CTT310: Digital Image Processing

Digital Image Fundamentals

Dr. Nguyen Ngoc Thao
Department of Computer Science, FIT, HCMUS

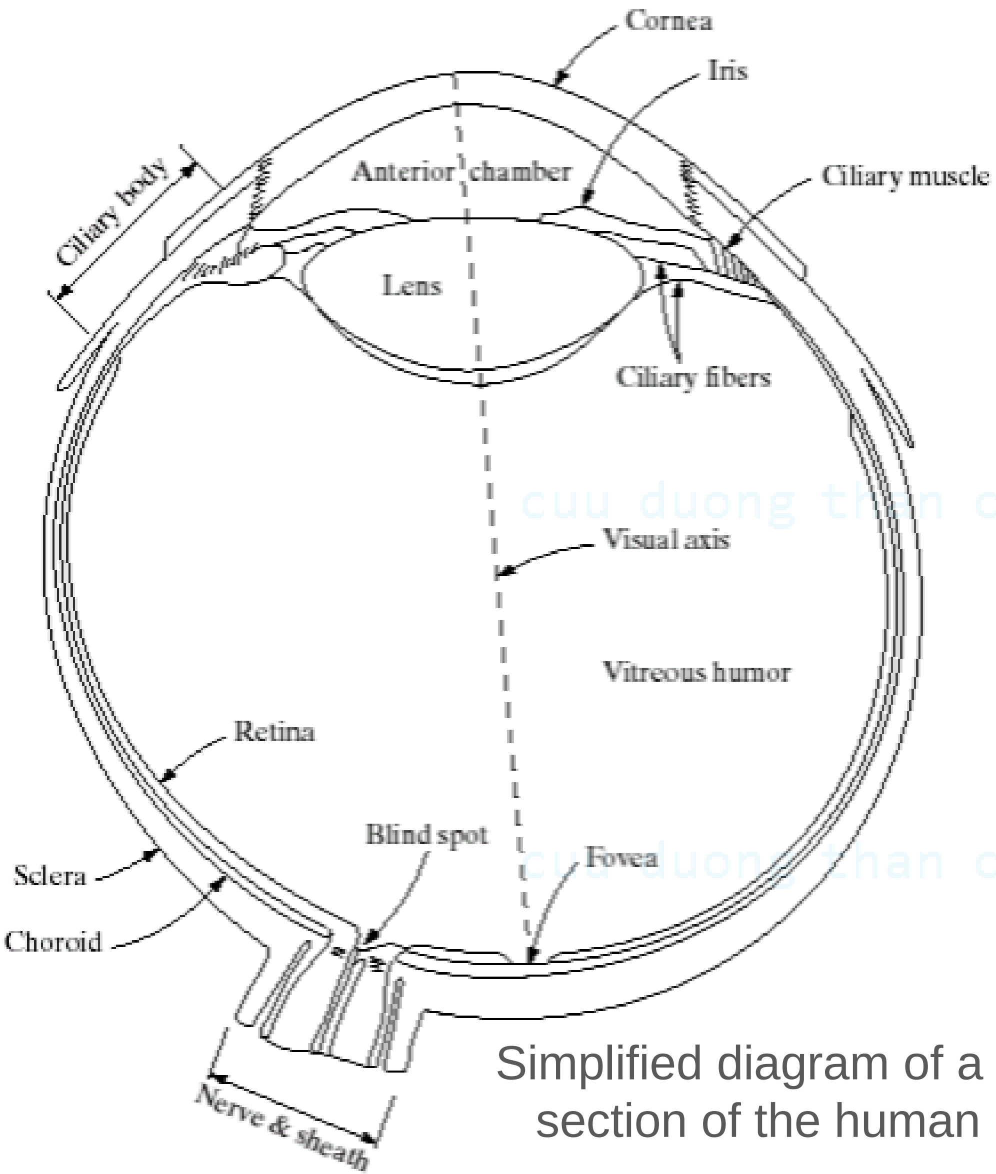
Outline

- Elements of visual perceptions
- Light and electromagnetic spectrum
- Image sensing and acquisition
- Image sampling and quantization
- Some basic relationships between pixels
- Mathematical tools used in digital image processing

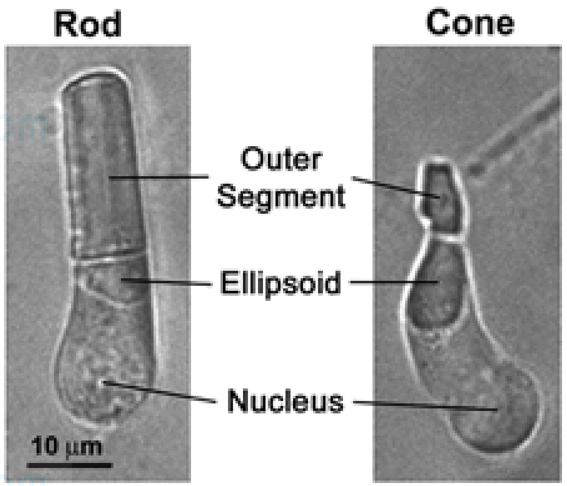
Section 2.1

ELEMENTS OF VISUAL PERCEPTIONS

Structure of the human eye



Simplified diagram of a cross section of the human eye



Light receptors

Structure of the human eye: Demo

Light receptor: cones and rods

Cones

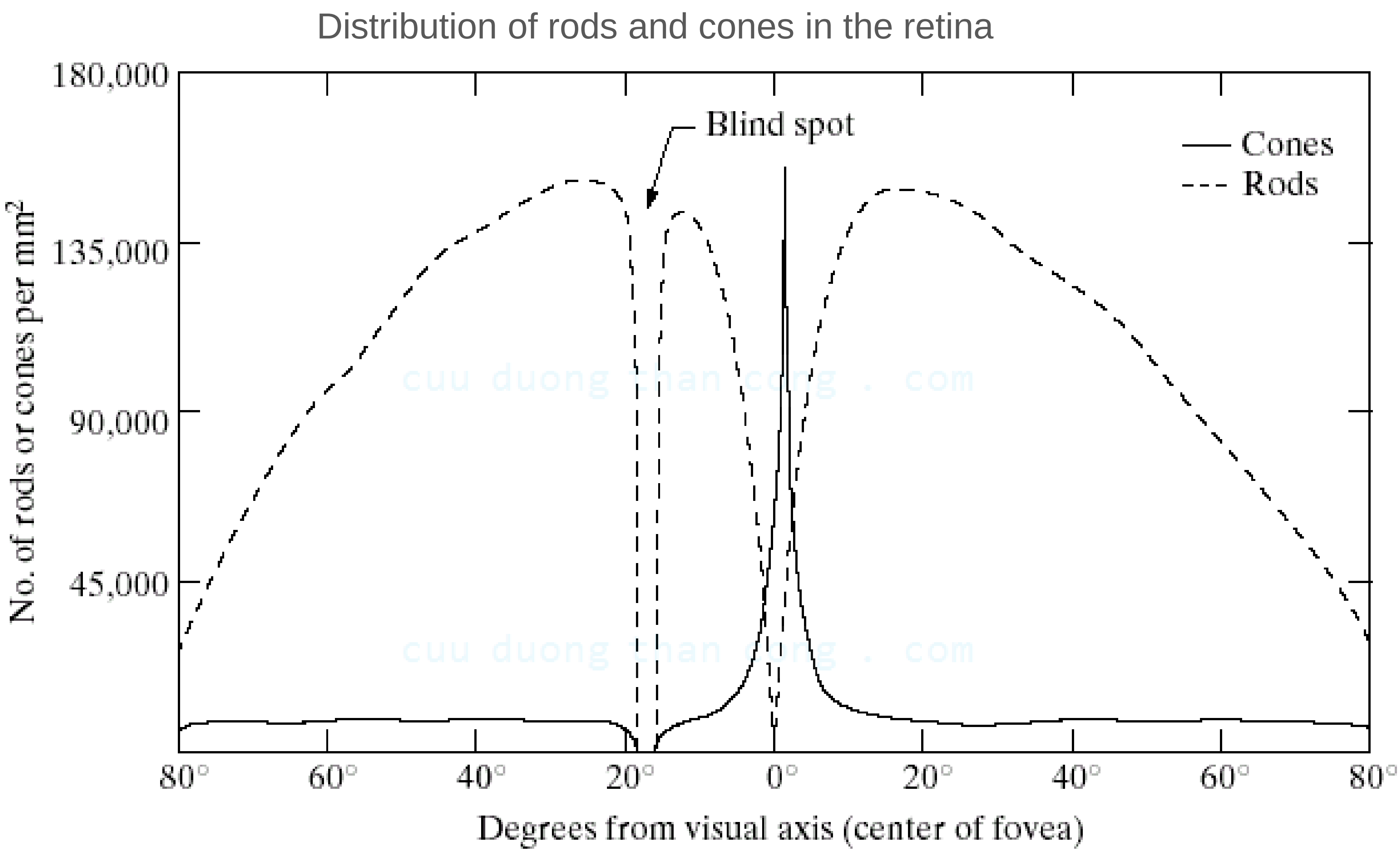
- **6 – 7 millions** in each eye, located primarily in the fovea
- Each cone is connected to its own nerve end
- Highly sensitive to color, fine detail discernible
- Cone vision is called **photopic** or **bright-light vision**

Rods

- **75 – 150 millions**, distributed over the retinal surface
- Several rods are connected to a single nerve end
- Not involved in color vision, sensitive to low levels of illumination
- Rod vision is called **scotopic** or **dim-light vision**

For example, objects that appear brightly colored in daylight when seen by moonlight appear as colorless forms because only the rods are stimulated

Light receptor: cones and rods



Demonstration of blind spot

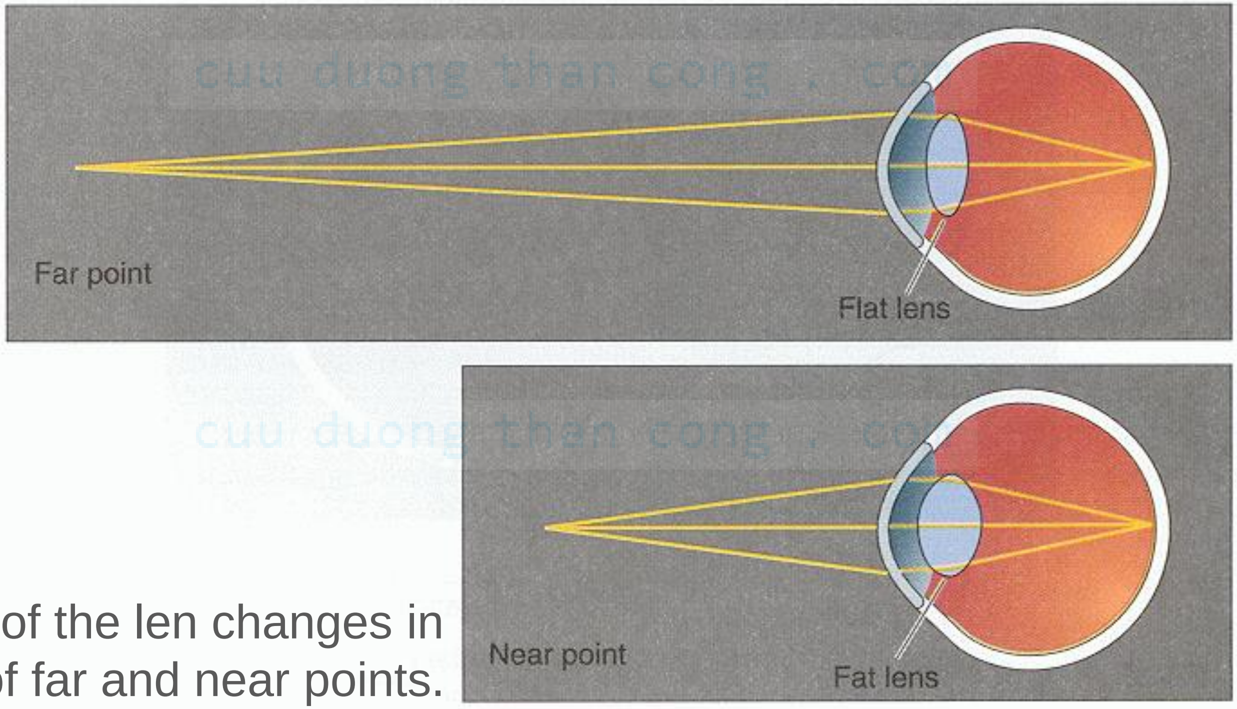
- **Blind spot** is the region of emergence of the optic nerve from the eye
- For example,

R **L**

Instructions: Close one eye and focus the other on the appropriate letter (R for right or L for left). Place your eye a distance from the screen approximately equal to 3× the distance between the R and the L. Move your eye towards or away from the screen until you notice the other letter disappear. For example, close your right eye, look at the "L" with your left eye, and the "R" will disappear.

