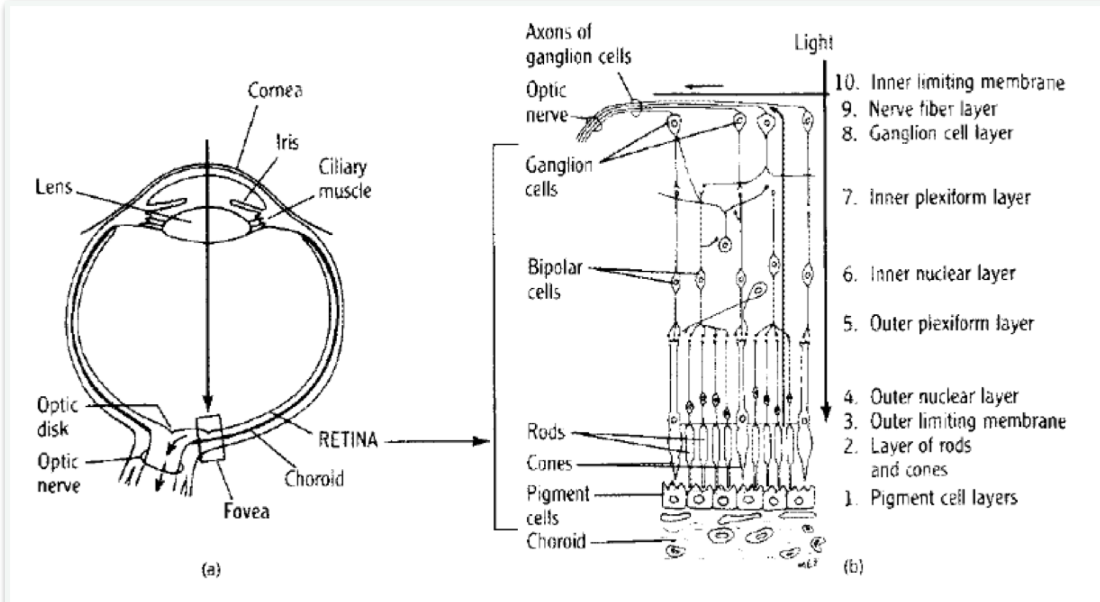
Course 2:

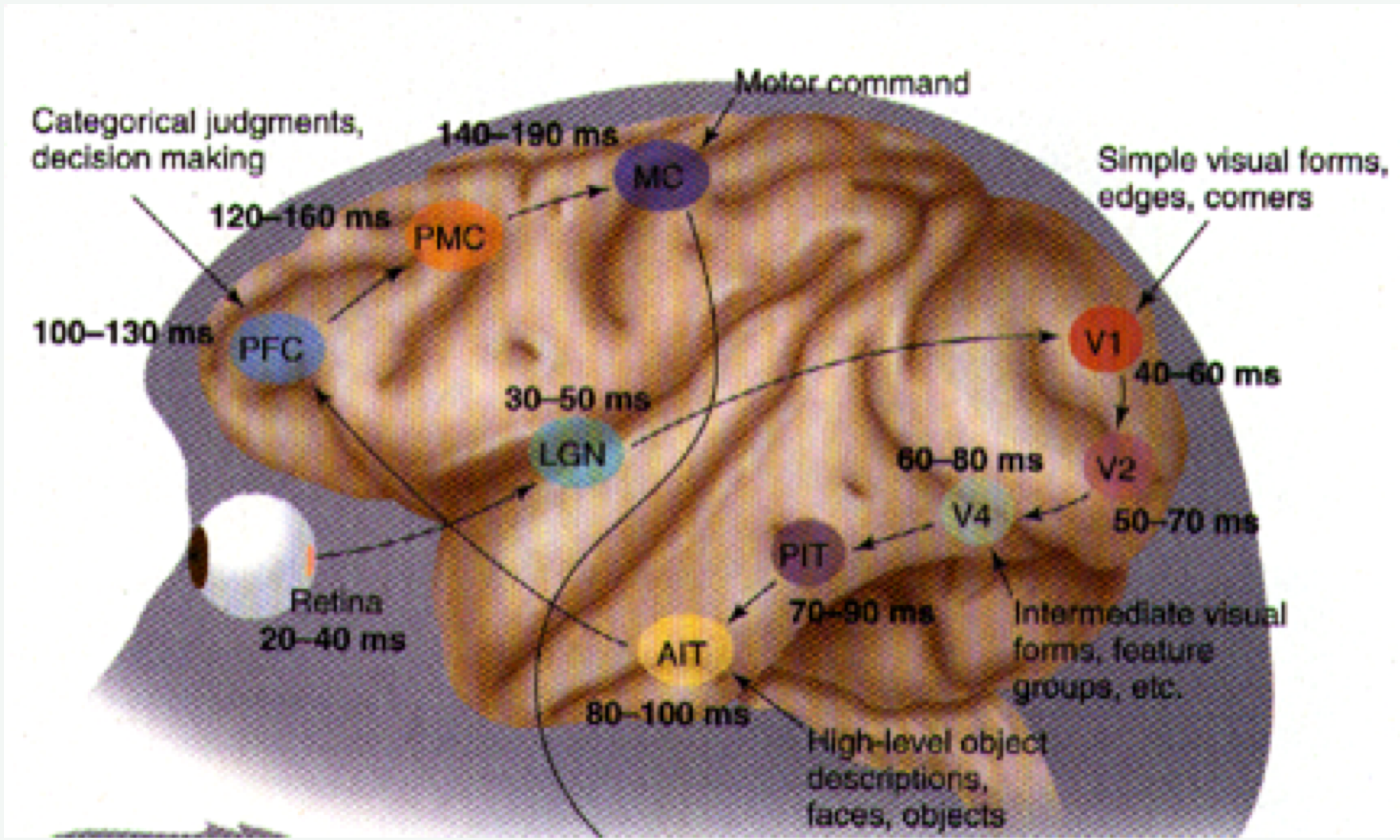
Fundamentals

Human visual system

- The eye's spatial resolution is about $0.01°$ over a $150°$ field of view (not evenly spaced, there is a fovea and a peripheral region).
- Intensity resolution is about 11 bits/element, spectral range is 400–700nm.
- Temporal resolution is about 100 ms (10 Hz).
- Two eyes give a data rate of about 3 GBytes/s!



- Retina measures about 5×5 cm and contains 10^8 sampling elements (rods and cones).

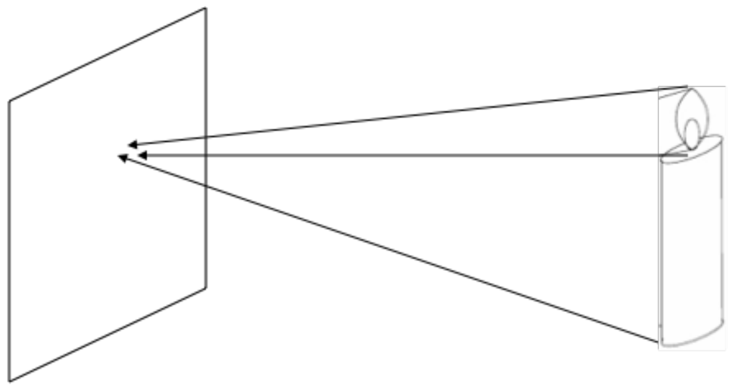


- Vision is the most powerful of our own senses.
- Around 1/3 of our brain is devoted to processing the signals from our eyes.
- The visual cortex has around $o(10^{11})$ neurons.

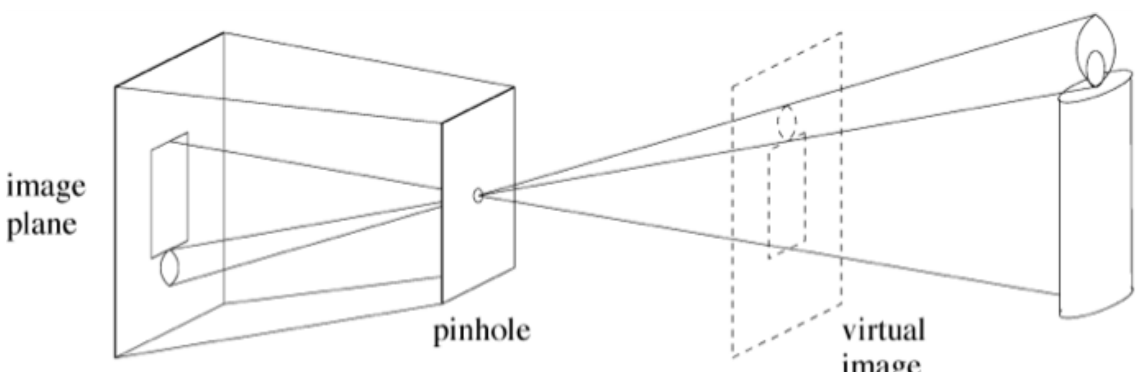
Image formation- camera model

- A camera is a mapping between the 3D world (object space) and a 2D image
- Find the projection equations

