Indian Institute of Information Technology, Surat

भारतीय सूचना प्रौद्योगिकी संस्थान,सूरत

(Institute of National Importance under Act of Parliament)

# EC 503:

# Image Processing and Computer Vision

By:

Dr. Hemant S. Goklani

IIIT, Surat

# UNIT - I

# Digital Image Fundamentals

# UNIT - I

- ***Digital Image Fundamentals***:

- Visual Perception — Human Eye, Brightness Adaptation And Discrimination, Electromagnetic Spectrum;
- Image Sensing And Acquisition — Single, Strip And Array Sensors,
- Image Formation Models; Image Sampling And Quantization — Basic Concepts, Representation Of Image, Special And Gray Level Resolution, Aliasing, Zooming And Shrinking;
- Relationships Between Pixels — Nearest Neighbor, Adjacency, Connectivity, Regions, And Boundaries; Distance Measures;
- Image Operations On A Pixel Basis; Linear And Nonlinear Operations.

# Human Visual System

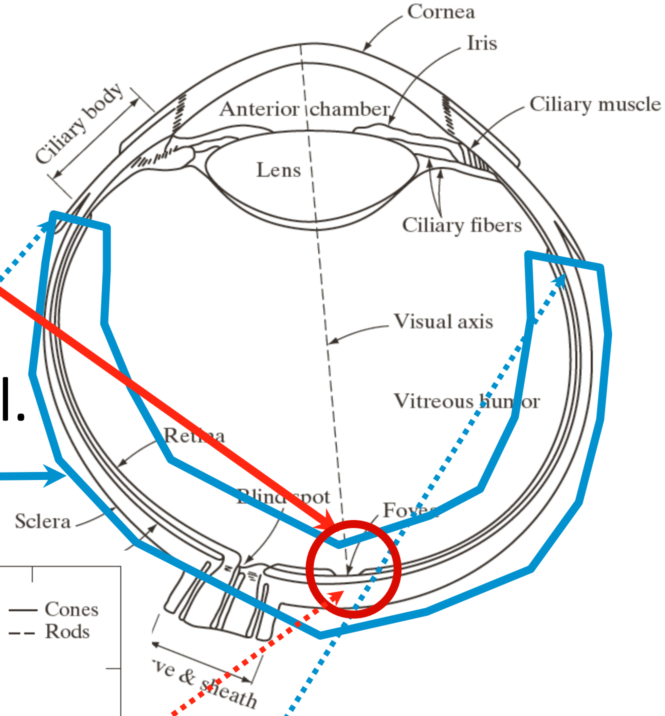
- The best visual system we have
- Knowledge of how images form in the eye can help us with processing digital images

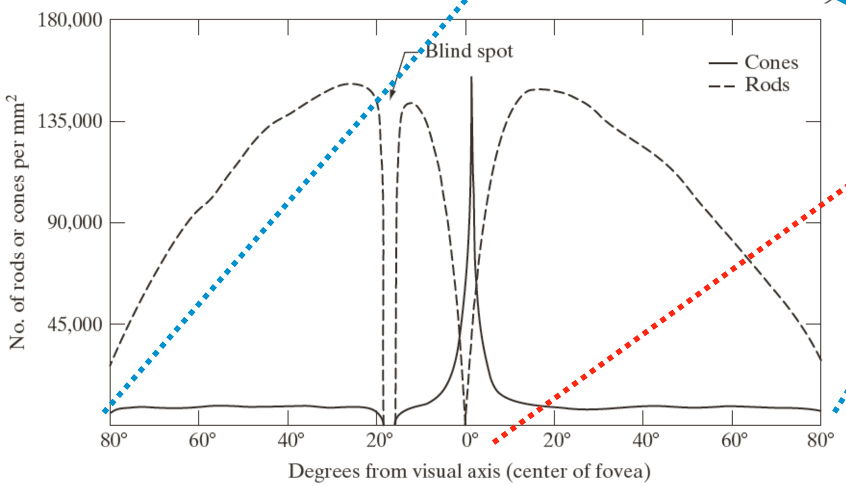
# Structure of the Human Eye

Retina: consist of receptors

- cones: highly sensitive to colors. *Photopic* or *bright-light* vision

-rods: give overall picture with reduced detail. *Scotopic* or *dim-light* vision



