Chapter 2: Digital Image Processing
Digital Image Fundamentals

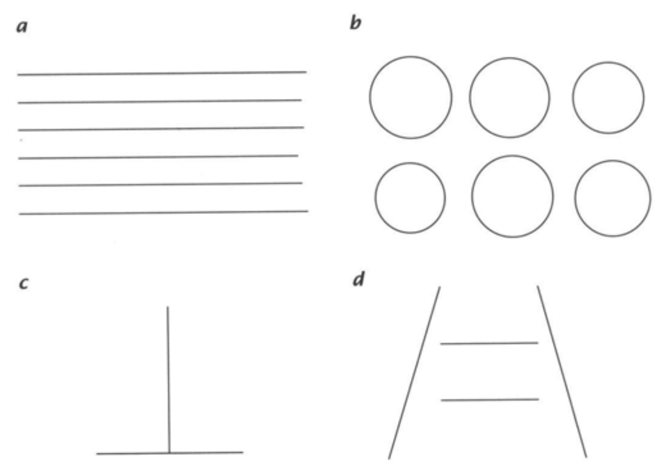
Human and Computer Vision

- 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.
- Without taking this elementary fact into consideration, we may be much misled in the interpretation of images.

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?

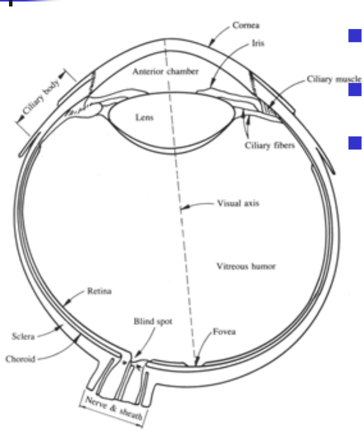
Test images



Test images

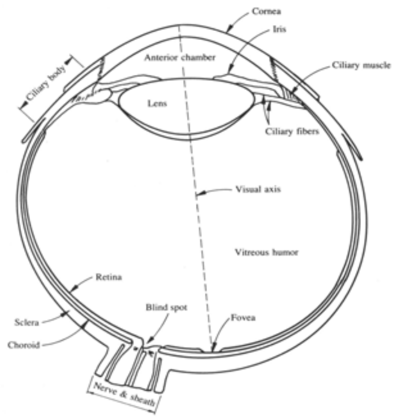
- Test images for distances and area estimation:
  a) Parallel lines with up to 5% difference in length.
  b) Circles with up to 10% difference in radius.
  c) The vertical line appears longer but actually has the same length as the horizontal line.
  d) Deception by perspective: the upper line appears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

Structure of the Human Eye



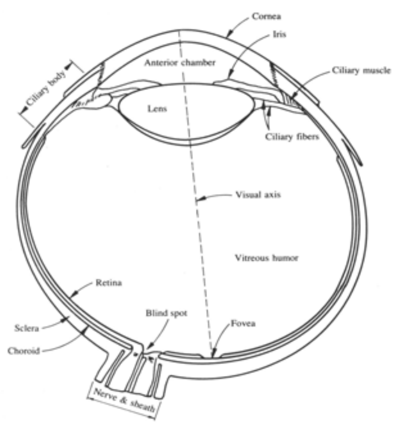
- Shape is nearly a sphere.
- Average diameter = 20 mm.
- 3 membranes:
  - Cornea and Sclera - outer cover
  - Choroid
  - Retina -enclose the eye

Structure of the Human Eye



- Cornea :
  - tough, transparent tissue, covers the anterior surface of the eye.
- Sclera :
  - Opaque membrane, encloses the remainder of the optic globe

Structure of the Human Eye



- Choroid :
  - Lies below the sclera, contains network of blood vessels that serve as the major source of nutrition to the eye.
  - Choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe.

Lens & Retina

- Lens

  both infrared and ultraviolet light are absorbed appreciably by proteins within the lens structure and, in excessive amounts, can cause damage to the eye.

- Retina

  Innermost membrane of the eye which lines inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina.

Receptors

- Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina.
- Receptors are divided into 2 classes:
  - Cones
  - Rods

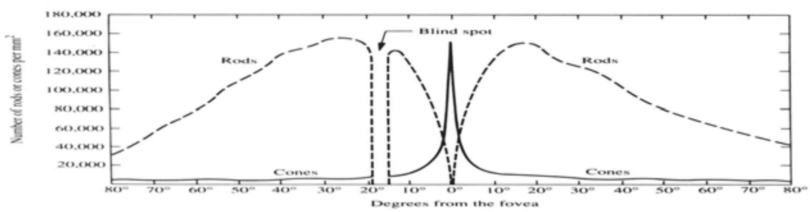
Cones

- 6-7 million, located primarily in the central portion of the retina (the fovea, 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.

