Chapter 2
影像處理基本介紹

2.1 眼球構造

- Cones (視錐細胞)
  - 約 6~7 百萬個視錐細胞
  - 主要分佈在視網膜中央窩(fovea)附近
  - 對顏色敏感
- Rods (視桿細胞)
  - 約 75~150百萬個視桿細胞
  - 分佈在除了盲點區域的整個視網膜上。
  - 對亮度敏感，低亮度也會有感覺。

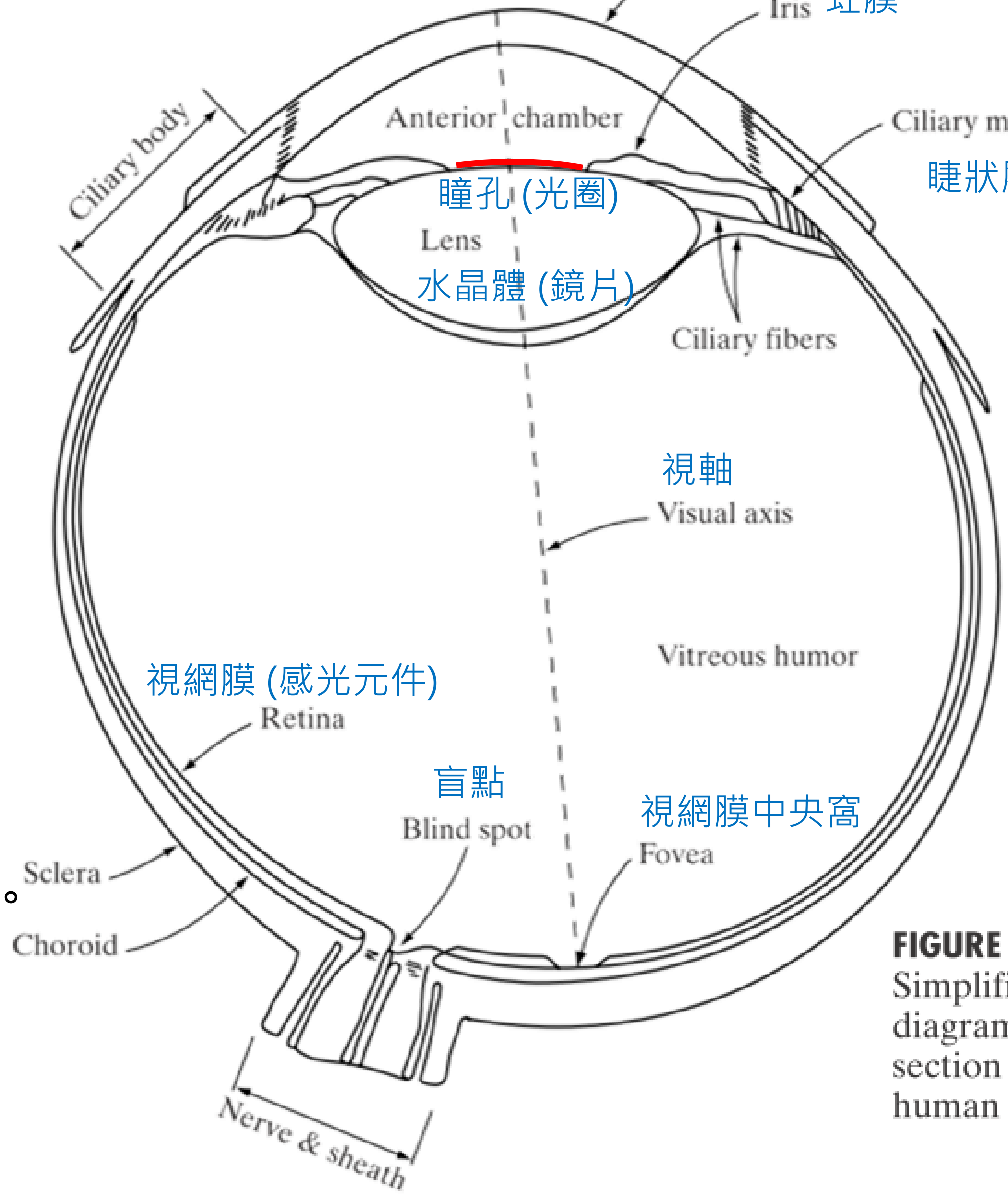


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

2.1 眼球構造

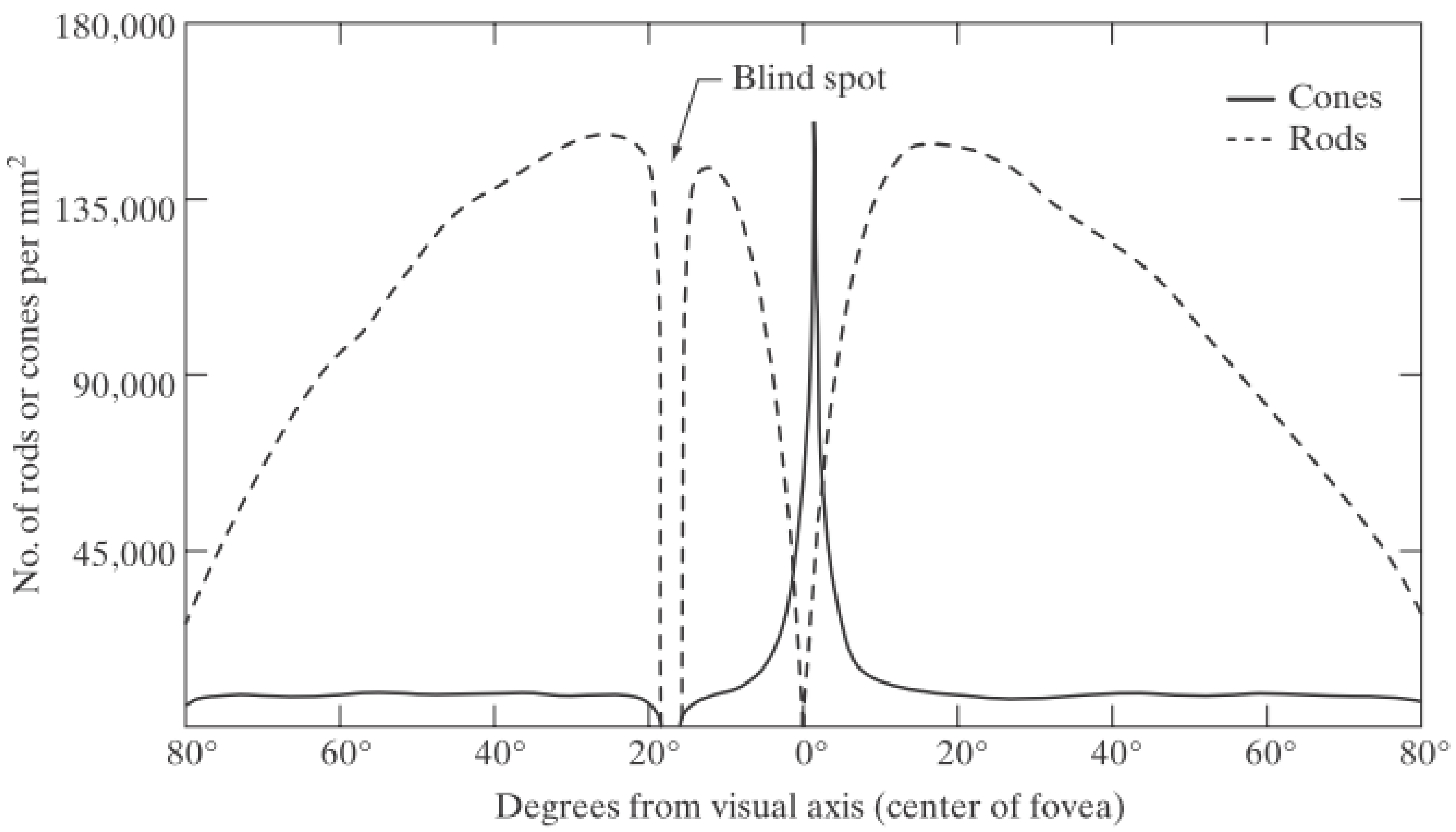
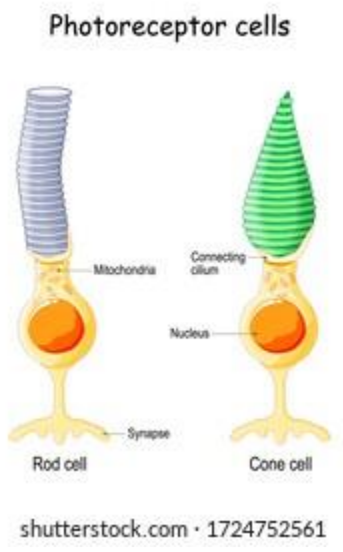


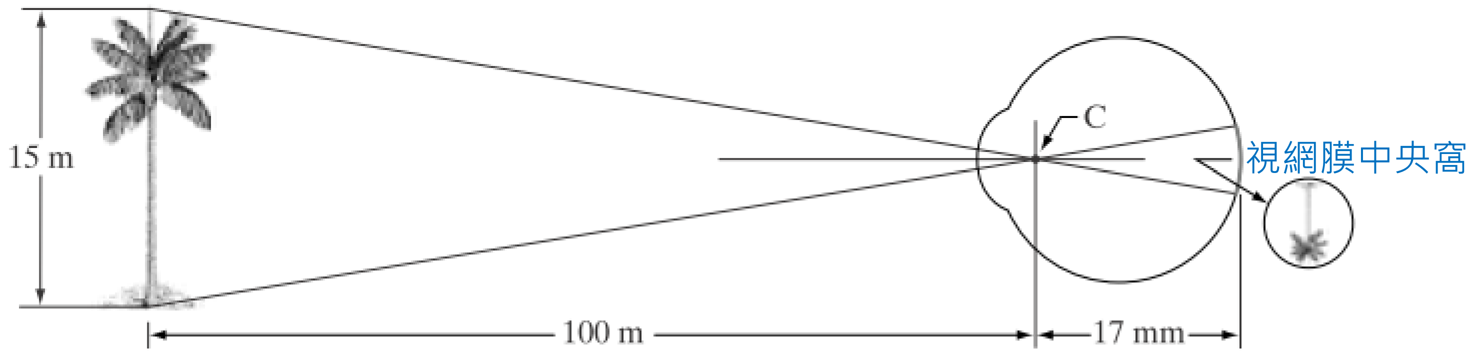
FIGURE 2.2
Distribution of rods and cones in the retina.



2.1 眼球構造

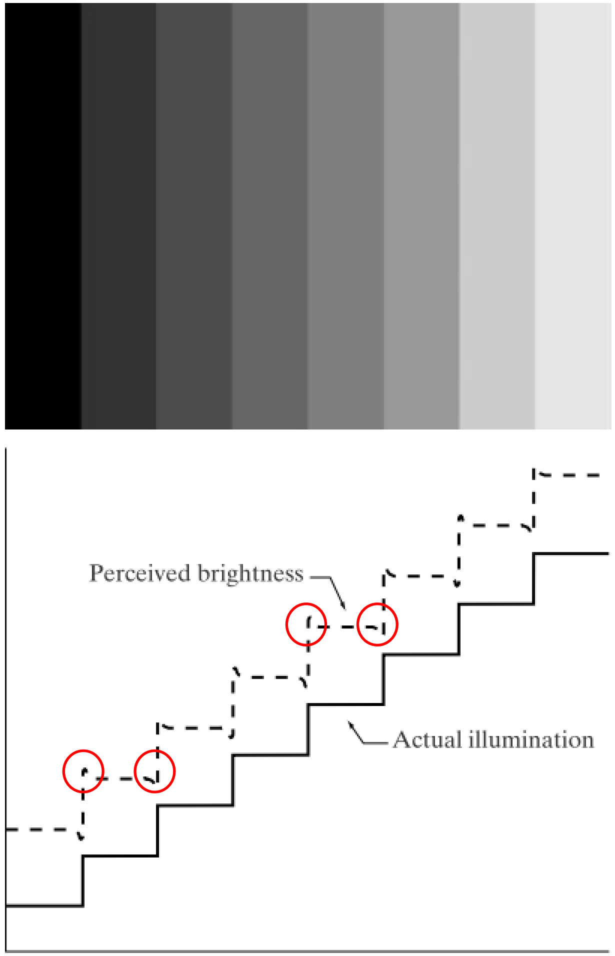
- Image formation in eye:

FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.



