Digital Image Processing

Chapter 2:
Digital Image Fundamentals

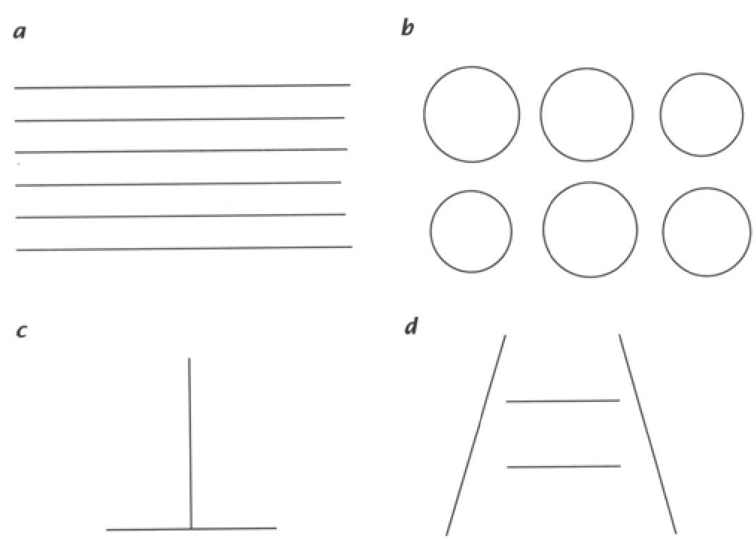
Human and Computer Vision

- We can't think of image processing without considering the human vision system.
- We observe and evaluate the images that we process with our visual system.

Simple questions

- What intensity differences can we distinguish?
- What is the spatial resolution of our eye?
- How accurately we estimate and compare distances and areas?
- How do we sense colors?
- By which features can we detect and distinguish objects?

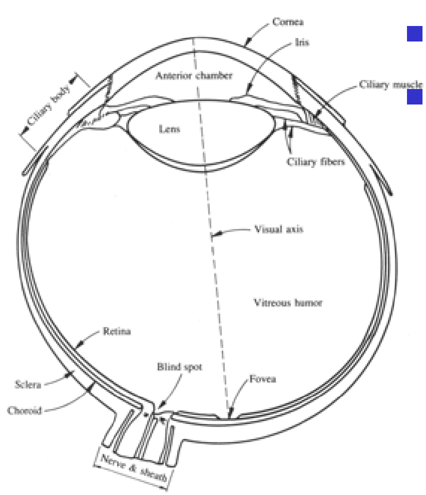
Test images



Test images

- Test images for distances and area estimation:
  a) Parallel lines with up to 5% difference in length.
  b) Circles with up to 10% difference in radius.
  c) The vertical line appears longer but actually has the same length as the horizontal line.
  d) Deception by perspective: the upper line appears longer than the lower one but actually have the same length.

Structure of the Human Eye



- Shape is nearly a sphere.
- Average diameter = 20 mm.

Lens & Retina

- Lens

  both infrared and ultraviolet light are absorbed, in excessive amounts, can cause damage to the eye.

- Retina

  Innermost membrane of the eye. When the eye is properly focused, light from an object outside the eye is imaged on the retina.

Receptors

- Receptors are divided into 2 classes:
  - Cones
  - Rods

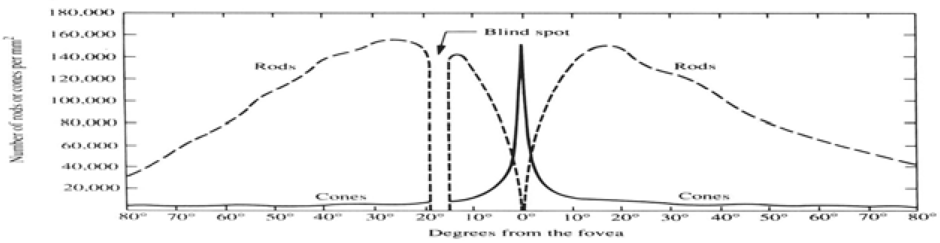
Cones

- 6-7 million, located primarily in the central portion of the retina (muscles controlling the eye rotate the eyeball until the image falls on the fovea).
- Highly sensitive to color.
- Each is connected to its own nerve end thus human can resolve fine details.
- Cone vision is called photopic or bright-light vision.

Rods

- 75-150 million, distributed over the retina surface.
- Several rods are connected to a single nerve end reduce the amount of detail discernible.
- Serve to give a general, overall picture of the field of view.
- Sensitive to low levels of illumination.
- Rod vision is called scotopic or dim-light vision.

