EE 5098 – Digital Image Processing

2. Digital Image Fundamentals

Source: Frank Liu

Outline

- Elements of visual perception
- Light and the electromagnetic spectrum
- Image sensing and acquisition
- Image sampling and quantization
- Basic relationships between pixels
- Basic mathematical tools for DIP

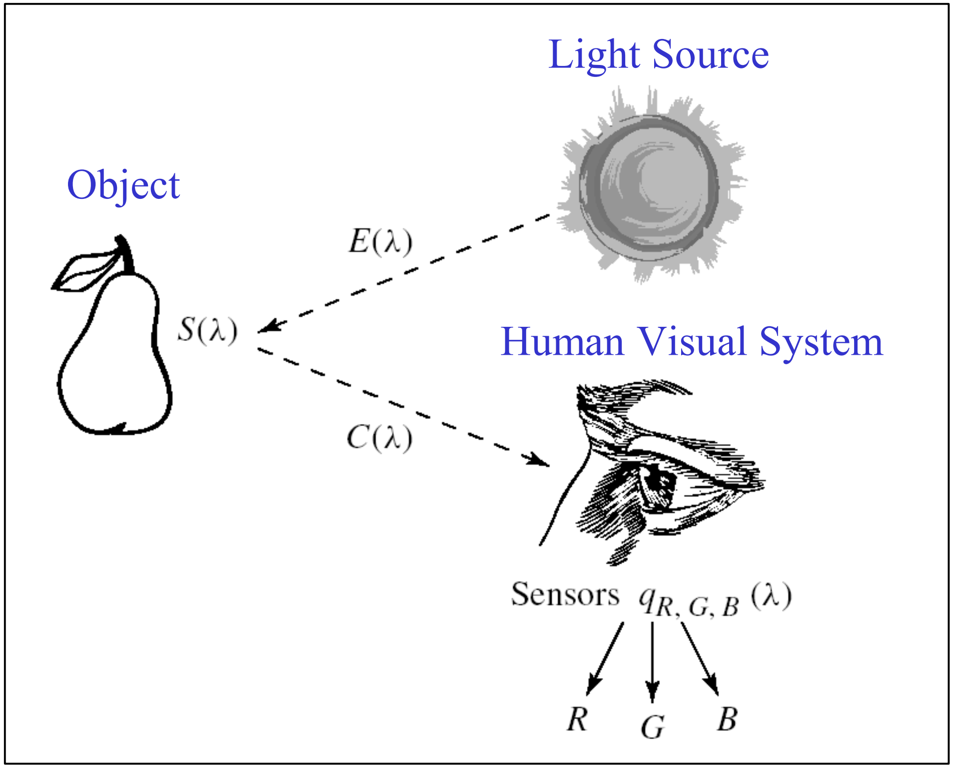
Visual Interpretation of Shadow

Charlie Chaplin Optic Illusion

https://youtu.be/QbKw0_v2clo

Visual Perception

- Vision is the most important sensor (more than 70% of our information is collected by vision) of a normal human being
- Our eyes can detect electromagnetic waves (400 to 700 nm) coming from the object



$$\boxed{C(\lambda) = E(\lambda)\, S(\lambda)}$$

$E(\lambda)$: spectral power distribution (SPD) or spectrum
$S(\lambda)$: surface spectral reflectance function
$C(\lambda)$: SPD of the perceived color signal
$q_R(\lambda)$, $q_G(\lambda)$, $q_B(\lambda)$: spectral sensitivity functions of HVS

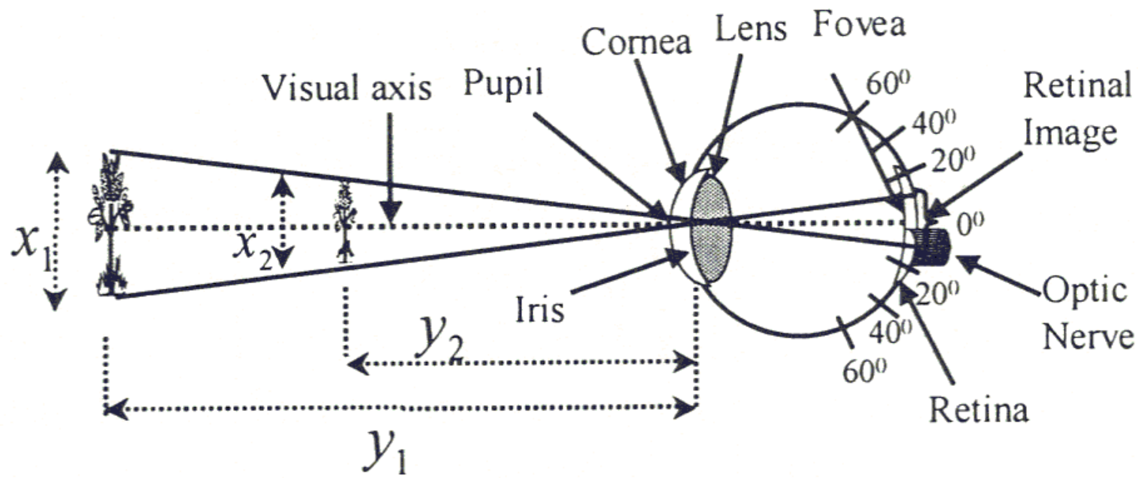
Visual Perception (cont’d)

- When the SPD $C(\lambda)$ is received by a viewer, a complex process starts in the HVS to form an image from light energy
- It is important to learn the characteristics of HVS to efficiently design multimedia systems for transmission, processing, and display of images
- For example, the systems need not transmit or display what the eye cannot see in order to reduce system cost
- Another example: unnoticeable split-up of light field

(a) (b)

(a) An image of the light field captured by placing a camera at the eye position of our near-eye AR display. The camera is configured to focus at the leftmost coke can (60 cm), and the AR content enclosed by the red box is displayed at about 33 cm. (b) Top: The defocused object captured by a camera. Bottom: The defocused object perceived by eye.

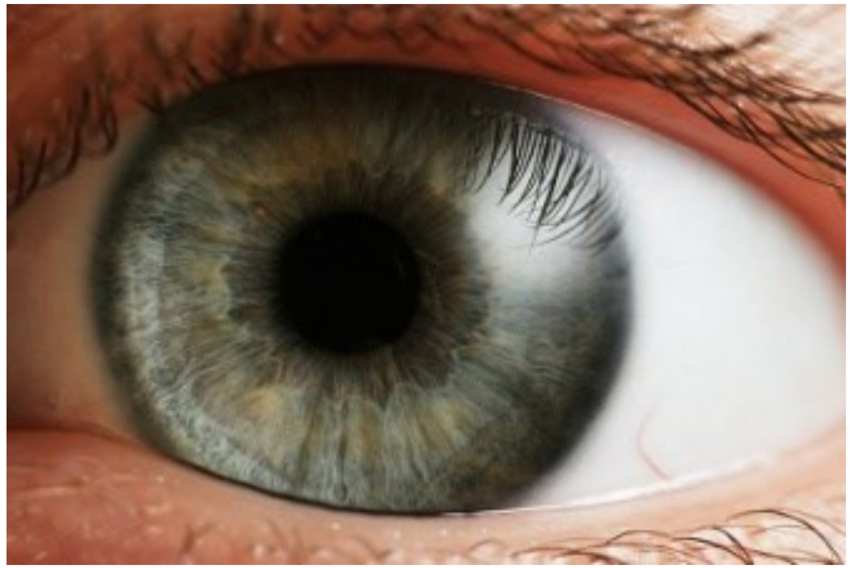
Human Visual System



- Can be considered an optical system
- The light from an object falls on our eye
- The pupil of the eye acts as an aperture
- An image is created on the *retina*
- We see the object
- The angle $\theta = 2\tan^{-1}(\frac{x_1}{2y_1}) = 2\tan^{-1}(\frac{x_2}{2y_2})$ determines the retinal image size

- Incident light is focused by the cornea and the lens into the retina
- The image on the retina is transformed into neural responses by photoreceptors
- Neural cells in the retina interconnect these responses providing some elementary "retinal processing"
- The neural responses are transformed into several neural representations in the optic nerve and proceed to the cortex of the brain

Human Visual System (cont'd)



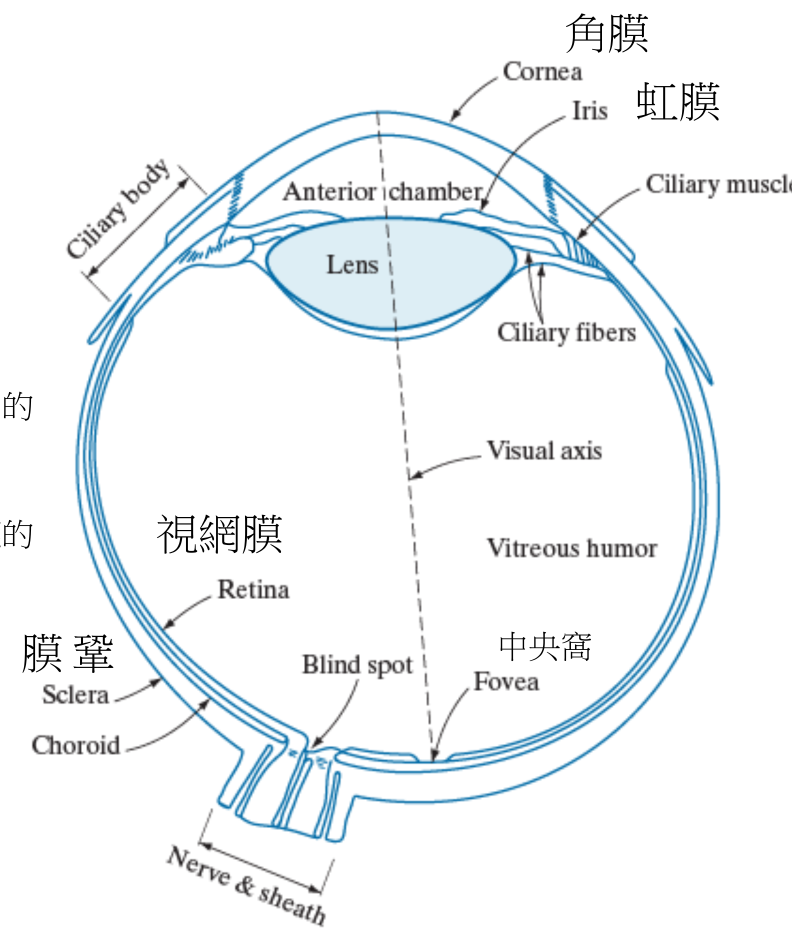
- The central transparent area (showing as black) is the pupil (瞳孔)
- The grey/blue area surrounding the pupil is the iris (虹膜)
- The white outer area is the sclera (鞏膜)
- The central transparent part of the sclera is the cornea (角膜)

Human Visual System (cont’d)

- Pupil
  - Controls the amount of light entering the eye
  - Size increases (decreases) at low (high) illumination; typical size 2 mm
- Retina
  - Contains two types of photoreceptors

    暗視力的
    - Rods (100 millions) provide *scotopic* vision at low illumination

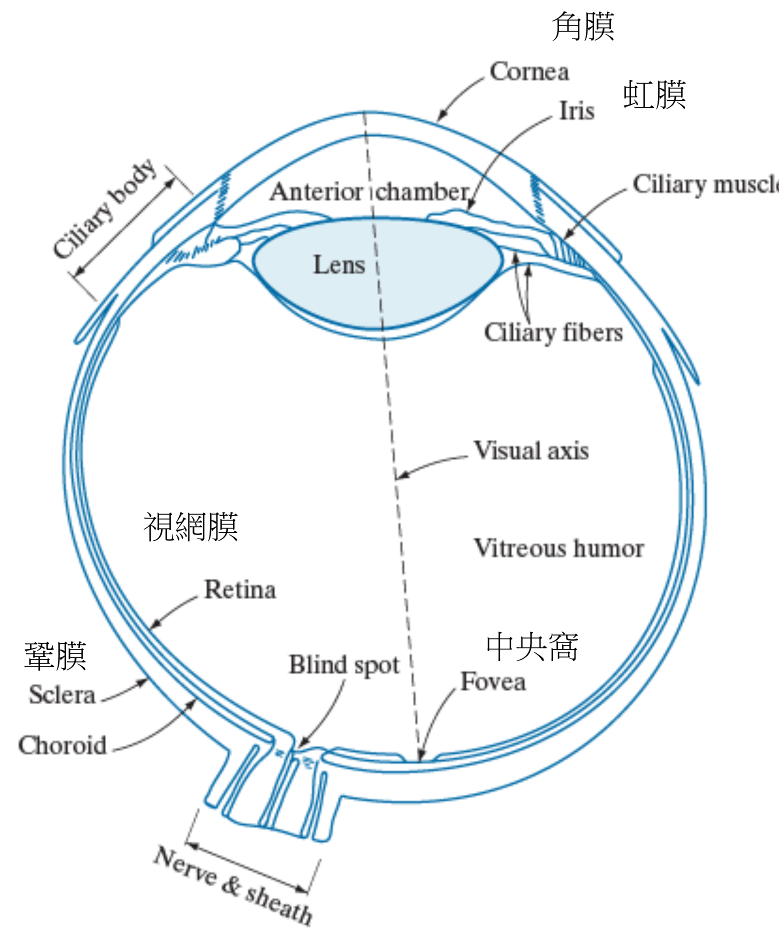
    光適應的
    - Cones (5 millions) provide *photopic* vision at higher illumination
  - In the *mesopic* vision, both rods and cones are active
- 0.8 millions nerve fibers
  - Compression!



Cross section of a human eye

Two Distinguished Regions in Retina

- Fovea
  - A small region of the retina densely packed with photoreceptors.
  - Has high visual acuity
  - The neurons in the central fovea are displaced, leaving a clear optical path from the lens to the photoreceptors
  - Fovea is relatively small compared to the rest of retina, but the fovea is the only area of the retina where 20/20 vision is attainable, and very important for seeing fine detail and color.
- Blind spot
  - This is where the optic nerve exists the eye
  - No photoreceptors here

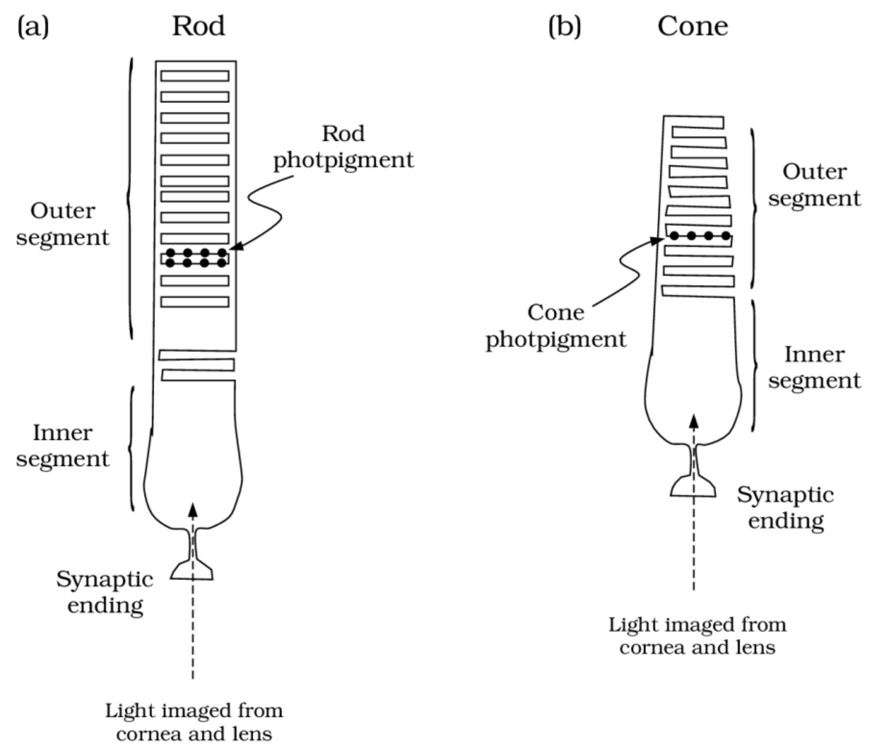


Two Types of Photoreceptors

- Rods
  - Long and thin
  - Sensitive to light
  - Responsible for scotopic vision
  - Used at low illumination

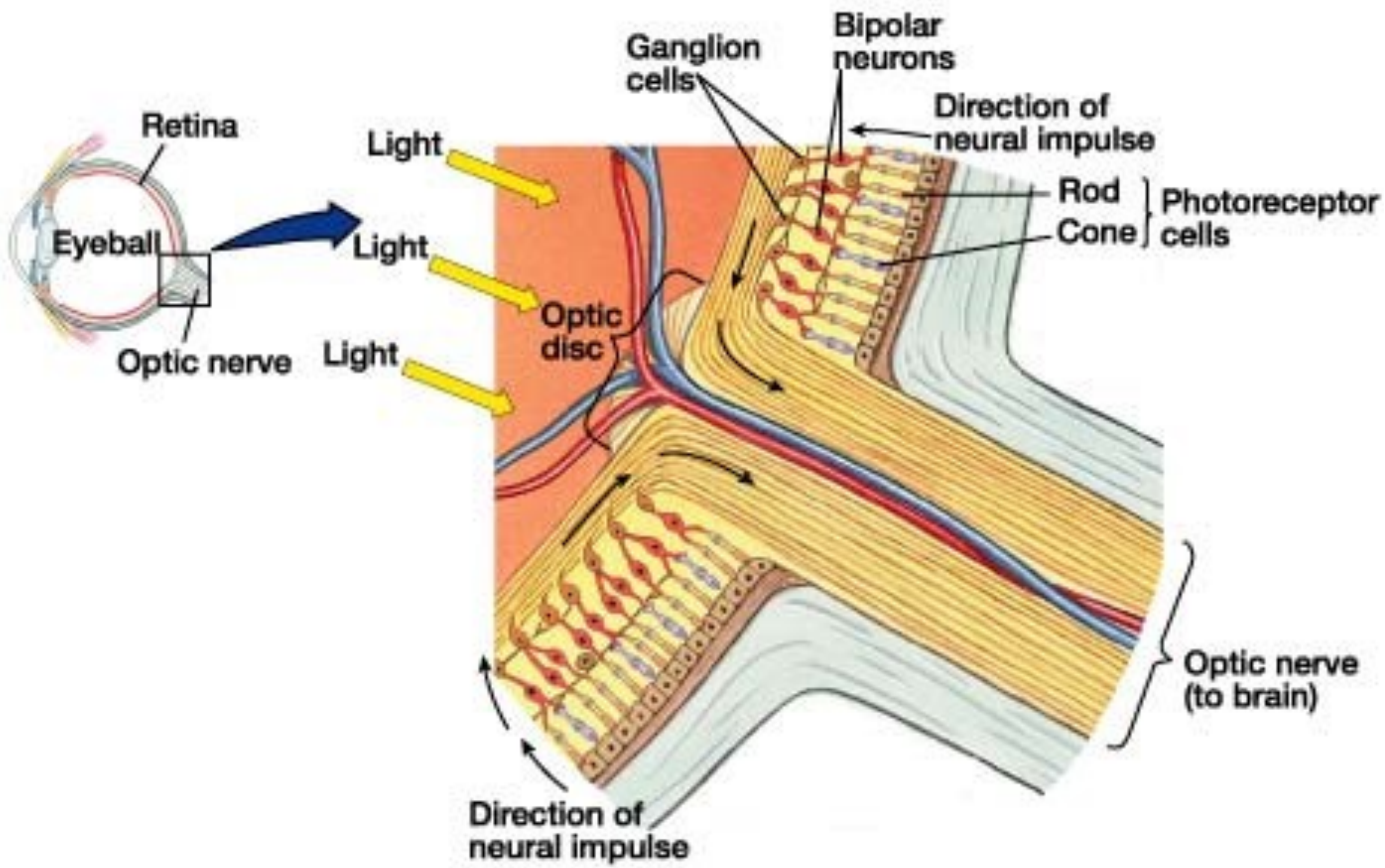
- Cones
  - Short and thick
  - Less sensitive to light
  - Responsible for photopic vision
  - Used at high illumination
  - Source of color vision

When a light stimulates a rod or cone, a photochemical transition occurs, producing a nerve impulse.



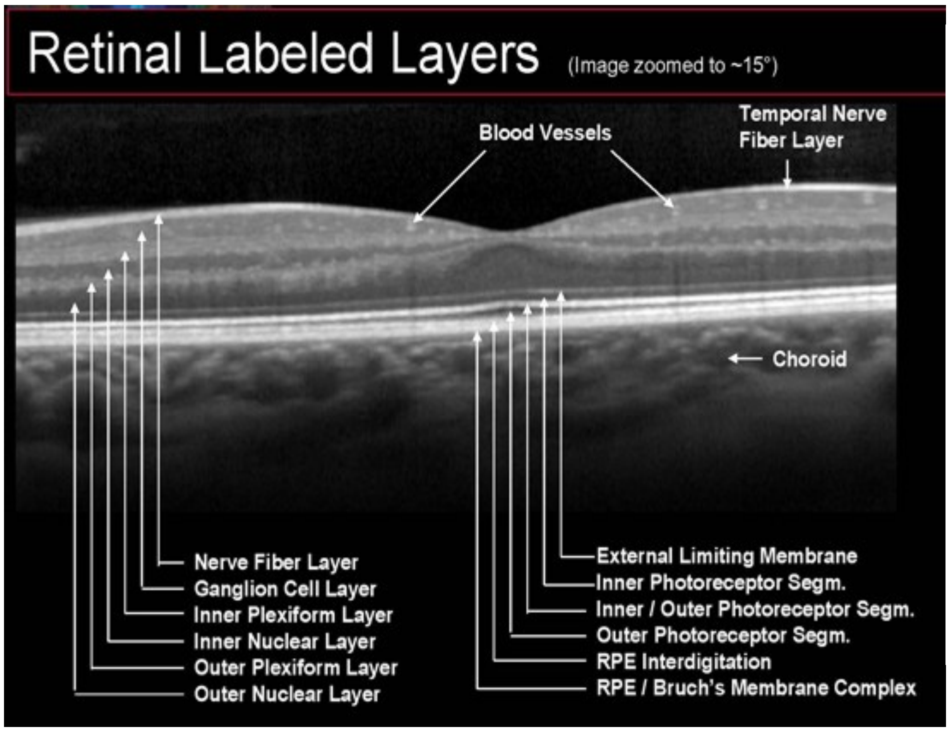
Only about 10% of the light entering the eye is absorbed by the photopigment

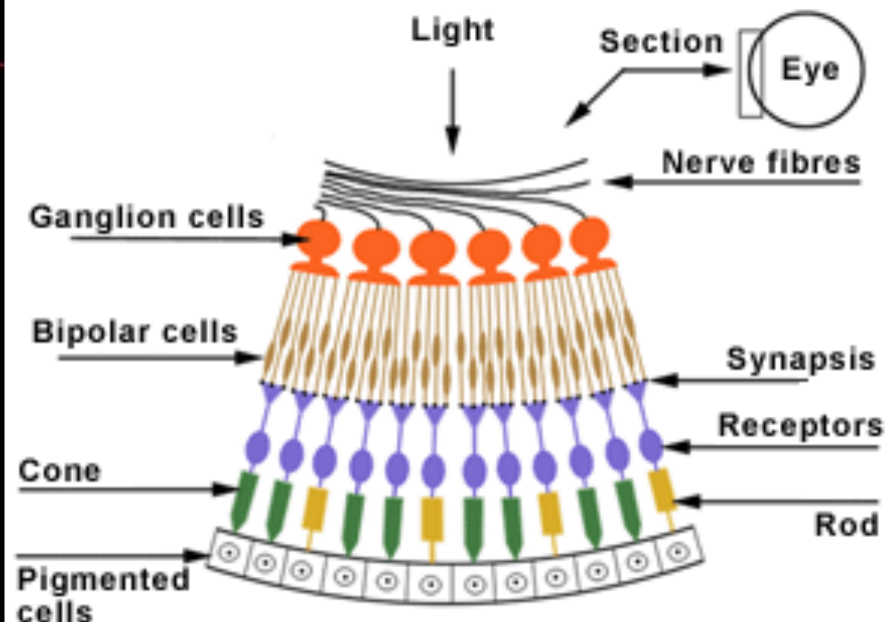
Optic Nerve



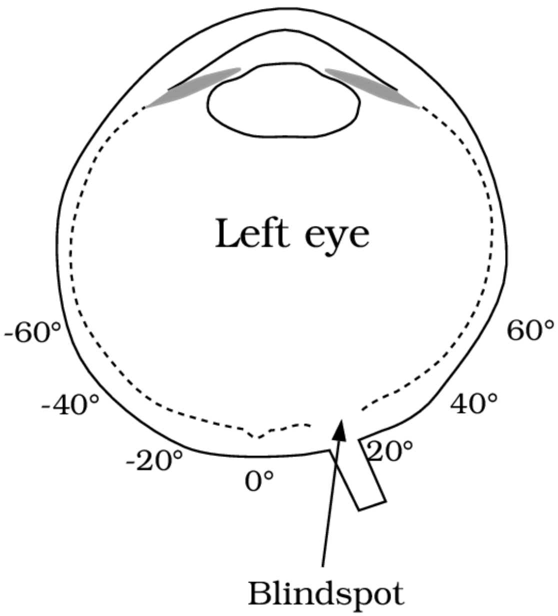
Retina
Eyeball
Optic nerve
Light
Light
Light
Ganglion cells
Bipolar neurons
Direction of neural impulse
Rod
Cone
Photoreceptor cells
Optic disc
Optic nerve (to brain)
Direction of neural impulse