Image formation in the human eye

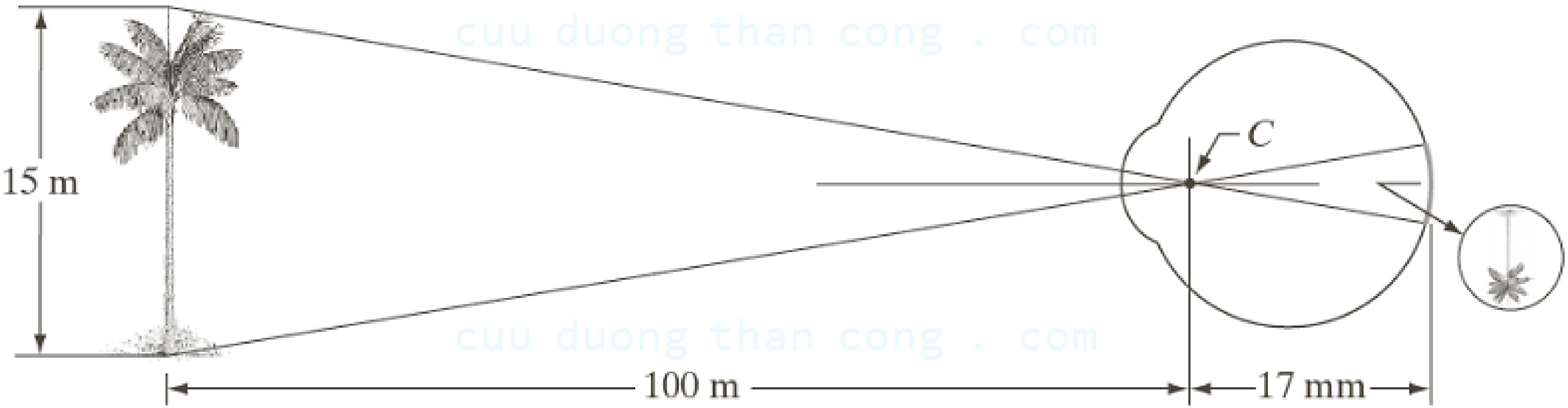
- The distance between the lens and the imaging region (the retina) is fixed.
- The focal length needed to achieve proper focus is obtained by varying the shape of the lens.



The shape of the len changes in the cases of far and near points.

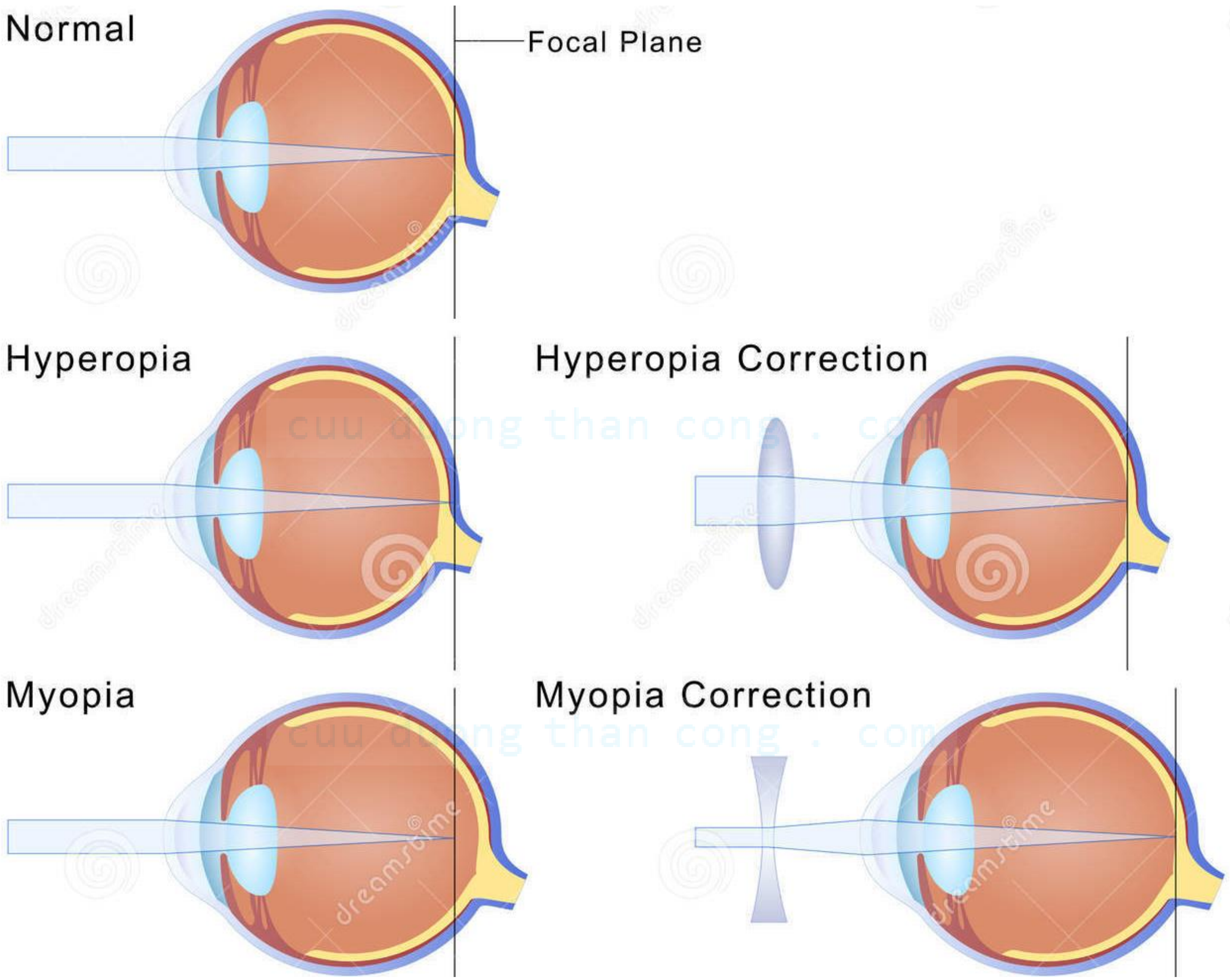
Image formation in the human eye

- The distance between the center of the lens and the retina along the visual axis is **approximately 17 mm**.
- The range of focal lengths is approximately **14 – 17 mm**.



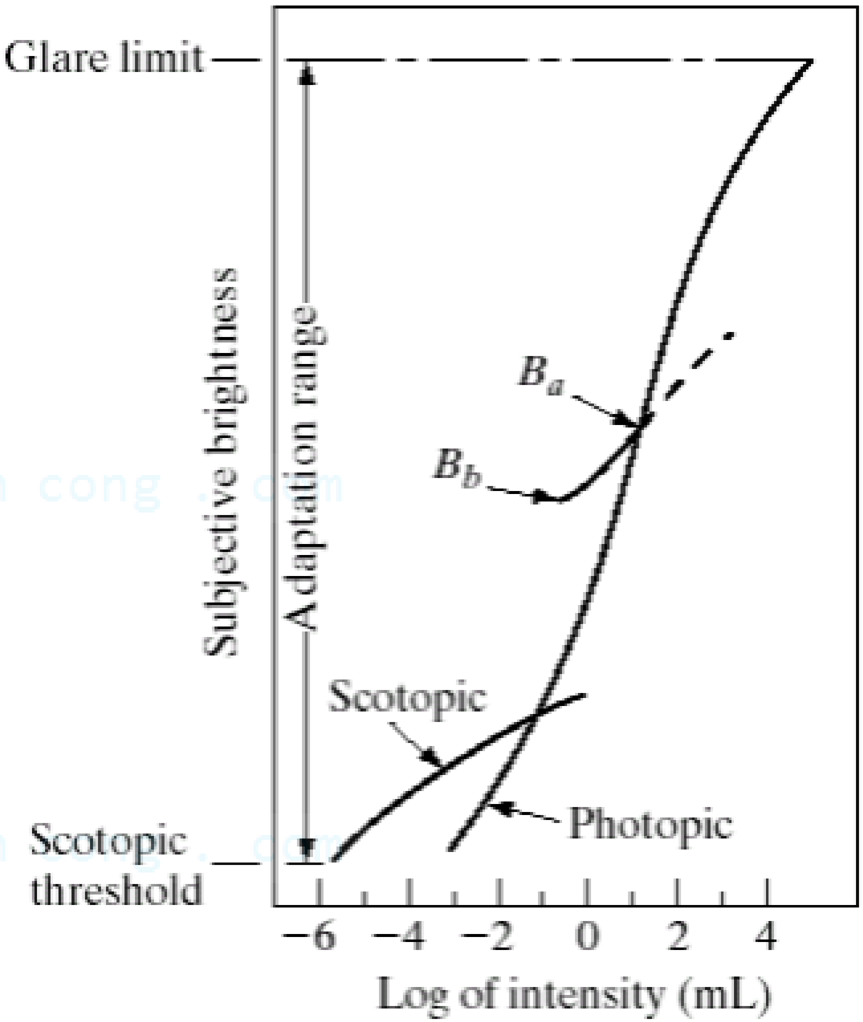
Graphical representation of the eye looking at a palm tree.
Point *C* is the optical center of the lens

Hyperopia and Myopia



Brightness adaptation

- **Subjective brightness**: intensity as perceived by the human visual system
- **Brightness adaptation**: the total range of distinct intensity levels the eye can *discriminate simultaneously* is rather small when compared with the total adaptation range

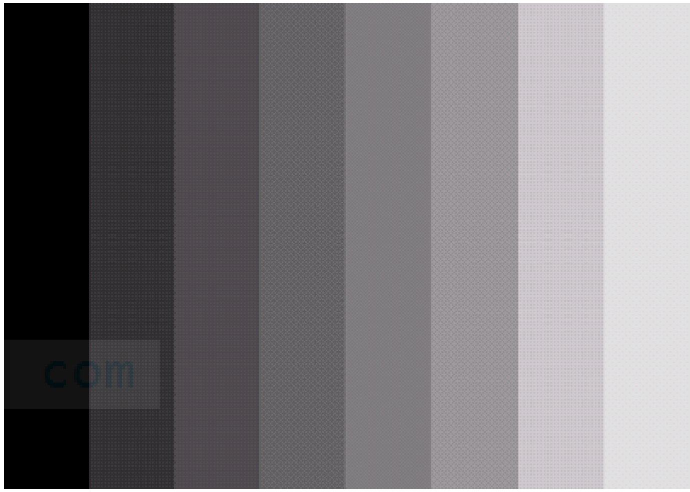


Range of subjective brightness sensations showing a particular adaptation level

Human perception phenomena

- Perceived brightness is not a simple function of intensity

- **Mach band effect** (Ernst Mach, described in 1865): the visual system tends to undershoot or overshoot around the boundary of regions of different intensities



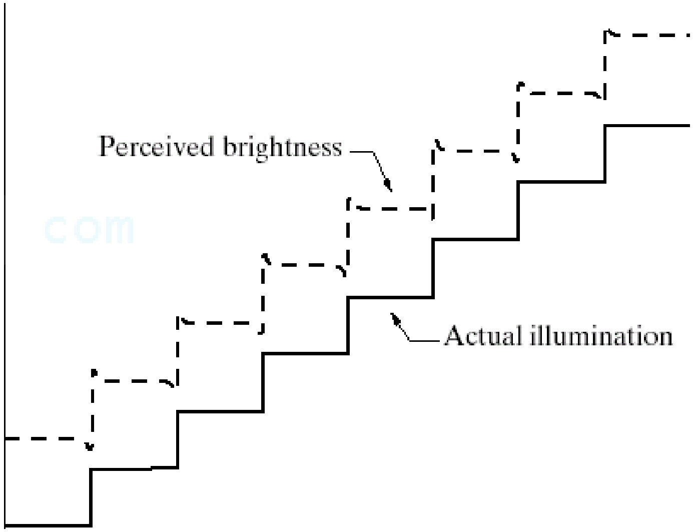
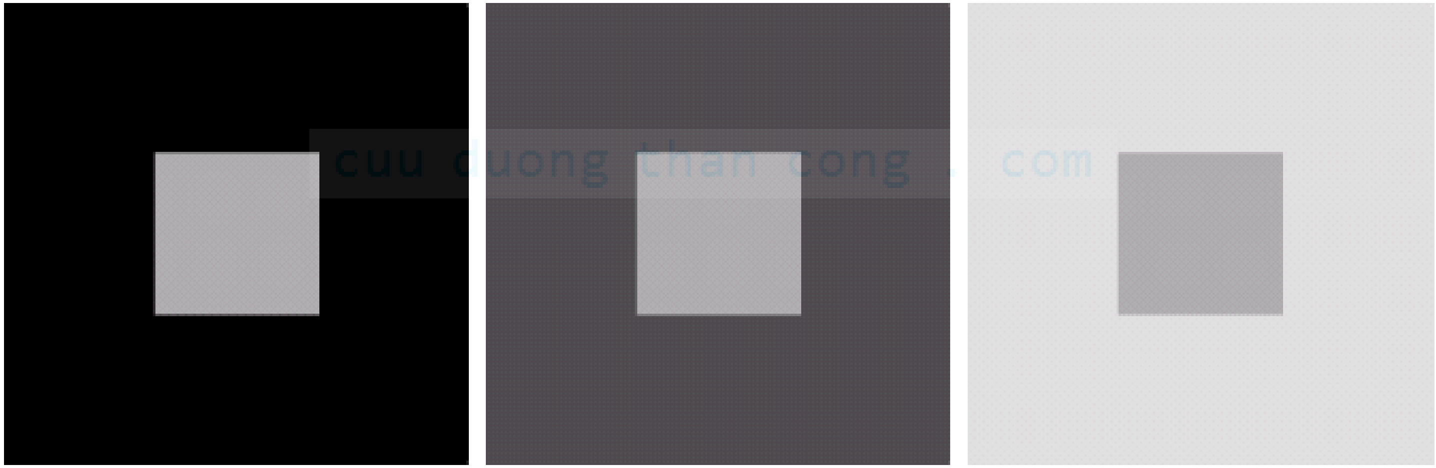


Illustration of the Mach band effect

Human perception phenomena

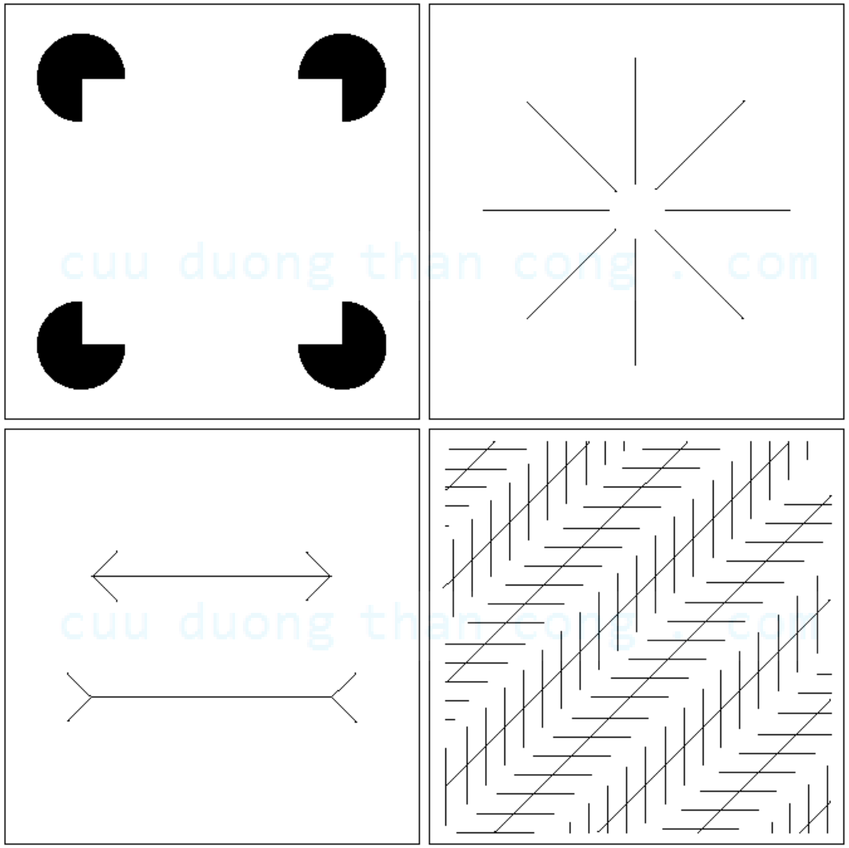
- **Simultaneous contrast:** the perceived brightness of a region does not depend simply on its intensity.