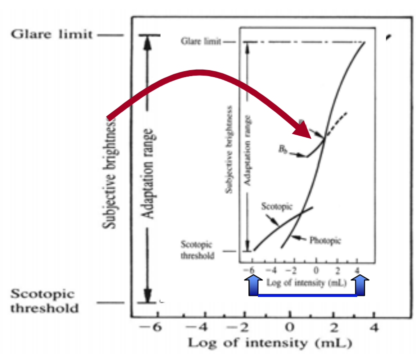
Cross section of the right eye



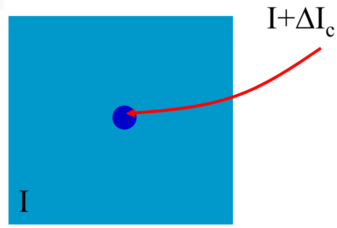
- Blind spot ⇨ the absence of receptors area.
- Receptor density is measured in degrees from the fovea.
- Cones are most dense in the center of the retina (in the area of the fovea).
- Rods increase in density from the center out to approx. 20° off axis and then decrease in density out to the extreme periphery of the retina.

Brightness adaptation and discrimination

- The total range of intensity levels it can discriminate simultaneously is rather small compared with the total adaptation range.
- B_a is a brightness adaptation level. The short intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level.



Contrast sensitivity



Weber's ratio: $\Delta I_c/I$

Good brightness discrimination

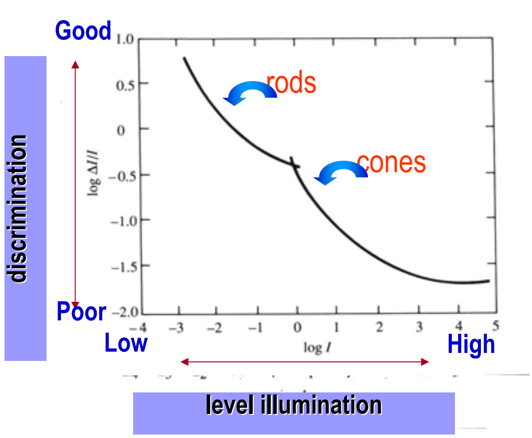
⇨ $\Delta I_c I$ small

Bad brightness discrimination

⇨ $\Delta I_c/I$ is large.

- The ability of the eye to discrimination b/w changes in brightness at any specific adaptation level is of considerable interest.
- I is uniformly illumination on a flat area large enough to occupy the entire field of view.
- ΔI_c is the change in the object brightness required to just distinguish the object from the background

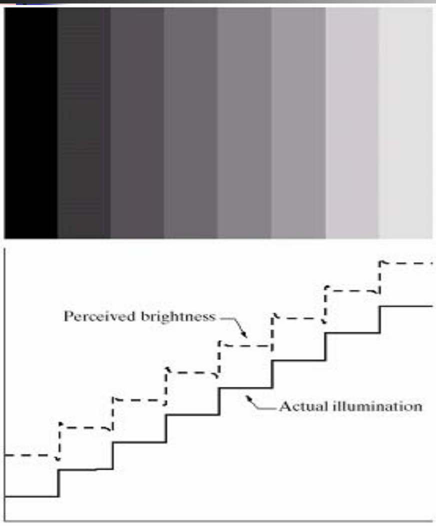
Weber ratio



brightness discrimination is poor (the Weber ratio is large) at low levels of illumination and improves significantly (the ratio decreases) as background illumination increases.

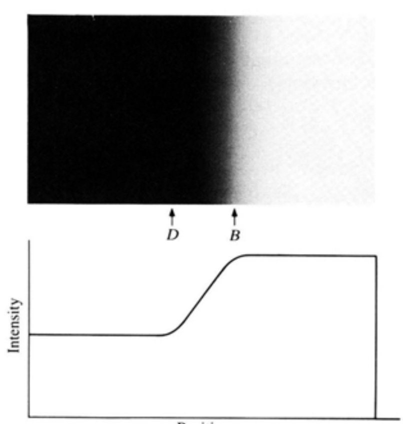
hard to distinguish the discrimination when it is bright area but easier when the discrimination is on a dark area.