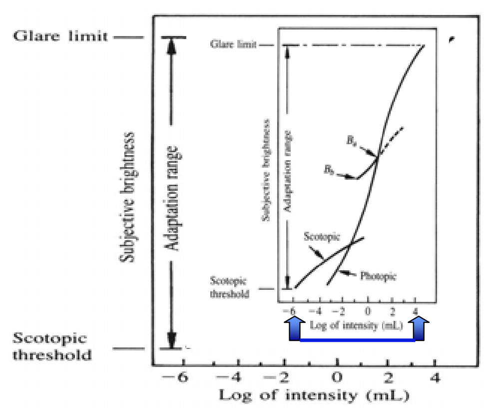
Cross section of the eye



- Blind spot ⇨ the absence of receptors area.
- Cones are most dense in the center of the retina (in the area of the fovea).

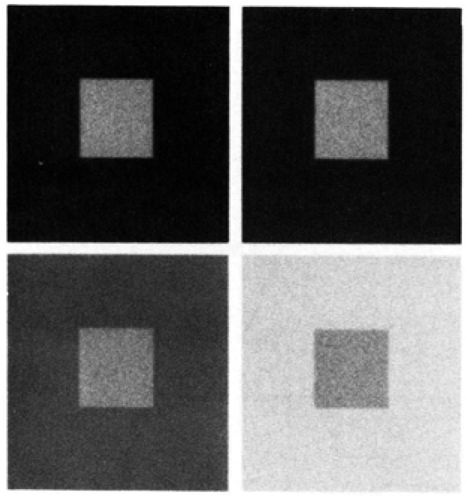
Brightness adaptation and discrimination

- The total range of intensity levels it can discriminate simultaneously is rather small compared with the total adaptation range.



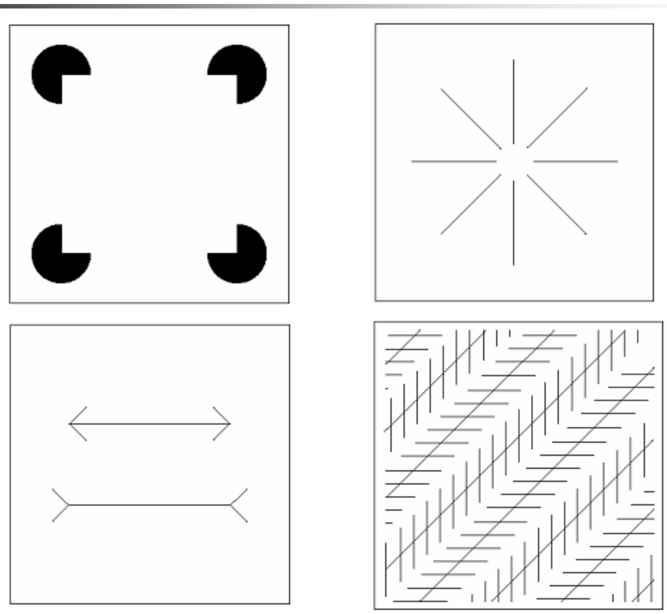
Simultaneous contrast

Which small square is the darkest one ?



- All the small squares have exactly the same intensity, but they appear to the eye progressively darker as the background becomes brighter.

Human Perception Phenomena



Signals

- a signal is a function that carries information.
- usually content of the signal changes over some set of spatiotemporal dimensions.

Vocabulary:

Spatiotemporal: existing in both space and time; having both spatial extension and temporal duration

Time-Varying Signals

- Some signals vary over time:

$$f(t)$$

for example: audio signal

may be thought at one level as a collection various tones of differing audible frequencies that vary over time.

Spatially-Varying Signals

- Signals can vary over space as well.
- An image can be thought of as being a function of 2 spatial dimensions:

$$f(x,y)$$

- for monochromatic images, the value of the function is the amount of light at that point.
- medical CAT and MRI scanners produce images that are functions of 3 spatial dimensions:

$$f(x,y,z)$$

Spatiotemporal Signals

What do you think a signal of this form is?

$$\boldsymbol{f(x,y,t)}$$

x and y are spatial dimensions;

t is time.

Perhaps, it is a video signal, animation, or other time-varying picture sequence.

Types of Signals

- most naturally-occurring signals are functions having a continuous domain.
- however, signals in a computer are discrete samples of the continuous domain.
- in other words, signals manipulated by computer have discrete domains.

