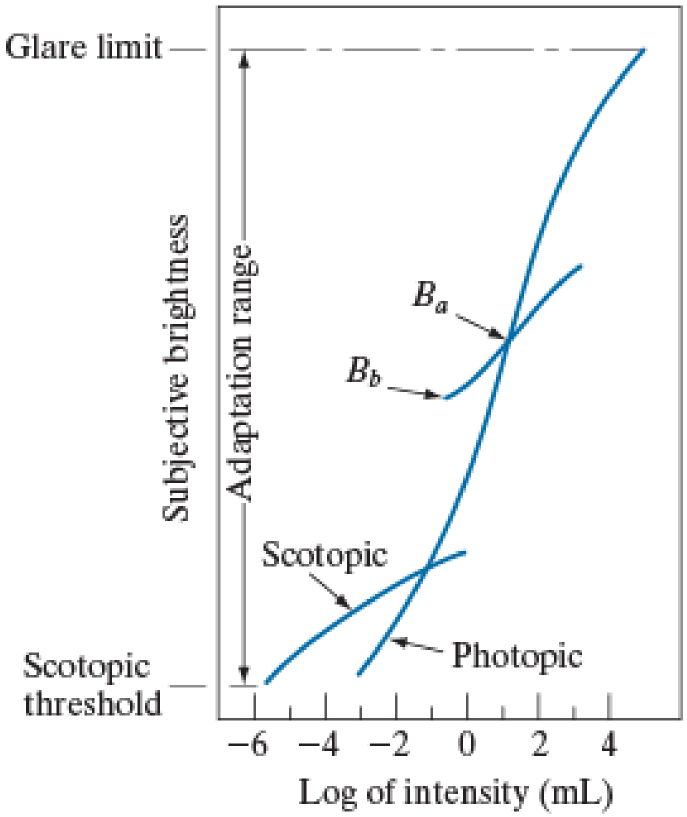


	Rod Cells	Cone Cells
Location in retina	Found around periphery	Found around centre (fovea)
Optimal light conditions	Dim light ('night' vision)	Bright light ('day' vision)
Visual acuity	Low resolution (many rods : one bipolar cell)	High resolution (one cone : one bipolar cell)
Colour sensitivity	All wavelengths	Certain wavelengths (red, green, blue)
Type of vision	Achromatic (black and white)	Colour
Number of types	One (all contain rhodopsin)	Three different iodopsin pigments
Relative abundance	Many	Fewer

ib.bioninja.com.au

Characteristics of Human Visual System

- ❑ **Brightness adaption**
- ❑ Simultaneous contrast
- ❑ Mach band effect
- ❑ Optical illusion



- Impressive total dynamic range ;)
- But, cannot operate over a long range simultaneously ;(
- Brightness adaption level (*Ba*)
- Below *Bb* is **black!**

Characteristics of Human Visual System

- **Brightness adaption**
- Simultaneous contrast
- Mach band effect
- Optical illusion

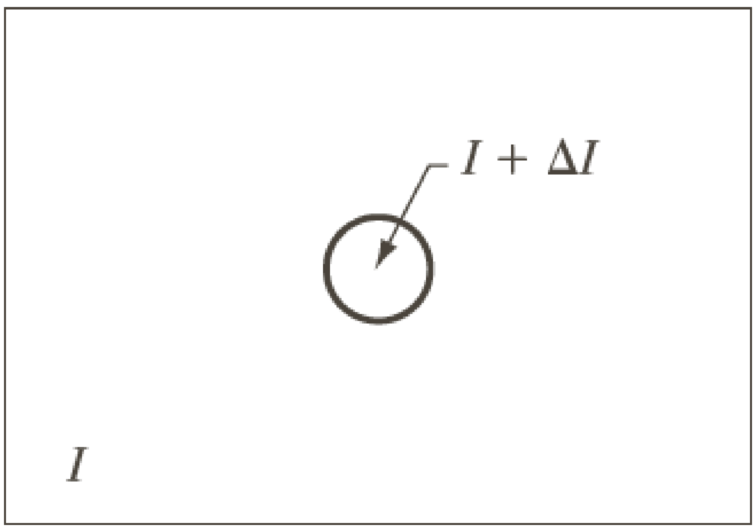


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

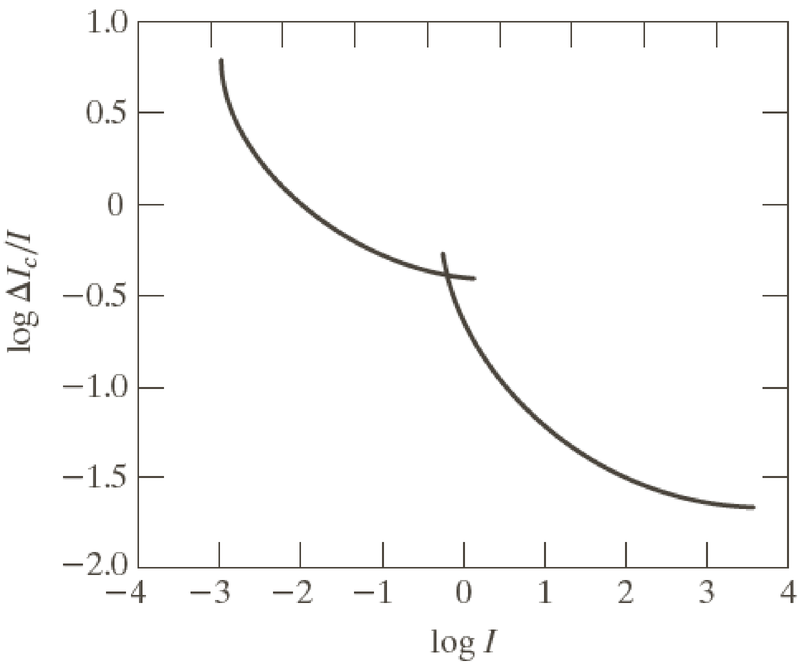


FIGURE 2.6 Typical Weber ratio as a function of intensity.

上海科技大学
ShanghaiTech University

Characteristics of Human Visual System

- Brightness adaption
- Simultaneous contrast
- **Mach band effect**
- Optical illusion