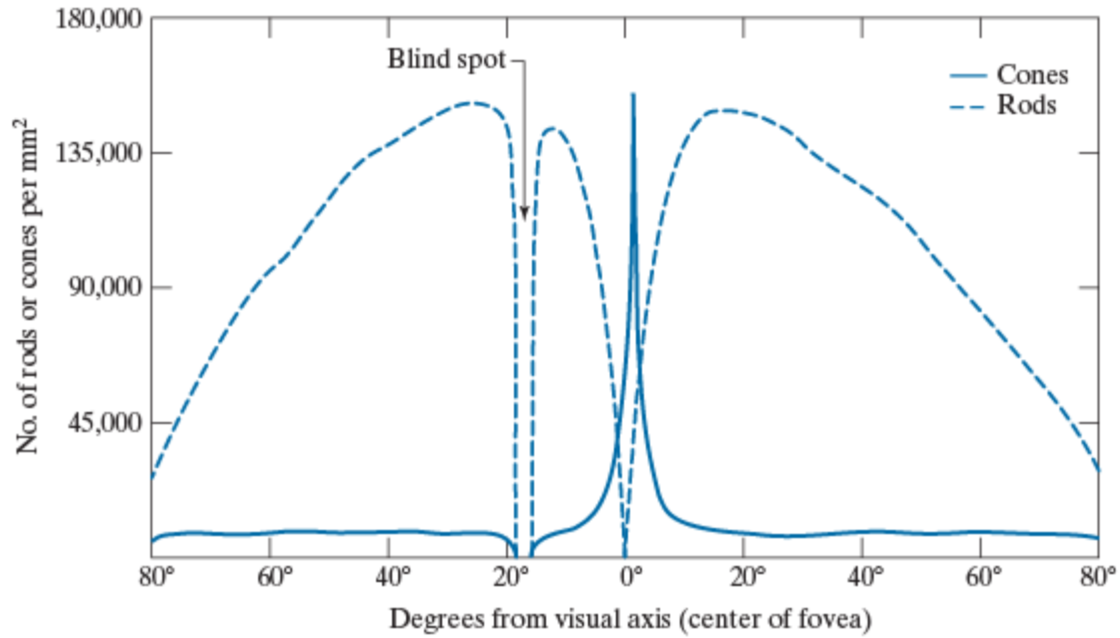
Cones and Rods





Distributions of Cones and Rods

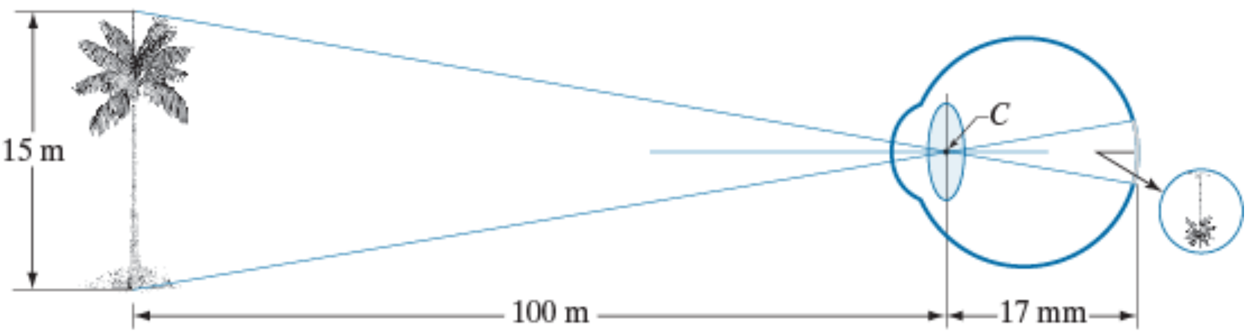




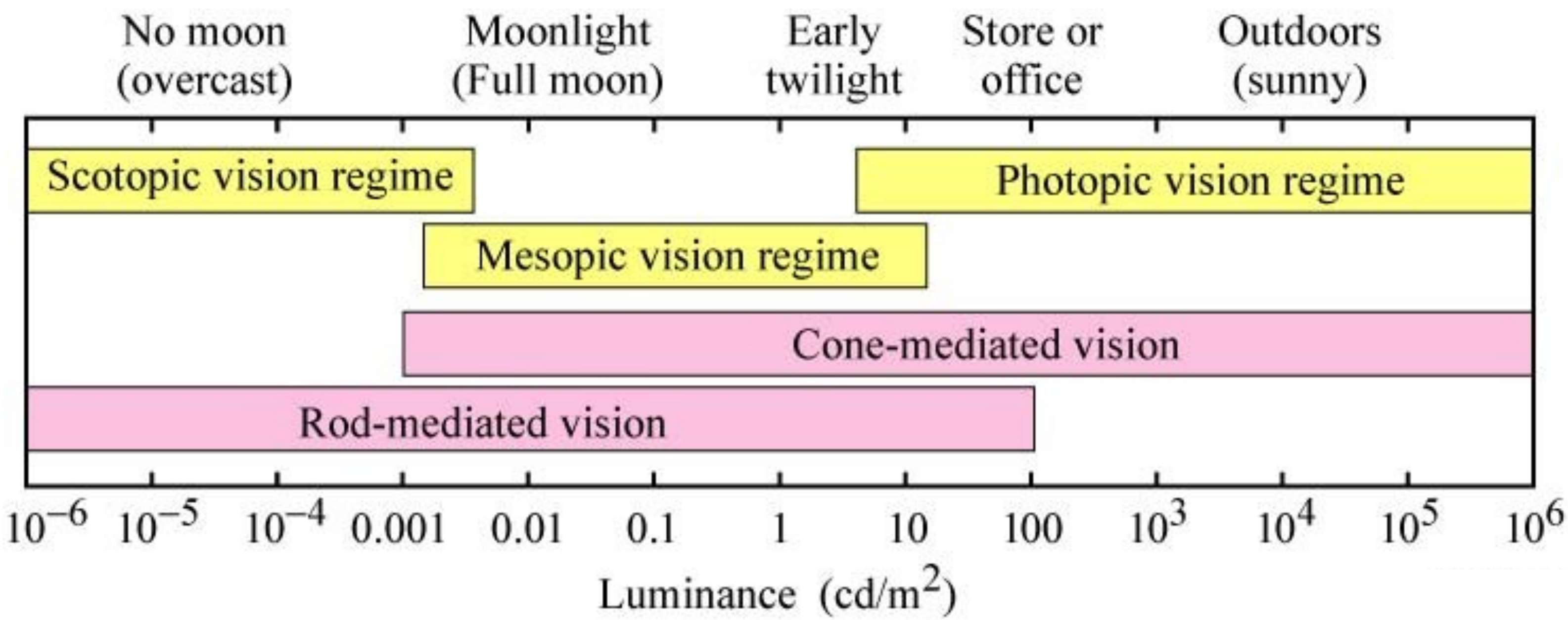
100 million rods and 5 million cones distributed very nonuniformly

Image Formation in the Eye

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



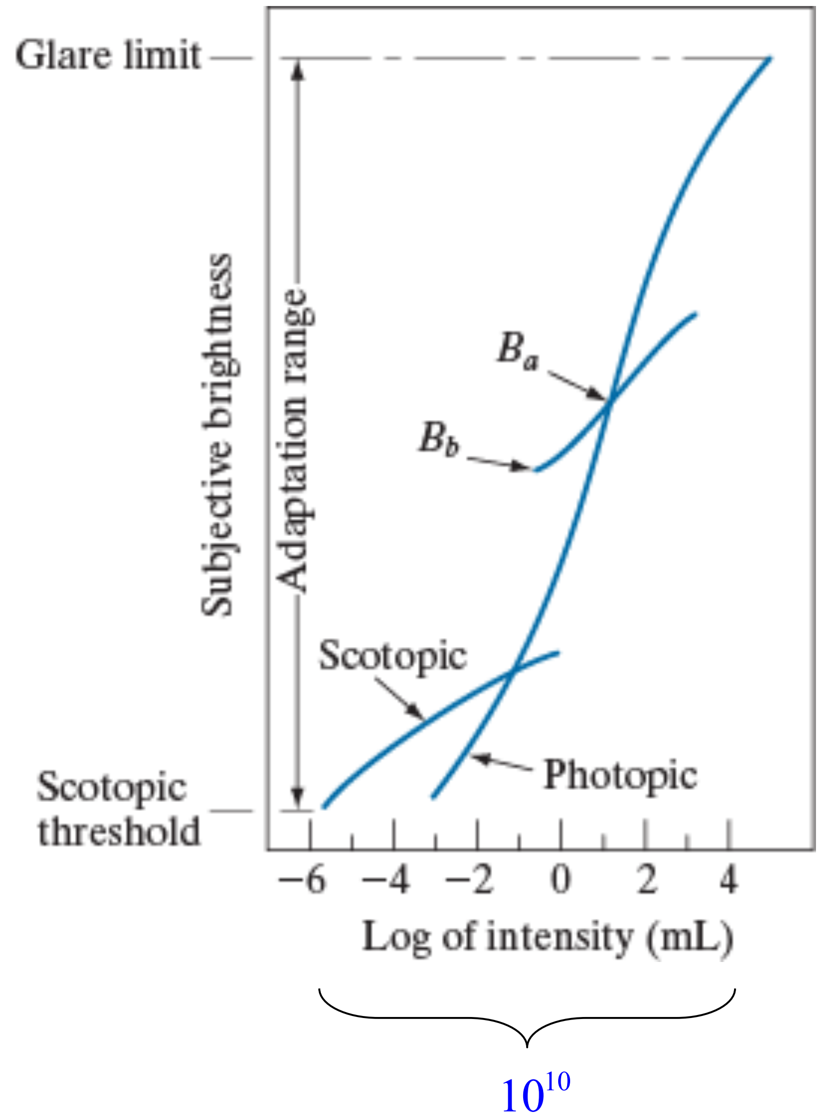
Human Vision



candela per square metre

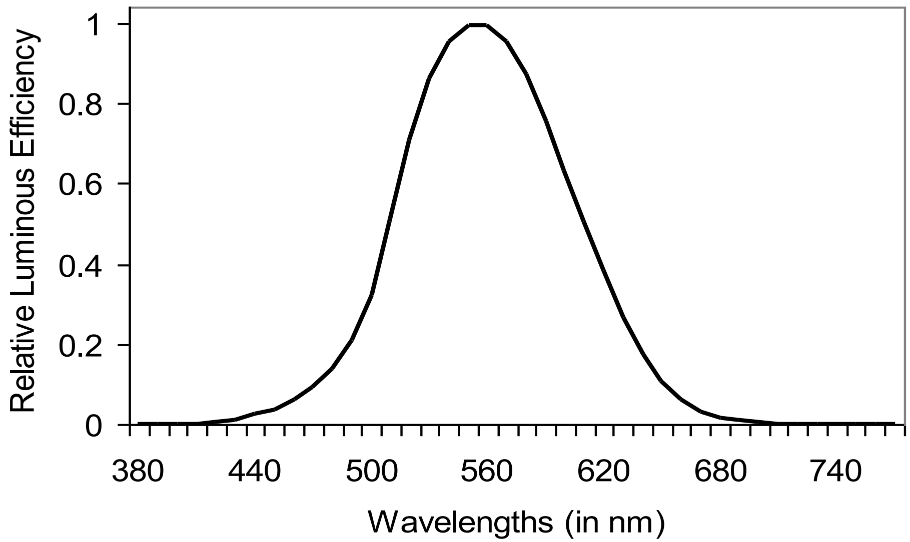
Brightness Adaptation

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



- The total range that human visual system can discriminate at a give time is rather small
- Brightness adaptation refers to our ability to adjust eye sensitivity over a wide range of adaptation levels.

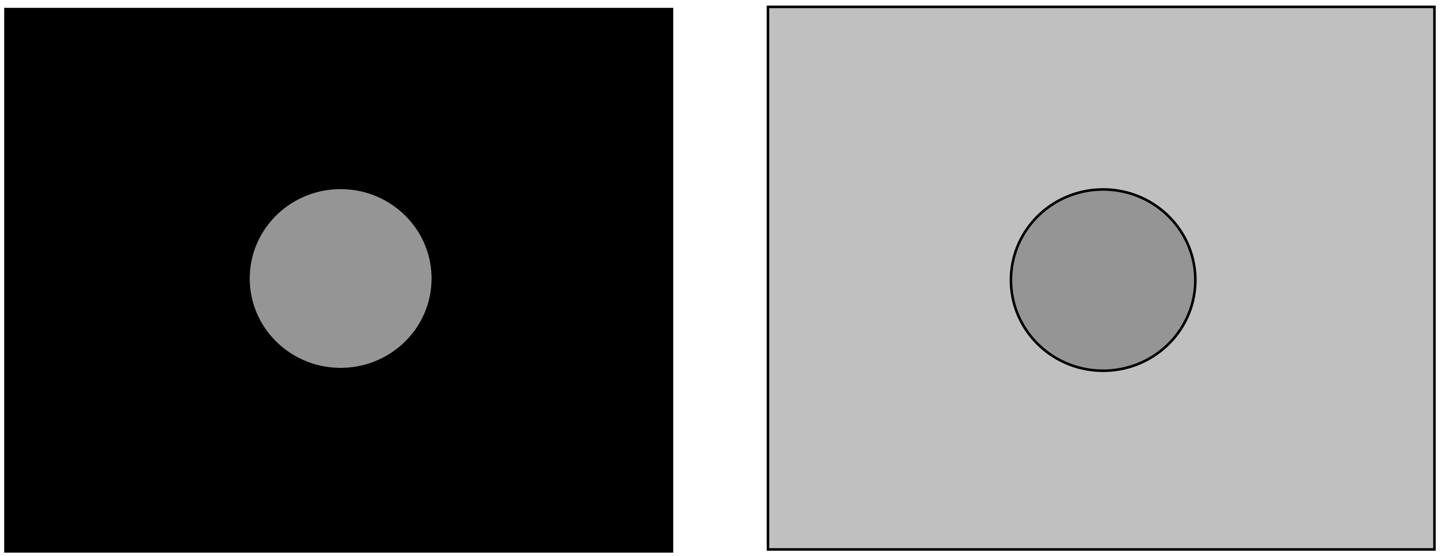
Relative Luminous Efficacy Function



- Luminous efficacy is a measure of how well a light source produces visible light. The eye acts as a **bandpass filter** to radiant energy.
- Not all wavelengths of light are equally visible, or equally effective at stimulating human vision, due to the spectral sensitivity of the human eye
- The peak response is obtained around **555 nm**
- Note: Some systems have the same units for luminous flux and radiant flux. Then the luminous efficacy of radiation is dimensionless and is often instead called the **luminous efficiency**.

Contrast

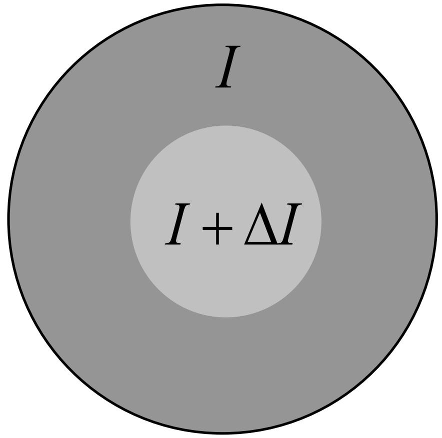
The **brightness** of an object is the perceived luminance and depends on the luminance of the surround.



Small circles in the middle of these two squares have equal luminance, but do not appear equally bright.

Brightness Discrimination Property

FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.



As ΔI gets stronger, the viewer may start to tell the brightness difference between the background and the foreground.

Weber's Law:

Our perception is sensitive to luminance contrast rather than the absolute luminace value. If a picture area has a high luminance value, a higher ΔI is required to notice the contrast between the two regions:

$$\frac{\Delta I}{I} = k \quad (k \text{ is known as the Weber constant and is about 0.1-0.2})$$

K代表人眼是否能区分(愈大愈能区分)

∴若I↑,則ΔI也要↑,才能区分

which can be rewritten as

$$c\frac{\Delta I}{I} = \Delta B$$

Therefore, we have

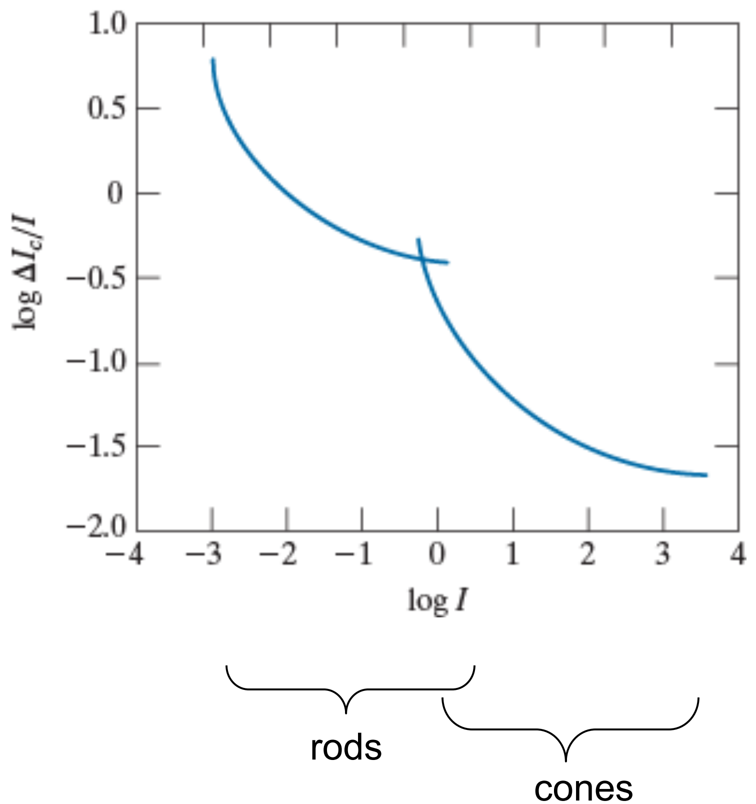
$$B = c \log I + d \qquad d\text{: a constant}$$

呈正比

Brightness is proportional to the logarithm of the luminance!

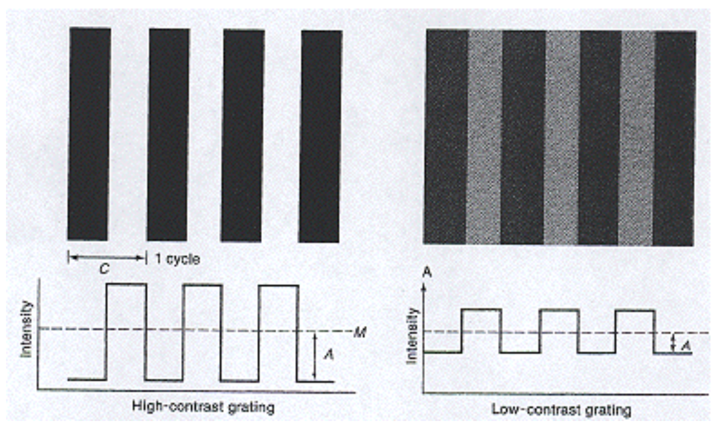
Weber Ratio

FIGURE 2.6
A typical plot of the Weber ratio as a function of intensity.

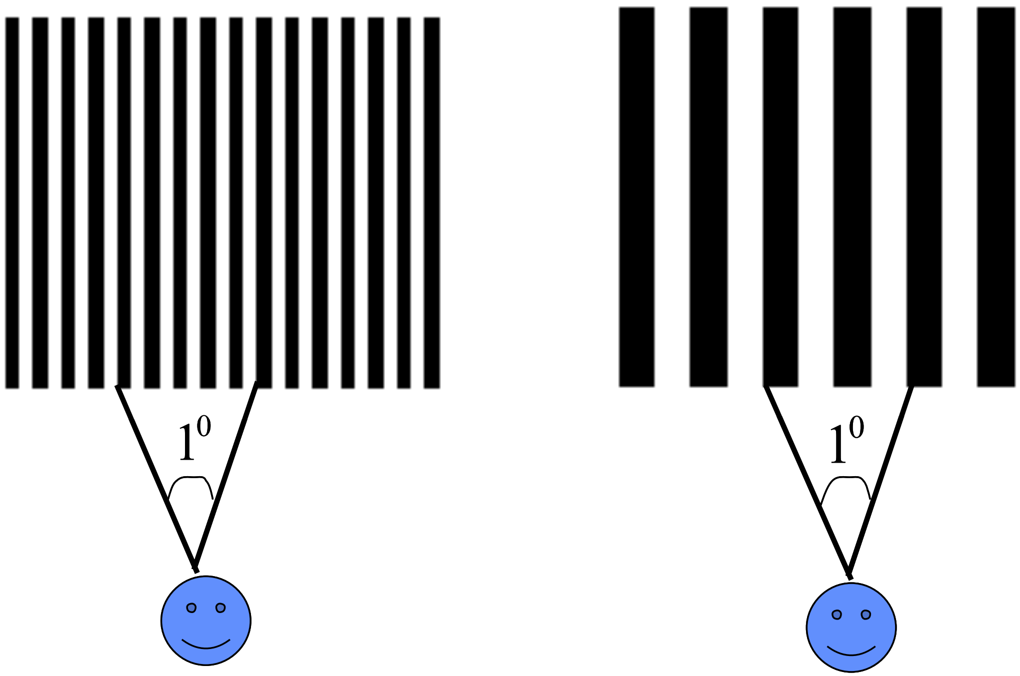


Question

How much contrast - how much difference between light and dark parts of objects - do we need to just see objects?



