Chapter 2: Digital Image Fundamentals
Structure of the Human Eye

The *iris* acts as a diaphragm to control the amount of light entering the eye.

Front of eye is covered by a transparent surface called the

Light entering the cornea is focused to the retina surface by a

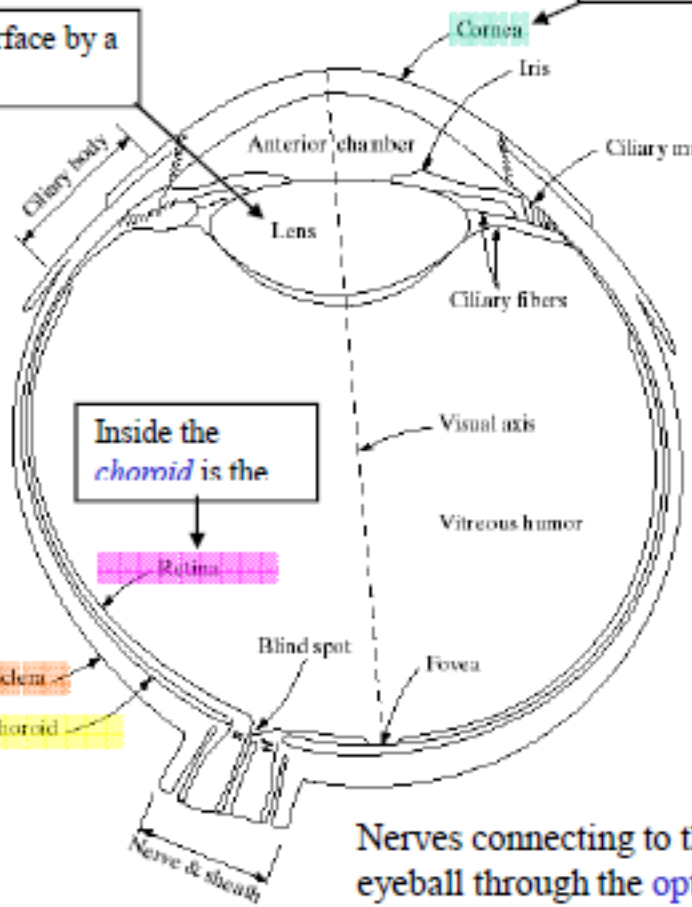


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

Inside the *choroid* is the

Three membranes enclose the eye:
Cornea and sclera, Choroid, Retina

ciliary body

iris diaphragm

The **retina** is composed of two types of receptors: ***rods*** and ***cones***.

Pupil size: 2-8mm

Eye color: melanin (pigment) in iris

Nerves connecting to the retina leave the eyeball through the optic nerve bundle.

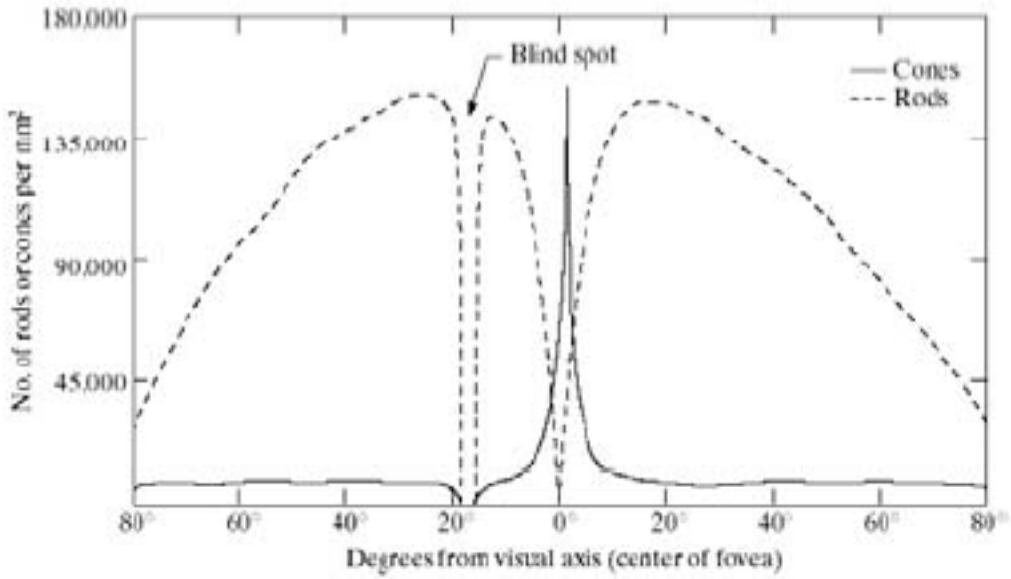


FIGURE 2.2 Distribution of rods and cones in the retina.

The distribution of rods and cones is radially symmetric wrt the fovea (central portion of the retina), except at the blind spot which includes no receptors.

Cones are responsible for photopic (color or bright-light) vision; while **rods** are for **scotopic** (dim-light) vision.

Fovea area in the retina is circular with 1.5 mm in diameter where most of the cones are concentrated with 150 000 cones/mm^2. This is easily achievable with medium resolution CCD imaging chip of size 5mm x 5mm!

How's an object seen at the back of the eye?

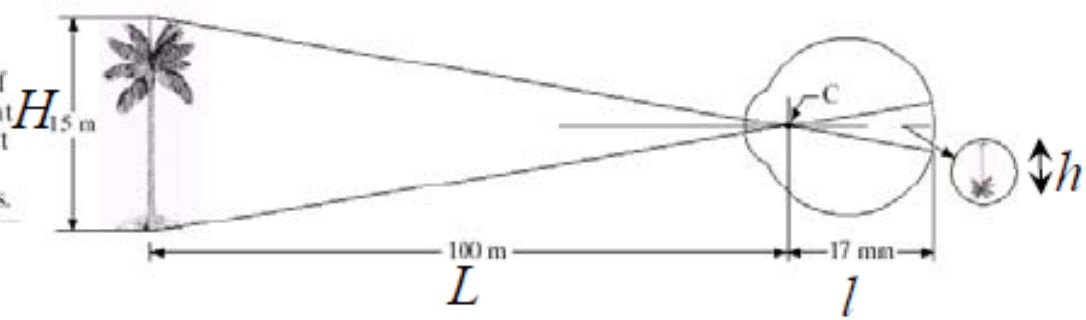


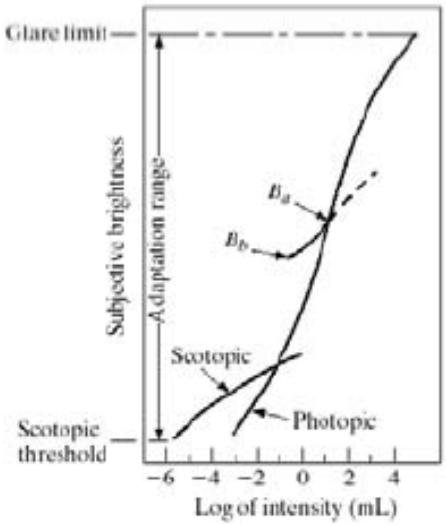
FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.

The focal lenght (distance bet center of the lens and the retina) varies from 17 mm to 14 mm (as the refractive power of the lens increases from its minimum to its maximum). Recall that $H/L = h/l$

Perception takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

Human eye can adapt to an enormous range (in the order of 10^{10}) of light intensity levels, from scotopic threshold to the glare limit.

FIGURE 2.4 Range of subjective brightness sensations showing a particular adaptation level.



Subjective brightness (i.e. perceived intensity) is a logarithmic function of the light intensity incident on the eye.

In photopic vision alone, the range is about 10^6 (-2 to 4 in the log scale). The transition from scotopic to photopic vision is gradual over the range (0.001, 0.1) millilambert[1] (-3 to -1 mL in the log scale).[2]

- HVS cannot operate over 10 ^ 10 orders of magnitude dynamic range *simultaneously*
- Total range of distinct intensity levels eye can discriminate simultaneously is small
- For given conditions, the current visual sensitivity of HVS is called “brightness adaptation”.

incremental illumination Δ I appears in the form of a short duration flash

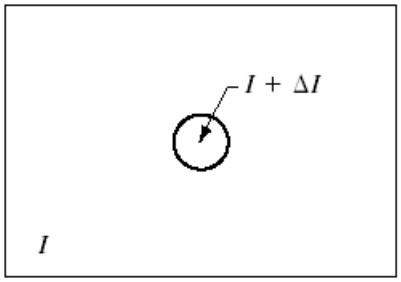
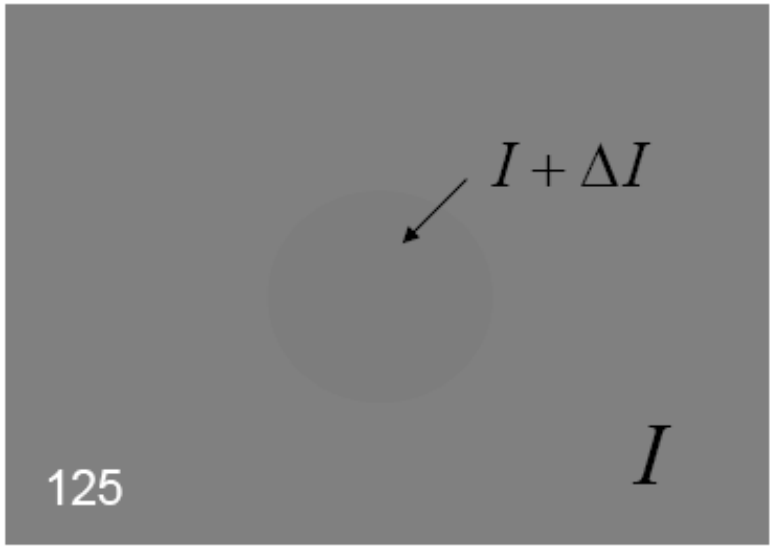


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.