Pinhole camera

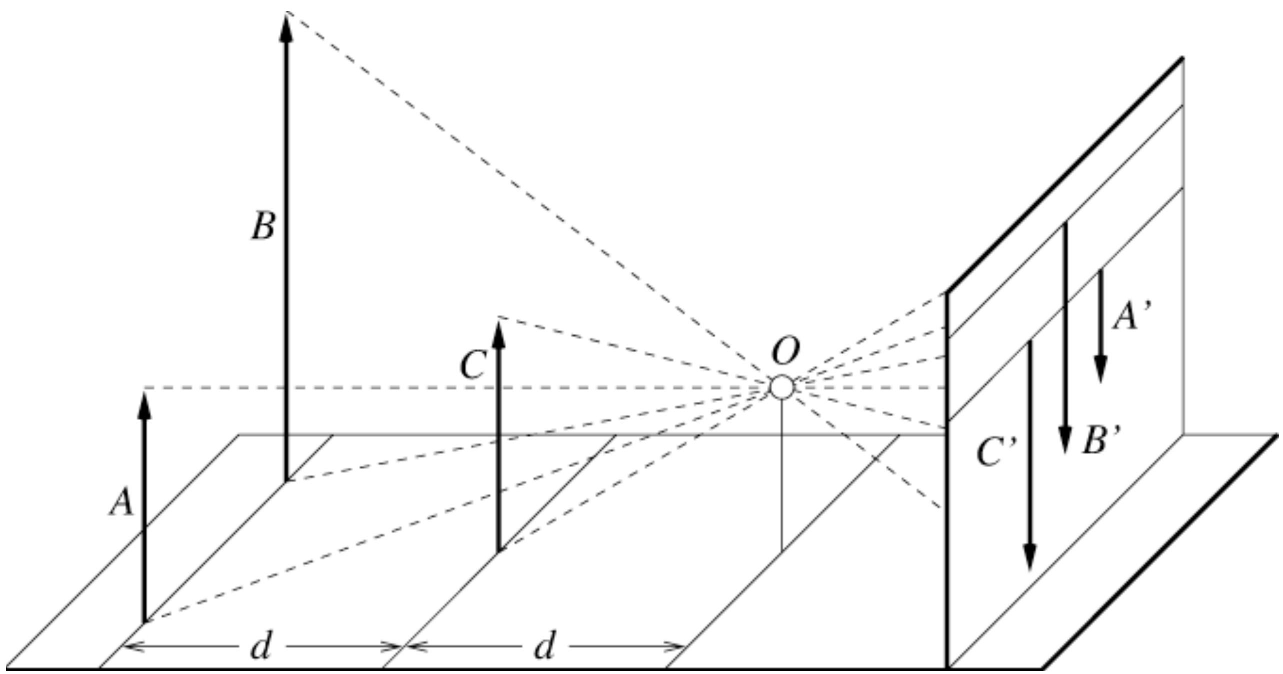


Light rays from many different parts of the scene strike the same point on the paper.



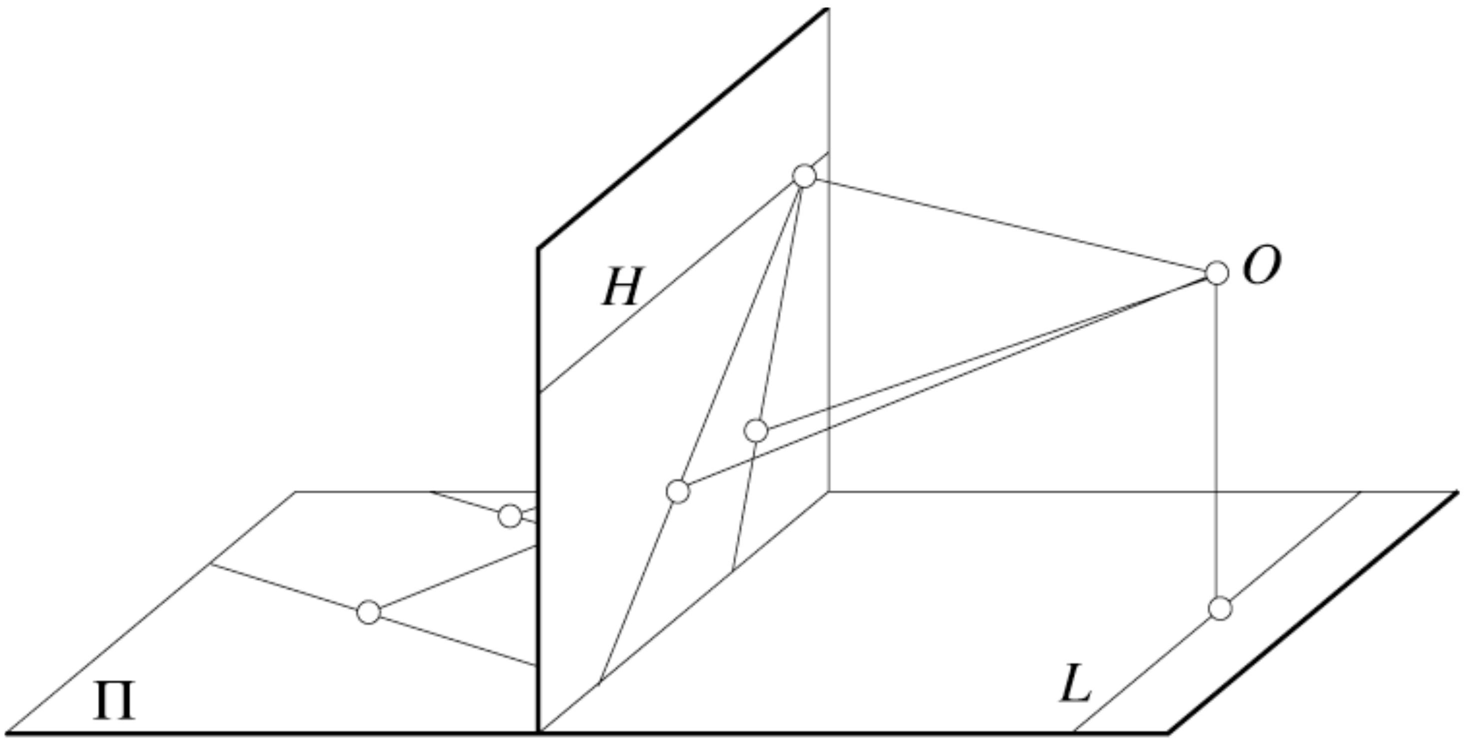
The pinhole camera only allows rays from one point in the scene to strike each point of the paper.