All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Human perception phenomena

- **Optical illusions**: the eye fills in nonexistence information or wrongly perceives geometrical properties of objects

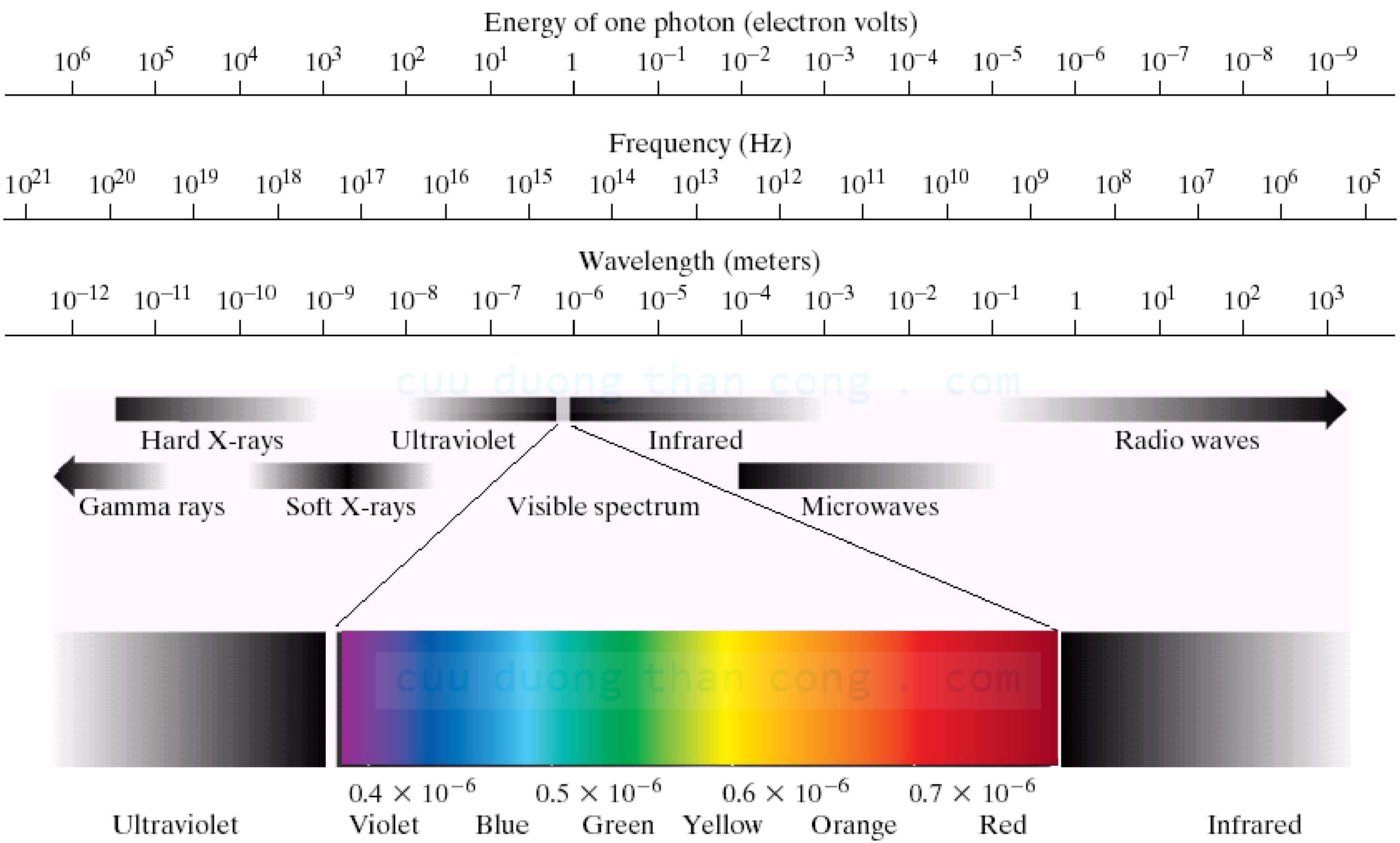


Some well-known optical illusions

Section 2.2

LIGHT AND THE ELECTROMAGNETIC SPECTRUM

The electromagnetic (EM) spectrum



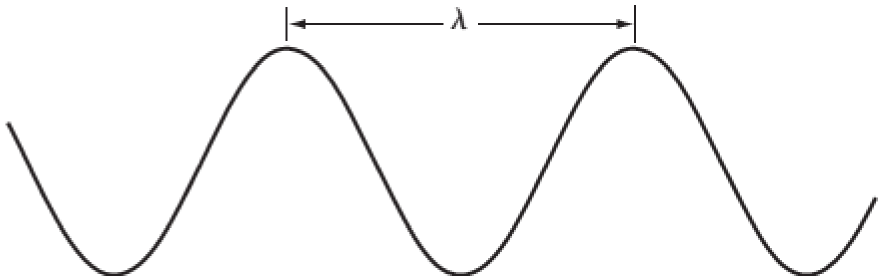
Note that the visible spectrum is a rather narrow portion of the EM spectrum.

Wavelength, frequency, and energy

- **Wavelength** λ (*in meters*) and **frequency** v (*in Hertz*) are related by the expression

$$\lambda = \frac{c}{v}$$

  - c is the speed of light $(2.998 \times 10^8\ m/s)$



Graphical representation of one wavelength

- The energy of the various components (*in electron-volt*) of the electromagnetic spectrum is given by the expression

$$\boldsymbol{E = hv}$$

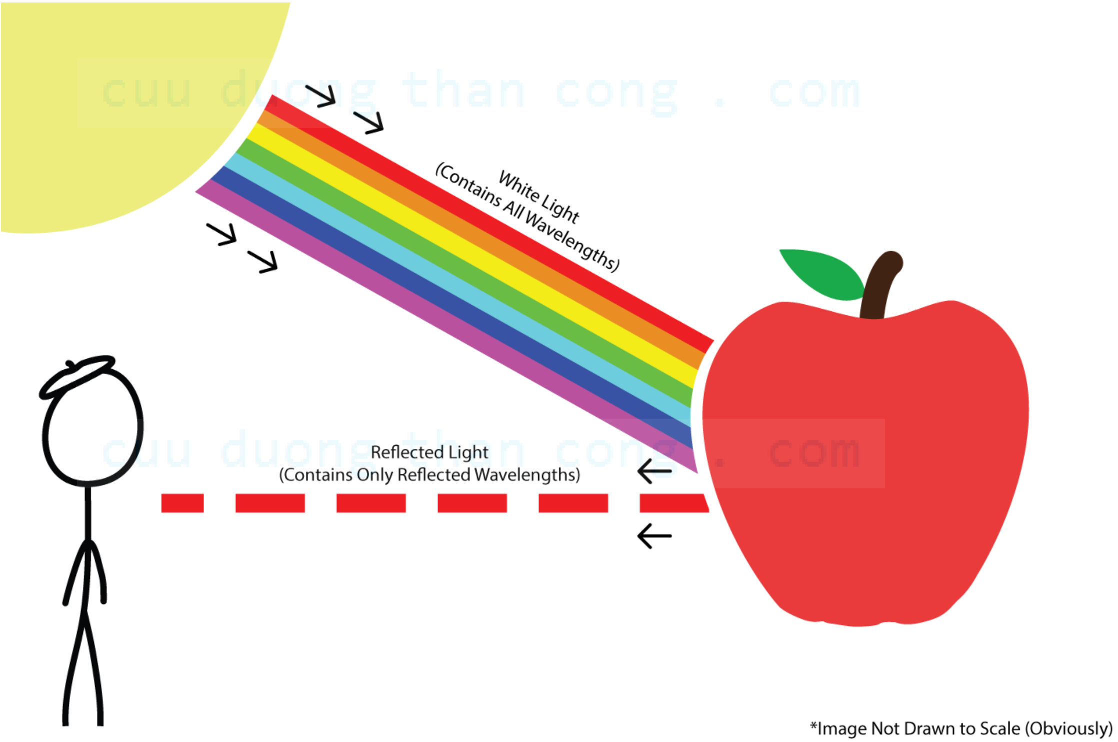
  - Where h is Planck's constant

Imaging a band of EM spectrum

- A sensor developed to detect energy radiated by a band of the electromagnetic spectrum is needed.
- The wavelength of an electromagnetic wave required to "see" an object must be of the same size as or smaller than the object.
  - E.g., a water molecule has a diameter on the order of 10^{-10}. Thus, to study molecules, we would need a source capable of emitting in the far ultraviolet or soft X-ray region.

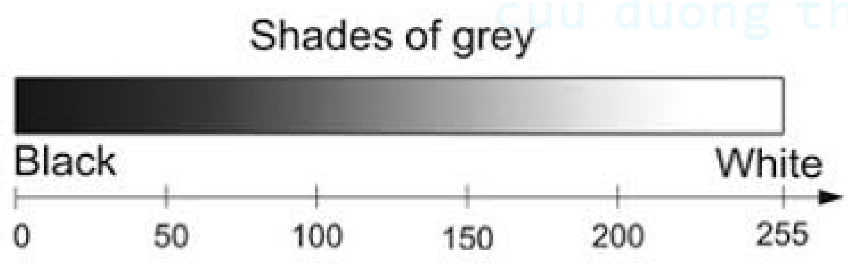
Visible light

- **Light** is a particular type of electromagnetic radiation that can be sensed by the human eye.
- The **colors** humans perceive in an object are determined by the nature of the light *reflected* from the object



Visible light: Monochromatic light

- Light that is void of color is called **monochromatic** (or achromatic) light
- Monochromatic intensity varies from black to gray and finally to white, which is called **intensity** (or **gray level**).



Gray scale: The range of measured values of monochromatic light from black to white

Gray-scale image: monochromatic images

Visible light: Chromatic light

- **Chromatic** (or color) light spans the electromagnetic energy spectrum from approximately 0.43 to 0.79 μm
- In addition to frequency, three following basic quantities are used to describe the quality of a chromatic light source
  - Radiance (in Watts): the amount of energy that flows from the light source
  - Luminance (in *lumens*): the amount of energy an observer perceives from a light source
  - Brightness: a subjective descriptor of light perception that is practically impossible to measure

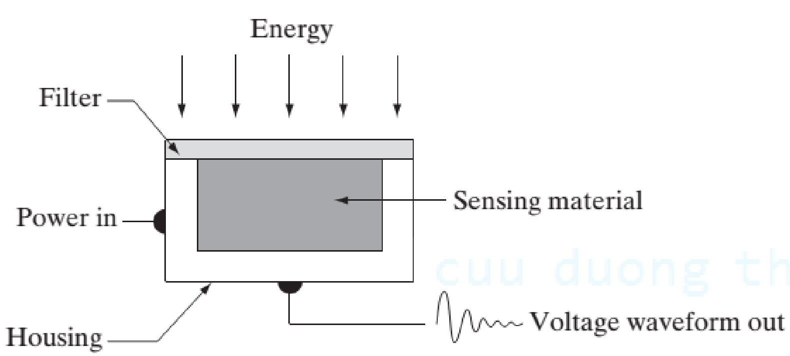
Chromatic (color) image

Section 2.3

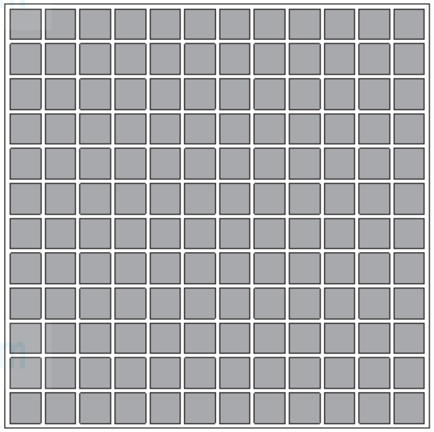
IMAGE SENSING AND ACQUISITION

Image sensing and generation

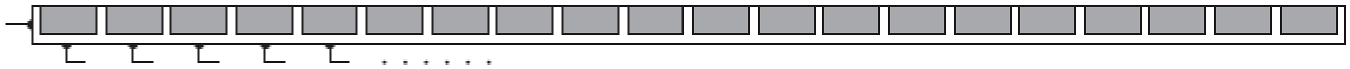
- Incoming energy is transformed into a voltage.
  - Combination of input electrical power and sensor material that is responsive to the particular type of energy being detected
  - The sensor(s) outputs voltage waveform.
- A digital quantity is obtained from each sensor by digitizing its response.