2.1 眼球構造

- Visual Perception:



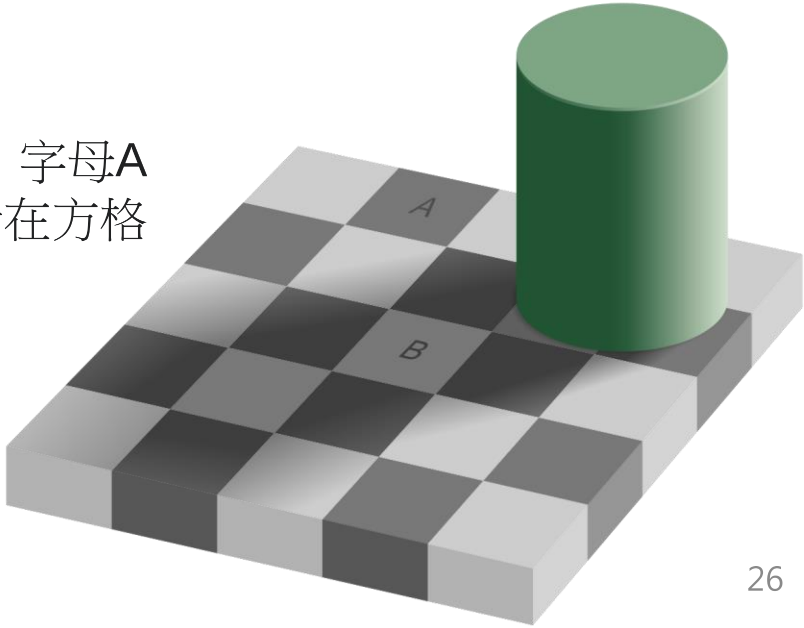
a
b

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.

a b c

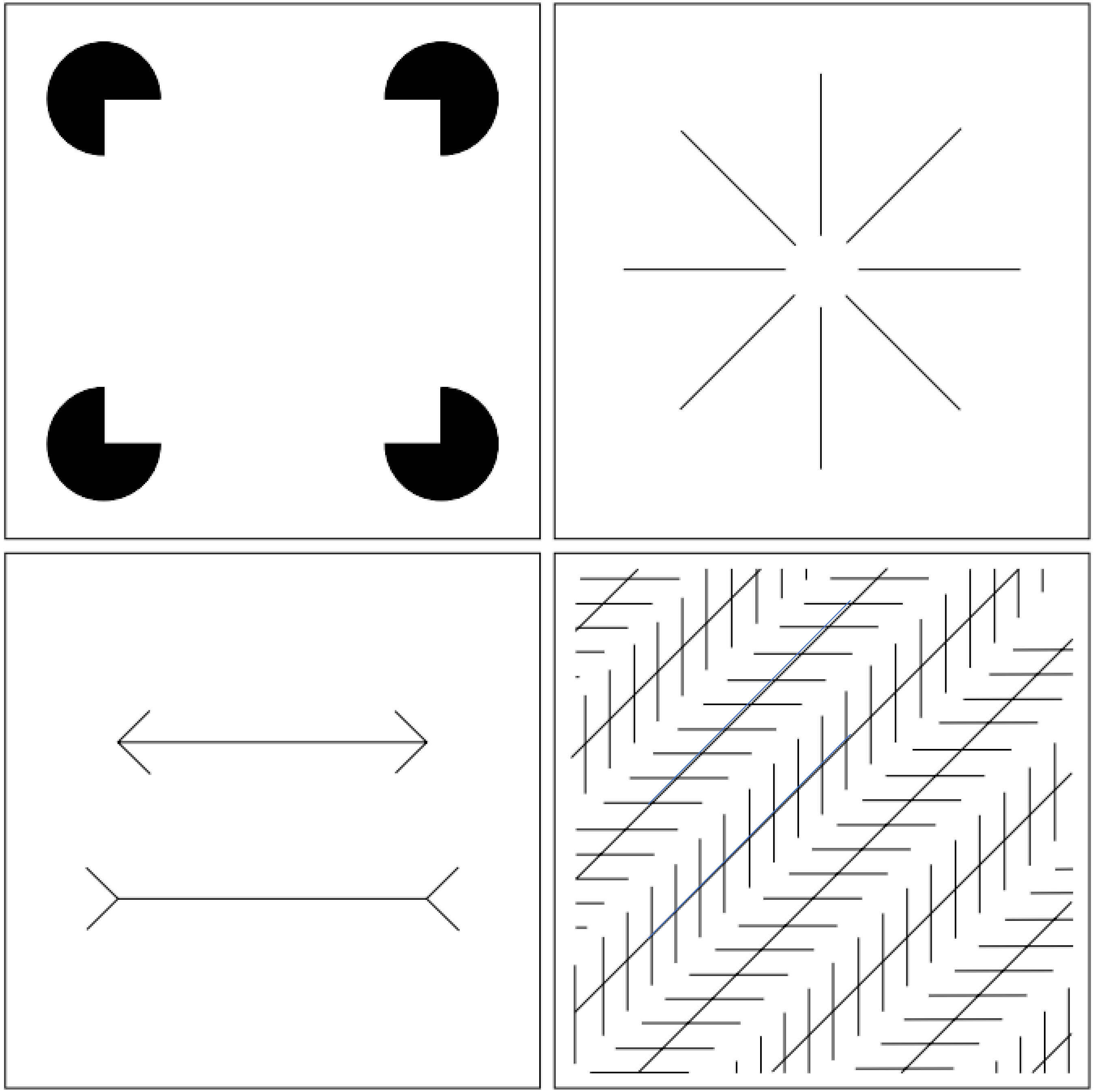
FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

[棋盤陰影]：字母A與字母B所在方格實為同色。



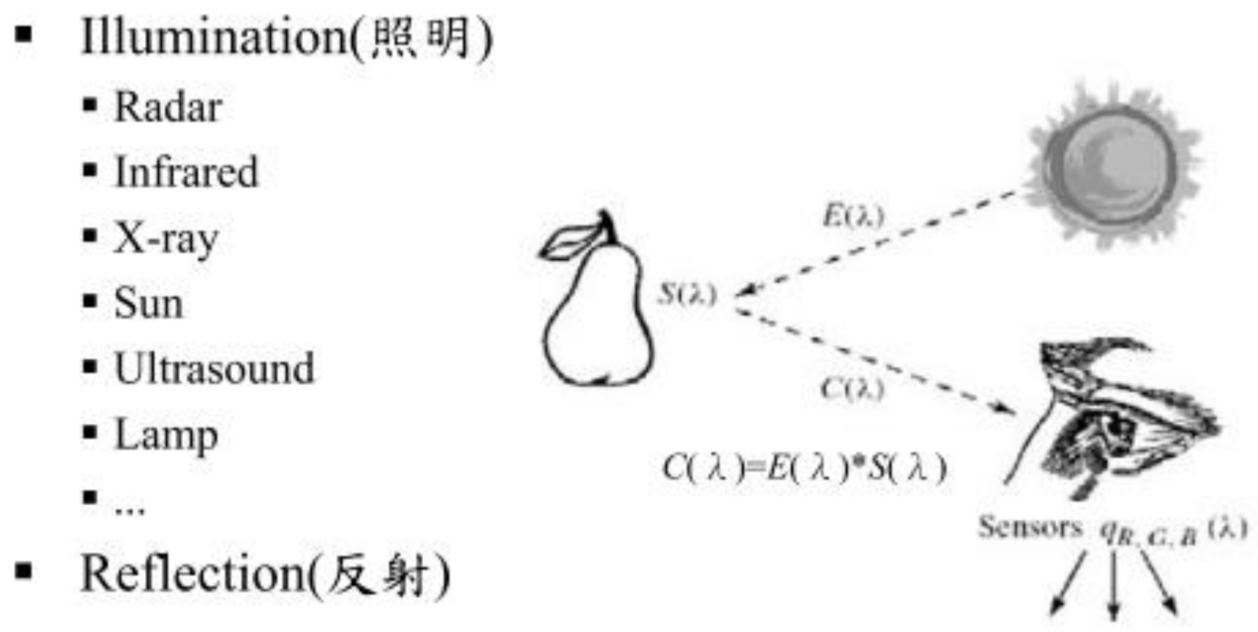
2.1 眼球構造

- optical illusions:

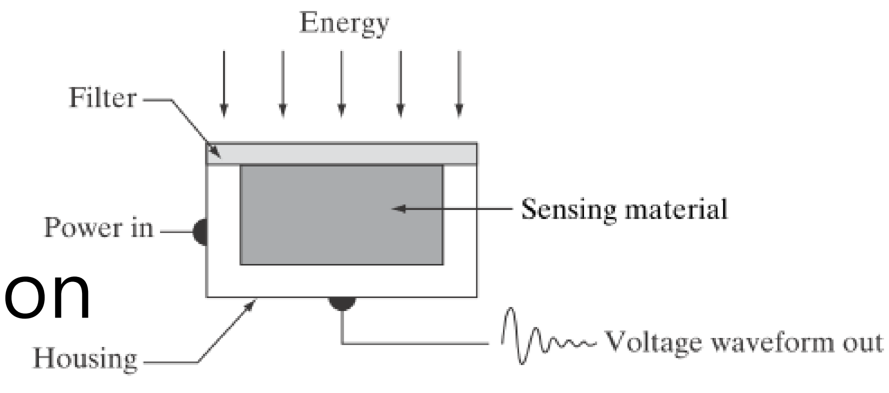


2.2 Image sensing

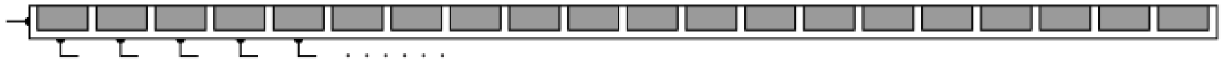
- 可見光大約在 430um~790um
  - Illumination (照度)- 單位lumens
  - Brightness (亮度)-感受程度, gray intensity
  - Reflection (反射)

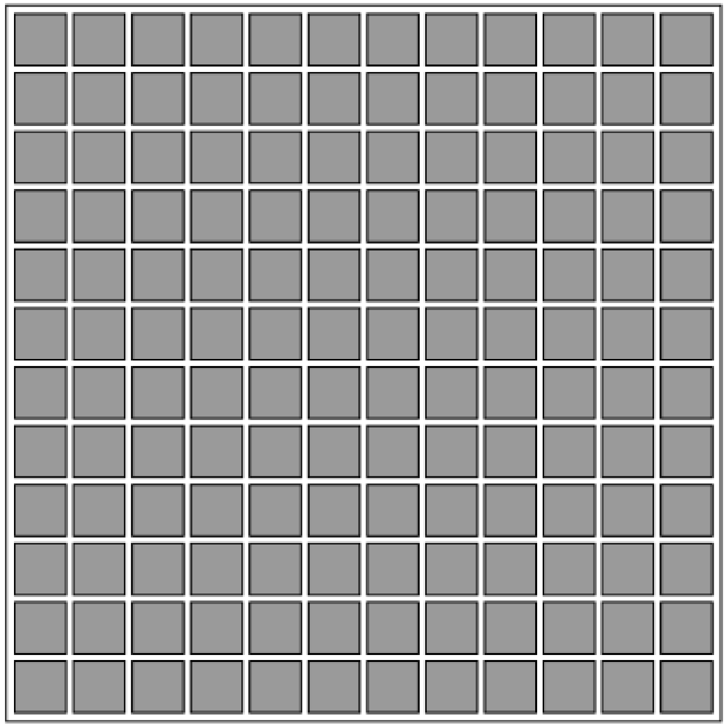


2.2 Image sensing



- Purpose: transform illumination energy into digital images
- Three principal sensor arrangements:
  1. Single sensor
  2. Sensor Stripes (Line sensor)
  3. Sensor Arrays (Array sensor)