Note: I is luminance, measured in cd/m^2

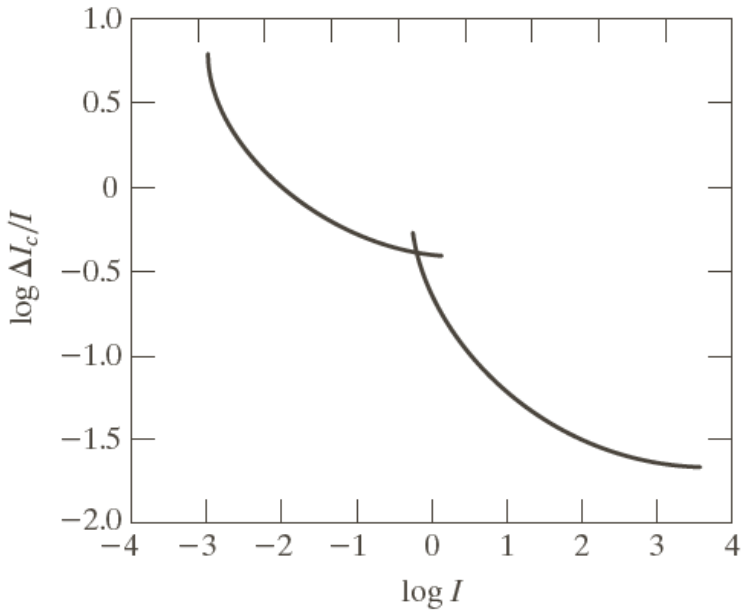
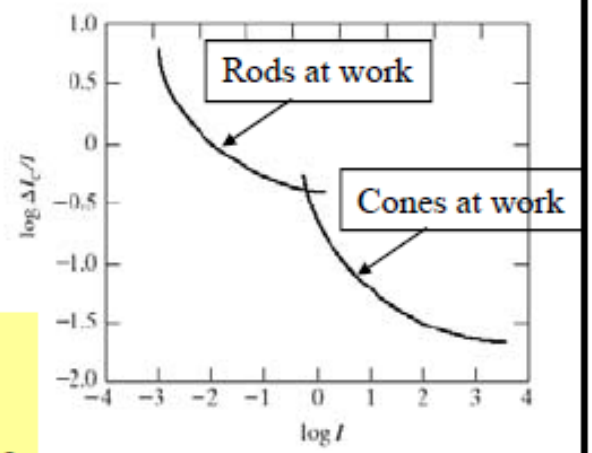


FIGURE 2.6 Typical Weber ratio as a function of intensity.

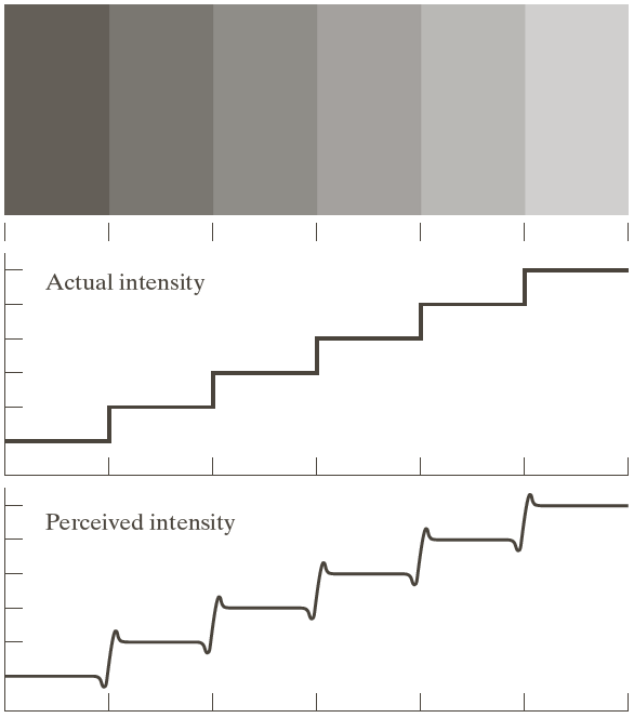
A small Weber ratio indicates "good" brightness where a small percentage change in illumination is discriminable. On the other hand, a large Weber ratio represents "poor" brightness indicating that a large percentage change in intensity is needed.

The curve shows that brightness discrimination is poor (large Weber ratio) at low level of illumination, and it improves significantly (Weber ratio decreases) as background illumination increases.



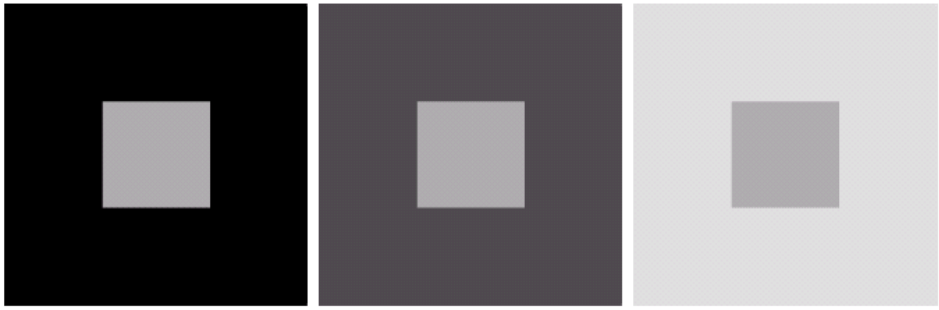
The two branches illustrate the fact that at low levels of illumination, vision is carried out by the rods, whereas at high levels (showing better discrimination), cones are at work.

2.11



a
b
c

FIGURE 2.7 Illustration of the Mach band effect. Perceived intensity is not a simple function of actual intensity.

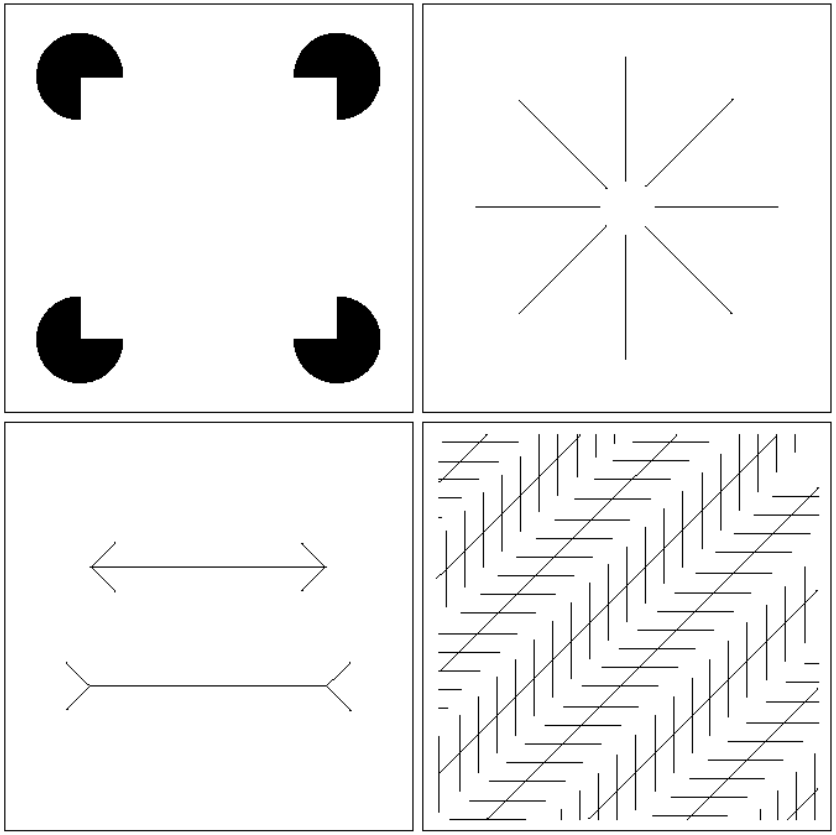


a b c

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.

a b
c d

FIGURE 2.9 Some well-known optical illusions.



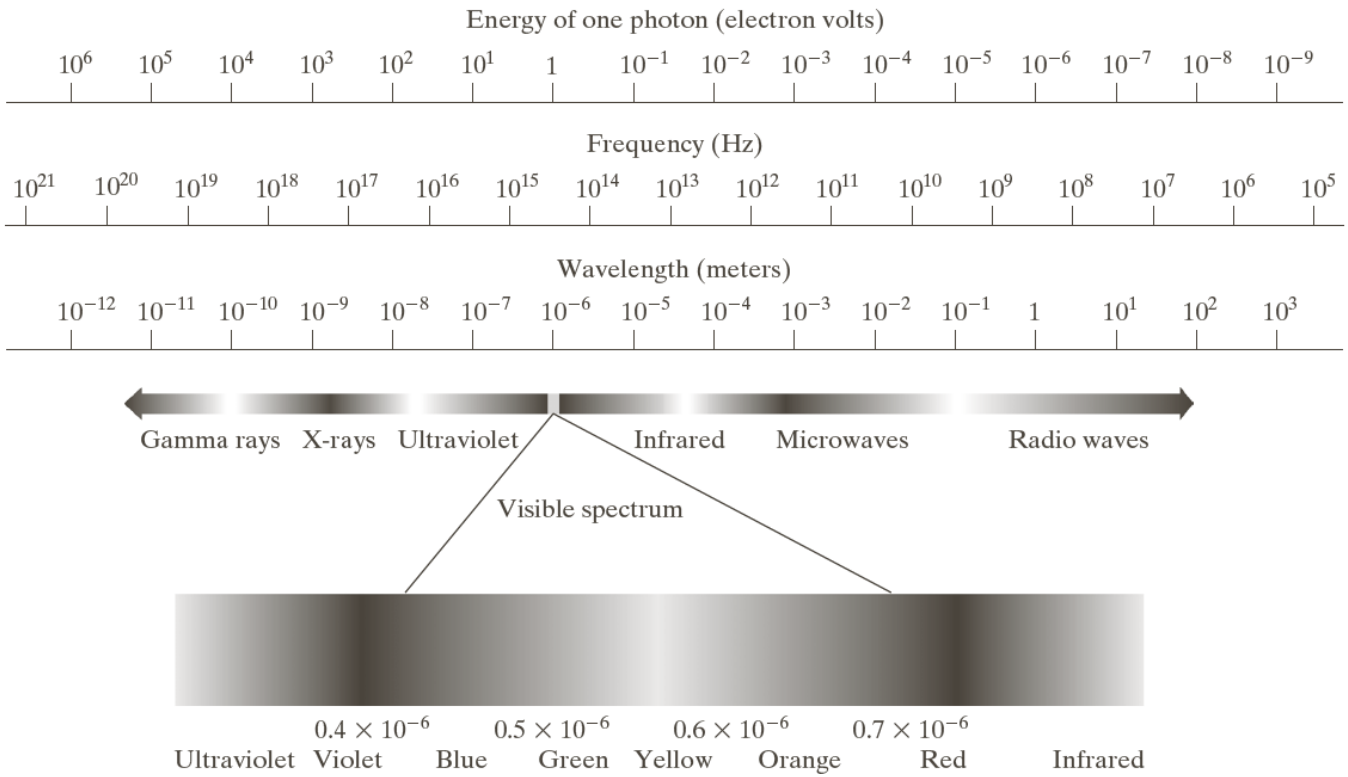


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

The electromagnetic spectrum can be expressed in terms of wavelength (λ), frequency (ν), or energy (E). Recall that

$$\lambda = c/\nu$$

where c is the speed of light (2.998×10^8 m/s).

The energy of the various components is given by:

$$E = h\nu$$

where h is Planck's constant ($6.62606891 \times 10^{-34}$ Joule-seconds (or m^2kg/s)). E is measured in electron-volt.