Perspective effect



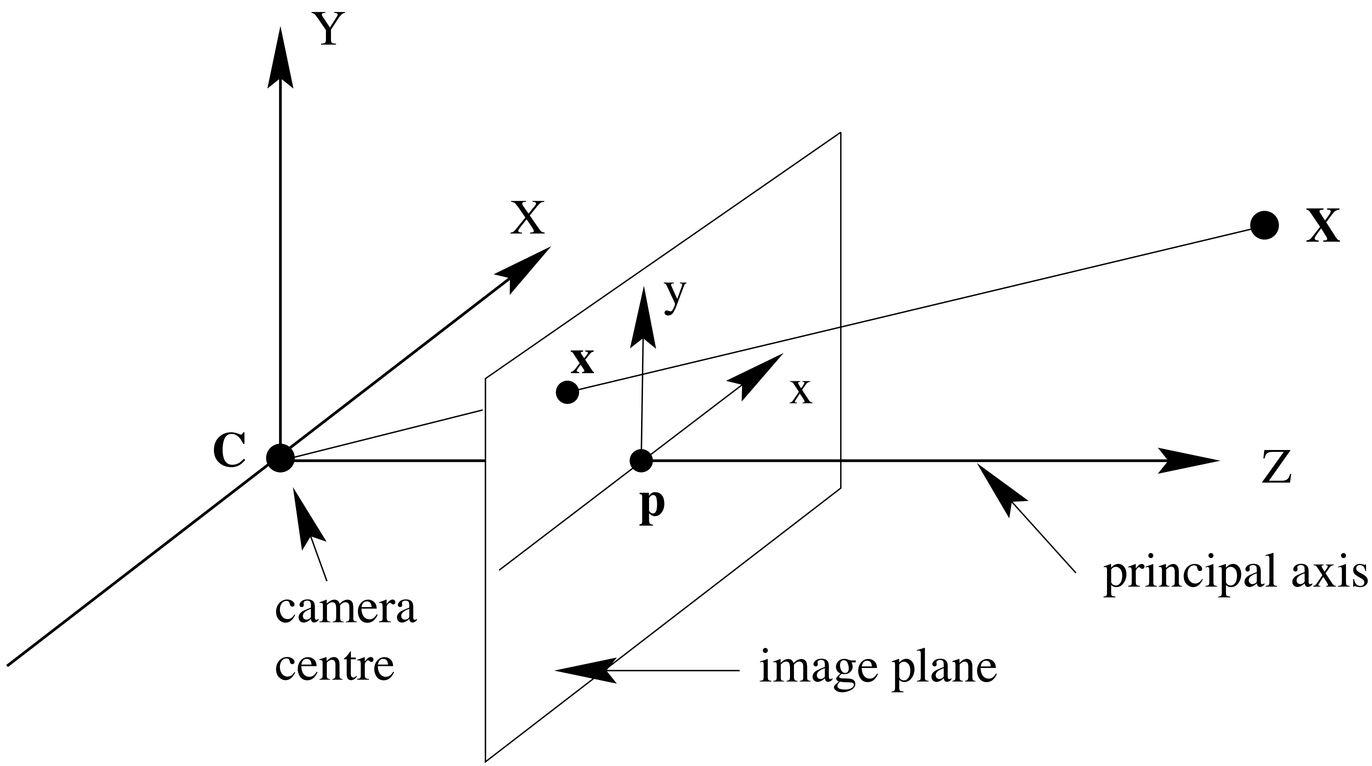
Forsyth & Ponce

Perspective effect



Forsyth & Ponce

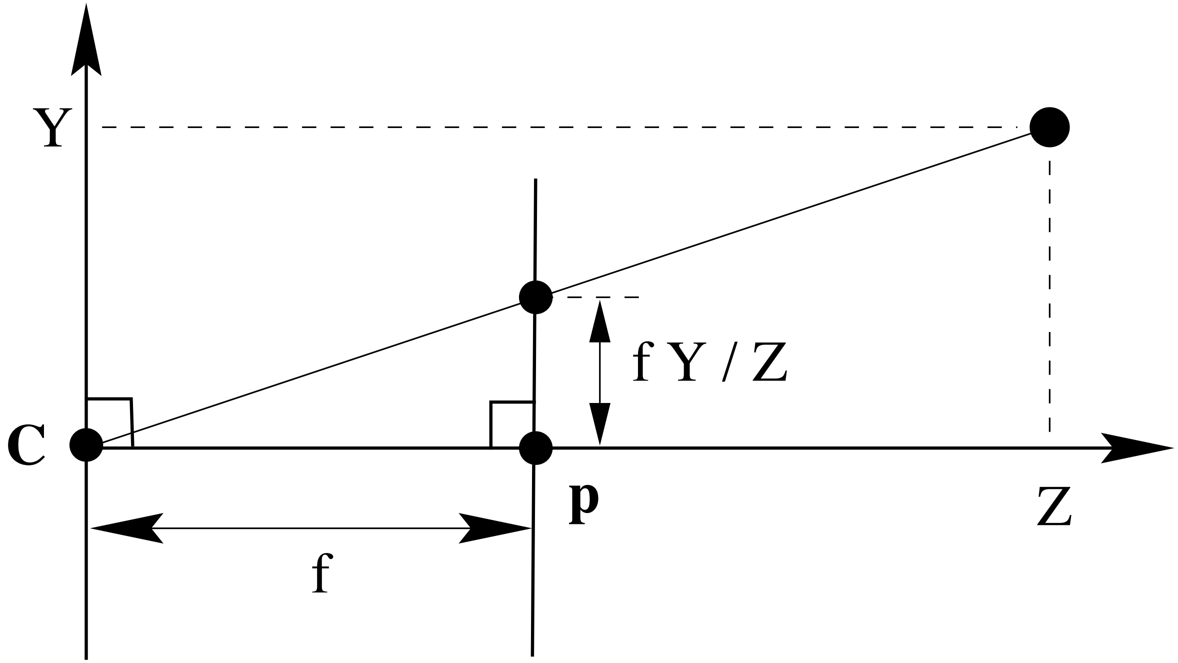
Pinhole model



Pinhole model

- Projection center (i.e., camera) is at the origin of scene coordinate system with optic axis of camera pointing to z-axis direction of coordinates.
- Origin of image coordinate system is located at (0, 0, f) of scene coordinate system (f is focus length of the camera), the x- and y-axes of image coordinates are parallel with X- and Y- axes of scene coordinates, respectively.
- Camera does not have any lens distortions.

Pinhole model



$$\boxed{(\mathrm{X}, \mathrm{Y}, \mathrm{Z})^{\mathsf{T}} \mapsto (f\mathrm{X}/\mathrm{Z}, f\mathrm{Y}/\mathrm{Z})^{\mathsf{T}}}$$

Image sensing and acquisition

- Different types of sensors

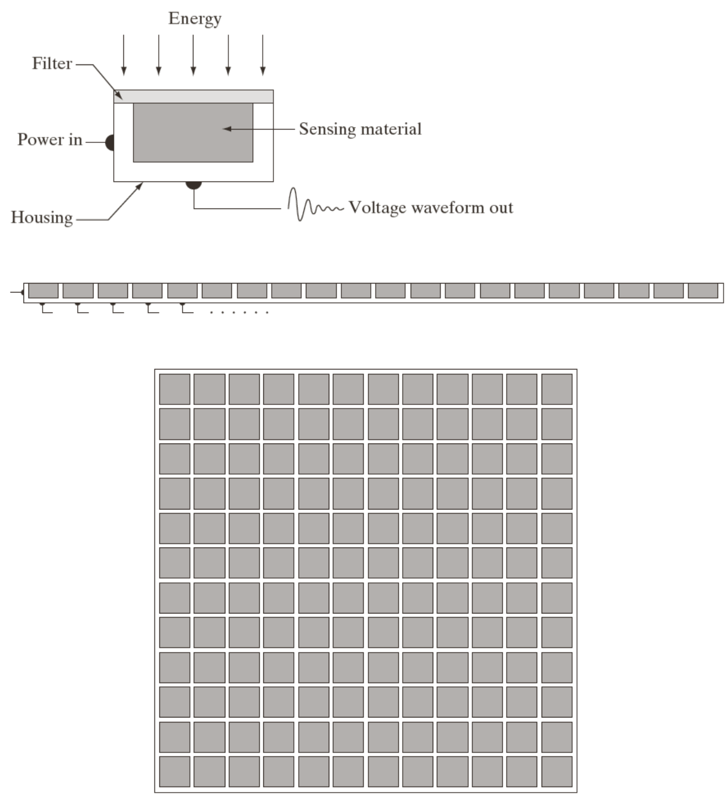


FIGURE 2.12
(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.

Single point sensor

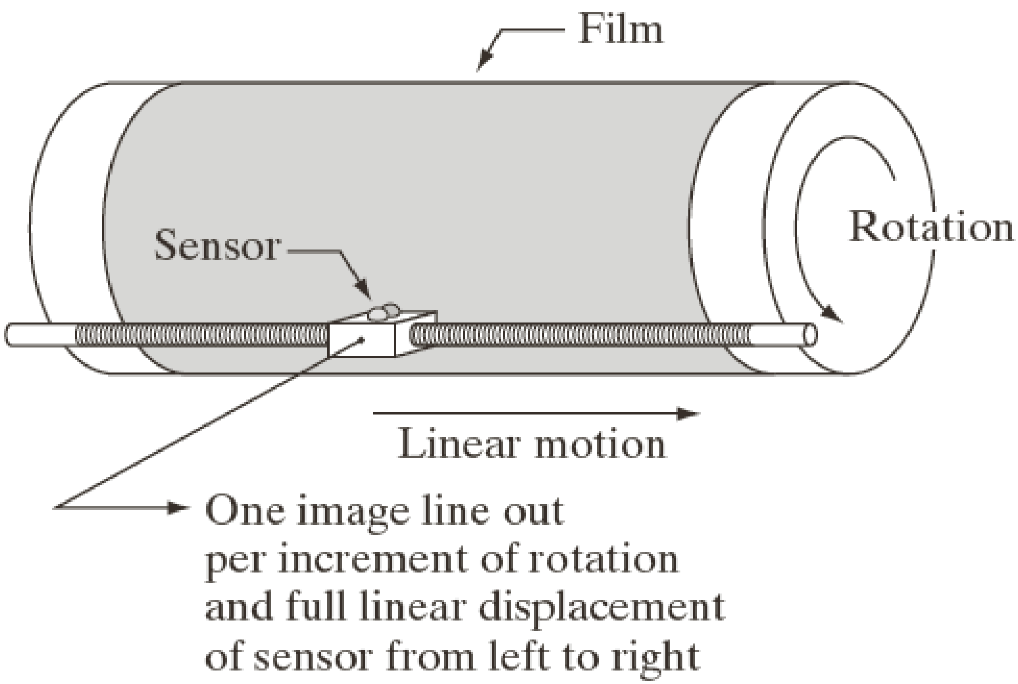
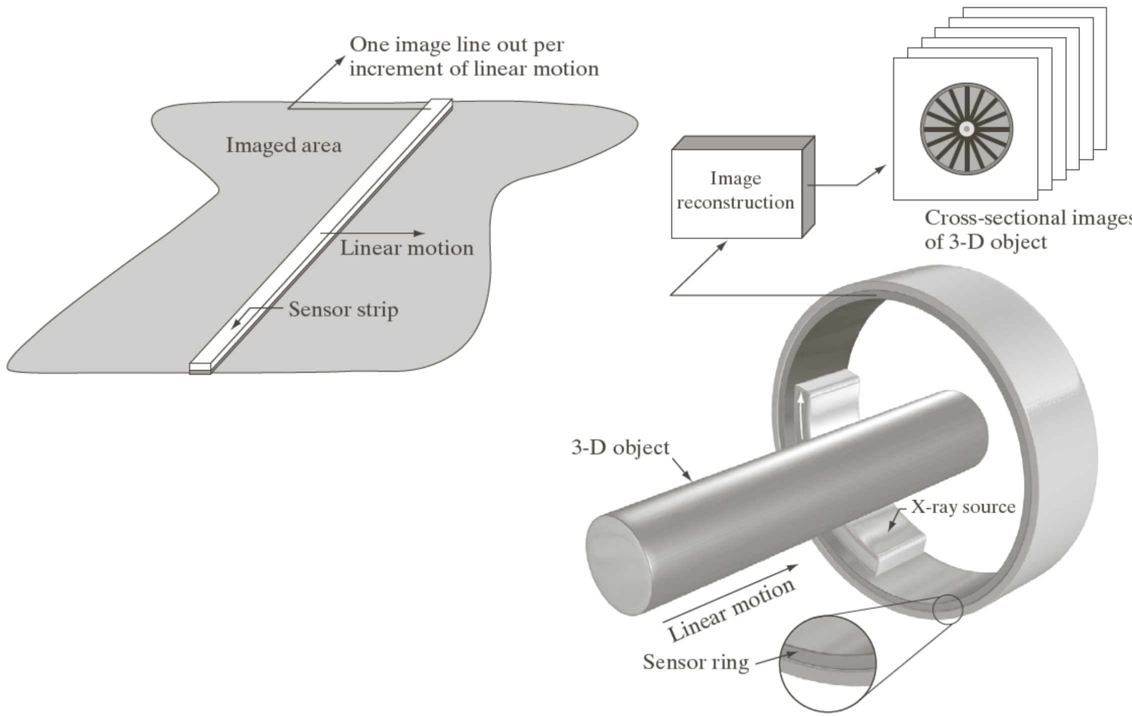


