Chapter 2: Digital Image Fundamentals

- Elements of Visual Perception
- Light and EM Spectrum
- Image Sensing and Acquisition
- Image Sampling and Quantization
- Basic Relationships between Pixels
- Mathematical Tools in DIP

2.1 Elements of Visual Perception

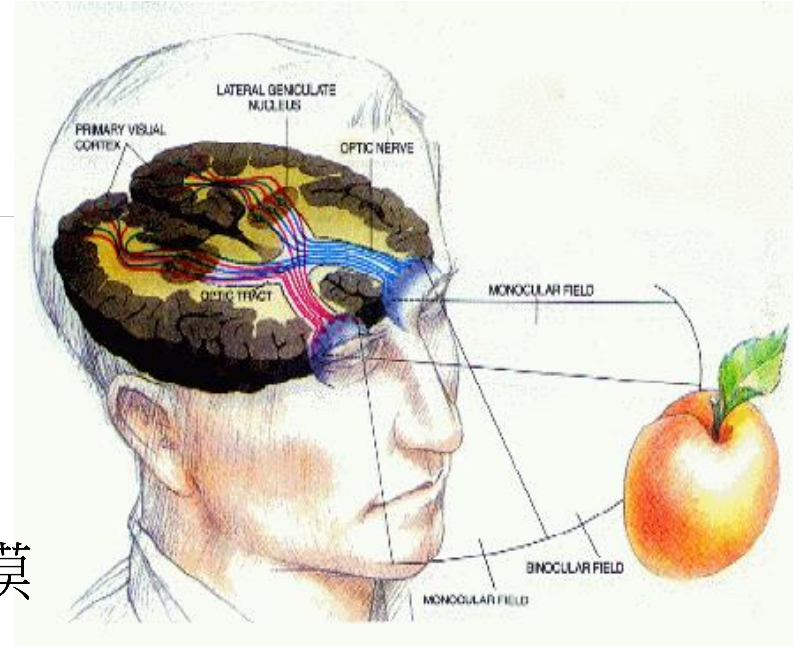
- Structure of Human Eye
- Image Formation in the Eye
- Brightness Adaptation and Discrimination

Why study human perception?

- DIP field is built on a foundation of mathematical and probabilistic formulation, but human intuition and analysis play a central role in choosing techniques.

➔ Want to learn the physical limitation of human vision.

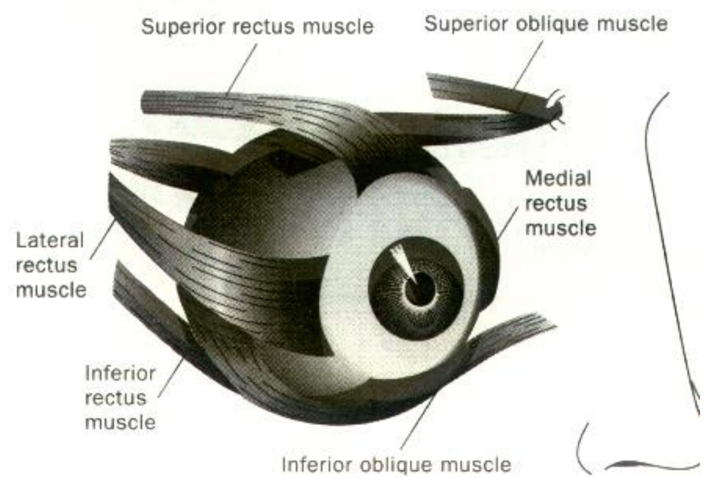
Structure of Human Eye



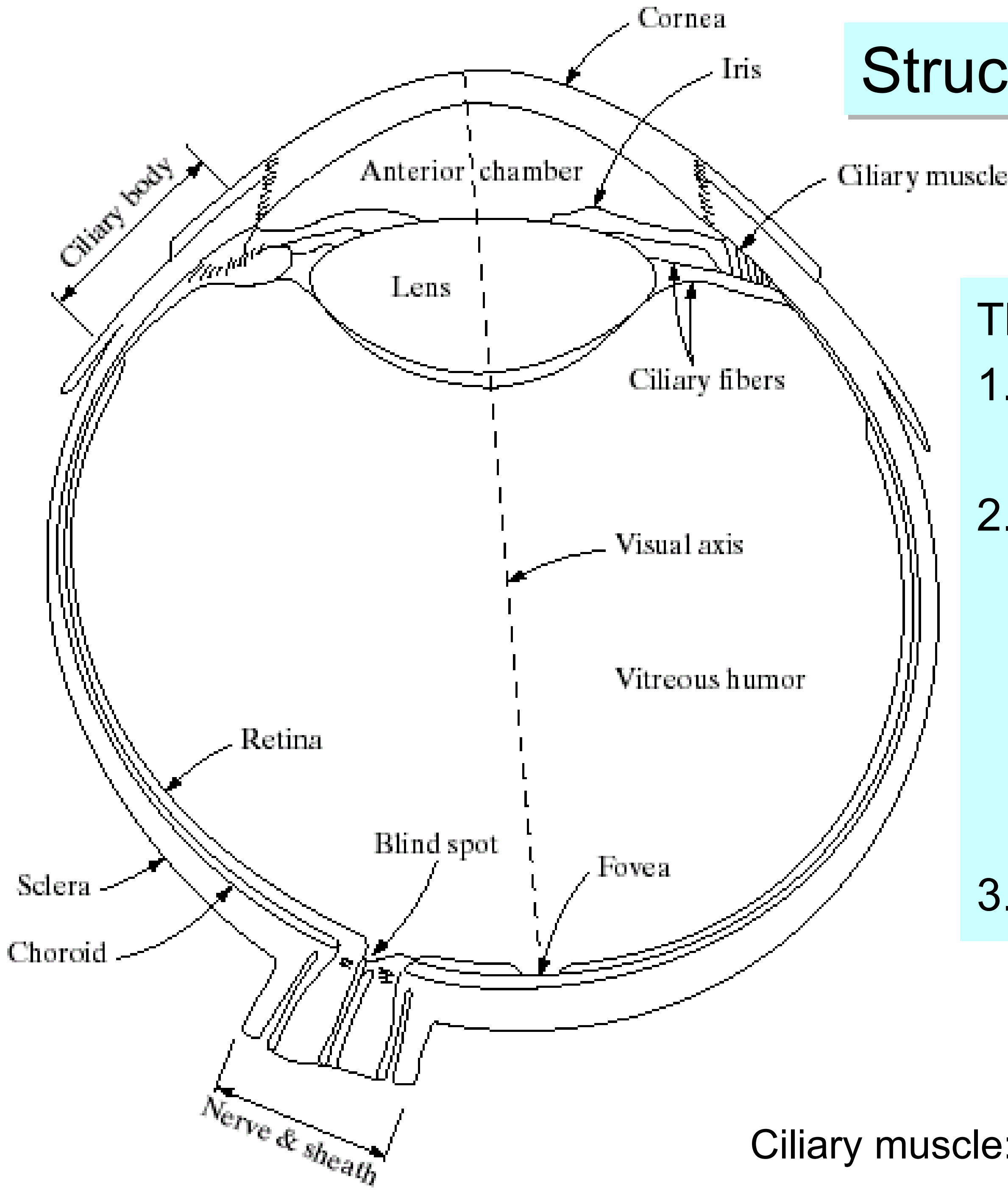
人類的眼球 (~20mm):

- 三層膜
  -- 鞏膜及角膜、脈絡膜、視網膜
- 兩個房
  -- 前房 Anterior chamber
  、後房 Posterior chamber
- 六條控制眼動的肌肉

-- 包附著厚厚的保護性脂肪層，使得眼球能安全地在眼窩裡轉動。



Structure of Human Eye



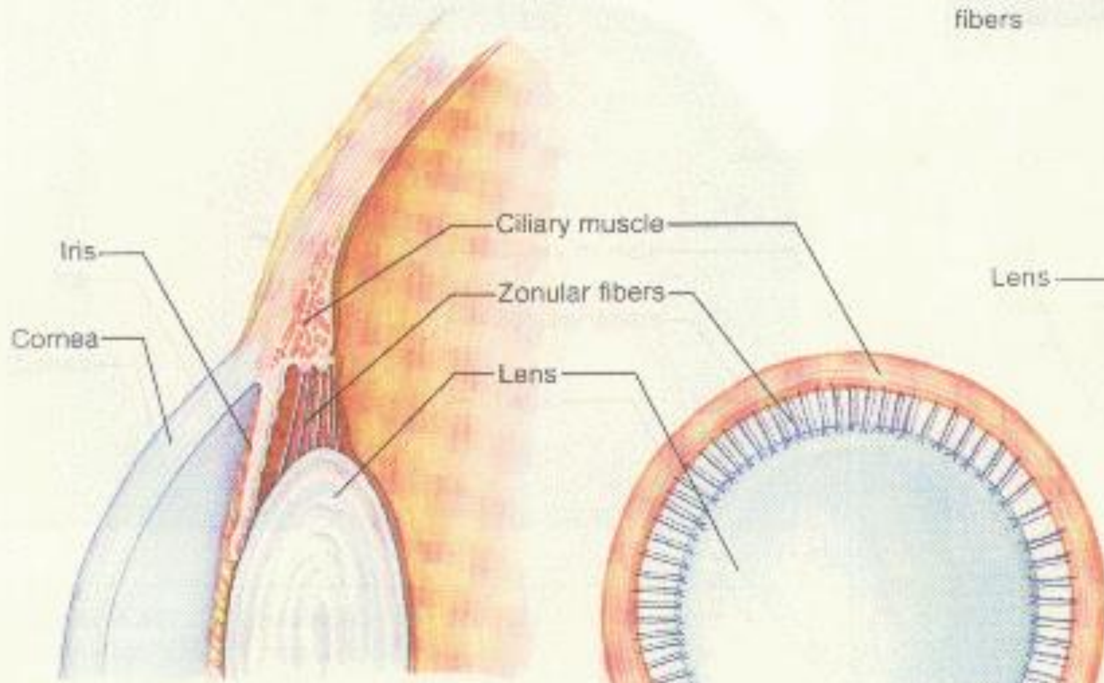
Three membranes(膜):

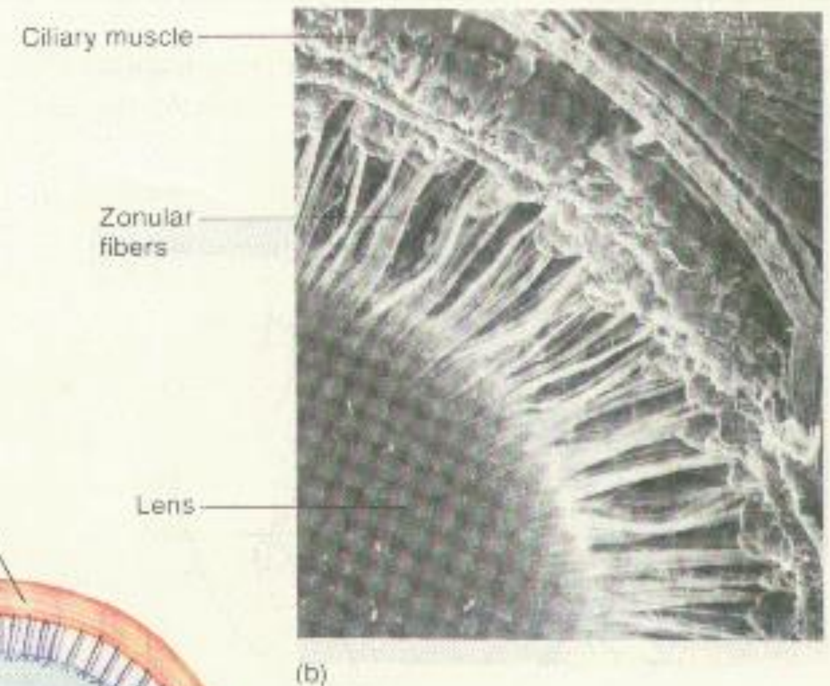
1. Sclera 鞏膜、眼白
   + Cornea 角膜、黑眼珠
2. Choroid 脈絡膜
   (blood vessel)
   (heavily pigmented)
   + Ciliary body 睫狀體
   + Iris diaphragm 虹膜
     * Pupil 瞳孔
3. Retina 視網膜

Ciliary muscle: control the shape of lens

Lens(水晶體), Cornea, Iris, Limbus

Corneal limbus 角膜緣
-- boundary between the white sclera and the dark iris



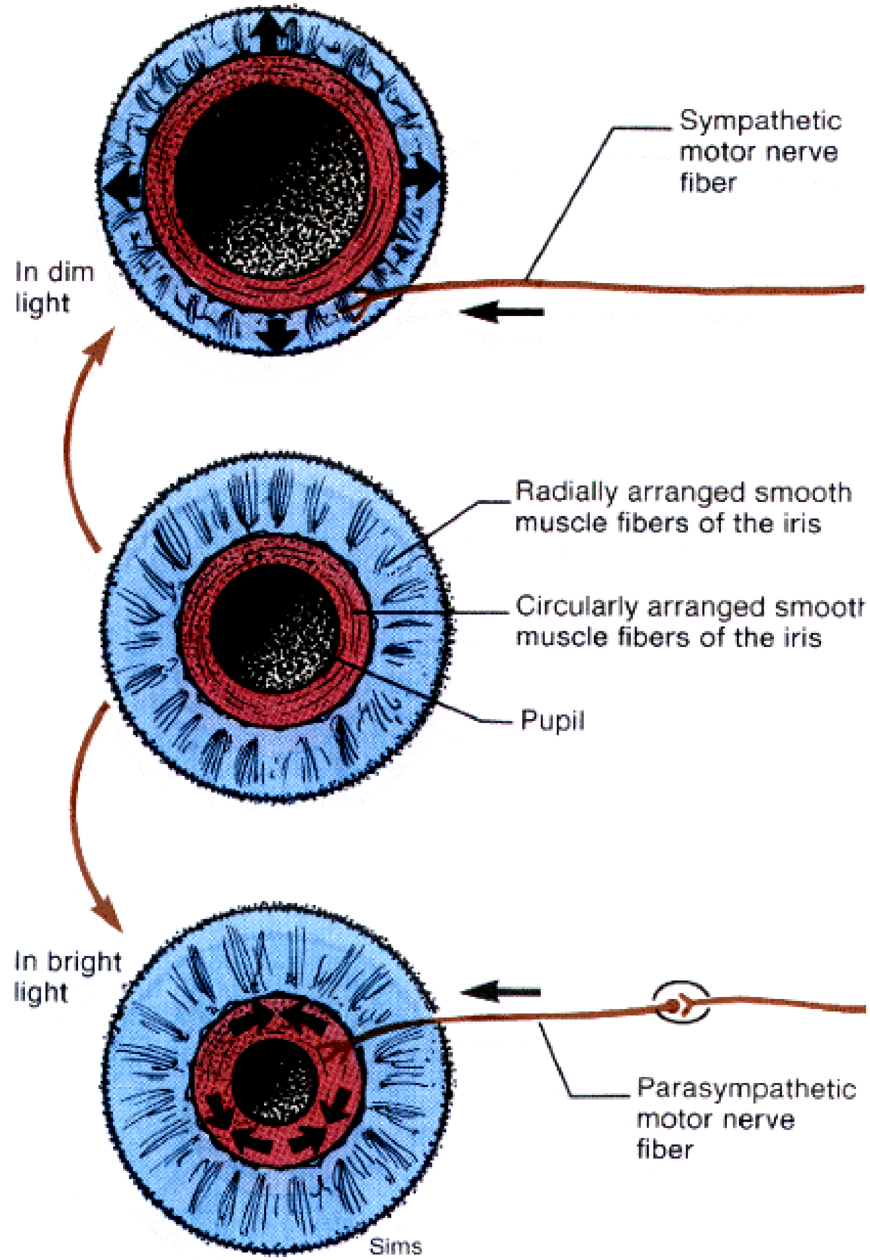


Iris and Pupil

- 瞳孔大小的調節機制是根據環境的光亮強度 (直徑2-8mm) :