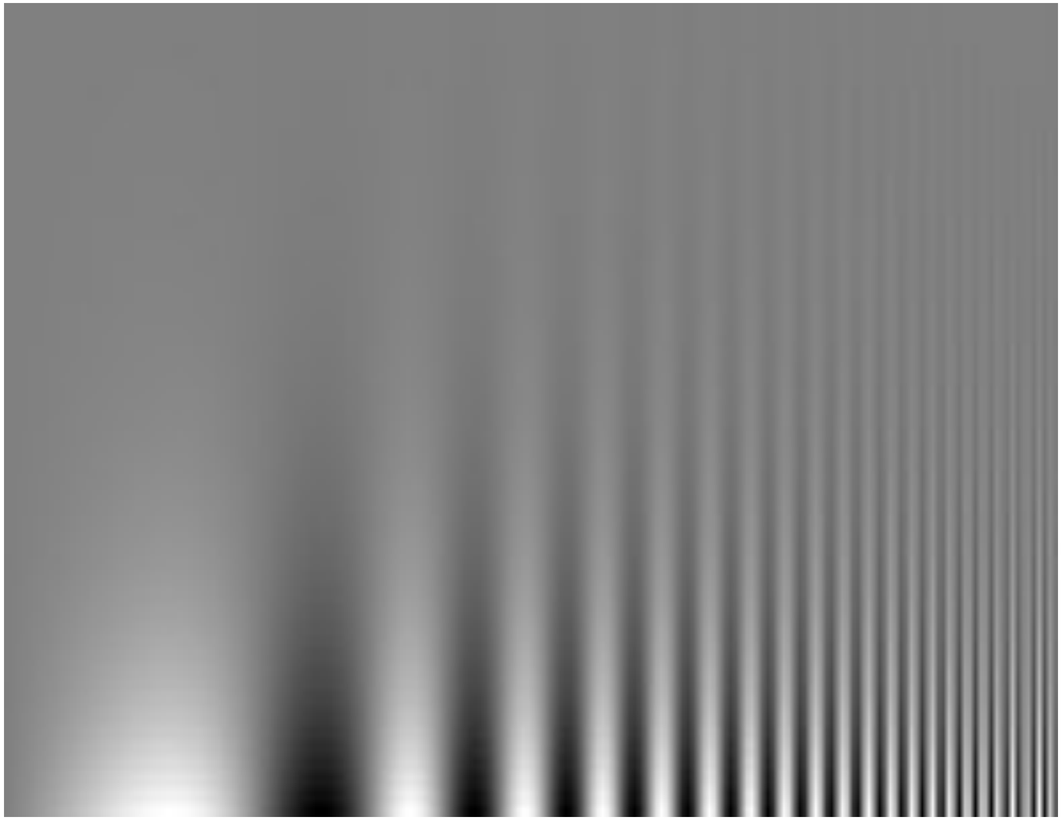
Square Wave Grating



- Square wave gratings
- The left grating (5 cycles/deg) has a higher frequency than the right (2 cycles/deg)

Sinusoidal Grating

- Sinusoidal grating of varying contrast and spatial frequency
- Frequency increases exponentially from left to right
- Amplitude decreases exponentially from bottom to top
- The sensitivity changes with the spatial frequency of the stripes
- The location of the highest sensitivity depends on the viewing distance

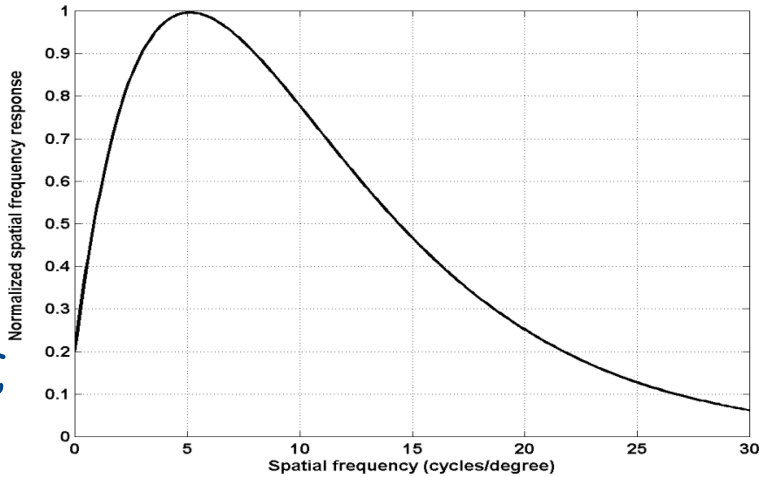


Contrast versus spatial frequency sinusoidal grating

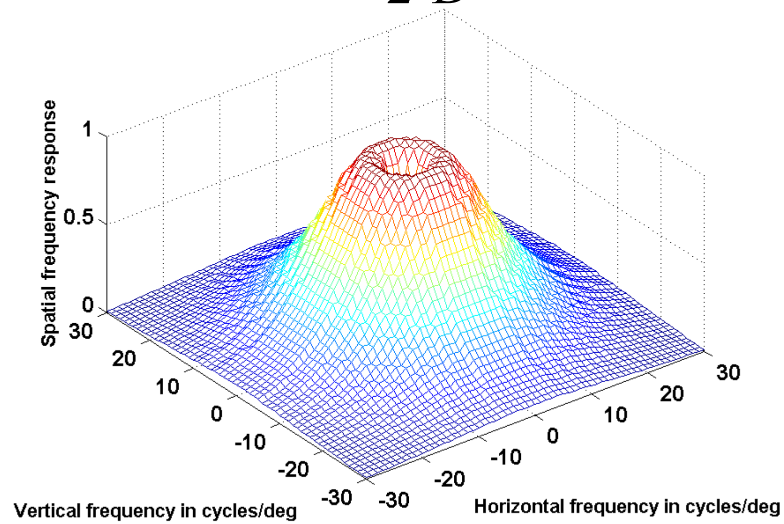
Frequency Sensitivity of HVS

1-D

(視覺反應)

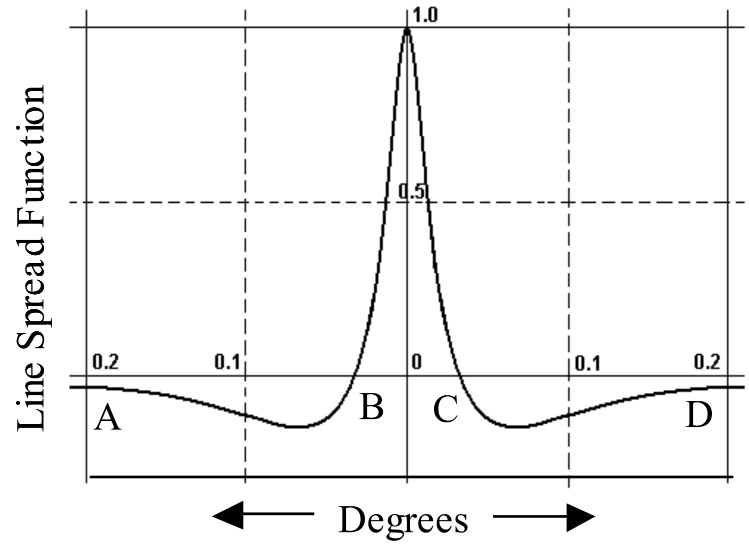


2-D



- The spatial frequency response of the eye is called the modulation transfer function (MTF)
- The inverse Fourier transform of the MTF is the impulse response of the system

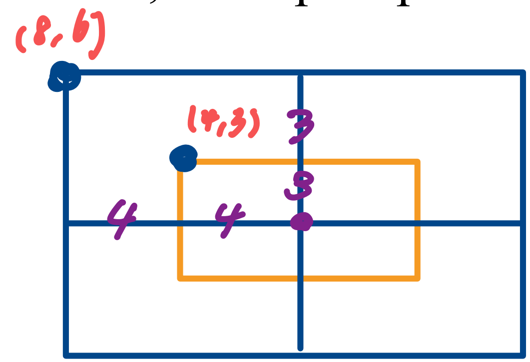
Line Spread Function of Human Eye



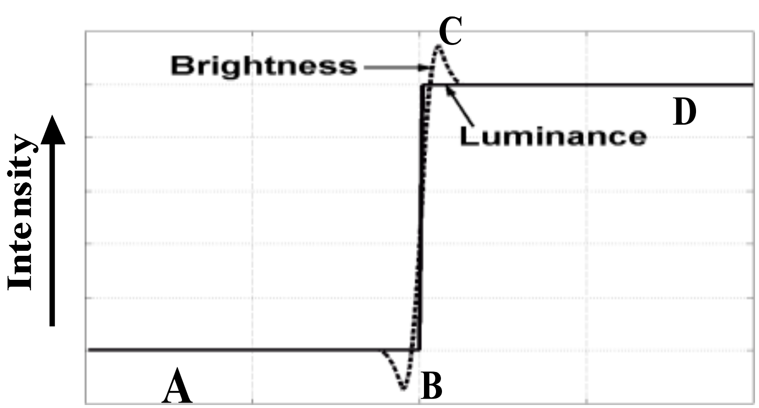
- The one-dimensional impulse response of the system is known as the line spread function of the system
- It is the response of the system to an infinitely thin light represented by $\delta(x)$ or $\delta(y)$
- Observations
    1. It has a finite width
    2. It has negative value (e.g. at 0.08°)

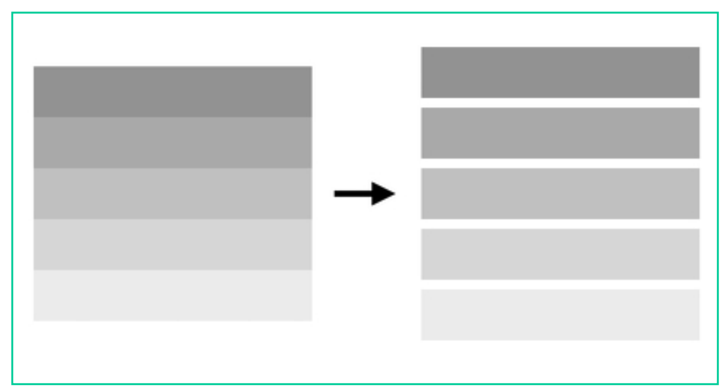
Mach Band Effect

- There is overshoot and undershoot in the middle boundary as we look at the image below
- Assuming the HVS is linear and time invariant, its step response is the integration of its impulse response.



Step function

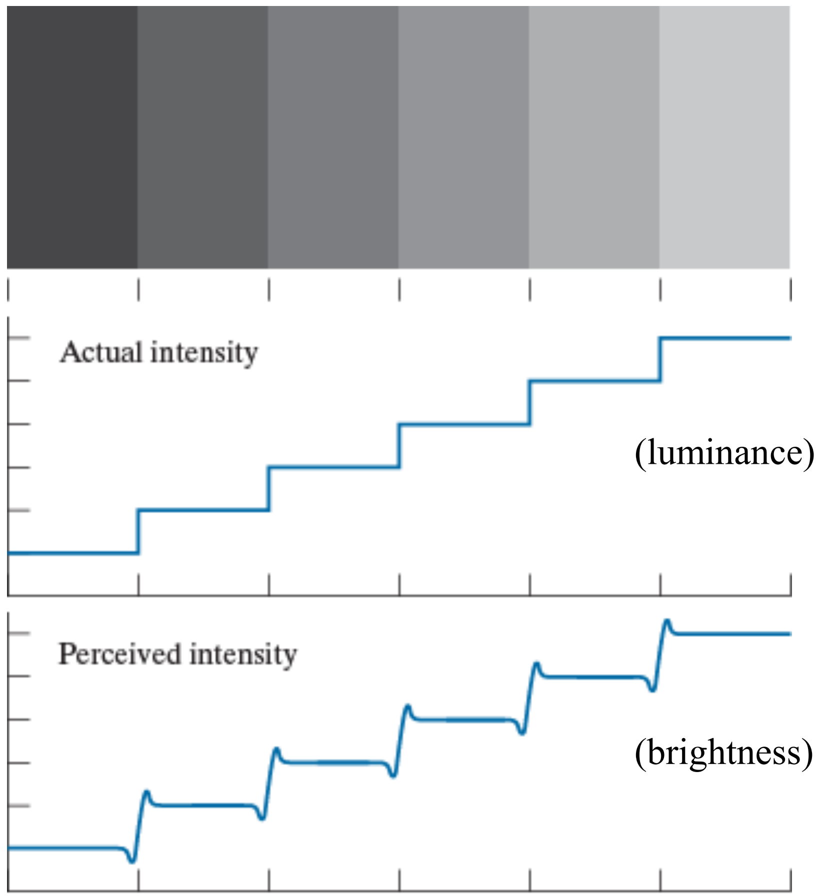




Mach Band Effect (cont'd)

a
b
c

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

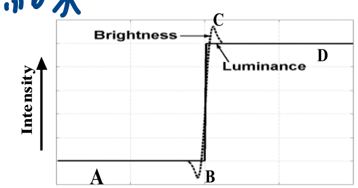


$$x(t) \rightarrow \boxed{h(t)} \rightarrow y(t)$$

$$y(t) = x(t) \underset{=}{*} h(t)$$

convolution

結果:



Simultaneous Contrast

(同時放一起比較)

A region's perceived brightness does not depend simply on its 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.

The term "contrast" refers to the fact that **the perceived color of the surfaces is "contrasted" by the color of the surround**.
The term "simultaneous" was introduced by Chevreul to "distinguish this phenomenon to the 'successive' contrast, where two colors appear in succession upon the same retinal area." [H. V. Helmholtz]

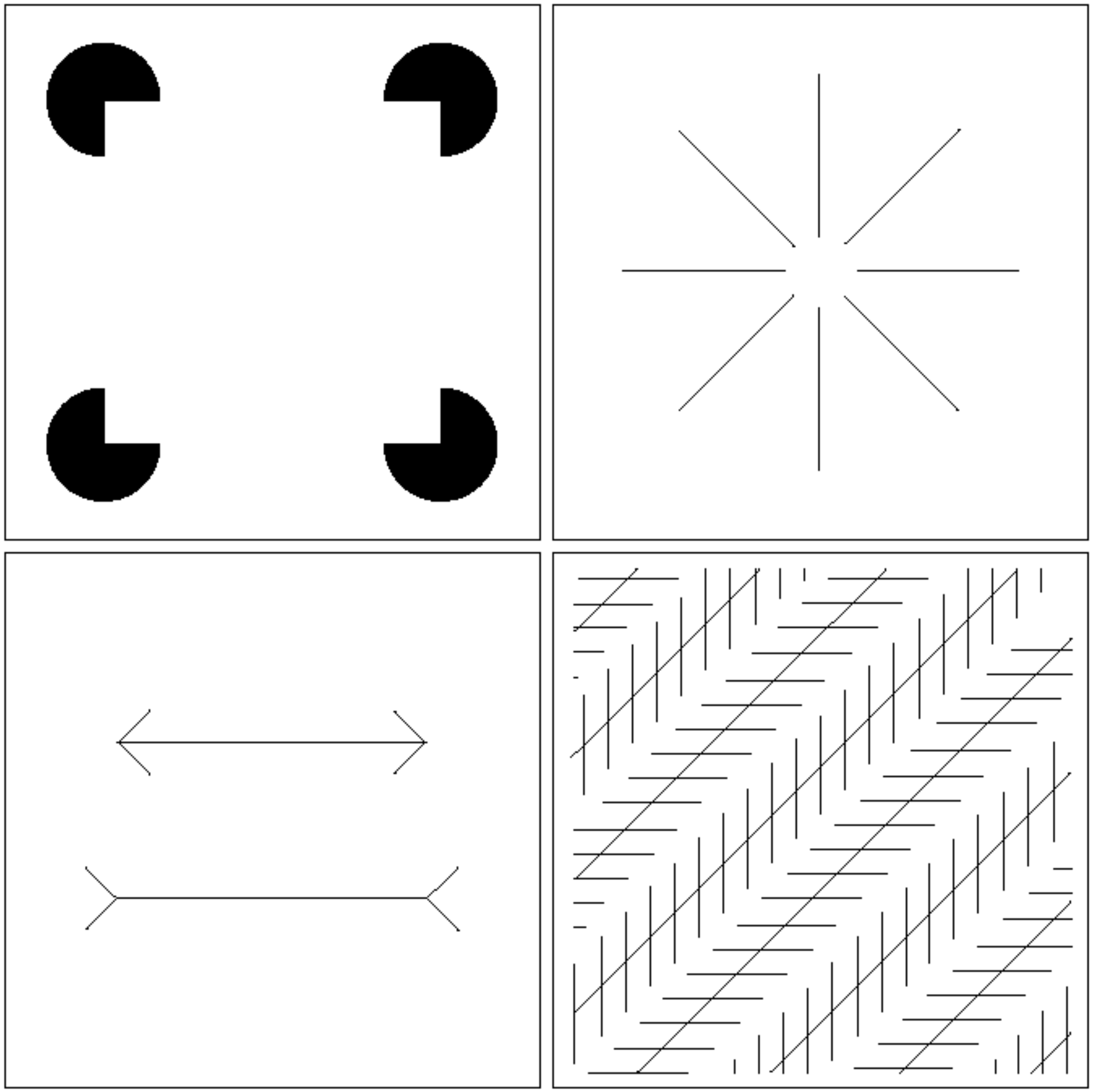
Simultaneous Contrast (cont’d)

… the tendency of a color to induce its opposite in hue, value and intensity upon an adjacent color and be mutually affected in return by the law of simultaneous contrast a light, dull red will make an adjacent dark, bright yellow seem darker, brighter and greener; in turn, the former will appear lighter, duller and ... [merriam-webster.com]

Two colors, side by side, interact with one another and change our perception accordingly. Simultaneous contrast is most intense when the two colors are complementary colors.

http://www.webexhibits.org/colorart/contrast.html
https://www.youtube.com/watch?v=rnVhD-ke9Qg

Optical Illusion

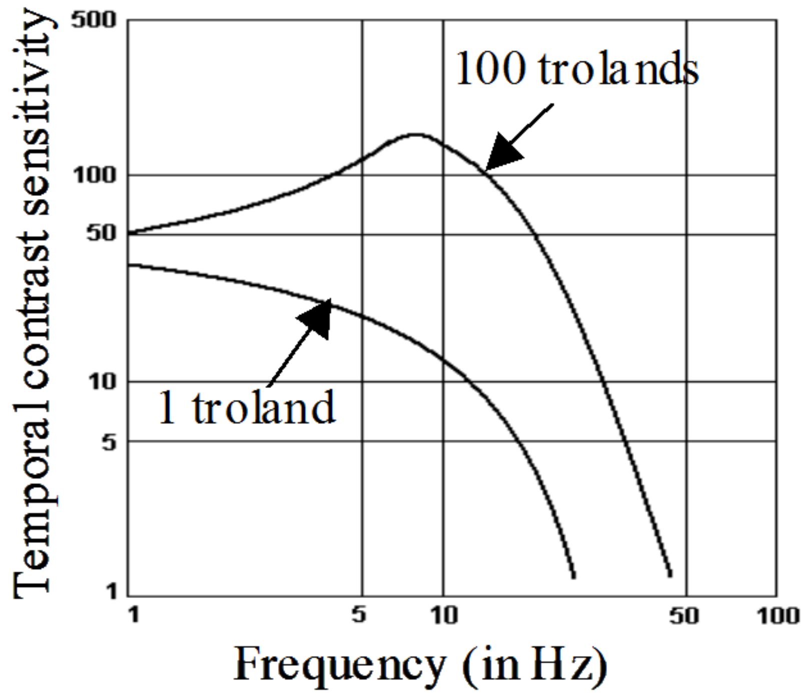


a b
c d

FIGURE 2.9 Some well-known optical illusions.

Temporal Contrast Sensitivity

- Test stimulus: $L+\Delta L\cos(2\pi f t)$
- For any given value of f, the value of ΔL that produces a threshold sensation of flicker is denoted ΔL_T
- Contrast sensitivity: $1/\Delta L_T$
- 100 trolands is equivalent to 10 cd/m^2
- The eye is more sensitive to flicker at high luminance
- Flicker sensitivity is negligible above 50/60 Hz



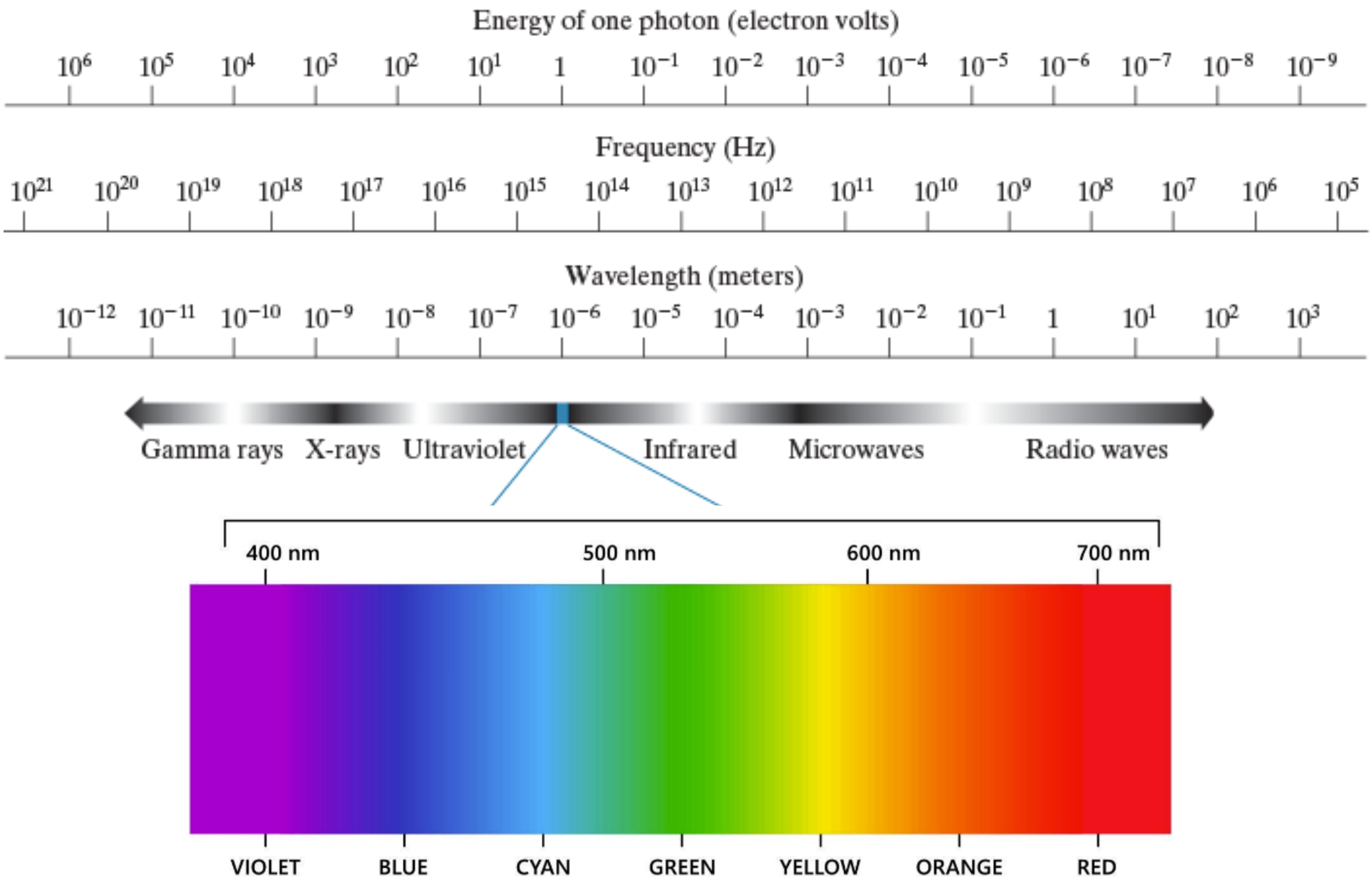
Flicker sensitivity of the eye for a 2° field at 1 troland and 100 trolands illuminance levels.

亮度低，对flicker的感受低，所以频率可以低

ex. 电影院是24Hz 但手机要60Hz↑

cd/m^2 (SI): candela per square meter. A non-SI term for the same unit is the nit.
stilb (CGS): one cd per square centimeter, or 10 kcd/m^2.