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Sensor strips



a b

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

Sensor arrays

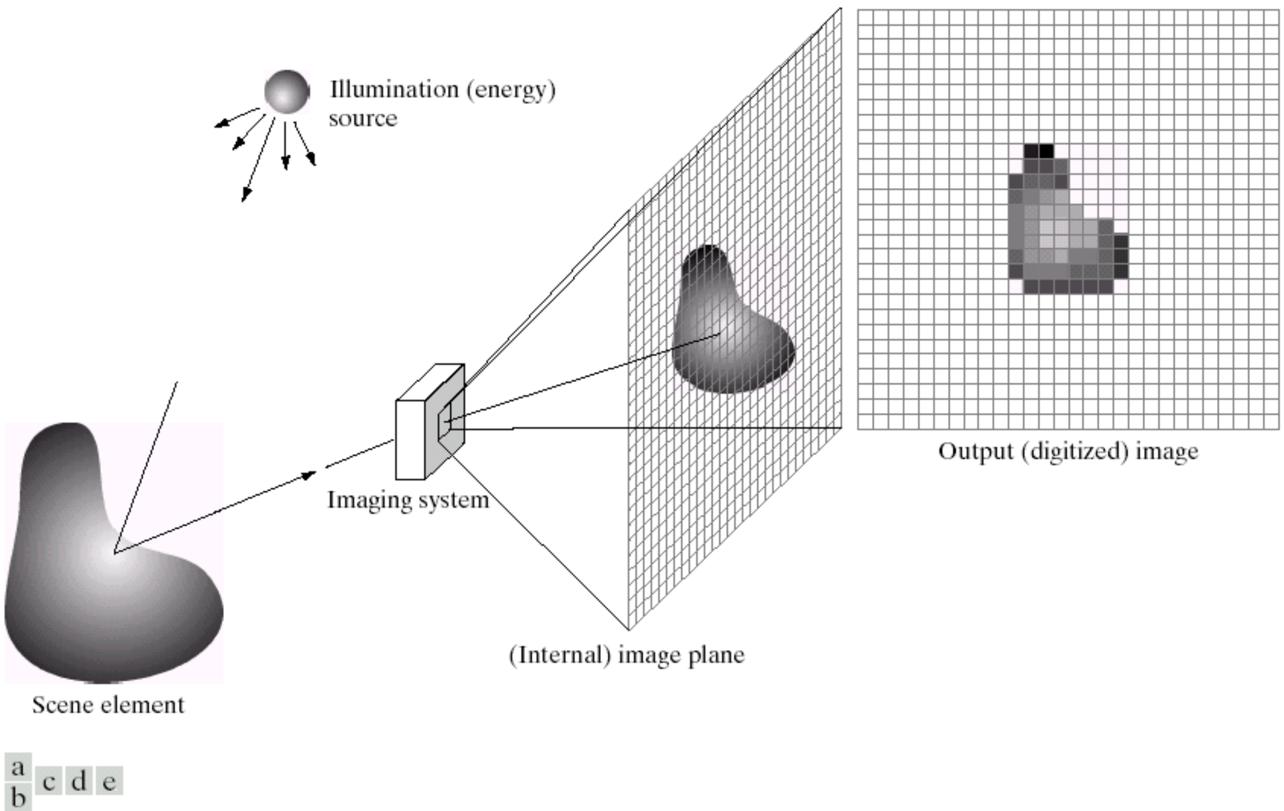


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Image sampling and quantization

- Digitizing the coordinate values is called *sampling*. Digitizing the amplitude values is called *quantization*.

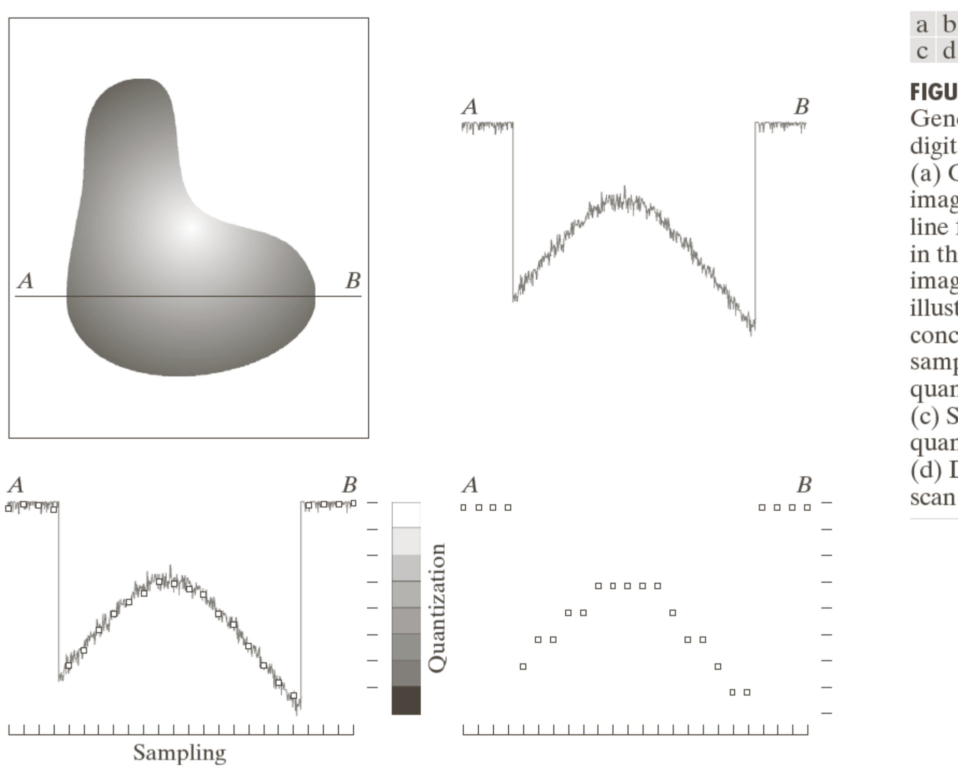
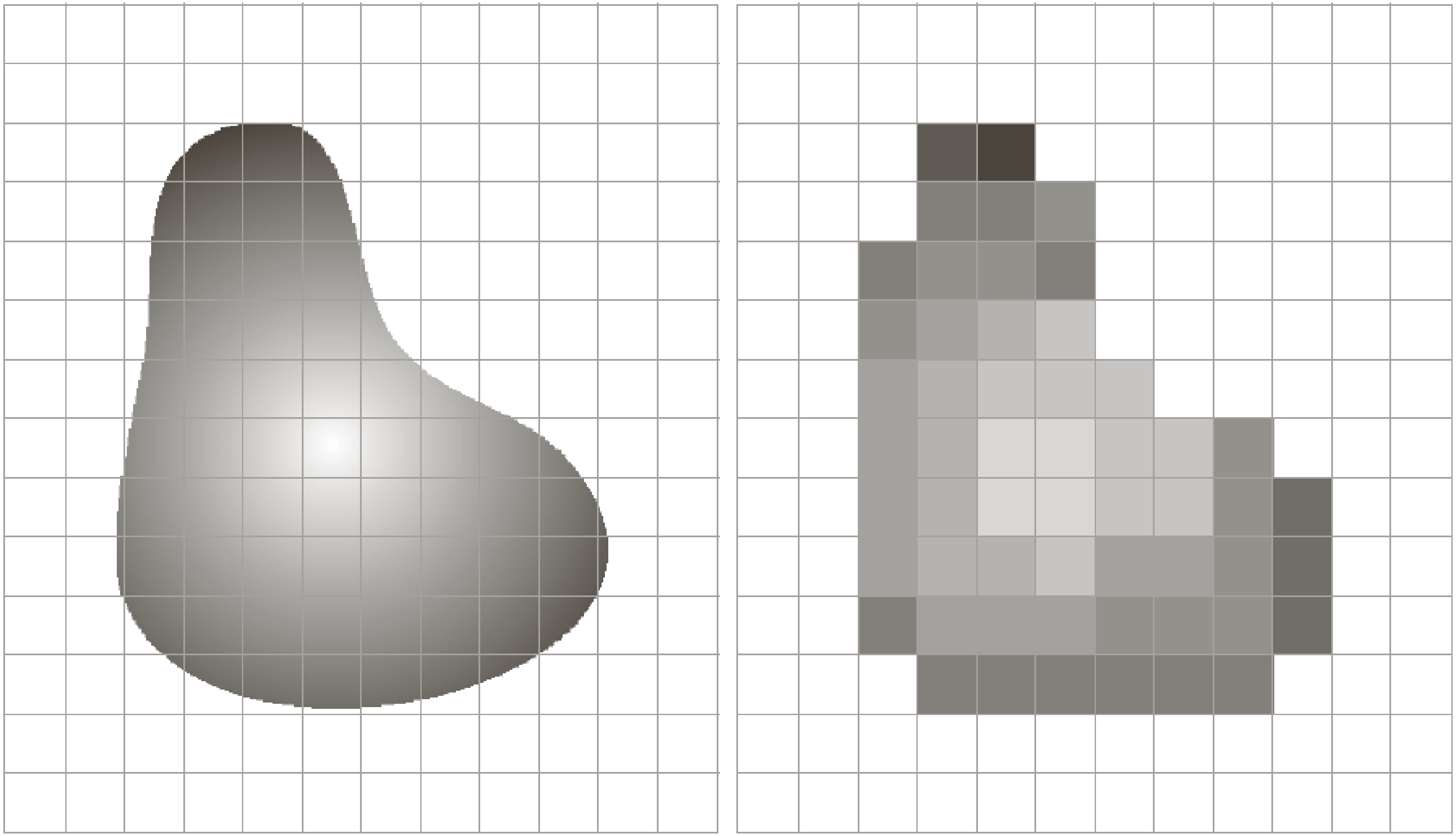