2.2 Image sensing

1. Single sensor

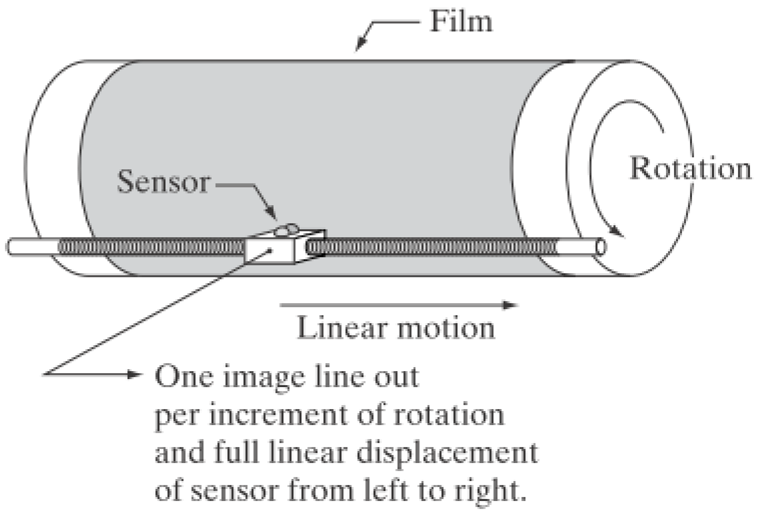


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

2.2 Image sensing

2. Sensor Stripes
   (Line sensor)

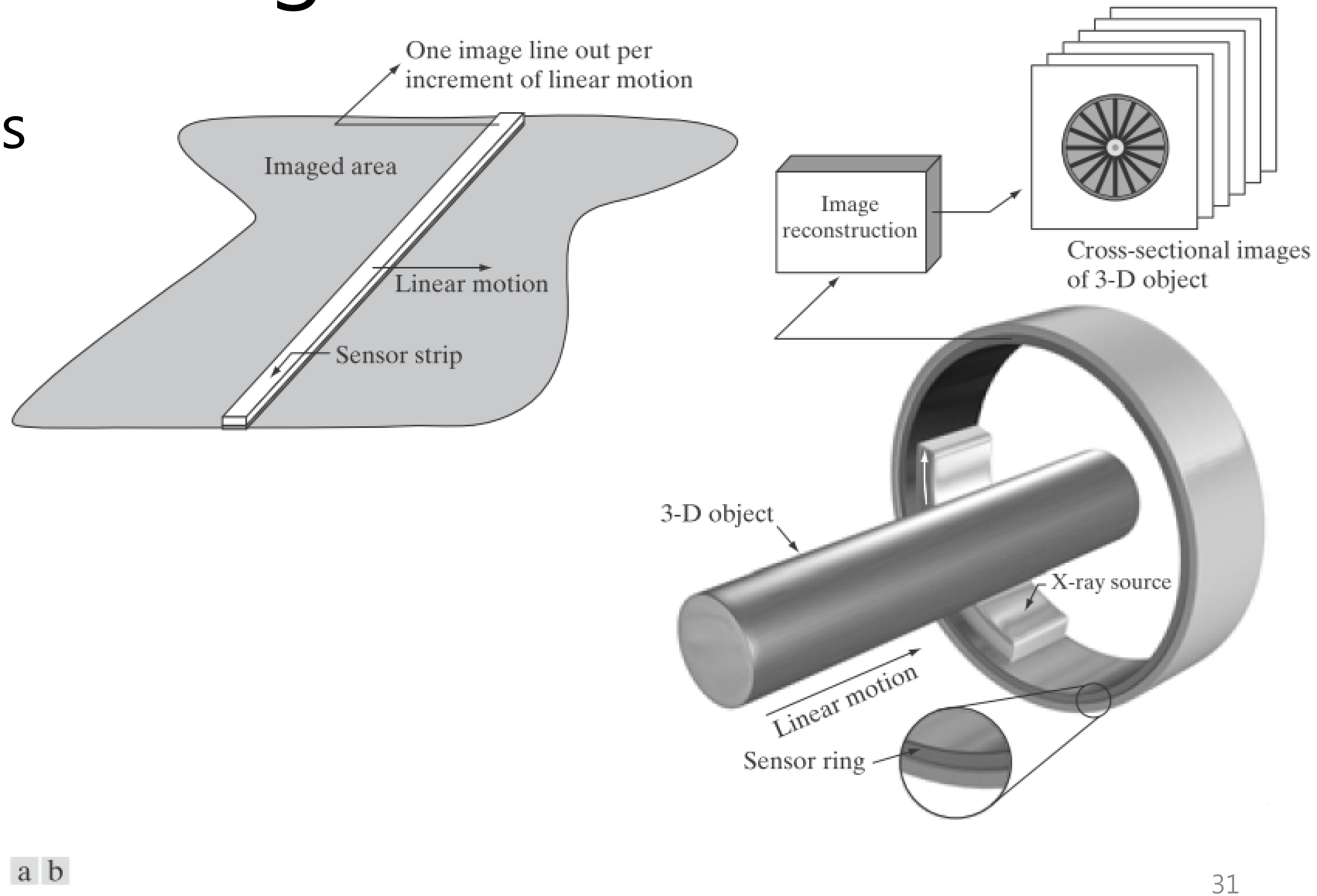


FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

2.2 Image sensing

3. Sensor Arrays
(Array sensor)

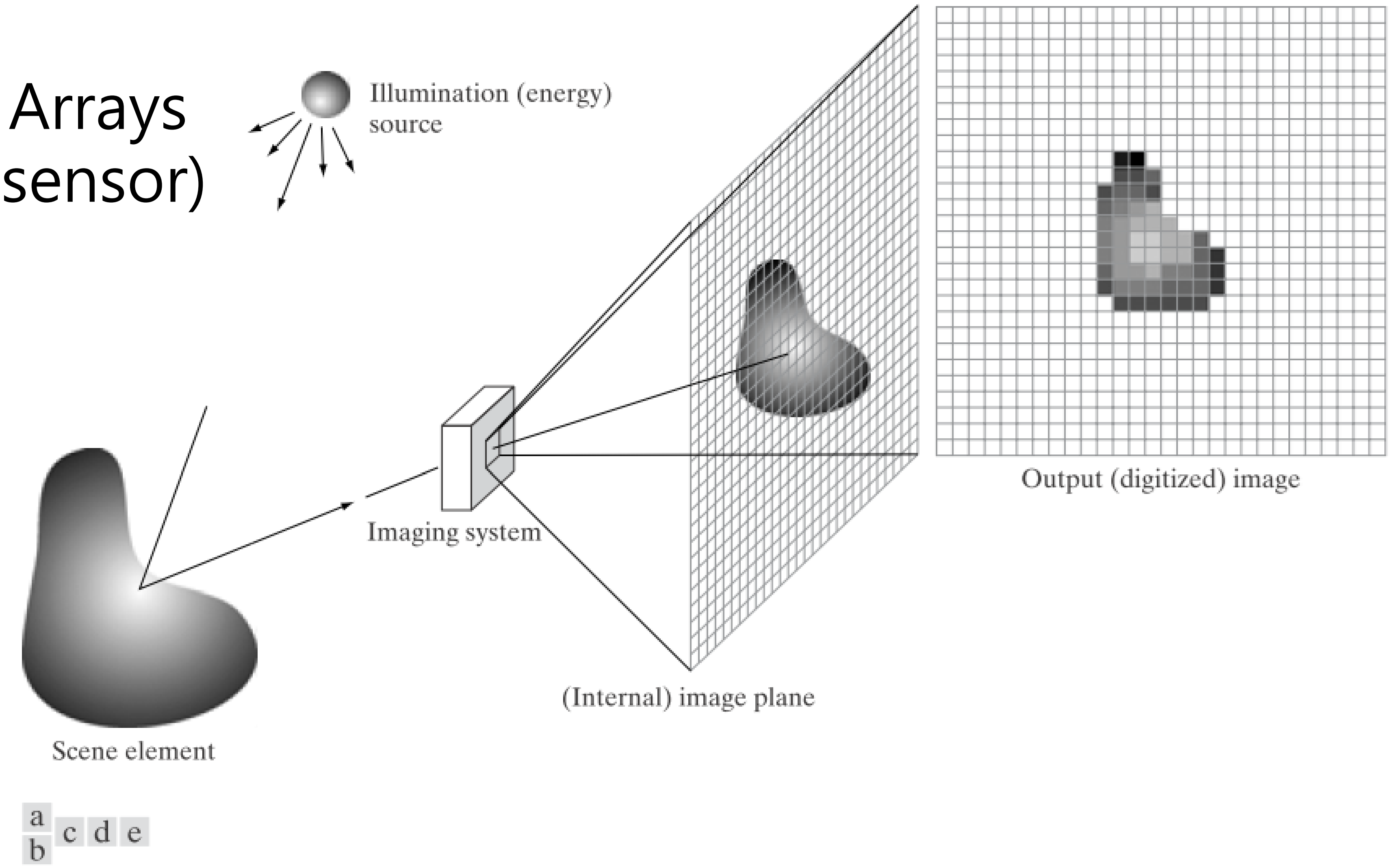


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

2.3 A simple image formation model

- Image: a two-dimensional function $f(x,y): 0 < f(x,y) < \infty$

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

where $0 < i(x,y) < \infty$ $i(x,y)$: illumination

$0 < r(x,y) < 1$ $r(x,y)$: reflectance

- 幾個照度的例子：$i(x,y)$
  - 晴天的太陽 $i = 90000\ \mathrm{lm/m^2}$
  - 陰天的太陽 $i = 10000\ \mathrm{lm/m^2}$
  - 晴天的滿月 $i = 0.1\ \mathrm{lm/m^2}$
  - 辦公室照明 $i = 1000\ \mathrm{lm/m^2}$
- 幾個反射率的例子：$r(x,y)$
  - 黑天鵝絨 $r = 0.01$
  - 不銹鋼 $r = 0.65$
  - 上漆白牆 $r = 0.8$
  - 銀盤 $r = 0.9$
  - 雪 $r = 0.93$

2.3 A simple image formation model

- Gray level: the intensity (luminance) of a monochrome image at (x_0, y_0)

$$\ell = f(x_0, y_0)$$

$$\text{where } L_{\min} < \ell < L_{\max}$$

- The interval $[L_{min}, L_{max}]$ is called the gray scale.

  ➔common practice $[0, L-1]$

  where 0 is black, L-1 is white, and intermediate values are shades of gray varying from black to white.

2.4 Image Sampling and Quantization

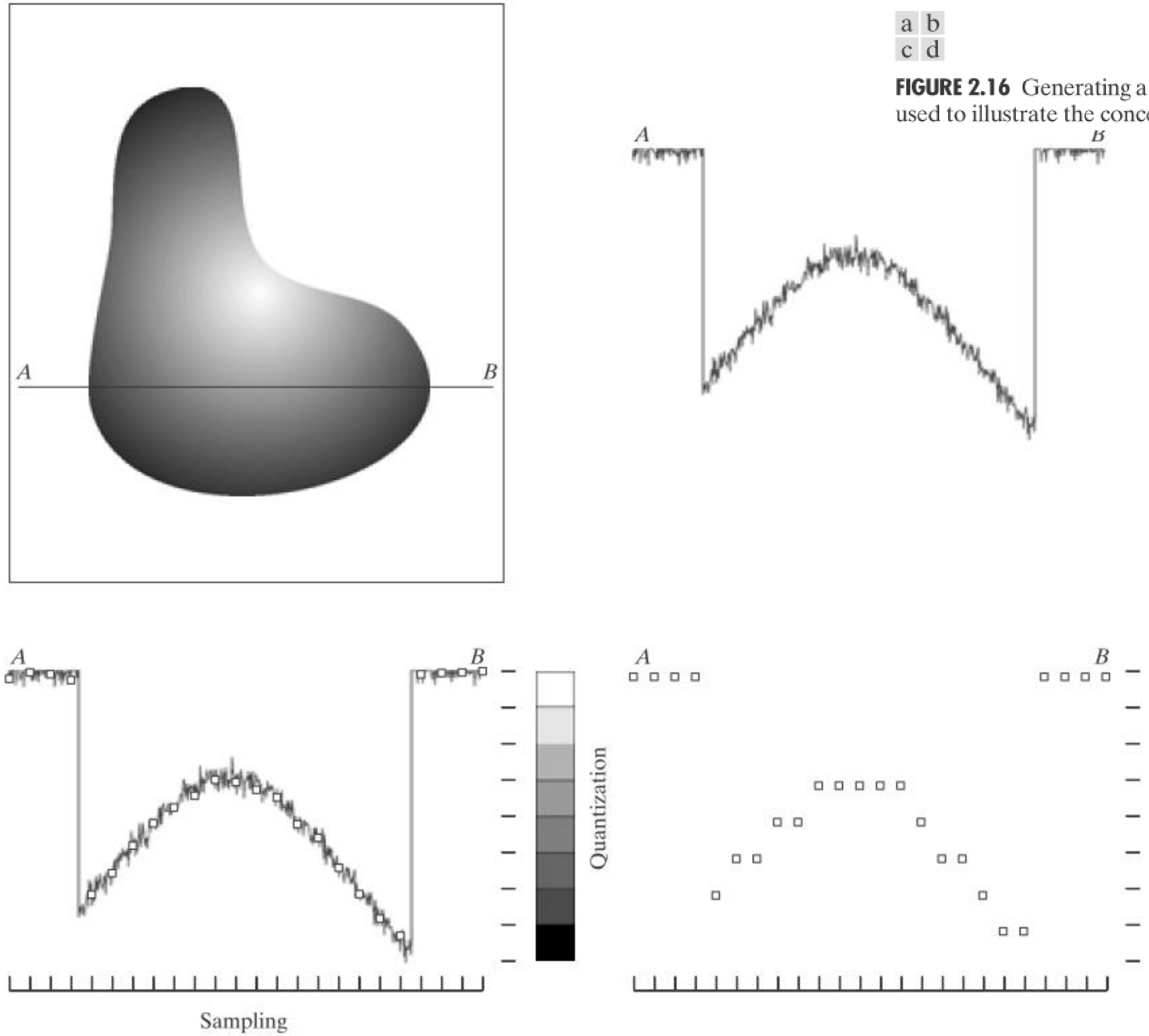
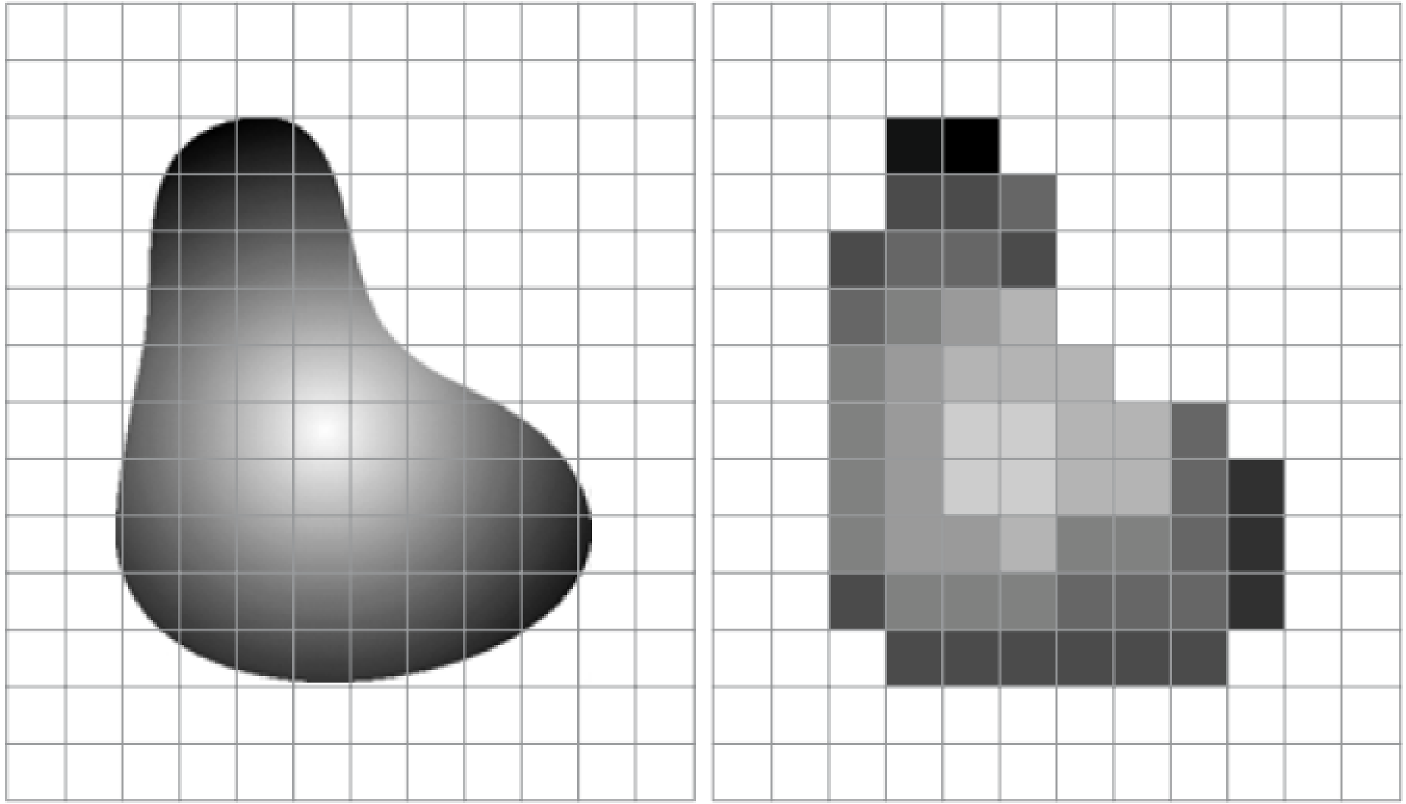