Analog & Digital

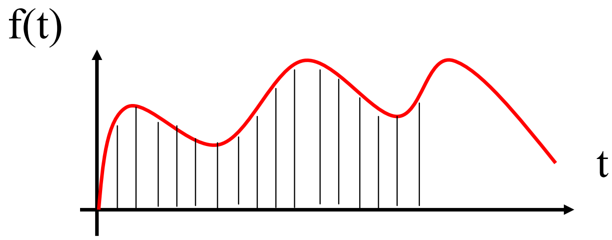
- most naturally-occurring signals also have a real-valued range in which values occur with infinite precision.
- to store and manipulate signals by computer we need to store these numbers with finite precision. thus, these signals have a discrete range.

signal has continuous domain and range = analog

signal has discrete domain and range = digital

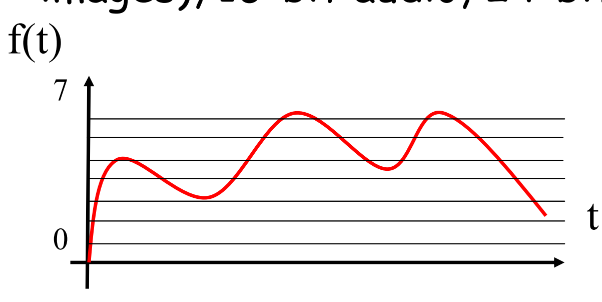
Sampling

- sampling = the spacing of discrete values in the domain of a signal.
- sampling-rate = how many samples are taken per unit of each dimension. e.g., samples per second, frames per second, etc.

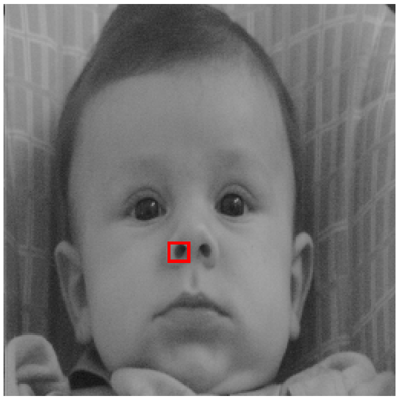


Quantization

- Quantization = spacing of discrete values in the range of a signal.
- usually thought of as the number of bits per sample of the signal. e.g., 1 bit per pixel (b/w images), 16-bit audio, 24-bit color images, etc.



Digital Image Representation



Pixel values in highlighted region

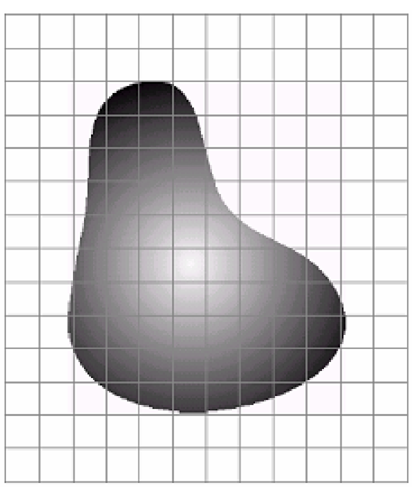
99	71	61	51	49	40	35	53	86	99
93	74	53	56	48	46	48	72	85	102
101	69	57	53	54	52	64	82	88	101
107	82	64	63	59	60	81	90	93	100
114	93	76	69	72	85	94	99	95	99
117	108	94	92	97	101	100	108	105	99
116	114	109	106	105	108	108	102	107	110
115	113	109	114	111	111	113	108	111	115
110	113	111	109	106	108	110	115	120	122
103	107	106	108	109	114	120	124	124	132

A set of number in 2D grid

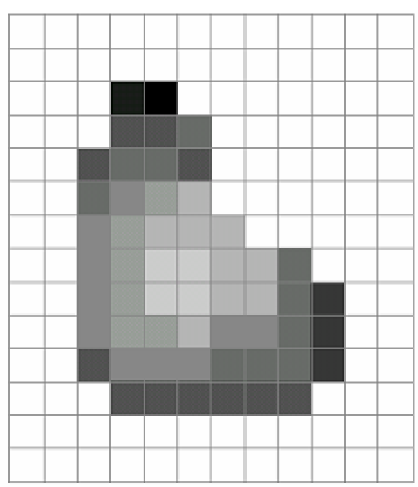
CAMERA → DIGITIZER

Samples the analog data and digitizes it.