Light and EM Spectrum



$$c = \lambda \nu \qquad E = h\nu, \quad h: \text{ Planck's constant.}$$

Light and EM Spectrum

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

  For example, green objects reflect light with wavelength primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths

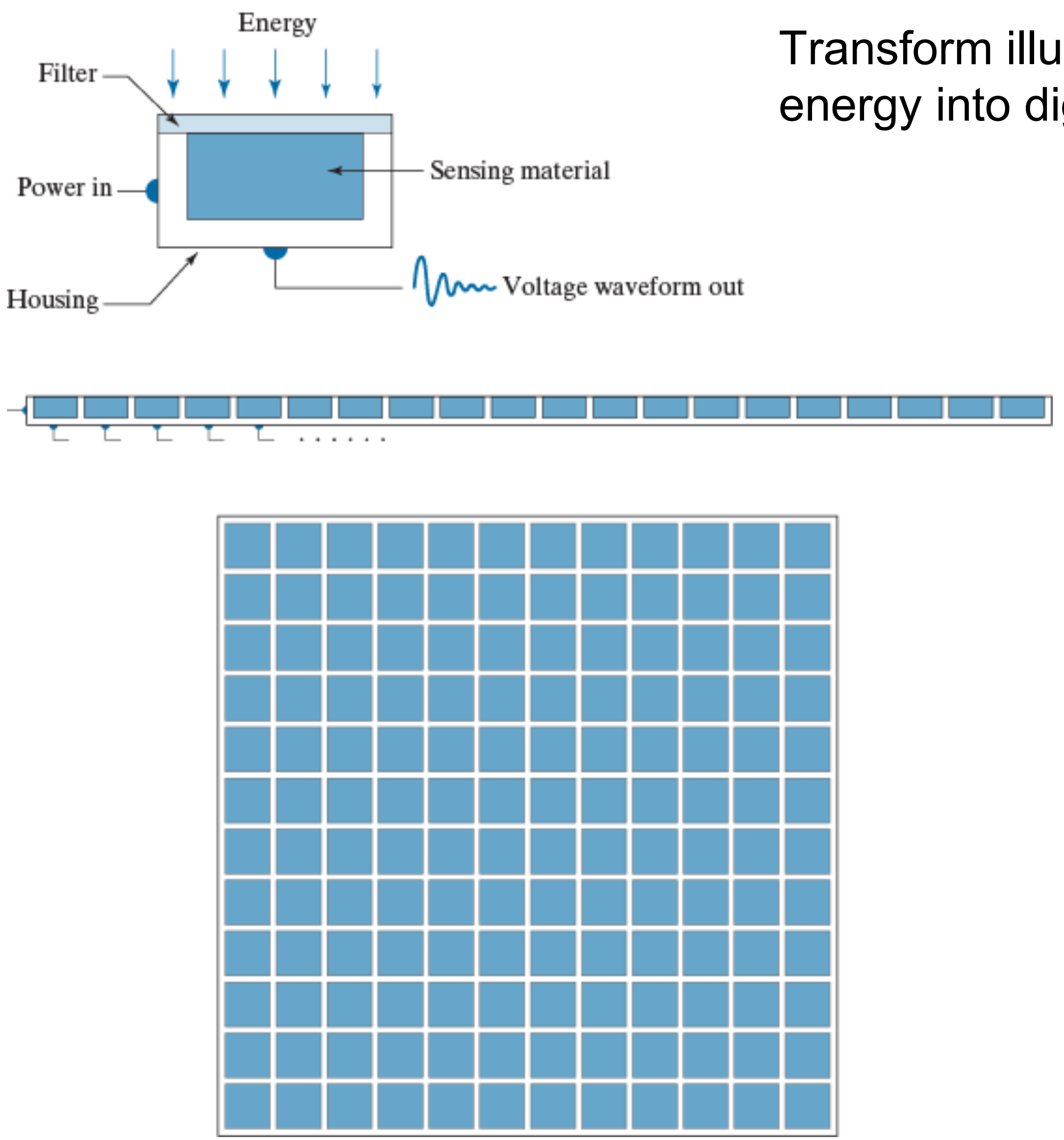
Light and EM Spectrum (cont'd)

- Monochromatic light
  - **Intensity** is the only attribute, from black to white
  - Monochromatic images are referred to as **gray-scale** images

- Chromatic light bands
  - 430 to 790 nm
  - **Radiance**: total amount of energy
  - **Luminance**: 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.

Single, Linear, and Areal Image Sensor

Transform illumination energy into digital images



a
b
c

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

Image Acquisition Using a Single Sensor

FIGURE 2.13
Combining a single sensing element with mechanical motion to generate a 2-D image.

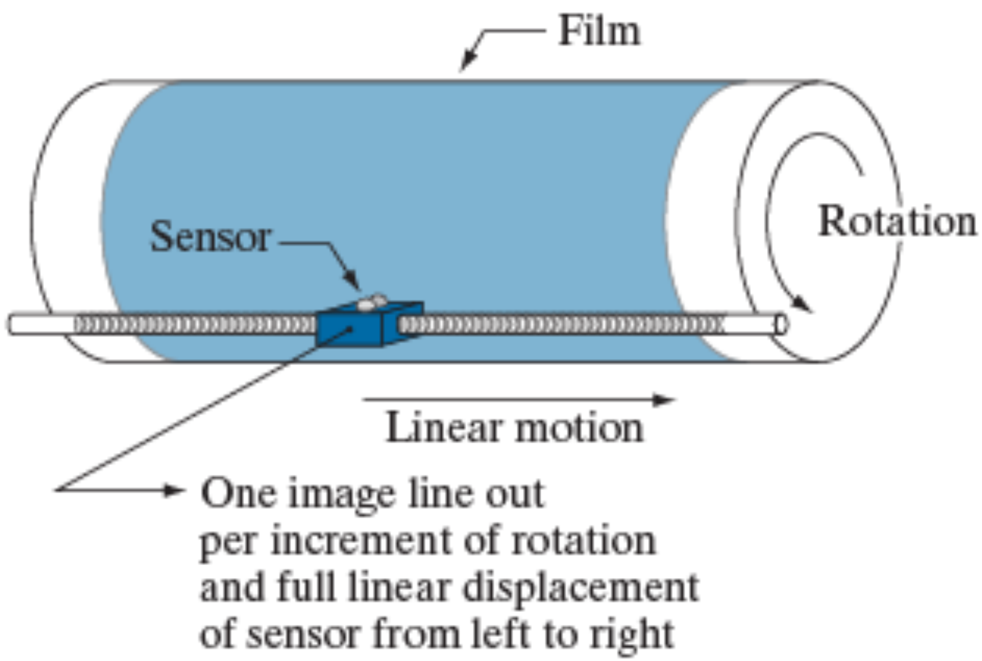
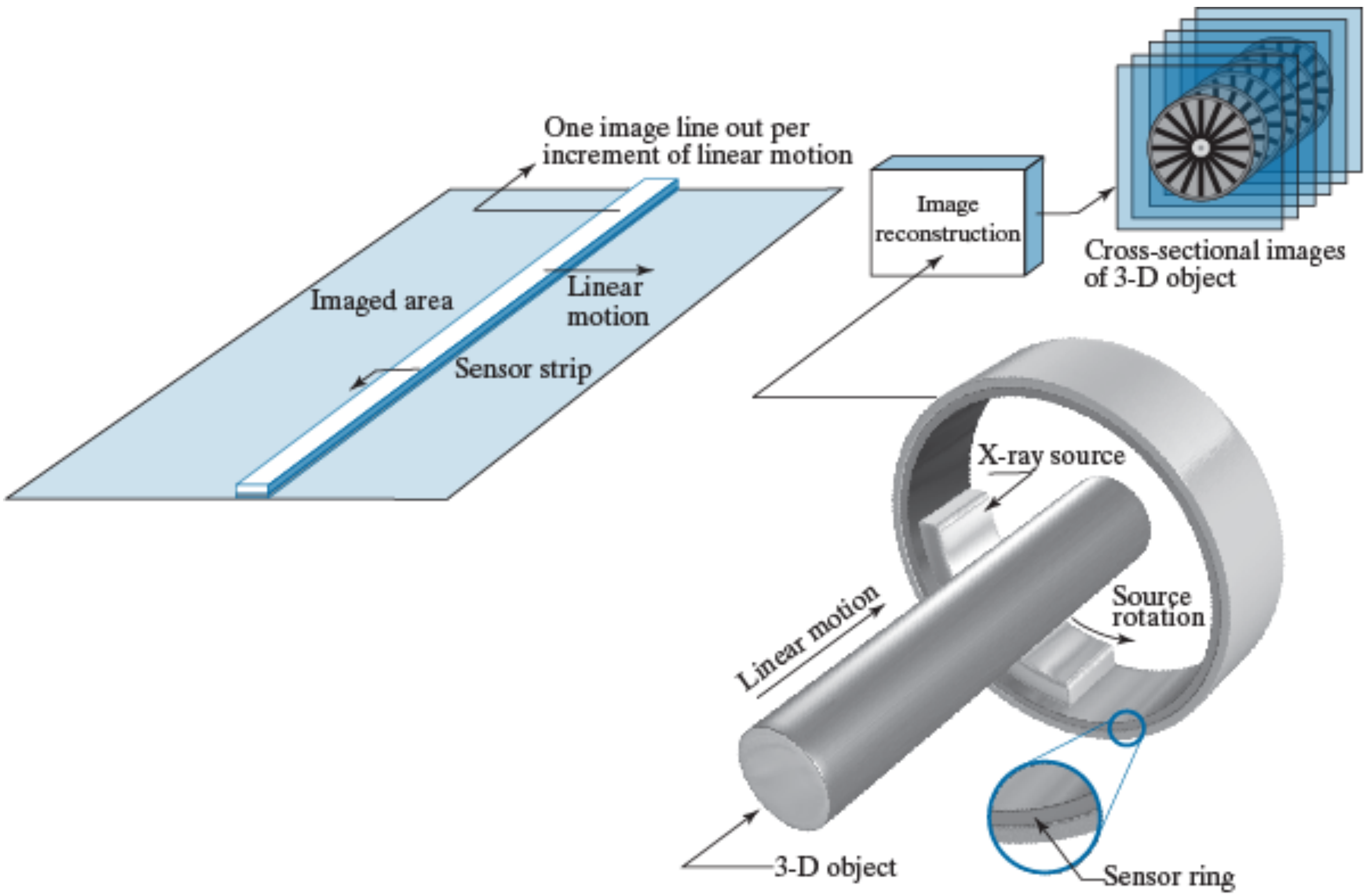


Image Acquisition Using Sensor Strips



a b

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

Image Acquisition Process

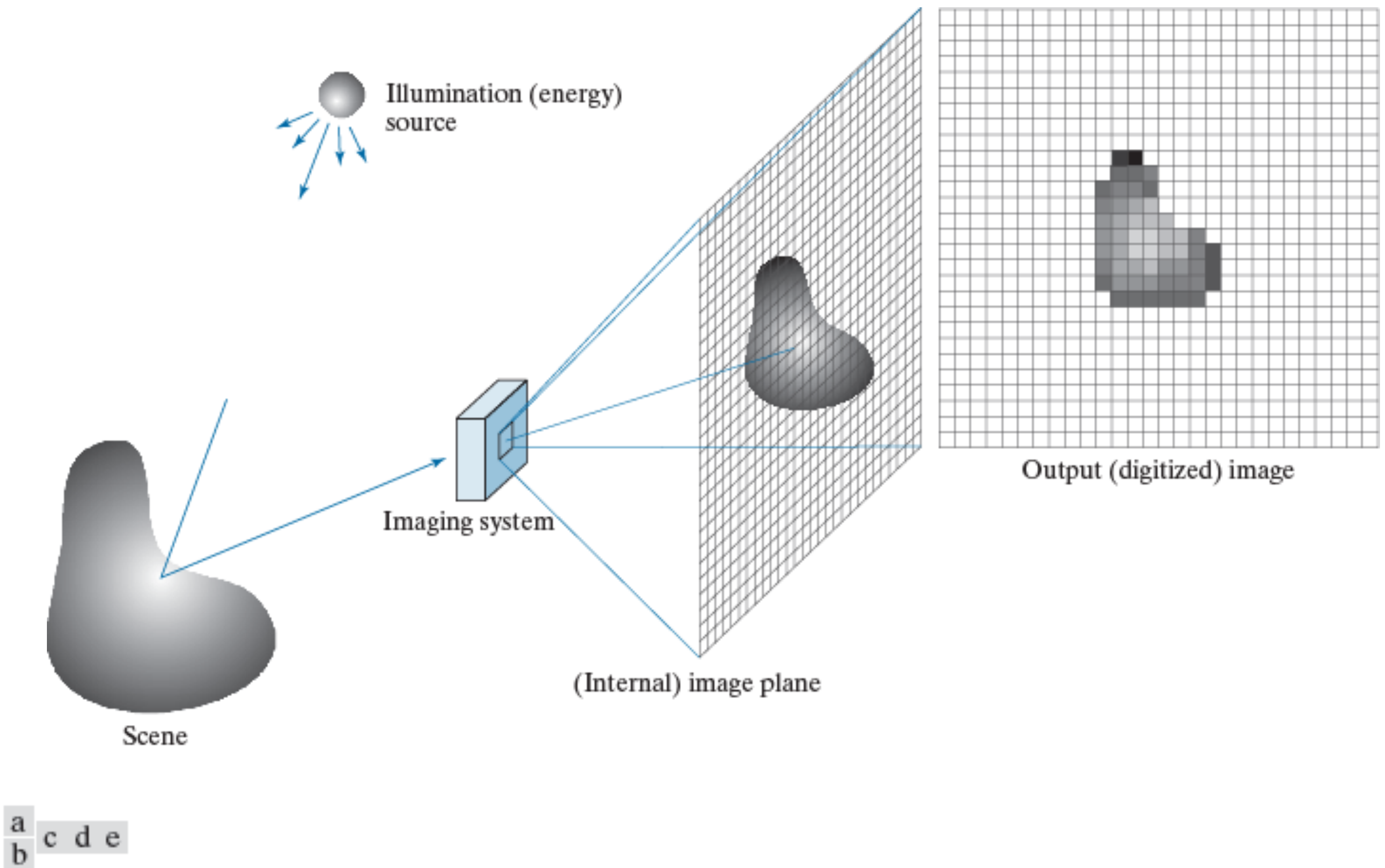


FIGURE 2.15 An example of digital image acquisition. (a) Illumination (energy) source. (b) A scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

A Simple Image Formation Model

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

$f(x, y)$: Image intensity at (x, y), $0 < f(x, y) < \infty$

$i(x, y)$: Amount of illumination incident on object at (x, y),

$0 < i(x, y) < \infty$

$r(x, y)$: Amount of illumination reflected by object at (x, y),

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

Some Typical Ranges of 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

- 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

Image Sampling and Quantization

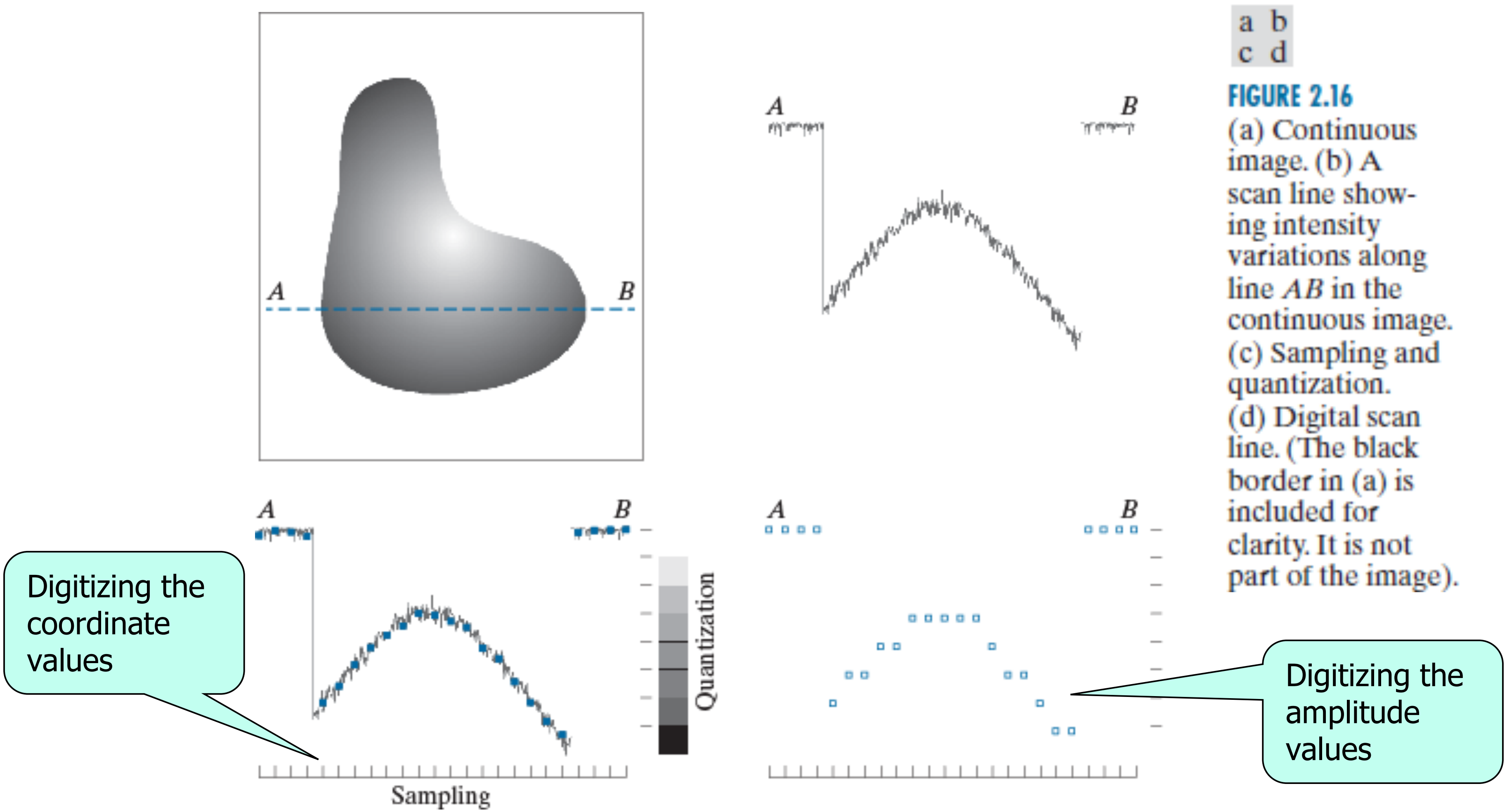
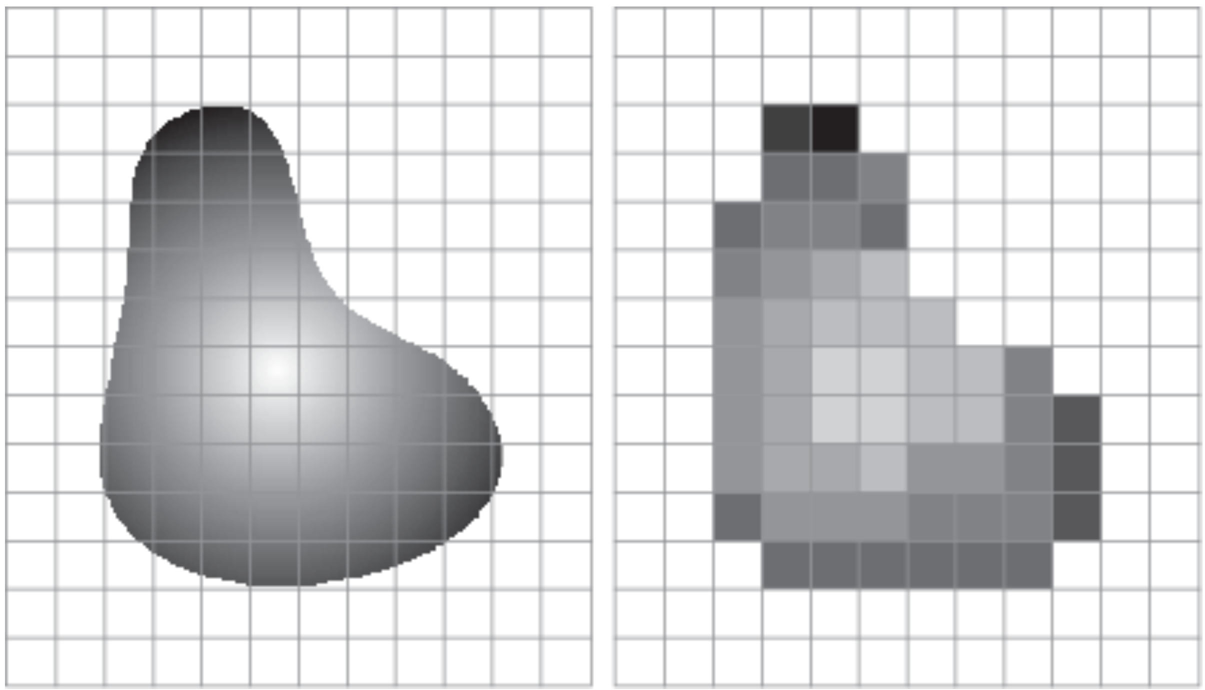


FIGURE 2.16
(a) Continuous image. (b) A scan line showing intensity variations along line *AB* in the continuous image. (c) Sampling and quantization. (d) Digital scan line. (The black border in (a) is included for clarity. It is not part of the image).

Sampled and Quantized Image



a b

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

Representing Digital Images

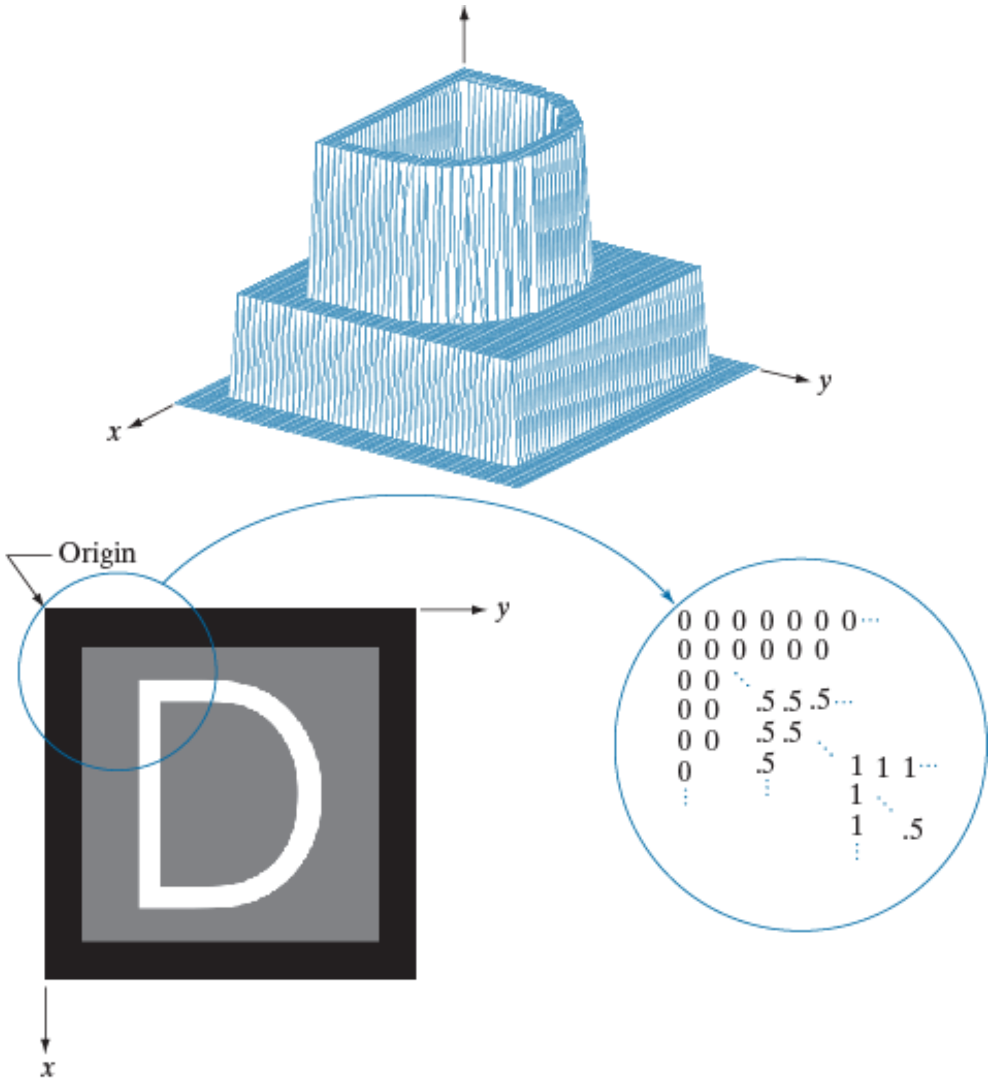
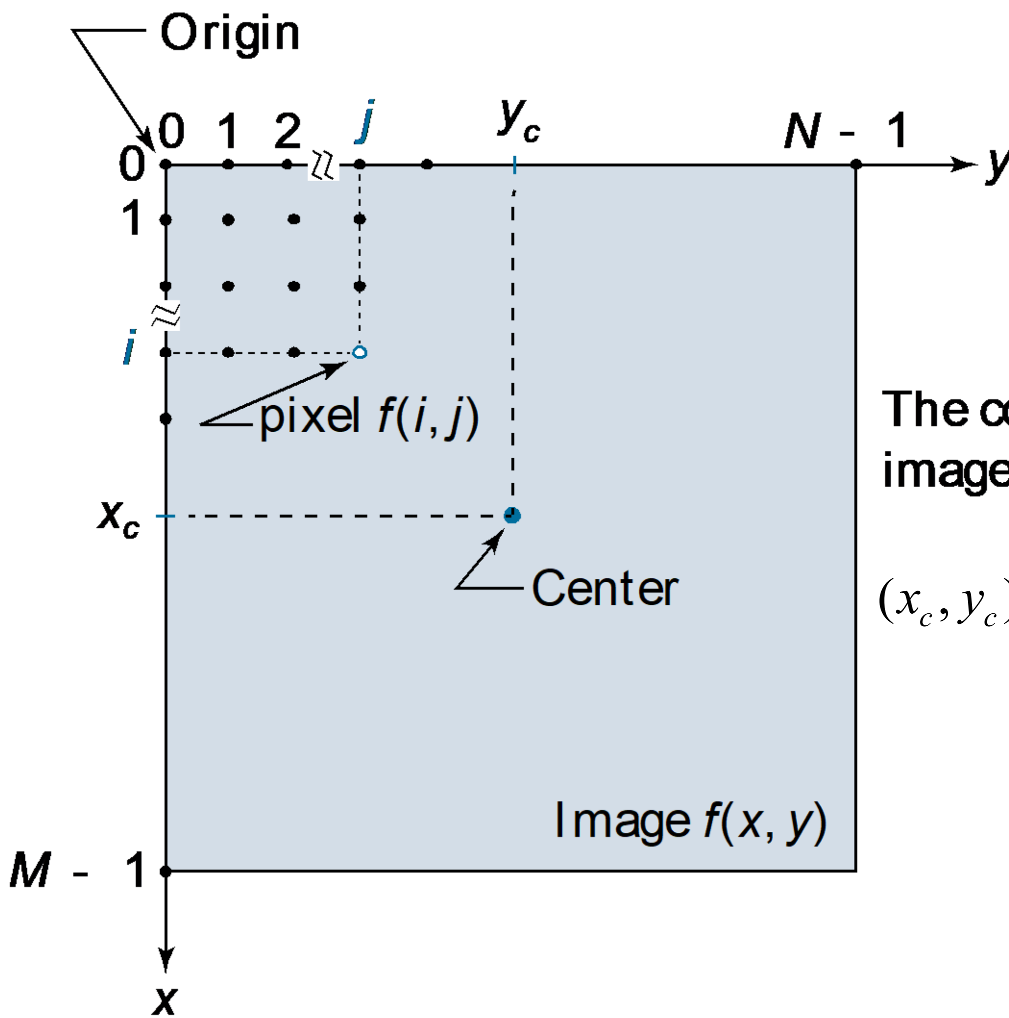


FIGURE 2.18
(a) Image plotted as a surface. (b) Image displayed as a visual intensity array. (c) Image shown as a 2-D numerical array. (The numbers 0, .5, and 1 represent black, gray, and white, respectively.)