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.



a b

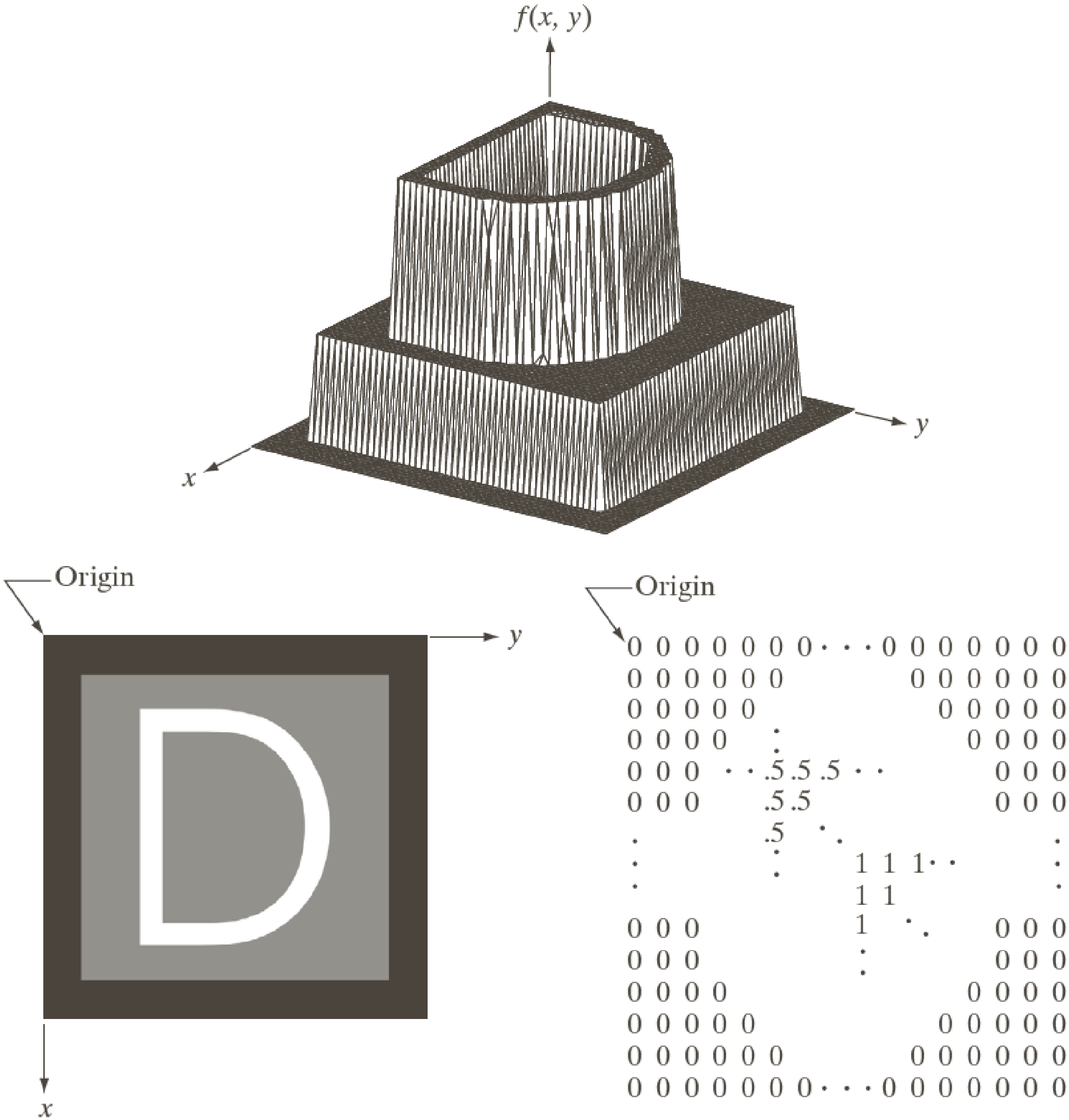
FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

Digitized image representation

- An image can be represented with a matrix:

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

- A digital image is a 2-D function whose coordinates and amplitude values are integers.



a
b c

FIGURE 2.18
(a) Image plotted as a surface.
(b) Image displayed as a visual intensity array.
(c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

Digitized image representation

- Due to storage and quantizing hardware considerations, the number of intensity levels typically is an integer power of 2:

$$L = 2^k$$

- The number, *b*, of bits required to store a digitized image is

$$b = M \times N \times k$$

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

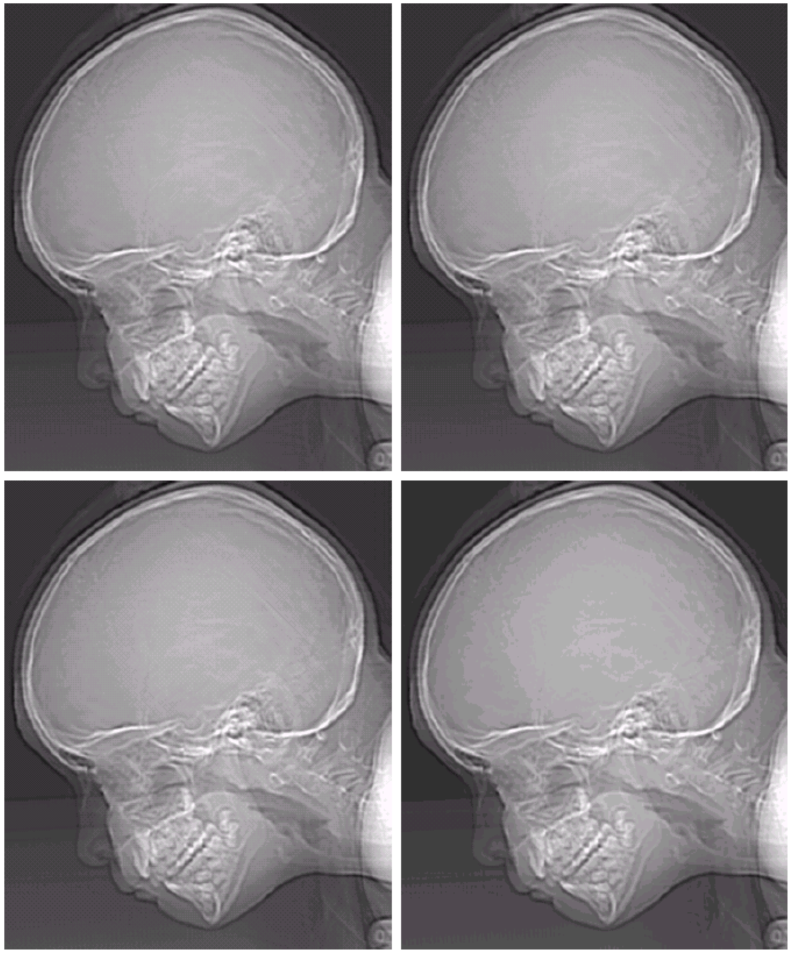
N/k	1 (L = 2)	2 (L = 4)	3 (L = 8)	4 (L = 16)	5 (L = 32)	6 (L = 64)	7 (L = 128)	8 (L = 256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680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