Example of Digital Image



Continuous image projected onto a sensor array



Result of image sampling and quantization

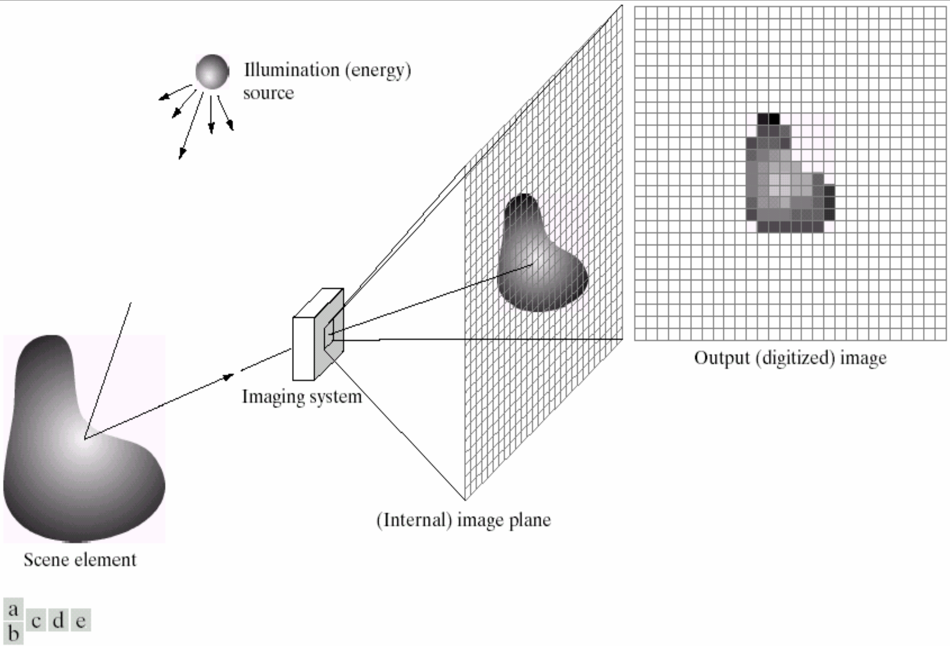


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

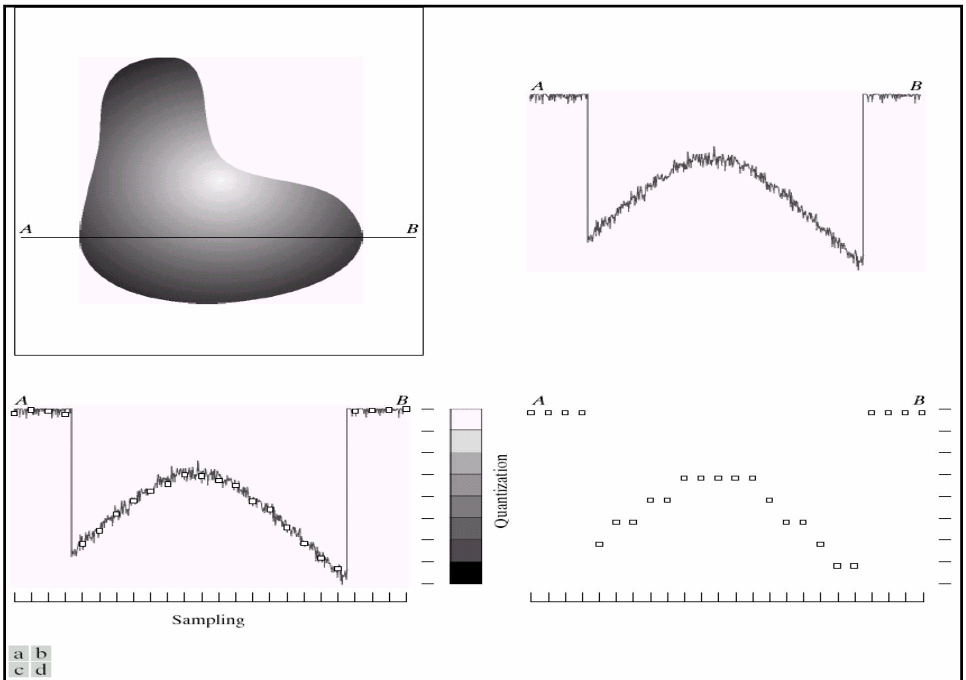
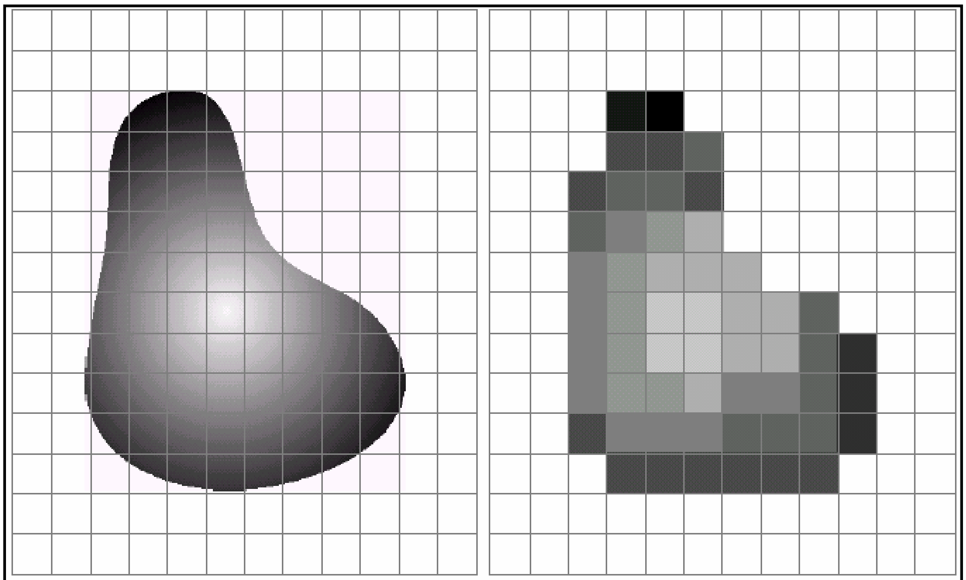