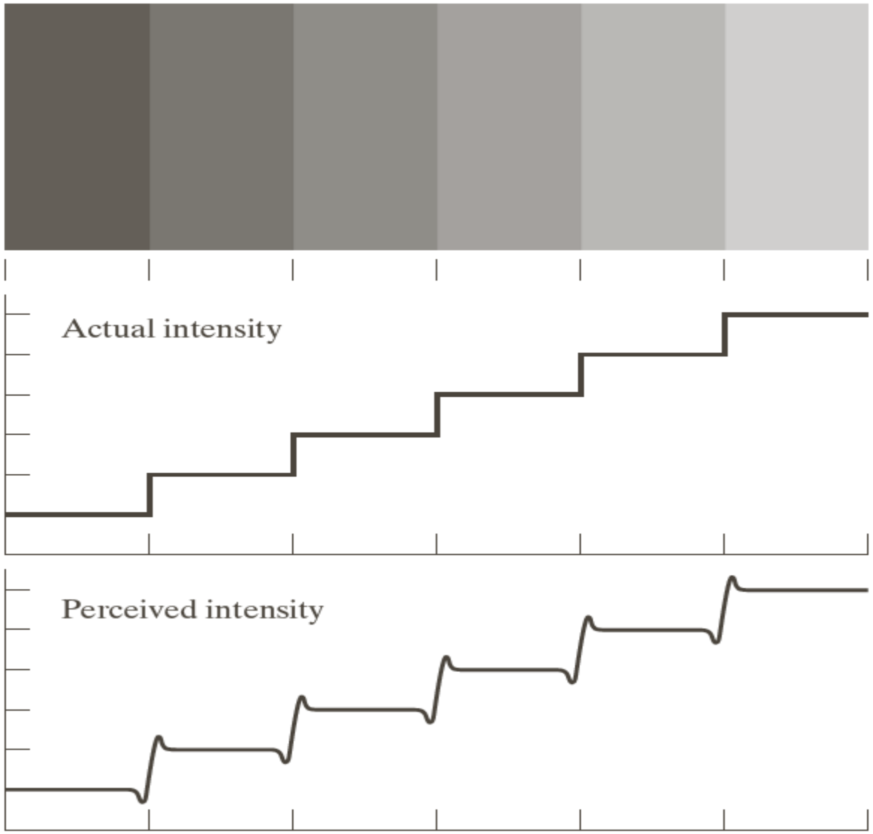
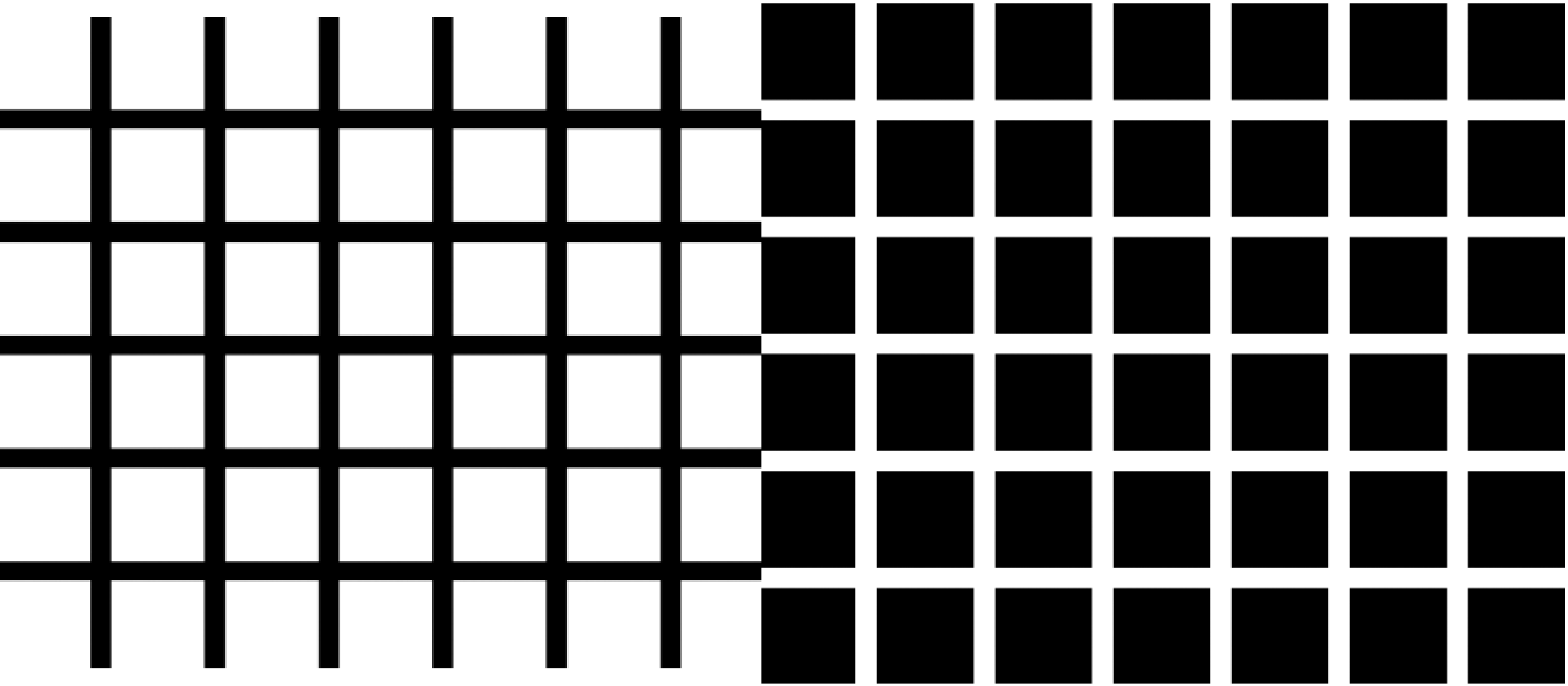


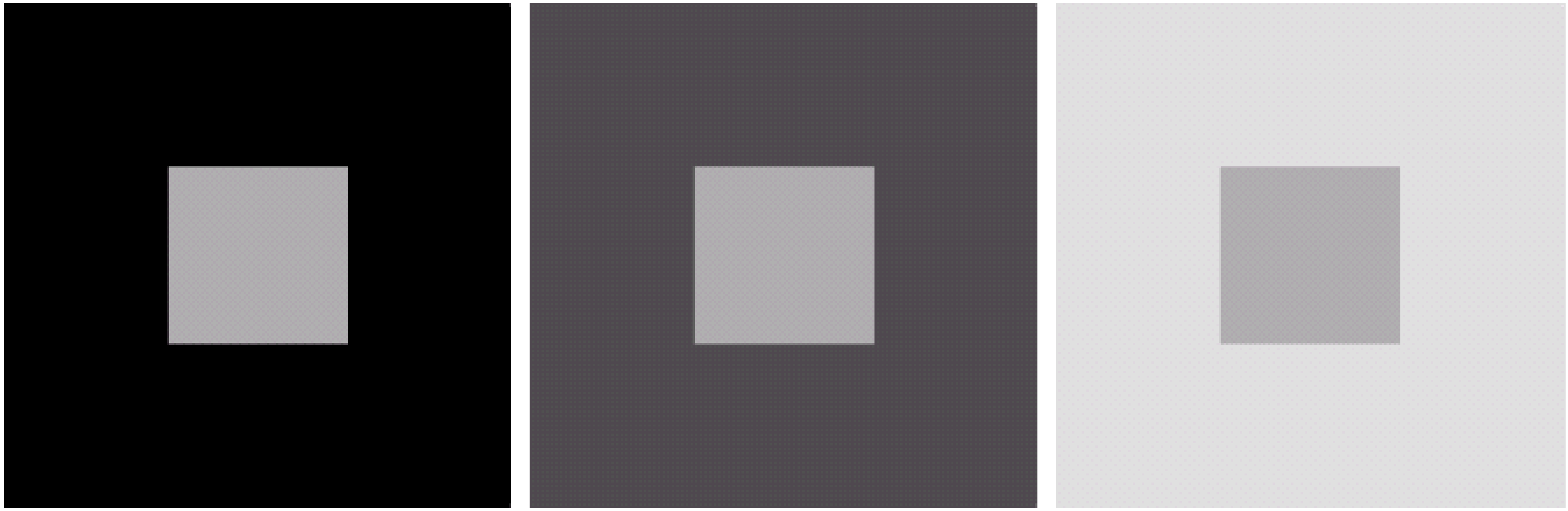
Figure 2.7: Perceived intensity is not a simple function of actual intensity

Characteristics of Human Visual System



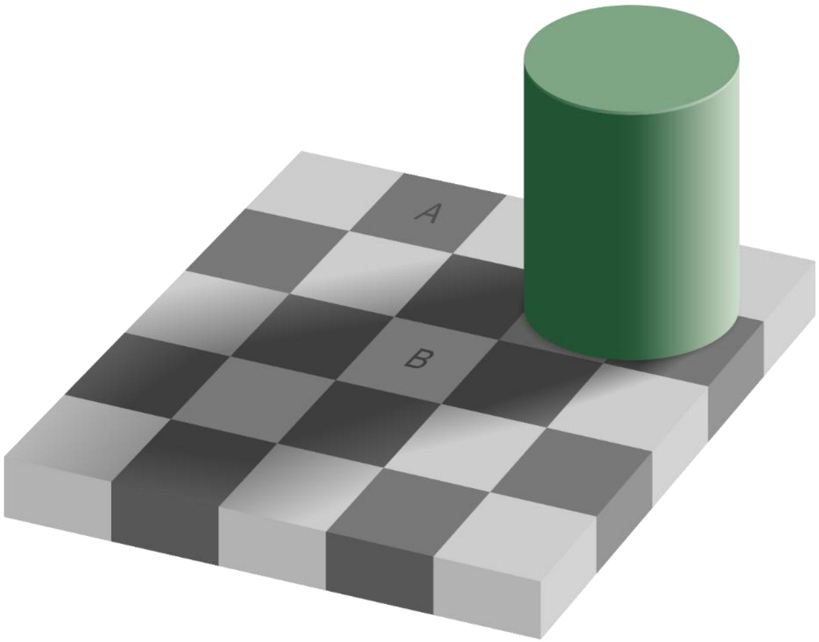
Characteristics of Human Visual System

- Brightness adaption
- **Simultaneous contrast**
- Mach band effect
- Optical illusion