Light and EM Spectrum

- The colors that humans perceive in an object are determined by the nature of the light reflected from the object.

  e.g. green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelength

- Monochromatic light: void of color

  Intensity is the only attribute, from black to white

  Monochromatic images are referred to as **gray-scale** images

- Chromatic light bands: 0.43 to 0.79 um

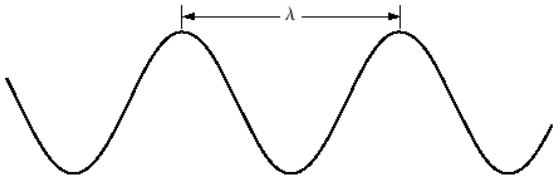
  The quality of a chromatic light source:

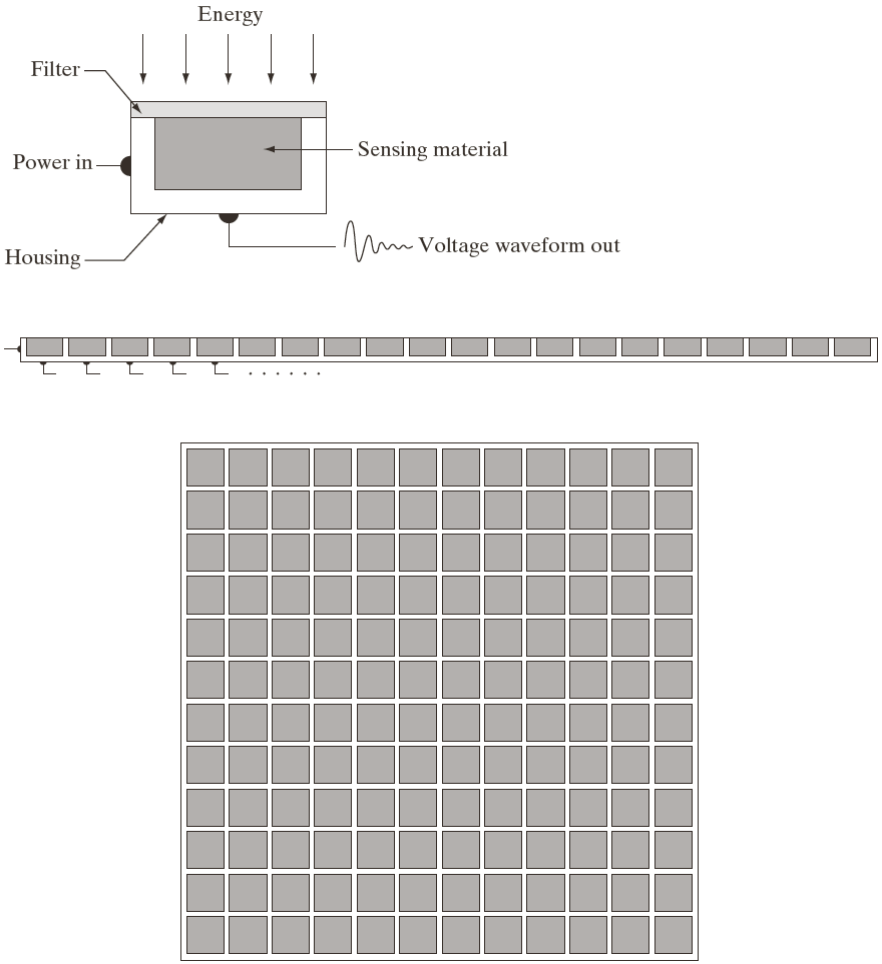
  Radiance: total amount of energy

  Luminance (lm): the amount of energy an observer perceives from a light source

  Brightness: a subjective descriptor of light perception that is impossible to measure. It embodies the achromatic notion of intensity and one of the key factors in describing color sensation.

FIGURE 2.11
Graphical representation of one wavelength.





a
b
c

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

Transform illumination energy into digital images

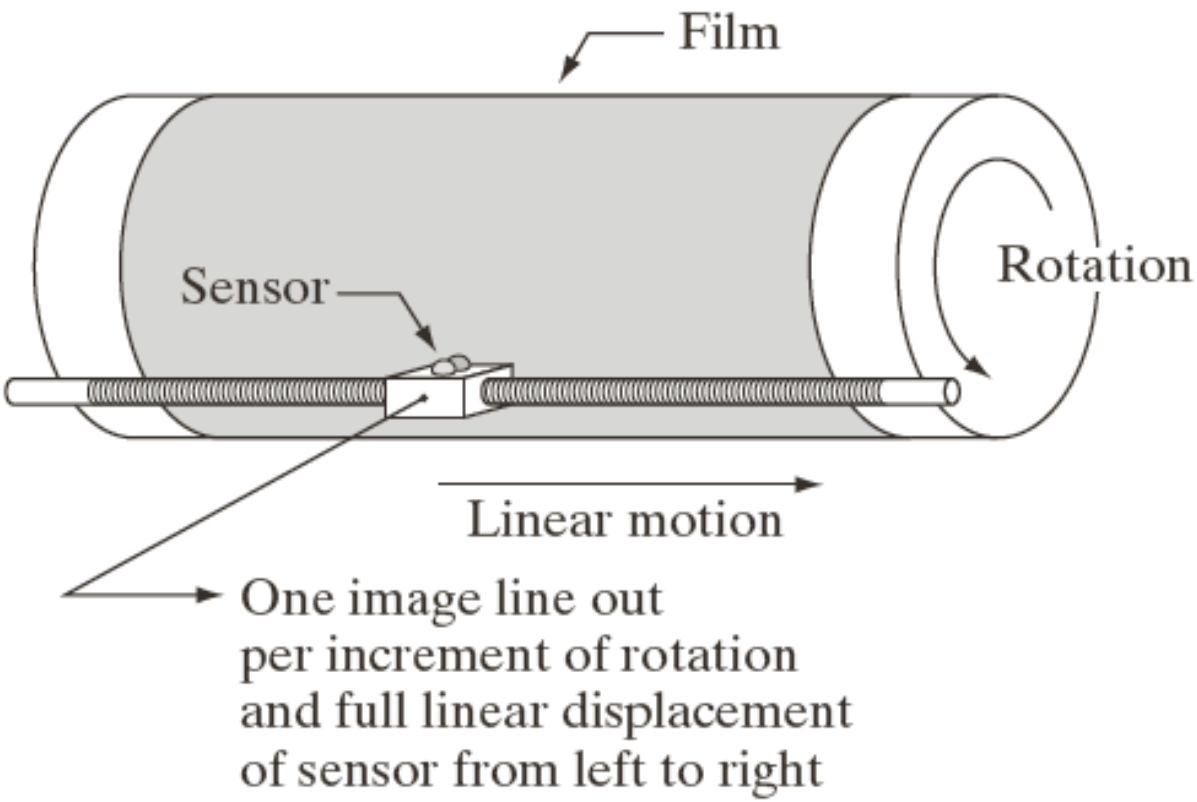
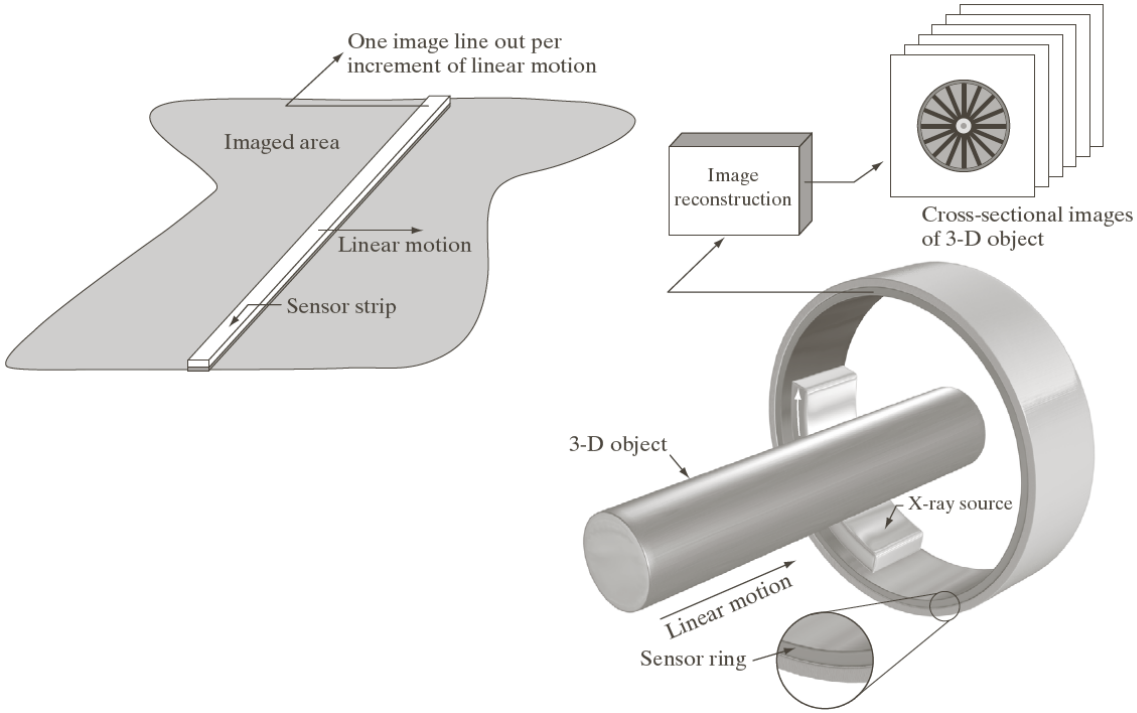
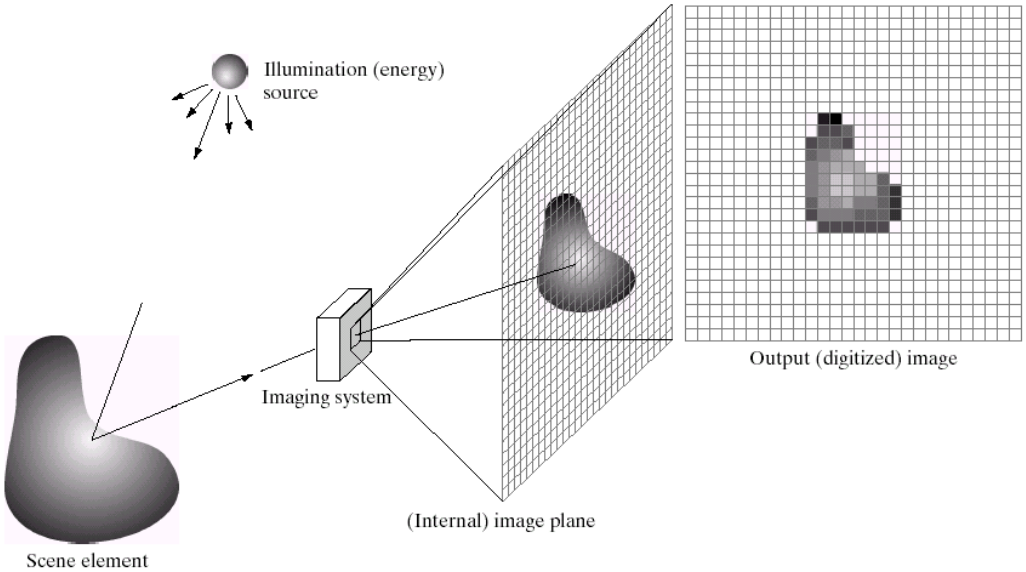


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.



a b

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



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.