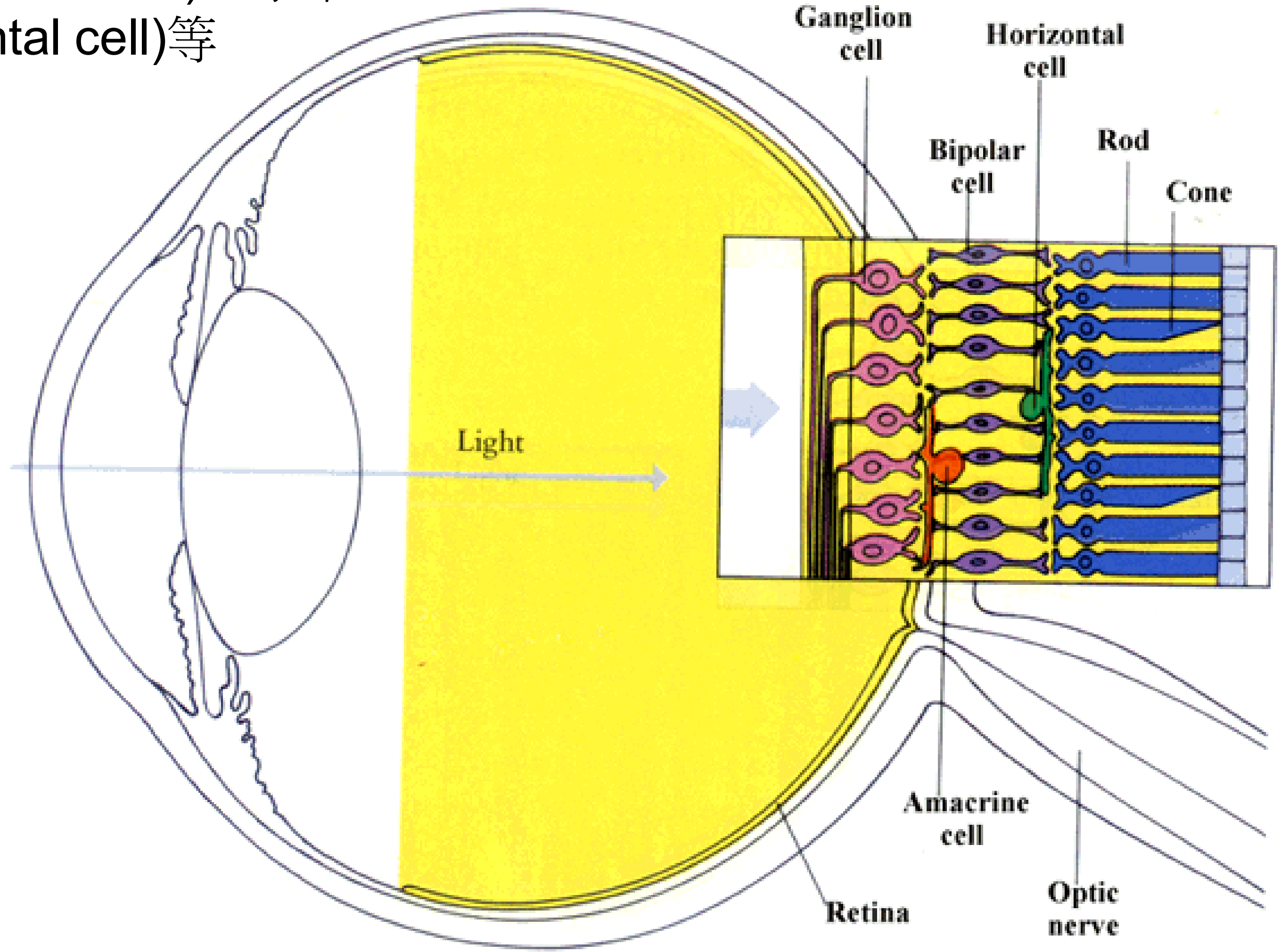
1. 瞳孔放大: 放射狀纖維 (縱走肌)收縮、環狀纖維 (環狀肌)舒張，使得中間的孔徑放大

2. 瞳孔縮小: 放射狀纖維舒張、環狀纖維向內收縮



視網膜上最重要的兩種視覺接受器：

- 錐細胞(cone): 6~7 million
- 桿細胞(rod):75~150 million
- 其他還有節細胞(ganglion cell)、無軸突細胞(amacrine cell)、兩極細胞(bipolar cell)、水平細胞(horizontal cell)等

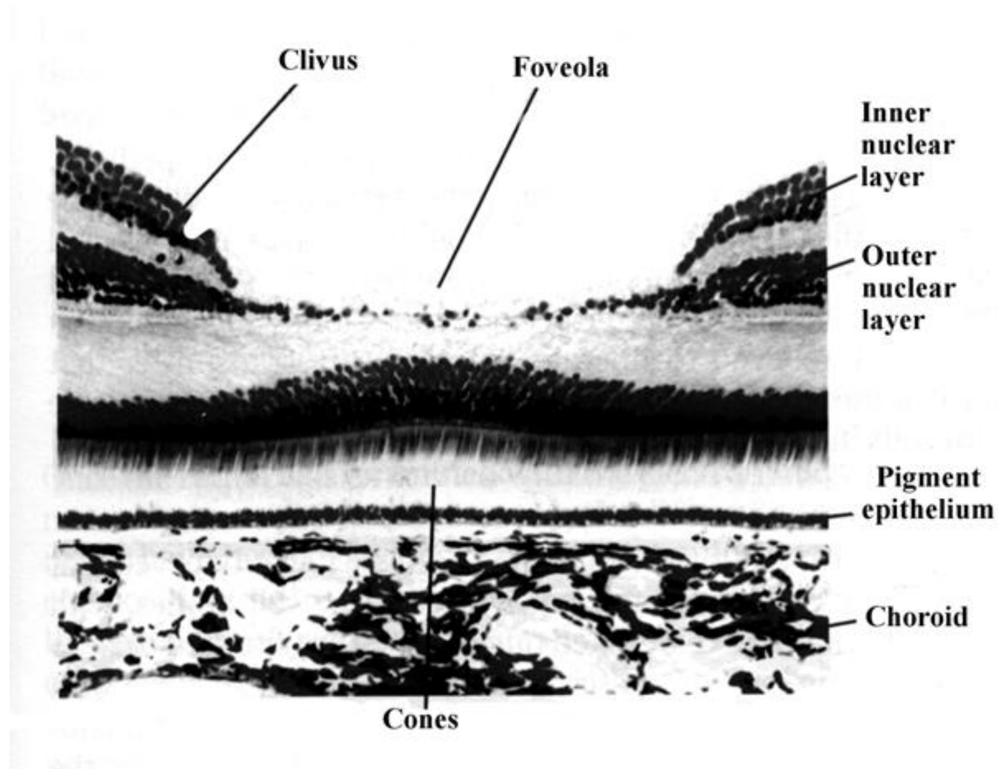


Fovea / Foveola 中央窩地區
(~1.5mm in diameter)

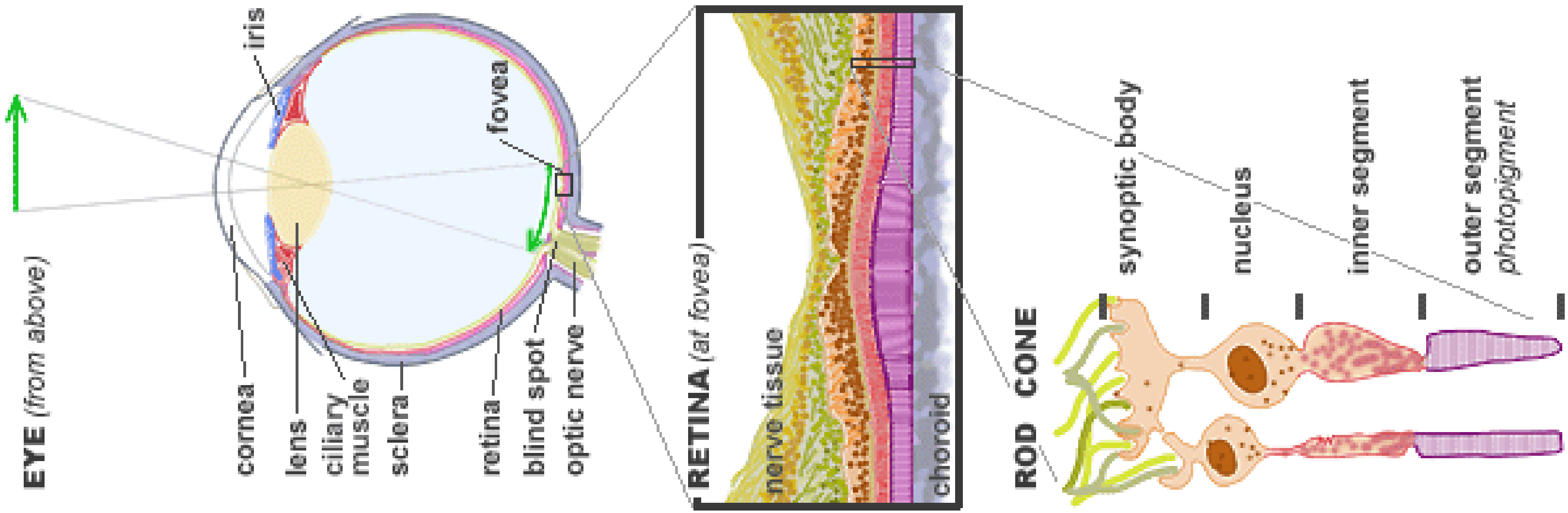
中央窩地區

- 神經層較薄，沒有阻礙光線的血管層
- 兩極細胞、水平細胞、與節細胞等神經元都盡量地移到兩旁

→ 此區的光線能更順利的達到視覺受器上。

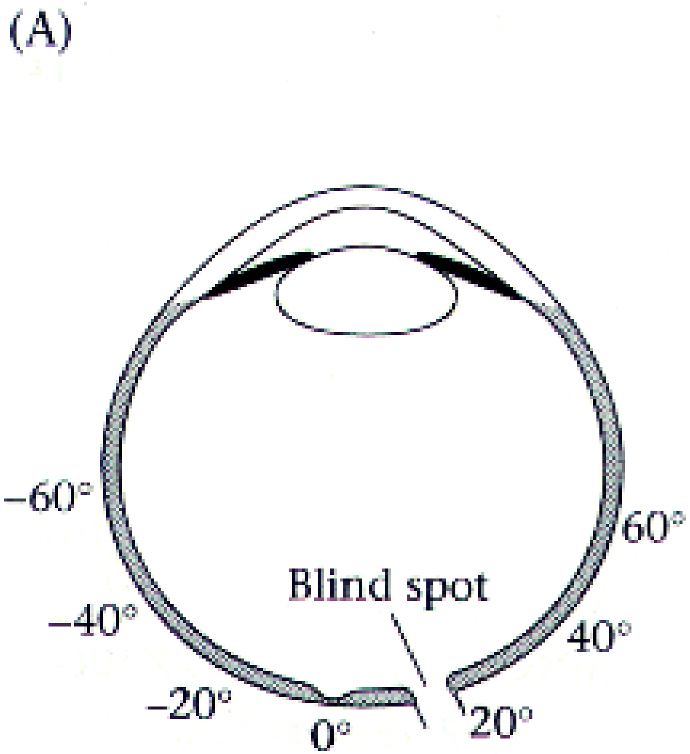


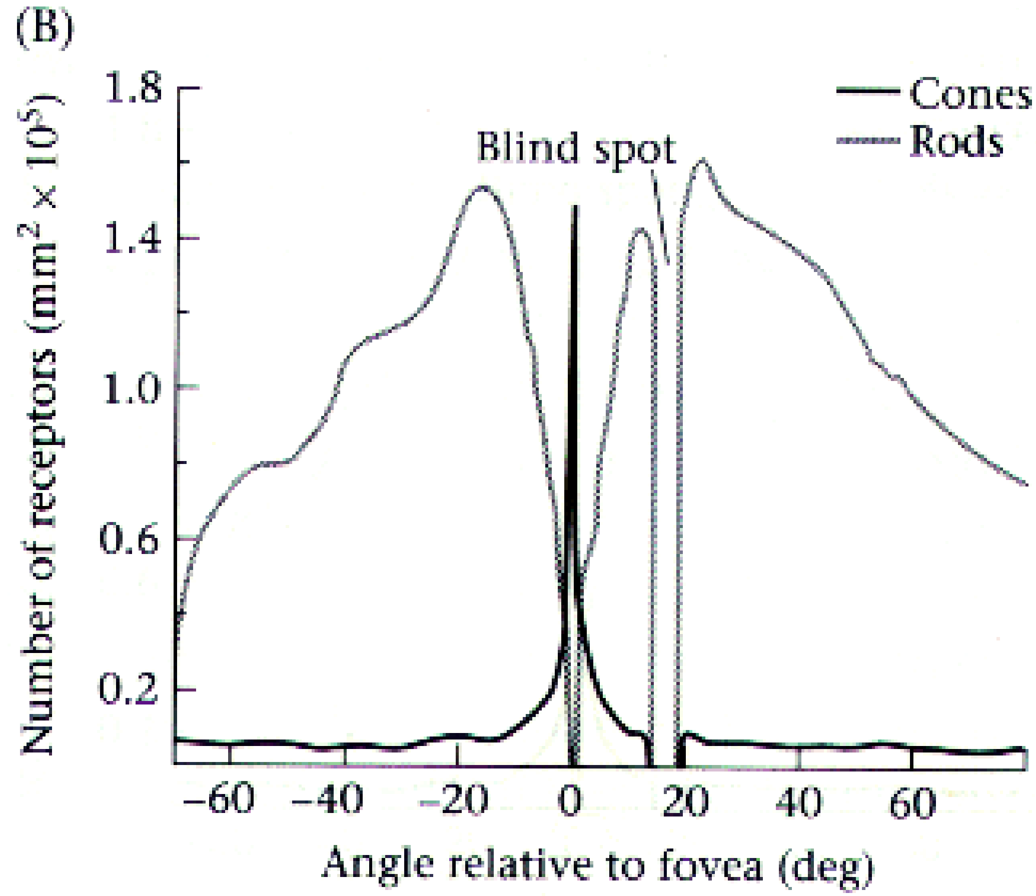
中央窩地區的解剖圖


EYE (from above)
iris
fovea
cornea
lens
ciliary muscle
sclera
retina
blind spot
optic nerve
RETINA (at fovea)
nerve tissue
choroid
ROD
CONE
synoptic body
nucleus
inner segment
outer segment
photopigment

Distribution of rods and cones in retina

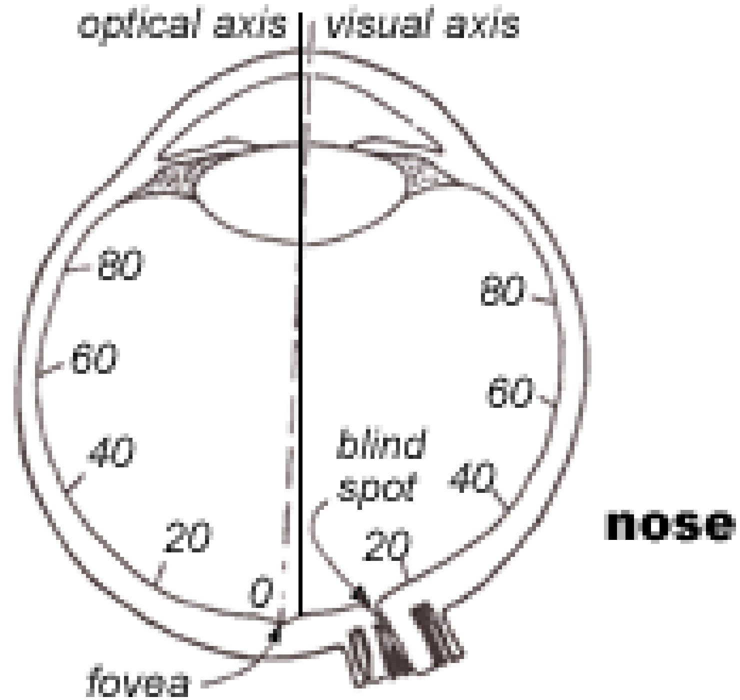
- 錐細胞(cone): 6~7 million
- 桿細胞(rod):75~150 million