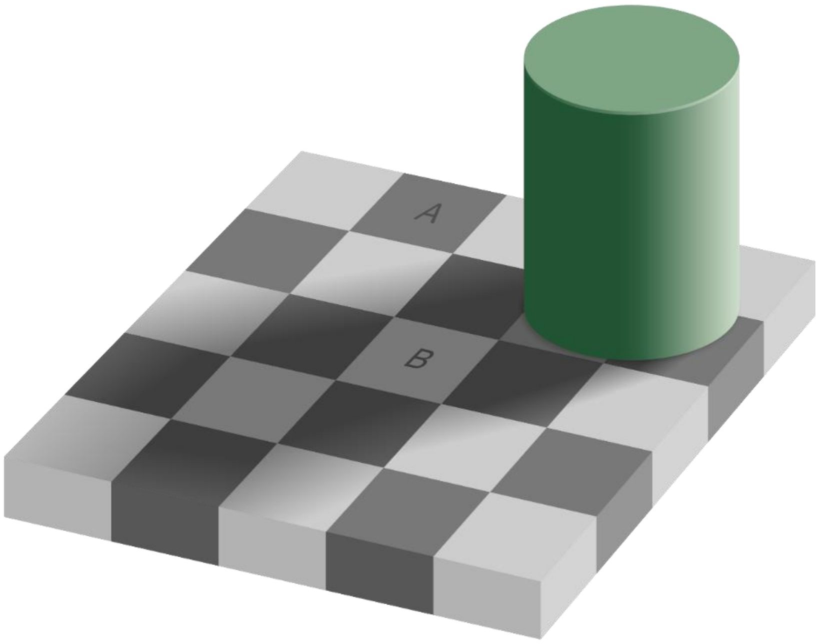


All the inner squares have the same intensity,
but they appear progressively darker as the background becomes lighter