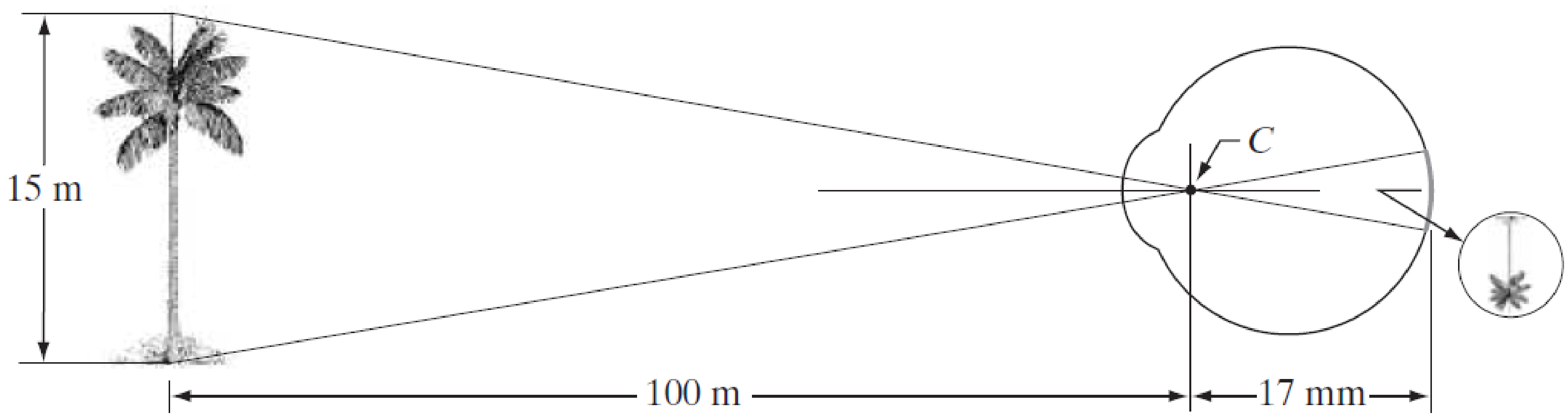
# Structure of the Human Eye

- The lens focuses light from objects onto the retina
- The retina is covered with light receptors called cones (6-7 million) and rods (75-150 million)
- Cones are concentrated around the fovea and are very sensitive to color
- Rods are more spread out and are sensitive to low levels of illumination

# Image Formation in the Eye

- Muscles within the eye can be used to change the shape of the lens allowing us to objects that near or far away
- An image is focused onto the retina causing rods and cones to become excited which ultimately send signals to the brain