Coordinate Cenvention

FIGURE 2.19 Coordinate convention used to represent digital images. Because coordinate values are integers, there is a one-to-one correspondence between x and y and the rows (r) and columns (c) of a matrix.



The coordinates of the image center are

$$(x_c, y_c) = \left(\text{floor}(\frac{M}{2}),\ \text{floor}(\frac{N}{2}) \right)$$

Representing Digital Images (cont'd)

- By an M×N numerical array

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

- In a traditional matrix form

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

Representing Digital Images (cont'd)

- In MATLAB

$$f(x,y)=\begin{bmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \dots & \dots & \dots & \dots \\ f(M,1) & f(M,2) & \dots & f(M,N) \end{bmatrix}$$

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

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

Representing Digital Images (cont'd)

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

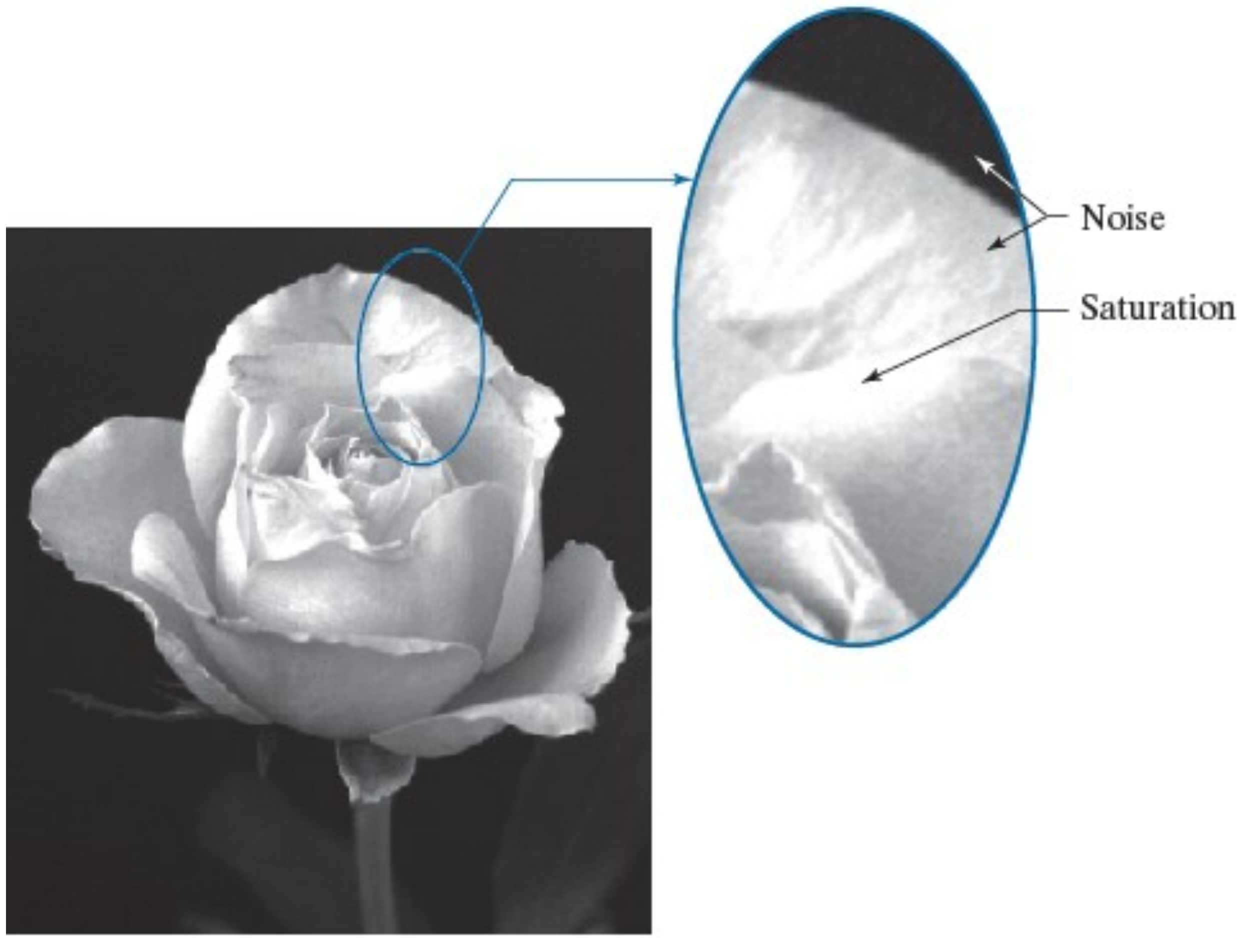
- Intensity resolution
  - — The smallest discernible change in intensity level
  - — *8 bits, 12 bits, 16 bits, etc.*

Dynamic Range and Contrast Measurements

- Dynamic Range 最亮的和最暗的 range
  - — The ratio of the maximum measureable intensity to the minimum detectable intensity level
  - — Upper limit determined by saturation and lower limit by noise
- Image contrast
  - — Difference in intensity between the highest and lowest intensity levels in an image
- Contrast ratio 还是和 Dynamic Range 不太一樣
  - — The ratio of the highest and lowest intensity levels

Saturation and Noise

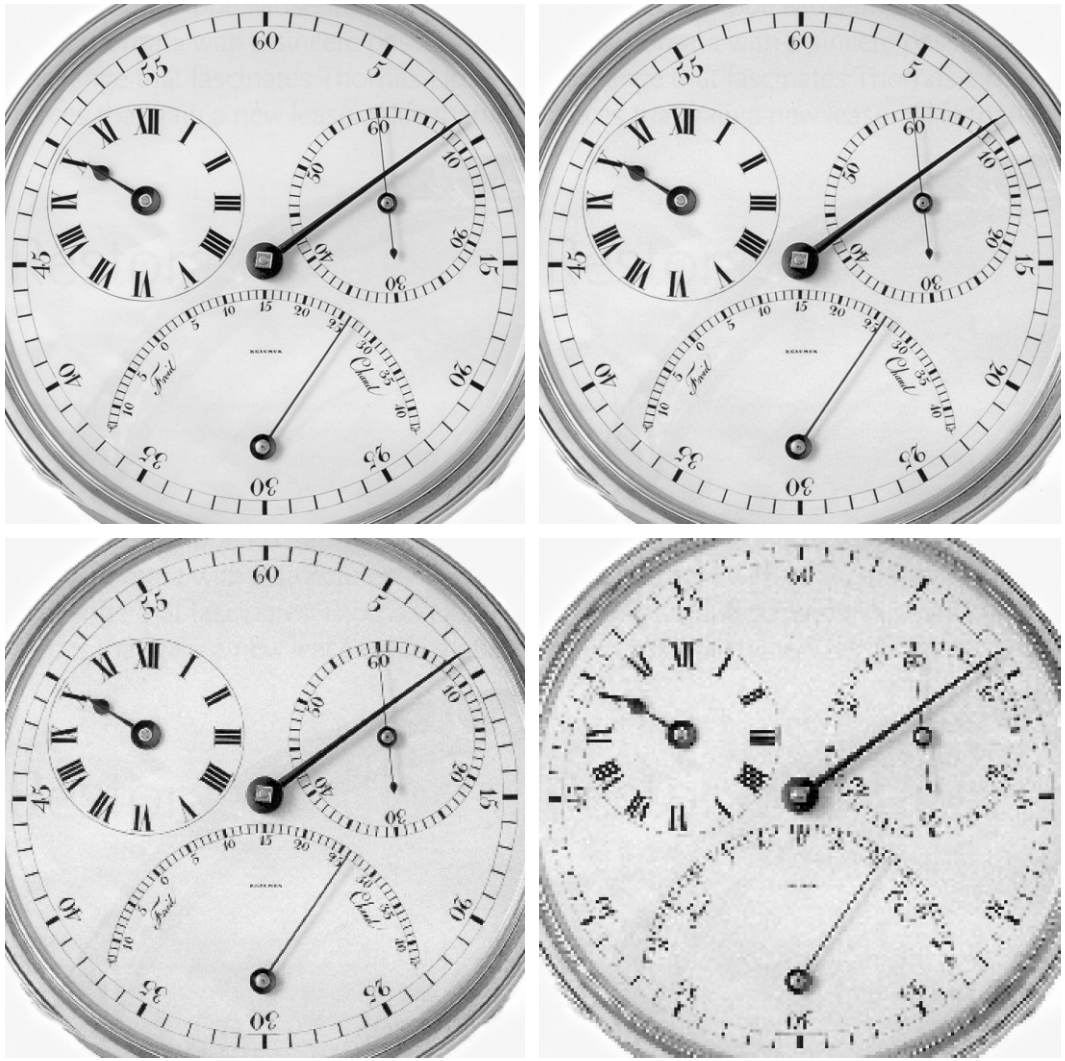
FIGURE 2.20 An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity values are clipped (note how the entire saturated area has a high, constant intensity level). Visible noise in this case appears as a grainy texture pattern. The dark background is noisier, but the noise is difficult to see.



Effect of Spatial Resolution Reduction

a b
c d

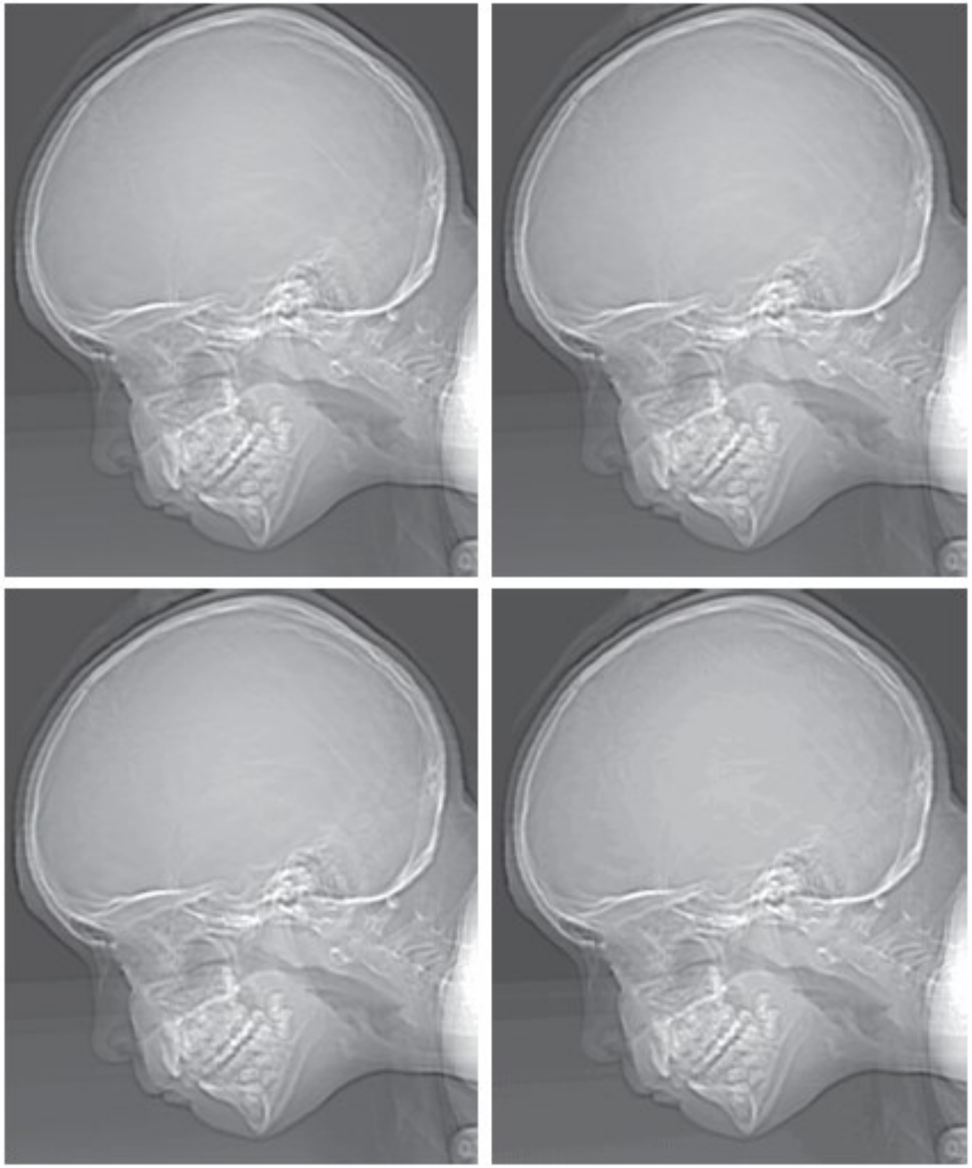
FIGURE 2.23 Effects of reducing spatial resolution. The images shown are at:
(a) 930 dpi,
(b) 300 dpi,
(c) 150 dpi, and
(d) 72 dpi.



Effect of Intensity Resolution Reduction

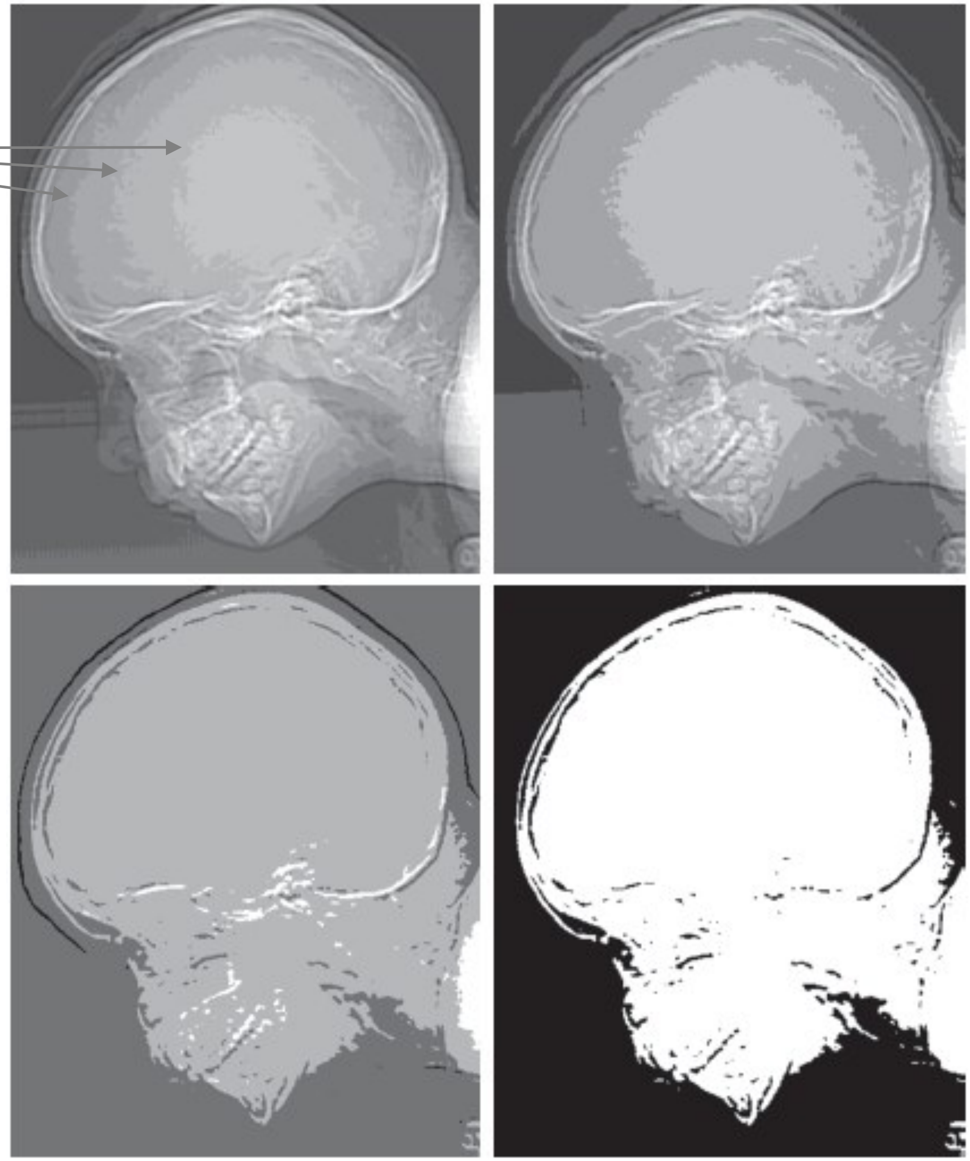
a b
c d

FIGURE 2.24
(a) 774×640, 256-level image. (b)-(d) Image displayed in 128, 64, and 32 intensity levels, while keeping the spatial resolution constant. (Original image courtesy of the Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Effect of Intensity Resolution Reduction (cont'd)

False contours



e f
g h

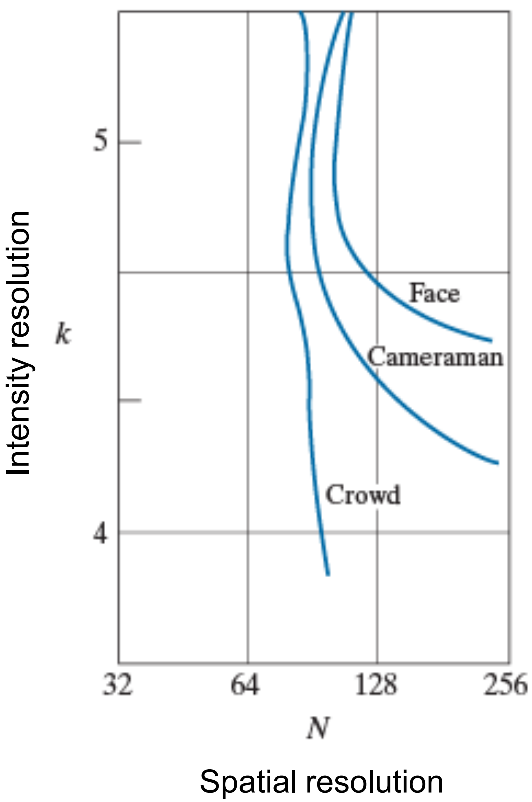
FIGURE 2.24 (*Continued*) (e)-(h) Image displayed in 16, 8, 4, and 2 intensity levels.

Isopreference Curves

Vertical part: For images with a large amount of details, only a few intensity levels may be needed.

N↑ k↓ part: A decrease in k increases the apparent contrast

FIGURE 2.26 Representative isopreference curves for the three types of images in Fig. 2.25.



a b c

FIGURE 2.25 (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.)