Single image sensor

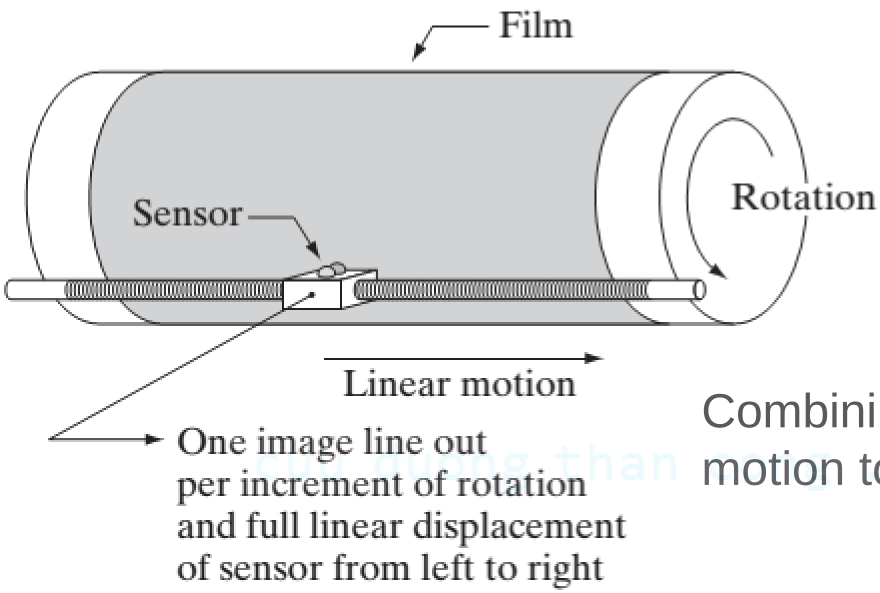


Array sensor



Line sensor

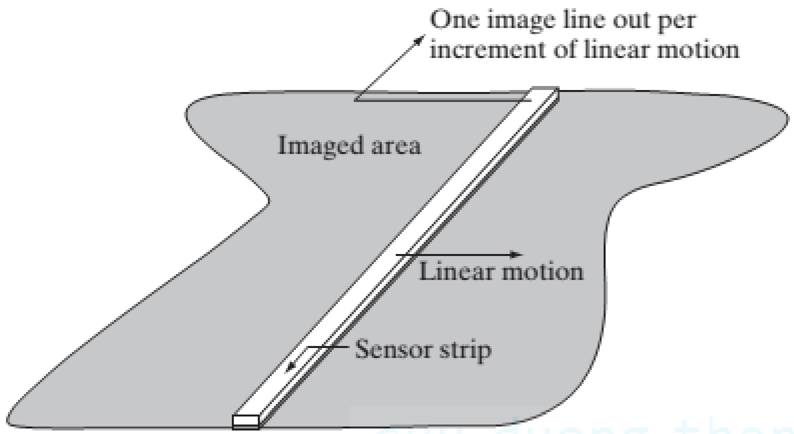
Image acquisition using a single sensor



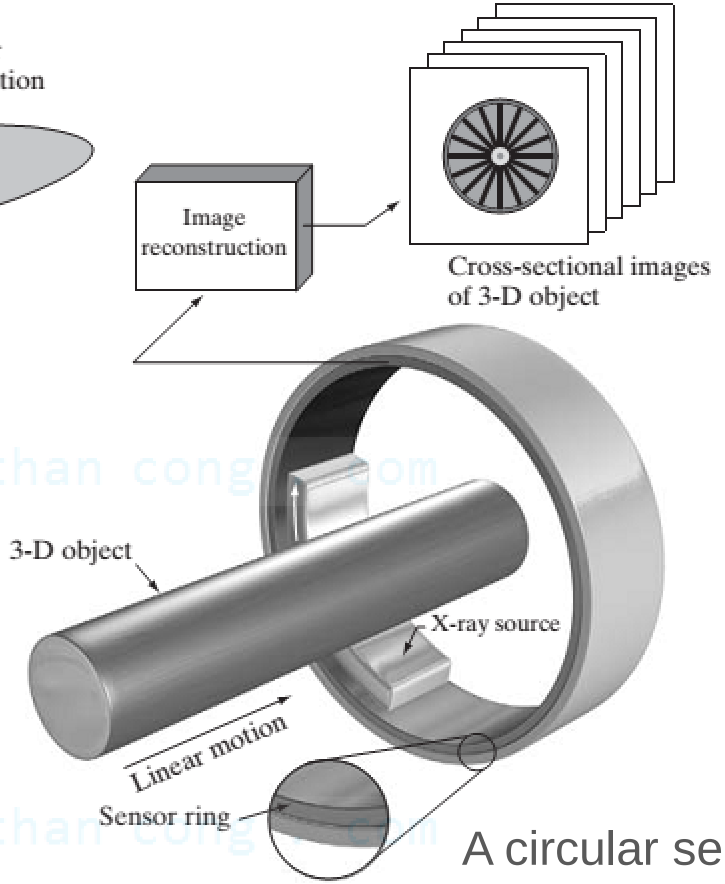
Combining a single sensor with motion to generate a 2-D image

- A negative film is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction
- Inexpensive but slow, high-resolution images, since mechanical motion can be controlled with high precision

Image acquisition using sensor strips



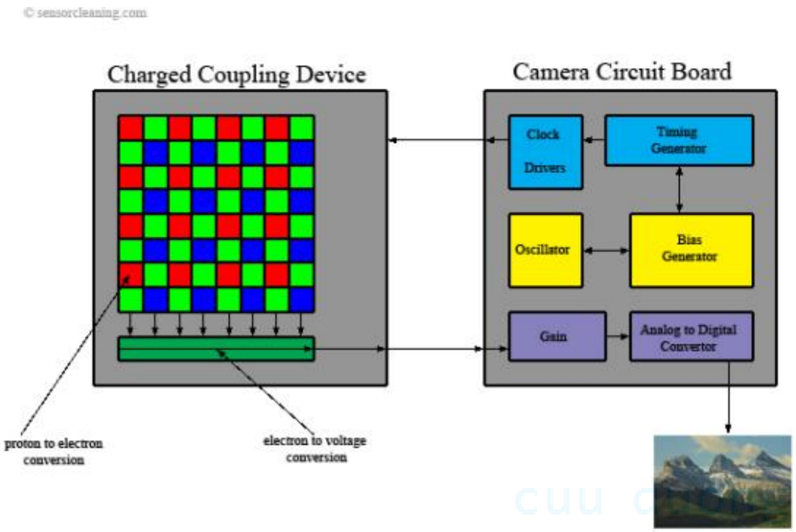
A linear sensor strip



A circular sensor strip

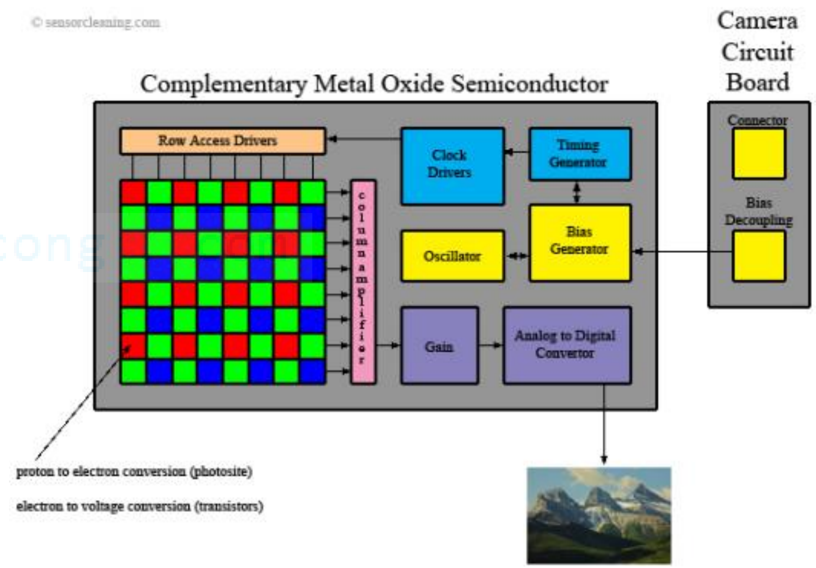
- The strip provides imaging elements in one direction.
- Motion perpendicular to the strip provides imaging in the other direction.

Image acquisition using sensor arrays



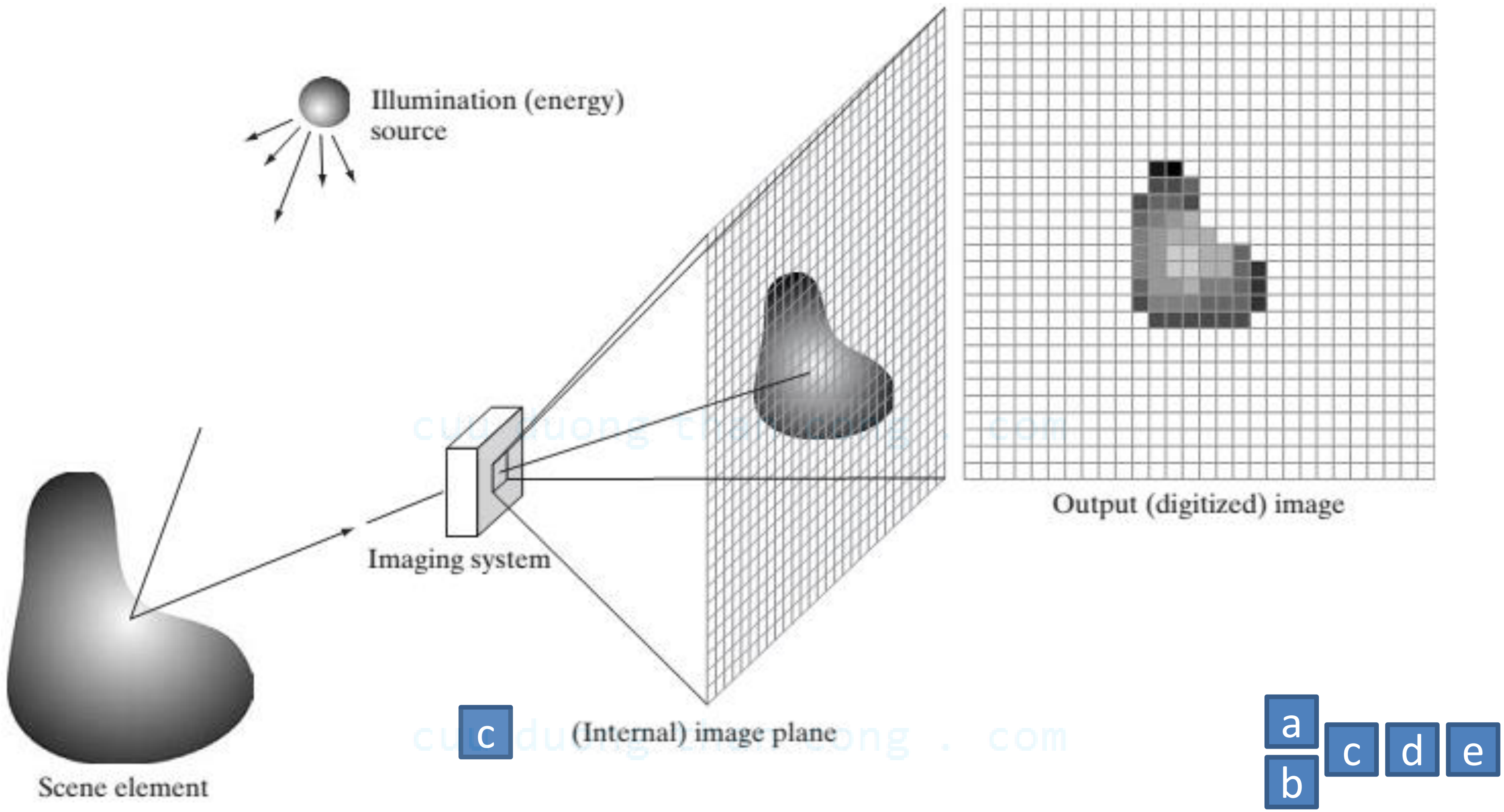
Charge-coupled Device (CCD)

Complementary Metal Oxide Semiconductor (CMOS)



- Key advantage: a complete image can be obtained by focusing the energy pattern onto the surface of the array while motion obviously is not necessary

Image sensing and generation



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.

A simple image formation model

- Images are denoted by two-dimensional functions $f(x, y)$ such that

$$\mathbf{0} < \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) < \infty$$

- The amplitude of f at spatial coordinates has its physical meaning determined by the source of the image
  - E.g., an image generated from a physical process has its intensity values proportional to energy radiated by a physical source

A simple image formation model

- $f(x, y)$ may be characterized by two components, including **illumination** and **reflectance**.