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

- Sampling: digitize the coordinate values
- Quantization: digitize the amplitude values

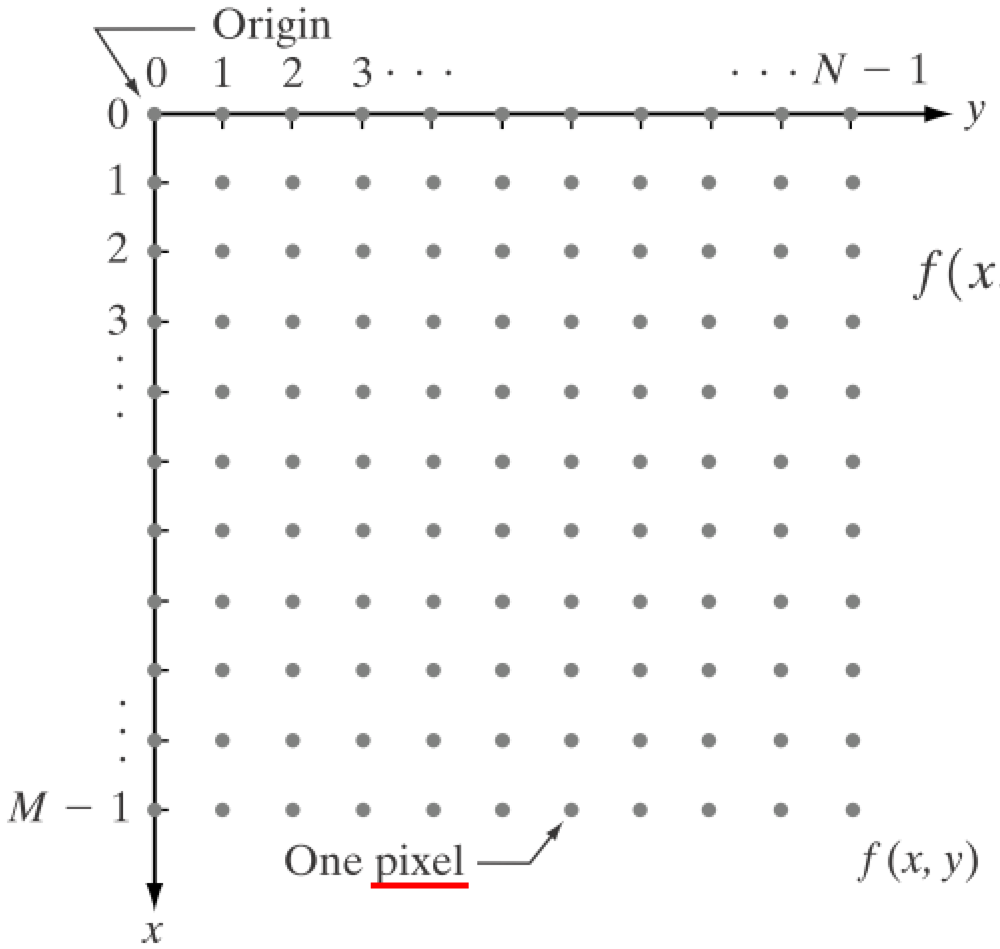
2.4 Image Sampling and Quantization



a b

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

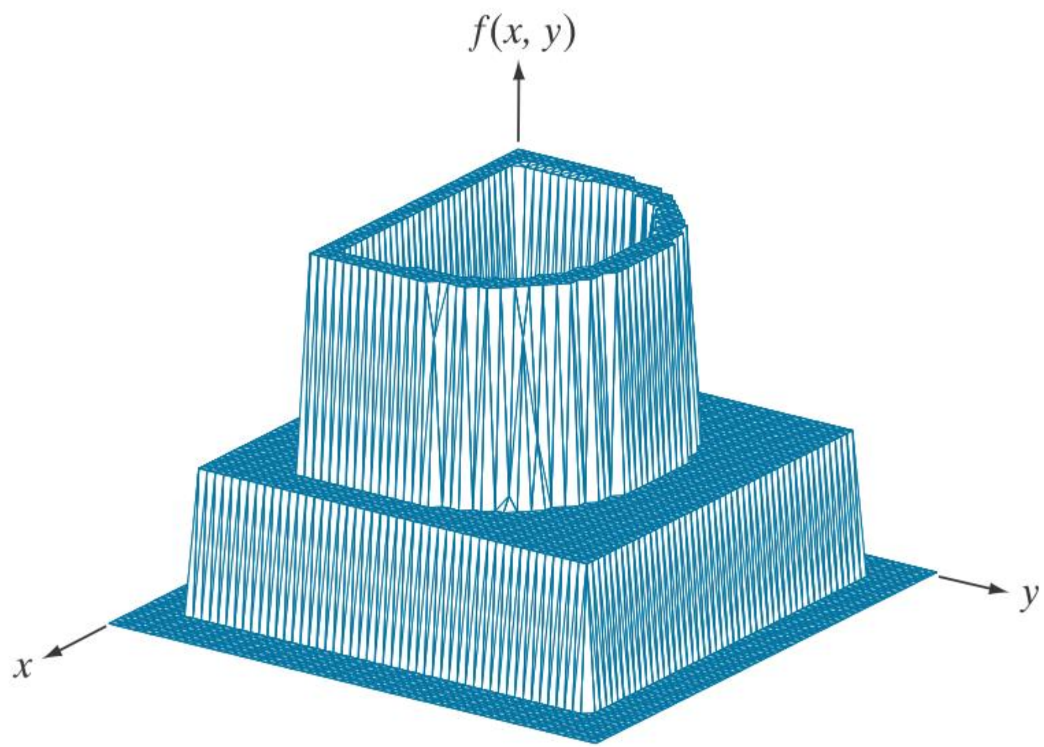
2.5 Representing digital images

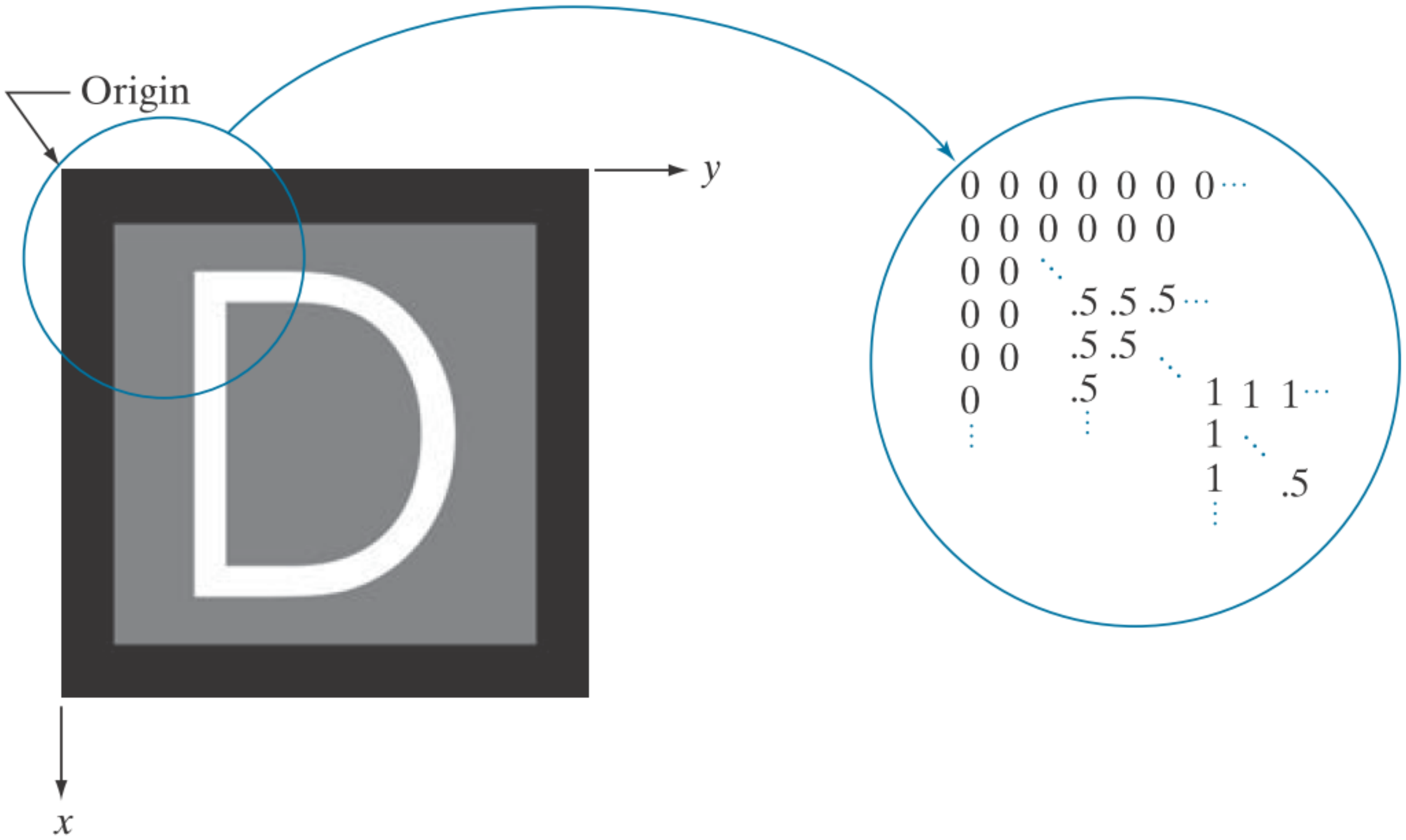


$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix}$$

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

2.5 Representing digital images





2.6 Dynamic range and image storage size

- The number of gray scale typically is an integer power of two.

  $L = 2^k$

- Dynamic range: the range of value spanned by the gray scales.