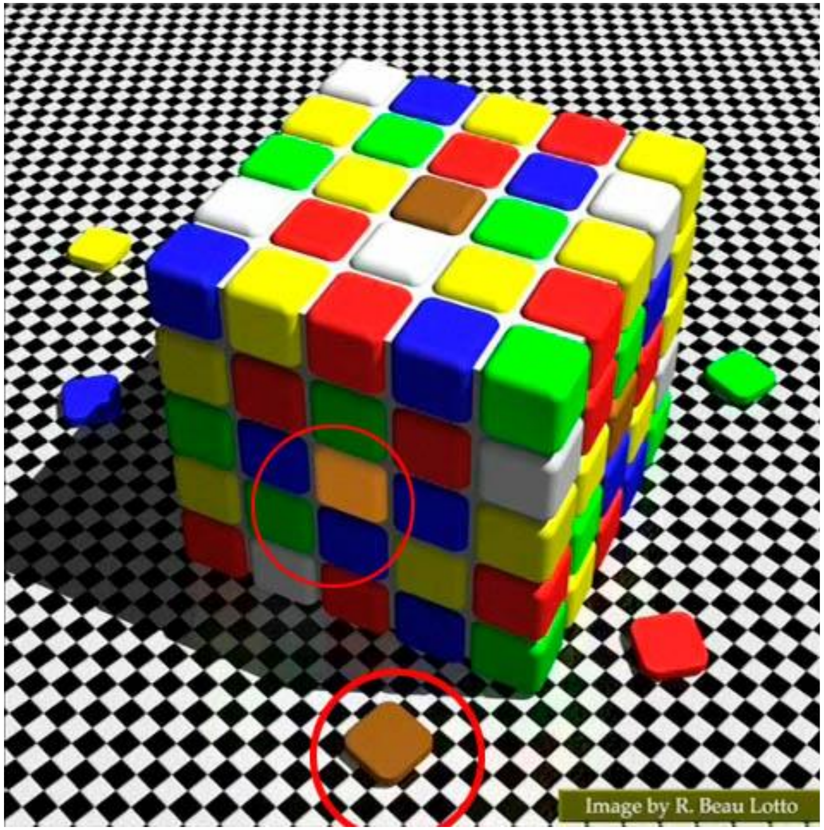
Checker shadow illusion



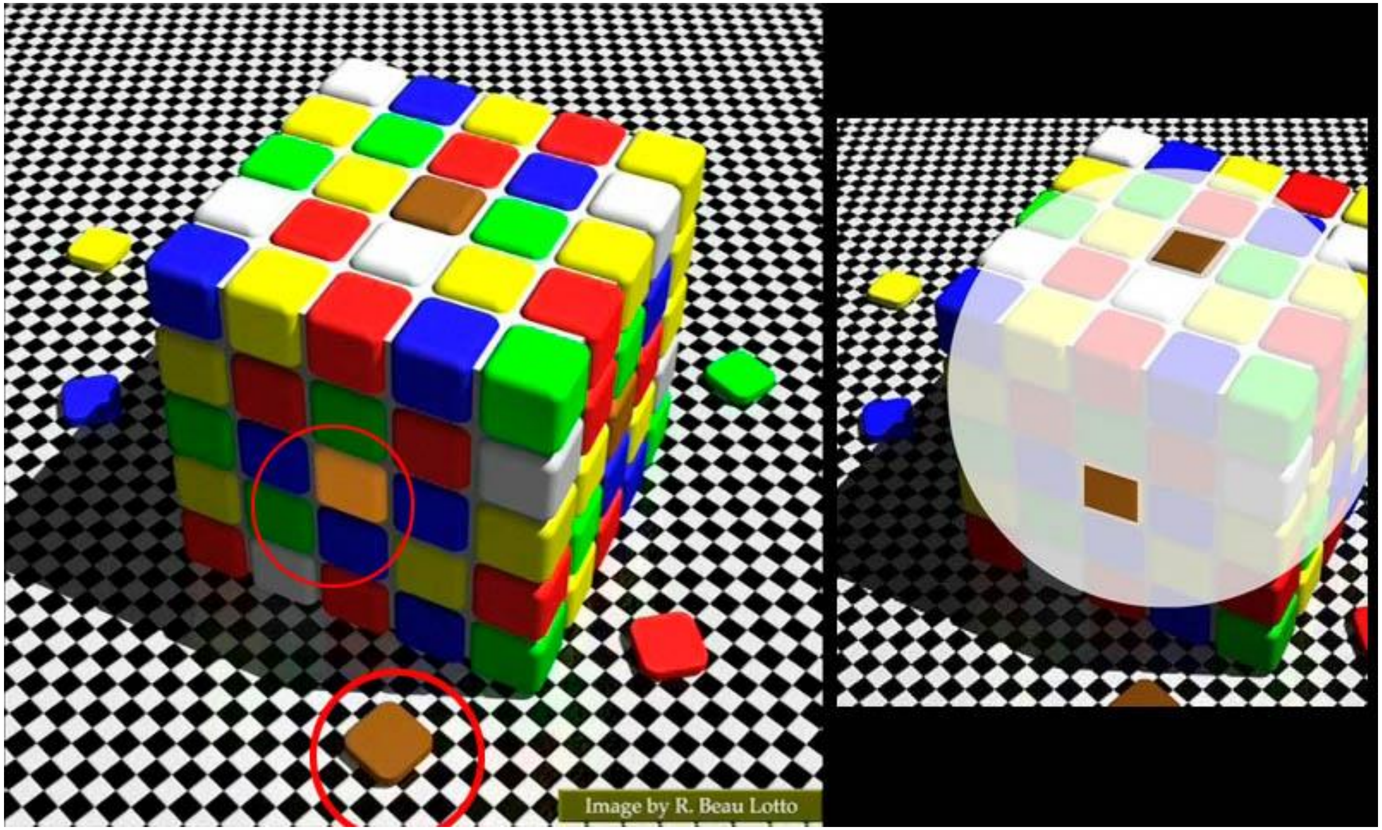
Checker shadow illusion



Cube

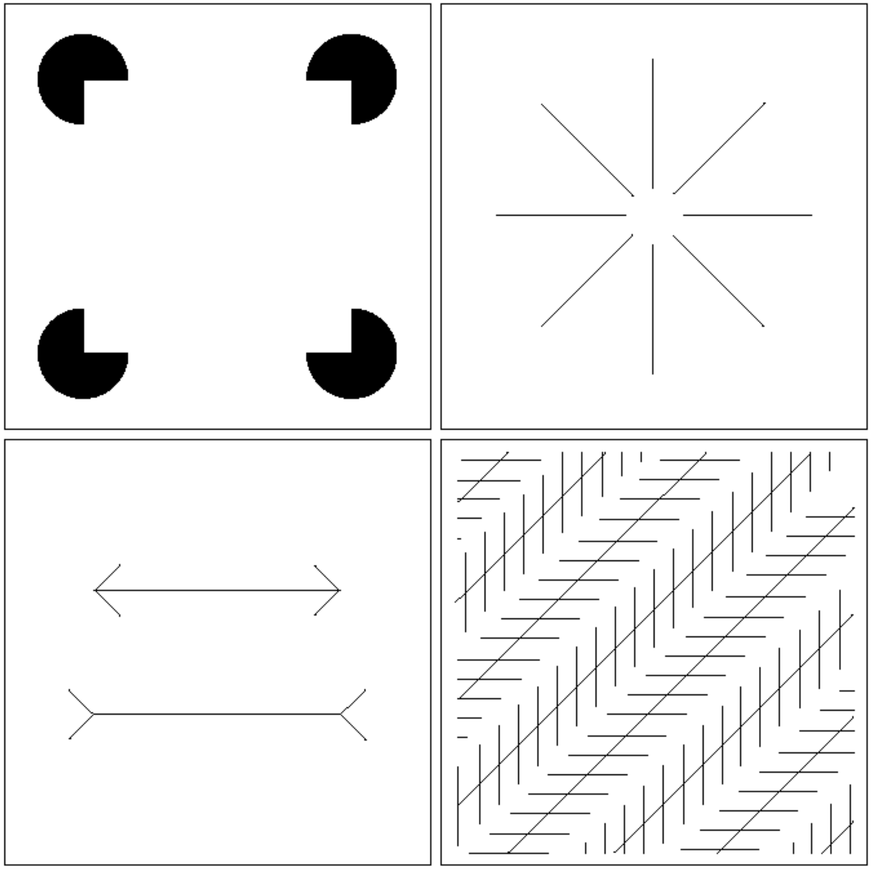


Cube

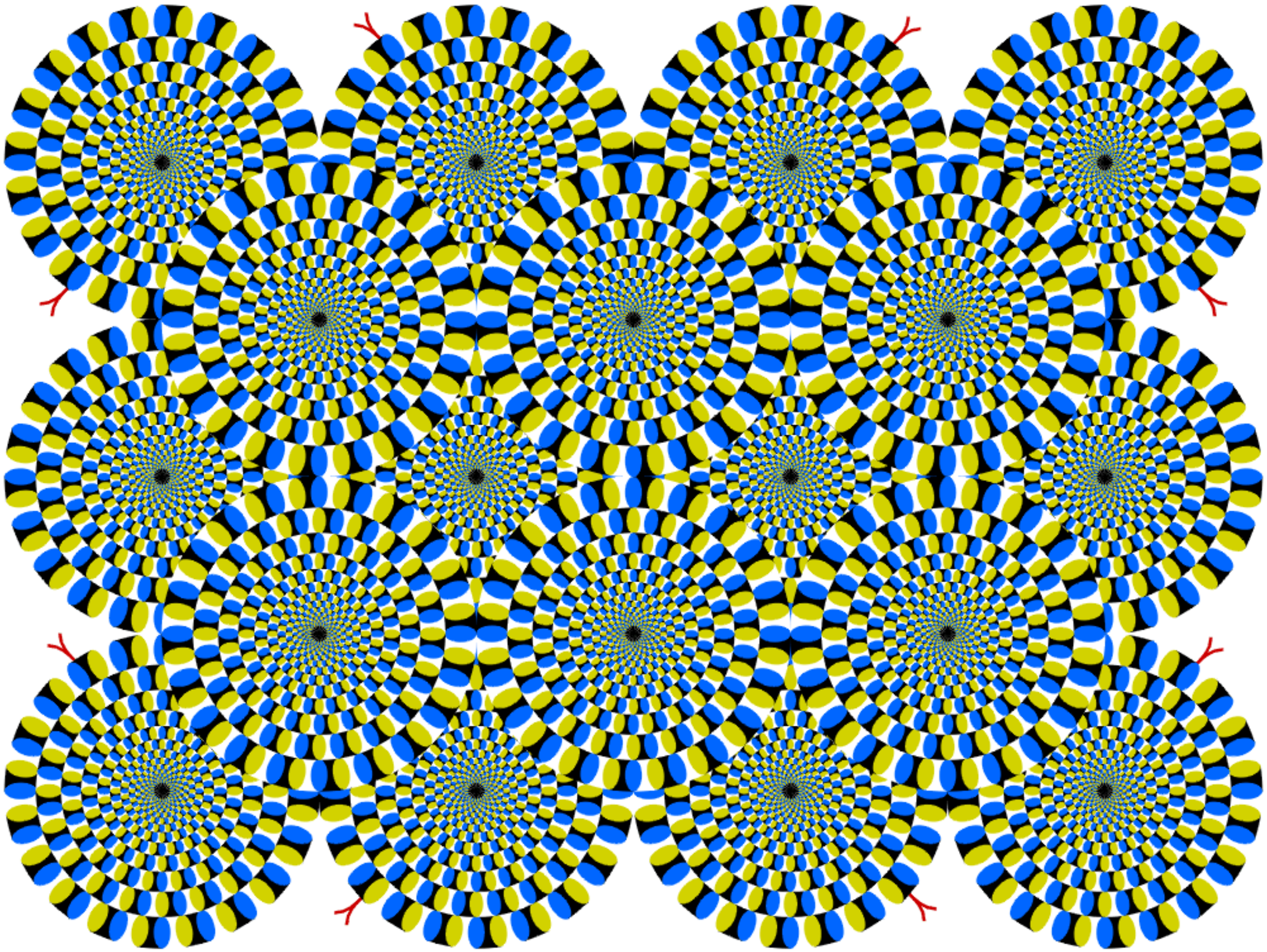


Characteristics of Human Visual System

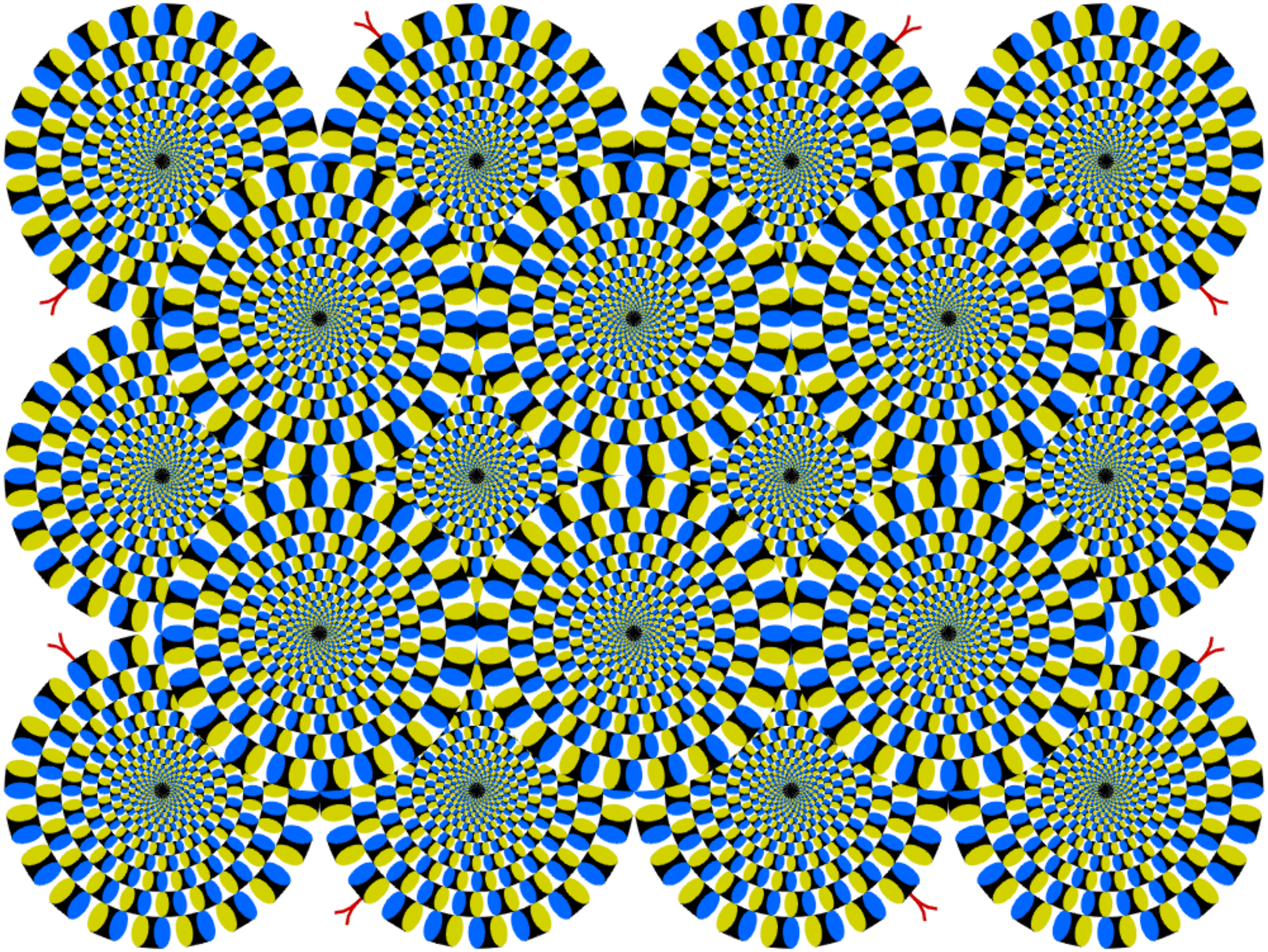
- Brightness adaption
- Simultaneous contrast
- Mach band effect
- **Optical illusion**