# Image Formation in the Eye




This distance varies between 14-17 mm depending on the lens' focusing

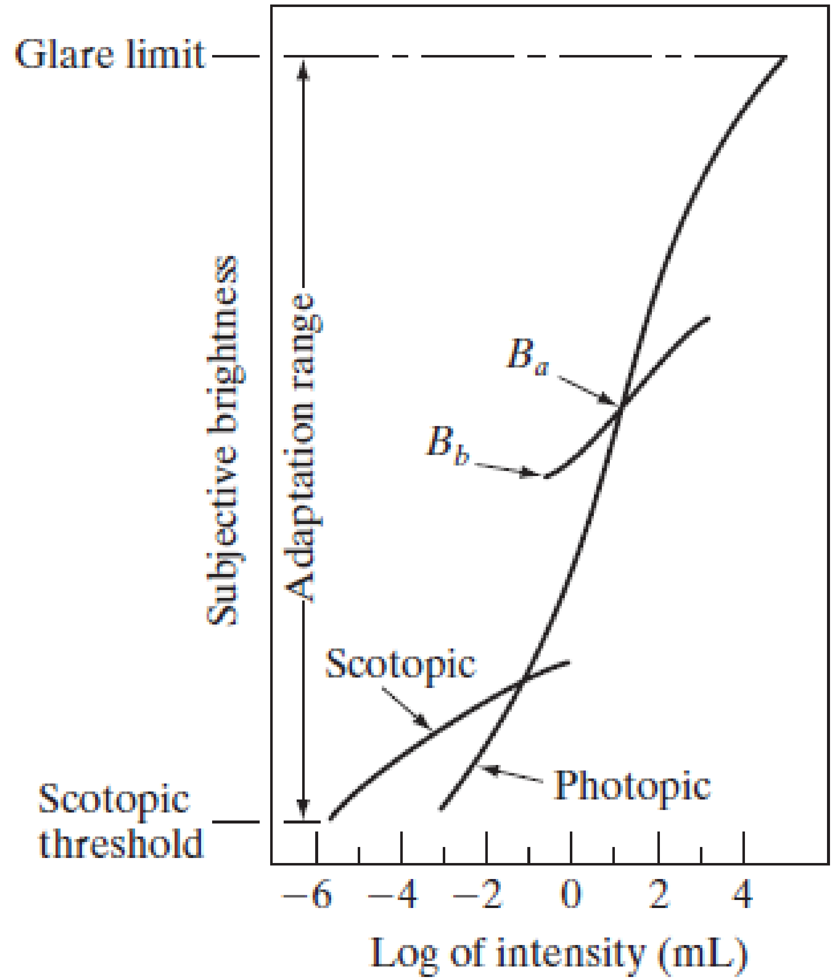
- Classical optical theory:
- A ray passes through the center C of the lens.
- The two triangle are proportional:
- $h$ is the height of the object on the retina (note that is located close to the fovea) $\frac{h}{17} = \frac{15}{100}$

# Brightness Adaptation and Discrimination

- The human visual system can perceive approximately $10^{10}$ different light intensity levels
- However, at any one time we can only discriminate between a much smaller number – Brightness Adaption
- Similarly, the perceived intensity of a region is related to the light intensities of the regions surrounding it

# Brightness Adaptation and Discrimination

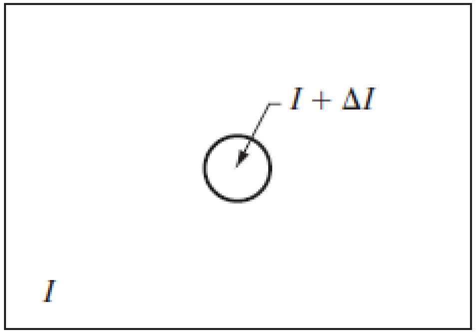
- The subjective brightness (intensity as perceived by the human visual system) is a logarithmic function of light intensity incident on the eye



# Weber Ratio

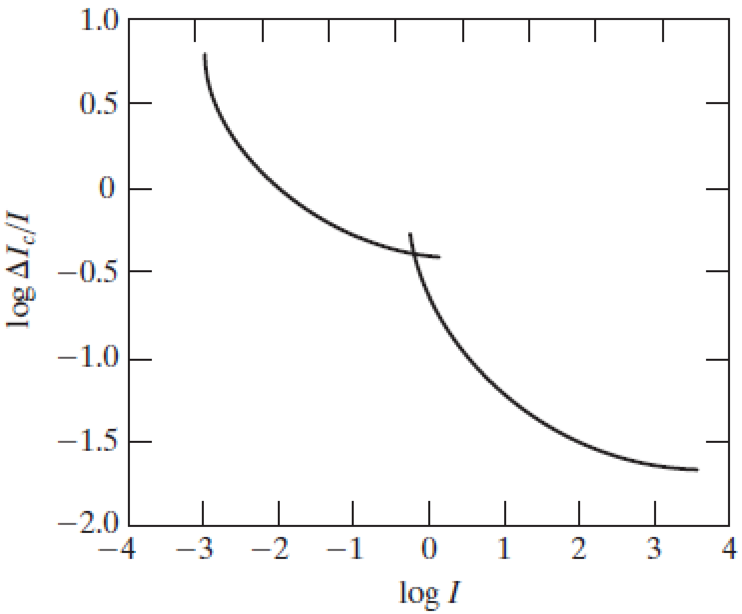
- Light source having intensity $I$ can be varied with the increment of illumination $\Delta I$
- If $\Delta I$ is not bright enough then no perceivable change
- As $\Delta I$ get stronger, there these is a perceived change
- Weber ratio = $\Delta I_c$ / $I$
- Where $\Delta I_c$ is the increment of illumination discriminable by 50% of the time with background $I$