  $[0, L-1], L = 2^k$

- Bits required to store a digitized image is $b = M \times N \times k$
- When M=N, the equation becomes $b = M^2 \times k$

2.6 Dynamic range and image storage size

- Assume M=N,

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

Usually, k=8, 10, 12, 14, 16

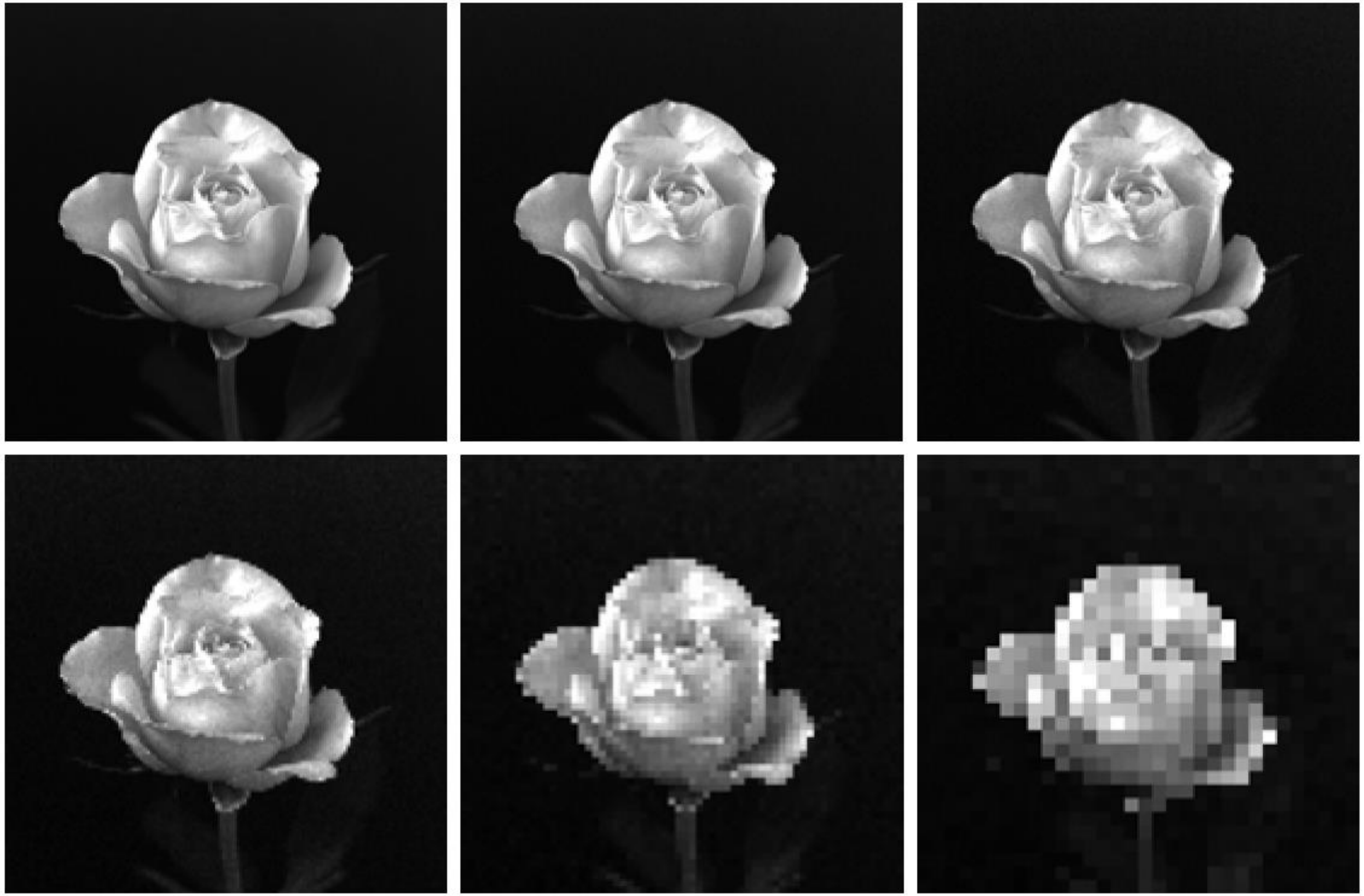
N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	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

2.7 Spatial and Intensity resolution

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.

downsample

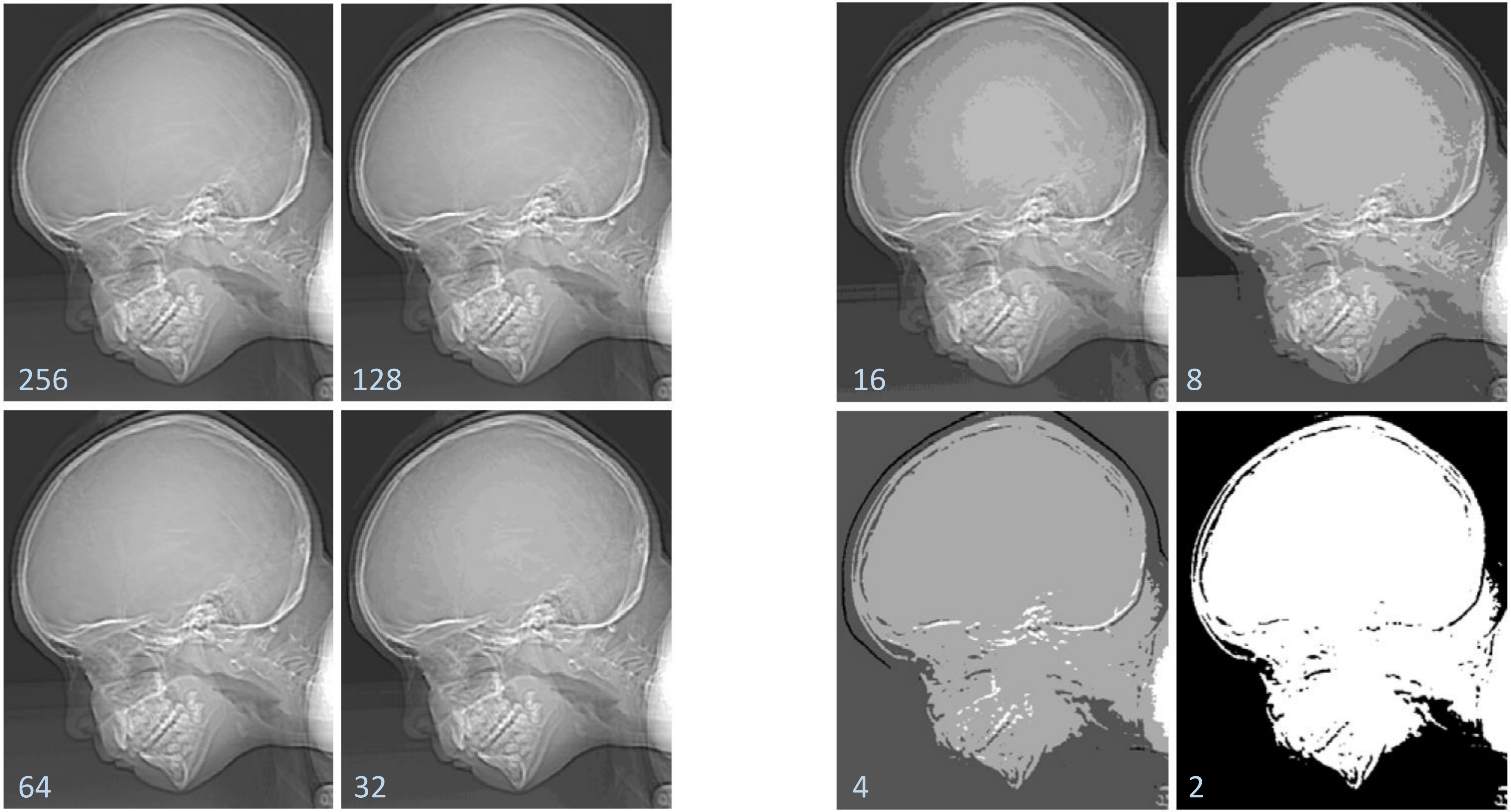
2.7 Spatial and Intensity resolution



a b c
d e f

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.

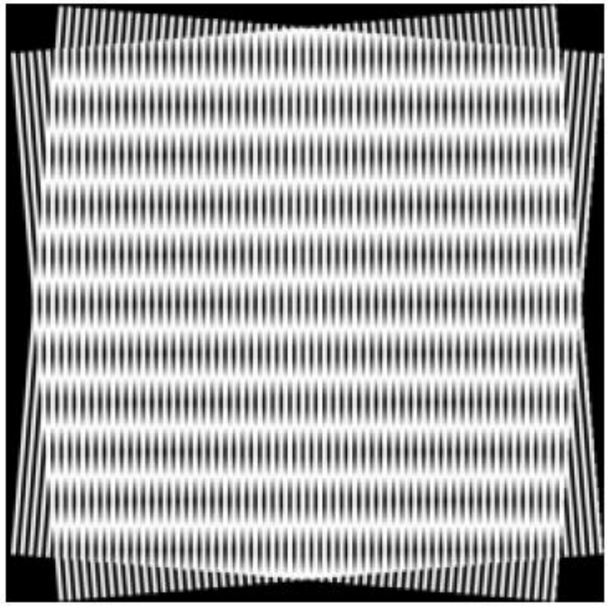
2.7 Spatial and Intensity resolution

- 

2.8 Aliasing

- Nyquist–Shannon sampling theorem: if the function is sampled at a rate equal to or greater than twice its highest frequency, it is possible to recover the original function.
- Aliasing: it the function is undersampled, then a phenomenon called aliasing corrupts the sampled image.
- Sampling rate: the sampling rate in images is the number of samples taken per unit distance. (PPI, DPI)

Moiré Patterns



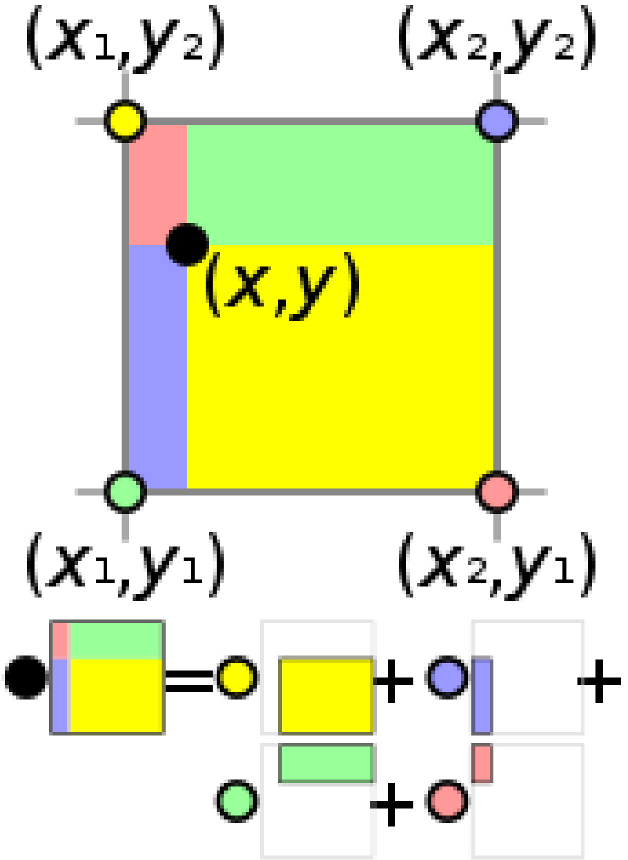
2.8 Aliasing

a b
c d

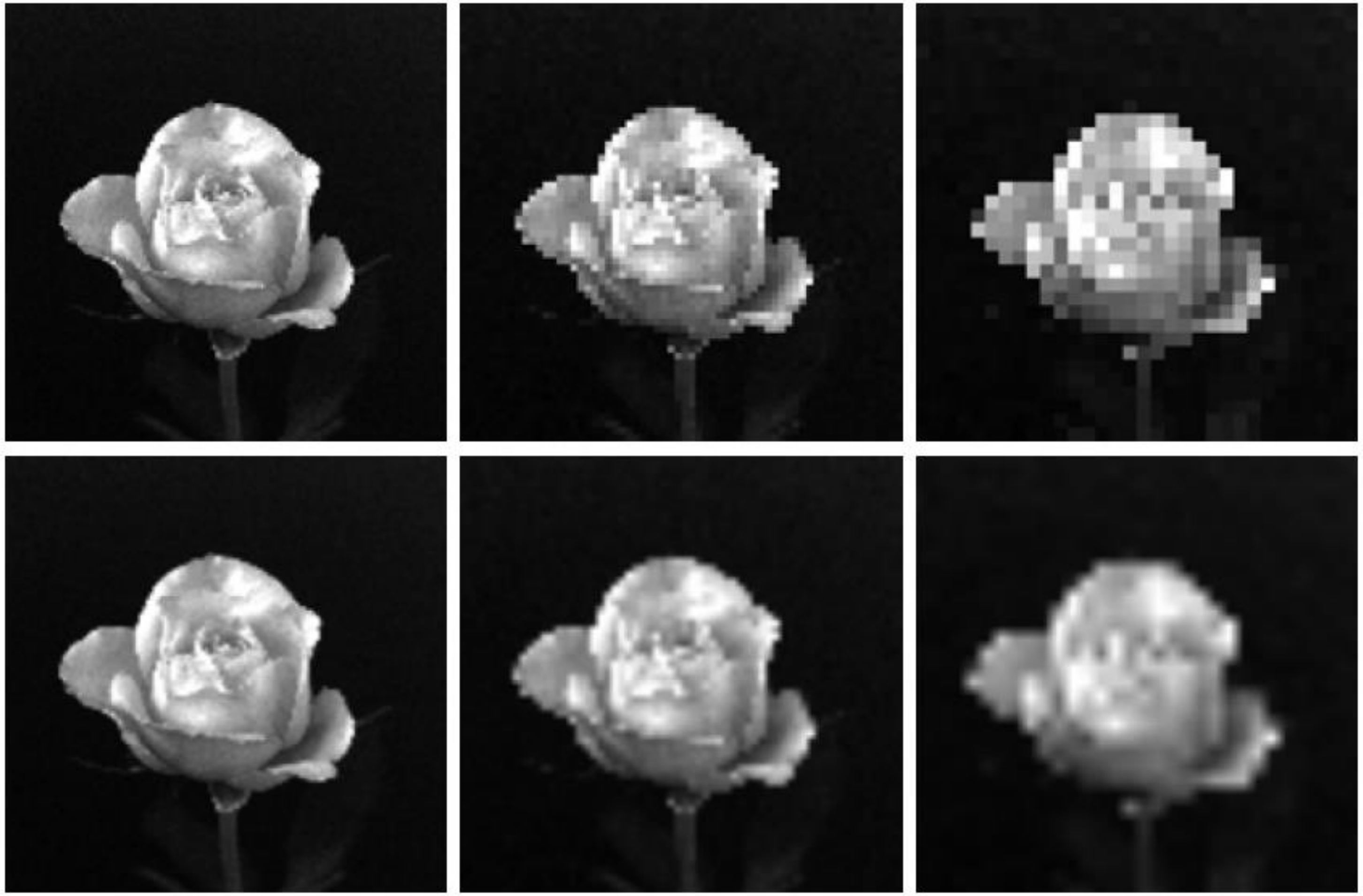
FIGURE 2.23
Effects of reducing spatial resolution. The images shown are at:
(a) 930 dpi,
(b) 300 dpi,
(c) 150 dpi, and
(d) 72 dpi.