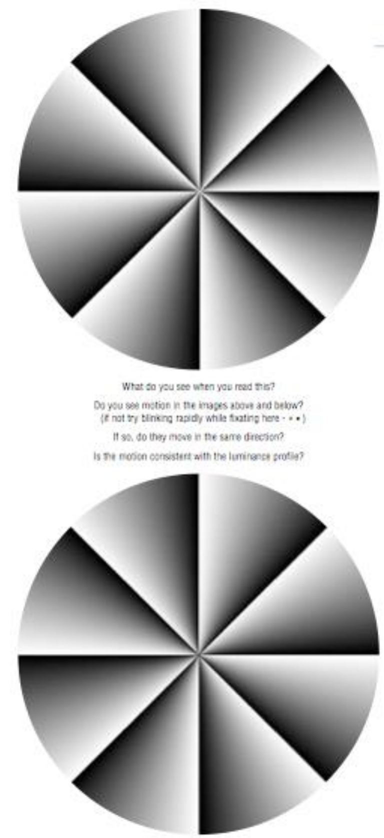


Characteristics of Human Visual System



Characteristics of Human Visual System





http://www.ritsumei.ac.jp/~akitaoka/index-e.html

Gender illusion

Your eyes can do more than you think...

- **McGurk effect**
- **https://www.youtube.com/watch?v=G-lN8vWm3m0**
- **https://www.youtube.com/watch?v=jtsfidRq2tw**

- **Rubber hand illusion**
- **https://www.youtube.com/watch?v=DphlhmtGRqI**

Take home message

- Be careful of your brain, it might be alien…

Pixel basics

- **Neighbors of Pixel**
- **Relationship between Pixels**
  - Adjacency
  - Connectivity
  - Regions
  - Boundaries
- **Distance measures**
  - Euclidean distance
  - City-block distance
  - Chessboard distance

Neighbors of Pixel

If a pixel p at coordinate (x, y)

- $N_4(p)$

- $N_D(p)$

- $N_8(p)$

Neighbors of Pixel

If a pixel p at coordinate (x, y)

- **$N_4(p)$**

(x+1, y), (x-1, y), (x, y+1), (x, y-1)

- **$N_D(p)$**

- **$N_8(p)$**

	q_1	
q_2	p	q_3
	q_4	

Neighbors of Pixel

If a pixel p at coordinate (x, y)

- **$N_4(p)$**

(x+1, y), (x-1, y), (x, y+1), (x, y-1)

- **$N_D(p)$**

(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)

- **$N_8(p)$**

r_1		r_2
	p	
r_3		r_4

Neighbors of Pixel

If a pixel p at coordinate (x, y)

- **$N_4(p)$**

(x+1, y), (x-1, y), (x, y+1), (x, y-1)

- **$N_D(p)$**

(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)

- **$N_8(p)$: $N_4(p) \cup N_D(p)$**

r_1	q_1	r_2
q_2	p	q_3
r_3	q_4	r_4

Adjacency

- **To define adjacency (邻接性) of pixels, we need identify**
- **Type of Neighbor**
  - $N_4(p)$, $N_D(p)$, $N_8(p)$
- **The set of intensity values V**
  - Binary image: $V = \{1\}$
  - Gray-scale image: $V = [L_{min}, L_{max}]$

e.g. V = [10, 15], which pixels are adjacent?

Why adjacency? (E.g. Used for edge detection)

q_1	p	q_2

0	0	0
0	1	1
0	0	0

Adjacency in a binary image

	q_1	
q_2	p	q_3

0	39	0
11	13	16
0	0	0

Adjacency in a gray-scale image

Adjacency

Types of Adjacency:

- 4-adjacency
- 8-adjacency
- M-adjacency (mixed adjacency)

Adjacency

Types of Adjacency:

- 4-adjacency
  - $p, q \in V$
  - $q \in N_4(p)$
- 8-adjacency
- M-adjacency (mixed adjacency)

r_{11}	r_{12}	r_{13}
r_{21}	r_{22}	r_{23}
r_{31}	r_{32}	r_{33}

0	1	1
0	1	0
0	0	1

Adjacency

Types of Adjacency:

- 4-adjacency
- 8-adjacency
  - $p, q \in V$
  - $q \in N_8(p)$
- M-adjacency (mixed adjacency)

$$\begin{array}{ccc} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{array}$$

$$\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}$$

Adjacency

Types of Adjacency:

- 4-adjacency
- 8-adjacency
- M-adjacency (mixed adjacency)
  - $p, q \in V$
  - $q \in N_4(p)$ or
  - $q \in N_D(p)$ and $N_4(p) \cap N_4(q) \notin V$

$$\begin{array}{ccc} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{array}$$

$$\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}$$

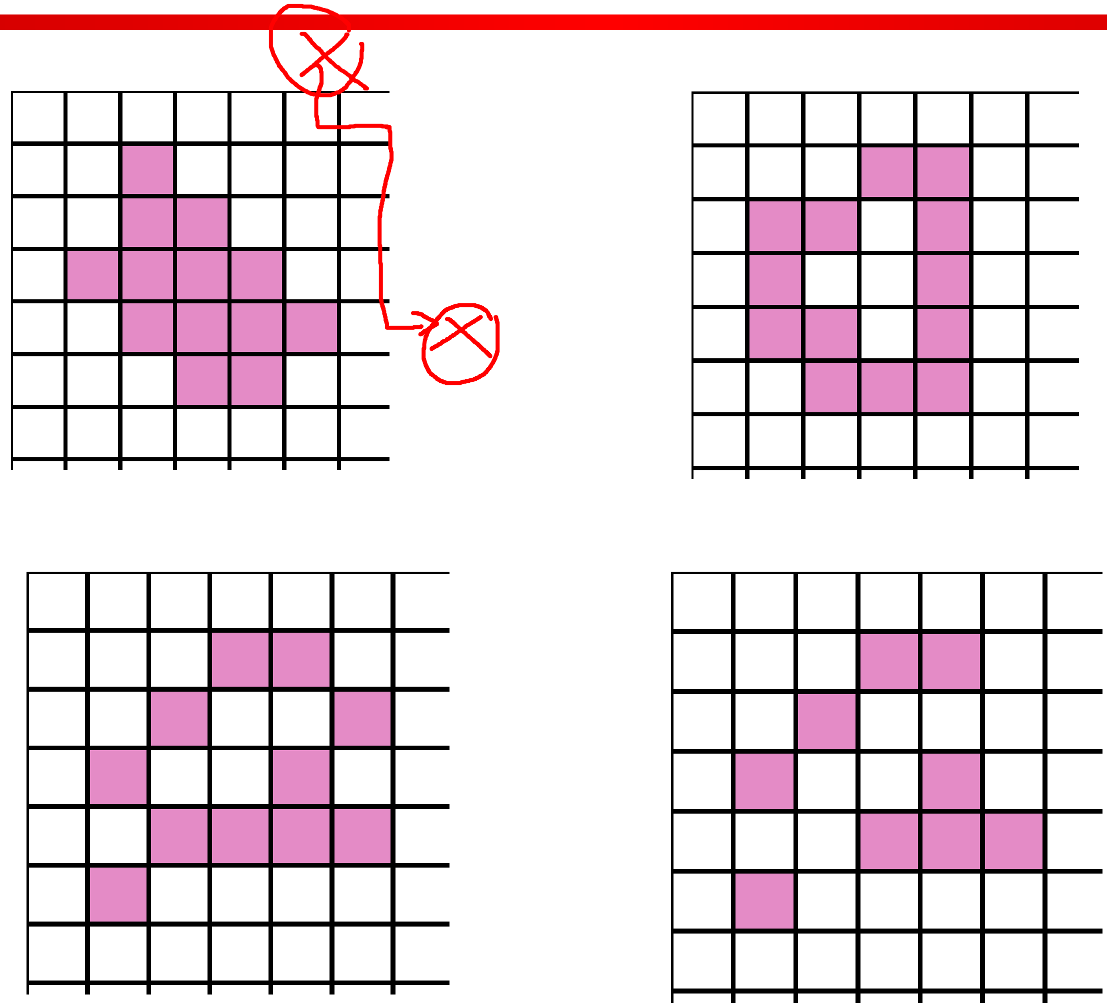
M-adjacency is a modification of 8-adjacency, to eliminate the ambiguities that may result from using 8-adjacency