- A small value weber ratio means that a small percentage change in intensity is discriminable - **good brightness discrimination**
- A large value of weber ratio means that a large percentage change in intensity is required. - **poor brightness discrimination**



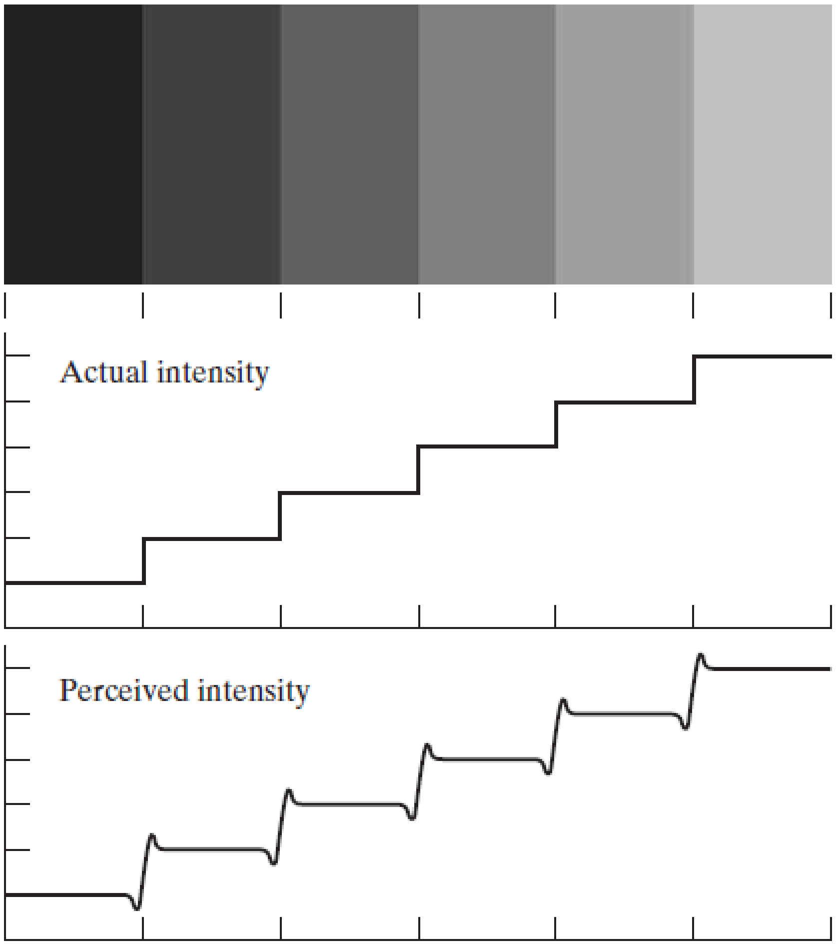
# Weber Ratio

- This curve shows that brightness discrimination is poor (the Weber ratio is large) at low levels of illumination
- And it improves significantly (the Weber ratio decreases) as background illumination increases
- The two branches in the curve reflect the fact that at low levels of illumination vision is carried out by the rods
- Whereas at high levels (showing better discrimination) vision is the function of cones