2.9 Zooming and shrinking digital images

- Nearest neighbor interpolation.
- Bilinear interpolation.
- Other higher order methods.
  - Gaussian
  - Bicubic
  - Bicubic-spline
  - Lanczos
  - Hermit
  - Mitchell
  - Bell



2.9 Zooming and shrinking digital images



a	b	c
d	e	f

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.

2.10 Some basic relationships between pixels

- Neighbors of a pixel
- Adjacency, connectivity, regions, and boundaries
- Distance measures
- Image operation on a pixel basis

2.10 Some basic relationships between pixels

- Neighbors of a pixel:
  - $N_4(p)$: 4-neighbors of $p(x, y)$
    - point set of the horizontal and vertical pixels
    - $(x\text{-}1, y), (x\text{+}1, y), (x, y\text{-}1), (x, y\text{+}1)$
  - $N_D(p)$: diagonal neighbors of $p(x, y)$
    - $(x\text{-}1, y\text{-}1), (x\text{+}1, y\text{+}1), (x\text{+}1, y\text{-}1), (x\text{-}1, y\text{+}1)$
  - $N_8(p)= N_4(p) \cup N_D(p)$: 8-neighbors of $p(x, y)$
    - $(x\text{-}1, y\text{-}1), (x, y\text{-}1), (x\text{+}1, y\text{-}1),$
      $(x\text{-}1, y), (x\text{+}1, y), (x\text{-}1, y\text{+}1), (x, y\text{+}1), (x\text{+}1, y\text{+}1)$

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

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

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

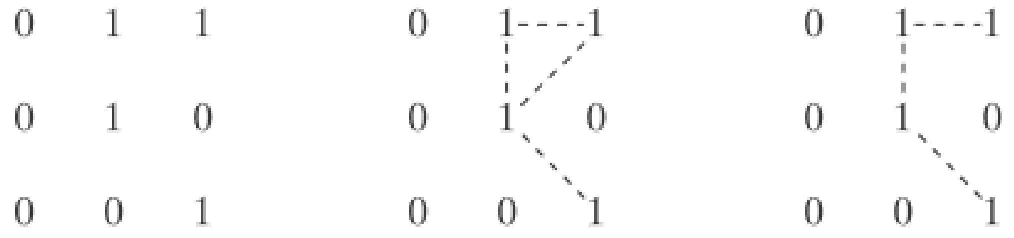
2.10 Some basic relationships between pixels

- Adjacency :

  Two pixels p and q with values are

  - 4-adjacency: q is in the set $N_4(p)$
  - 8-adjacency: q is in the set $N_8(p)$.
  - Mixed-adjacency: q is in the set N_4 (p) or
    q is in the set N_D (p) and $N_4 (p) \cap N_4 (q)$
    has no pixels with value

  -To eliminate the ambiguities in 8-adjacency



a b c

FIGURE 2.26 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

2.10 Some basic relationships between pixels

- Connectivity :
  - A digital path (or curve) from pixel p with coordinates (x, y) to pixel q with coordinates (s, t) is a sequence of distinct pixels with coordinates

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

    where $(x_0, y_0)=(x, y)$, $(x_n, y_n)=(s, t)$, and pixels(x_i, y_i) and $(x_i\text{-}1, y_i\text{-}1)$ are adjacent for $1 \leq \mathrm{I} \leq n$.
  - n: the length of the path.
  - If $(x_0, y_0) = (x_n, y_n)$, the path is a closed path.

2.10 Some basic relationships between pixels

- Connectivity :
  - *S*: a subset of pixels in an image.
  - Connected: *p* and *q* are said to be connected if there exists a path between them consisting entirely of pixels in *S*.
  - Connected component: for any pixel *p* in *S*, the set of pixels that are connected to it in *S* is called a connected component of *S*.
  - Connected set: If it only has one connected component, then set *S* is called a connected set.

2.10 Some basic relationships between pixels

- Regions and boundaries :
  - *R*: a subset of pixels in an image.
  - Region: *R* is a region of the image if *R* is a connected set.
  - Boundary: (also called border or contour) The boundary of a region R is the set of pixels in the region that have one or more neighbors that are not in R.

2.10 Some basic relationships between pixels

- Distance measures:
  - Distance function (metric)
  - Euclidean distance
  - D_4 distance (city-block distance)
  - D_8 distance (chessboard distance)
  - D_m distance .

2.10 Some basic relationships between pixels

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

      (a) $D(p, q) \geq 0 \quad (D(p, q) = 0 \quad \text{iff} \quad p = q)$,
      (b) $D(p, q) = D(q, p)$, and
      (c) $D(p, z) \leq D(p, q) + D(q, z)$.
  - Euclidean distance:
    - The Euclidean distance between p and q is defined as

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

2.10 Some basic relationships between pixels

- Distance measures:
  - D_4 distance (city-block distance)
    - The D_4 distance (also called city-block distance) between p and q is defined as
      $$D_4(p,q) = |(x-s)| + |(y-t)|$$
    - In this case, the pixels having a D_4 distance from (*x*, *y*) less than or equal to some value *r* form a diamond centered at (x, y).
    - The pixels with D_4 distance ≦ 2 from (*x*, *y*) (the center point) form the following contours of constant distance:

$$\begin{array}{ccccc} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{array}$$

The pixels with D_4 =1 are the 4-neighbors of (*x*, *y*).

2.10 Some basic relationships between pixels

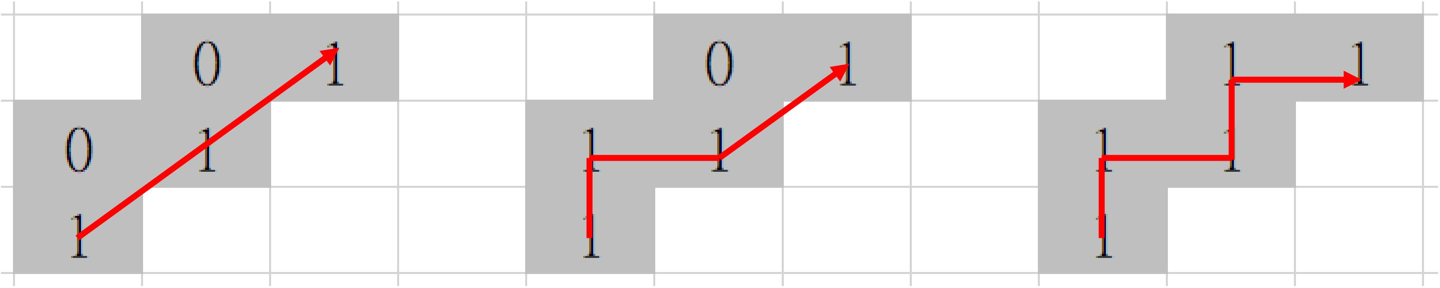
- Distance measures:
  - D_8 distance (chessboard distance)
    - The D_8 distance (also called chessboard distance) between p and q is defined as
      $$D_8(p,q) = \max\left(\left|(x-s)\right|, \left|(y-t)\right|\right)$$
    - In this case, the pixels having a D_8 distance from (x, y) less than or equal to some value r form a square centered at (x, y).
    - The pixels with D_8 distance ≦ 2 from (x, y) (the center point) form the following contours of constant distance:

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

The pixels with D_8 =1 are the 8-neighbors of (x, y).

2.10 Some basic relationships between pixels

- Distance measures:
  - D_m distance
    - If we elect to consider the *m*-adjacency, the D_m distance between two points is defined as the shortest m-path between the points.



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

Miin Wu School of Computing 敏求智慧運算學院

2.11 Image operation on a pixel basis

- Linear operation and nonlinear operations:
  - H: an operator whose input and output are images. *H* is said to be linear if

    $$H(af+bg)=aH(f)+bH(g)$$

    where *f* and *g* are images, *a* and *b* are scaling factors.
  - Ex: Sum operation is linear.

$$f=\begin{bmatrix}0 & 5\\ 10 & 20\end{bmatrix}, g=\begin{bmatrix}10 & 5\\ 1 & 2\end{bmatrix}$$

$$\begin{aligned}sum(2f+3g) &= sum\left(2\times\begin{bmatrix}0 & 5\\ 10 & 20\end{bmatrix}+3\times\begin{bmatrix}10 & 5\\ 1 & 2\end{bmatrix}\right)\\ &= sum\left(\begin{bmatrix}30 & 25\\ 21 & 46\end{bmatrix}\right)\\ &= sum(2f)+sum(3g)\end{aligned}$$

max operation is not linear.

$$\begin{aligned}\max(2f+3g) &= \max\left(2\times\begin{bmatrix}0 & 5\\ 10 & 20\end{bmatrix}+3\times\begin{bmatrix}10 & 5\\ 1 & 2\end{bmatrix}\right)\\ &= \max\left(\begin{bmatrix}30 & 25\\ 23 & 46\end{bmatrix}\right)=46\\ &\neq \max(2f)+\max(3g)\end{aligned}$$

2.11 Image operation on a pixel basis

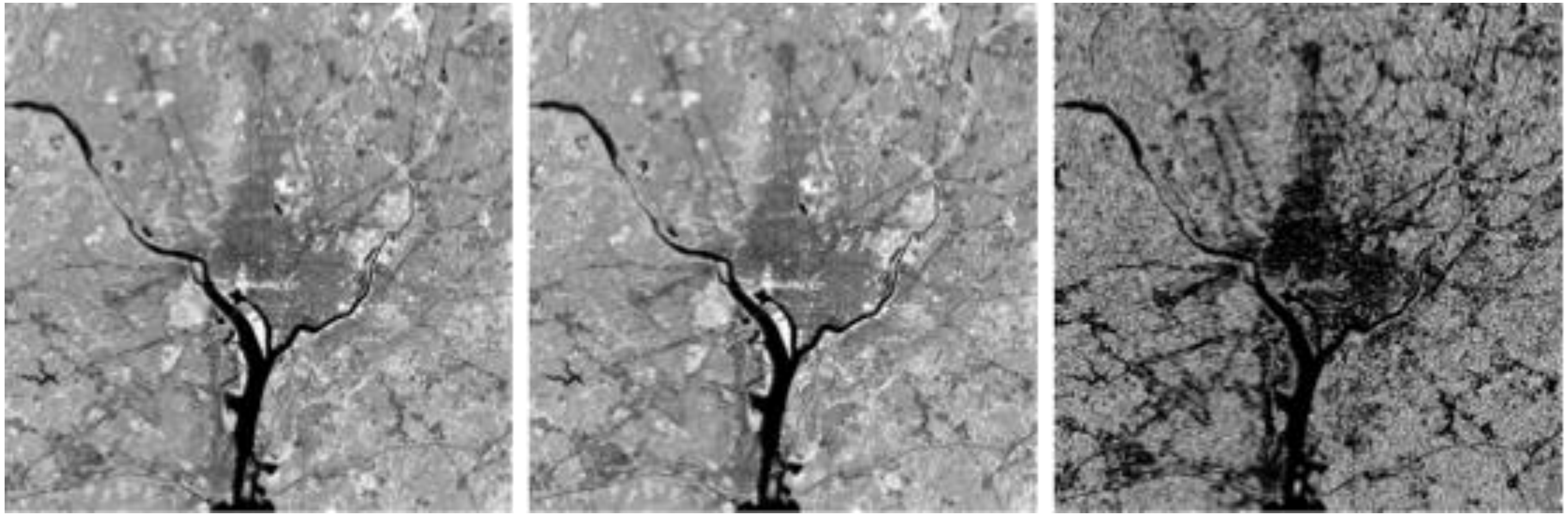
- Arithmetic operations:

$$s(x,y) = f(x,y) + g(x,y)$$
$$d(x,y) = f(x,y) - g(x,y)$$
$$p(x,y) = f(x,y) \times g(x,y)$$
$$v(x,y) = f(x,y) \div g(x,y)$$

These are elementwise operations.

2.11 Image operation on a pixel basis

- Comparing images using 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.)