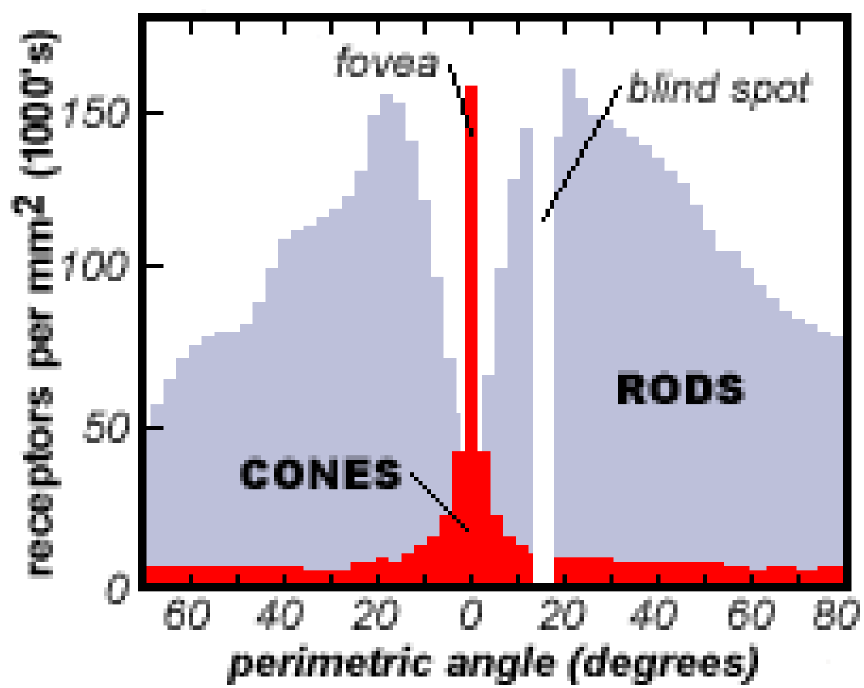




- Cone Vision
    - Photopic Vision
    - Bright-Light Vision

- Rod Vision
    - Scotopic Vision
    - Dim-Light Vision





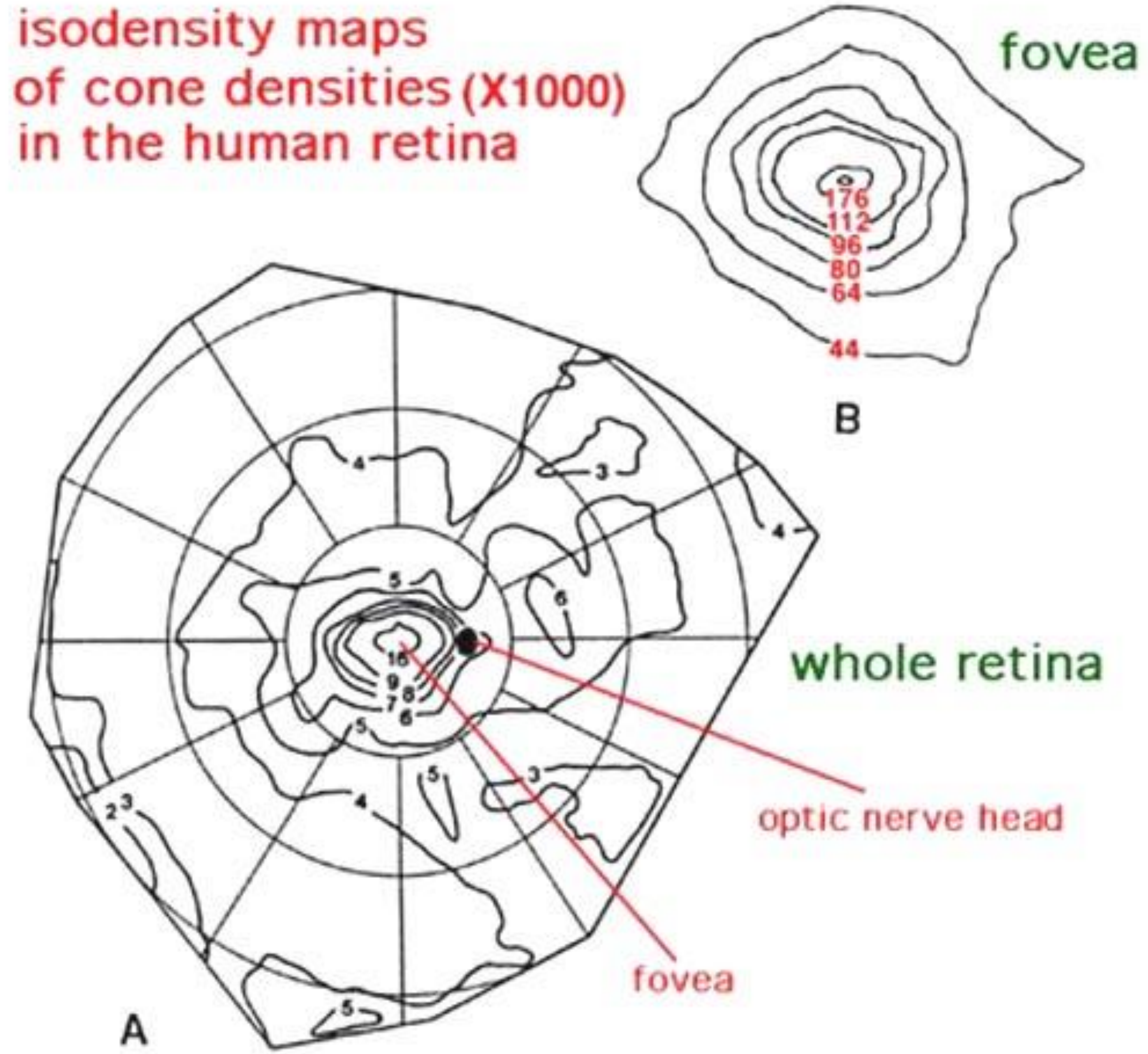
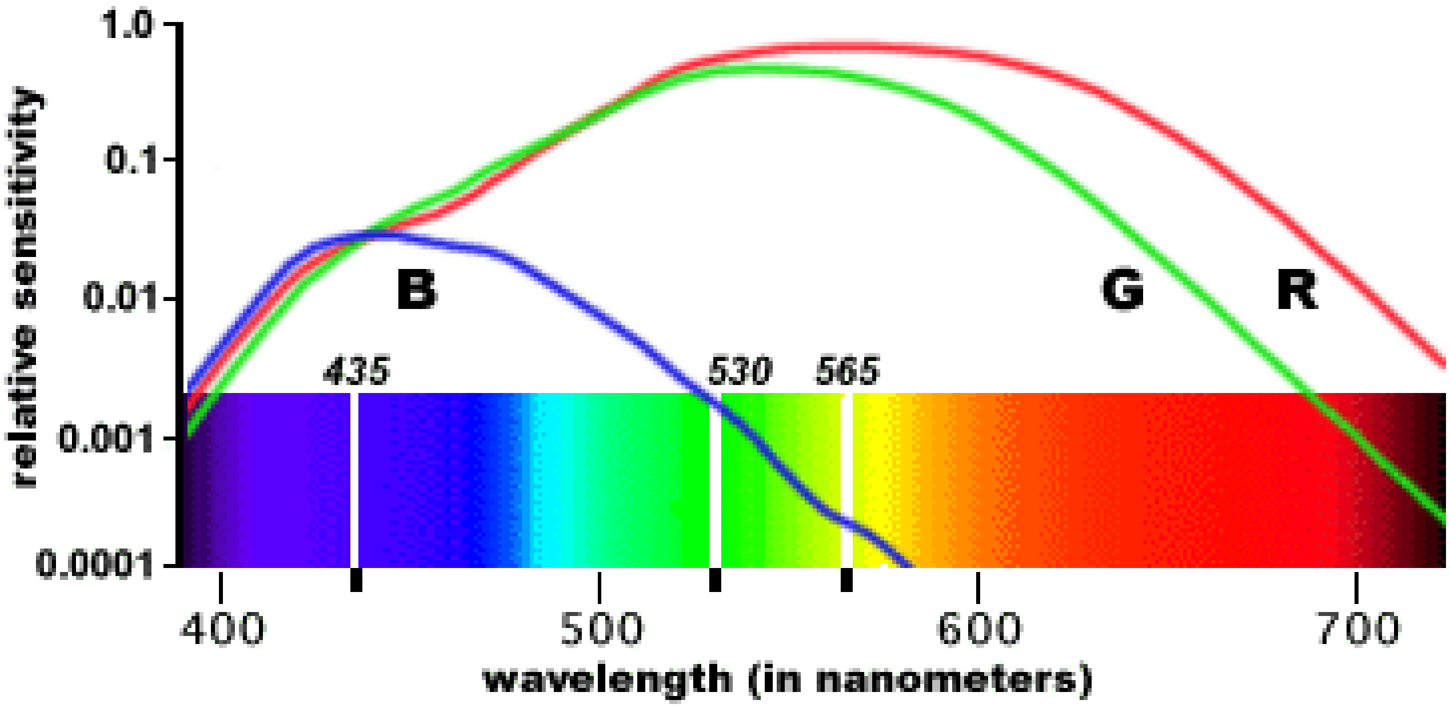


Fig. 21. Cone densities in human retina as revealed in whole mount. The foveal area is enlaged in B. (from Curcio et al., 1987).

錐細胞分為對 長波、中波與短波敏感的三種類型 (L-cone, M-cone, S-cone)，可感知色彩



More in Chapter 6

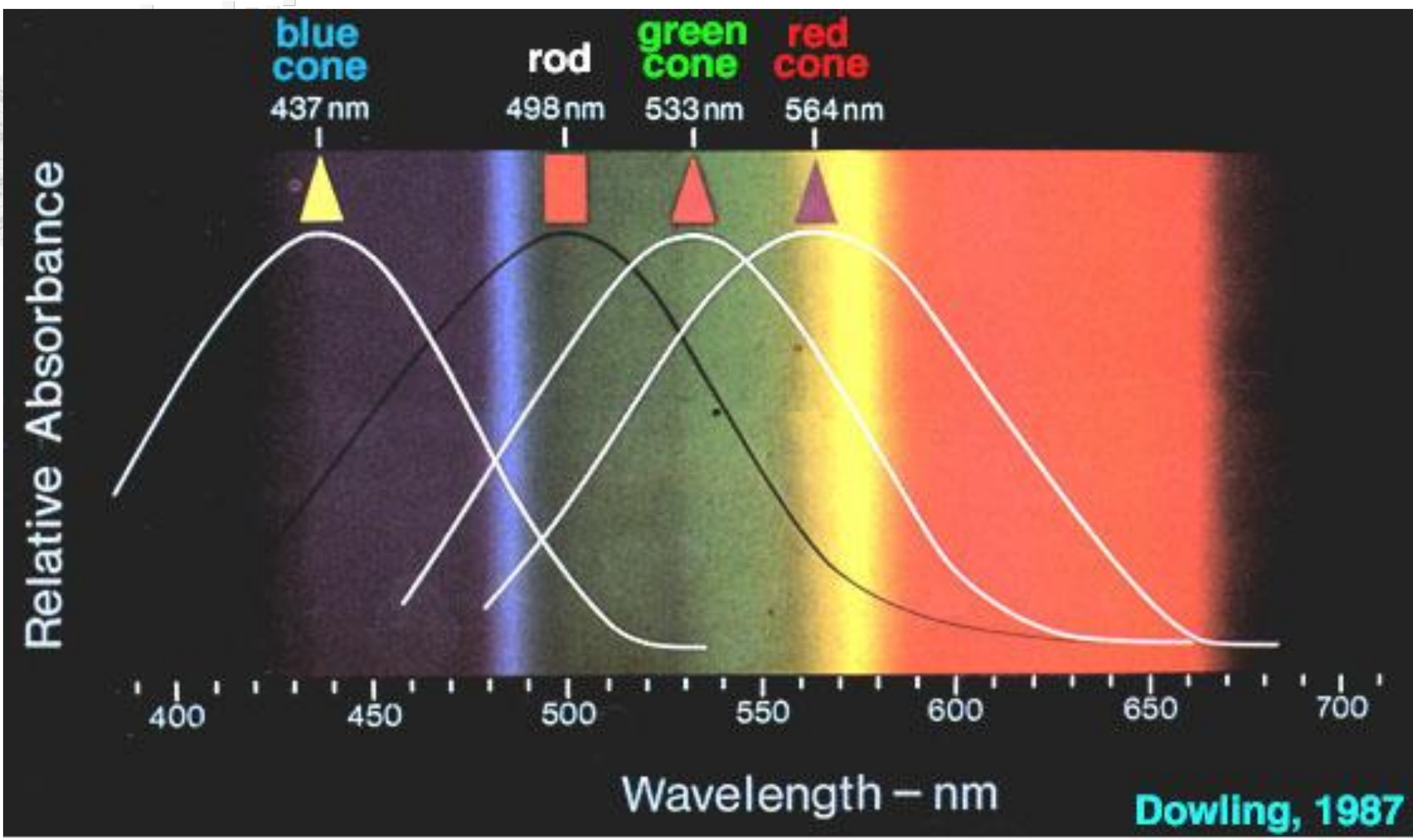


Fig. 14.The peak spectral sensitivities of the the 3 cone types and the the rods in the primate retina (Brown and Wald, 1963). From Dowling's book (1987).

人類的視覺神經路徑

網膜→視交叉→側膝核→視皮質

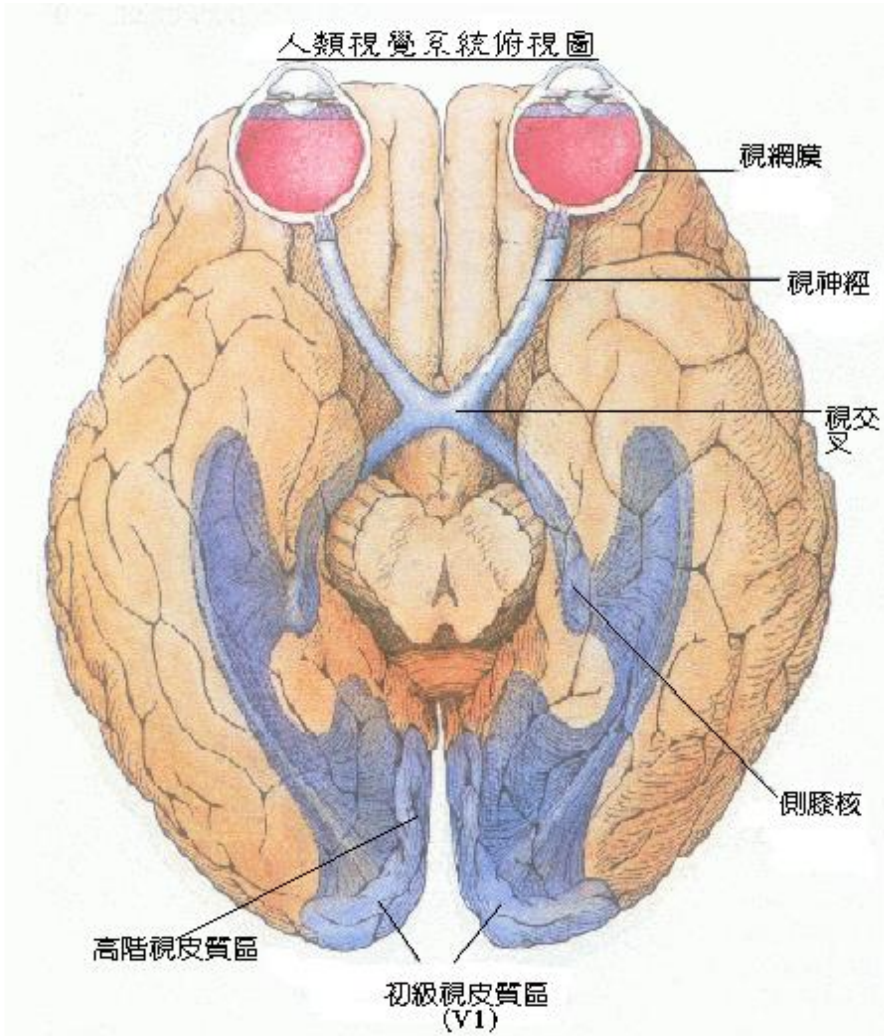
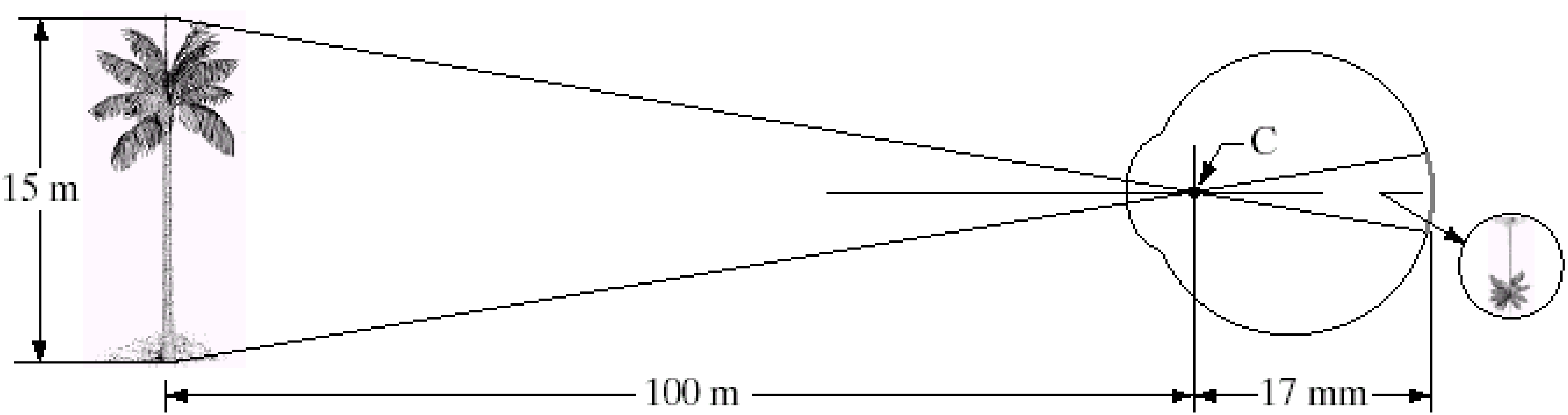