- Spatial resolution
  - *Line pairs per unit distance:* a widely used definition of image resolution is the largest number of *discernible* line pairs per unit distance
  - *dots* (*pixels*) *per unit distance: dots per inch* (dpi)
- Intensity resolution: it is common practice to refer to the number of bits used to quantize intensity as the *intensity resolution.*

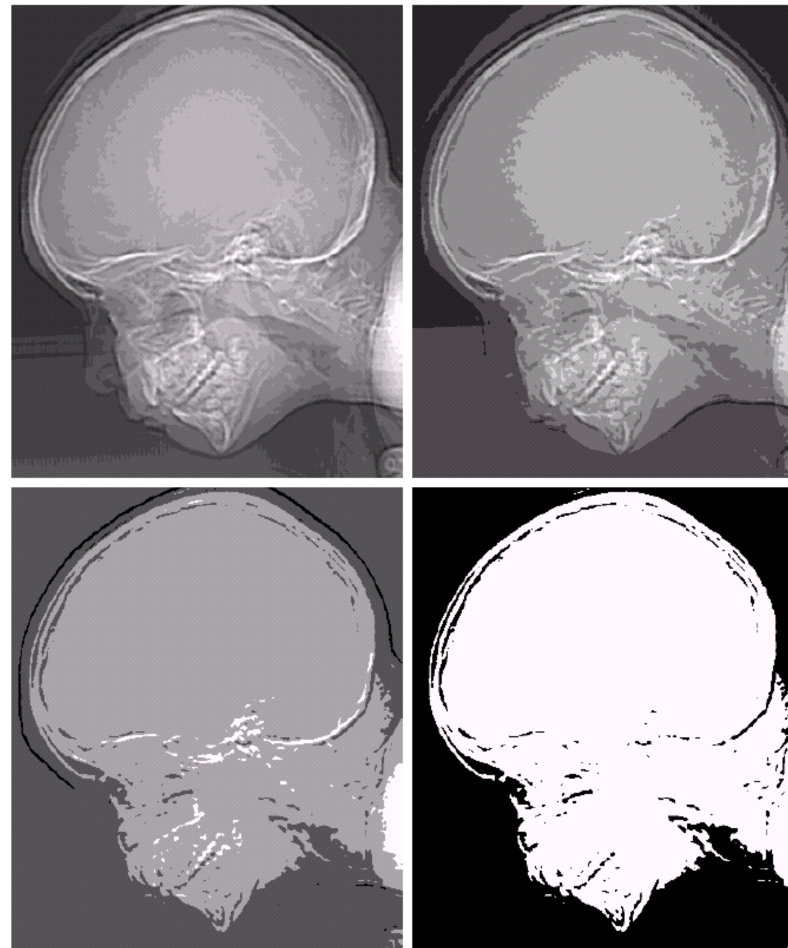
a b
c d

FIGURE 2.20 Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.



a b
c d

FIGURE 2.21
(a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.



e f
g h

FIGURE 2.21
(Continued)
(e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

false contouring

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

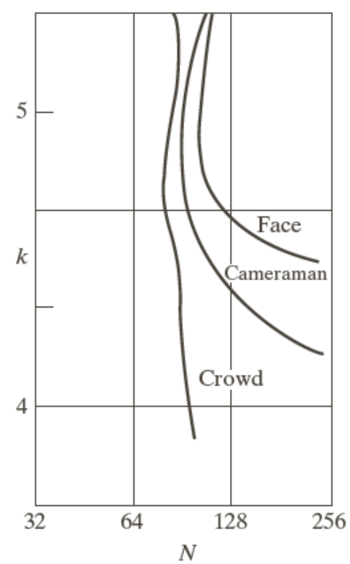


FIGURE 2.23 Typical isopreference curves for the three types of images in Fig. 2.22.

Interpolation

- *Interpolation* is the process of using known data to estimate values at unknown locations.
- Nearest neighbor
- Bilinear interpolation

$$v(x, y) = ax + by + cxy + d$$

- Cubic interpolation

$$v(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j$$

a b c
d e f

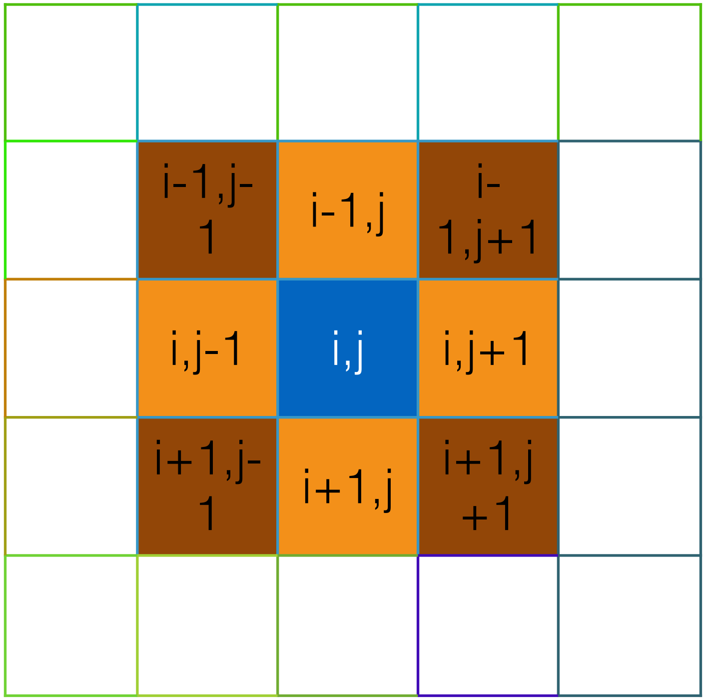
FIGURE 2.24 (a) Image reduced to 72 dpi and zoomed back to its original size (3692 × 2812 pixels) using nearest neighbor interpolation. This figure is the same as Fig. 2.20(d). (b) Image shrunk and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation. (d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. 2.24(d) is the same as Fig. 2.20(c)]. Compare Figs. 2.24(e) and (f), especially the latter, with the original image in Fig. 2.20(a).

Pixel relationships

- Neighbors of a pixel

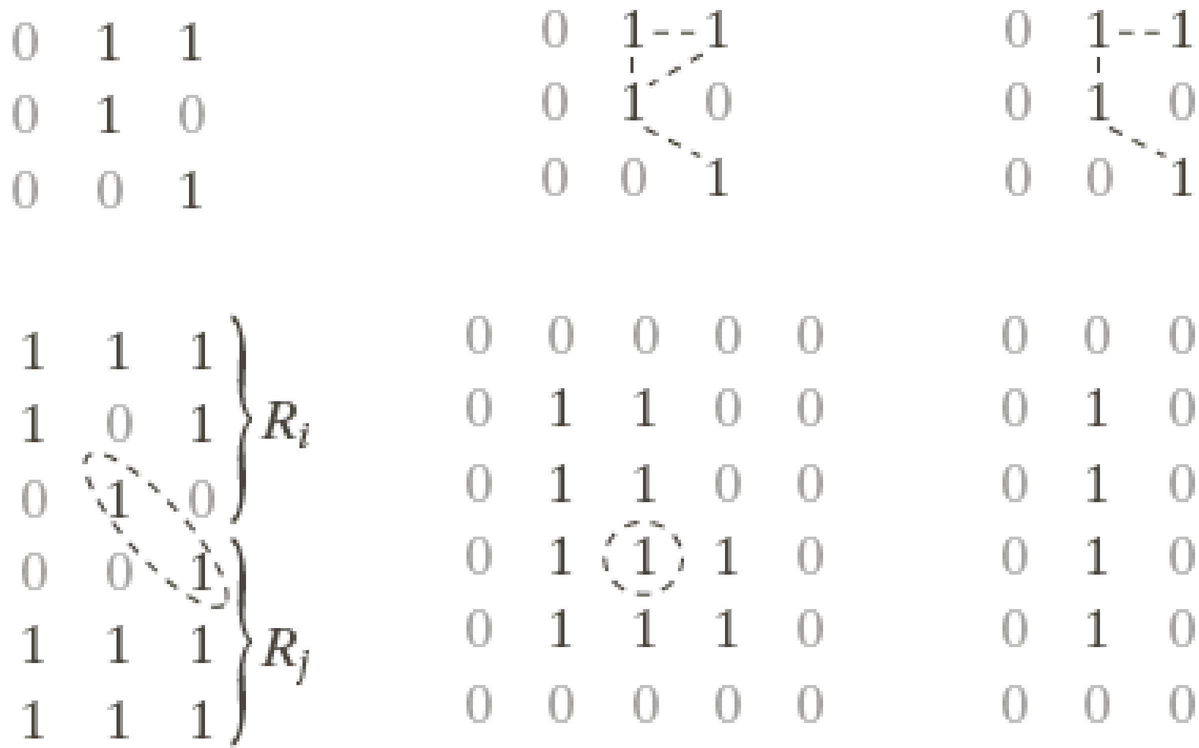
4-neighbours: the upper, lower, left and right neighbour pixels of a pixel.