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.



a b

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

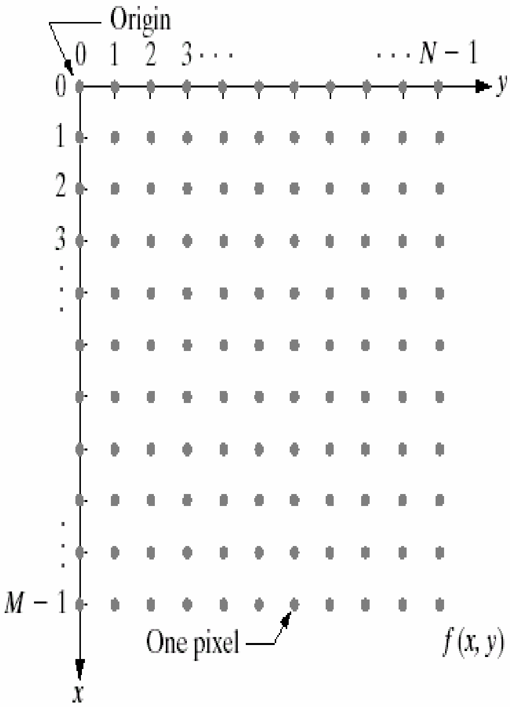
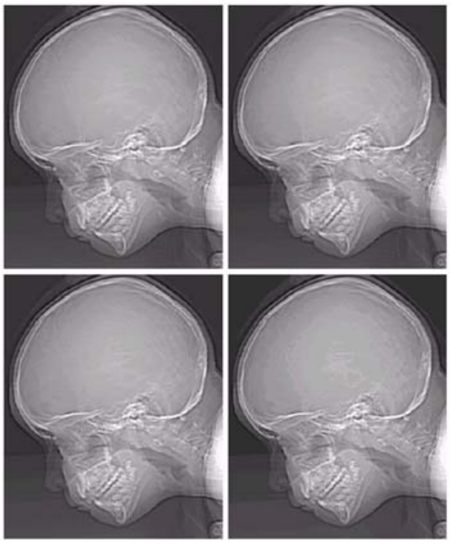


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

Chapter 2: Digital Image Fundamentals



a b
c d

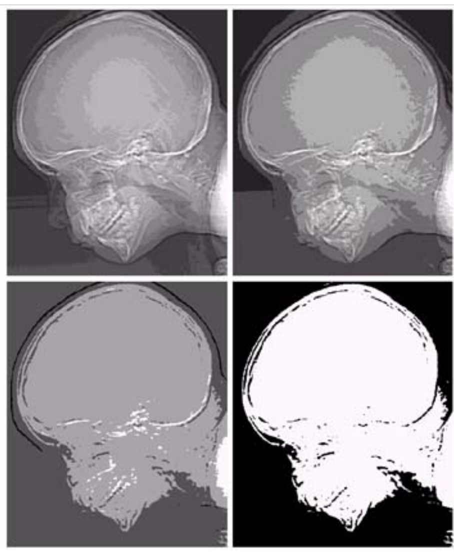
FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

Digital Image Processing, 2nd ed.
www.imageprocessingbook.com

Chapter 2: Digital Image Fundamentals

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.)



Light-intensity function

- image refers to a 2D light-intensity function, $f(x,y)$
- the amplitude of f at spatial coordinates (x,y) gives the intensity (brightness) of the image at that point.
- light is a form of energy thus $f(x,y)$ must be nonzero and finite.