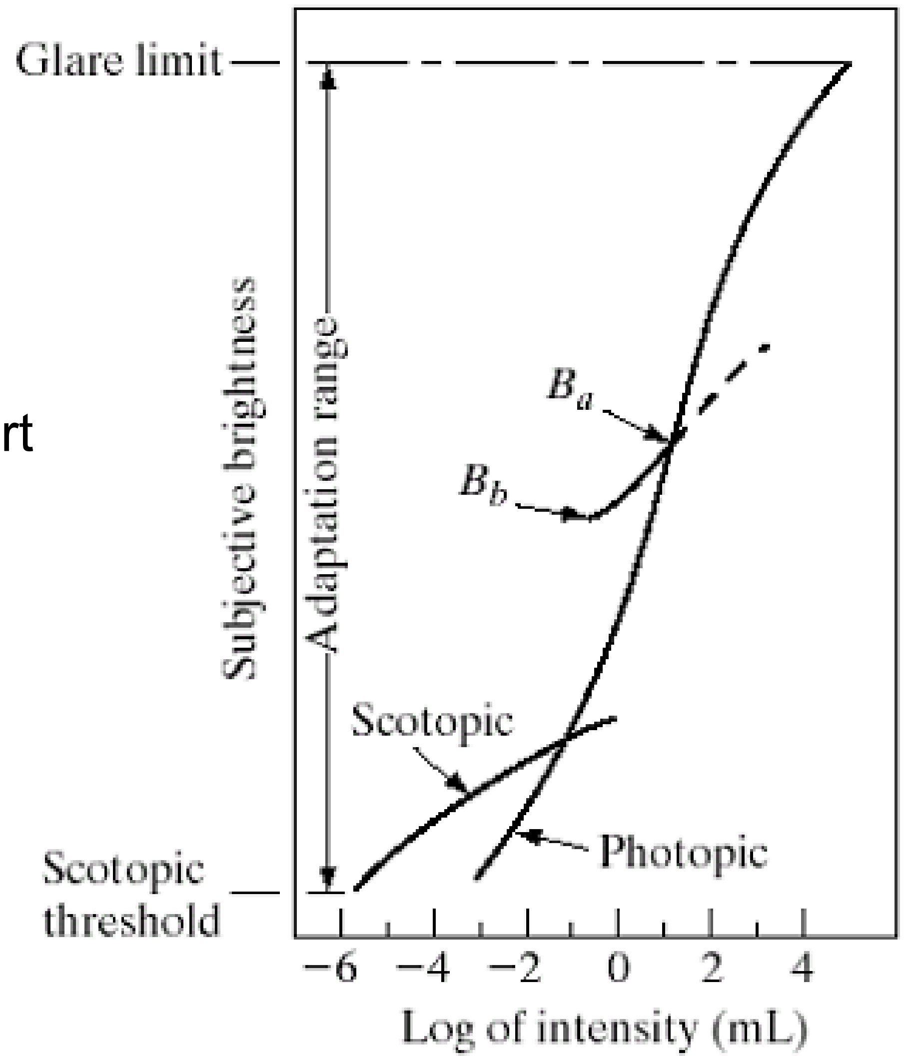
Image Formation in the Eye

- Lens shape controlled by the tension of ciliary fibers
- Focal length: 17mm ~ 14mm



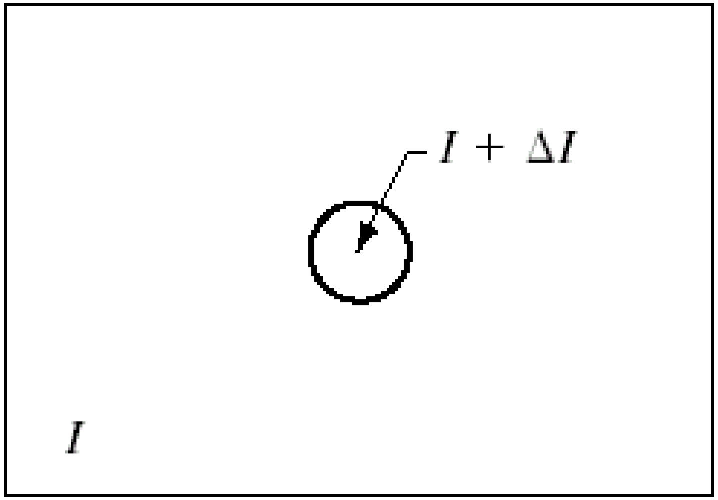
Brightness Adaptation and Discrimination

- Range of light intensity that human can adapt
  ➔ 10^{10}
- Transition from scotopic to photopic vision is gradual
  ➔ from 0.001 to 0.1 millilambert (-3 to –1 in log)
- Vision System cannot operate over such a range simultaneously
- B_a : Brightness adaptation level
- B_b : at and below this level all stimuli are perceived as indistinguishable black



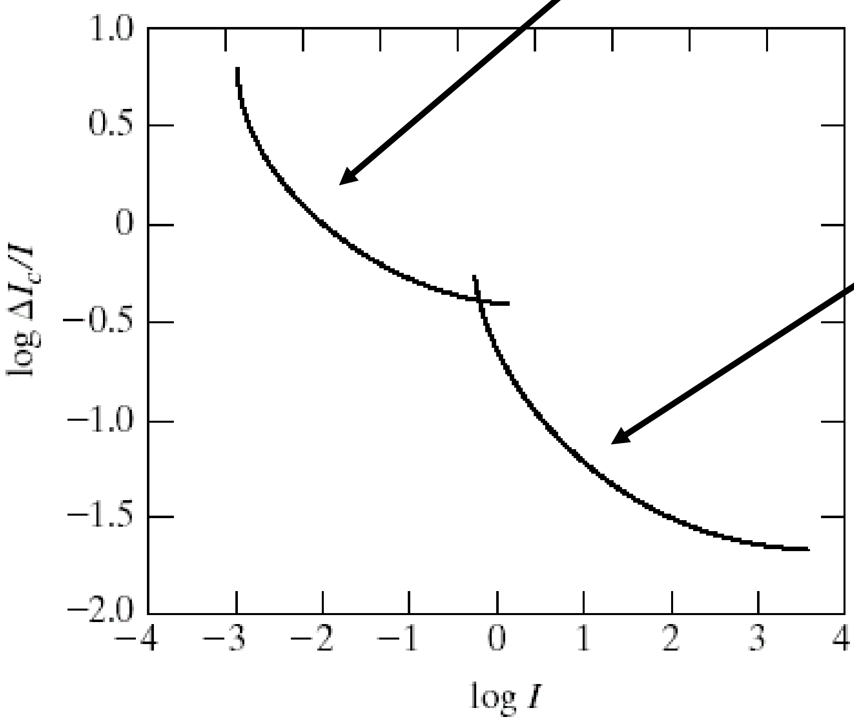
Weber Ratio, $\Delta I_c / I$

- Uniformly illuminated background, occupying entire field of view
- Add a short duration of flash ΔI
- ΔI_c: the increment of illumination discriminable 50% of the time



Weber Ratio as a Function of Intensity

Rod vision: weber ratio large
→ brightness discrimination poor



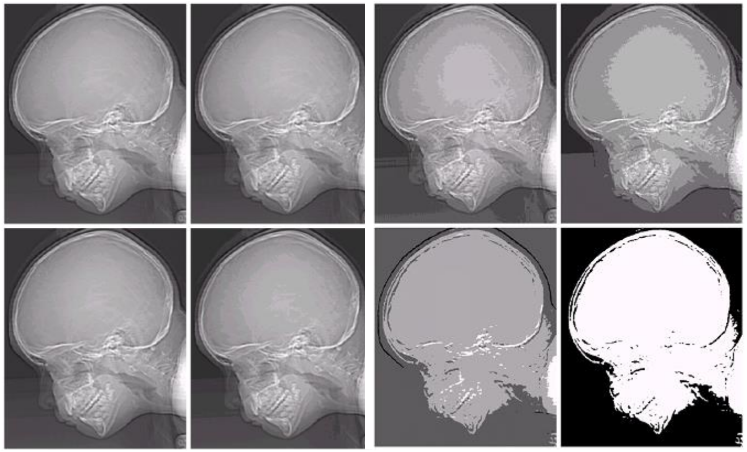
Cone vision: weber ratio smaller
→ brightness discrimination better

Intensity Discrimination of Human Vision

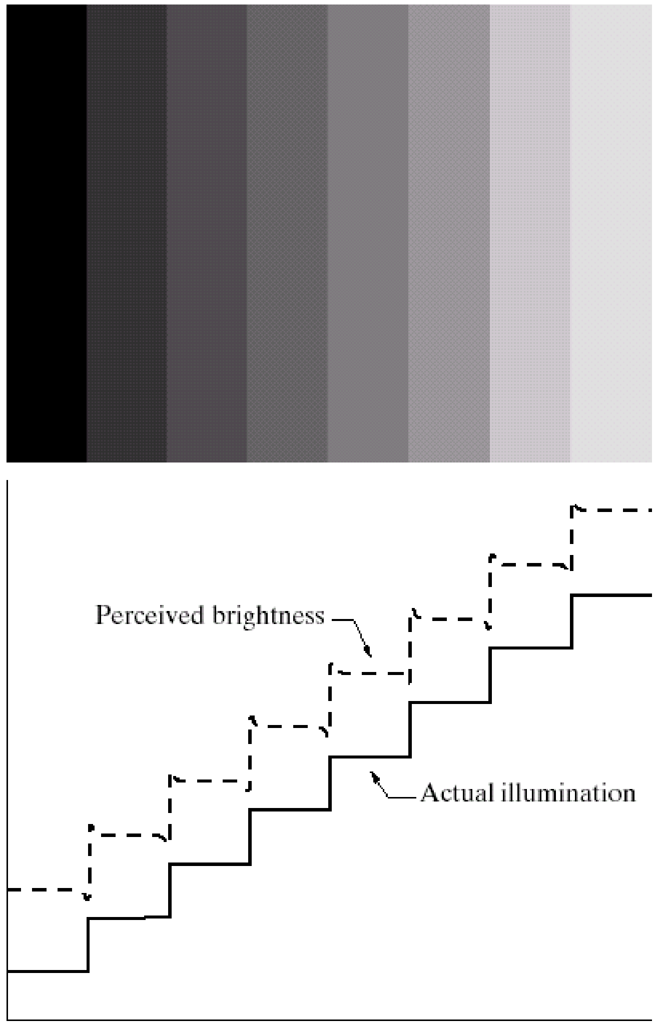
- Instead of flashing, ΔI varies incrementally from never being perceived to always being perceived
- The total number of distinct intensity levels the observer can discriminate is rather small: ~20

➢ Because the eye can roam about the image
→ capable of a much broader range of overall intensity discrimination

see Sec. 2.4.3



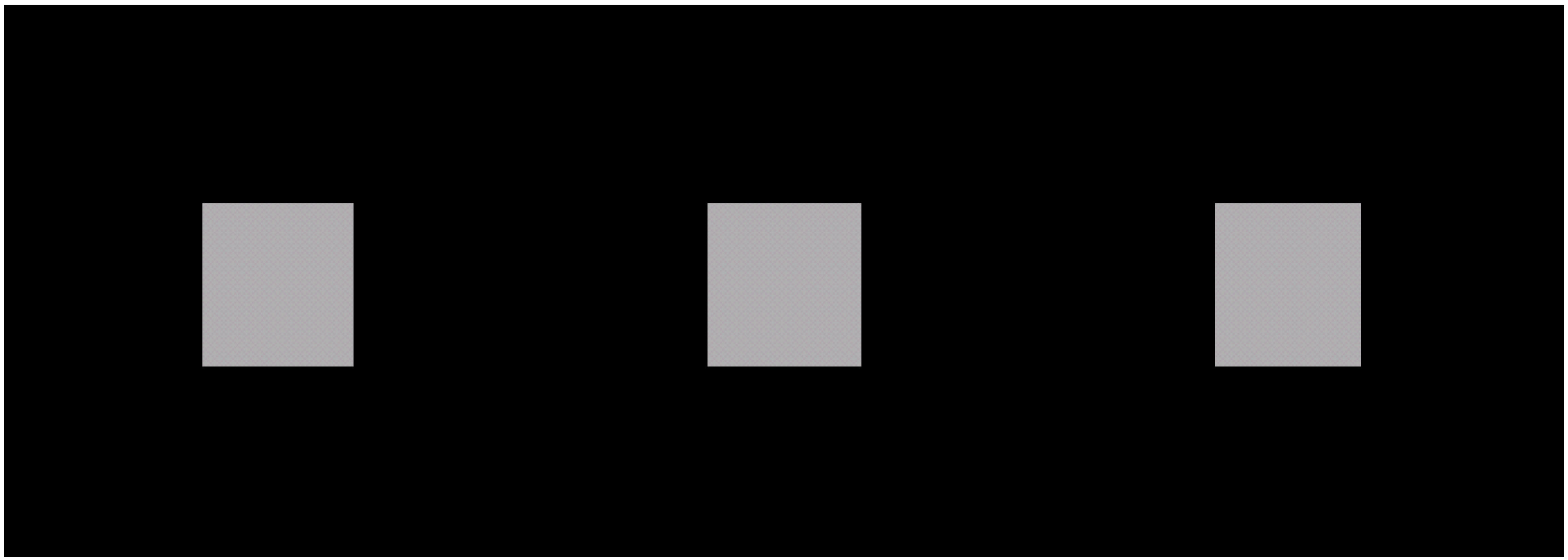
Human Vision Phenomenon: Mach Band



a
b

FIGURE 2.7
(a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

Human Vision Phenomenon: Simultaneous Contrast



a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Human Vision Phenomenon: Optical Illusions

http://dragon.uml.edu/psych/illusion.html

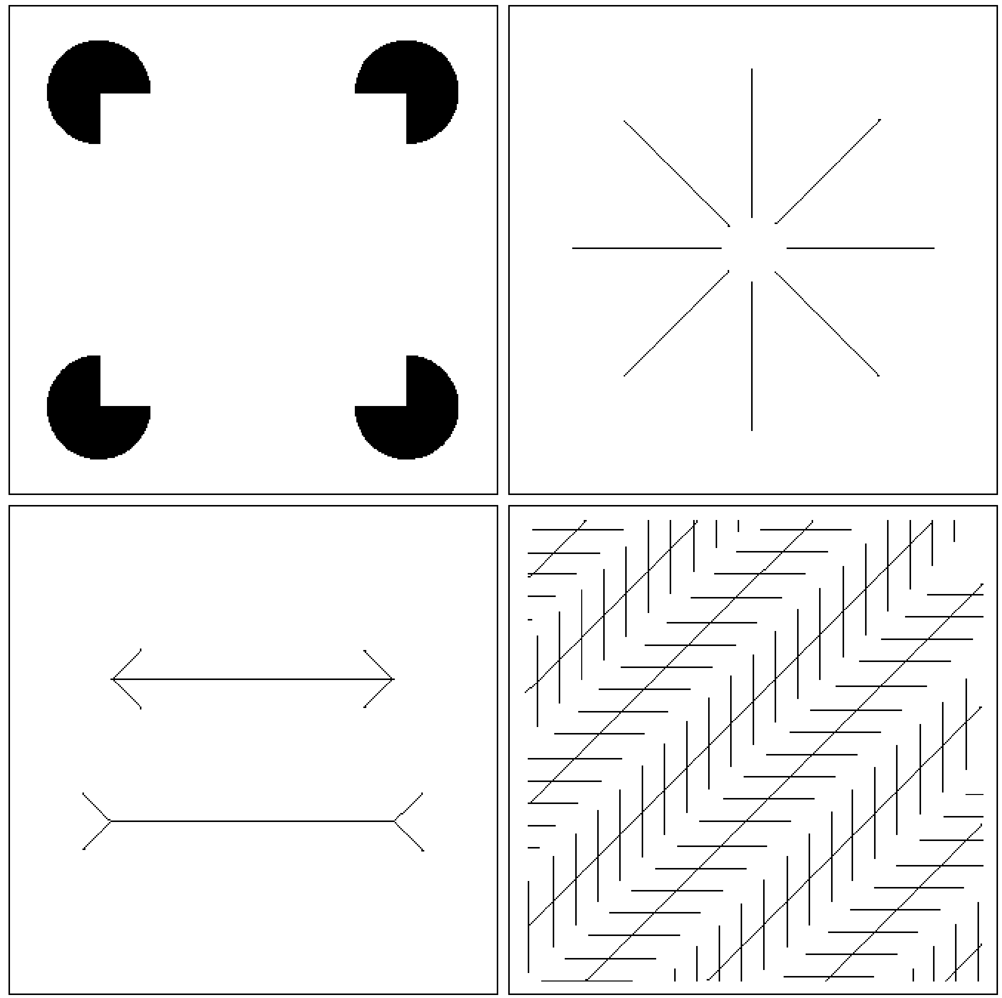
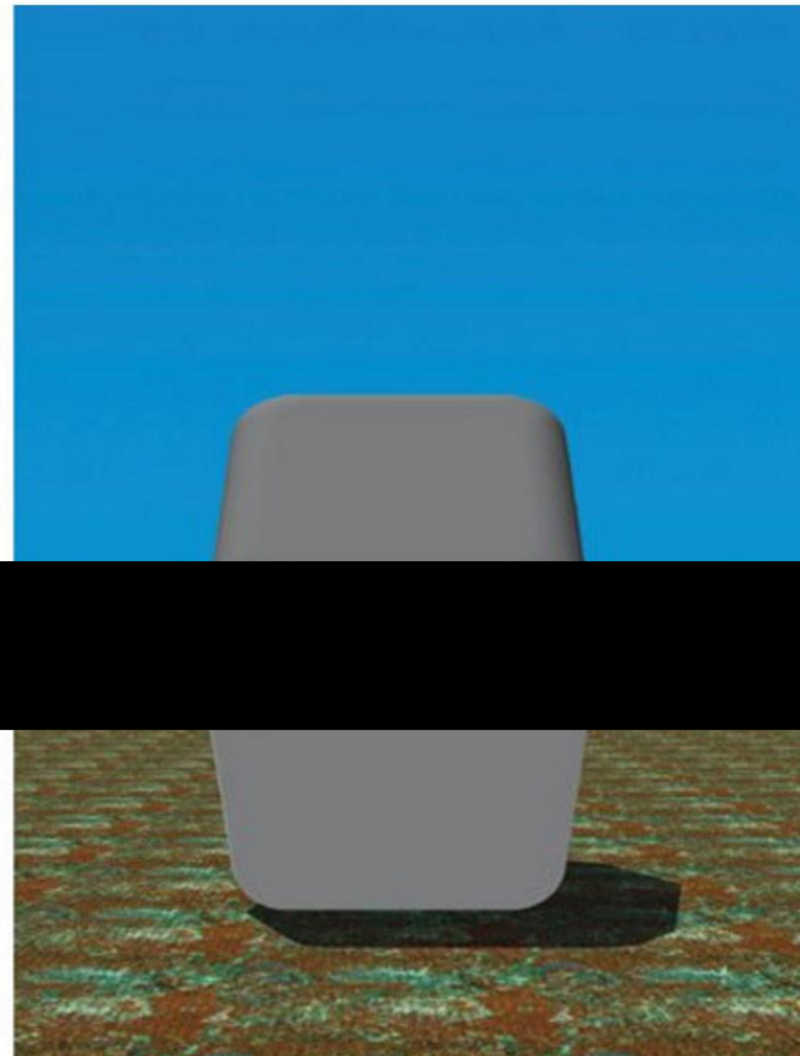


FIGURE 2.9 Some well-known optical illusions.