Image formation model

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

$f(x, y)$: intensity at the point (x, y)

$i(x, y)$: illumination at the point (x, y)

(the amount of source illumination incident on the scene)

$r(x, y)$: reflectance/transmissivity at the point (x, y)

(the amount of illumination reflected/transmitted by the object)

where $0 < i(x, y) < \infty$ and $0 < r(x, y) < 1$

Some Typical Ranges of illumination

- Illumination

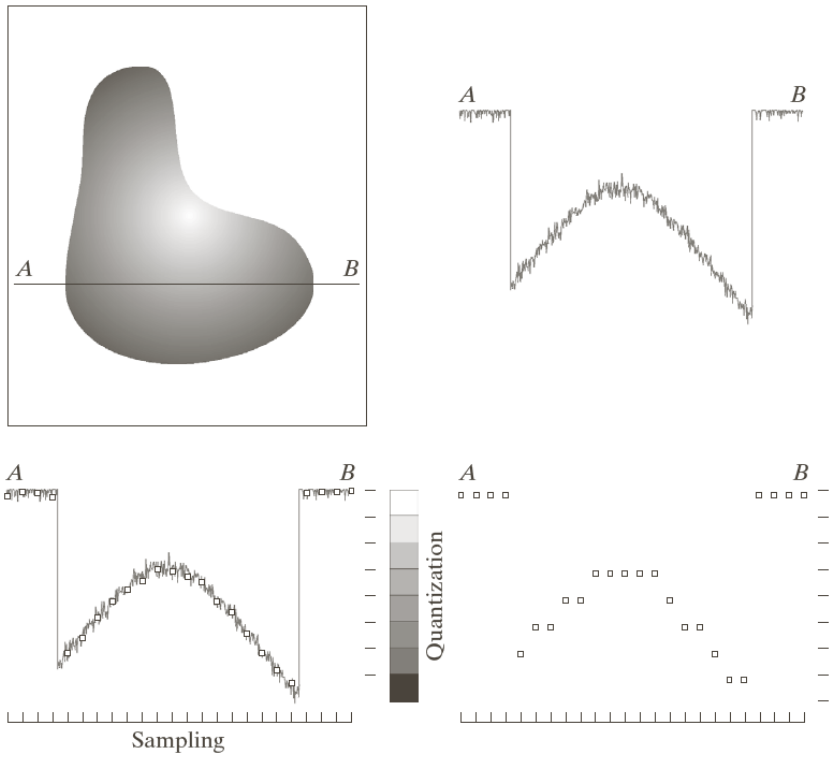
Lumen — A unit of light flow or luminous flux

Lumen per square meter (lm/m^2) — The metric unit of measure for illuminance of a surface

- On a clear day, the sun may produce in excess of 90,000 lm/m^2 of illumination on the surface of the Earth
- On a cloudy day, the sun may produce less than 10,000 lm/m^2 of illumination on the surface of the Earth
- On a clear evening, the moon yields about 0.1 lm/m^2 of illumination
- The typical illumination level in a commercial office is about 1000 lm/m^2

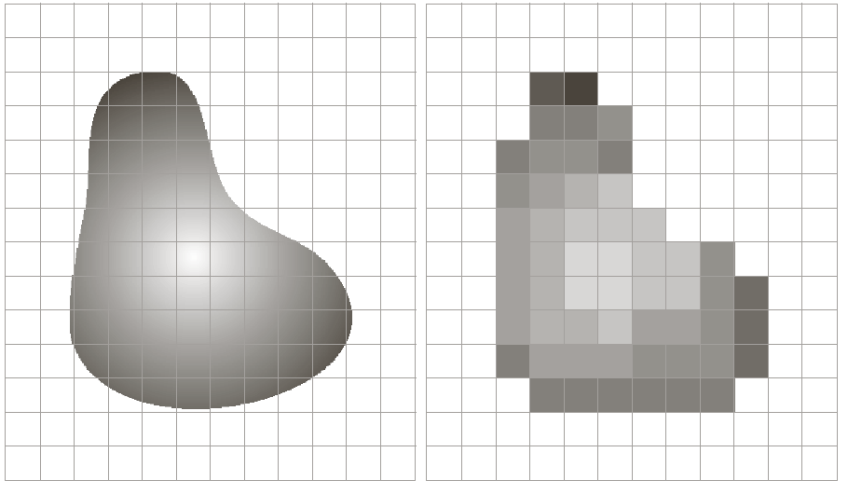
Some Typical Ranges of Reflectance

- Reflectance
  - 0.01 for black velvet
  - 0.65 for stainless steel
  - 0.80 for flat-white wall paint
  - 0.90 for silver-plated metal
  - 0.93 for snow



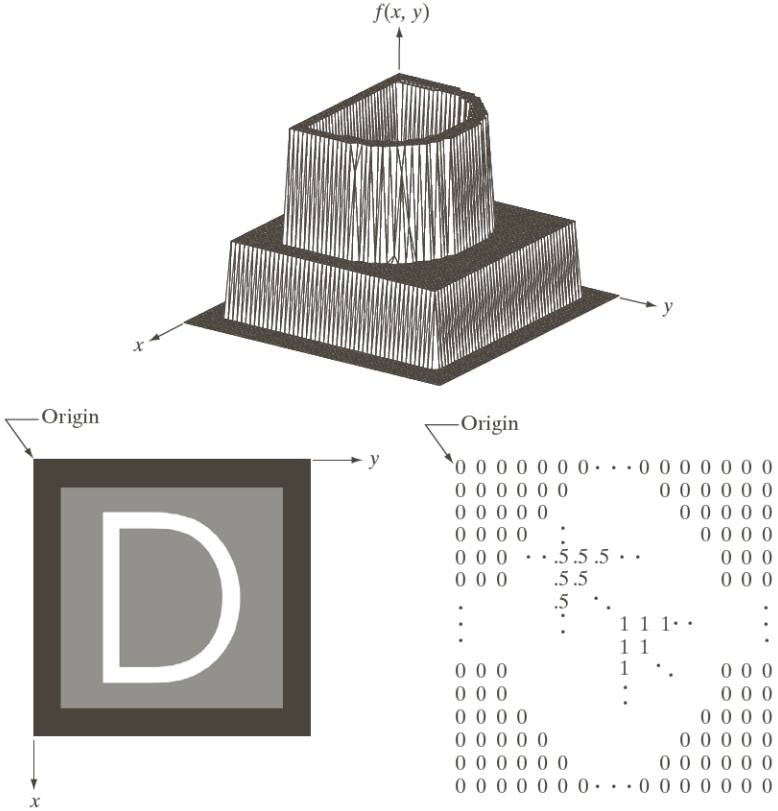
a b
c d

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.



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).

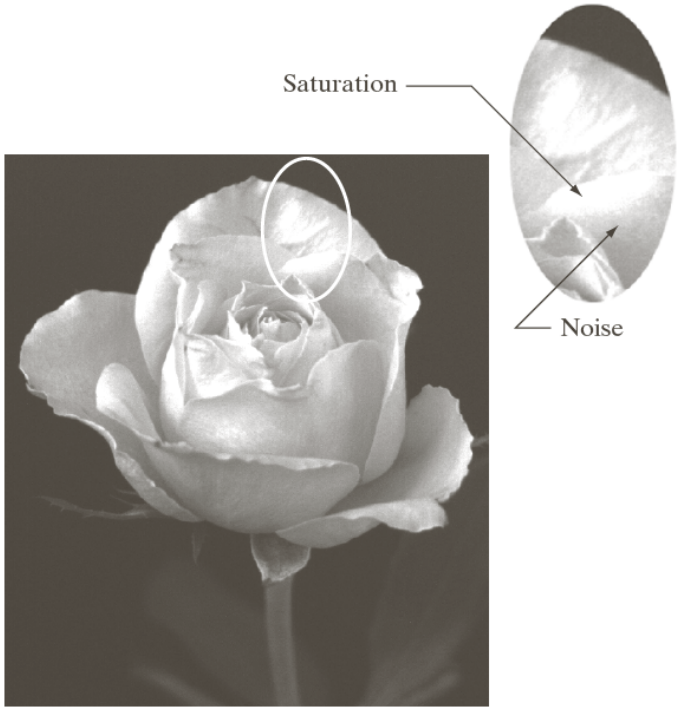


FIGURE 2.19 An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity levels are clipped (note how the entire saturated area has a high, *constant* intensity level). Noise in this case appears as a grainy texture pattern. Noise, especially in the darker regions of an image (e.g., the stem of the rose) masks the lowest detectable true intensity level.

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

- Discrete intensity interval [0, L-1], $L=2^k$
- The number b of bits required to store a $M \times N$ digitized image

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

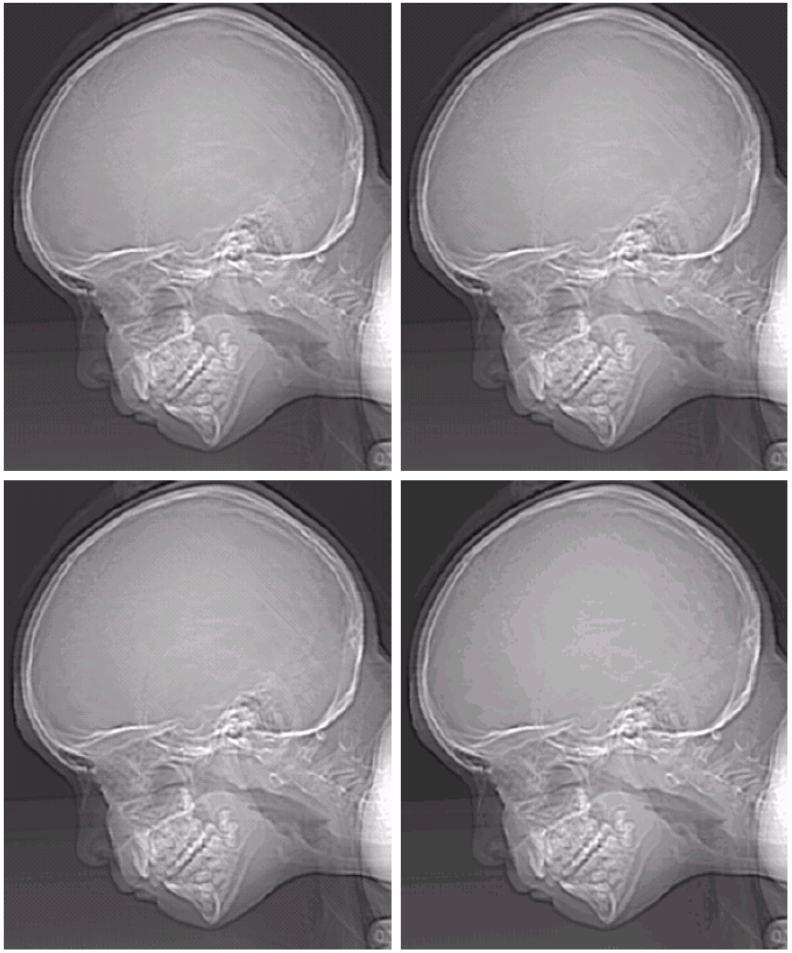
Spatial and Intensity Resolution

- Spatial resolution
  - — A measure of the smallest discernible detail in an image
  - — stated with *line pairs per unit distance, dots (pixels) per unit distance, dots per inch (dpi)*