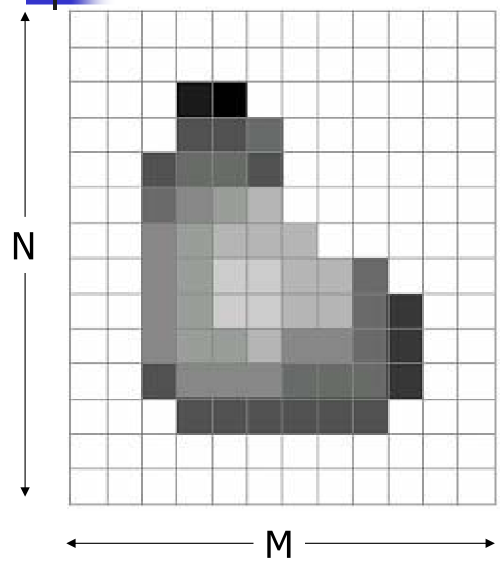
$$0 < f(x,y) < \infty$$

Gray Level

- We call the intensity of a monochrome image f at coordinate (x,y)the gray level (l)of the image at that point.

  - l lies in the range Lmin <= l <= Lmax
  - Lmin is positive and Lmax is finite
  - gray scale =[Lmin,Lmax]
  - 0 = Black, L = White

Number of bits



- The number of gray levels typically is an integer power of 2

$$L = 2^k$$

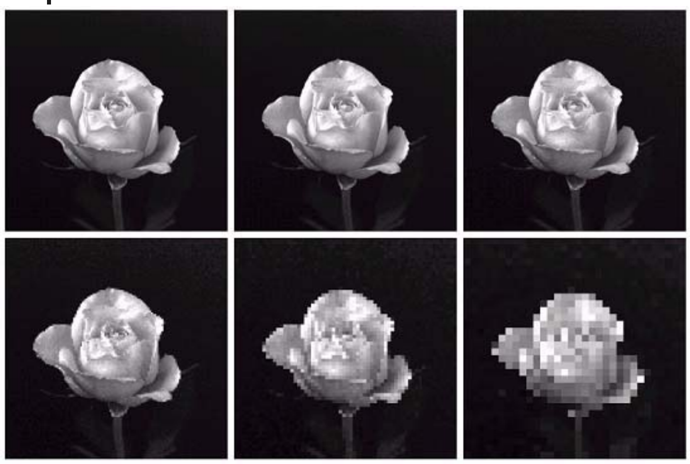
- Number of bits required to store a digitized image

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

Resolution

- Resolution (how much you can see the detail of the image) depends on sampling and gray levels.
- the bigger the sampling rate (n) and the gray scale (g), the better the approximation of the digitized image from the original.
- the more the quantization scale becomes, the bigger the size of the digitized image.

Checkerboard effect



a	b	c
d	e	f

(a) 1024x1024

(b) 512x512

(c) 256x256

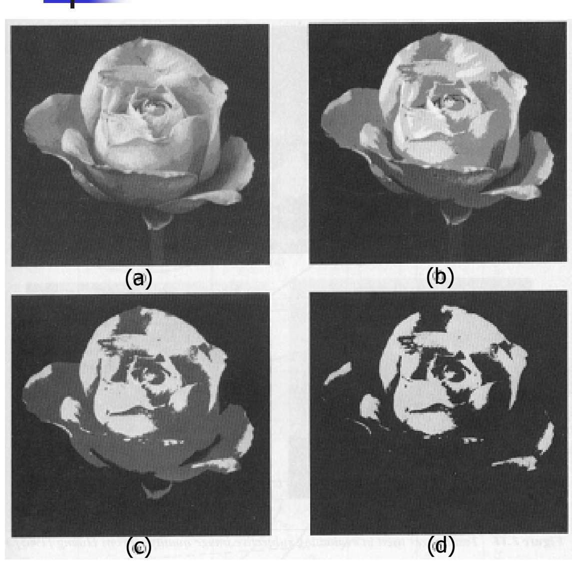
(d) 128x128

(e) 64x64

(f) 32x32

- if the resolution is decreased too much, the checkerboard effect can occur.

False contouring



(a) Gray level = 16
(b) Gray level = 8
(c) Gray level = 4
(d) Gray level = 2

- if the gray scale is not enough, the smooth area will be affected.
- False contouring can occur on the smooth area which has fine gray scales.

Nonuniform sampling

- for a fixed value of spatial resolution, the appearance of the image can be improved by using adaptive sampling rates.
  - fine sampling ⇨ required in the neighborhood of sharp gray-level transitions.
  - coarse sampling ⇨ utilized in relatively smooth regions.

Example

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.)

Example

- an image with a face superimposed on a uniform background.
  - background ⇨ little detailed information
    ⇨ coarse sampling is enough.
  - face ⇨ more detail ⇨ fine sampling.
- if we can use adaptive sampling, the quality of the image is improved.
- Moreover, we should care more around the boundary of the object ⇨ sharp gray-level transmission from object to background.