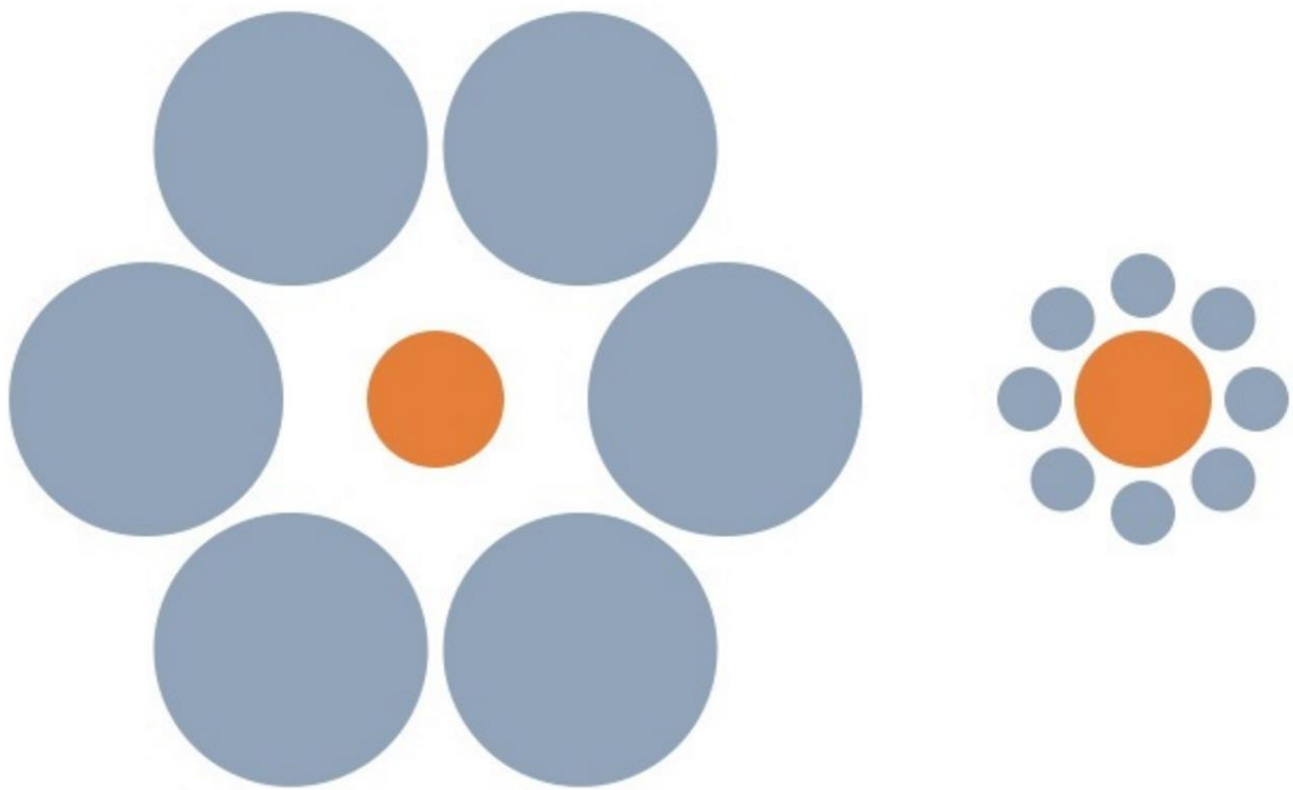
Teepr：32個就算你知道是錯覺但還是一定會被騙的終極錯覺



[redacted]

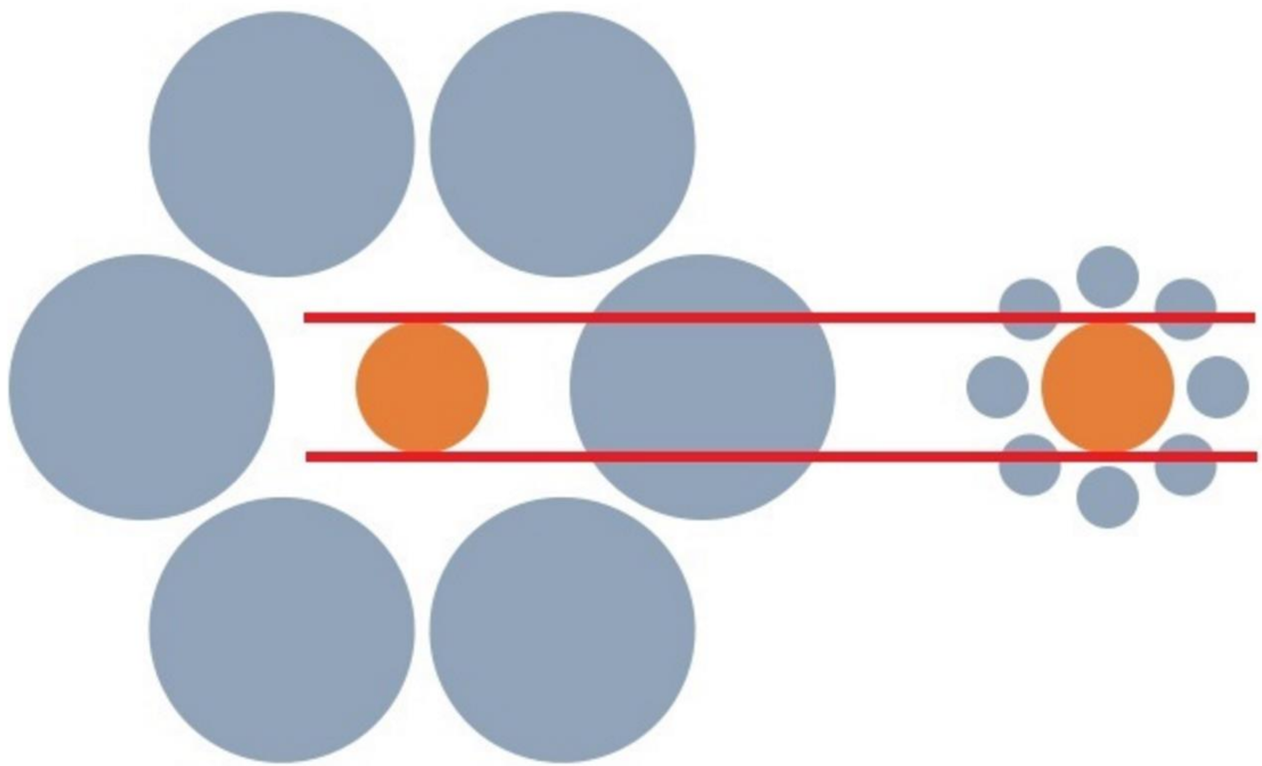
Teepr : 32個就算你知道是錯覺但還是一定會被騙的終極錯覺

8. 哪一個橘圈大呢？



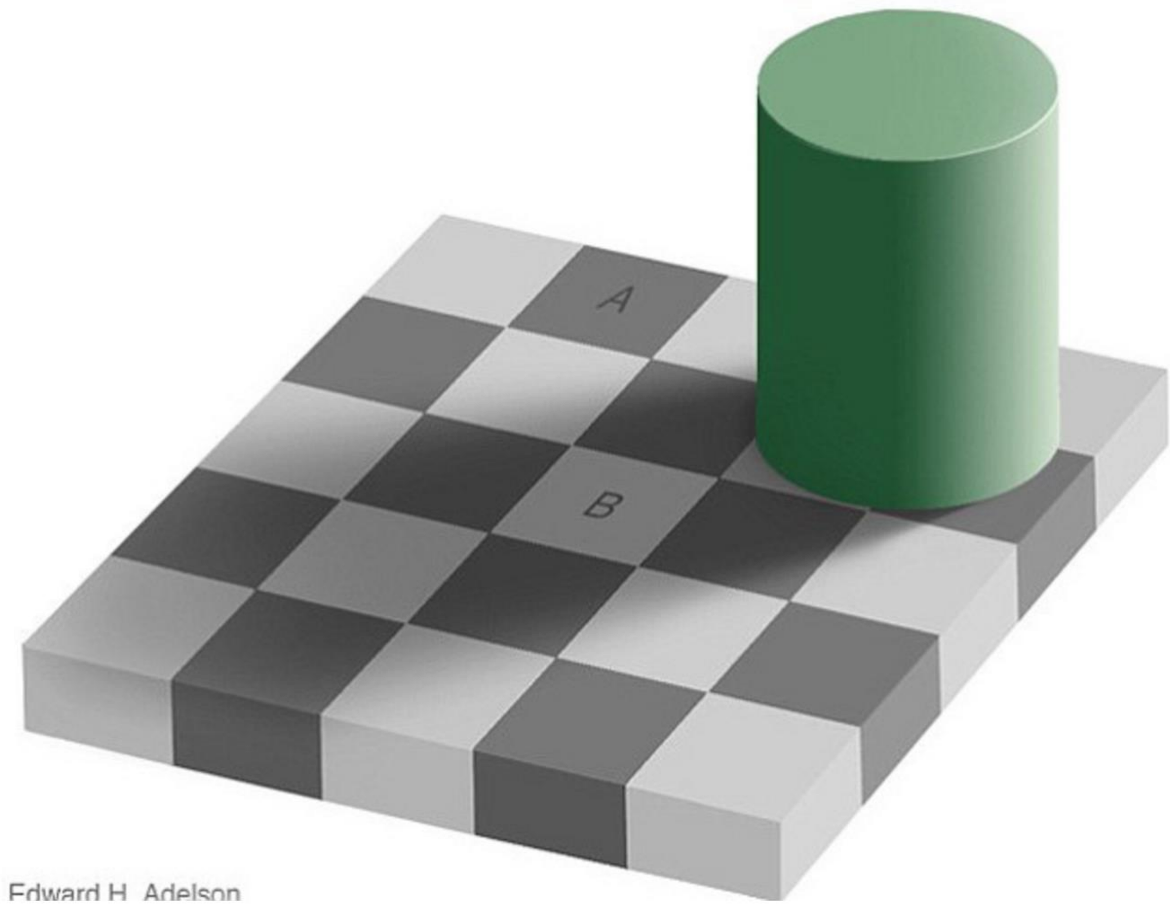
Teepr：32個就算你知道是錯覺但還是一定會被騙的終極錯覺

我想你已經猜到答案了。(但是你的眼睛還是告訴你是左邊大...)



Teepr : 32個就算你知道是錯覺但還是一定會被騙的終極錯覺

10. 已經告訴你了，圖中A和B兩格的顏色是一樣的。



2.2 Light and EM Spectrum

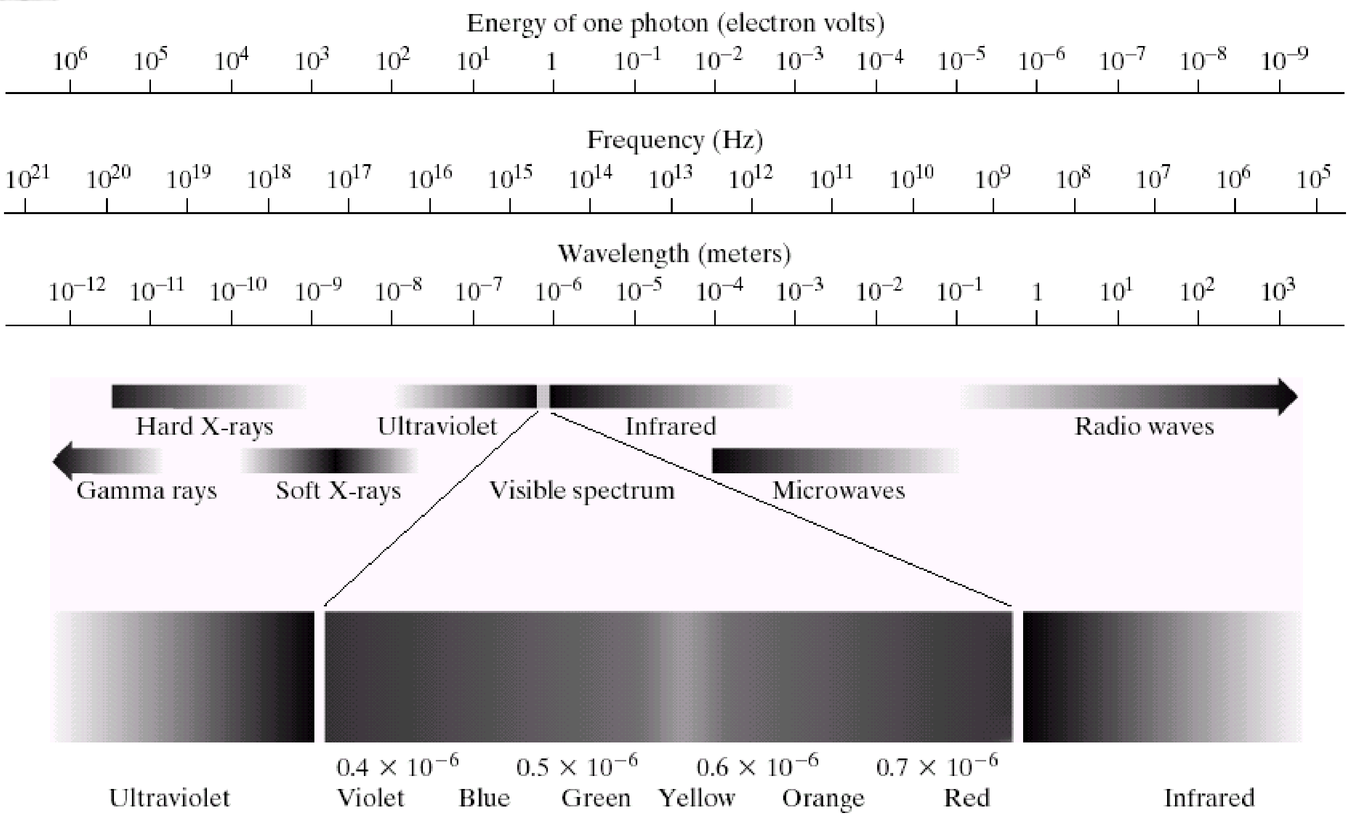
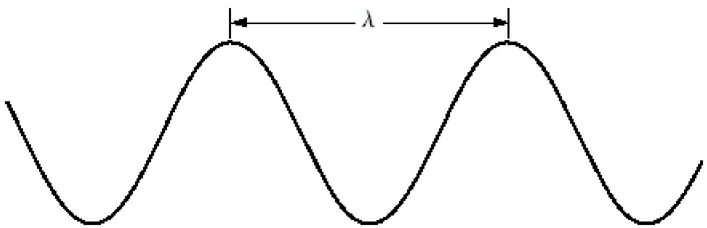


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Visible band: 0.43 μm (violet) ~ 0.79 μm (red)



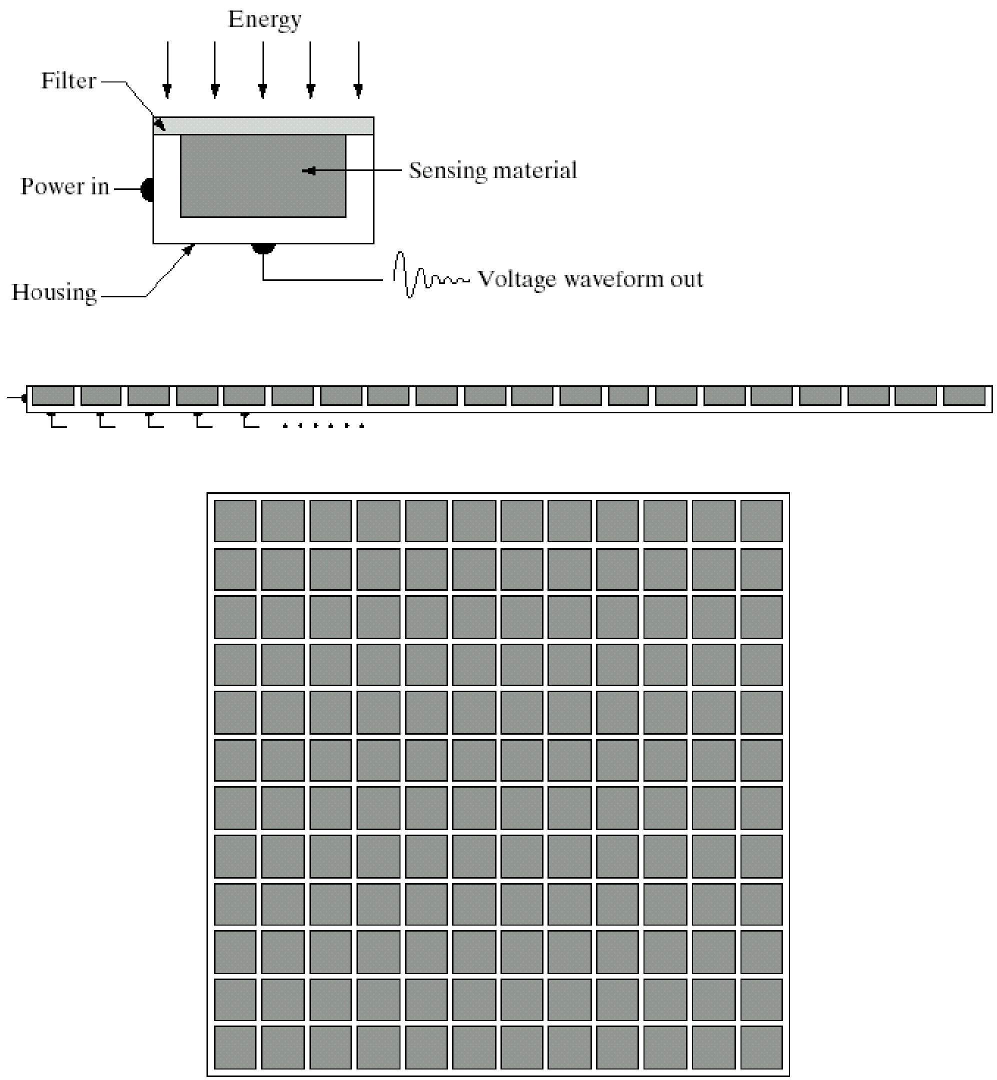
- Achromatic or monochromatic light
  → gray level, intensity

- Radiance: total energy from light source → watts
- Luminance: energy that an observer perceives → lumens
- Brightness: subjective descriptor of light perception
  → hard to measure

2.3 Image Sensing and Acquisition

- Image Acquisition Using a Single Sensor
- Image Acquisition Using Sensor Strips
- Image Acquisition Using Sensor Arrays
- A simple Image Formation Model

Sensors



a
b
c

FIGURE 2.12
(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.