Image Interpolation

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

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

- **Image resampling** — Application of interpolation to image shrinking or zooming

  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

ex. warping 扭曲 (ex. 變笑臉)

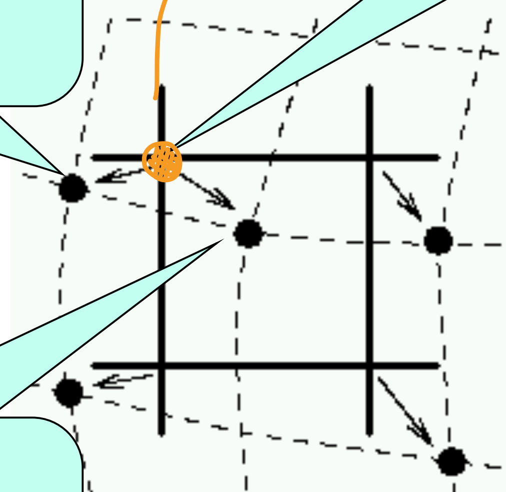
morphing 影像轉換 (ex. 從人臉 smoothly 轉成動物)

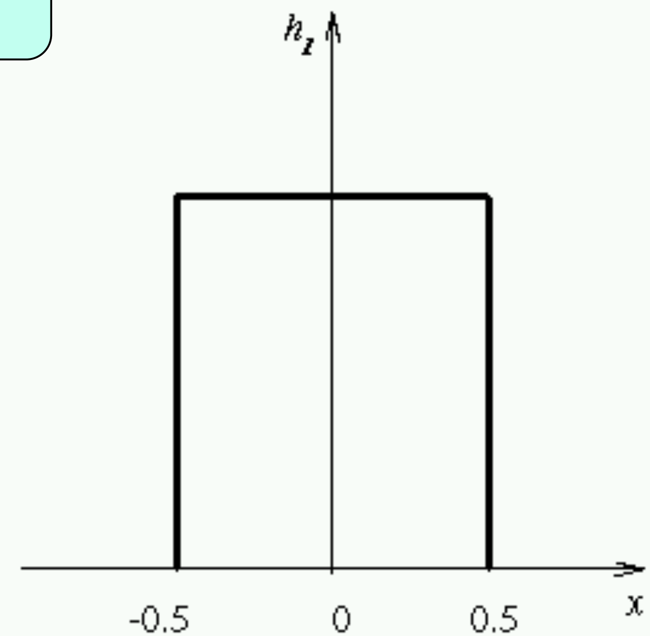
Nearest Neighbor Interpolation

old 點找最近的
new 點移動

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

$f(x_1,y_1)$



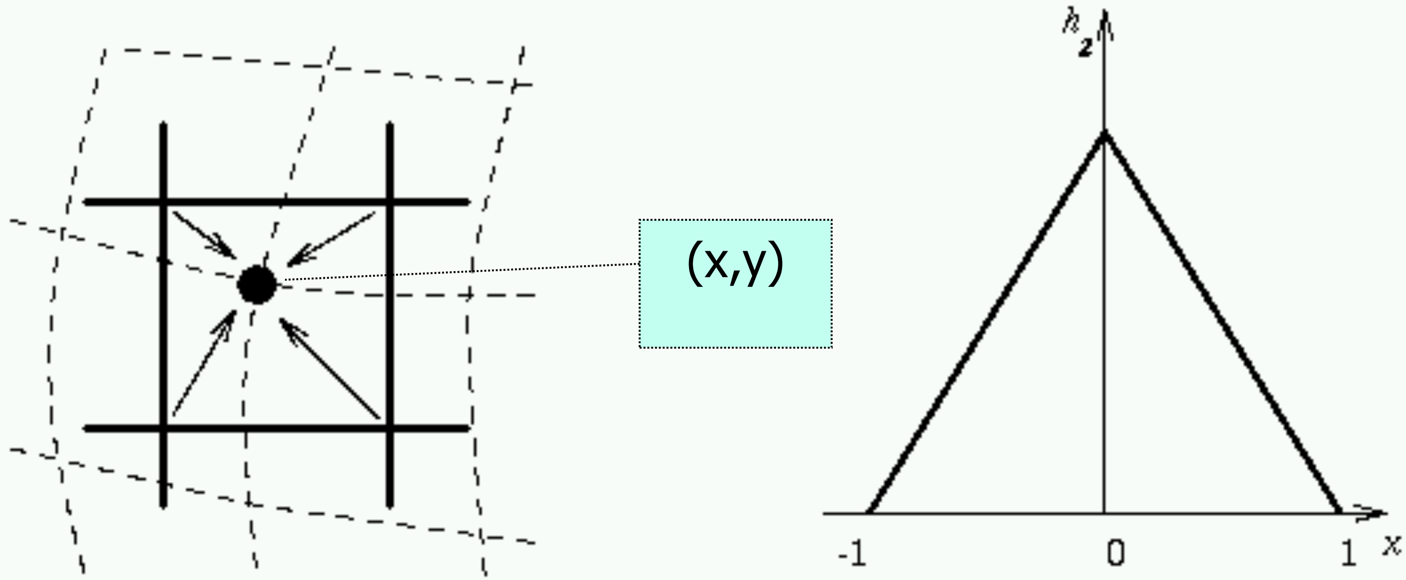


$f_1(x_3,y_3) =$
$f(\mathrm{round}(x_3), \mathrm{round}(y_3))$
$=f(x_1,y_1)$

實線 old，虛線 new

Bilinear Interpolation

距離愈近，weight就愈高



$$f_2(x,y)$$

$$= (1-a)\bullet(1-b)\bullet f(l,k) \quad \bullet\, a(1-b)\bullet f(l+1,k)$$

$$+(1-a)\bullet b\bullet f(l,k+1)+a \quad \bullet\, b\bullet f(l+1,k+1)$$

$$l = floor(x), k = floor(y), a = x-l, b = y-k.$$

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

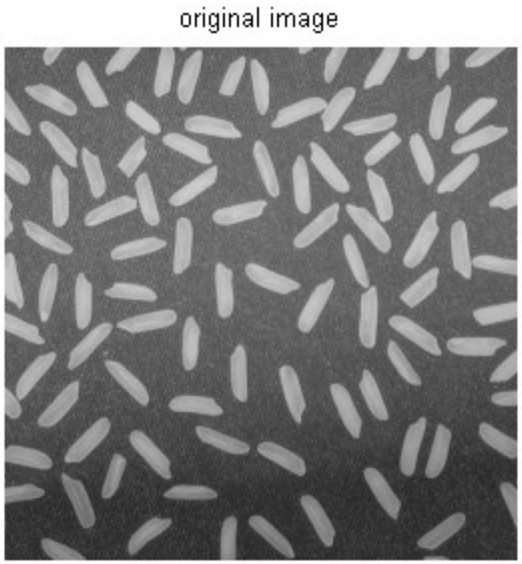
Comparison

Nearest neighbor interpolation | Bilinear interpolation | Bicubic interpolation

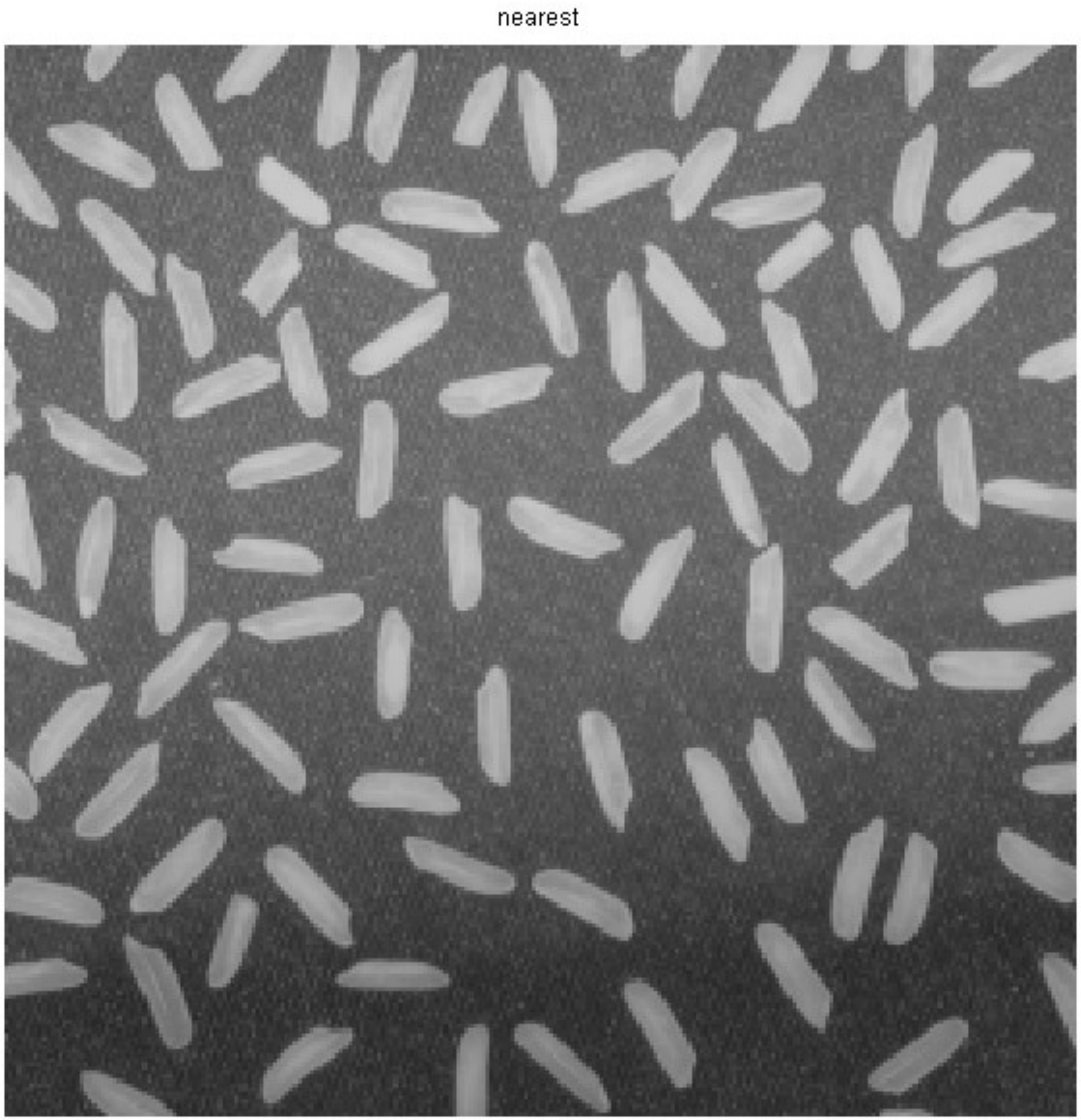
a b c

FIGURE 2.27 (a) Image reduced to 72 dpi and zoomed back to its original 930 dpi using nearest neighbor interpolation. This figure is the same as Fig. 2.23(d). (b) Image reduced to 72 dpi and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation.

Examples: Interpolation

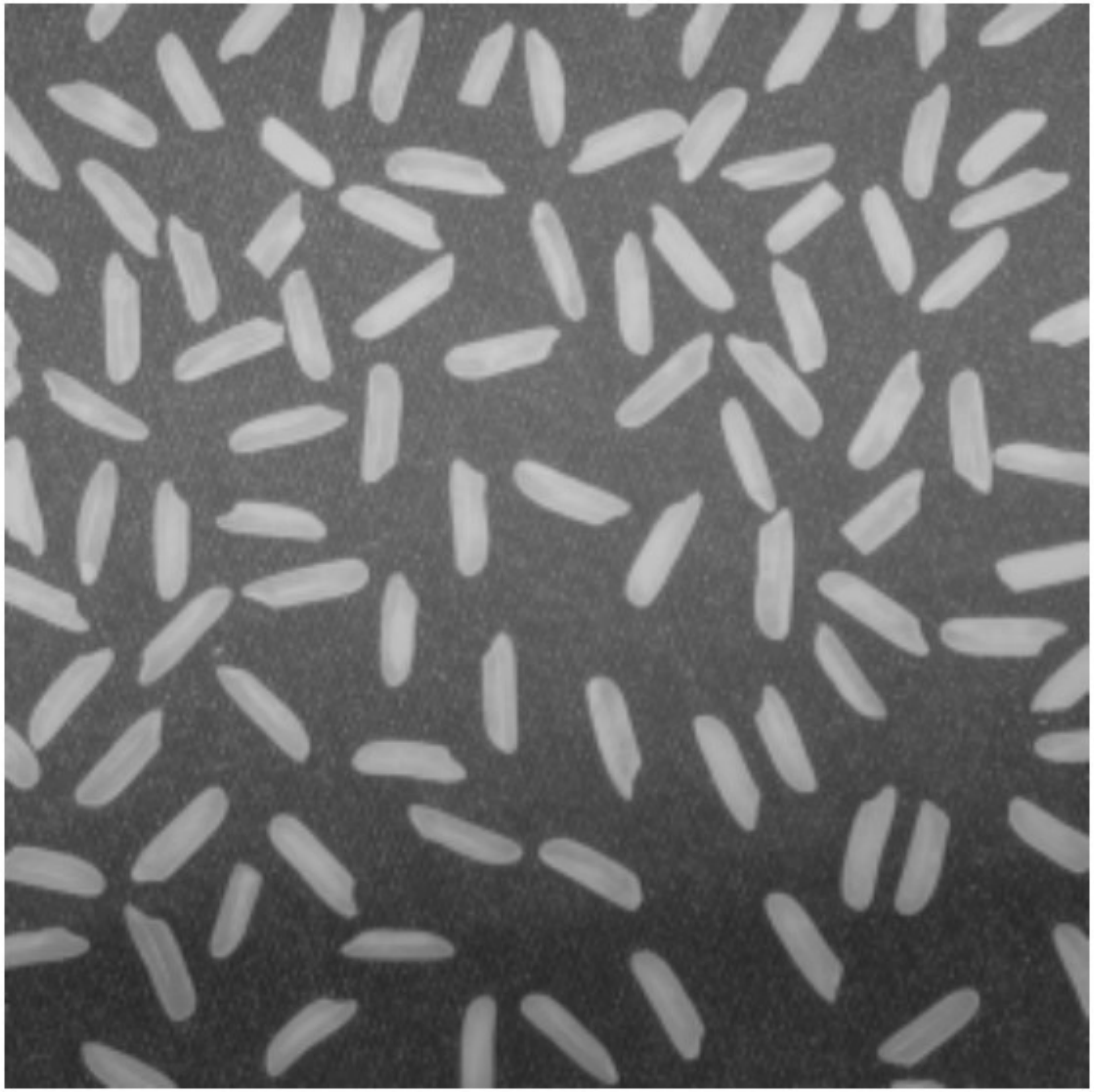


Examples: Interpolation



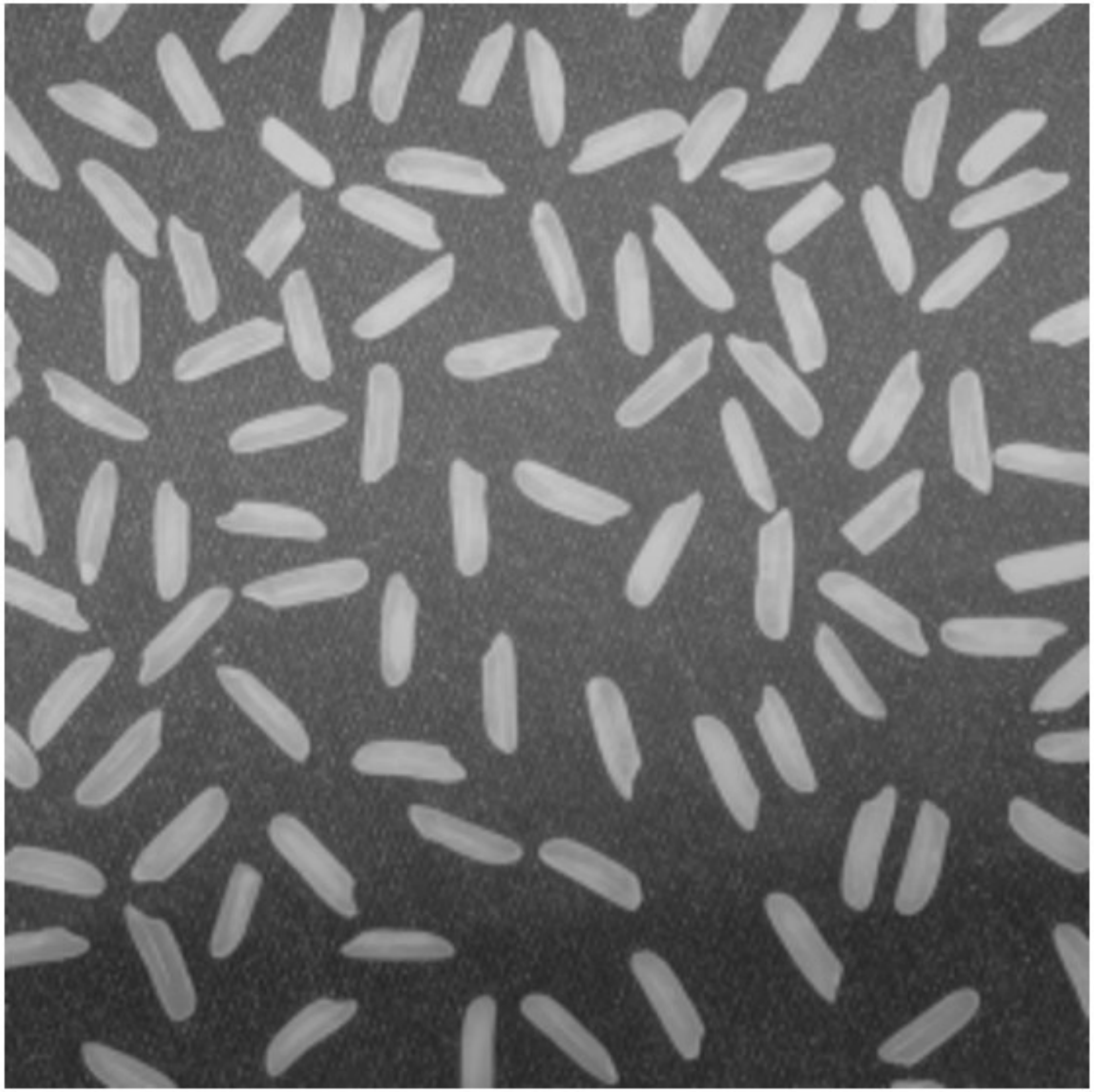
Examples: Interpolation

bilinear



Examples: Interpolation

bicubic



Examples: Interpolation

Original Image

Examples: Interpolation

Nearest Neighbor Interpolation

Examples: Interpolation

Bilinear Interpolation

Examples: Interpolation

Bicubic Interpolation

Basic Relationships Between Pixels

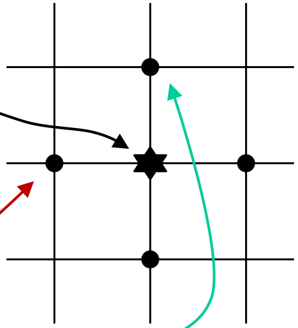
- Neighborhood
- Adjacency
- Connectivity
- Paths
- Regions and boundaries

Outline

- Elements of visual perception
- Light and the electromagnetic spectrum
- Image sensing and acquisition
- Image sampling and quantization
- Basic relationships between pixels
- Basic mathematical tools for DIP

Neighborhood

- **Neighbors** of a pixel p at coordinates (x,y)



- **4-neighbors** of p, denoted by $N_4(p)$:
  (x-1, y), (x+1, y), (x,y-1), and (x, y+1).
- **4 diagonal neighbors** of p, denoted by $N_D(p)$:
  (x-1, y-1), (x+1, y+1), (x+1,y-1), and (x-1, y+1).
- **8 neighbors** of p, denoted $N_8(p)$
  $N_8(p) = N_4(p) \cup N_D(p)$

Adjacency

Let V be the set of intensity values

- **4-adjacency**: Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.
- **8-adjacency**: Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.
- **m-adjacency (mixed adjacency):** Two pixels p and q with values from V are m-adjacent if
  - q is in the set $N_4(p)$, or
  - q is in the set $N_D(p)$ and the set $N4(p) \cap N4(q)$ has no pixels whose values are from V.

(To resolve the ambiguity that may result from using the 8-adjacency)

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

8-adjacent (ambiguous) m-adjacent

Path

- A (digital) path (or curve) from pixel p with coordinates (x_0, y_0) to pixel q with coordinates (x_n, y_n) is a sequence of distinct pixels with coordinates

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

  where (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$.

- *n:* the *length* of the path.

- If $(x_0, y_0) = (x_n, y_n)$, the path is ***closed*** path.

- We can define 4-, 8-, and m-paths based on the type of adjacency used.

Examples: Adjacency and Path

V = {1, 2}

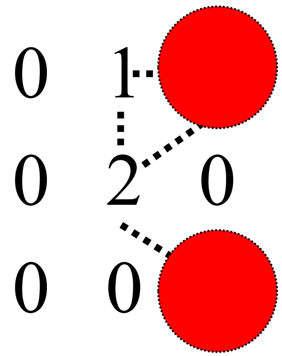
$$
\begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}
\qquad\qquad
\begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}
\qquad\qquad
\begin{matrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}
$$

8-adjacent **m-adjacent**

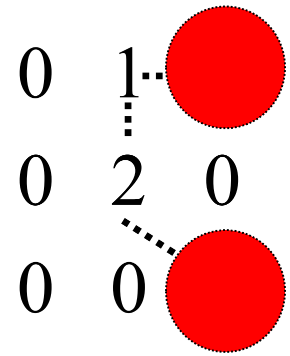
Examples: Adjacency and Path

V = {1, 2}

$$
\begin{array}{ccc}
0_{1,1} & 1_{1,2} & 1_{1,3} \\
0_{2,1} & 2_{2,2} & 0_{2,3} \\
0_{3,1} & 0_{3,2} & 1_{3,3}
\end{array}
$$



8-adjacent



m-adjacent

The 8-path from (1,3) to (3,3):
(i) (1,3), (1,2), (2,2), (3,3)
(ii) (1,3), (2,2), (3,3)

Ambiguous!

The m-path from (1,3) to (3,3):
(1,3), (1,2), (2,2), (3,3)

Connectivity

- **Connected in S**

  Let S represent a subset of pixels in an image. Two pixels p with coordinates (x_0, y_0) and q with coordinates (x_n, y_n) are said to be **connected in S** if there exists a path

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

  between these two pixels, where

$$\forall i, 0 \le i \le n, (x_i, y_i) \in S$$

Connectivity and Region

Let S represent a subset of pixels in an image.

For every pixel *p* in S, the set of pixels in S that are connected to *p* is called a ***connected component*** of S.

If S has only one connected component, then S is called ***Connected Set*.**

We call R a **region** of the image if R is a connected set

Two regions, R_i and R_j are said to be ***adjacent*** if their union forms a connected set. Regions that are not to be adjacent are said to be ***disjoint.***

Boundary, Foreground, and Background

- **Boundary (or border)**
  - The set of pixels in R that are adjacent to pixels in the complement of R. (That is, the set of pixels in R that have at least one background neighbor.)
  - If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns of the image.
- **Foreground and background**
  - An image contains K disjoint regions, R_k, $k = 1, 2, \ldots, K$. Let R_u denote the union of all the K regions and $(R_u)^c$ denote its complement.

  All the points in R_u is called **foreground**;

  All the points in $(R_u)^c$ is called **background**.

Question 1

- In the following image, are the two regions of 1s adjacent if 8-adjacency is used? Are they adjacent if 4-adjacency is used?

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

Question 2

- **In the following arrangement of pixels, the circled point is part of the boundary of the 1-valued pixels if 8-adjacency is used, true or false?**

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

Question 3

- In the following arrangement of pixels, the circled point is part of the boundary of the 1-valued pixels if 4-adjacency is used, true or false?

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

Distance Measures

- Given pixels *p, q* and *z* with coordinates (x, y), (s, t), (u, v) respectively, the distance function D has following properties:

  - $D(p, q) \geq 0$ [$D(p, q) = 0$, iff $p = q$]
  - $D(p, q) = D(q, p)$
  - $D(p, z) \leq D(p, q) + D(q, z)$

Distance Measures

Three common distance measures

- Euclidean Distance

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

- City Block Distance

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

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

- Chess Board Distance

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

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

Question 4

- In the following arrangement of pixels, what's the value of 1) chessboard distance, 2) city-block, and 3) m-path distances between the two circled points?

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

Question 5

- **In the following arrangement of pixels, what's the value of the length of the m-path between the circled two points?**

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

Introduction to Mathematical Operations in DIP

- **Array vs. Matrix Operation**

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

$$A \odot B \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

Array product

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

Matrix product

Introduction to Mathematical Operations in DIP

- **Linear vs. Nonlinear Operation**

$$H\left[f(x,y)\right]=g(x,y)$$

$$H\left[a_i f_i(x,y)+a_j f_j(x,y)\right]$$
$$=H\left[a_i f_i(x,y)\right]+H\left[a_j f_j(x,y)\right]$$
$$=a_i H\left[f_i(x,y)\right]+a_j H\left[f_j(x,y)\right]$$
$$=a_i g_i(x,y)+a_j g_j(x,y)$$

Additivity

Homogeneity

H is said to be a **linear operator;**

H is said to be a **nonlinear operator** if it does not meet the above condition.