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

- $illumination\ i(x, y)$ ($0 < i(x, y) < \infty$): amount of source illumination incident on the scene being viewed
- $reflectance\ r(x, y)$($0 < r(x, y) < 1$): amount of illumination reflected by the objects in the scene

Values of illumination and reflectance

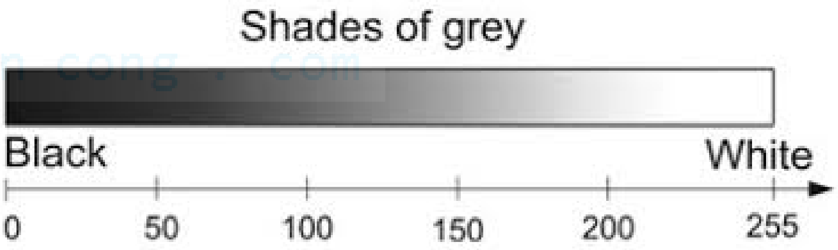
- Some typical values of illumination
  - 90000 lm/m^2 on a clear day
  - 10000 lm/m^2 on a cloudy day
  - 0.1 lm/m^2 on a clear evening with fullmoon
  - 1000 lm/m^2 in commercial office
- Some typical values of reflectance
  - 0.01 for black velvet
  - 0.65 for stainless steel
  - 0.80 for flat-white wall paint
  - 0.90 for silver-plated metal
  - 0.93 for snow

Gray-scale image formation model

- Let the intensity (gray level) of a monochrome image at any coordinates (x_0, y_0) be denoted by

$$\boldsymbol{l = f(x_0, y_0)}$$

- It is evident that l lies in the range $L_{min} \leq l \leq L_{max}$
  - $[L_{min}, L_{max}]$ (usually shifted to $[0, L-1]$): gray (or intensity) scale (from black to gray and finally to white)
  - In practice, $L_{min} = i_{min} r_{min}$ and $L_{max} = i_{max} r_{max}$



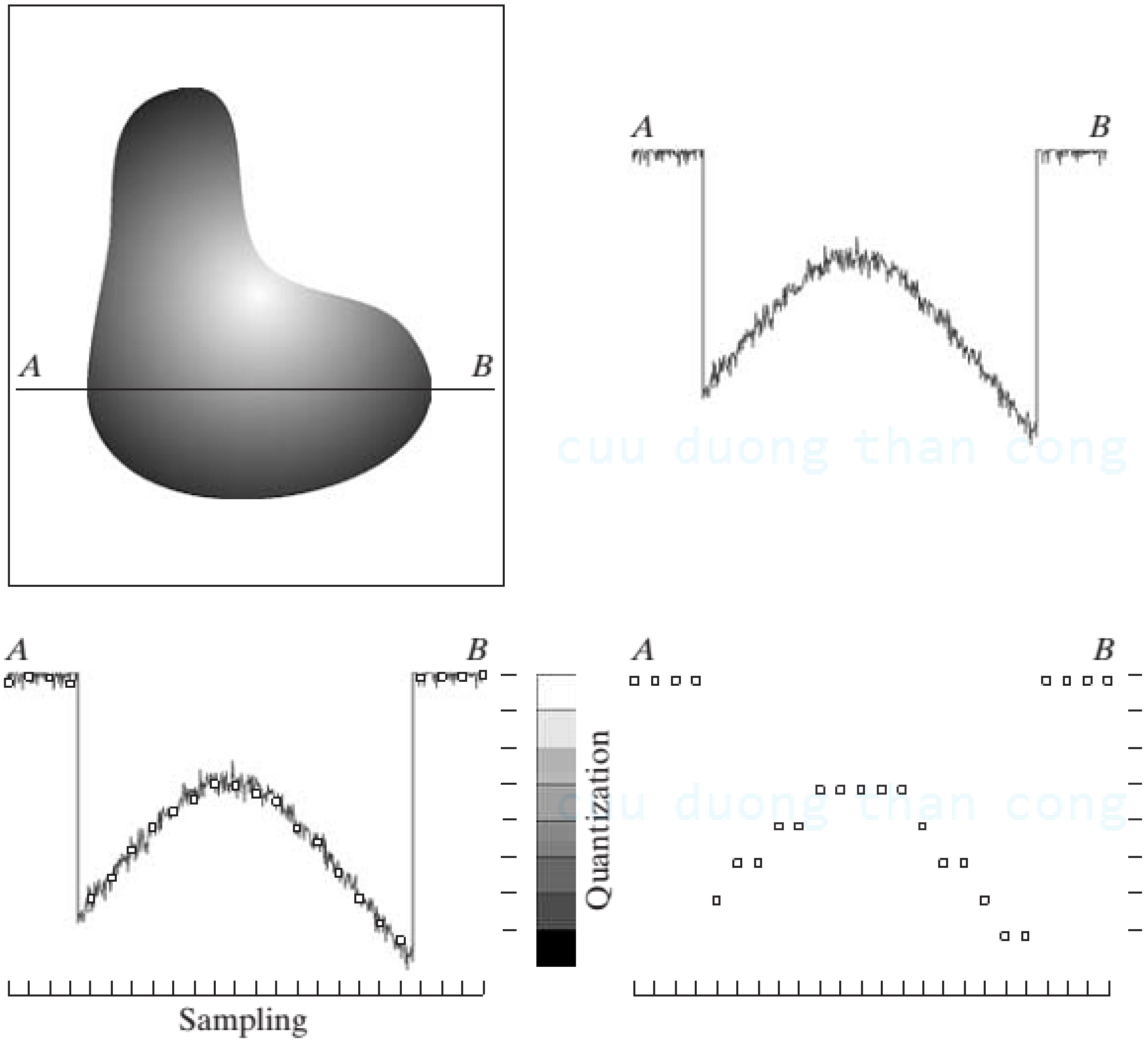
Section 2.4

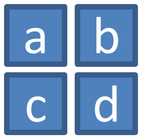
IMAGE SAMPLING AND QUANTIZATION

Generating a digital image

- An image may be continuous with respect to the x- and y-coordinates, and also in amplitude.
- To convert it to digital form, sample the function in both coordinates (*sampling*) and in amplitude (*quantization*)
- The quality of a digital image is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization
  - Image content is an important consideration in choosing parameters

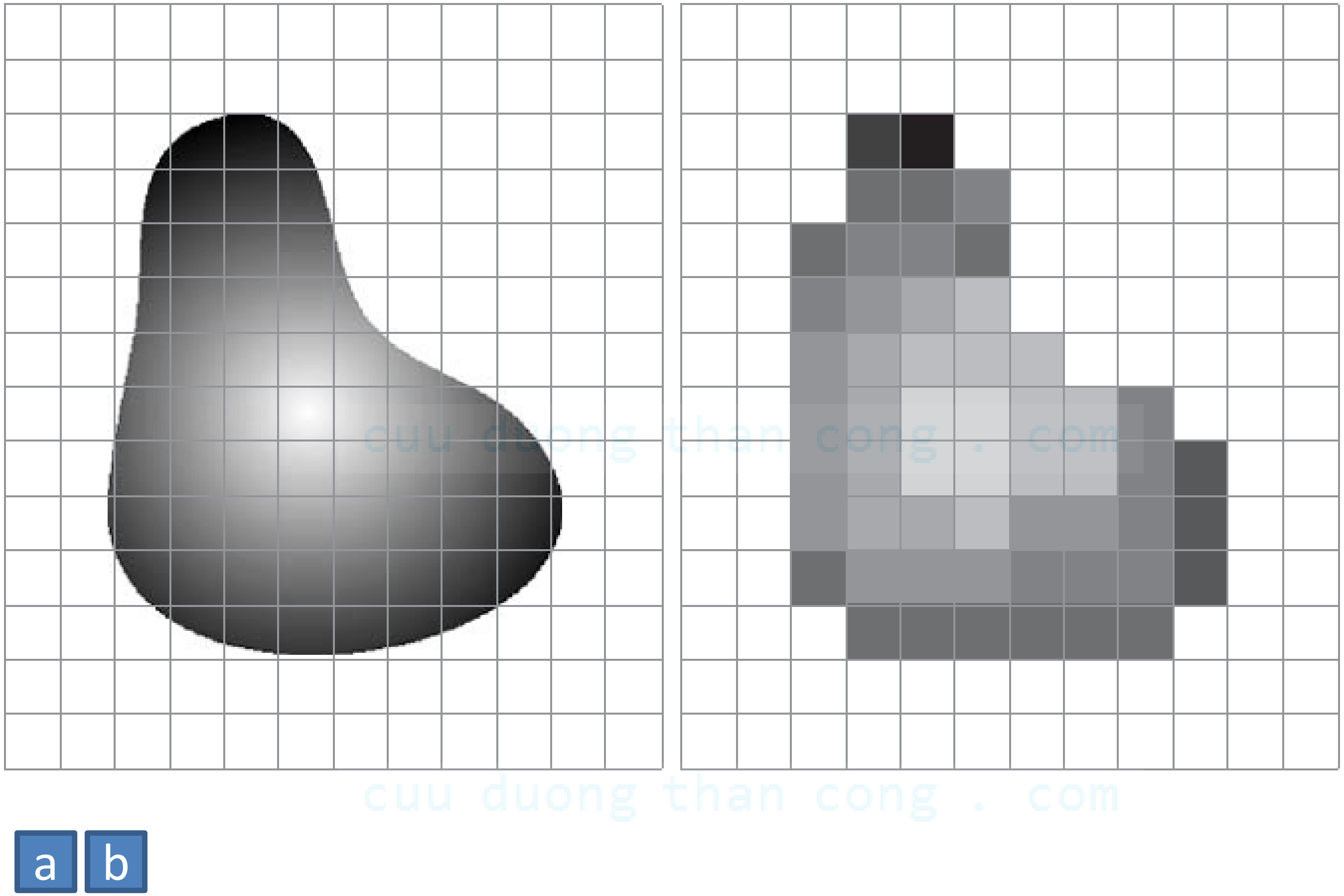
Generating a digital image: An example





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.

Generating a digital image: An example

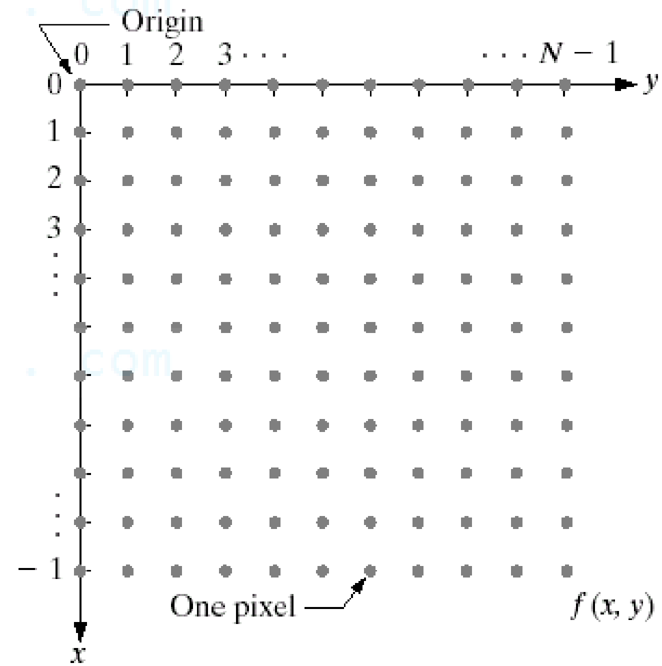


(a) Continuous image projected onto a sensor array.
(b) Result of image sampling and quantization

Representing a digital image

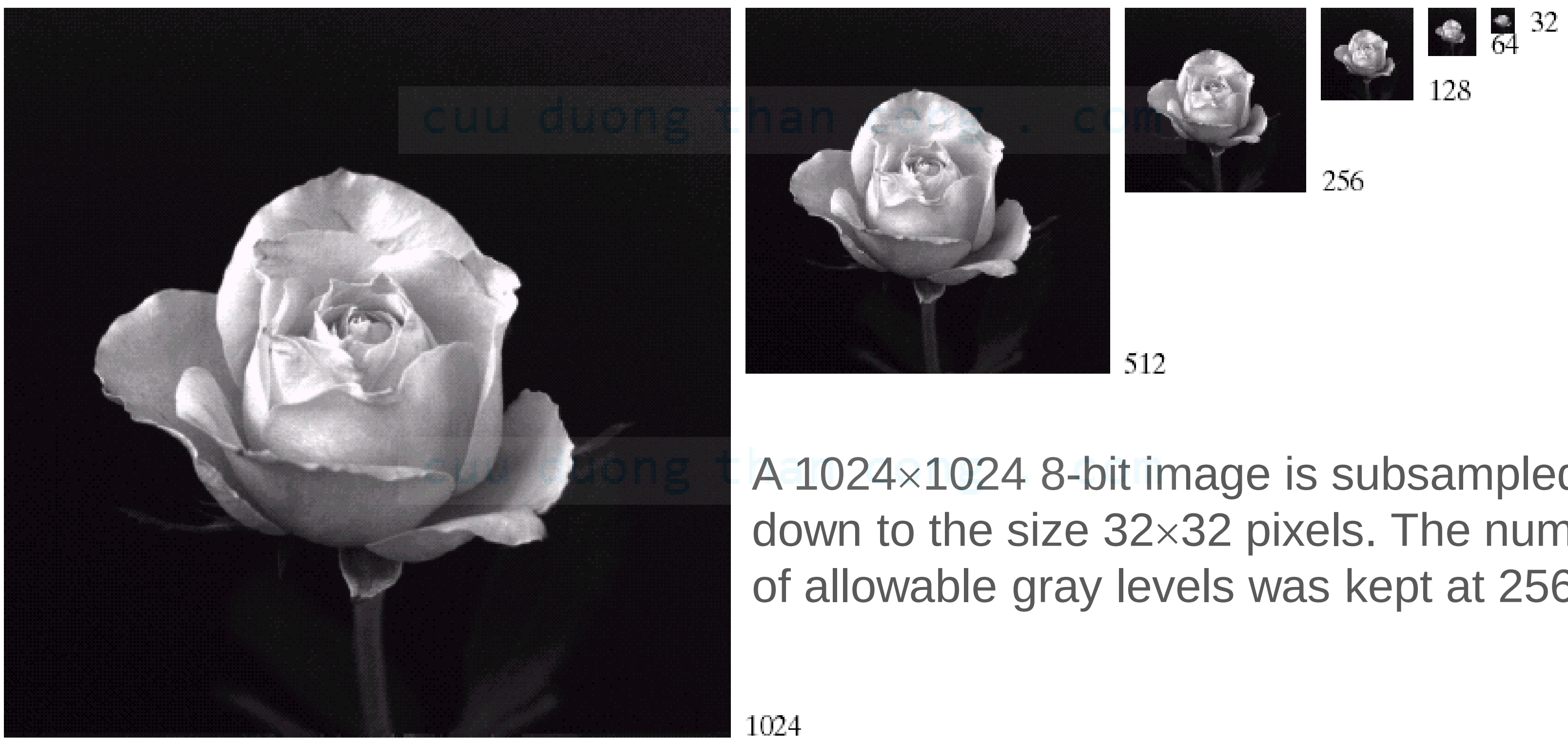
- Let $f(s,t)$ represent a continuous image function of two continuous variables s and t.
- The continuous image is sampled into a 2-D array $f(x,y)$ of M rows and N columns
  - Discrete coordinates (x,y), $x = 0,1,2,\ldots M-1$ and $y = 0,1,2,\ldots N-1$

- **Spatial domain:** The section of the real plane spanned by the coordinates of an image
  - x and y are spatial variables or spatial coordinates.