- Intensity resolution
  - — The smallest discernible change in intensity level
  - — stated with *8 bits, 12 bits, 16 bits, etc.*

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

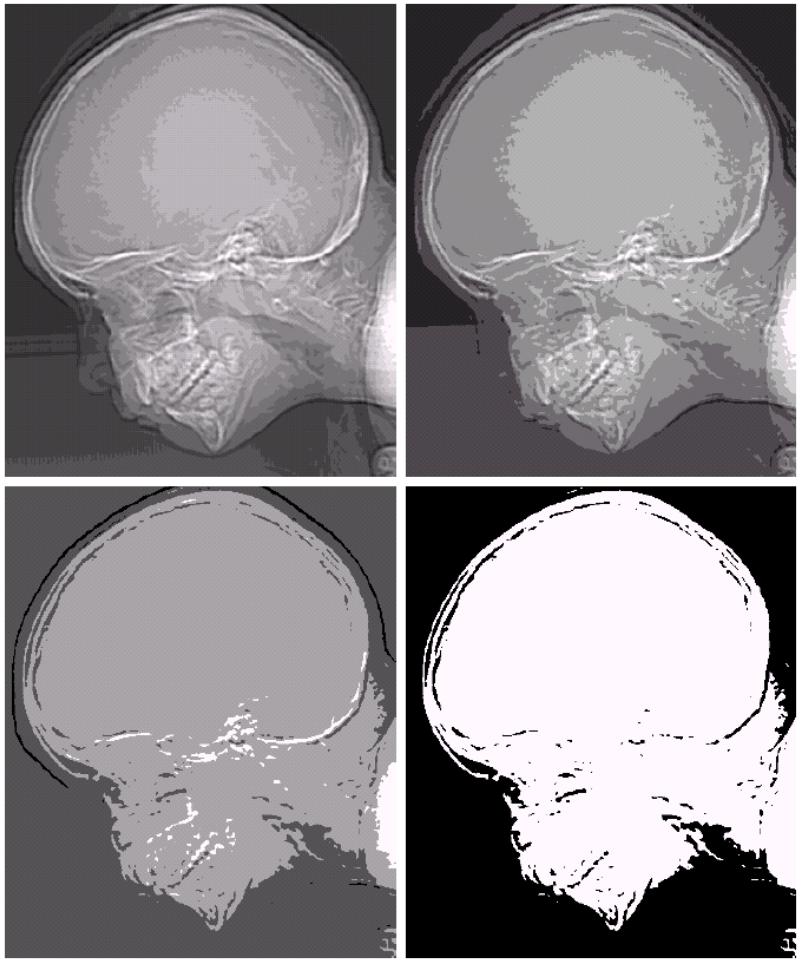


FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

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.)

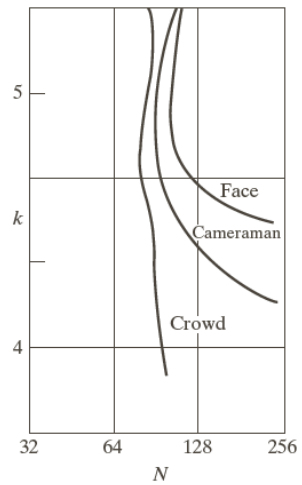


FIGURE 2.23 Typical isopreference curves for the three types of images in Fig. 2.22.

- **Curves show isopreference points**
- **For low-, moderate-detail images, increasing K or N will help perceptual quality**
- **For high-detail images, increasing N is more important.**

Image Interpolation

- **Interpolation** — Process of using known data to estimate unknown values

  e.g., zooming, shrinking, rotating, and geometric correction

- **Interpolation** (sometimes called ***resampling***) — an imaging method to increase (or decrease) the number of pixels in a digital image.

  Some digital cameras use interpolation to produce a larger image than the sensor captured or to create digital zoom

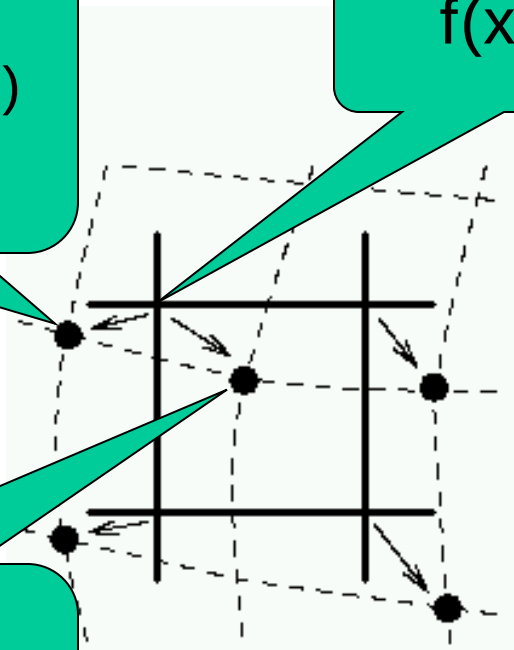
  http://www.dpreview.com/learn/?/key=interpolation

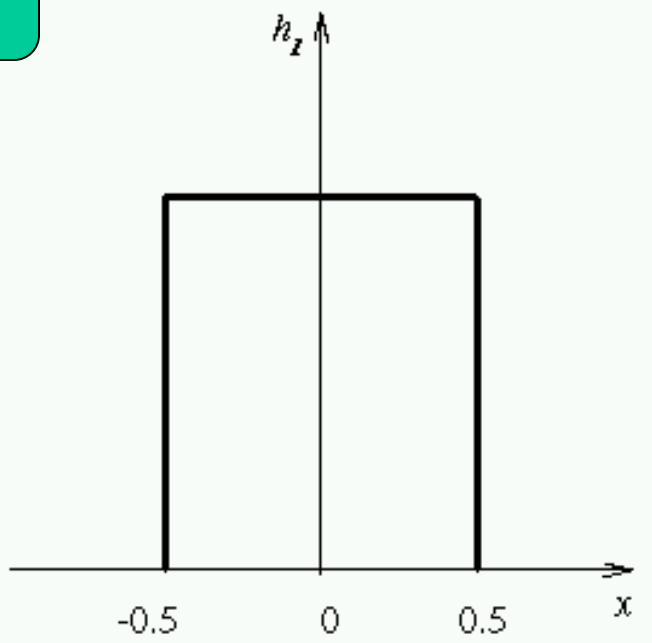
Image Interpolation:

Nearest Neighbor Interpolation

$f_1(x_2,y_2) =$
$f(\text{round}(x_2), \text{round}(y_2))$
$= f(x_1,y_1)$

$f(x_1,y_1)$





$f_1(x_3,y_3) =$
$f(\text{round}(x_3), \text{round}(y_3))$
$= f(x_1,y_1)$

Image Interpolation:

Bilinear Interpolation

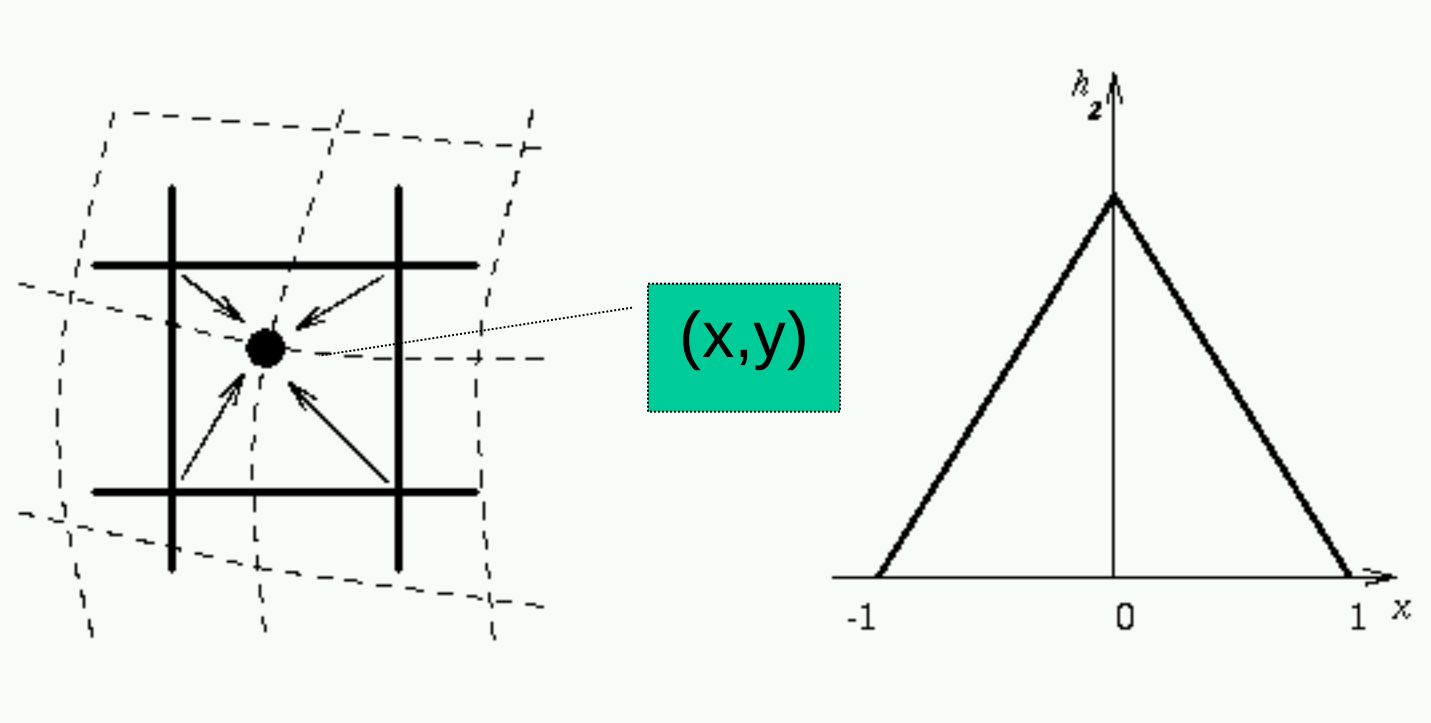


Image Interpolation:
Bicubic Interpolation

- The intensity value assigned to point (x,y) is obtained by the following equation

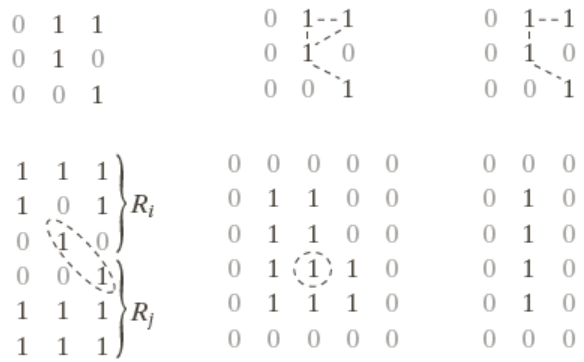
$$f_3(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} x^i y^j$$

- The sixteen coefficients are determined by using the sixteen nearest neighbors.

http://en.wikipedia.org/wiki/Bicubic_interpolation

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).