Brightness vs.Function of intensity



- Brightness is not a simple function of intensity.
- visual system tends to undershoot or overshoot around the boundary of regions of different intensities.
- the intensity of the stripes is constant but we actually perceive a brightness pattern is strongly scalloped near the boundaries.

Mach band pattern



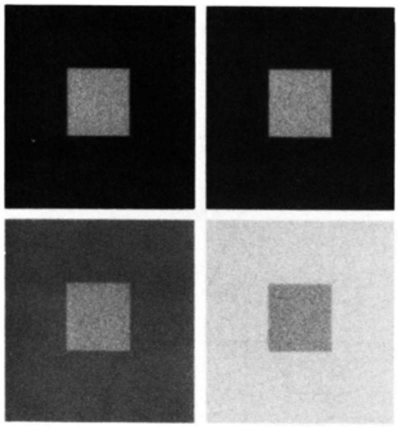
Is it the same level of darkness around area D and B ?

- The brightness pattern perceived is a darker stripe in region D and a brighter one in the region B whereas actually the region from D to B has the same intensity.

Note: it was named for Ernst Mach who discovered the phenomenon in 1865.

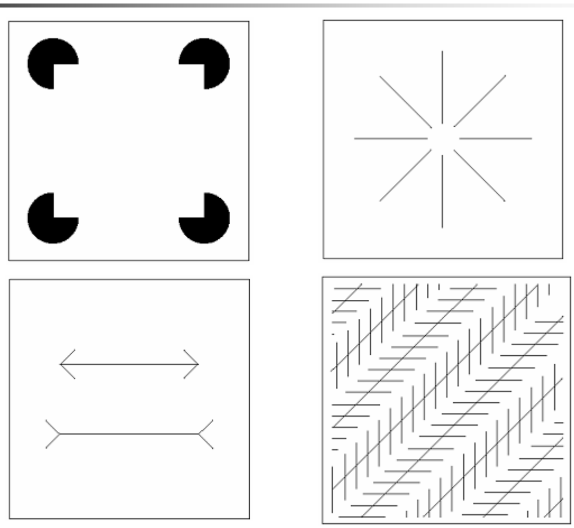
Simultaneous contrast

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.
- Region's perceived brightness does not depend simply on its intensity.

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?

$$f(x,y,t)$$

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 have are discrete samples of the continuous domain.
- in other words, signals manipulated by computer have discrete domains.

Analog & Digital

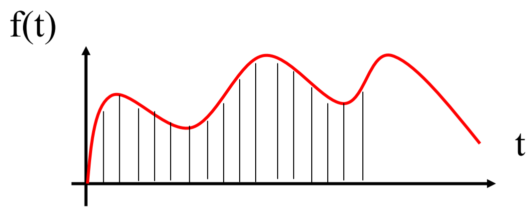
- most naturally-occurring signals also have a real-valued range in which values occur with infinite precision.
- to store and manipulate signals by computer we need to store these numbers with finite precision. thus, these signals have a discrete range.

signal has continuous domain and range = analog

signal has discrete domain and range = digital

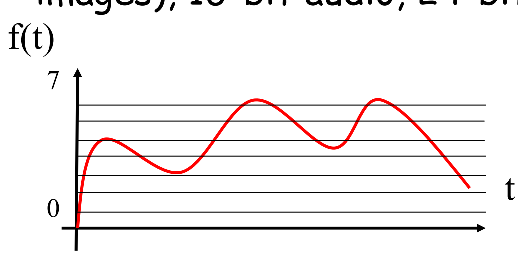
Sampling

- 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.



8 levels = 2^3 : uses 3 bits to represent the value of the function.

Digital Image Representation

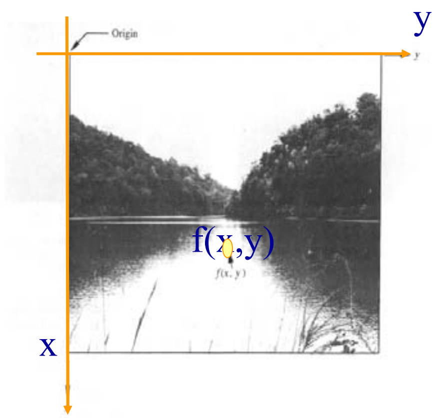


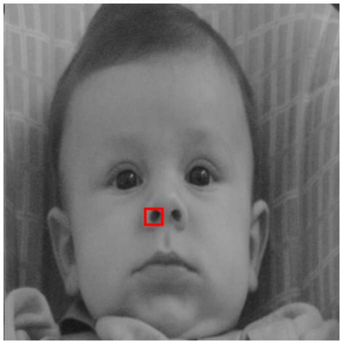
Figure 1.5 Axis convention used for digital image representation.