Effect of spatial resolution

- **Spatial resolution** is a measure of the smallest discernible detail in an image
  - Line pairs per unit distance, dots (pixels) per unit distance

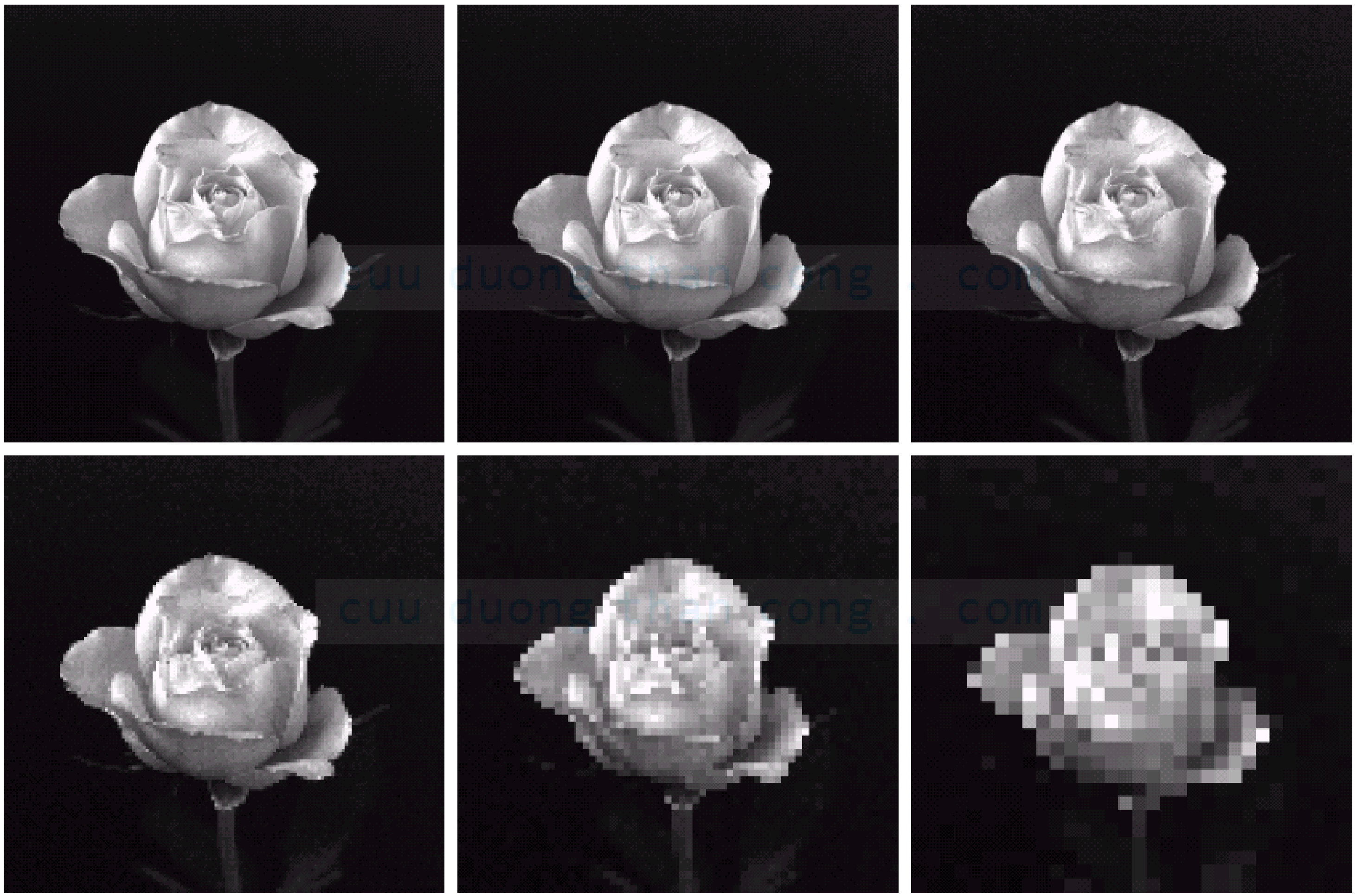


A 1024×1024 8-bit image is subsampled down to the size 32×32 pixels. The number of allowable gray levels was kept at 256

Effect of spatial resolution: An example

a	b	c
d	e	f

(a) 1024×1024 8-bit image. (b) 512 × 512 image resampled into 1024×1024 pixels by row and column duplication. (c) through (f) 256×256, 128×128, 64×64, 32×32 images resampled into 1024×1024 pixels



Effect of intensity resolution

- **Intensity resolution** refers to the smallest discernible change in intensity level
  - Common measure: the number of bits k used to quantize intensity
  - The number of intensity levels L usually is an integer power of two based on hardware considerations: $L = 2^k$

- An image whose intensity is quantized into 256 levels is said to have 8 bits of intensity resolution.

Effect of intensity resolution: An example

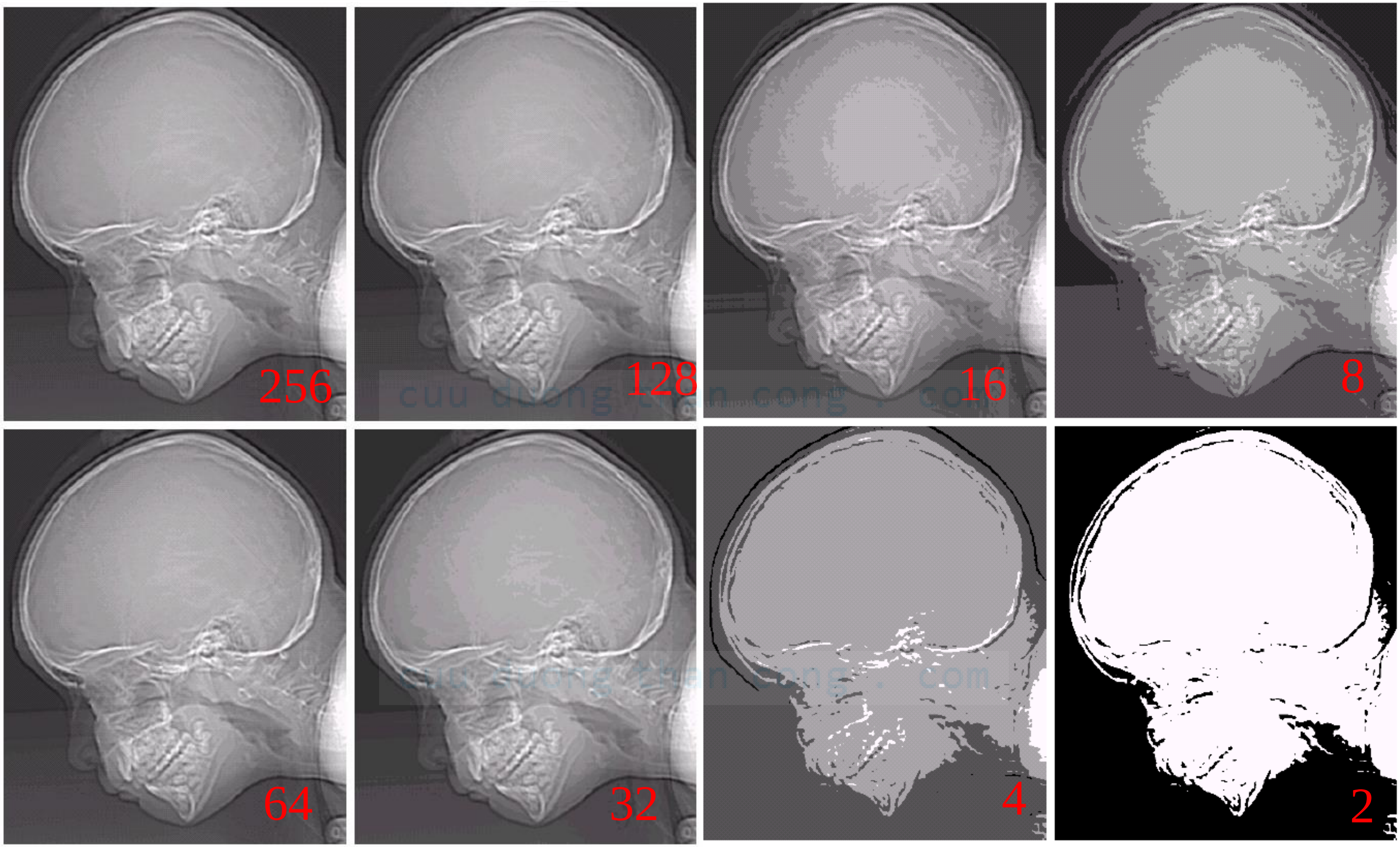


Image displayed in 256 down to 2 intensity levels, while keeping the image size constant

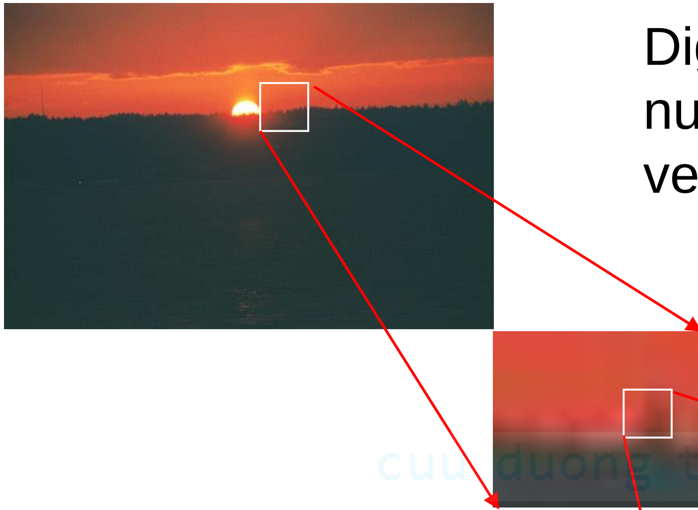
Storing (gray-scale) digital image

- The number of bits b required to store a digital image of size $M \times N$ is $b = M \times N \times k$
- Gray-scale images having 2^k gray levels are called **k-bit (gray-scale) images.**
- For example, a 8-bit image has 256 gray levels

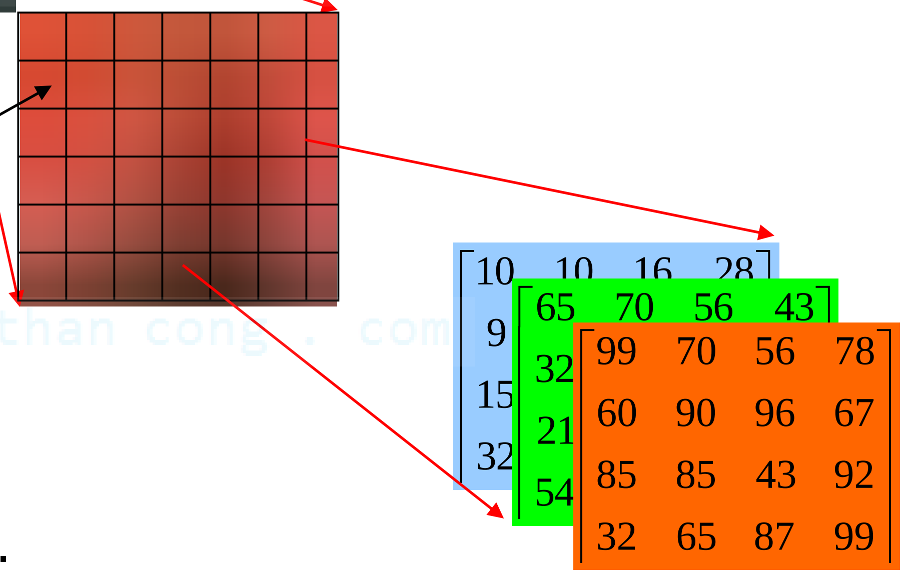
N/k	$1\ (L = 2)$	$2\ (L = 4)$	$3\ (L = 8)$	$4\ (L = 16)$	$5\ (L = 32)$	$6\ (L = 64)$	$7\ (L = 128)$	$8\ (L = 256)$
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072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

Number of storage bits for various values of N and k. L is the number of intensity levels

Digital image



Digital image = a multidimensional array of numbers (such as intensity image) or vectors (such as color image)



Each component in the image, called pixel, associates with the pixel value (a single number in the case of intensity images or a vector in the case of color images).