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.

Example: Addition of Noisy Images for Noise Reduction

Noiseless image: f(x,y)

Noise: n(x,y) (at every pair of coordinates (x,y), the noise is uncorrelated and has zero average value)

Corrupted image: g(x,y)

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

Reducing the noise by adding a set of noisy images, $\{g_i(x,y)\}$

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

Example: Addition of Noisy Images for Noise Reduction

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

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

$$= E\left\{\frac{1}{K}\sum_{i=1}^{K}\left[f(x,y)+n_i(x,y)\right]\right\}$$

$$= f(x,y) + E\left\{\frac{1}{K}\sum_{i=1}^{K} n_i(x,y)\right\}$$

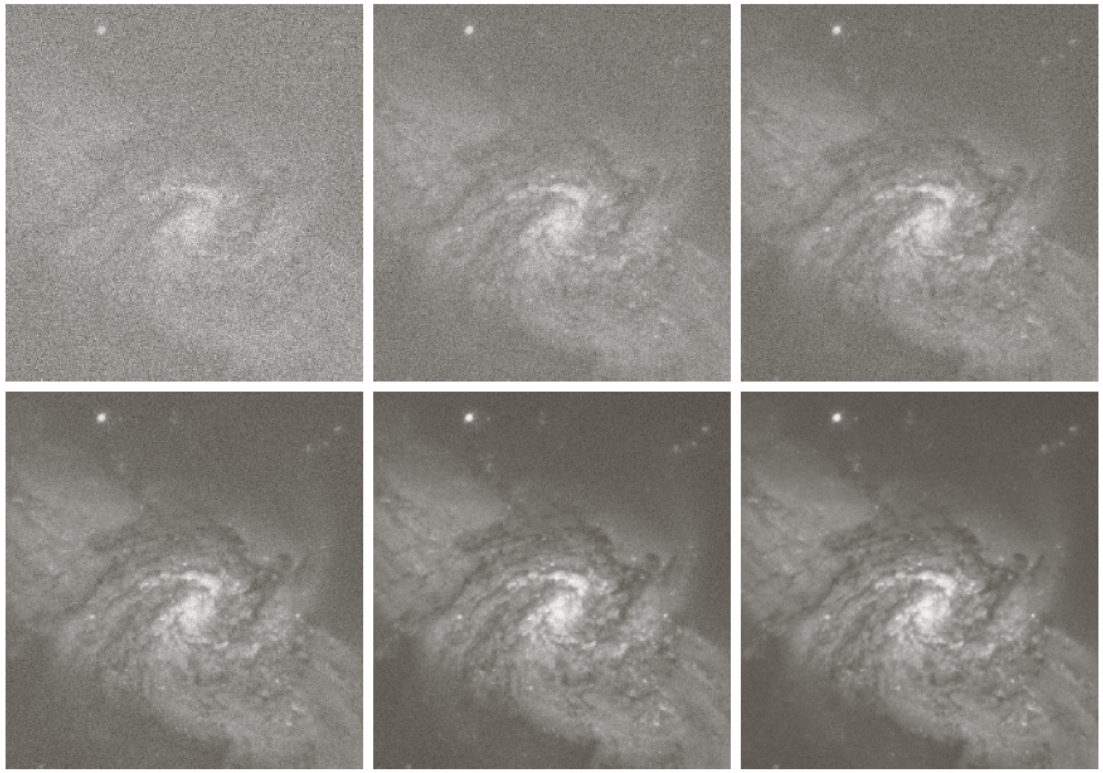
$$= f(x,y)$$

$$\sigma^2_{\bar{g}(x,y)} = \sigma^2_{\frac{1}{K}\sum_{i=1}^{K} g_i(x,y)}$$

$$= \sigma^2_{\frac{1}{K}\sum_{i=1}^{K} n_i(x,y)} = \frac{1}{K}\sigma^2_{n(x,y)}$$

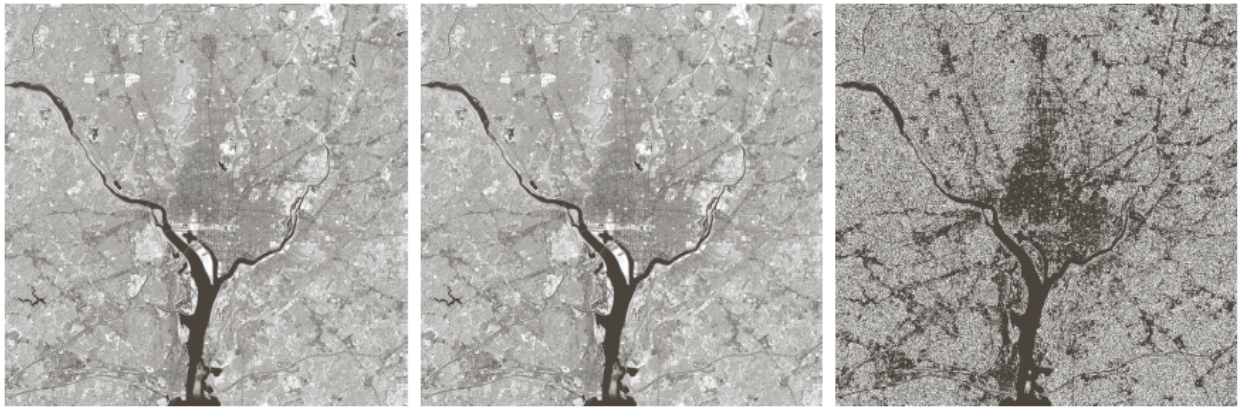
Example: Addition of Noisy Images for Noise Reduction

- In astronomy, imaging under very low light levels frequently causes sensor noise to render single images virtually useless for analysis.
- In astronomical observations, similar sensors for noise reduction by observing the same scene over long periods of time. Image averaging is then used to reduce the noise.



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

FIGURE 2.27 (a) Infrared image of the Washington, D.C. area. (b) Image obtained by 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.

Black regions in (c) shows no difference
Non black is regions where they are different

An Example of Image Subtraction: Mask Mode Radiography

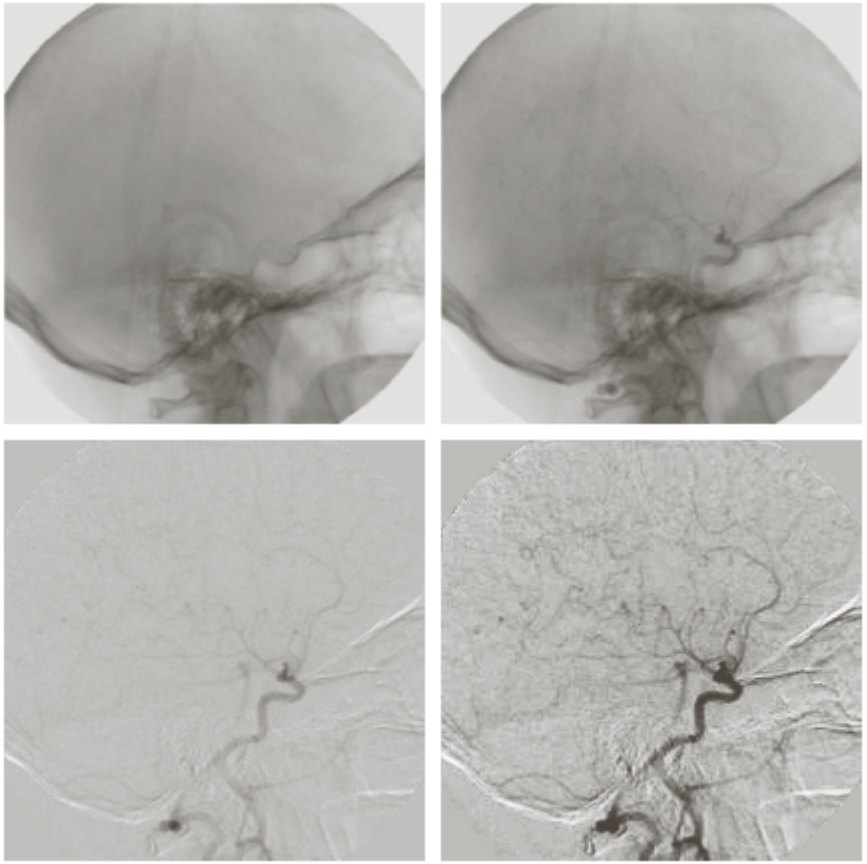
Mask h(x,y): an X-ray image of a region of a patient's body

Live images f(x,y): X-ray images captured at TV rates after injection of the contrast medium

Enhanced detail g(x,y)

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

The procedure gives a movie showing how the contrast medium propagates through the various arteries in the area being observed.



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.)

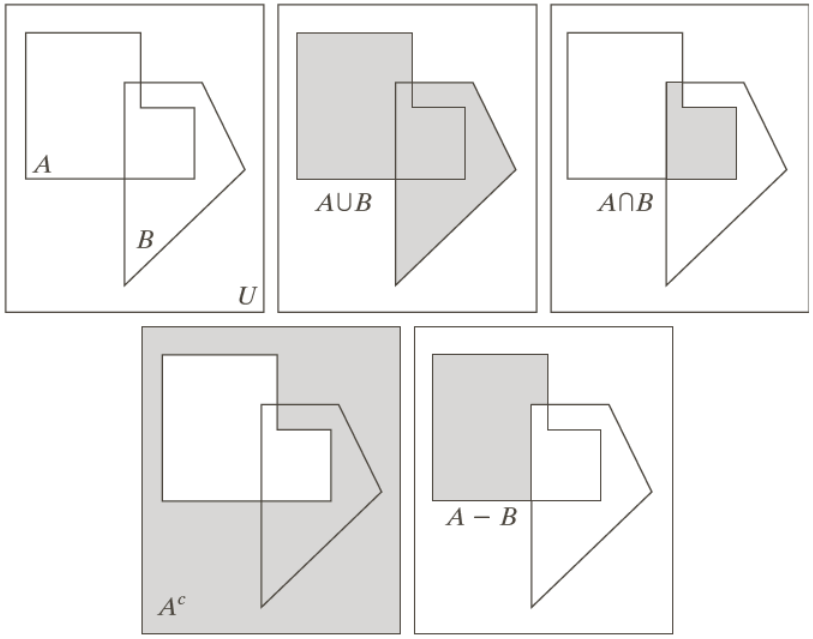
An Example of Image Multiplication

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.)