Connectivity

- Important concept used in establishing boundaries of objects and components of regions in an image (连通性)
  - Path
  - Connected
  - Connected component
  - Connected set

$$\begin{matrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{matrix}$$

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \qquad \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$



Region

- ***R***: a subset of an image which is also a connected set
- Adjacent region
- Disjoint region

$$
\begin{array}{l}
\left.\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right\} R_i \\
\left.\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{array}\right\} R_j
\end{array}
$$

Boundary

A set of pixels that are adjacent to pixels in the complement of ***R***.

- **Inner border and outer border**
- **Image border**
- **Edge**

0	0	0	0	0
0	1	1	0	0
0	1	1	0	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0

Boundary

- A set of pixels that are adjacent to pixels in the complement of ***R***.
- Inner border and outer border
- Image border
- **Edge**

Pixel basics

- **Neighbors of Pixel**
- **Relationship between Pixels**
  - Adjacency
  - Connectivity
  - Regions
  - Boundaries
- **Distance measures**
  - Euclidean distance
  - City-block distance
  - Chessboard distance

Distance Measures

- For pixels p, q and z, with coordinates (x, y), (s, t) and (v, w), D is a distance function or metric if

  - $D(p, q) \geqq 0$ ($D(p, q) = 0$ only if $p = q$)
  - $D(p, q) = D(q, p)$
  - $D(p, z) \leq D(p, q) + D(q, z)$

Distance Measures

- **Euclidean distance:**

  $$D_e(p, q) = [(x-s)^2+(y-t)^2]^{1/2}$$

- **City-block distance:**
- **Chessboard distance**

		2		
	1.4	1	1.4	
2	1	0	1	2
	1.4	1	1.4	
		2		

Distance Measures

- Euclidean distance:

  $D_e(p, q) = [(x-s)^2+(y-t)^2]^{½}$

- City-block distance:

  $D_4(p, q) = |x-s|+|y-t|$

- Chessboard distance

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

Distance Measures

- **Euclidean distance:**

  $D_e(p, q) = [(x-s)^2+(y-t)^2]^{1/2}$

- **City-block distance:**

  $D_4(p, q) = |x-s|+|y-t|$

- **Chessboard distance**

  $D_8(p, q) = \max(|x-s|, |y-t|)$

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Distance Measures

- D_m distance is defined as the shortest m-path between the point

	r_{12}	1
r_{21}	r_{22}	
1		

$$\mathbf{D_m = ?}$$

Distance Measures

- No m-path between the point

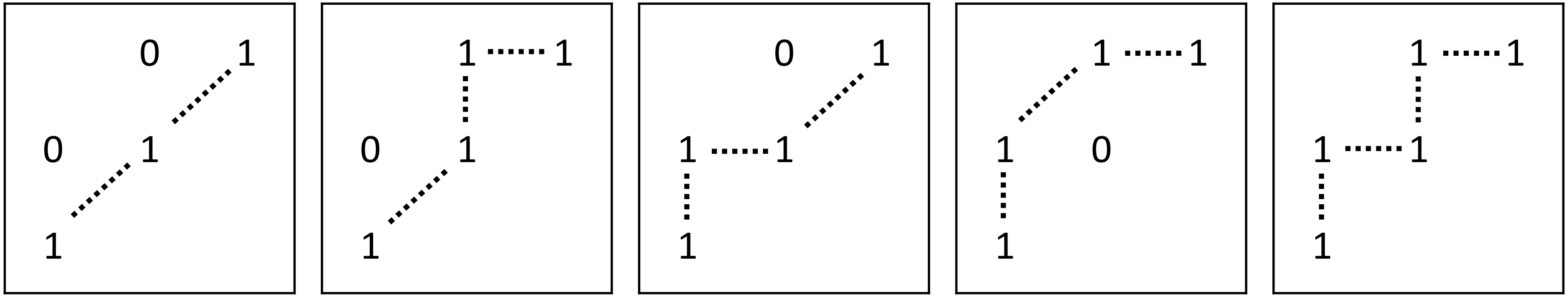
	0	1
0	0	
1		

	1	1
0	0	
1		

	0	1
1	0	
1		

Distance Measures

- **D_m distance** is different by the values of r_{12}, r_{21} and r_{22}



$D_m = 2$ $D_m = 3$ $D_m = 3$ $D_m = 3$ $D_m = 4$

Arithmetic Operation

- Addition

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

- Subtraction

$$d(x,y) = f(x,y) - g(x,y)$$

- Multiplication

$$p(x,y) = f(x,y) \times g(x,y)$$

- Division

$$v(x,y) = f(x,y) \div g(x,y)$$

Array and Matrix Operation

Consider two 2 x 2 image

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

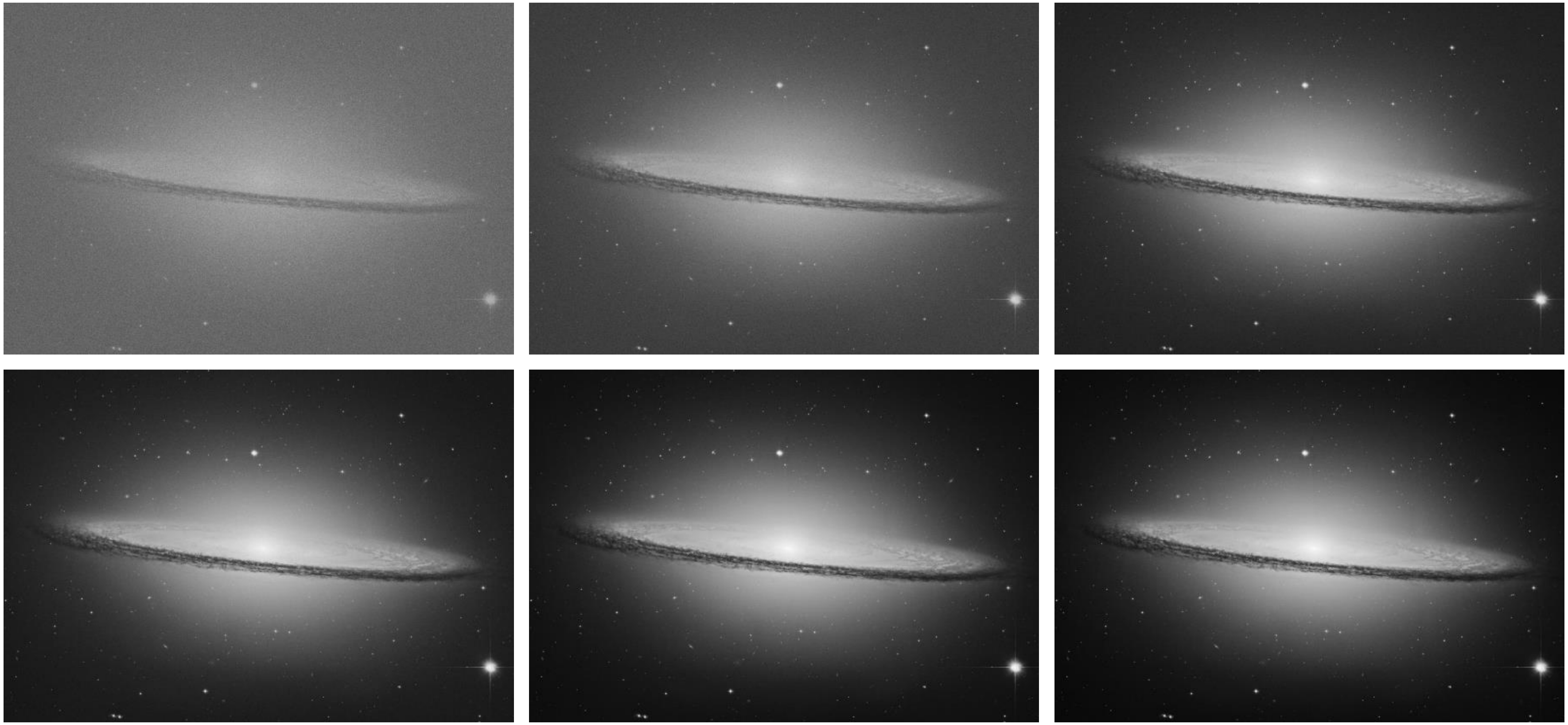
- Array product

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{11} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