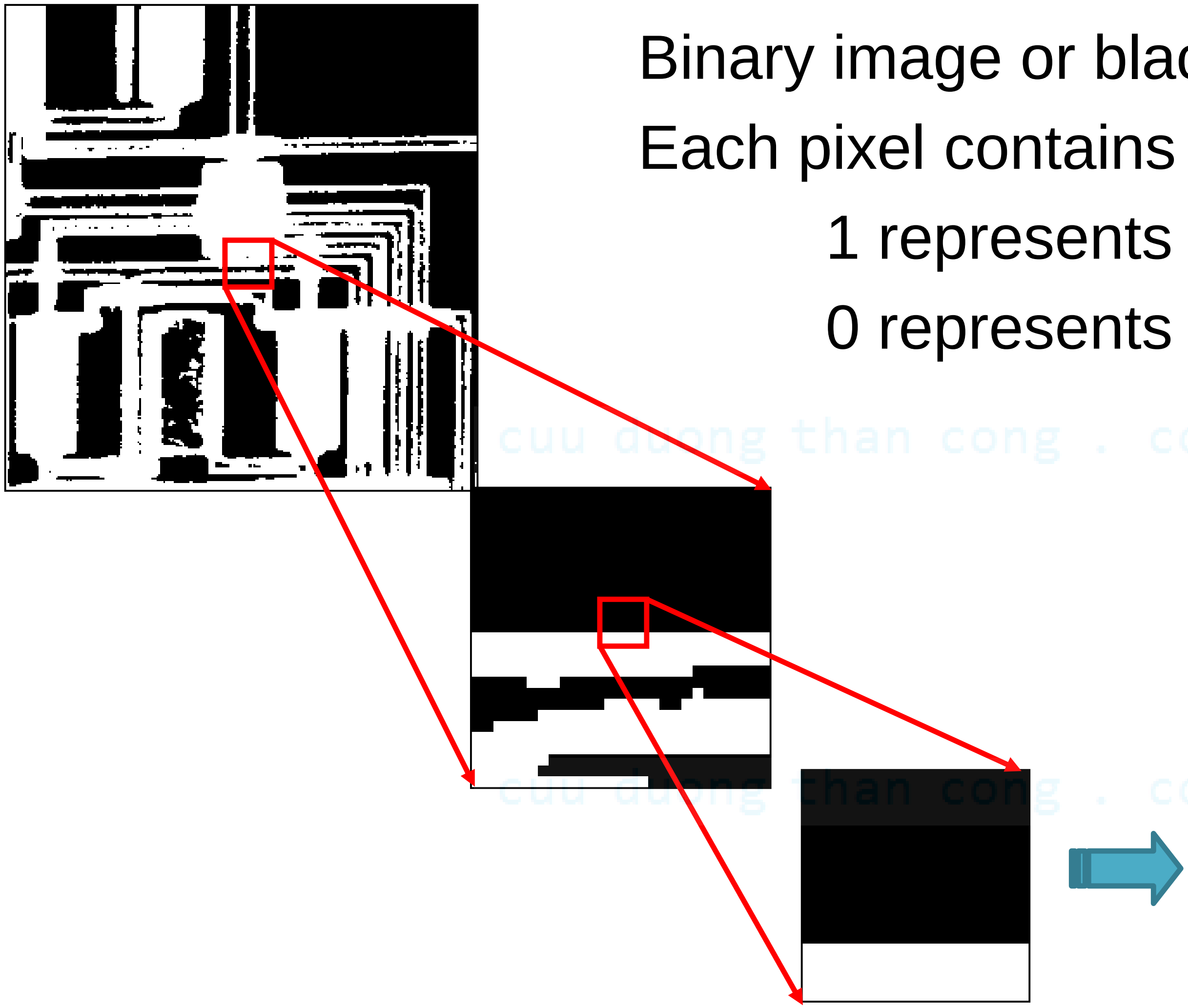
Digital image: Binary image

Binary image or black and white image

Each pixel contains one bit:

1 represents white

0 represents black



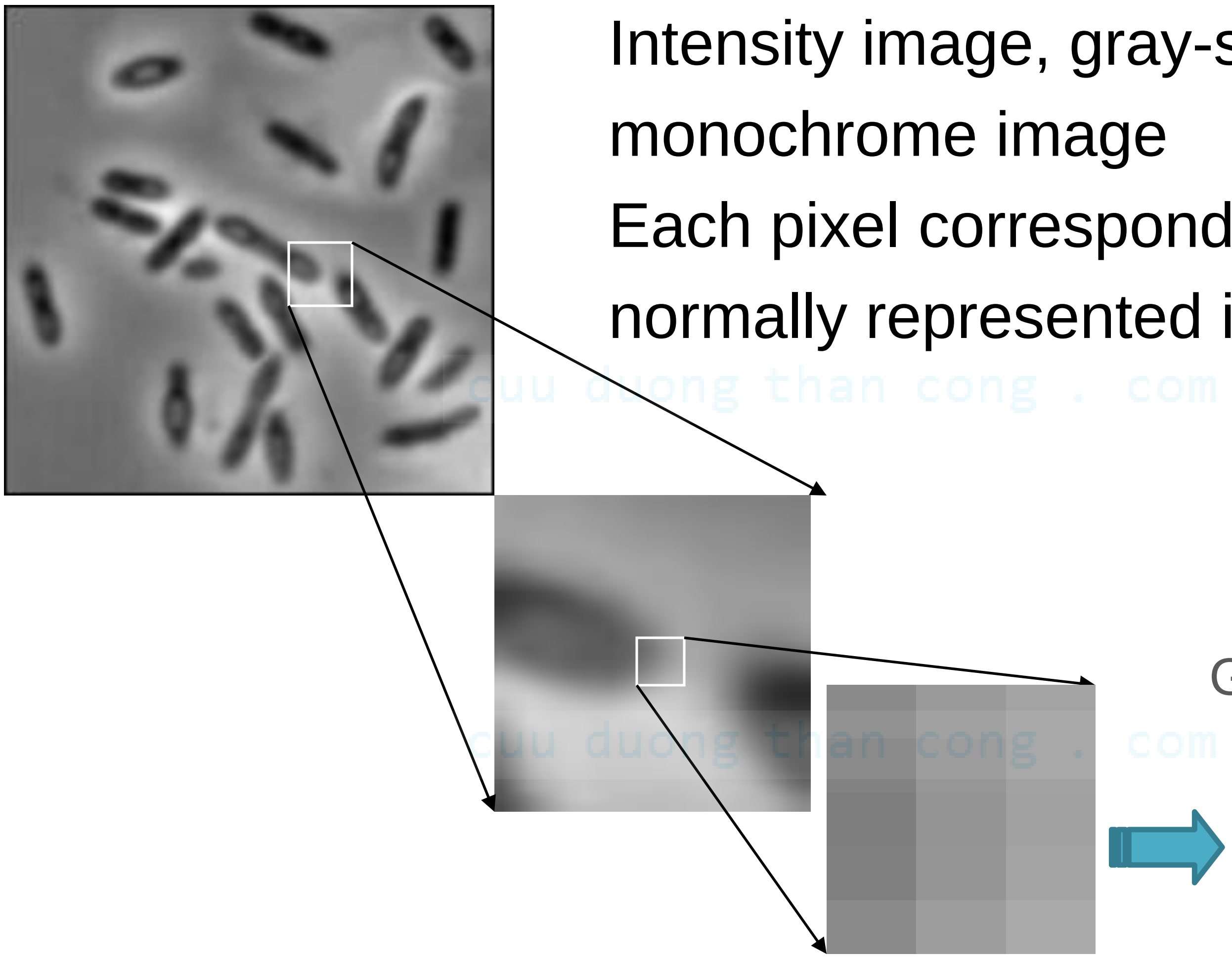
Binary data

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Digital image: Gray-scale image

Intensity image, gray-scale image or monochrome image

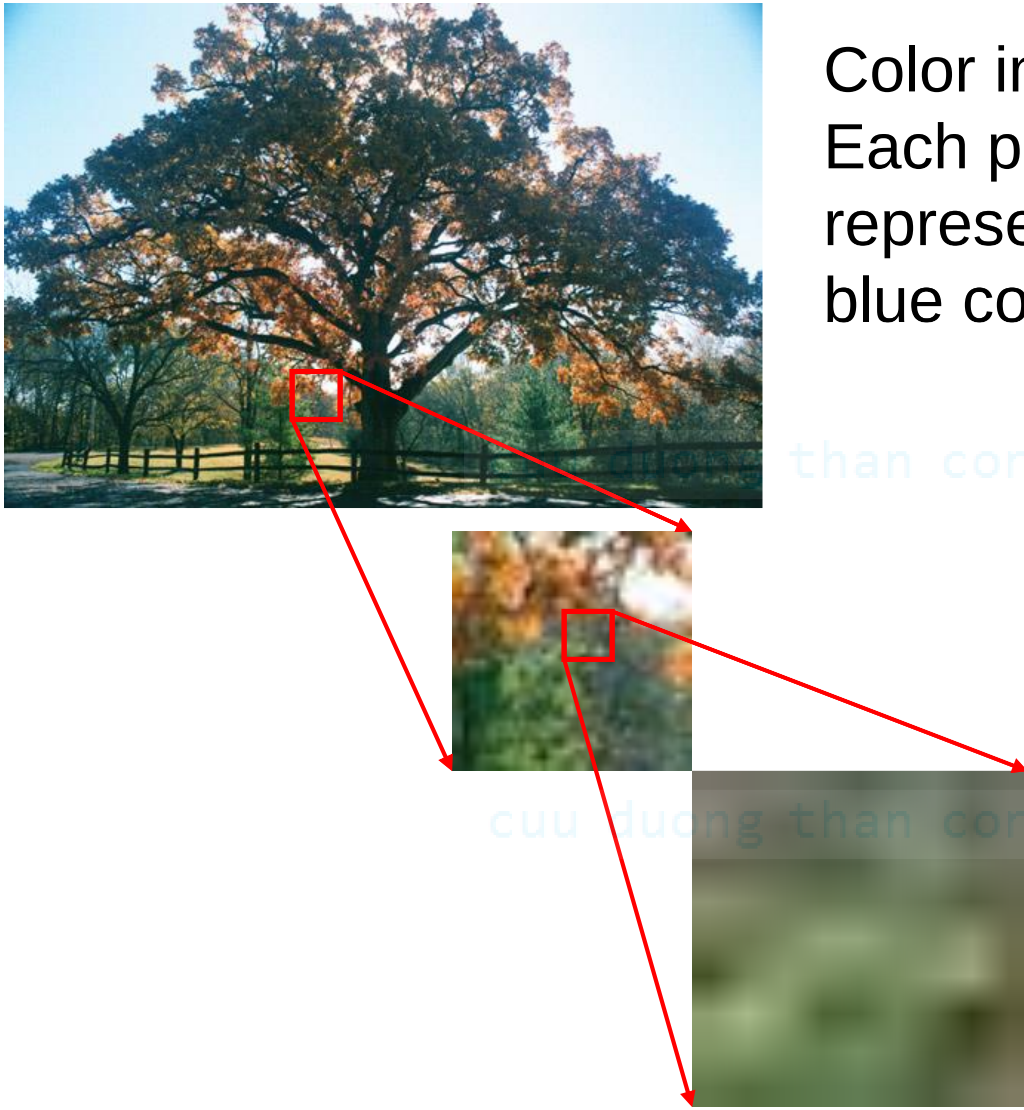
Each pixel corresponds to light intensity normally represented in gray scale



Gray scale values

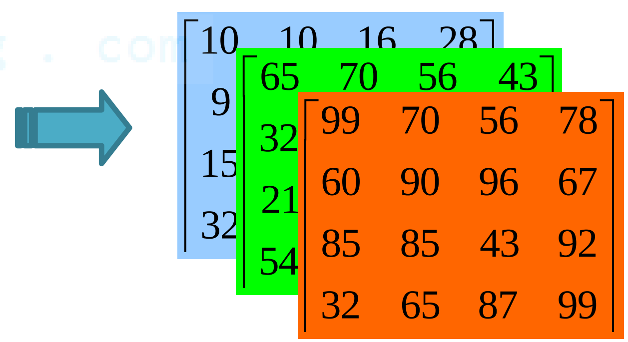
$$\begin{bmatrix} 10 & 10 & 16 & 28 \\ 9 & 6 & 26 & 37 \\ 15 & 25 & 13 & 22 \\ 32 & 15 & 87 & 39 \end{bmatrix}$$

Digital image: Color image



Color image or RGB image
Each pixel contains a vector representing red, green and blue components.