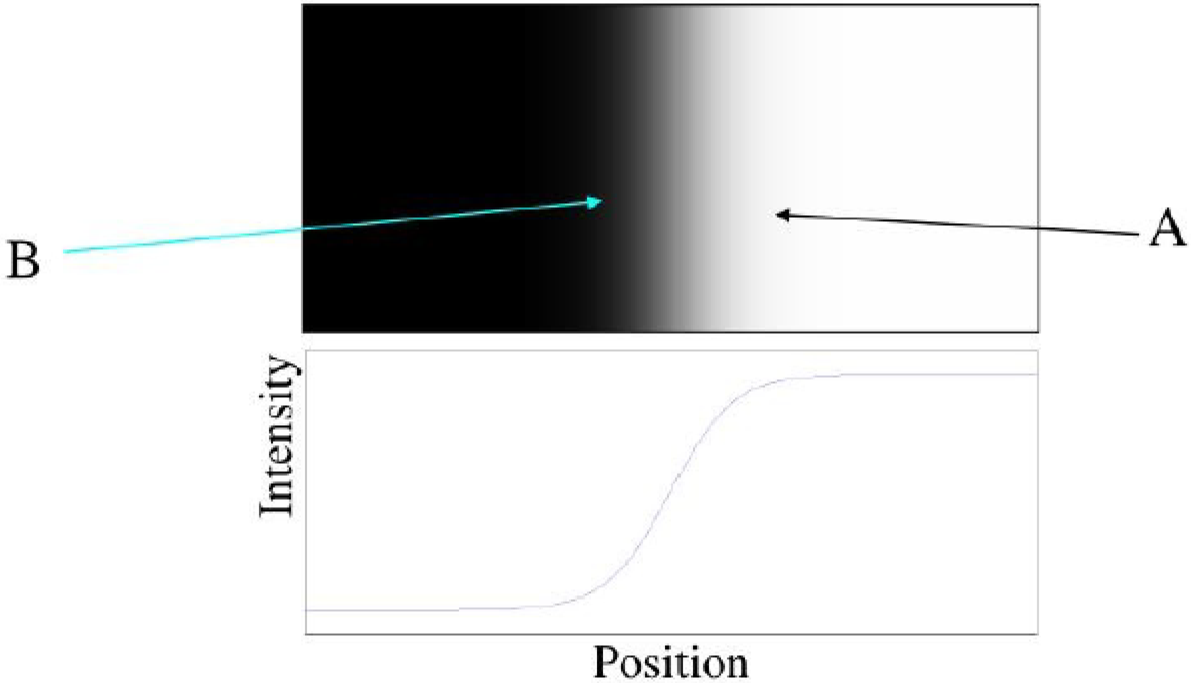
# Brightness Adaptation and Discrimination

- The visual system tends to undershoot or overshoot around the boundary of regions of different intensities
- Fig (a) shows a striking example of this phenomenon
- Although the intensity of the stripes is constant, we actually perceive a brightness pattern that is strongly scalloped near the boundaries [Fig. (c)]