- A digital image is an image *f(x,y)* that has been digitized both in spatial coordinates and brightness.
- the value of f at any point *(x,y)* is proportional to the brightness (or gray level) of the image at that point.

Digital Image Representation

A digital image can be considered a matrix whose row and column indices identify a point in the image and the corresponding matrix element value identifies the gray level at that point.

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

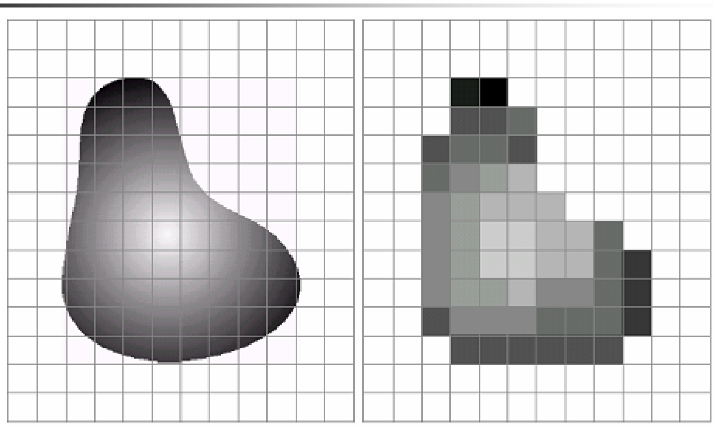
CAMERA ⇨ DIGITIZER

Samples the analog data and digitizes it.

spacing of discrete values in the … of signal

signal -> domain
quantization -> range

Example of Digital Image



Continuous image projected onto a sensor array

Result of image sampling and quantization

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.

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

Illumination and Reflectance

- the basic nature of $f(x,y)$ may be characterized by 2 components:
  - the amount of source light incident on the scene being viewed ⇨ Illumination, $i(x,y)$
  - the amount of light reflected by the objects in the scene ⇨ Reflectance, $r(x,y)$

Illumination and Reflectance

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

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

determined by the nature of the light source

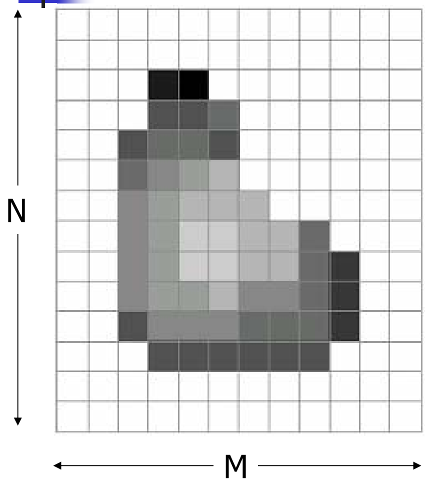
$$0 < r(x,y) < 1$$

determined by the nature of the objects in a scene ⇨ bounded from total absorption to total reflectance.

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.
- thus, l lies in the range $L_{\min} \le l \le L_{\max}$
- L_{min} is positive and L_{max} is finite.
- gray scale = [L_{min}, L_{max}]
- common practice, shift the interval to [0, L]
- $0 = black$, $L = white$

Number of bits



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

$$L = 2^k$$

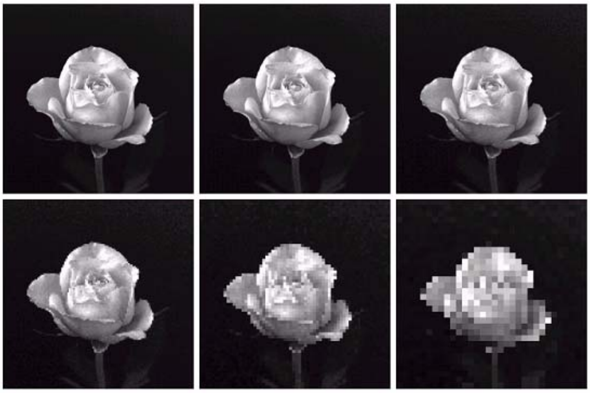
- Number of bits required to store a digitized image