Nonuniform quantization

- unequally spaced levels in quantization process influences on the decreasing the number of gray level.
  - use few gray levels in the neighborhood of boundaries. Why ? eye is relatively poor at estimate shades of gray near abrupt level changes.
  - use more gray levels on smooth area in order to avoid the "false contouring".

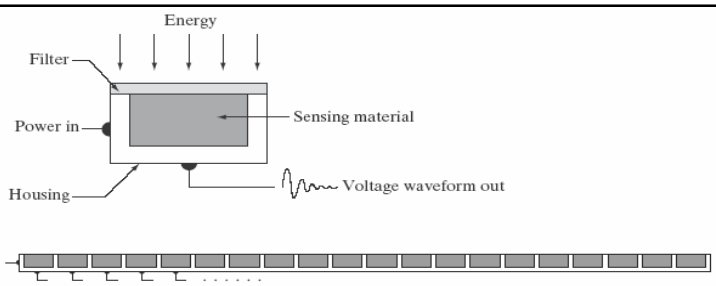
Definitions

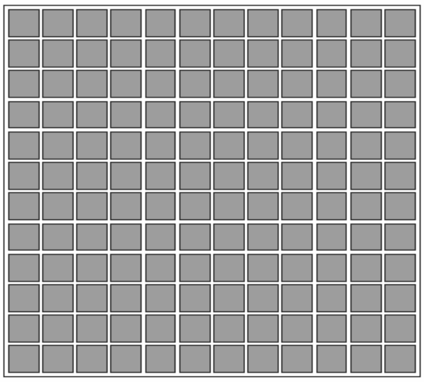
- Monochromatic (achromatic) light: Light that is void of color
  - Attribute: Intensity (amount) .
  - Gray level is used to describe monochromatic intensity.

- Chromatic light: To describe it, three quantities are used:
  - Radiance: The total amount of energy that flows from the light source (measured in Watts).
  - Luminance: The amount of energy an observer perceives from a light source (measured in lumens).
  - Brightness: A subjective descriptor of light perception that is impossible to measure (key factor in describing color sensation)

a
b
c

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





- Incoming energy (reflected or transmitted through) is transformed into a voltage.
- Sensor material is responsive to a particular type of energy being detected.

Single Sensor

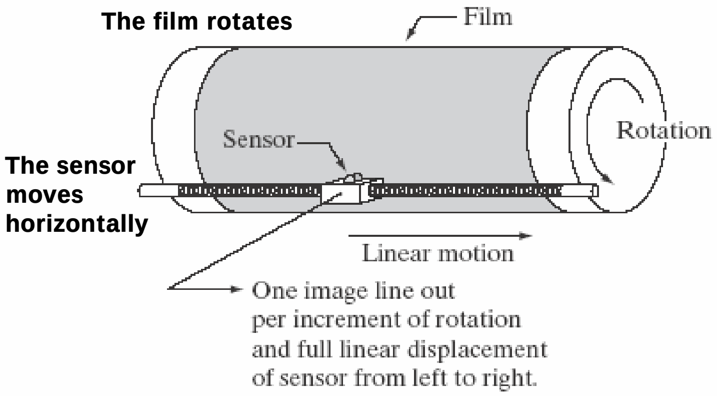


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Spatial And Gray Level Resolution

- An L-level image of size $M \times N$
- Spatial resolution:
  - # of samples per unit length or area.
  - Dots Per Inch (DPI): specifies the size of an individual pixel.
- Gray level resolution:
  - Number of bits per pixel.
  - Usually 8 bits.
  - Color image has 3 image planes to yield 8 x 3 = 24 bits/pixel.
  - Too few levels may cause false contour.

Basic Relationship b/w pixels

- Neighbors of a pixel
- Connectivity
- Labeling of Connected Components
- Relations, Equivalences, and Transitive Closure
- Distance Measures
- Arithmetic/Logic Operations

Neighbors of a pixel

- a pixel p at coordinate (x,y) has
  - $N_4(p)$: 4-neighbors of p
    (x+1, y), (x-1,y),(x,y+1), (x,y-1)

```
   x
x  p  x
   x
```

  - $N_D(p)$: 4-diagonal neighbors of p
    (x+1, y+1), (x+1,y-1),(x-1,y+1), (x-1,y-1)

```
x     x
   p
x     x
```