2.11 Image operation on a pixel basis

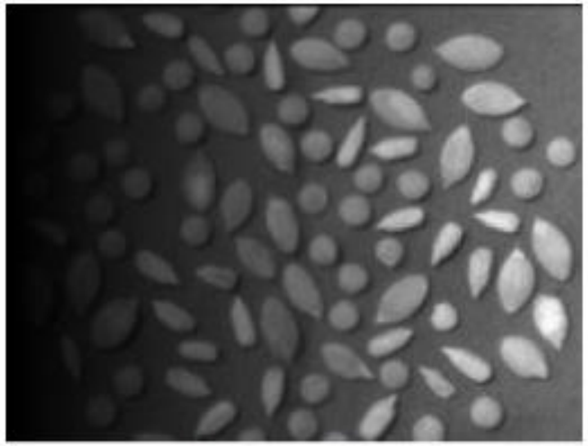
- Using image multiplication and division for shading correction and for masking.

a b c

FIGURE 2.33 Shading correction. (a) Shaded test pattern. (b) Estimated shading pattern. (c) Product of (a) by the reciprocal of (b). (See Section 3.5 for a discussion of how (b) was estimated.)

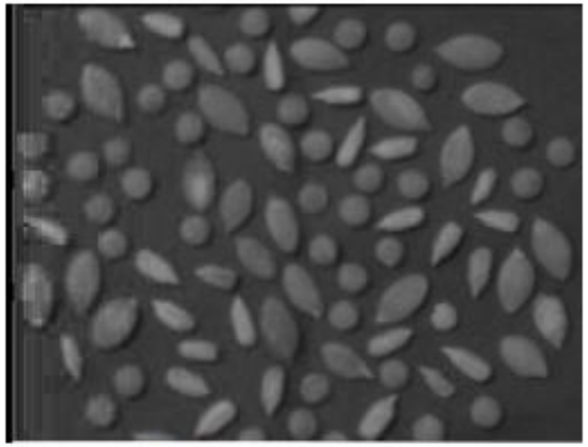
2.11 Image operation on a pixel basis

- Shading correction https://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction/index-en.html
  - Correction from a Bright Image



Input image.

Background image.



Output image.

Profile of the row 132.

Profile of the row 132.

Profile of the row 132.

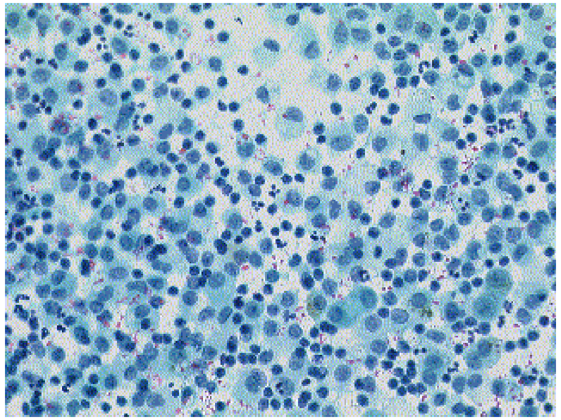
$$C = \mathrm{mean}(f(x,y)) \cdot \frac{1}{\mathrm{mean}\left(\frac{f(x,y)}{b(x,y)}\right)}$$

$$g(x,y) = \frac{f(x,y)}{b(x,y)} \cdot C$$

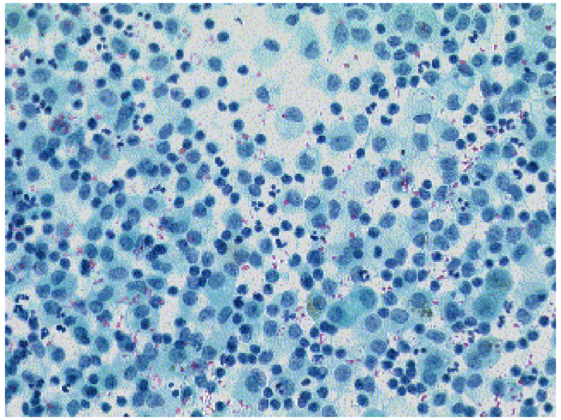
mean: mean filter

2.11 Image operation on a pixel basis

- Shading correction https://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction/index-en.html
    - Correction from a Dark Image

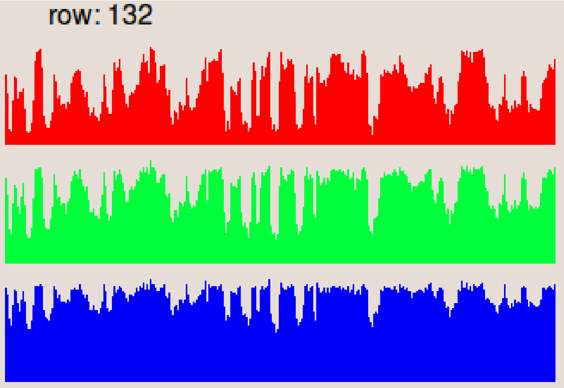

Input image.


Dark image.

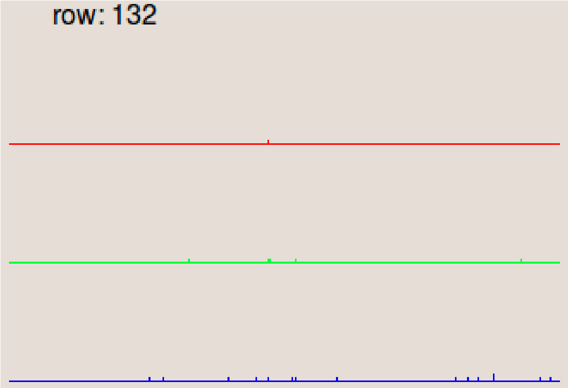

Output image.

g(x,y) = f(x,y) - d(x,y) + mean(d(x,y))

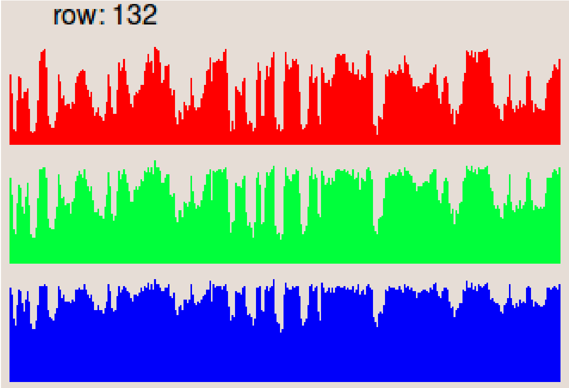
mean: mean filter



Profile of the row 132.



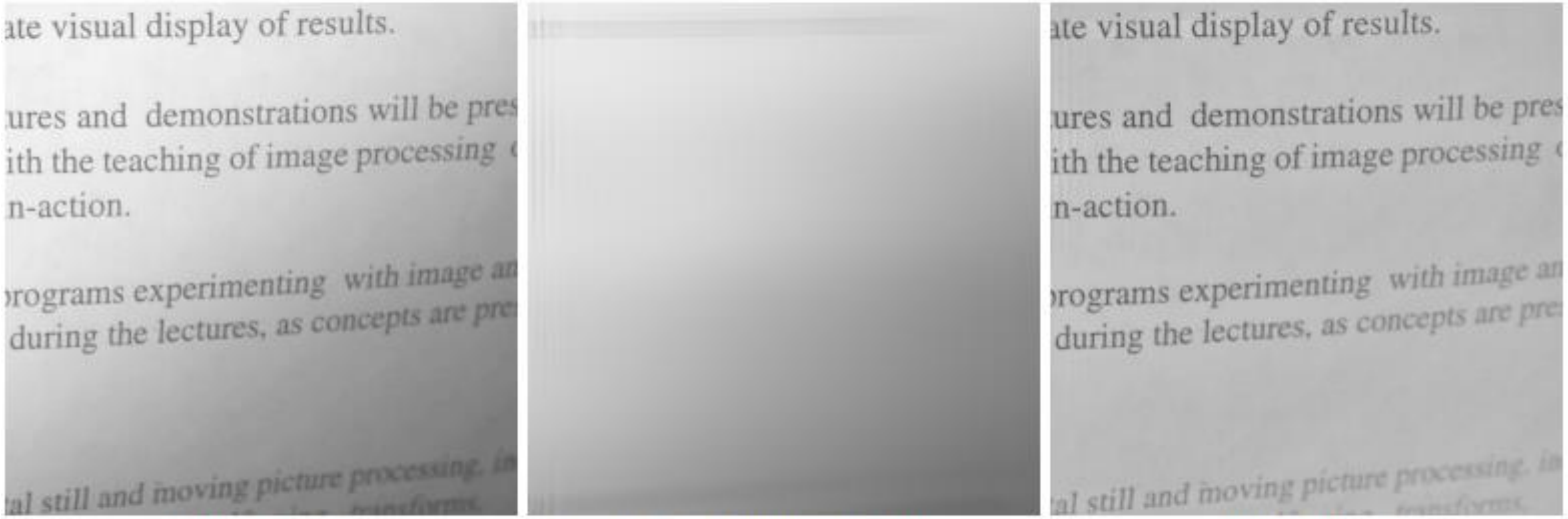
Profile of the row 132.



Profile of the row 132.

2.11 Image operation on a pixel basis

- Shading correction https://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction/index-en.html
    - Correction from Single Image: Low-pass Filtering



Input image. Estimated background image. Output image.

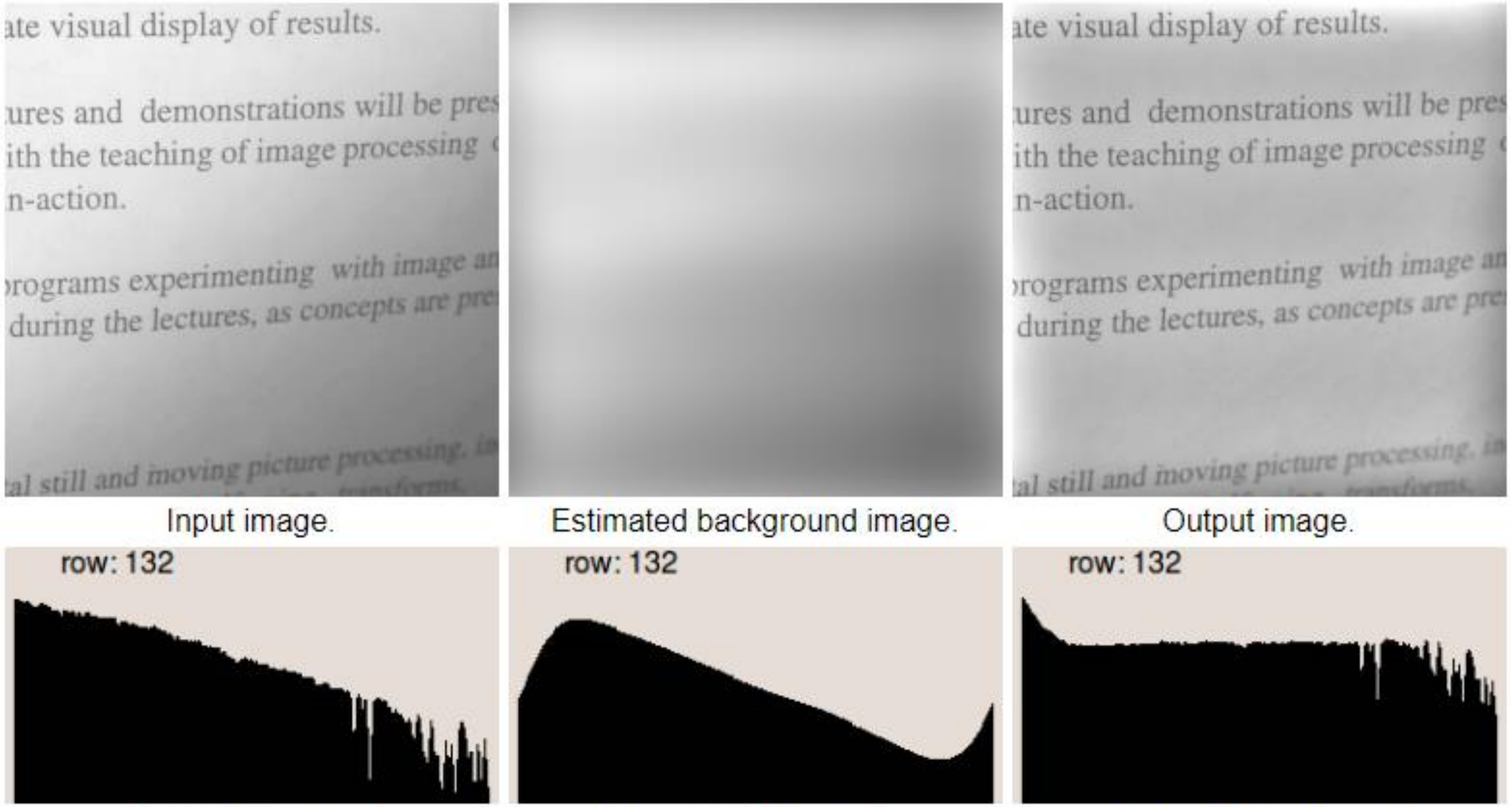
Profile of the row 132. Profile of the row 132. Profile of the row 132.

```
g(x,y) = f(x,y) - LPF(f(x,y)) + mean(LPF(f(x,y)))
```

LPF: low pass filter
mean: mean filter

2.11 Image operation on a pixel basis

- Shading correction https://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction/index-en.html
  - Correction from Single Image: Homomorphic Filtering



Input image. Estimated background image. Output image.

Profile of the row 132. Profile of the row 132. Profile of the row 132.