- Matrix product

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

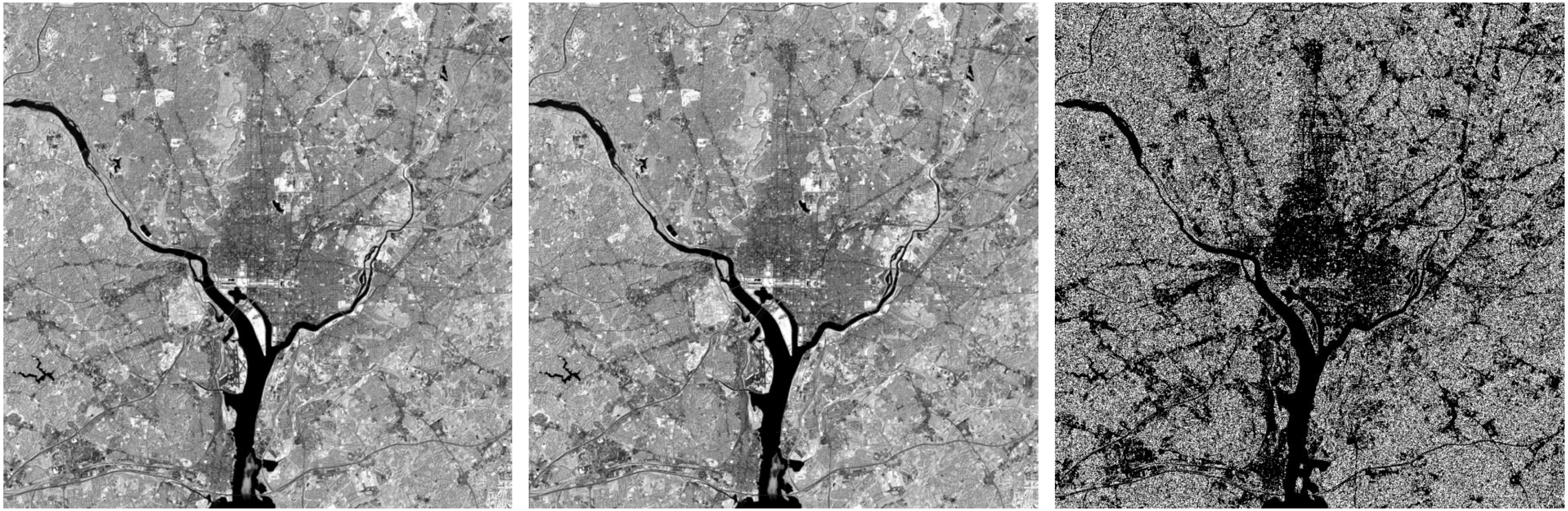
Image Addition



a b c
d e f

FIGURE 2.29 (a) Sample noisy image of the Sombrero Galaxy. (b)-(f) Result of averaging 10, 50, 100, 500, and 1,000 noisy images, respectively. All images are of size 1548×2238 pixels, and all were scaled so that their intensities would span the full [0, 255] intensity scale. (Discovered in 1767, the Sombrero Galaxy is 28 light years from Earth. Original image courtesy of NASA.)

Image Subtraction



a b c

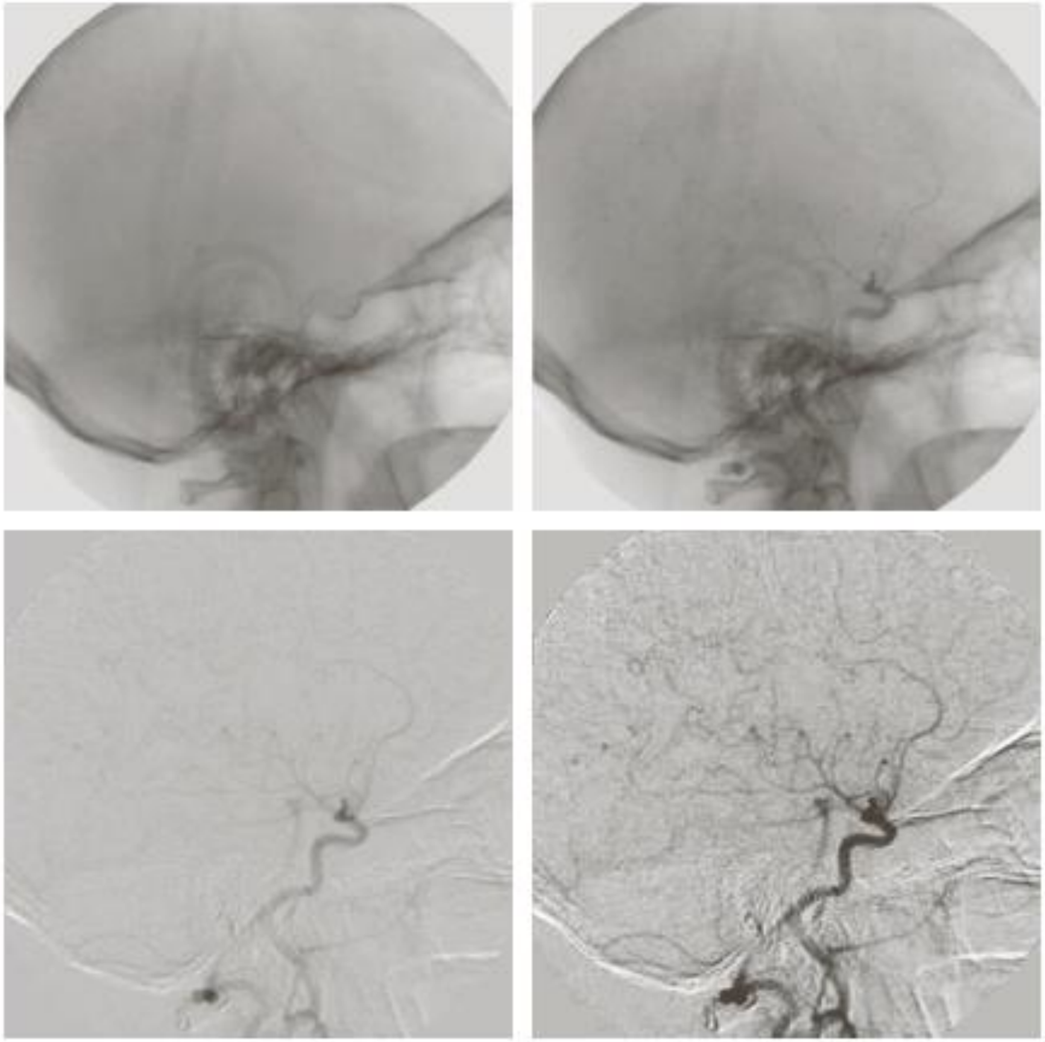
FIGURE 2.30 (a) Infrared image of the Washington, D.C. area. (b) Image resulting from setting to zero the least significant bit of every pixel in (a). (c) Difference of the two images, scaled to the range [0, 255] for clarity. (Original image courtesy of NASA.)

Image Subtraction

a b c

FIGURE 2.31 (a) Difference between the 930 dpi and 72 dpi images in Fig. 2.23. (b) Difference between the 930 dpi and 150 dpi images. (c) Difference between the 930 dpi and 300 dpi images.