8-neighbours: the 4-neighbours pixels plus the diagonal neighbors.



Several definitions

- Adjacency: 4-, 8-, and m-adjacency
- Path: 4-, 8-, or m-paths
- Connected component
- Region
- Boundary



a	b	c
d	e	f

FIGURE 2.25 (a) An arrangement of pixels. (b) Pixels that are 8-adjacent (adjacency is shown by dashed lines; note the ambiguity). (c) m-adjacency. (d) Two regions that are adjacent if 8-adjecency is used. (e) The circled point is part of the boundary of the 1-valued pixels only if 8-adjacency between the region and background is used. (f) The inner boundary of the 1-valued region does not form a closed path, but its outer boundary does.

Distance measures

- For pixels *p*, *q*, and *z*, with coordinates (*x*, *y*), (*s*, *t*), and (v, w), respectively, *D* is a *distance function* or *metric* if
  ① D(p,q) >= 0 (D(p,q)=0 iff p=q)
  ② D(p, q) = D(q, p), and
  ③ D(p, z) <= D(p, q) + D(q, z).
- *Euclidean distance* between *p* and *q* is defined as

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

- The D_4 distance (*city-block distance*)
- The D_8 distance (*chessboard distance*)

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

D_4

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

D_8

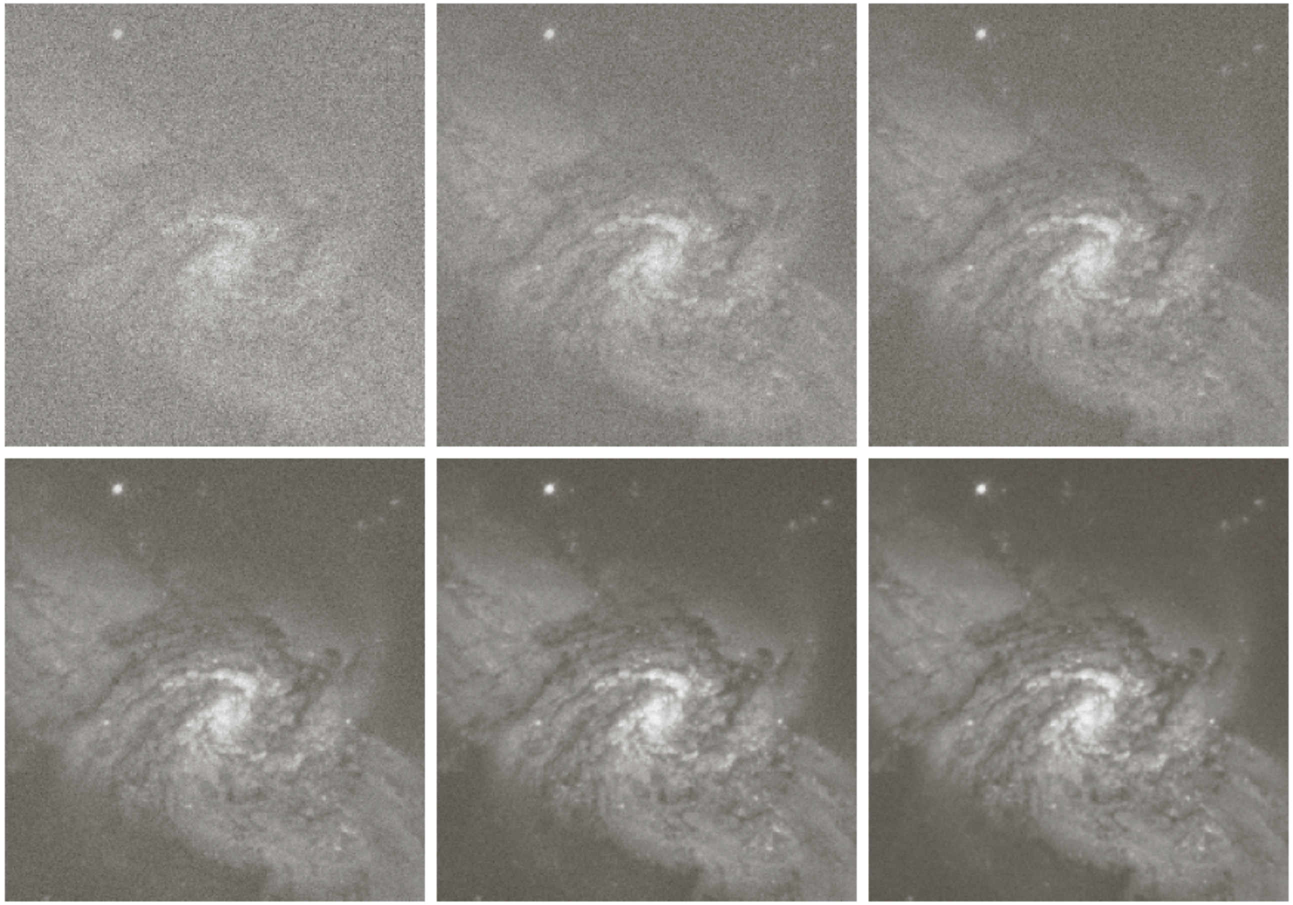
Mathematical tools

- Linear operations:

$$
\begin{aligned}
H\left[a_i f_i(x, y) + a_j f_j(x, y)\right] &= a_i H\left[f_i(x, y)\right] + a_j H\left[f_j(x, y)\right] \\
&= a_i g_i(x, y) + a_j g_j(x, y)
\end{aligned}
$$

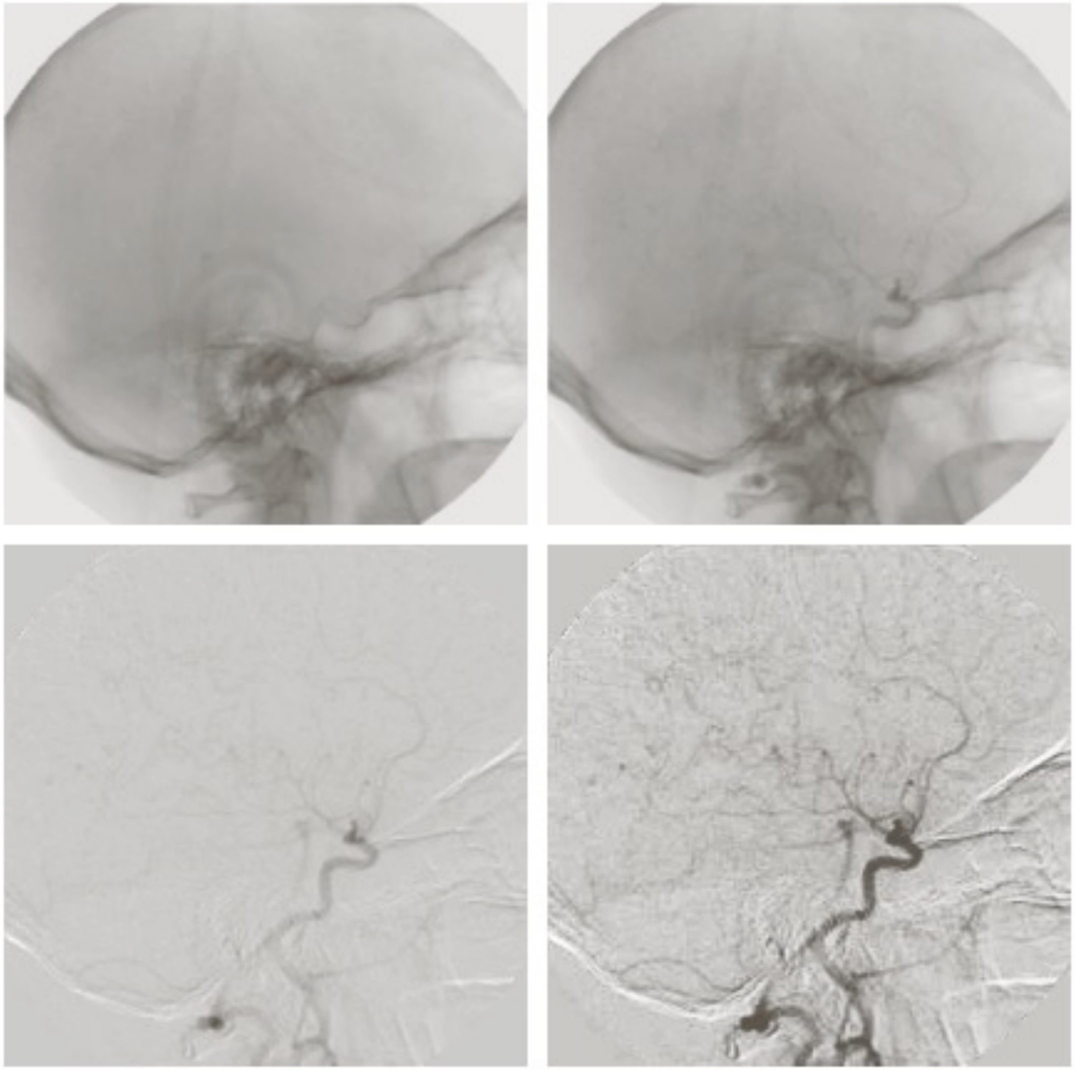
- Arithmetic operations

$$
\begin{aligned}
s(x, y) &= f(x, y) + g(x, y) \\
d(x, y) &= f(x, y) - g(x, y) \\
p(x, y) &= f(x, y) \times g(x, y) \\
v(x, y) &= f(x, y) \div g(x, y)
\end{aligned}
$$



a	b	c
d	e	f

FIGURE 2.26 (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)–(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)



a b
c d

FIGURE 2.28 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.)