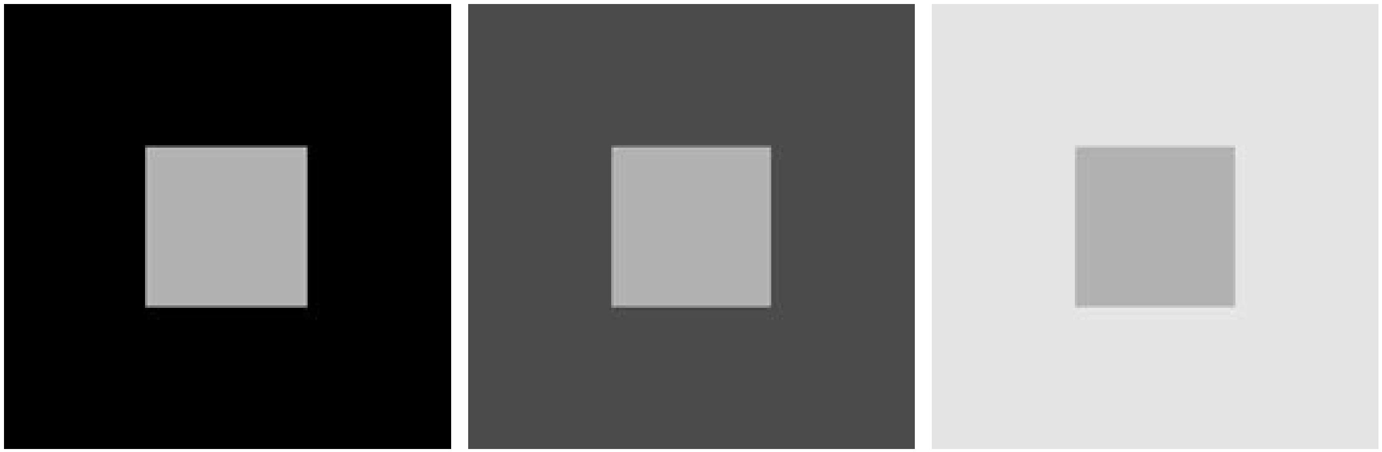
# Mach Band Effect

- In area A, brightness perceived is darker while area B is brighter
- This phenomenon is called Mach Band Effect
- These seemingly scalloped bands are called Mach bands after Ernst Mach, who first described the phenomenon in 1865



# SIMULTANEOUS CONTRAST

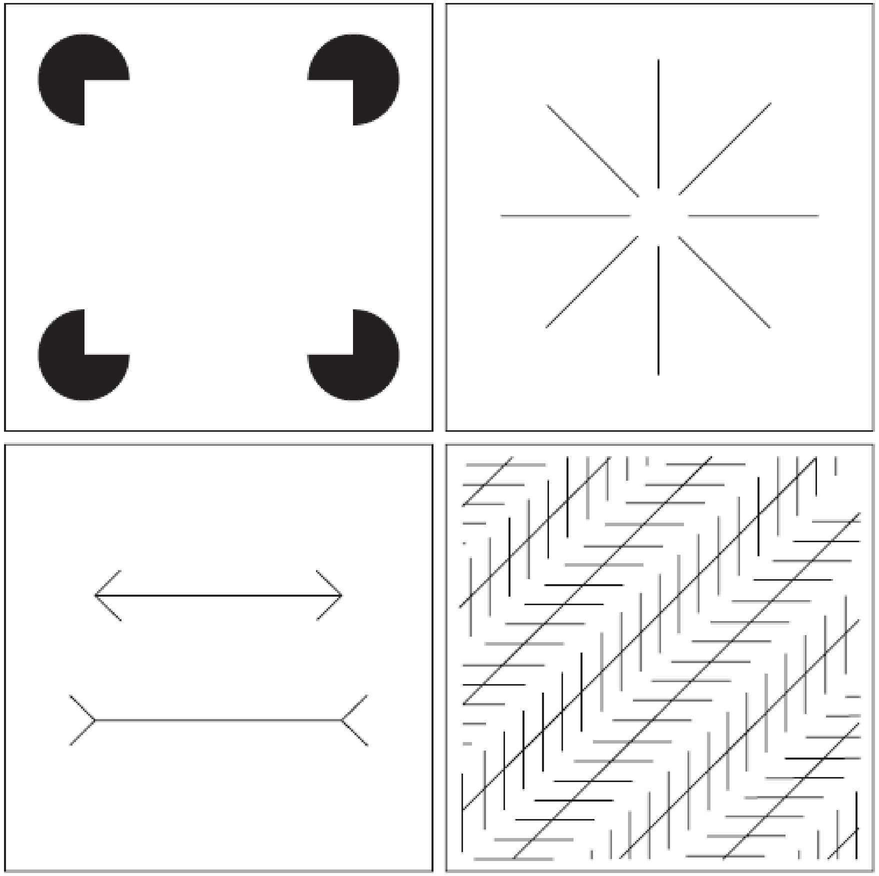
- A region's perceived brightness does not depend simply on its intensity



- All small squares have exactly the same intensity but they appear progressively darker as background becomes lighter

# OPTICAL ILLUSIONS

- The eye "fills in" non-existing information or wrongly perceives geometrical properties of objects