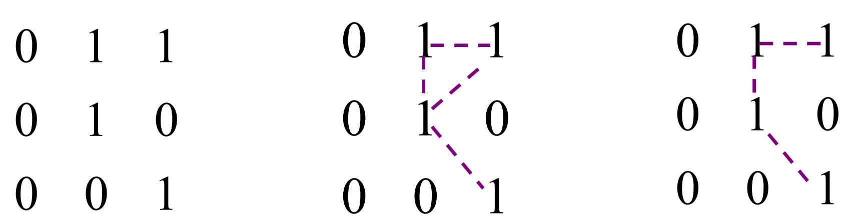
  - $N_8(p)$: 8-neighbors of p :
    a combination of $N_4(p)$ and $N_D(p)$

```
x  x  x
x  p  x
x  x  x
```

Connectivity

- Let V be the set of gray-level values used to defined connectivity
  - 4-connectivity :
    - 2 pixels p and q with values from V are 4-connected if q is in the set $N_4(p)$
  - 8-connectivity :
    - 2 pixels p and q with values from V are 8-connected if q is in the set $N_8(p)$
  - m-connectivity (mixed connectivity):
    - 2 pixels p and q with values from V are m-connected if
      - q is in the set $N_4(p)$ or
      - q is in the set $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ is empty.

Example



arrangement of pixels | 8-neighbors of the center pixel | m-neighbors of the center pixel

- m-connectivity eliminates the multiple path connections that arise in 8-connectivity.

Adjacent

- a pixel p is adjacent to a pixel q if they are connected.
- two image area subsets S1 and S2 are adjacent if some pixel in S1 is adjacent to some pixel S2.

Exercise

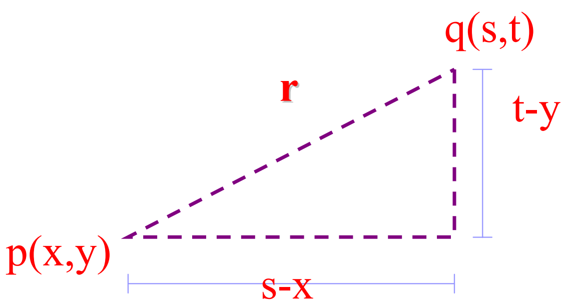
- Consider the two image subsets S_1 and S_2 :

		S_1					S_2		
0	0	0	0	0	0	0	1	1	0
1	0	0	1	0	0	1	0	0	1
1	0	0	1	0	1	1	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	1	1	1

- For $V=\{1\}$, determine whether S_1 and S_2 are
  - 4-connected
  - 8-connected
  - m-connected

Euclidean distance between p and q

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



City-block distance: D_4 distance

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

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

- diamond centered at (x,y)
- D_4 =1 are 4-neighbors of (x,y)

Chessboard distance: D_8 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

square centered at (x,y)

Arithmetic Operators

- used extensively in most branches of image processing.
- Arithmetic operations for 2 pixels p and q :
  - Addition : p+q used in image average to reduce noise.
  - Subtraction : p-q basic tool in medical imaging.
  - Multiplication : pxq
    - to correct gray-level shading result from non-uniformities in illumination or in the sensor used to acquire the image.
  - Division : p÷q
- Arithmetic Operation entire images are carried out pixel by pixel.

Logic operations

- AND : p AND q (p•q)
- OR : p OR q (p+q)
- COMPLEMENT : NOT q ($\bar{q}$)
- logic operations apply only to binary images.
- arithmetic operations apply to multivalued pixels.
- logic operations used for tasks such as masking, feature detection, and shape analysis.
- logic operations perform pixel by pixel.

Mask Operation

- Besides pixel-by-pixel processing on entire images, arithmetic and Logical operations are used in neighborhood oriented operations.

		⋮		
	Z_1	Z_2	Z_3	
…	Z_4	Z_5	Z_6	…
	Z_7	Z_8	Z_9	
		⋮		

Mask Operation

- Let the value assigned to a pixel be a function of its gray level and the gray level of its neighbors.
- e.g., replace the gray value of pixel Z_5 with the average gray values of it's neighborhood within a 3x3 mask.

$$Z = \frac{1}{9}(Z_1 + Z_2 + Z_3 + \ldots + Z_9)$$

Mask operator

- In general term:

$$Z = \frac{1}{9}Z_1 + \frac{1}{9}Z_2 + \frac{1}{9}Z_3 + \ldots + \frac{1}{9}Z_9$$
$$= w_1 Z_1 + w_2 Z_2 + w_3 Z_3 + \ldots w_9 Z_9$$
$$= \sum_{i=1}^{9} w_i Z_i$$

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

→

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Mask coefficient

- Proper selection of the coefficients and application of the mask at each pixel position in an image makes possible a variety of useful image operations
  - noise reduction
  - region thinning
  - edge detection
- Applying a mask at each pixel location in an image is a computationally expensive task.