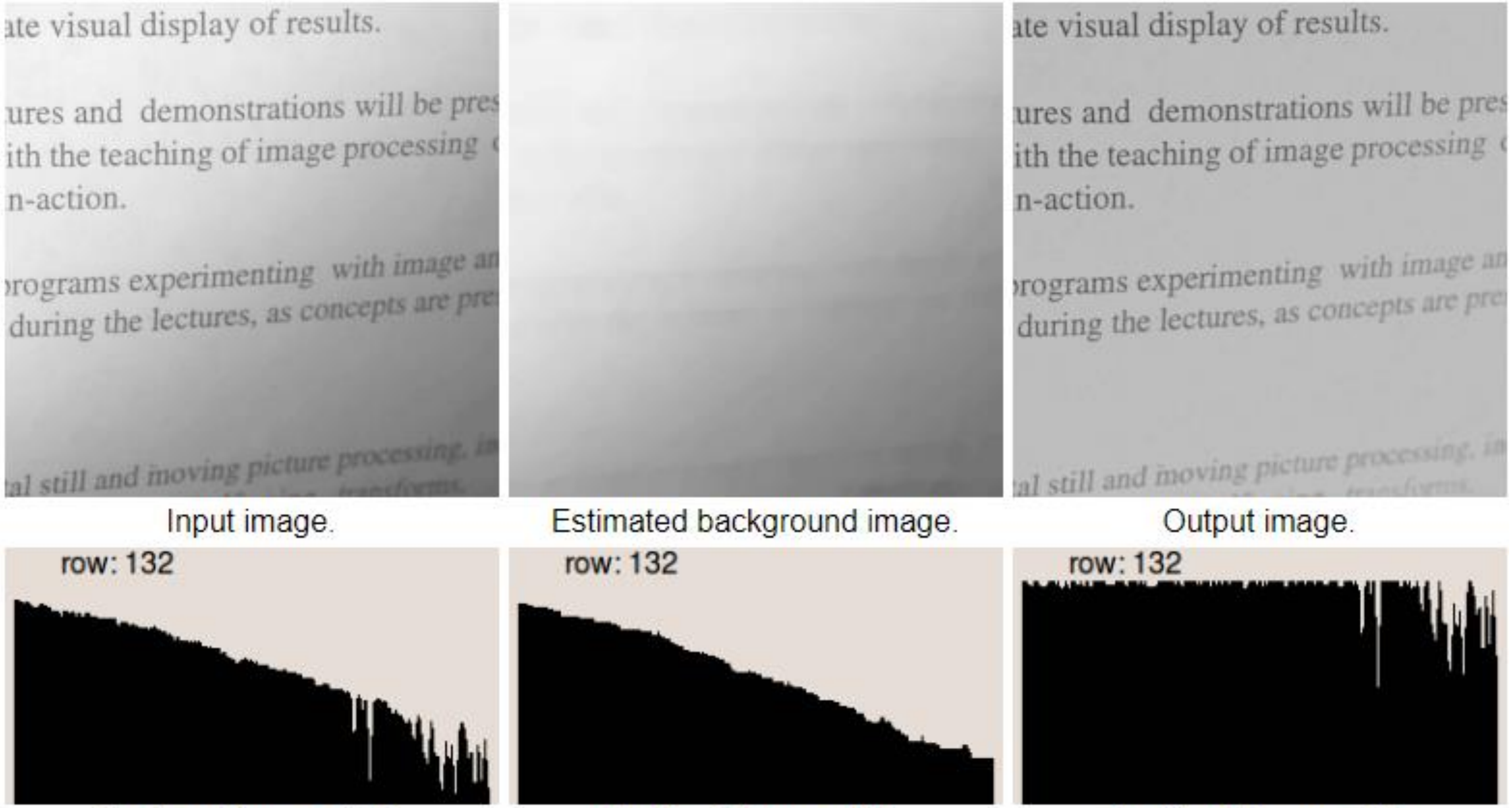
$$C = \frac{\text{mean}(f(x,y))}{\text{mean}\left(\dfrac{f(x,y)}{\exp(\text{LPF}(\log(f(x,y))))}\right)}$$

$$g(x,y) = \exp(\text{LPF}(\log(f(x,y)))) \cdot C$$

LPF: low pass filter
mean: mean filter

2.11 Image operation on a pixel basis

- Shading correction https://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction/index-en.html
  - Correction from Single Image: Morphological Filtering



Input image. Estimated background image. Output image.

Profile of the row 132. Profile of the row 132. Profile of the row 132.

```
g(x,y) = BTH[f(x,y)] + mean(closing(f(x,y)))
g(x,y) = [f(x,y) - closing(f(x,y))] + mean(closing(f(x,y)))
```

BTH: Black Top Hat
closing: morphological closing

2.11 Image operation on a pixel basis

- Image Averaging:
  - Consider a noisy image $g(x, y)$ formed by the addition of noise eta $\eta(x, y)$ to an original image $f(x, y)$:

    $$g(x,y) = f(x,y) + \eta(x,y)$$

    - where the assumption is that at every pair of coordinates (x, y) the noise is uncorrelated* and has zero average value.
  - If an image $\bar{g}(x,y)$ is formed by averaging K different noisy images

    $$\bar{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

    then it follows that $E\{\bar{g}(x,y)\} = f(x,y)$ and $\sigma^2_{\bar{g}(x,y)} = \frac{1}{K}\sigma^2_{\eta(x,y)}$

    - where $E\{\bar{g}(x,y)\}$ is the expected value of $\bar{g}$, and $\sigma^2_{\bar{g}(x,y)}$ and $\sigma^2_{\eta(x,y)}$ are the variance of $\bar{g}$ and η.

*:The covariance of two random variables x_i and x_j is defined as $E[(x_i-m_i)(x_j-m_j)]$.
If the variables are uncorrelated, their covariance is 0.

2.11 Image operation on a pixel basis

- Image Averaging:
  - The standard deviation at any point in the average image is

$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{K}} \sigma_{\eta(x,y)}$$

  - As *K* increases, equation above indicate that the variability (noise) of the pixel values at each location (*x*, *y*) decreases.

2.11 Image operation on a pixel basis

- Image Averaging:

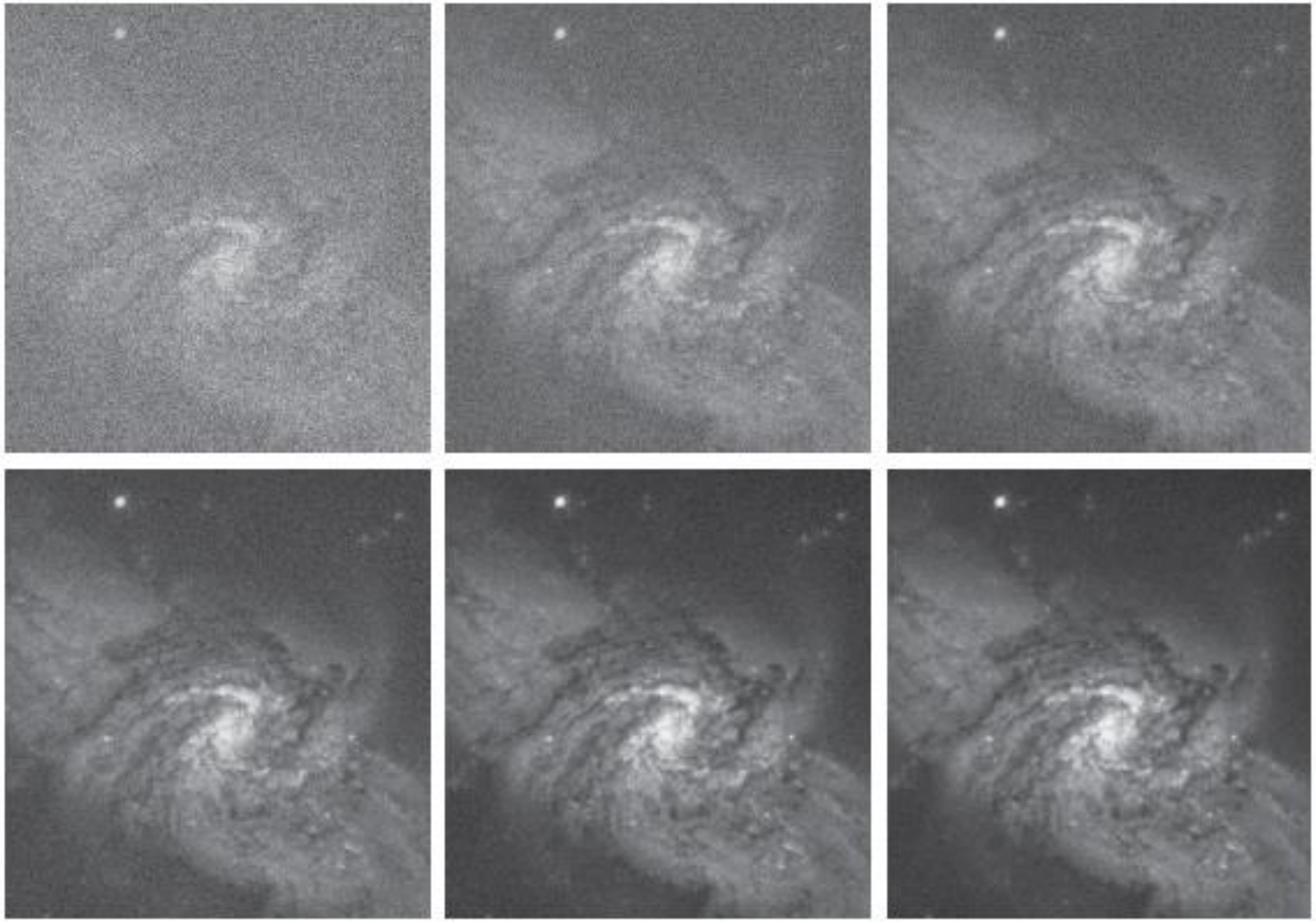


FIGURE 2.29 (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)-(f) Result of averaging 5, 10, 20, 50, and 1,00 noisy images, respectively. All images are of size 566×598 pixels, and all were scaled so that their intensities would span the full [0, 255] intensity scale. (Original image courtesy of NASA.)

2.11 Image operation on a pixel basis

- Logic Operation:

a	b	a AND b	a OR b	NOT(a)
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0

2.11 Image operation on a pixel basis

- Logic Operation:

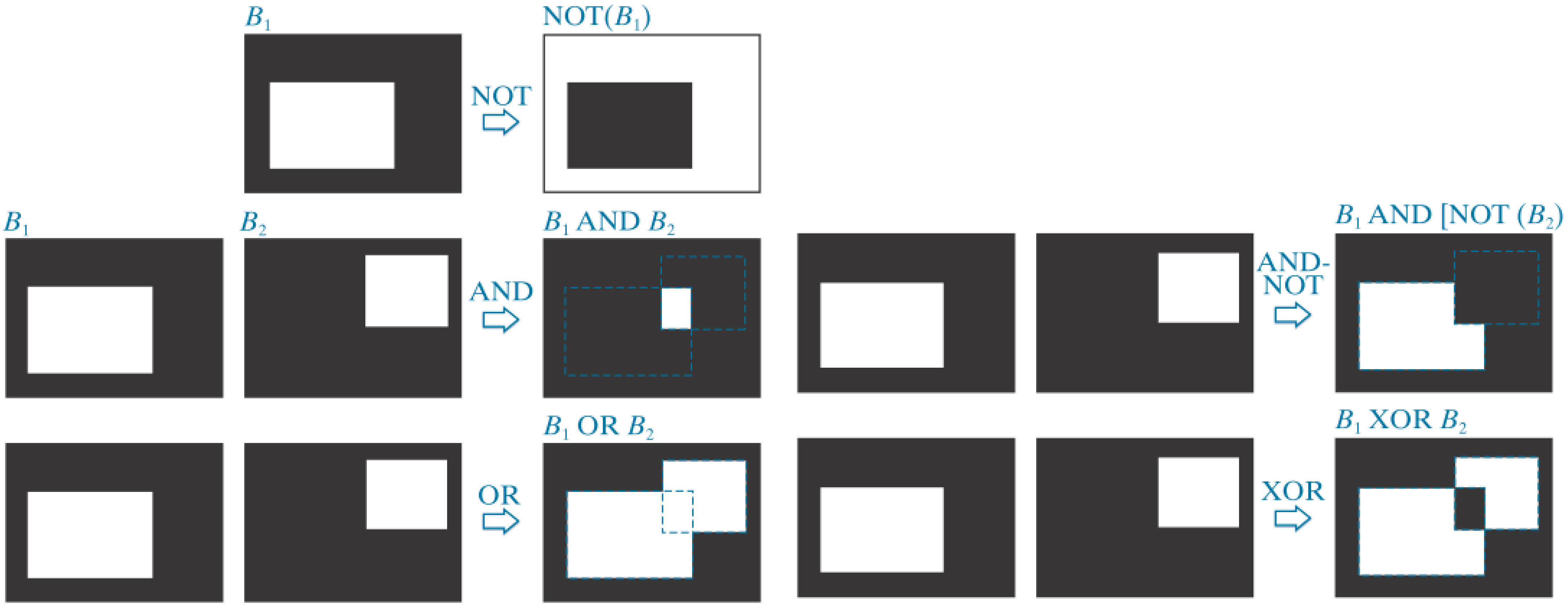


FIGURE 2.37 Illustration of logical operations involving foreground (white) pixels. Black represents binary 0's and white binary 1's. The dashed lines are shown for reference only. They are not part of the result.