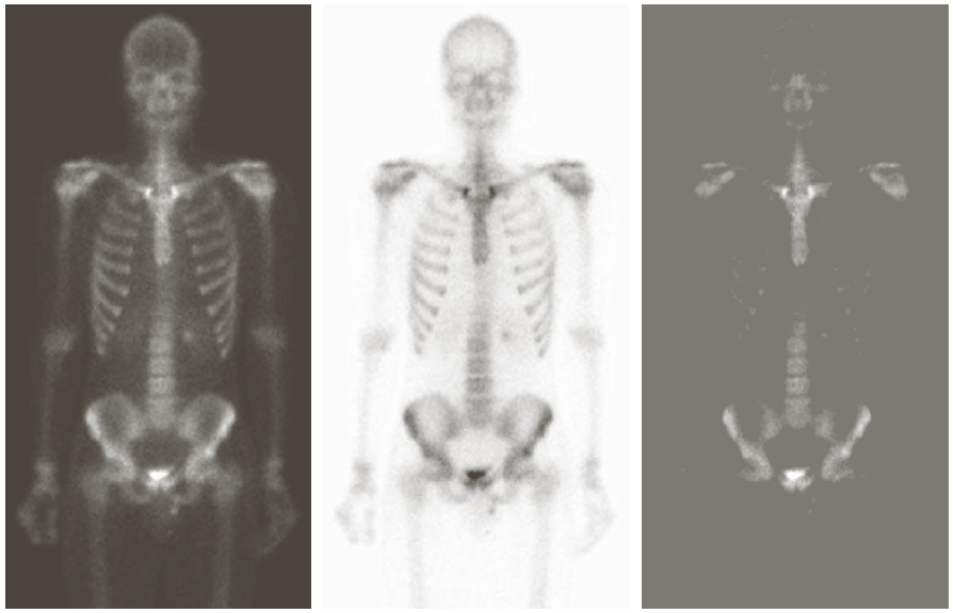
a b c

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).



a b c
d e

FIGURE 2.31
(a) Two sets of coordinates, A and B, in 2-D space. (b) The union of A and B. (c) The intersection of A and B. (d) The complement of A. (e) The difference between A and B. In (b)–(e) the shaded areas represent the member of the set operation indicated.



a b c

FIGURE 2.32 Set operations involving gray-scale images. (a) Original image. (b) Image negative obtained using set complementation. (c) The union of (a) and a constant image. (Original image courtesy of G.E. Medical Systems.)

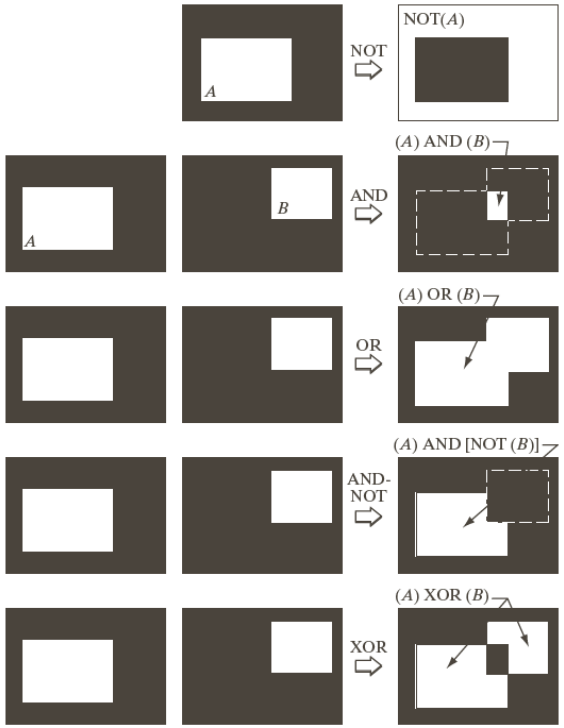


FIGURE 2.33 Illustration of logical operations involving foreground (white) pixels. Black represents binary 0s and white binary 1s. The dashed lines are shown for reference only. They are not part of the result.

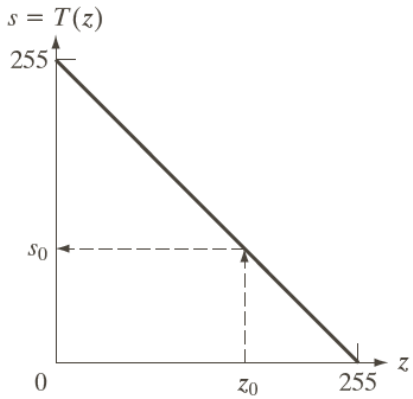
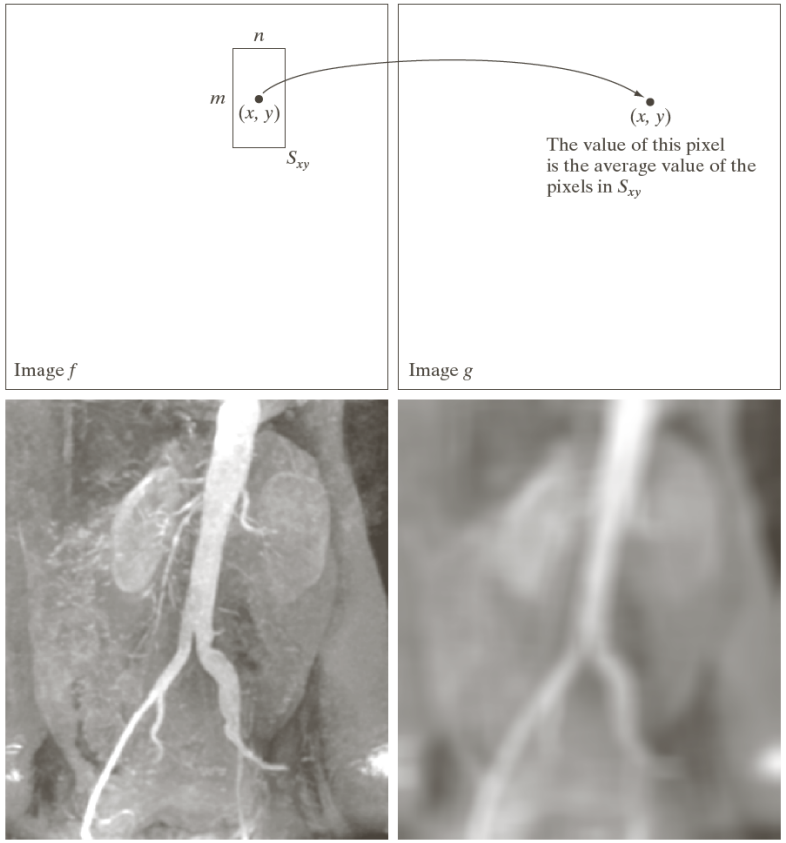


FIGURE 2.34 Intensity transformation function used to obtain the negative of an 8-bit image. The dashed arrows show transformation of an arbitrary input intensity value z_0 into its corresponding output value s_0.



a b
c d

FIGURE 2.35 Local averaging using neighborhood processing. The procedure is illustrated in (a) and (b) for a rectangular neighborhood. (c) The aortic angiogram discussed in Section 1.3.2. (d) The result of using Eq. (2.6-21) with $m = n = 41$. The images are of size 790×686 pixels.

Geometric Spatial Transformations

- Geometric transformation (rubber-sheet transformation)

  — A spatial transformation of coordinates

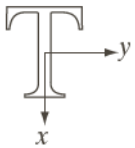
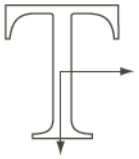
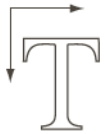
$$(x, y) = T\{(v, w)\}$$

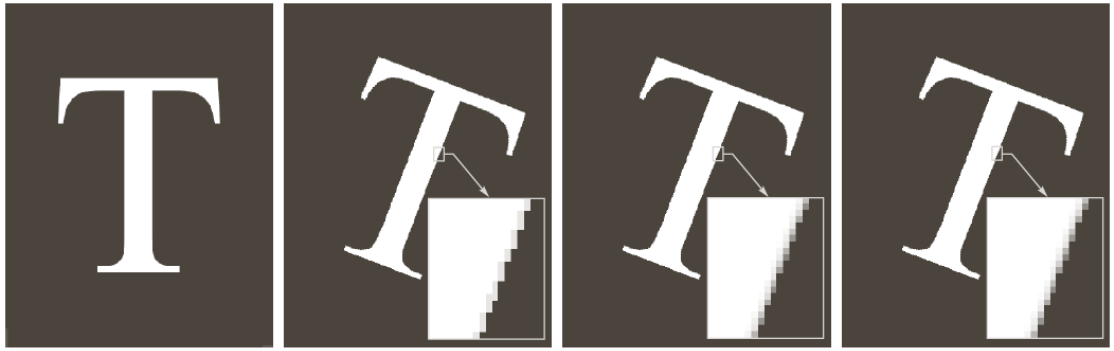
  — intensity interpolation that assigns intensity values to the spatially transformed pixels.

- Affine transform

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} v & w & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix}$$

TABLE 2.2
Affine transformations based on Eq. (2.6.–23).

Transformation Name	Affine Matrix, T	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = w$	
Scaling	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = c_x v$ $y = c_y w$	
Rotation	$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v\cos\theta - w\sin\theta$ $y = v\cos\theta + w\sin\theta$	
Translation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$	$x = v + t_x$ $y = w + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v + s_v w$ $y = w$	
Shear (horizontal)	$\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = s_h v + w$	



a b c d

FIGURE 2.36 (a) A 300 dpi image of the letter T. (b) Image rotated 21° clockwise using nearest neighbor interpolation to assign intensity values to the spatially transformed pixels. (c) Image rotated 21° using bilinear interpolation. (d) Image rotated 21° using bicubic interpolation. The enlarged sections show edge detail for the three interpolation approaches.

Image Registration

- Input and output images are available but the transformation function is unknown.

  Goal: estimate the transformation function and use it to register the two images.