a b c

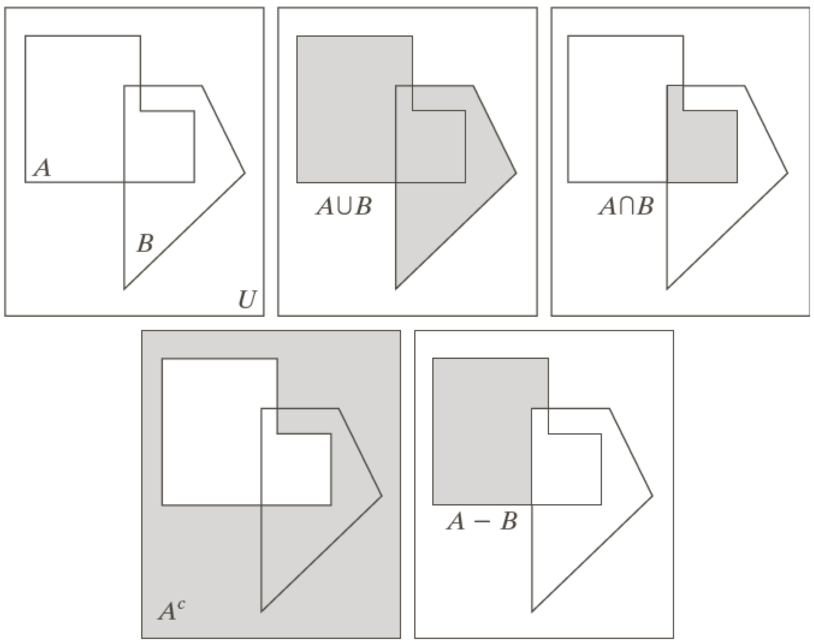
FIGURE 2.29 Shading correction. (a) Shaded SEM image of a tungsten filament and support, magnified approximately 130 times. (b) The shading pattern. (c) Product of (a) by the reciprocal of (b). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

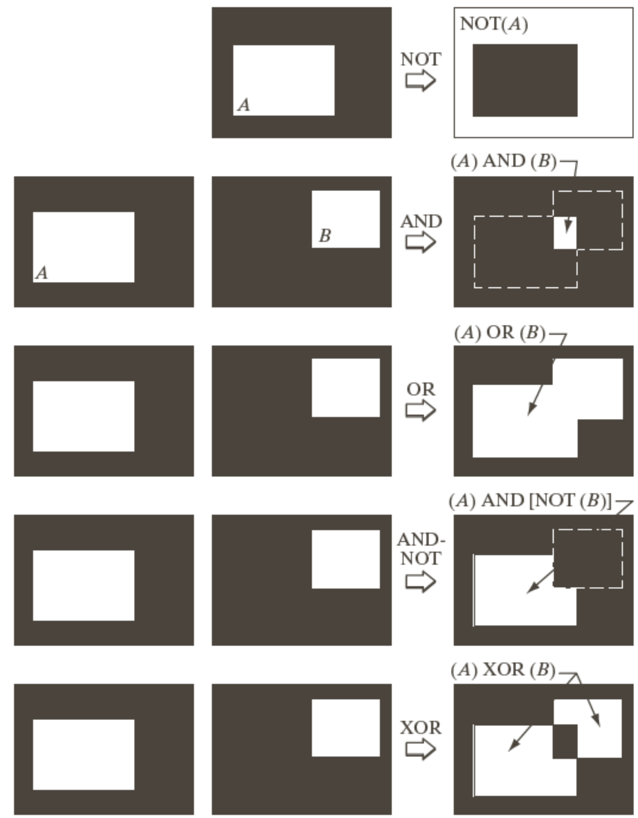
a b c

FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

Mathematical tools

- Set and Logical Operations





Mathematical tools

- Geometric transformations
  ① Isometries are transformations that preserve Euclidean distance. Invariants: length (the distance between two points), angle (the angle between two lines), and area

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \epsilon\cos\theta & -\sin\theta & t_x \\ \epsilon\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

where $\epsilon = \pm 1$.

$$\mathbf{x}' = \mathrm{H}_{\mathrm{E}}\mathbf{x} = \begin{bmatrix} \mathrm{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \mathbf{x}$$

Mathematical tools

- Geometric transformations
  ② Similarity is an isometry composed with an isotropic scaling. Invariants: angle (the angle between two lines), and area

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\mathbf{x}' = \mathrm{H_S}\mathbf{x} = \begin{bmatrix} s\mathrm{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \mathbf{x}$$

Mathematical tools

- Geometric transformations
  - ③ Affine is a non-singular linear transformation followed by a translation. Invariants: parallel lines; ratio of lengths of parallel line segments; ratio of area

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

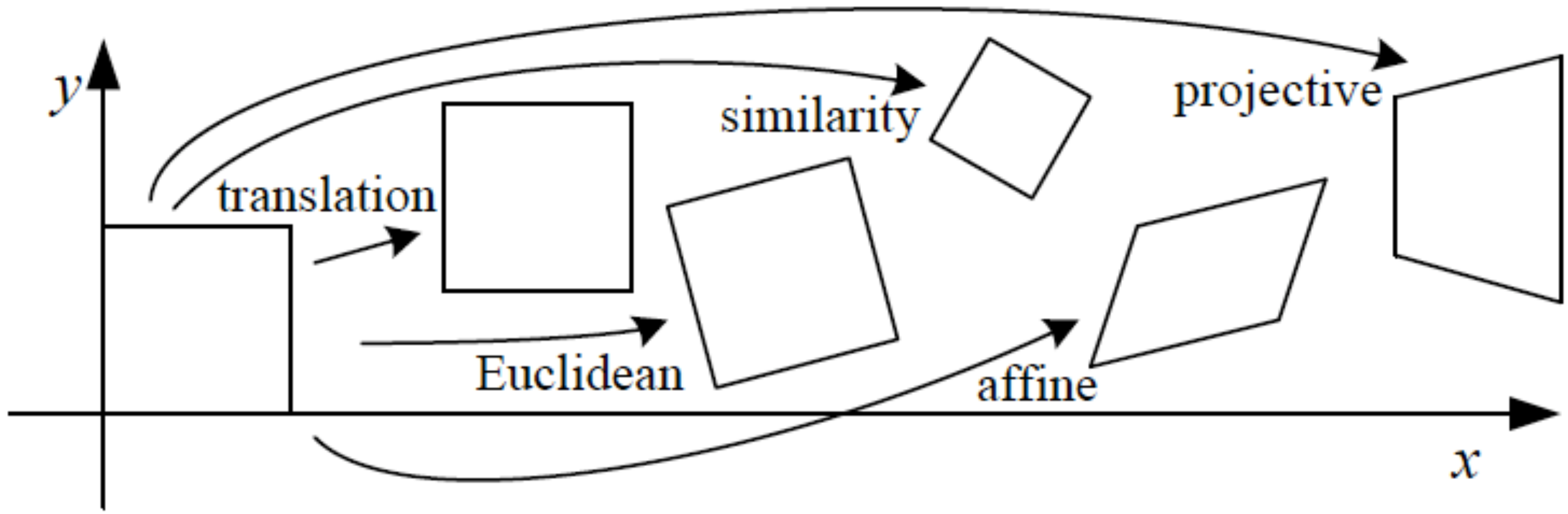
$$\mathbf{x}' = \mathrm{H}_{\mathrm{A}}\mathbf{x} = \begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \mathbf{x}$$

Mathematical tools

- Geometric transformations
  ④ Projective (Homogragy) is a general non-singular linear transformation of homogeneous coordinates. Invariants: cross ratio of four collinear points

$$\mathbf{x}' = \mathrm{H}_{\mathrm{P}}\mathbf{x} = \begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{v}^{\mathsf{T}} & v \end{bmatrix} \mathbf{x}$$

$$\mathbf{v} = (v_1, v_2)^{\mathsf{T}}$$



Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	