Image Acquisition Using a Single Sensor

Microdensitometers

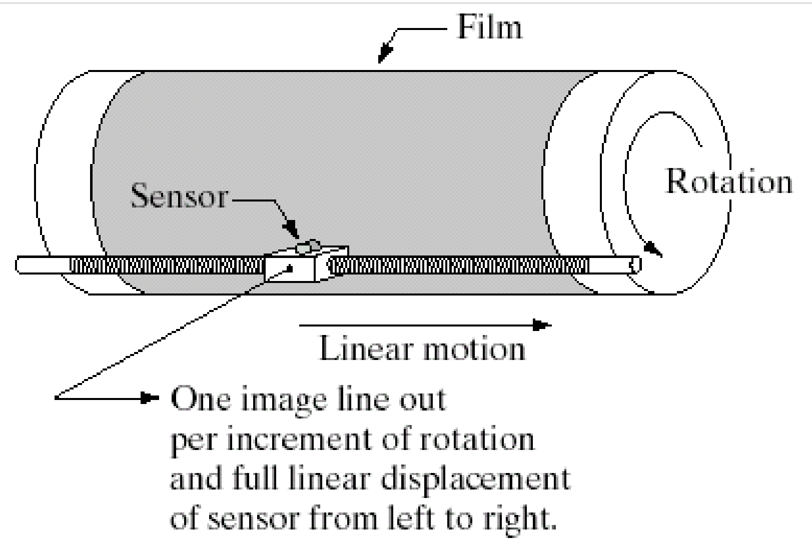


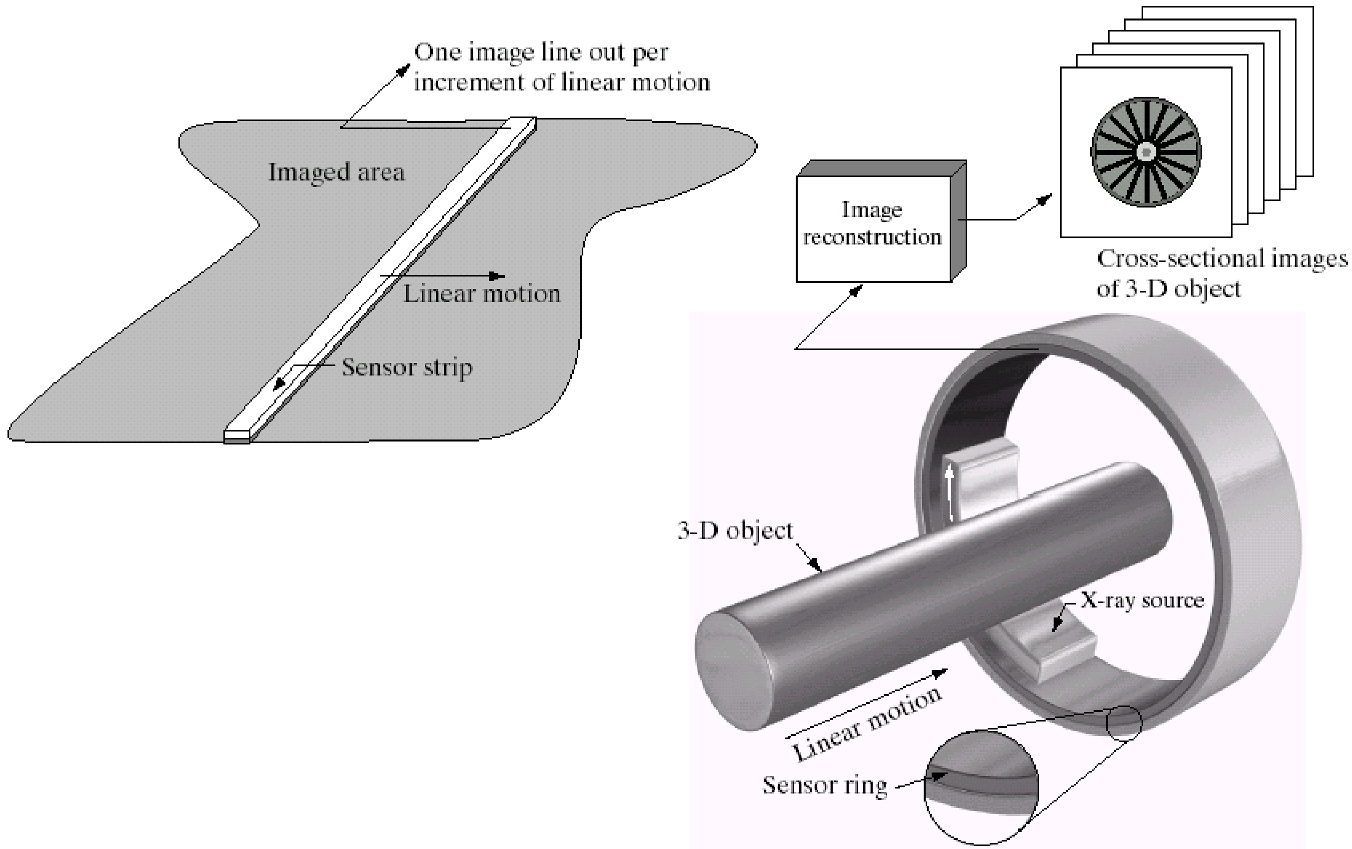
FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Laser Rangefinders

Cyra Leica

Trimble Callidus

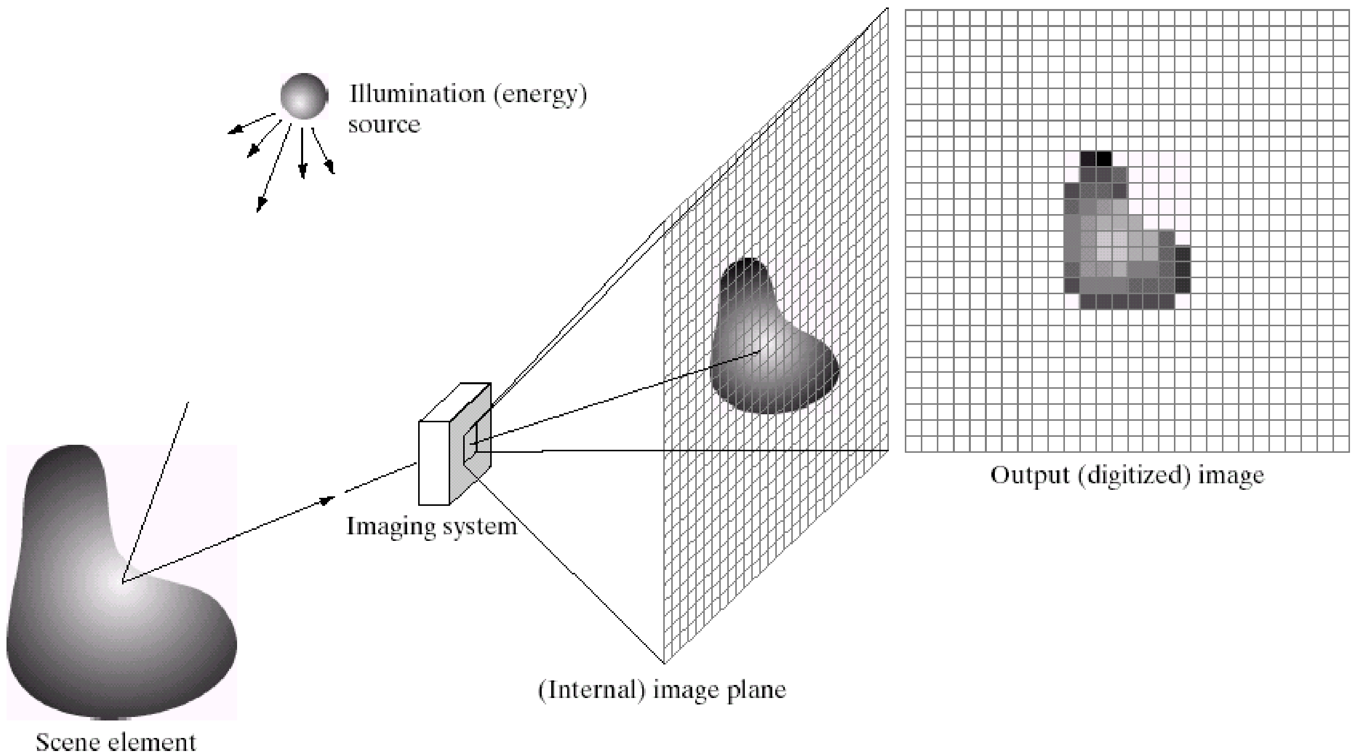
Image Acquisition Using Sensor Strips



a b

FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

Image Acquisition Using Sensor Arrays



a b c d e

FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

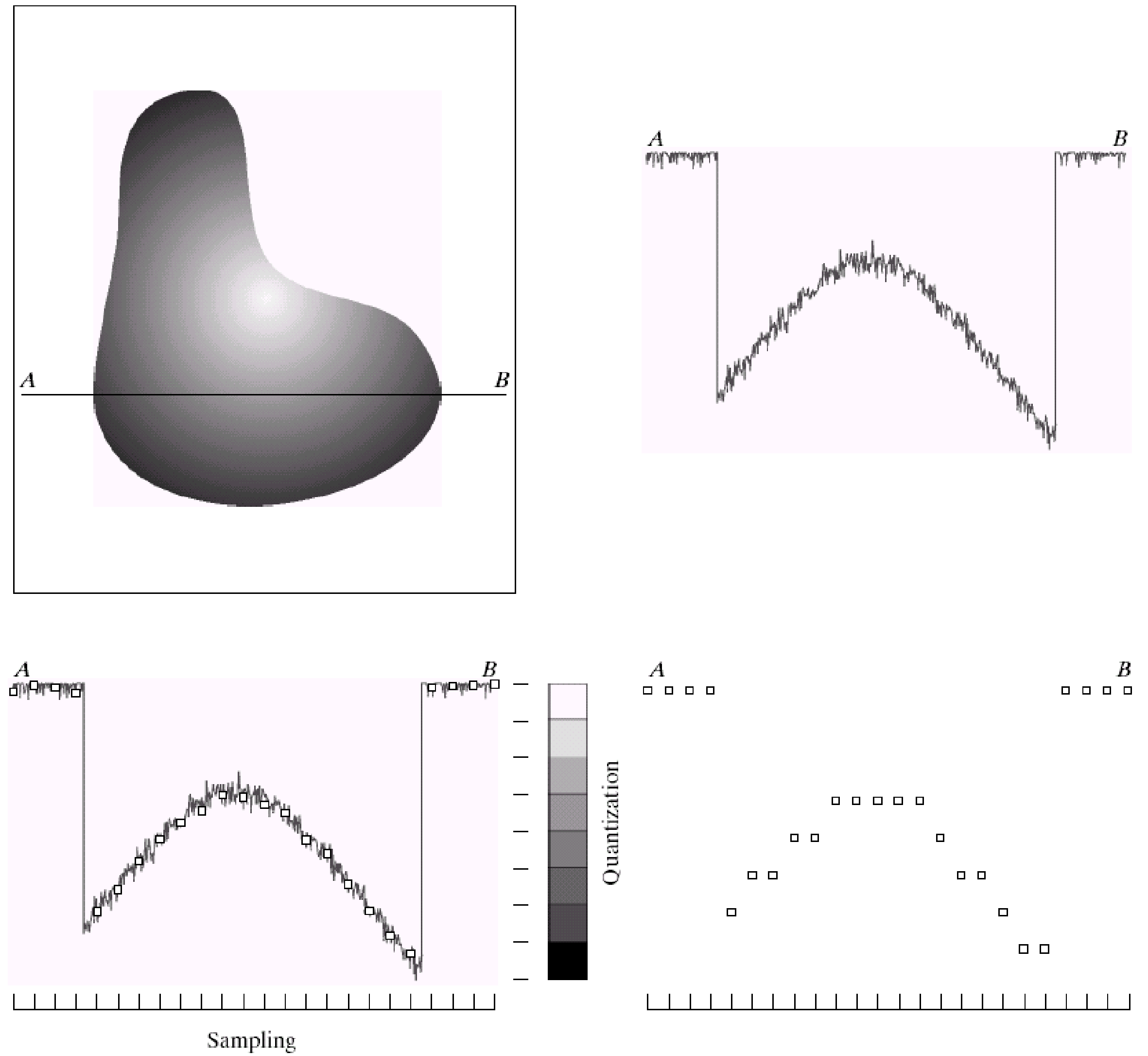
A simple Image Formation Model

$$f(x, y) = i(x, y)\, r(x, y)$$

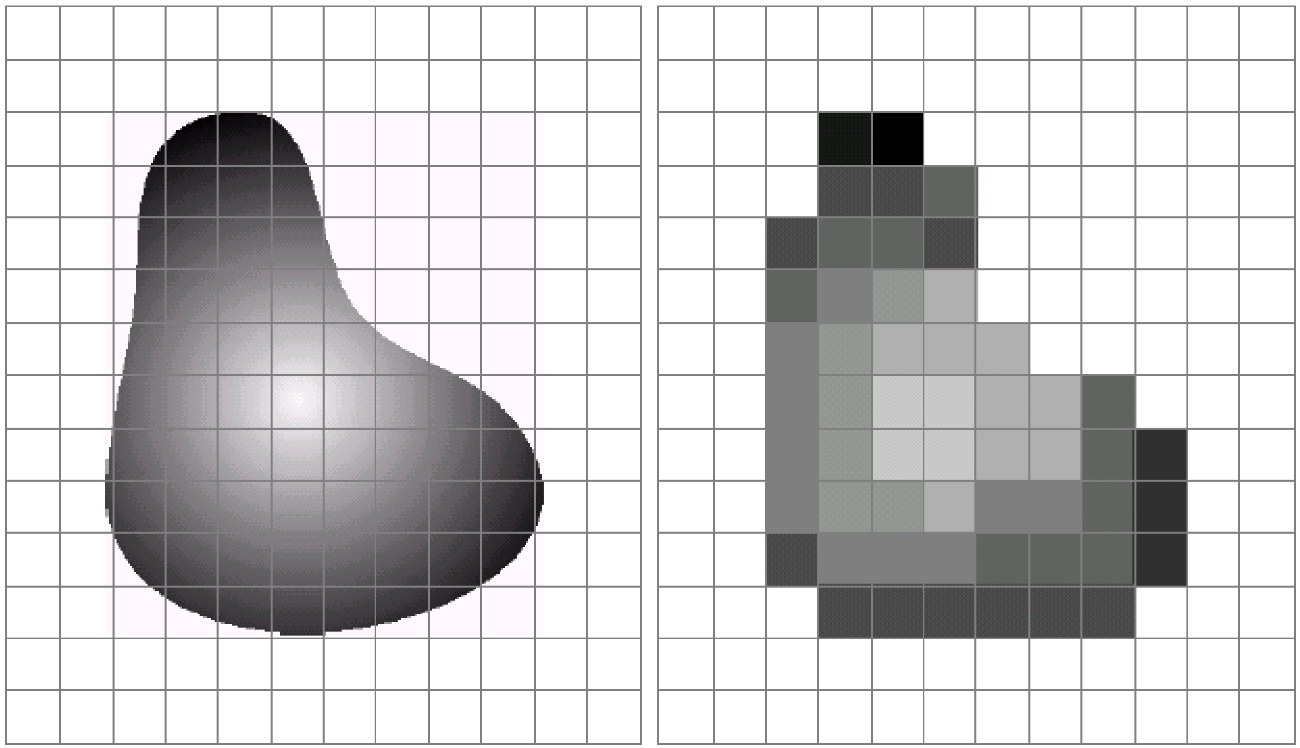
$i(x,y)$: the amount of source illumination incident on the scene