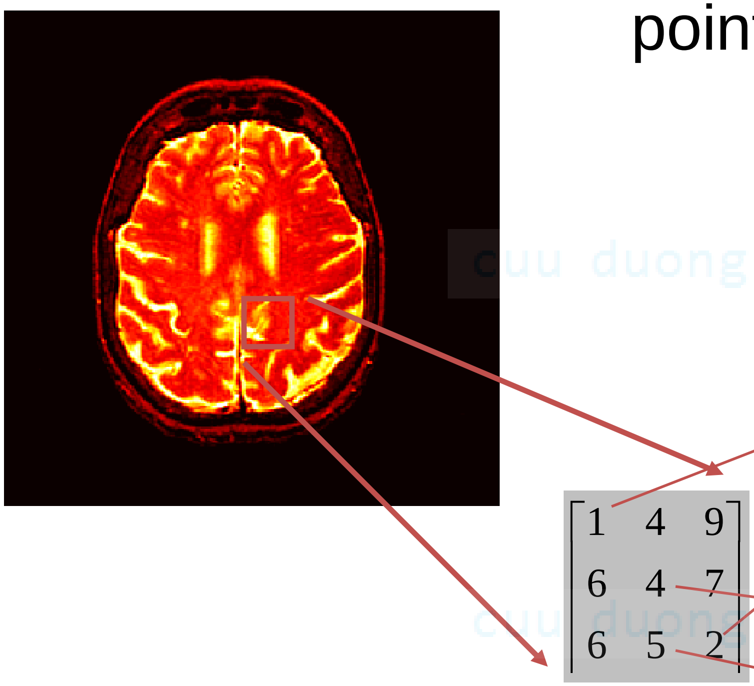
RGB components



Digital image: Index image

Index image

Each pixel contains index number pointing to a color in a color table



Index value

Color Table

Index No.	Red component	Green component	Blue component
1	0.1	0.5	0.3
2	1.0	0.0	0.0
3	0.0	1.0	0.0
4	0.5	0.5	0.5
5	0.2	0.8	0.9
…	…	…	…

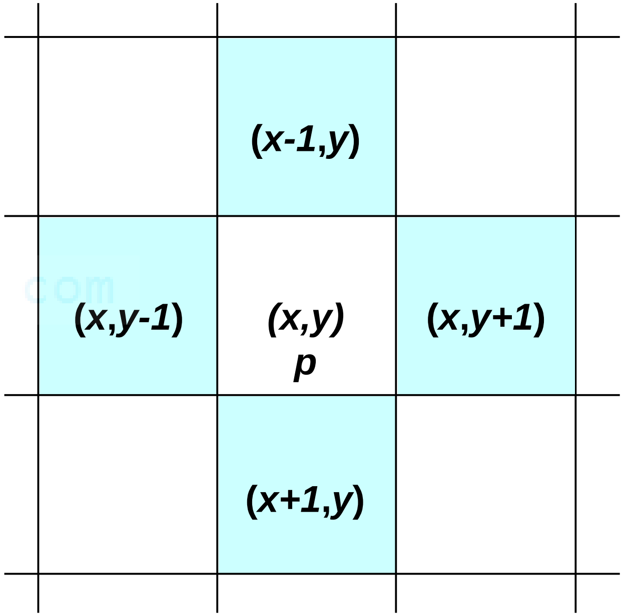
Section 2.5

SOME BASIC RELATIONSHIPS BETWEEN PIXELS

Neighbors of a pixel

- A pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by

$$\left\{ \begin{matrix} (x+1, y), (x-1, y), \\ (x, y+1), (x, y-1) \end{matrix} \right\}$$



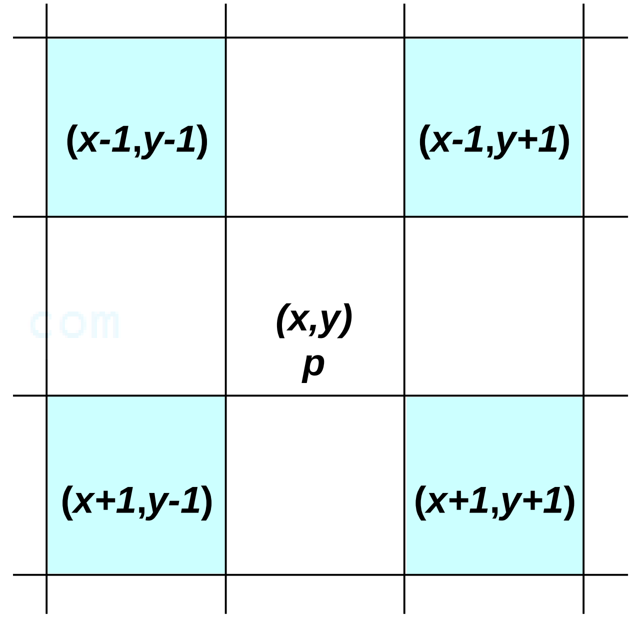
cuu duong than cong . com

- **4-neighbors of p**, denoted by $N_4(p)$.

Neighbors of a pixel

- A pixel p at coordinates (x, y) has four diagonal neighbors whose coordinates are given by

$$\begin{Bmatrix} (x+1, y+1), \\ (x+1, y-1), \\ (x-1, y+1), \\ (x-1, y-1) \end{Bmatrix}$$



- Denoted by $N_D(p)$

Neighbors of a pixel

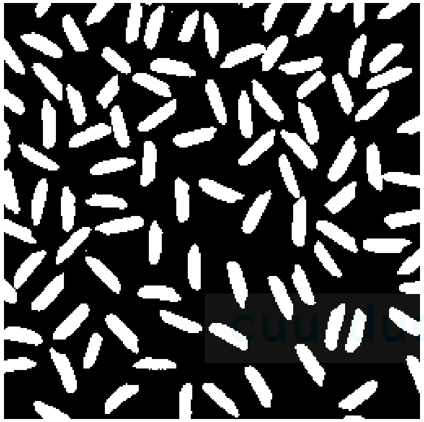
- Diagonal neighbors, together with the 4-neighbors, of a pixel p are called the **8-neighbors of p**, denoted by $\boldsymbol{N_8(p)}$.

$$\left\{\begin{matrix} (x-1, y-1), \\ (x, y-1), \\ (x+1, y-1) \\ (x-1, y) \\ (x+1, y) \\ (x-1, y+1), \\ (x, y+1), \\ (x+1, y+1) \end{matrix}\right\}$$

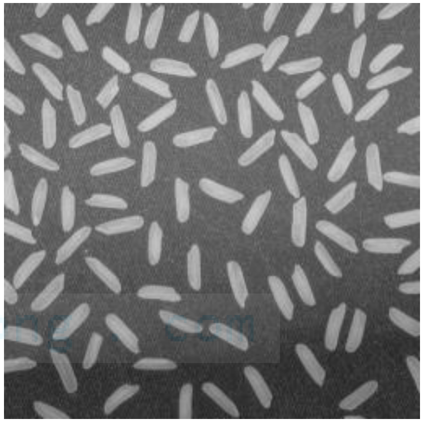
(x-1,y-1)	***(x-1,y)***	***(x-1,y+1)***
(x,y-1)	***(x,y)*** ***p***	***(x,y+1)***
(x+1,y-1)	***(x+1,y)***	***(x+1,y+1)***

Adjacency

- Let V be the set of intensity values used to define adjacency.
- For example, in binary image, $V = \{1\}$ if referring to adjacency of pixels with value 1; in a gray-scale image whose intensities are in [0, 255], $V \subseteq I_{256}$



Binary image, $V = \{1\}$



Gray-scale image, $V \subseteq I_{>100}$

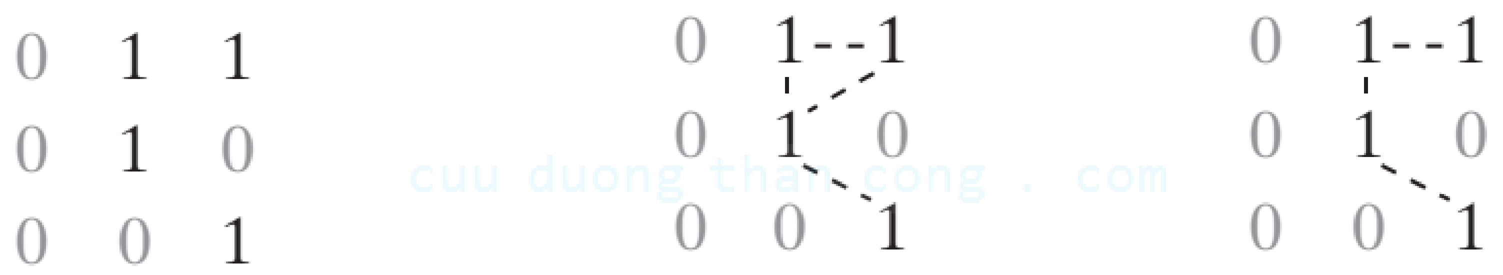
Adjacency

- **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
  - *i.* q is in $N_4(p)$ or
  - *ii.* q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V

Adjacency: An example

a b c

(a) An arrangement of pixels. (b) Pixels that are 8-adjacent (adjacency is shown by dashed lines; note the ambiguity). (c) m-adjacency. For $V = \{1\}$



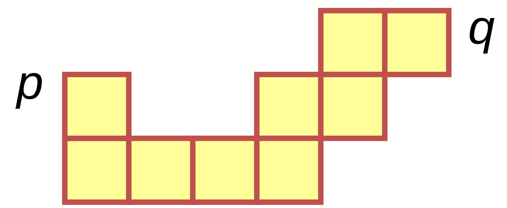
- m-adjacency is a modification of 8-adjacency to eliminate the ambiguities that often arise when 8-adjacency is used

Path

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

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

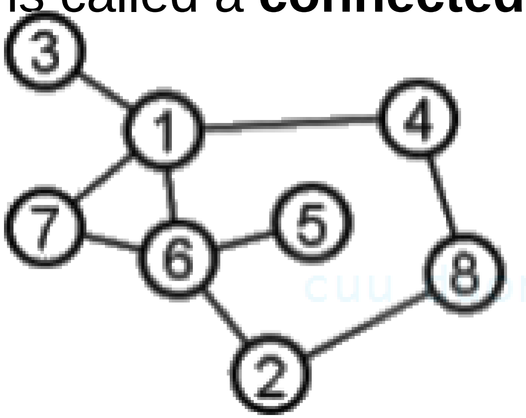
  - where $(x_0, y_0) = (x, y), (x_n, y_n) = (s, t)$ and pixels at (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$



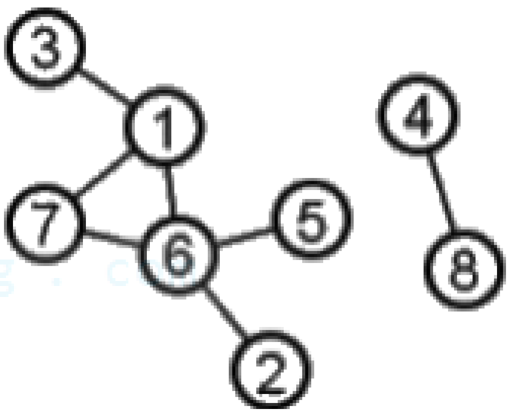
- A path could be **4-, 8-, or m-paths** depending on the type of adjacency specified.

Connected component

- Let S represent a subset of pixels in an image.
- Two pixels p and q are said to be **connected** in S if there exists a path between them consisting entirely of pixels in S
- For any pixel p in S, the set of pixels that are connected to it in S is called a **connected component of S**.
  - S is called a **connected set** if S only has one connected component.



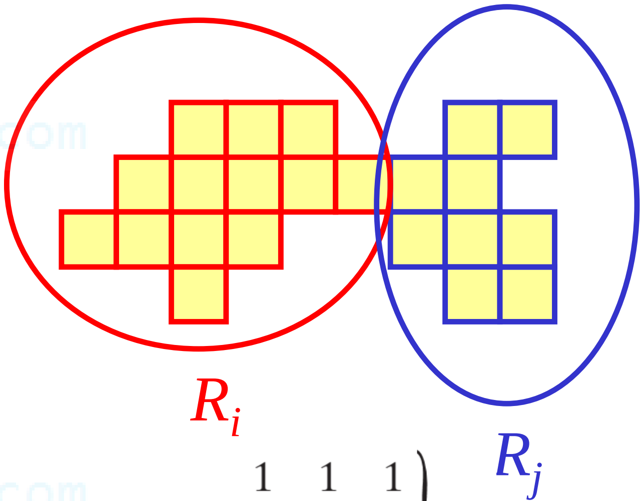
S is a connected set

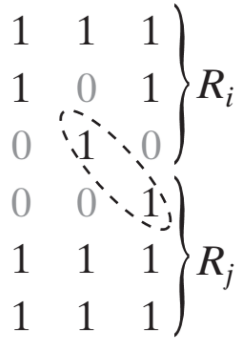


S has two connected components

Region

- Let R be a subset of pixels in an image.
- R is 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 adjacent are said to be *disjoint.*
- **4- and 8-adjacency** are considered when referring to regions





Distance measures

For pixels p, q and z with coordinates $(x, y), (s, t)$ and (v, w), respectively, D is a **distance function** (or *metric*) if

(a) $D(p, q) \geq 0$ $\quad$ $(D(p, q) = 0$ iff $p \equiv q)$,

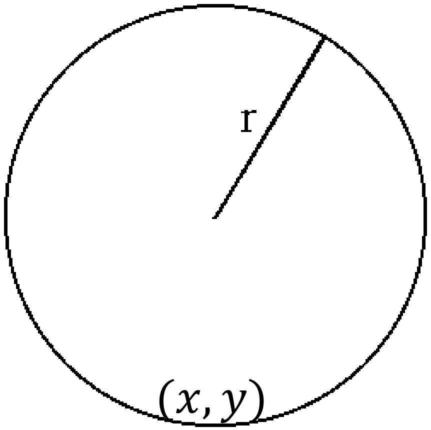
(b) $D(p, q) = D(q, p)$, and

(c) $D(p, z) \leq D(p, q) + D(q, z)$

Distance measures

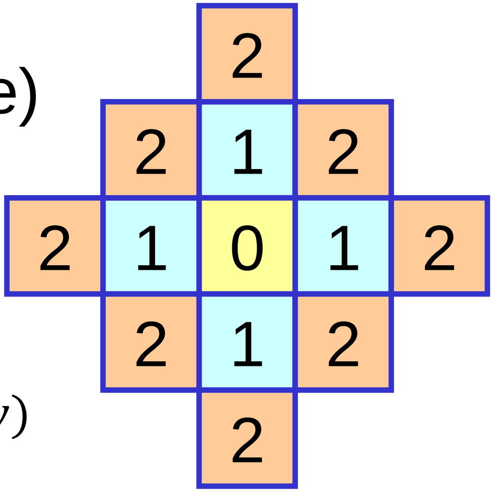
- **Euclidean distance**

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



- **D_4 distance** (Manhattan, city-block distance)

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



The pixels with $D_4 = 1$ are the 4-neighbors of (x, y)

Distance measures

- **D_8 distance** (chessboard 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

The pixels with $D_8 = 1$ are the 8-neighbors of (x, y)

Section 2.6

MATHEMATICAL TOOLS USED IN DIGITAL IMAGE PROCESSING

Mathematical tools used in Digital Image Processing

- Read Section 2.6, Chapter 2 in the textbook Digital Image Processing

Reference

- Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, 3rd edition, 2008. Chapter 2
- http://gear.kku.ac.th/~nawapak/178353/Chapter02.ppt
- Images are obtained from the above materials and Google