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

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



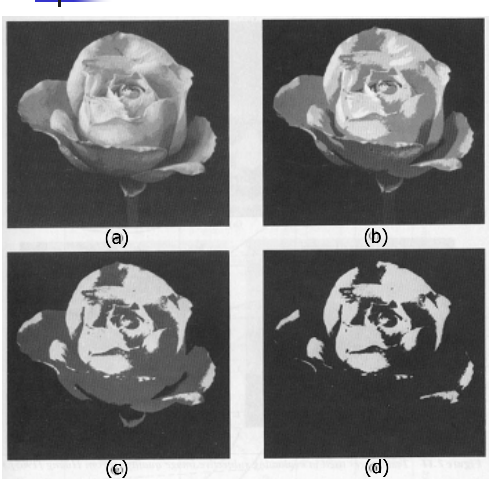
a	b	c
d	e	f

(a) 1024x1024
(b) 512x512
(c) 256x256
(d) 128x128
(e) 64x64
(f) 32x32

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

checkerboard effect : insufficient sampling
false contouring : insufficient quantization

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.

fine sampling : applied in sharp regions
coarse : applied in 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.

few gray levels in boundry
more gray levels in smooth area

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".

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.
      - (the set of pixels that are 4-neighbors of both p and q whose values are from V)

Example

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

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.

Path

- a path from pixel p with coordinates (x,y) to pixel q with coordinates (s,t) is a sequence of distinct pixels with coordinates

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

  where $(x_0,y_0) = (x,y)$, $(x_n,y_n) = (s,t)$ and (x_i,y_i) is adjacent to (x_{i-1},y_{i-1})
- n is the length of the path
- we can define 4-,8-, or m-paths depending on type of adjacency specified.

Exercise

- Consider the two image subsets S_1 and S_2 :

$$
\begin{array}{c|cccc|cccc|c}
 & \multicolumn{4}{c|}{S_1} & \multicolumn{4}{c|}{S_2} & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1
\end{array}
$$

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

Labeling of Connected Components

- scan the image from left to right
- Let p denote the pixel at any step in the scanning process.
- Let r denote the upper neighbor of p.
- Let t denote the left-hand neighbors of p, respectively.
- when we get to p, points r and t have already been encountered and labeled if they were 1's.

$$
\begin{array}{cc}
 & r \\
t & p
\end{array}
$$

Labeling of Connected Components

- if the value of p = 0, move on.
- if the value of p = 1, examine r and t.
  - if they are both 0, assign a new label to p.
  - if only one of them is 1, assign its label to p.
  - if they are both 1
    - if they have the same label, assign that label to p.
    - if not, assign one of the labels to p and make a note that the two labels are equivalent. (r and t are connected through p).
- at the end of the scan, all points with value 1 have been labeled.
- do a second scan, assign a new label for each equivalent labels.

What shall we do with 8-connected components ?

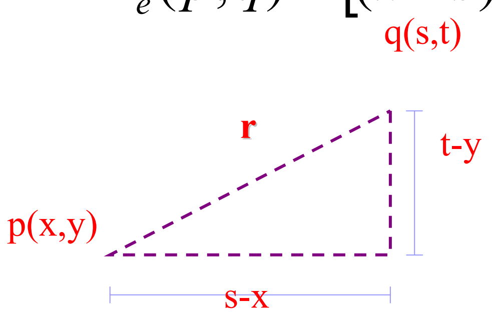
- do the same way but examine also the upper diagonal neighbors of p.
  - if p is 0, move on.
  - if p is 1
    - if all four neighbors are 0, assign a new label to p.
    - if only one of the neighbors is 1, assign its label to p.
    - if two or more neighbors are 1, assign one of the label to p and make a note of equivalent classes.
    - after complete the scan, do the second round and introduce a unique label to each equivalent class.

Distance Measures

- for pixel p, q and z with coordinates (x,y), (s,t) and (u,v) respectively,
- D is a distance function or metric if
  - (a) $D(p,q) \geq 0$; $D(p,q) = 0$ iff $D=q$
  - (b) $D(p,q) = D(q,p)$
  - (c) $D(p,z) \leq D(p,q) + D(q,z)$

Euclidean distance between p and q

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



radius (r) centered at (x,y)

City-block distance: D_4 distance

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

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

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

Chessboard distance: D_8 distance

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

max(lx-sl , ly-tl)

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

square centered at (x,y)

D_4 and D_8 distances

- D_4 distance and D_8 distance between pixel p and q = length 4- and 8-path between those 2 points, respectively.
- we can consider both D_4 and D_8 distances b/w p and q regardless of whether a connected path exists between them because the definitions of these distances involve only the coordinates.

m-connectivity's distance

- distances of m-connectivity of the path between 2 pixels depends on values of pixels along the path.
- e.g., if only connectivity of pixels valued 1 is allowed. find the m-distance b/w p and p_4

	p_3	p_4		0	1		1	1		1	1
p_1	p_2		0	1		0	1		1	1	
p			1			1			1		

distance = 2 distance = 3 distance = 4

Arithmetic Operators

- used extensively in most branches of image processing.
- Arithmetic operations b/w 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 nonuniformities 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.