Arithmetic Operations

- Arithmetic operations between images are array operations. The four arithmetic operations are denoted as

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

All elementwise operations

Example: Addition of Noisy Images for Noise Reduction

f(x,y): Noiseless image

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

g(x,y): Corrupted image

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

相加可以 reduce noice! → stablelization 也是同理

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

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

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

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

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

$$= f(x,y)$$

$$\sigma^2_{\bar{g}(x,y)}$$

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

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

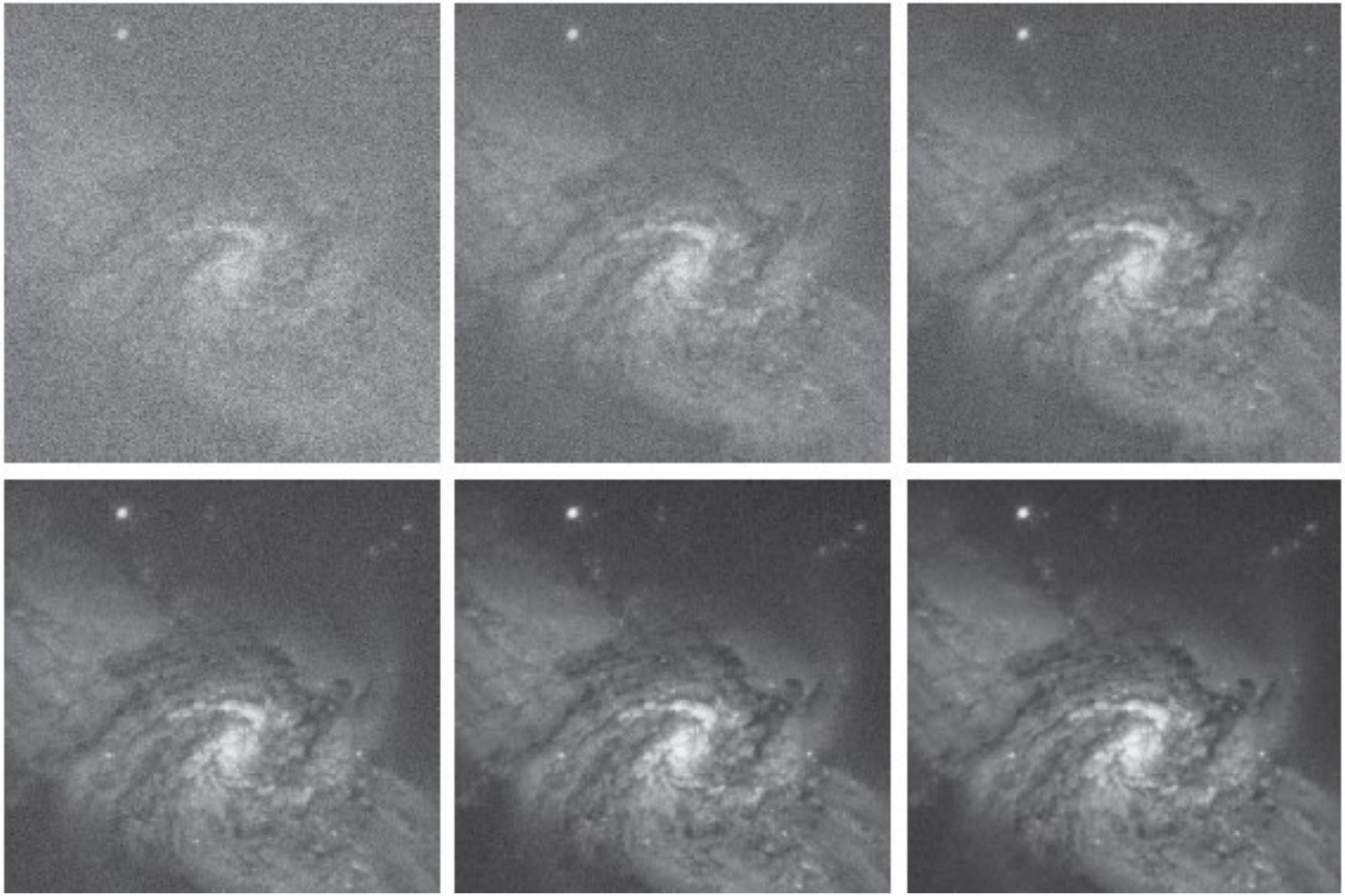
$$= \frac{1}{K}\sigma^2_{n(x,y)}$$

$K\uparrow$ ==> pixel variability $\downarrow$

Example: Addition of Noisy Images for Noise Reduction

- Imaging under low light levels frequently causes sensor noise to render single images virtually useless for analysis.

- In astronomical observations, similar sensors are used for noise reduction by observing the same scene over a long period of time. Image averaging is then applied to reduce the noise.



a b c
d e f

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.)

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 $g(x,y)$ 加减可以突出detail

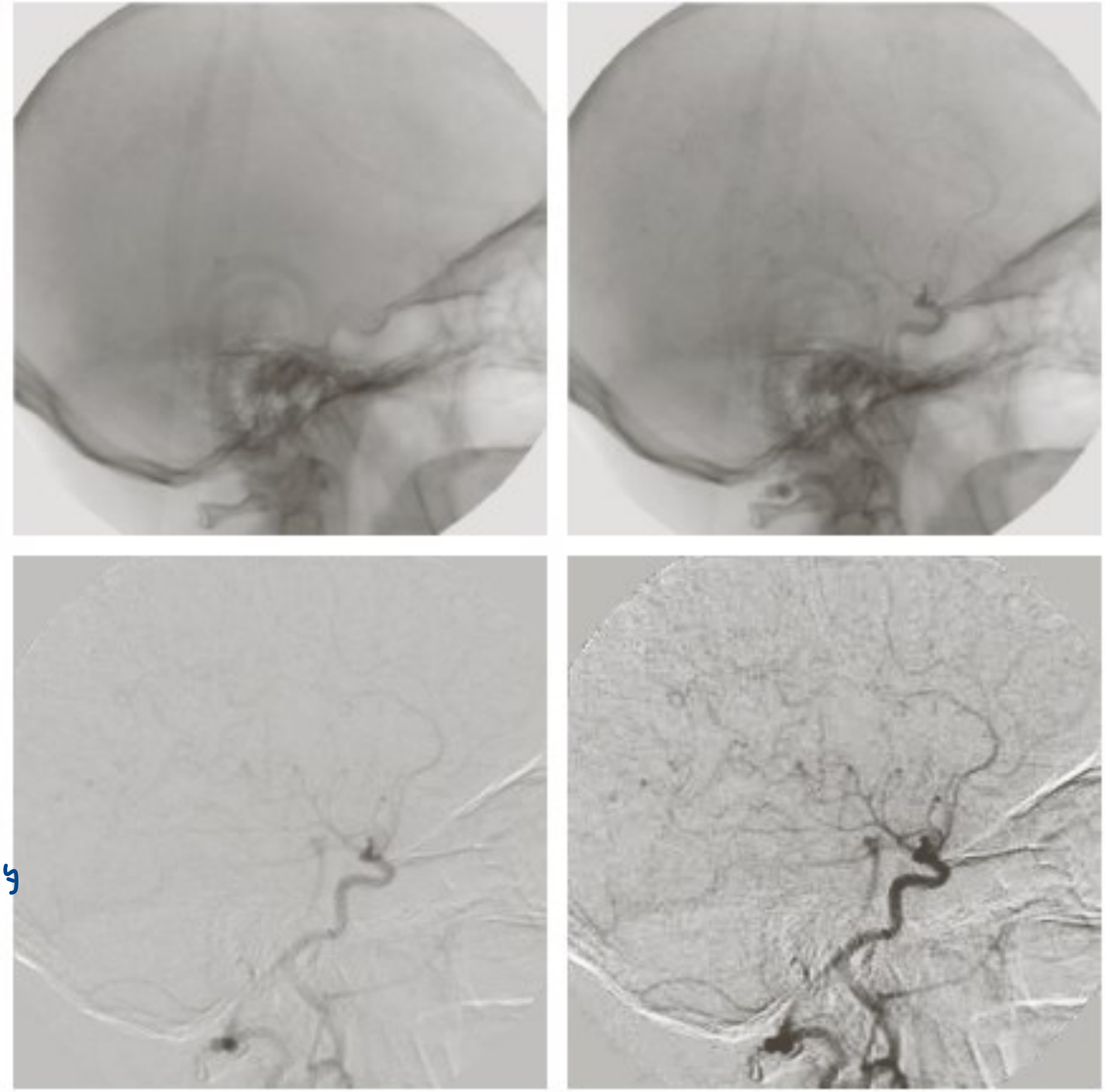
$$g(x,y) = f(x,y) - h(x,y)$$

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

血管造影法

a b
c d

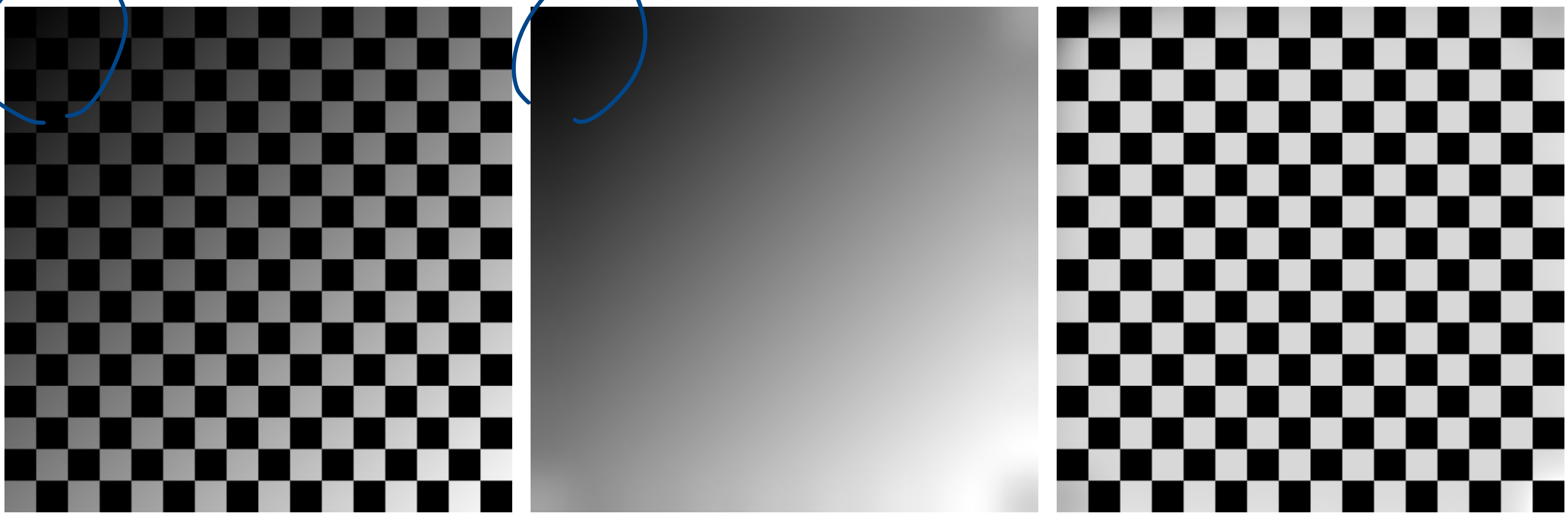
FIGURE 2.32 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of the Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)



ex. 注射顯影劑
後減掉未注射的

Example of Image Multiplication: Shading Correction

相除後相消的感覺

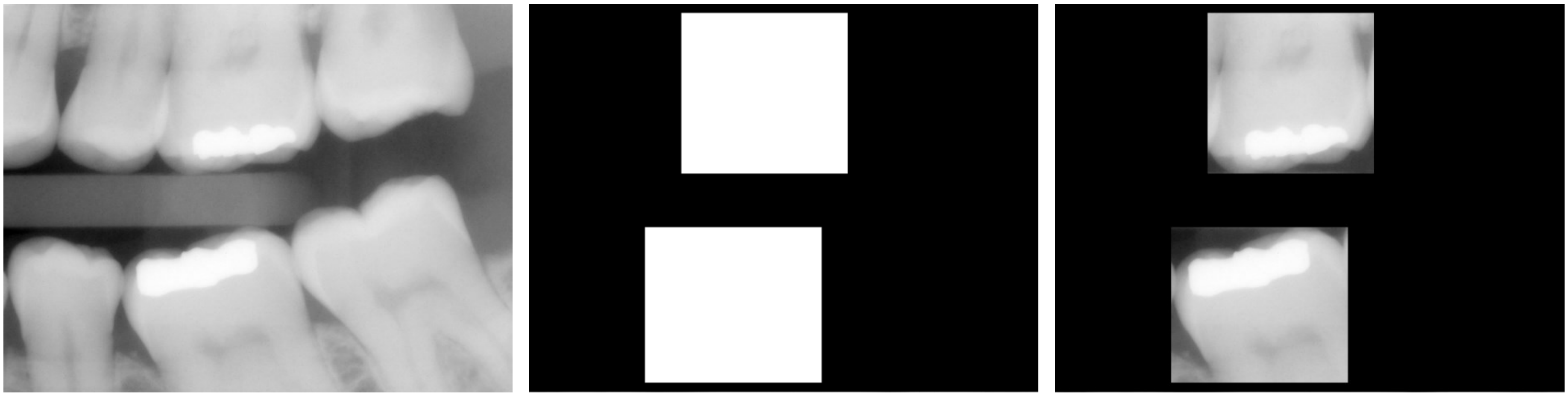


a b c

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.)

Example of Image Multiplication: Masking

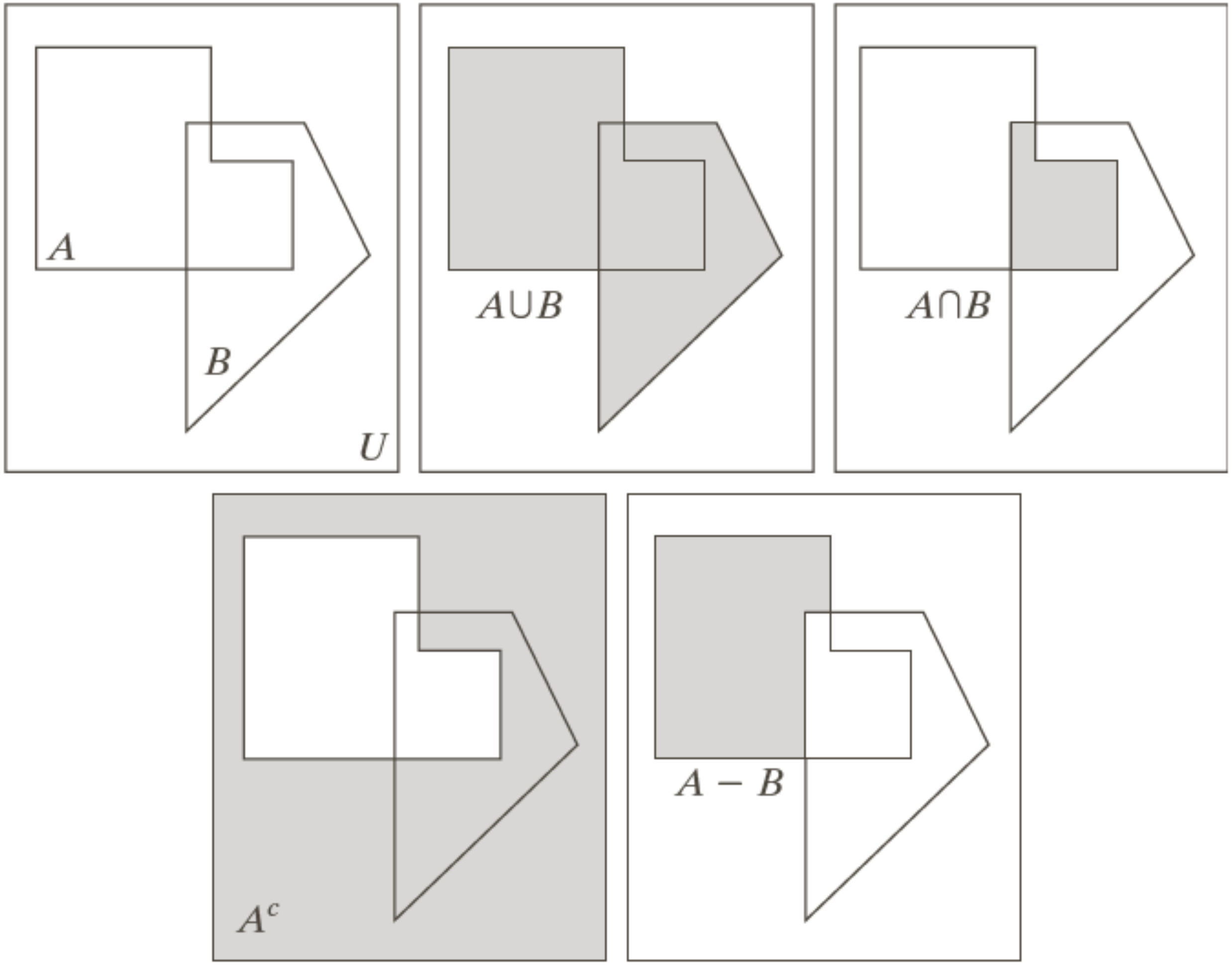
相乘



a b c

FIGURE 2.34 (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).

Set and Logical Operations



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.

Set and Logical Operations

- Let A be the elements of a gray-scale image

  The elements of A are triplets of the form (x, y, z), where x and y are spatial coordinates and z denotes the intensity at the point (x, y).

$$A = \{(x, y, z) \mid z = f(x, y)\}$$

- The complement of A is denoted by A^c

$$A^c = \{(x, y, K - z) \mid (x, y, z) \in A\}$$

$K = 2^k - 1; k$ is the number of intensity bits used to represent z

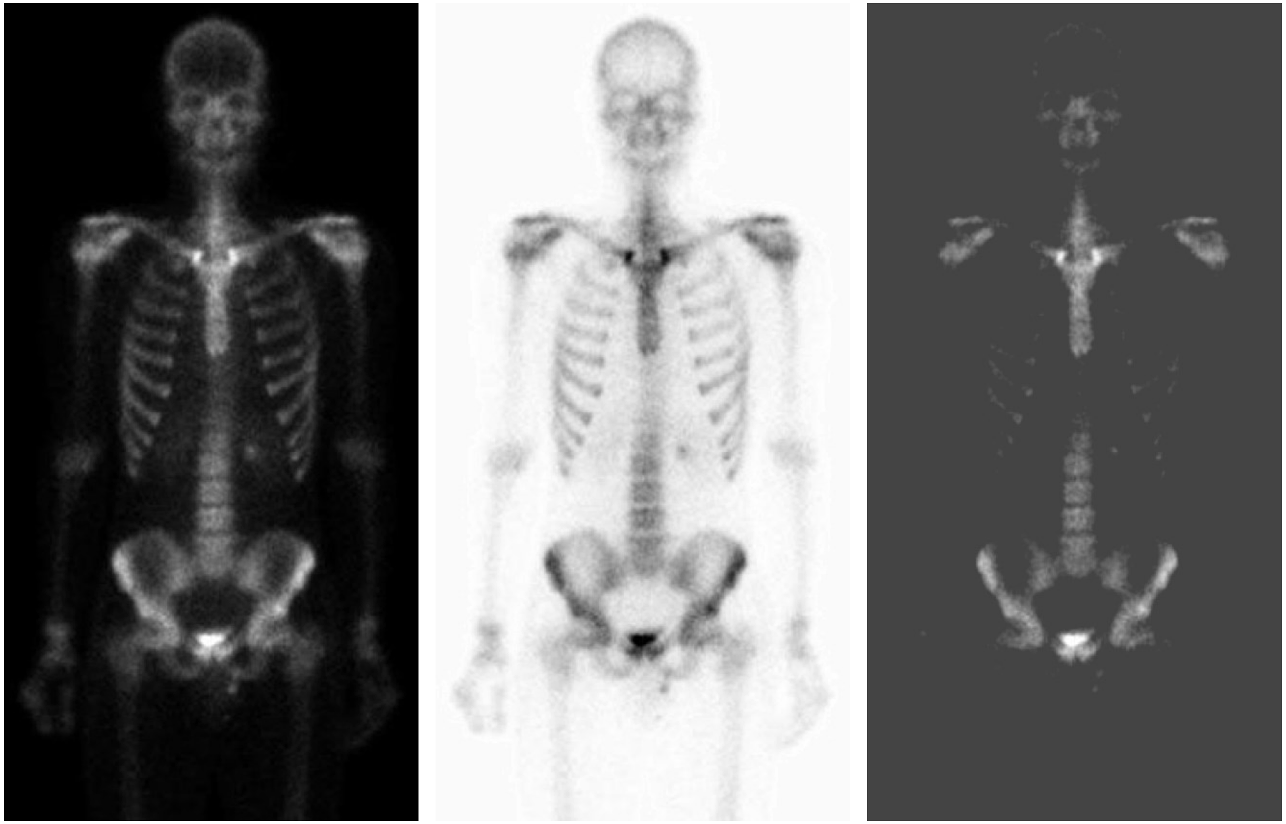
I_{max}

Set and Logical Operations

- The union of two gray-scale images (sets) A and B is defined as the set

$$A \cup \qquad \mathrm{x}(a, b) \mid a \in A, b \in B\}$$

Set and Logical Operations



a b c

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

(c) All values exceeding $3\bar{z}$ appears as values from A and all other pixels have value $3\bar{z}$, which is a mid-gray value.

Set and Logical Operations

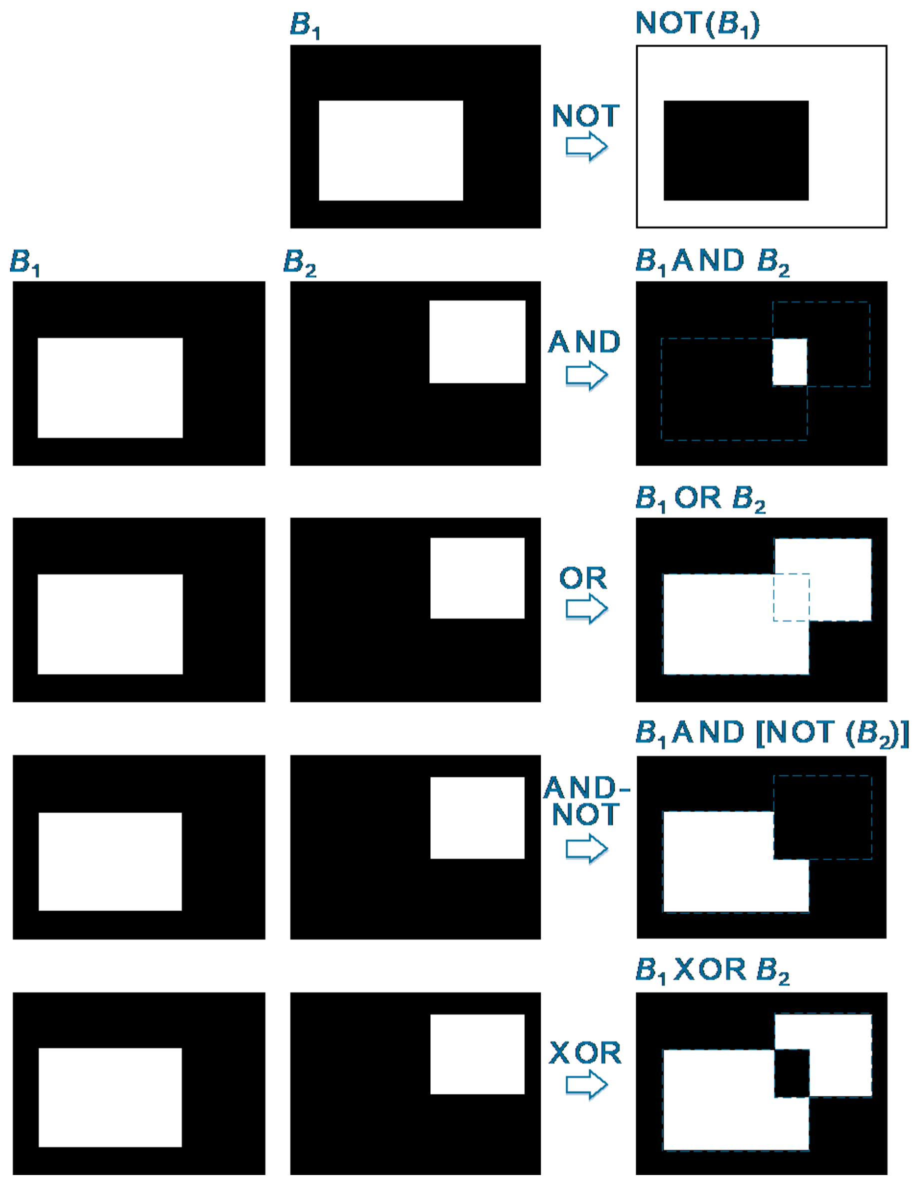


FIGURE 2.37 Illustration of logical operations involving foreground (white) pixels. Black represents binary 0's and white binary 1's. The dashed lines are shown for reference only. They are not part of the result.