$r(x,y)$: the reflectivity function (or transmissivity function)

2.4 Image Sampling and Quantization



Sampling and Quantization with a Sensing array



a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Representing Digital Images

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}.$$

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}.$$

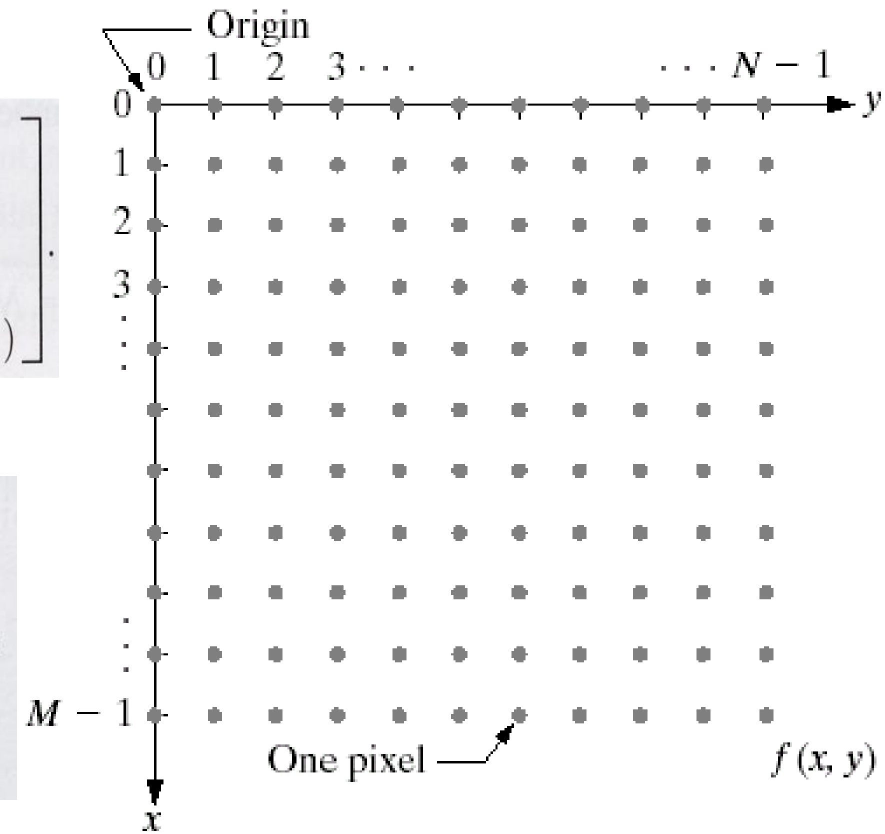


Image Size

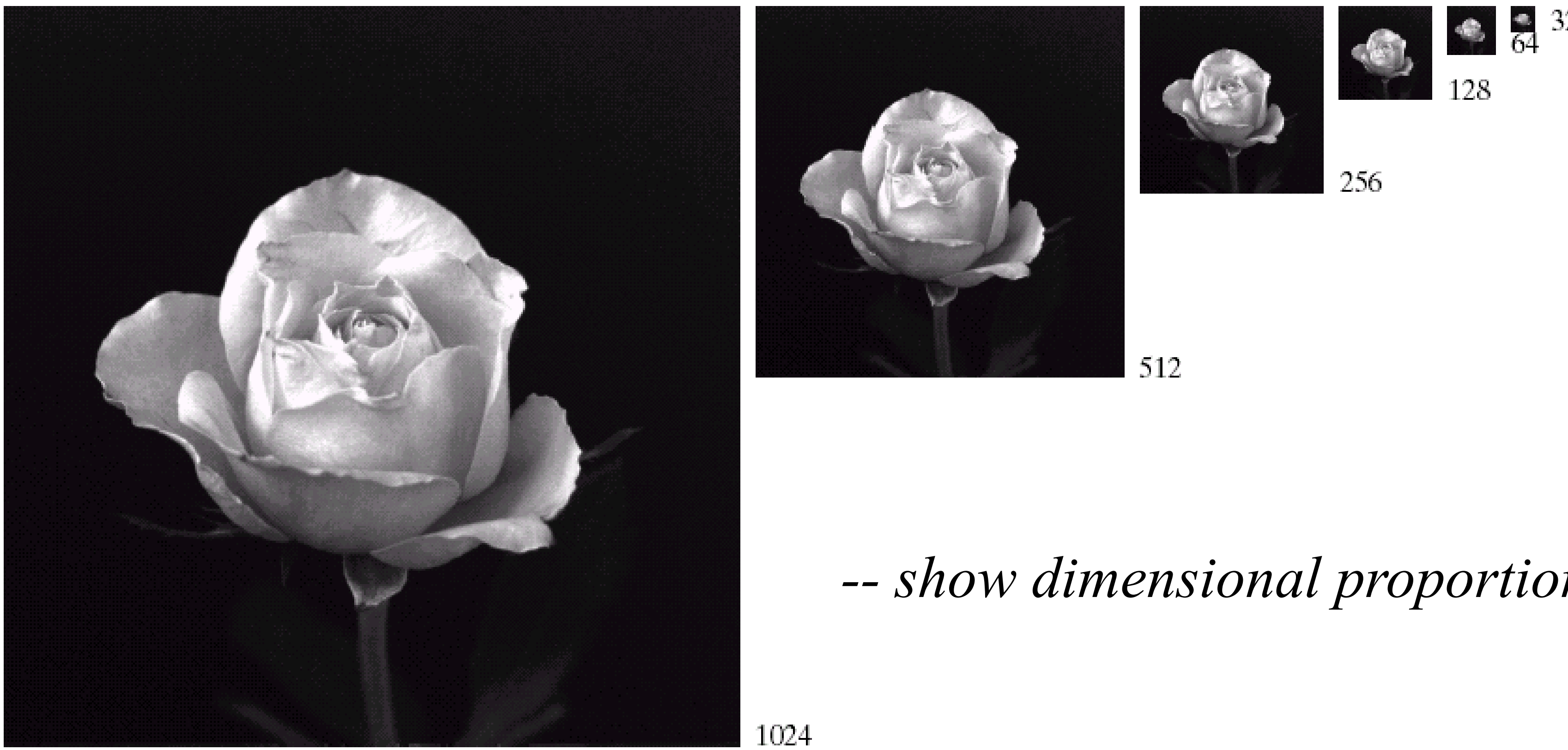
Dynamic Range: the range of values spanned by gray scale

$b = N^2 k$ $\qquad$ $0, \ldots, L\text{-}1, \quad L = 2^k$

TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Spatial Resolution



-- show dimensional proportion

FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

-- zoom-in to show the effects of subsampling

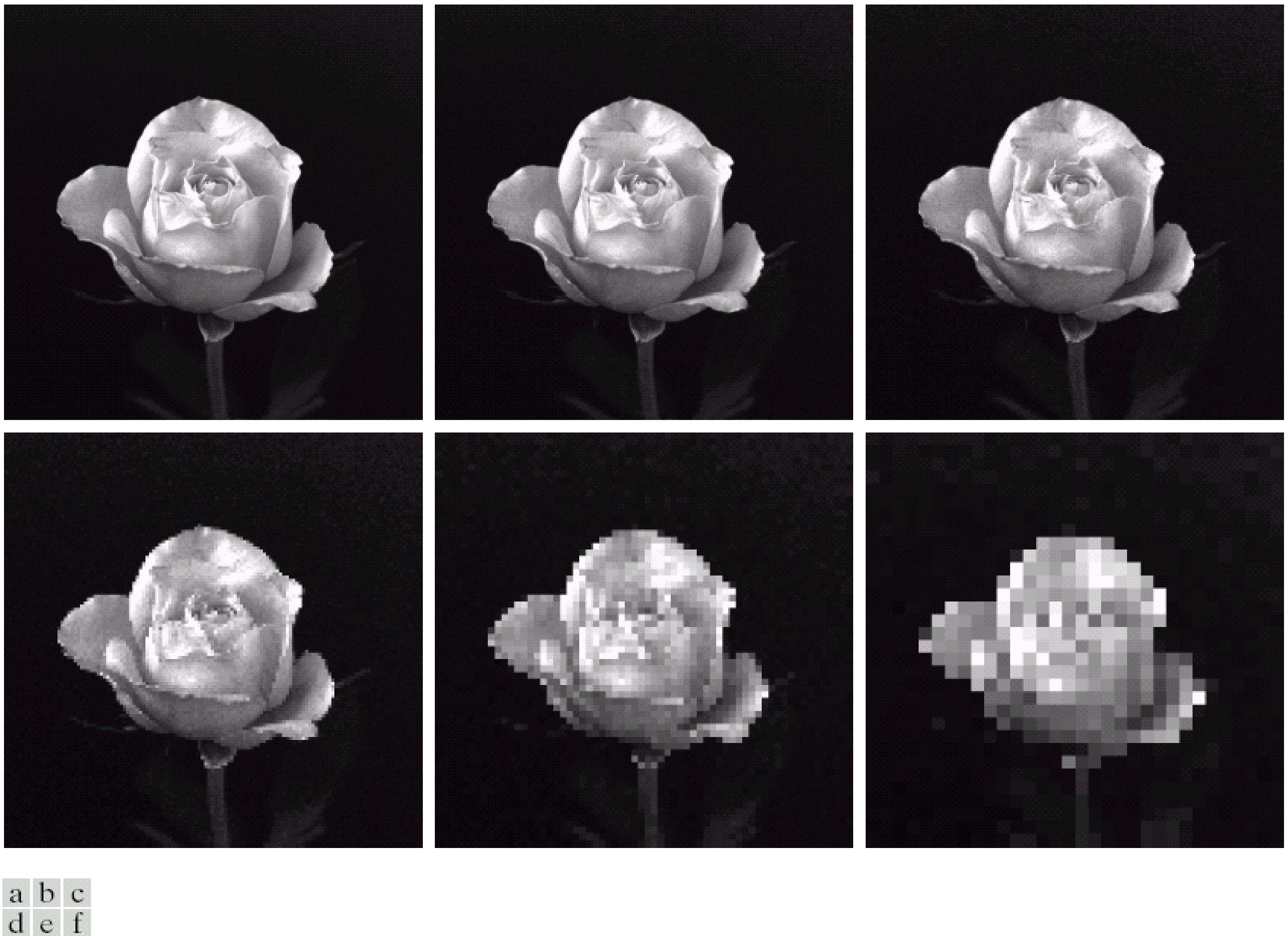
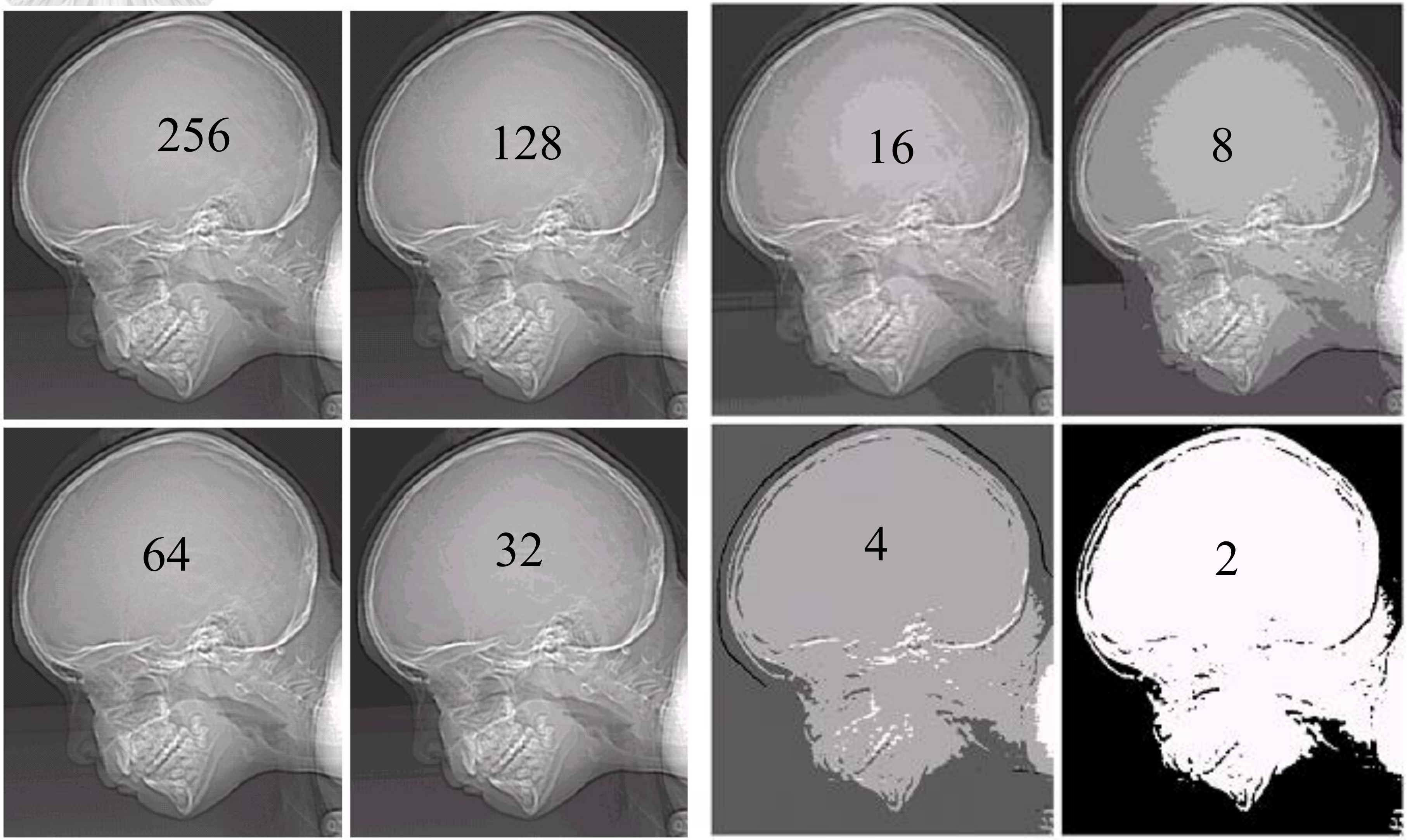


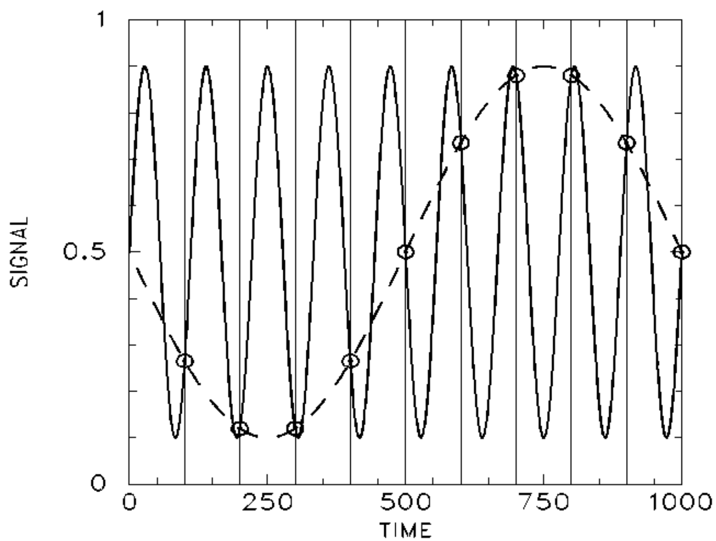
FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

L = 256, 128, 64, 32, 16, 8, 4, 2



Y

Aliasing and Moire Pattern



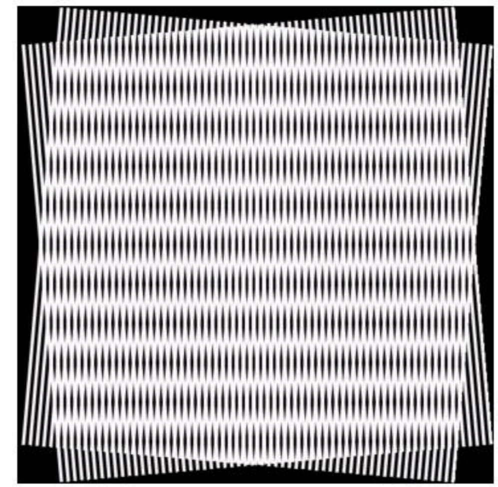
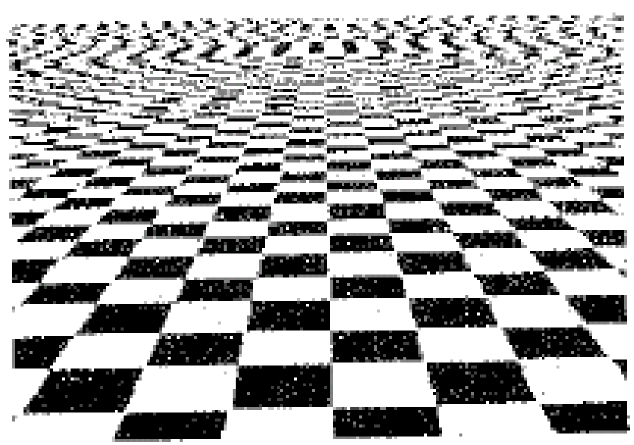
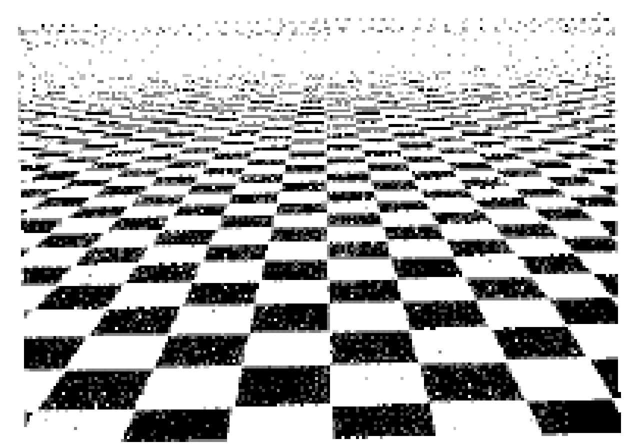
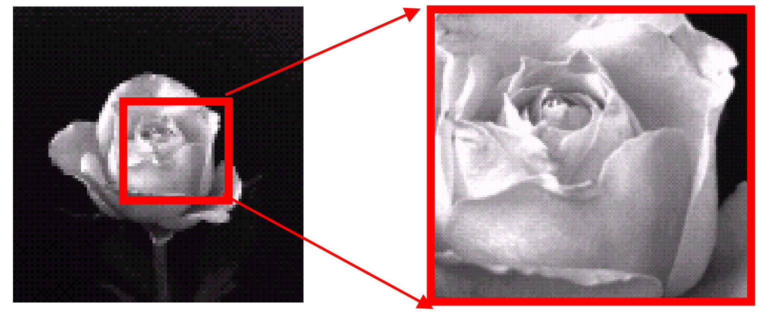


FIGURE 2.24 Illustration of the Moiré pattern effect.