		$\vdots$		
	Z_1	Z_2	Z_3	
$\cdots$	Z_4	Z_5	Z_6	$\cdots$
	Z_7	Z_8	Z_9	
		$\vdots$		

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.

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

Mark 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.

Image Geometry

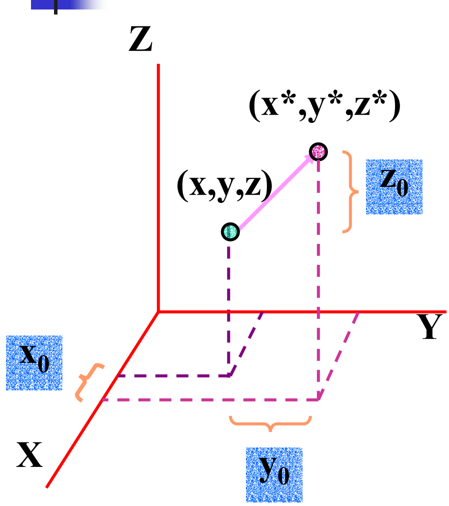
- Basic Transformations : expressed in 3D Cartesian coordinate system (x,y,z)
  - Translation
  - Scaling
  - Rotation
  - Concatenation and inverse transformation

$x^* = x + x_0$

$y^* = y + y_0$

$z^* = z + z_0$

Translation



$$\begin{bmatrix} x^* \\ y^* \\ z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Geometry General Form

$$\mathbf{V^* = AV}$$

A : 4x4 transformation matrix

V : column vector containing the original coordinates

V* : column vector whose components are the transformed coordinates

Translation matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$V^* = TV$$

Scaling

- scaling by factors S_x, S_y and S_z along the X, Y, Z axes

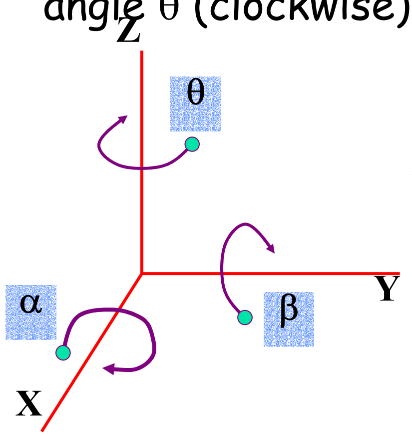
$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$x^* = S_x x$
$y^* = S_y y$
$z^* = S_z z$

Rotation

$x^* = x \cos\theta + y \sin\theta$
$y^* = -x \sin\theta + y \cos\theta$
$z^* = z$

- rotation of a point about Z axis by an angle θ (clockwise)



$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation

- rotation of a point about X axis by an angle α

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation

- rotation of a point about Y axis by an angle β

$$R_\beta = \begin{bmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

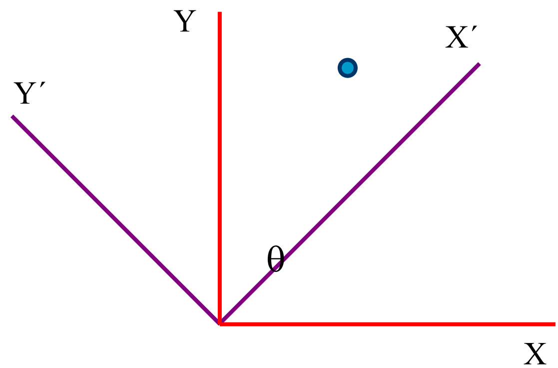
if we rotate a point about one axis, it will affect on the coordination of the other 2 axes.

Exercise

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Rotate a point about Z axis.

Proof



Concatenation and inverse transformation

- Several transformations can be represented by a single 4x4 transformation matrix.
- e.g., translation, scaling, and rotation about the z axis of a point V is given by

$$V^* = R_\theta(S(TV))$$
$$= AV$$

Order is important!!!