Spatial Operations

- Single-pixel operations

Alter the values of an image's pixels based on the intensity.

$$s = T(z)$$

e.g.,

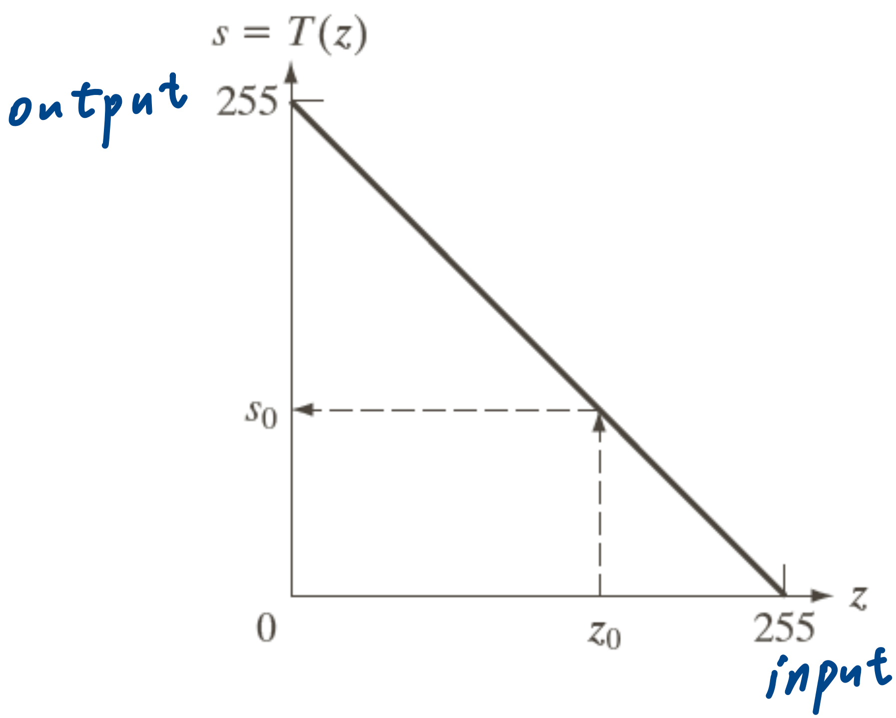
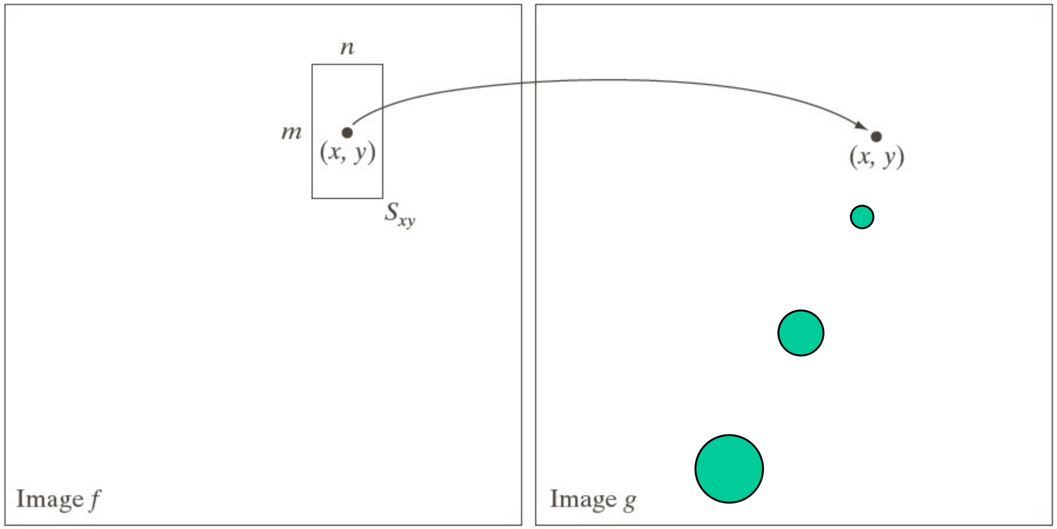


FIGURE 2.34 Intensity transformation function used to obtain the negative of an 8-bit image. The dashed arrows show transformation of an arbitrary input intensity value z_0 into its corresponding output value s_0.

Spatial Operations

- Neighborhood operations



The value of this pixel is determined by a specified operation involving the pixels in the input image with coordinates in S_{xy}

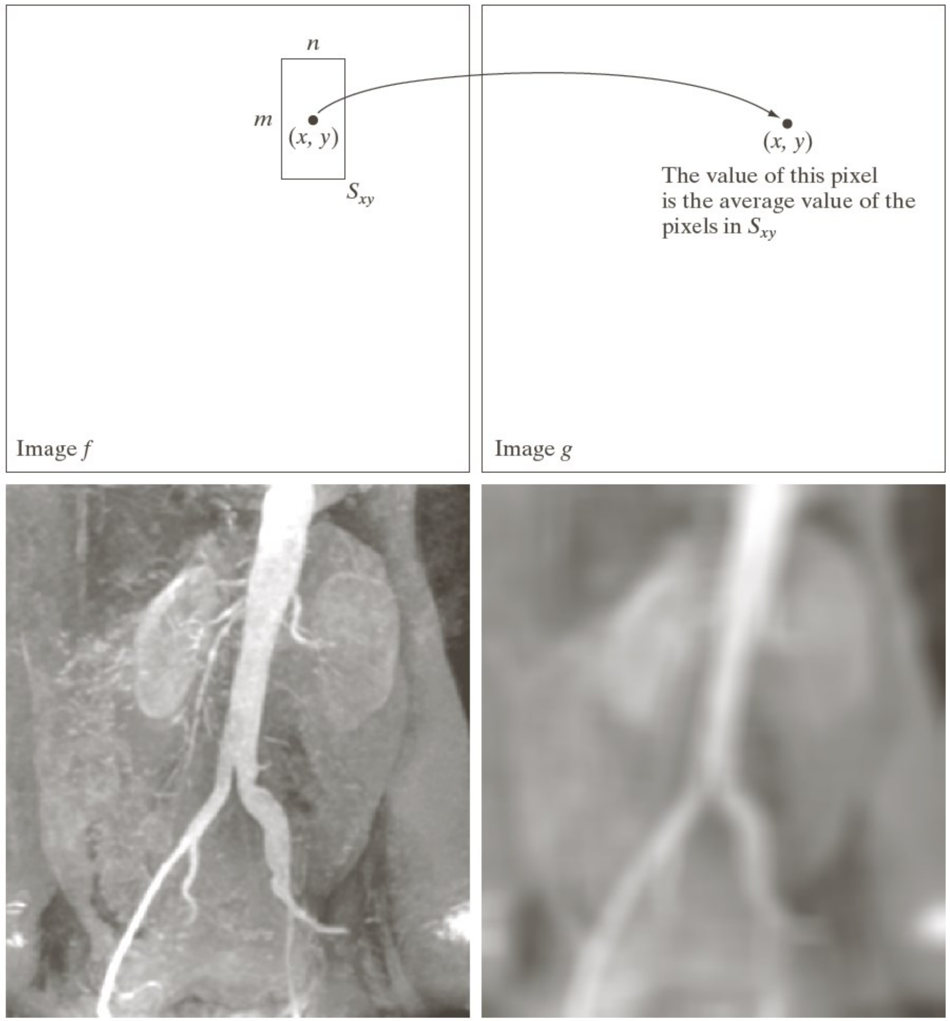
Spatial Operations

- Neighborhood operations

a b
c d

FIGURE 2.39
Local averaging using neighborhood processing. The procedure is illustrated in (a) and (b) for a rectangular neighborhood. (c) An aortic angiogram (see Section 1.3). (d) The result of using Eq. (2-43) with $m = n = 41$. The images are of size 790×686 pixels. (Original image courtesy of Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School.)

大主動脈的血管造影



Geometric Spatial Transformations

- Geometric transformation (rubber-sheet transformation)
  — A spatial transformation of coordinates

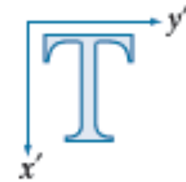
$$(x, y) = T\{(v, w)\}$$

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

- Affine transform (preserving parallel relationships)

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

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

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

Intensity Assignment

- Forward Mapping

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

It's possible that two or more pixels can be transformed to the same location in the output image.

- Inverse Mapping

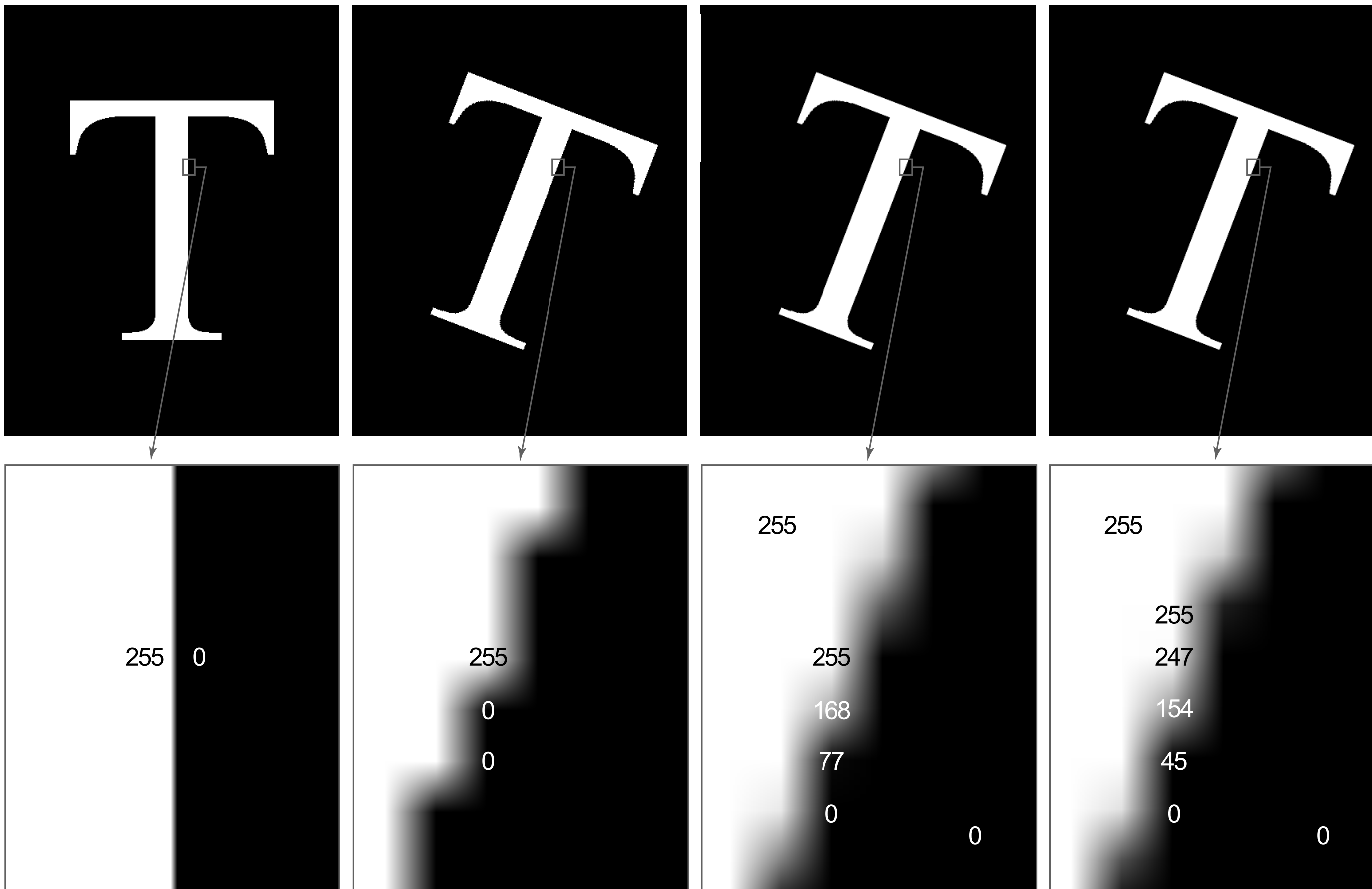
$$(v, w) = T^{-1}\{(x, y)\}$$

The nearest input pixels to determine the intensity of the output pixel value.

Inverse mappings are more efficient to implement than forward mappings.

Example: Image Rotation and Intensity Interpolation

要会区分



a b c d
e f g h

FIGURE 2.40 (a) A 541 × 421 image of the letter T. (b) Image rotated −21° using nearest-neighbor interpolation for intensity assignments. (c) Image rotated −21° using bilinear interpolation. (d) Image rotated −21° using bicubic interpolation. (e)-(h) Zoomed sections (each square is one pixel, and the numbers shown are intensity values).

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.

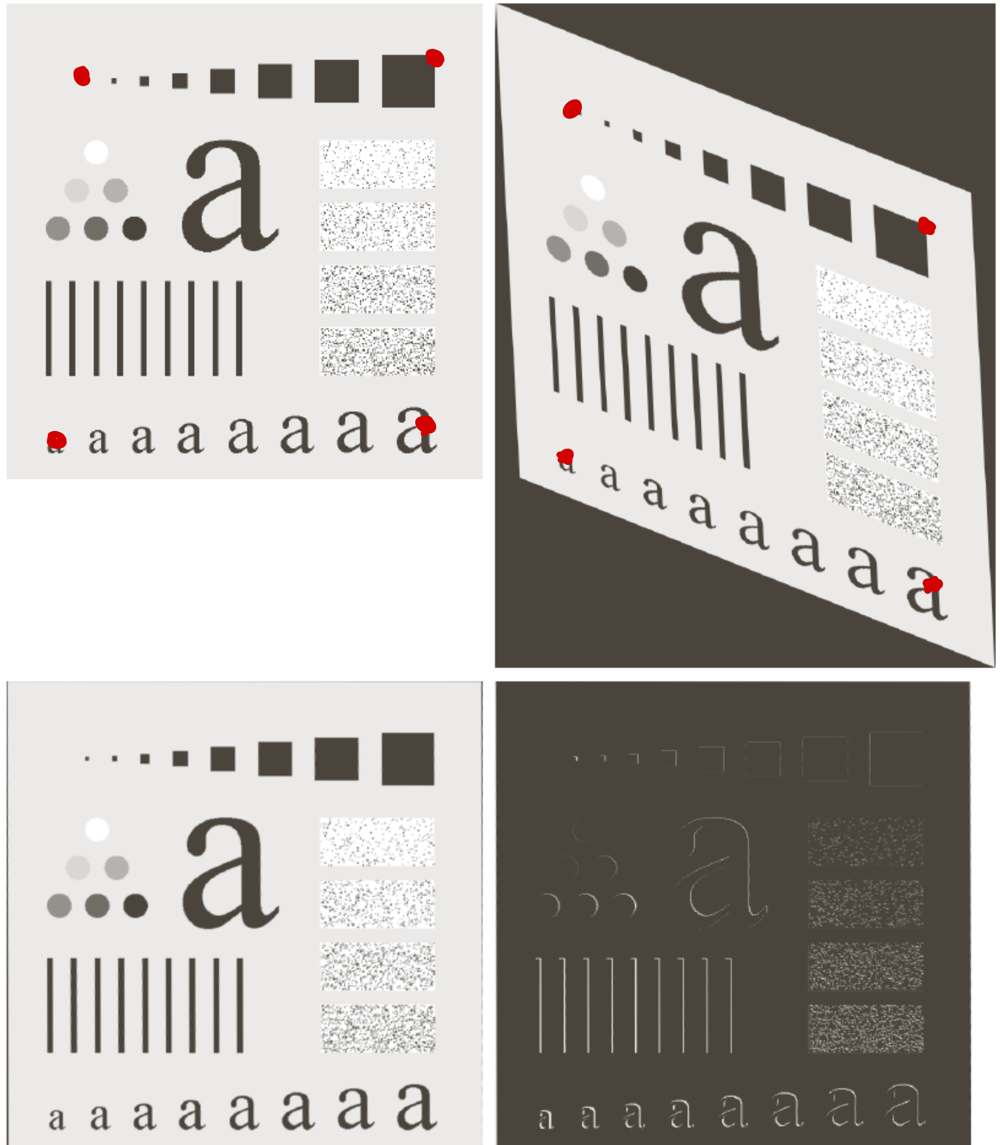
Image Registration

- 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.

Image Registration



a b
c d

FIGURE 2.42 Image registration. (a) Reference image. (b) Input (geometrically distorted image). Corresponding tie points are shown as small white squares near the corners. (c) Registered (output) image (note the errors in the border). (d) Difference between (a) and (c), showing more registration errors.

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*,

u and v are the transform variables,

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

Image Transform (cont'd)

- Given T(u, v), the original image f(x, y) can be recovered using the inverse transformation 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 kernel*,

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

Image Transform (cont'd)

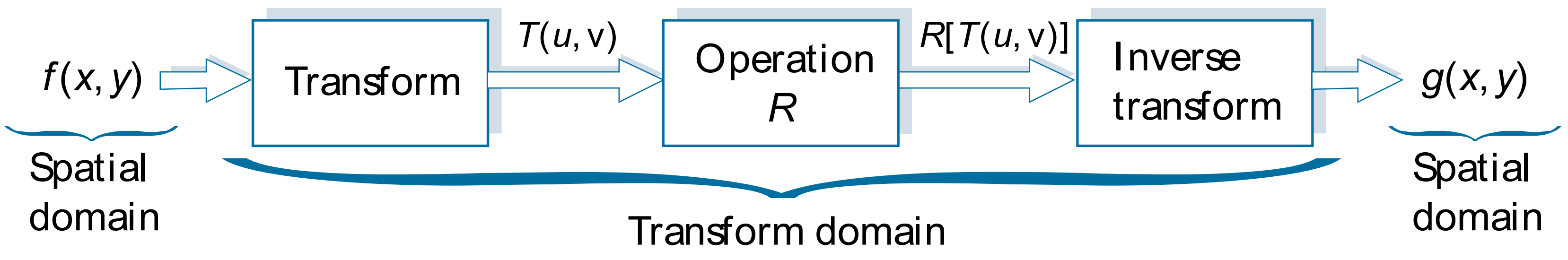
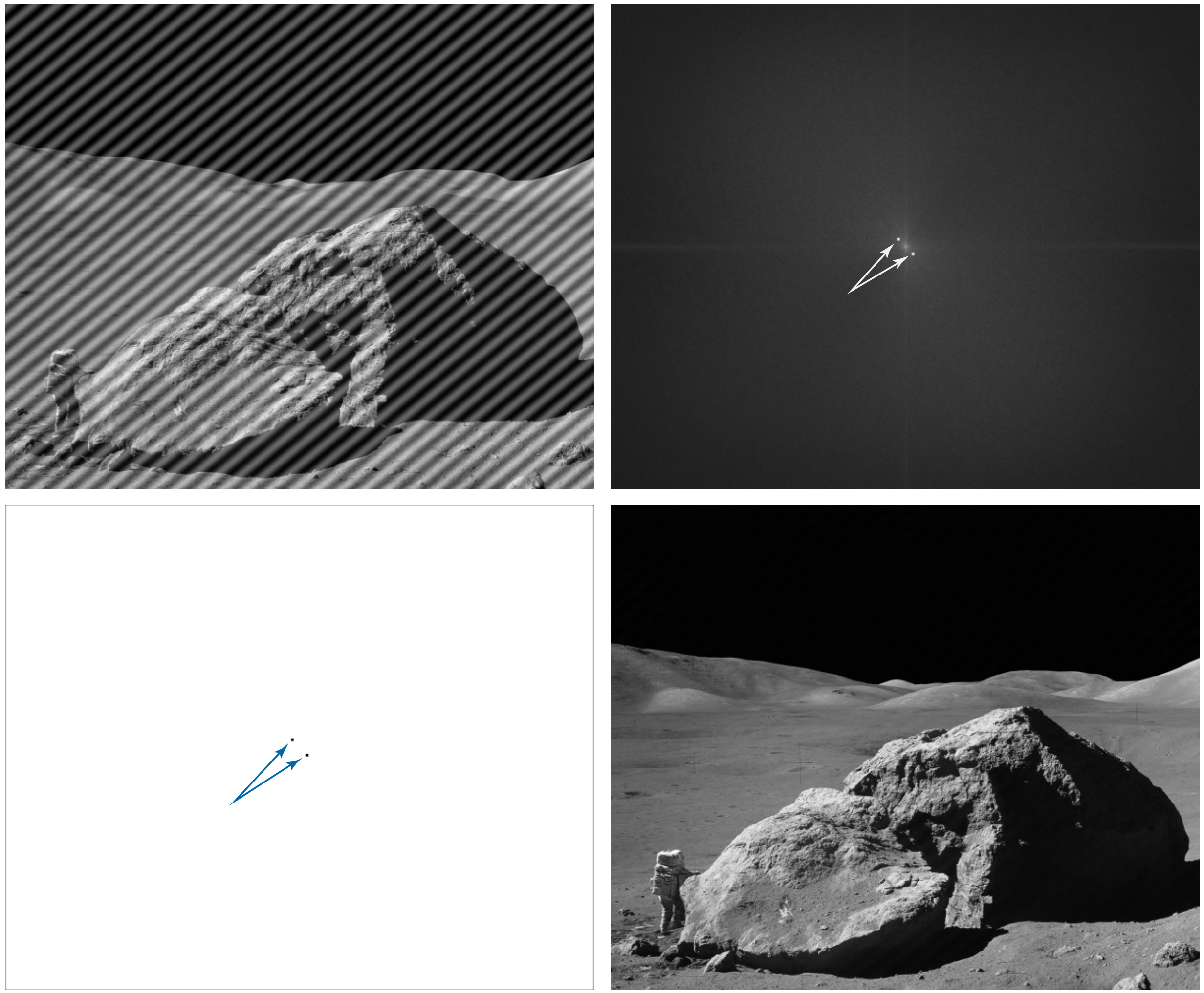


FIGURE 2.44 General approach for working in the linear transform domain.

Image Denoising in the Transform Domain



a b
c d

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

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 (cont'd)

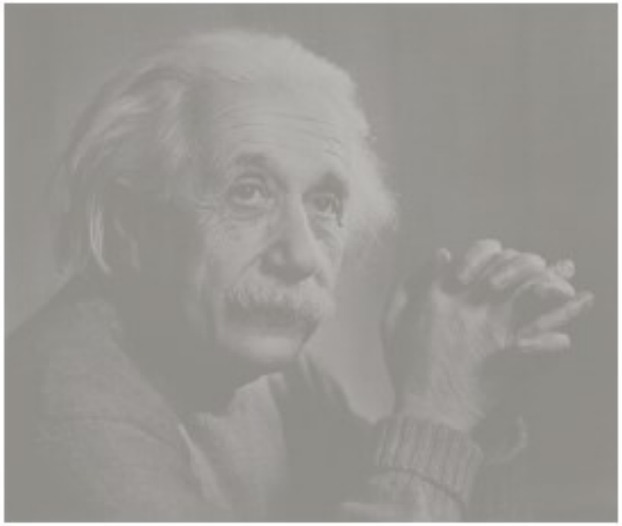
The variance of the intensities is given by

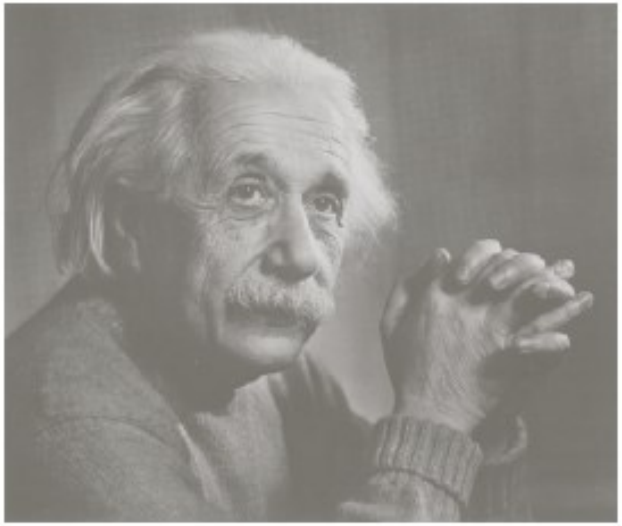
$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k)$$

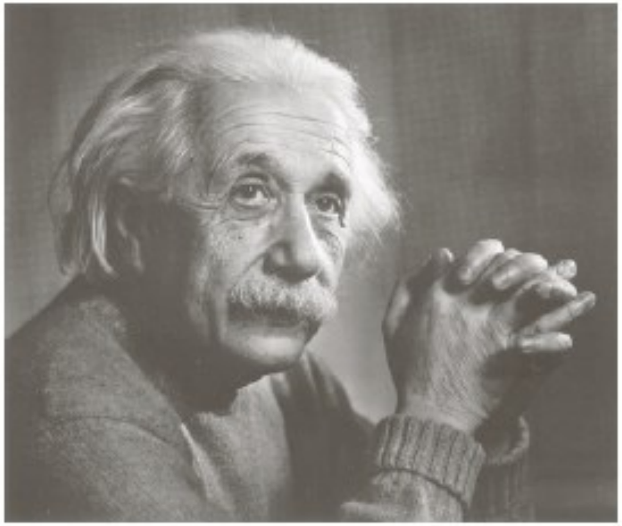
The n^{th} moment of the intensity variable z is

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

Example: Comparison of Standard Deviation Values







$\sigma = 14.3$

$\sigma = 31.6$

$\sigma = 49.2$