Image Subtraction



a b
c d

FIGURE 2.32 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of the Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

Image Multiplication

$$p(x,y) = f(x,y) \times g(x,y)$$

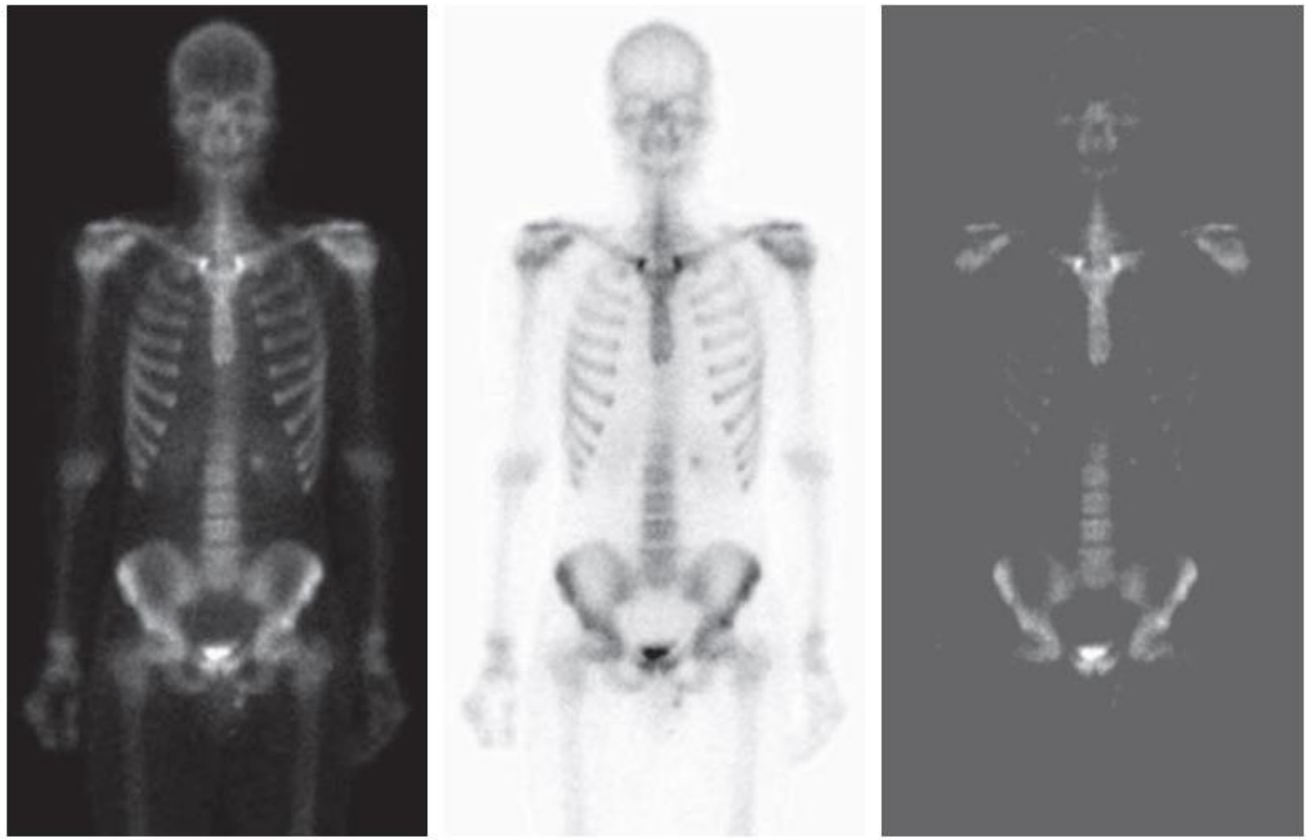
Image Division

$g(x,y) = f(x,y)\,h(x,y)$ $\qquad$ $h(x,y)$ $\qquad$ $f(x,y)$

$$f(x,y) = g(x,y)\,/\,h(x,y)$$

Background removal

The negative

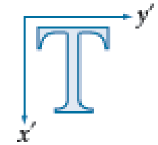
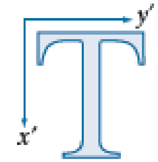
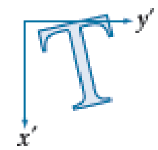
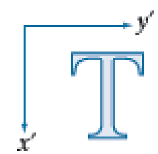
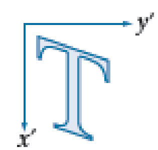
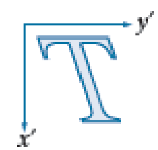


a b c

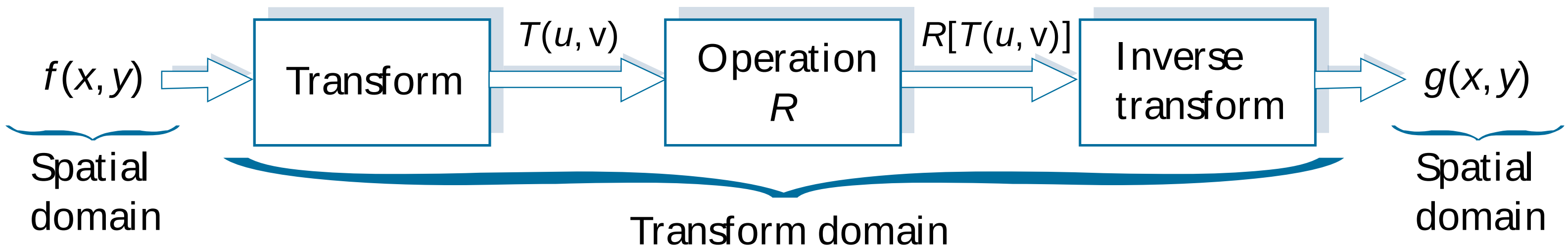
FIGURE 2.36
Set operations involving grayscale images. (a) Original image. (b) Image negative obtained using grayscale set complementation. (c) The union of image (a) and a constant image. (Original image courtesy of G.E. Medical Systems.)

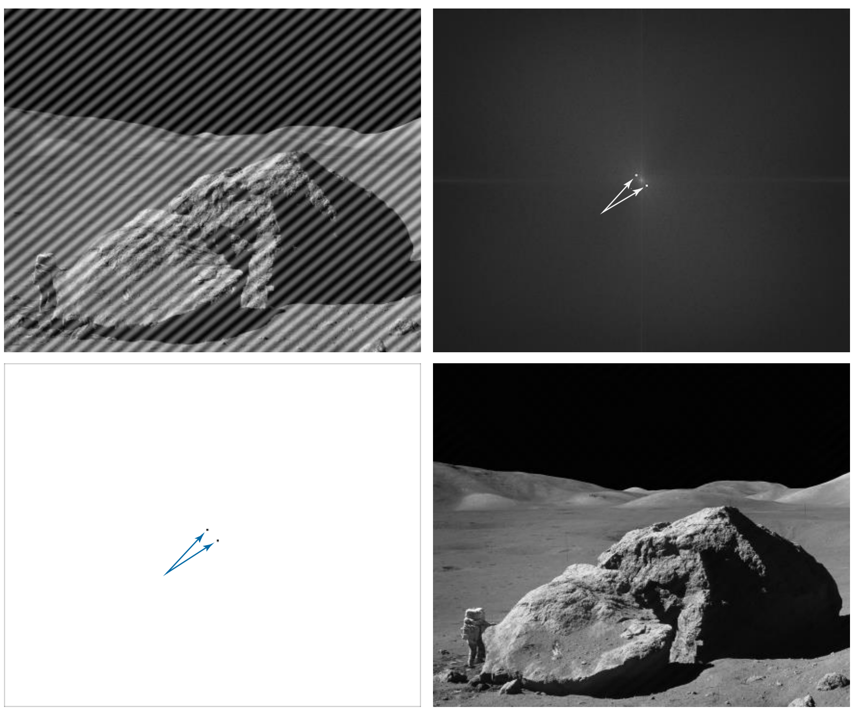
Affine transformation

TABLE 2.3
Affine transformations based on Eq. (2-45).

Transformation Name	Affine Matrix, A	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
Scaling/Reflection (For reflection, set one scaling factor to −1 and the other to 0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
Rotation (about the origin)	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x\cos\theta - y\sin\theta$ $y' = x\sin\theta + y\cos\theta$	
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

Transform domain





a b
c d

FIGURE 2.45 (a) Image corrupted by sinusoidal interference. (b) Magnitude of the Fourier transform showing the bursts of energy caused by the interference (the bursts were enlarged for display purposes). (c) Mask used to eliminate the energy bursts. (d) Result of computing the inverse of the modified Fourier transform. (Original image courtesy of NASA.)

上海科技大学
ShanghaiTech University