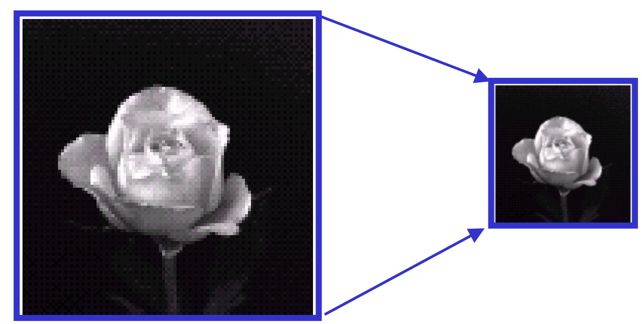




Zooming and Shrinking of Digital Images



- Zooming: oversampling
  - Need *interpolation*
  - "superresolution"



- Shrinking: undersampling
  - Better to apply LPF before subsampling to avoid aliasing
  - Need *interpolation* for non-integer factor

Image Interpolation

- Nearest Neighbor Interpolation
- Bilinear Interpolation
- Bicubic Interpolation
- etc

Applications

- Image Scaling/Resize
- Image Rotation
- Image Warping
- Image Morphing
- ……

Nearest-neighbor Interpolation

Bilinear Interpolation

FIGURE 2.25 Top row: images zoomed from 128 × 128, 64 × 64, and 32 × 32 pixels to 1024 × 1024 pixels, using nearest neighbor gray-level interpolation. Bottom row: same sequence, but using bilinear interpolation.

A geometric image operation generally requires two steps. The first is a spatial mapping of the coordinates of an original image f to define a new image g:

$$g(\mathbf{n}) = f(\mathbf{n}') = f[\mathbf{a}(\mathbf{n})]. \qquad (45)$$

7.1 Nearest-Neighbor Interpolation

Here, the geometrically transformed coordinates are mapped to the nearest integer coordinates of f:

$$g(\mathbf{n}) = f\{\mathrm{INT}[a_1(n_1, n_2) + 0.5], \mathrm{INT}[a_2(n_1, n_2) + 0.5]\}, \quad (46)$$

where $\mathrm{INT}[R]$ denotes the nearest integer that is less than or equal to R. Hence, the coordinates are *rounded* prior to assigning them to g. This certainly solves the problem of finding integer coordinates of the input image, but it is quite simplistic, and, in

7.2 Bilinear Interpolation

Bilinear interpolation produces a smoother interpolation than does the nearest-neighbor approach. Given four neighboring image coordinates $f(n_{10}, n_{20})$, $f(n_{11}, n_{21})$, $f(n_{12}, n_{22})$, and $f(n_{13}, n_{23})$ — these can be the four nearest neighbors of $f[\mathbf{a}(\mathbf{n})]$ — then the geometrically transformed image $g(n_1, n_2)$ is computed as

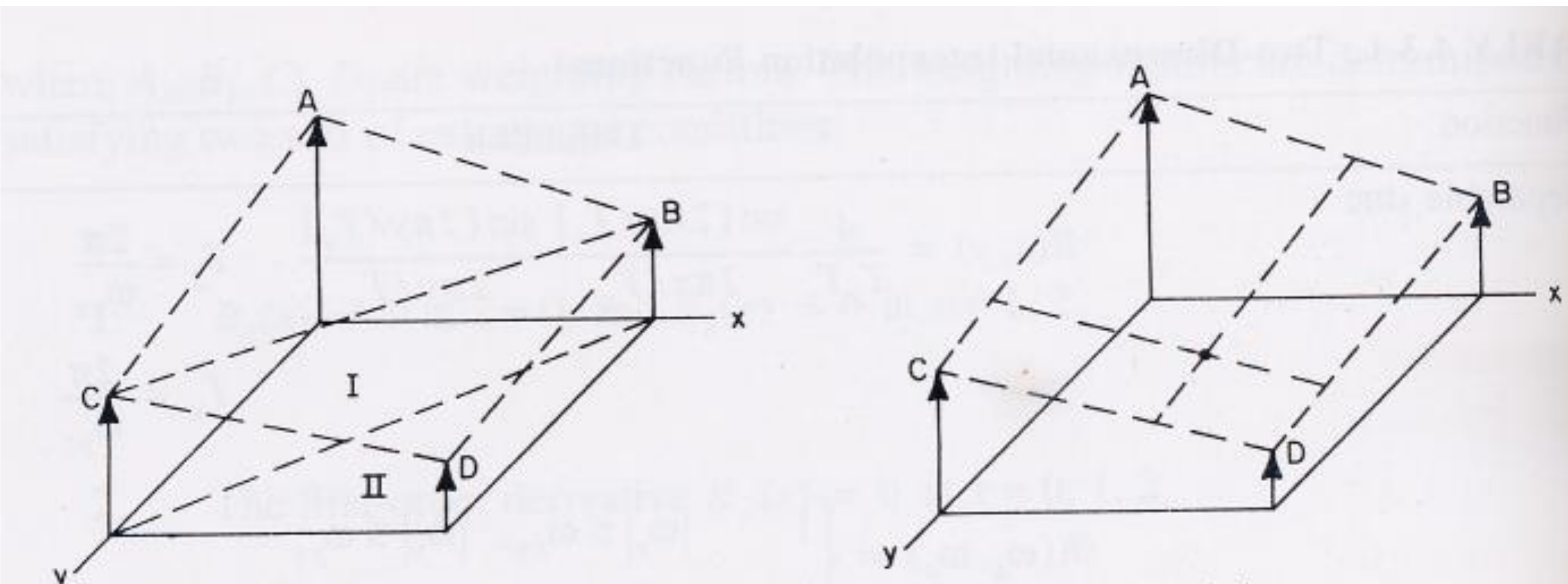
$$g(n_1, n_2) = \mathrm{A}_0 + \mathrm{A}_1 n_1 + \mathrm{A}_2 n_2 + \mathrm{A}_3 n_1 n_2, \tag{47}$$

which is a bilinear function in the coordinates (n_1, n_2). The bilinear weights A_0, A_1, A_2, and A_3 are found by solving

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 1 & n_{10} & n_{20} & n_{10}n_{20} \\ 1 & n_{11} & n_{21} & n_{11}n_{21} \\ 1 & n_{12} & n_{22} & n_{12}n_{22} \\ 1 & n_{13} & n_{23} & n_{13}n_{23} \end{bmatrix}^{-1} \begin{bmatrix} f(n_{10}, n_{20}) \\ f(n_{11}, n_{21}) \\ f(n_{12}, n_{22}) \\ f(n_{13}, n_{23}) \end{bmatrix}. \tag{48}$$

Two-Dimensional Image Interpolation

Piecewise Linear Interpolation *vs.* Bilinear Interpolation

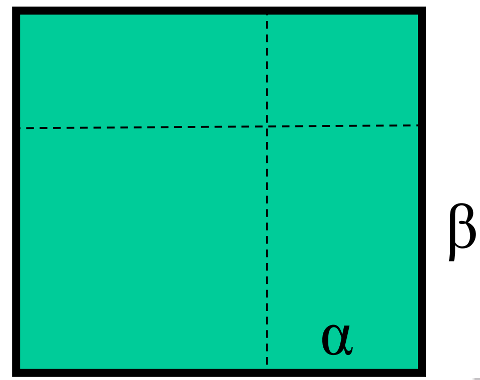


(a) Piecewise linear interpolation (b) Bilinear interpolation

Piecewise linear ： 同一個正方形pixel可以分成兩中三角形切法造成最後結果不一致
Bilinear：並非平面，四點可能構成曲面

Digital zoom by integer factors C x D

$(\lfloor m/C \rfloor, \lfloor n/D \rfloor)$



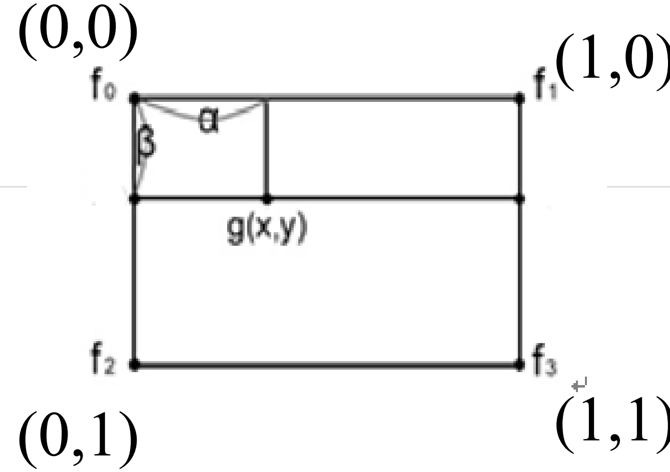
$(\lceil m/C \rceil, \lceil n/D \rceil)$

$$
\begin{aligned}
g(m, n) = {} & \alpha\beta f(\lfloor m/C \rfloor, \lfloor n/D \rfloor) \\
& + (1-\alpha)\beta f(\lceil m/C \rceil, \lfloor n/D \rfloor) \\
& + \alpha(1-\beta) f(\lfloor m/C \rfloor, \lceil n/D \rceil) \\
& + (1-\alpha)(1-\beta) f(\lceil m/C \rceil, \lceil n/D \rceil),
\end{aligned}
$$

where $\alpha = \lceil m/C \rceil - m/C$ and $\beta = \lceil n/D \rceil - n/D$.

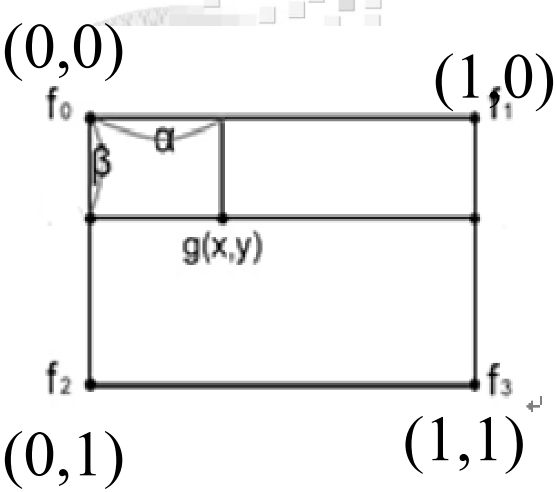
解A_0, A_1, A_2, A_3法一

$$g(\alpha,\beta) = A_0 + A_1\alpha + A_2\beta + A_3\alpha\beta$$



$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix}$$



解A_0, A_1, A_2, A_3法二

$$g(\alpha,\beta) = A_0 + A_1\alpha + A_2\beta + A_3\alpha\beta$$

$$f_0 = A_0 + A_1\cdot 0 + A_2\cdot 0 + A_3\cdot 0\cdot 0 = A_0$$

$$f_1 = A_0 + A_1\cdot 1 + A_2\cdot 0 + A_3\cdot 1\cdot 0 = A_0 + A_1$$

$$f_2 = A_0 + A_1\cdot 0 + A_2\cdot 1 + A_3\cdot 0\cdot 1 = A_0 + A_2$$

$$f_3 = A_0 + A_1\cdot 1 + A_2\cdot 1 + A_3\cdot 1\cdot 1 = A_0 + A_1 + A_2 + A_3$$

$\therefore$

$$A_0 = f_0$$

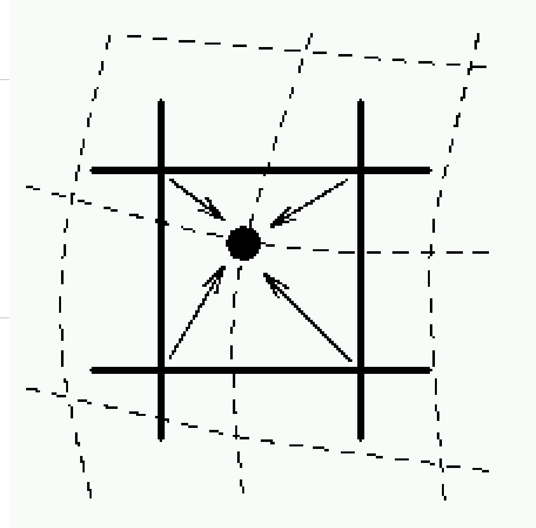
$$A_1 = f_1 - A_0 = f_1 - f_0$$

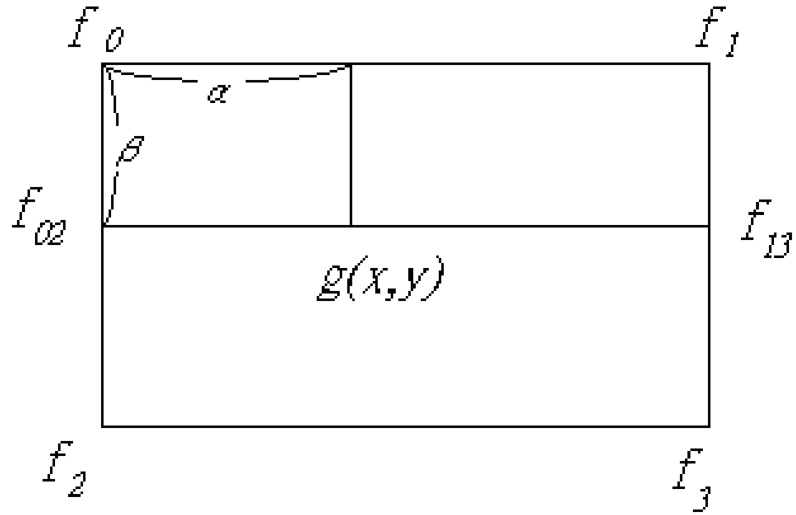
$$A_2 = f_2 - A_0 = f_2 - f_0$$

$$\begin{aligned} A_3 &= f_3 - A_0 - A_1 - A_2 = f_3 - f_0 - (f_1 - f_0) - (f_2 - f_0) \\ &= f_0 - f_1 - f_2 + f_3 \end{aligned}$$

將求得的A_0, A_1, A_2, A_3代入

$$
\begin{aligned}
g(\alpha,\beta) &= A_0 + A_1\alpha + A_2\beta + A_3\alpha\beta \\
&= f_0 + (f_1 - f_0)\alpha + (f_2 - f_0)\beta + (f_0 - f_1 - f_2 + f_3)\alpha\beta \\
&= (1-\alpha-\beta+\alpha\beta)f_0 + (\alpha-\alpha\beta)f_1 + (\beta-\alpha\beta)f_2 + \alpha\beta f_3 \\
&= (1-\alpha)(1-\beta)f_0 + \alpha(1-\beta)f_1 + \beta(1-\alpha)f_2 + \alpha\beta f_3
\end{aligned}
$$





$$g(\alpha,\beta) = (1-\alpha) * [\ (1-\beta) * f_0 + \beta * f_2]$$
$$+\ \alpha\ *\ [(1-\beta) * f_1 + \beta * f_3]$$

$$\Rightarrow\quad g(\alpha,\beta) = (1-\alpha)(1-\beta) f_0 + \alpha(1-\beta) f_1 + (1-\alpha)\ \beta\ f_2 + \alpha\beta\ f_3$$

$$\Rightarrow\quad g(\alpha,\beta) = f0 + \alpha(-f0+f1) + \beta(-f0+f2) + \alpha\beta(f0-f1-f2+f3)$$

where $A0 = f0$, $A1 = -f0+f1$, $A2 = -f0+f2$, $A3 = f0-f1-f2+f3$

Chapter 2: Digital Image Fundamentals

- Elements of Visual Perception
- Light and EM Spectrum
- Image Sensing and Acquisition
- Image Sampling and Quantization
- Basic Relationships between Pixels
- Mathematical Tools in DIP

2.5 Basic Relationships Between Pixels

- 4-neighbors of p, $N_4(p)$
- 8-neighbors of p, $N_8(p)$
- four diagonal neighbors of p, $N_D(p)$

Distance Measure

- Euclidean distance
- D_4 distance (city-block distance)
- D_8 distance (chessboard distance)