- One of the principal approaches for image registration is to use ***tie points*** (also called ***control points***)

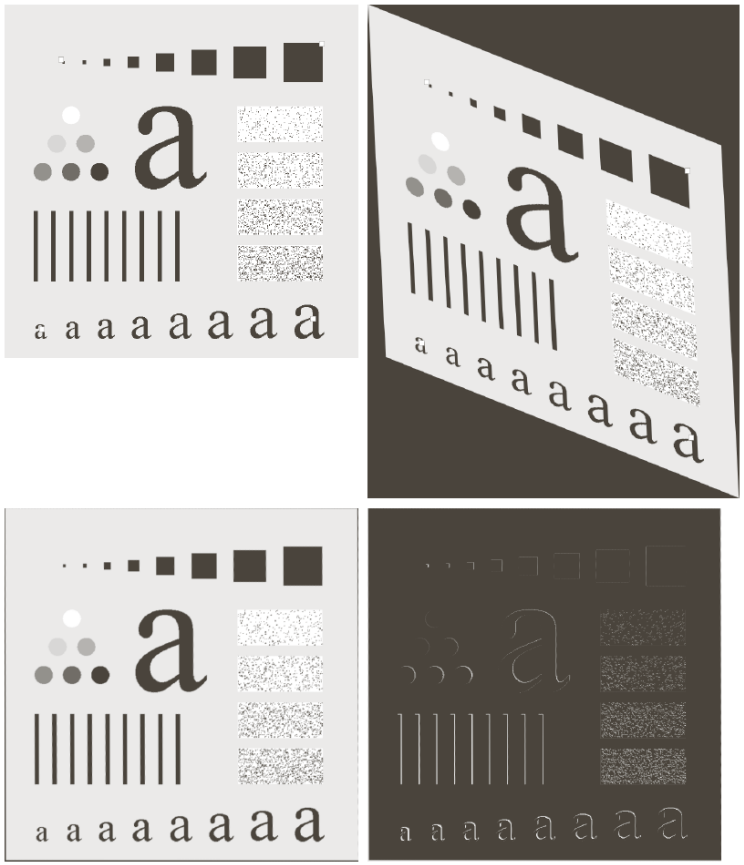
➢ The corresponding points are known precisely in the input and output (**reference**) images.

Image Registration

- A simple model based on bilinear approximation:

$$\begin{cases} x = c_1 v + c_2 w + c_3 vw + c_4 \\ y = c_5 v + c_6 w + c_7 vw + c_8 \end{cases}$$

Where (v, w) and (x, y) are the coordinates of tie points in the input and reference images.



a b
c d

FIGURE 2.37 Image registration. (a) Reference image. (b) Input (geometrically distorted image). Corresponding tie points are shown as small white squares near the corners. (c) Registered image (note the errors in the borders). (d) Difference between (a) and (c), showing more registration errors.

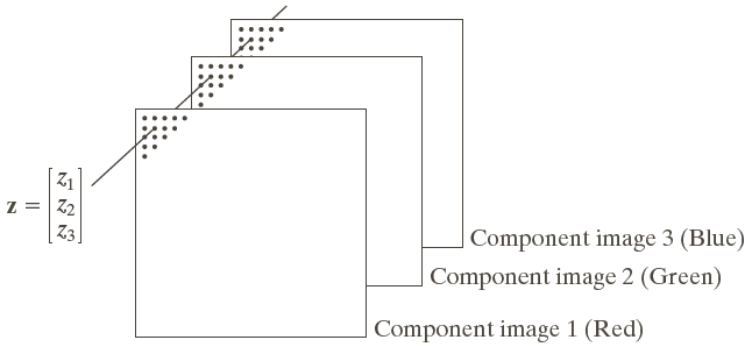


FIGURE 2.38 Formation of a vector from corresponding pixel values in three RGB component images.

Image Transform

- A particularly important class of 2-D linear transforms, denoted T(u, v)

$$T(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) r(x,y,u,v)$$

where $f(x,y)$ is the input image,

$r(x,y,u,v)$ is the *forward transformation* kernel,

variables u and v are the transform variables,

u = 0, 1, 2, ..., M-1 and v = 0, 1, ..., N-1.

Image Transform

- Given T(u, v), the original image f(x, y) can be recoverd using the inverse tranformation of T(u, v).

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u,v) s(x,y,u,v)$$

where $s(x,y,u,v)$ is the *inverse transformation* $\ker nel$,

x = 0, 1, 2, ..., M-1 and y = 0, 1, ..., N-1.

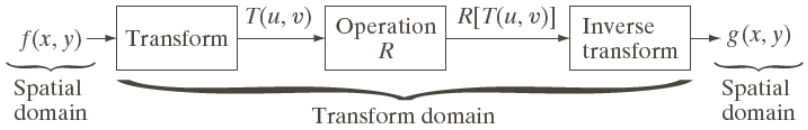
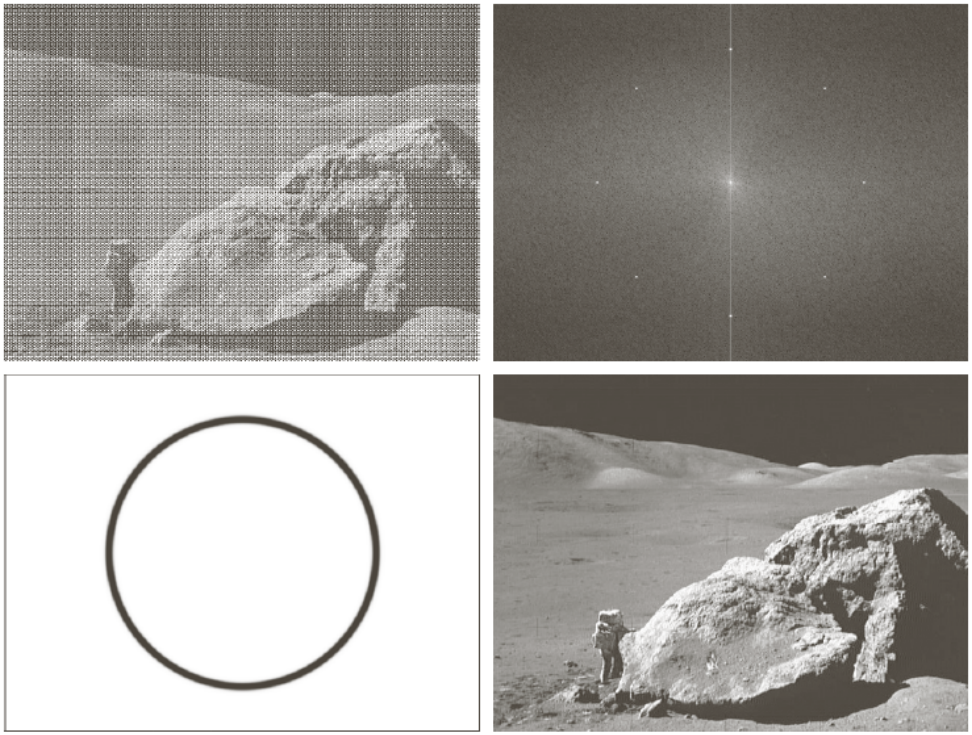


FIGURE 2.39 General approach for operating in the linear transform domain.



a b
c d

FIGURE 2.40
(a) Image corrupted by sinusoidal interference. (b) Magnitude of the Fourier transform showing the bursts of energy responsible for the interference. (c) Mask used to eliminate the energy bursts. (d) Result of computing the inverse of the modified Fourier transform. (Original image courtesy of NASA.)

Forward Transform Kernel

$$T(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) r(x,y,u,v)$$

The kernel $r(x,y,u,v)$ is said to be SEPERABLE if

$$r(x,y,u,v) = r_1(x,u) r_2(y,v)$$

In addition, the kernel is said to be SYMMETRIC if

$r_1(x,u)$ is functionally equal to $r_2(y,v)$, so that

$$r(x,y,u,v) = r_1(x,u) r_1(y,u)$$

The Kernels for 2-D Fourier Transform

The *forward* kernel

$$r(x, y, u, v) = e^{-j2\pi(ux/M + vy/N)}$$

Where $j = \sqrt{-1}$

The *inverse* kernel

$$s(x, y, u, v) = \frac{1}{MN} e^{j2\pi(ux/M + vy/N)}$$

2-D Fourier Transform

$$T(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M + vy/N)}$$

$$f(x,y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u,v) e^{j2\pi(ux/M + vy/N)}$$

Probabilistic Methods

Let z_i, $i = 0, 1, 2, ..., L-1$, denote the values of all possible intensities in an $M \times N$ digital image. The probability, $p(z_k)$, of intensity level z_k occurring in a given image is estimated as

$$p(z_k) = \frac{n_k}{MN},$$

where n_k is the number of times that intensity z_k occurs in the image.

$$\sum_{k=0}^{L-1} p(z_k) = 1$$

The mean (average) intensity is given by

$$m = \sum_{k=0}^{L-1} z_k p(z_k)$$

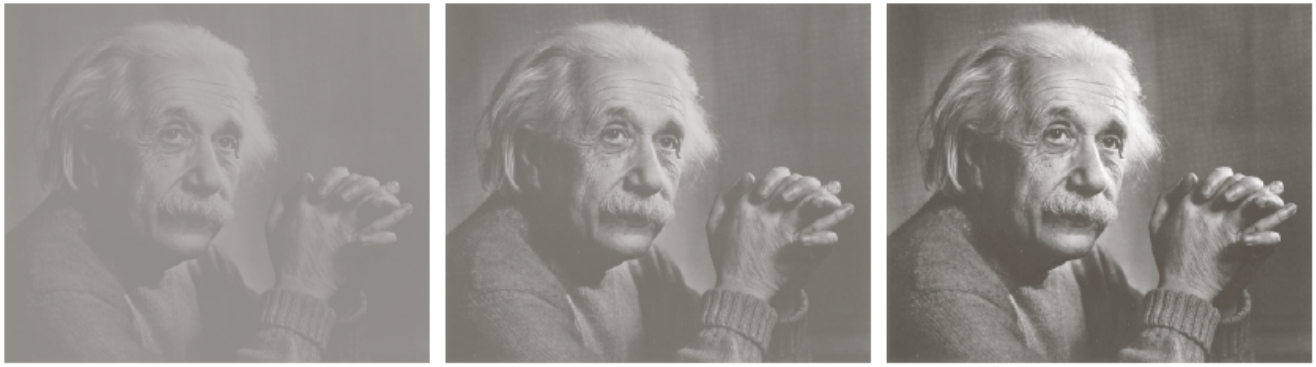
Probabilistic Methods

The variance of the intensities is given by

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k)$$

The nth moment of the intensity variable z is

$$u_n(z) = \sum_{k=0}^{L-1} (z_k - m)^n p(z_k)$$



a b c

FIGURE 2.41
Images exhibiting (a) low contrast, (b) medium contrast, and (c) high contrast.

$\sigma = 14.3$ $\sigma = 31.6$ $\sigma = 49.2$