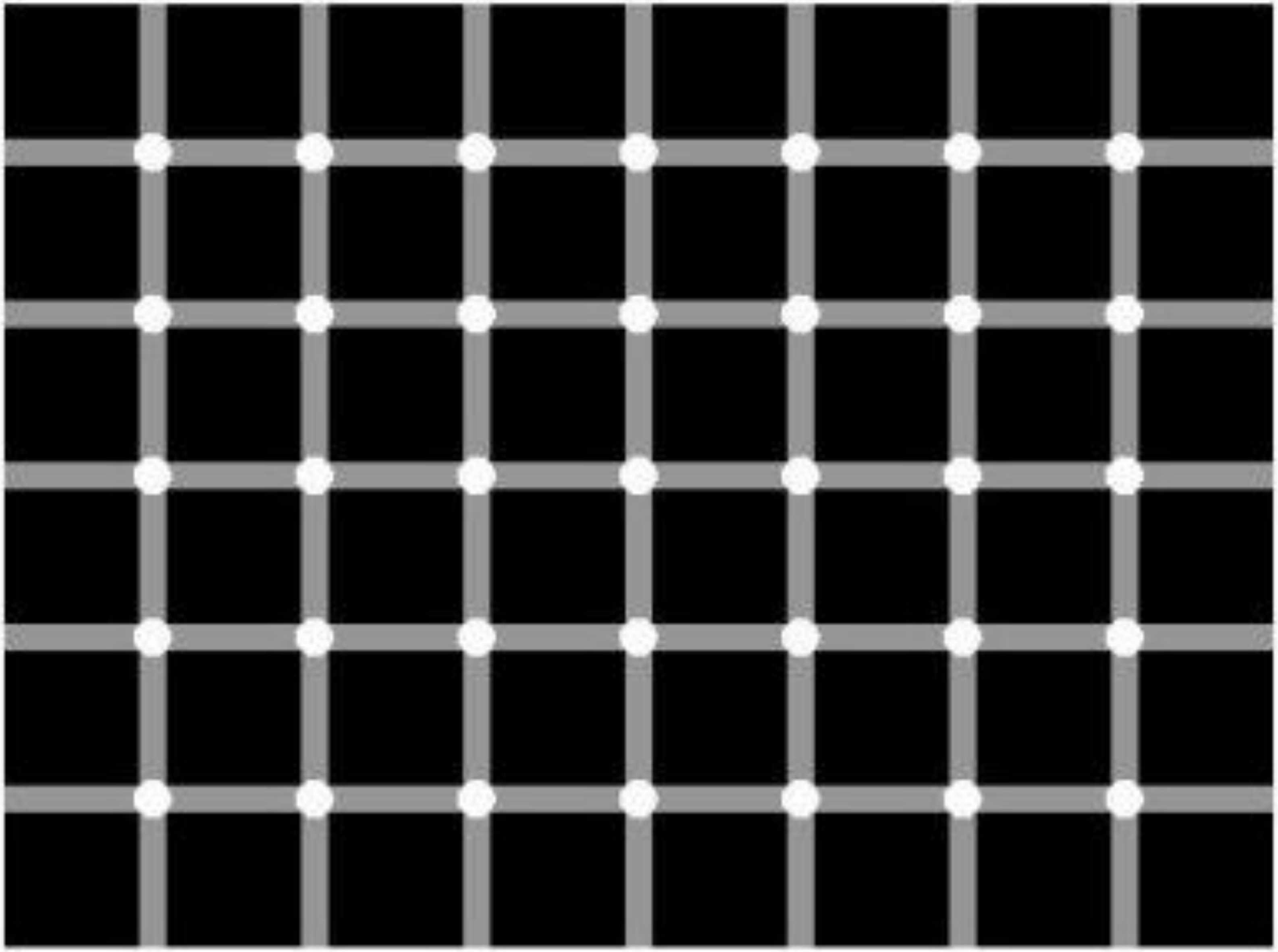


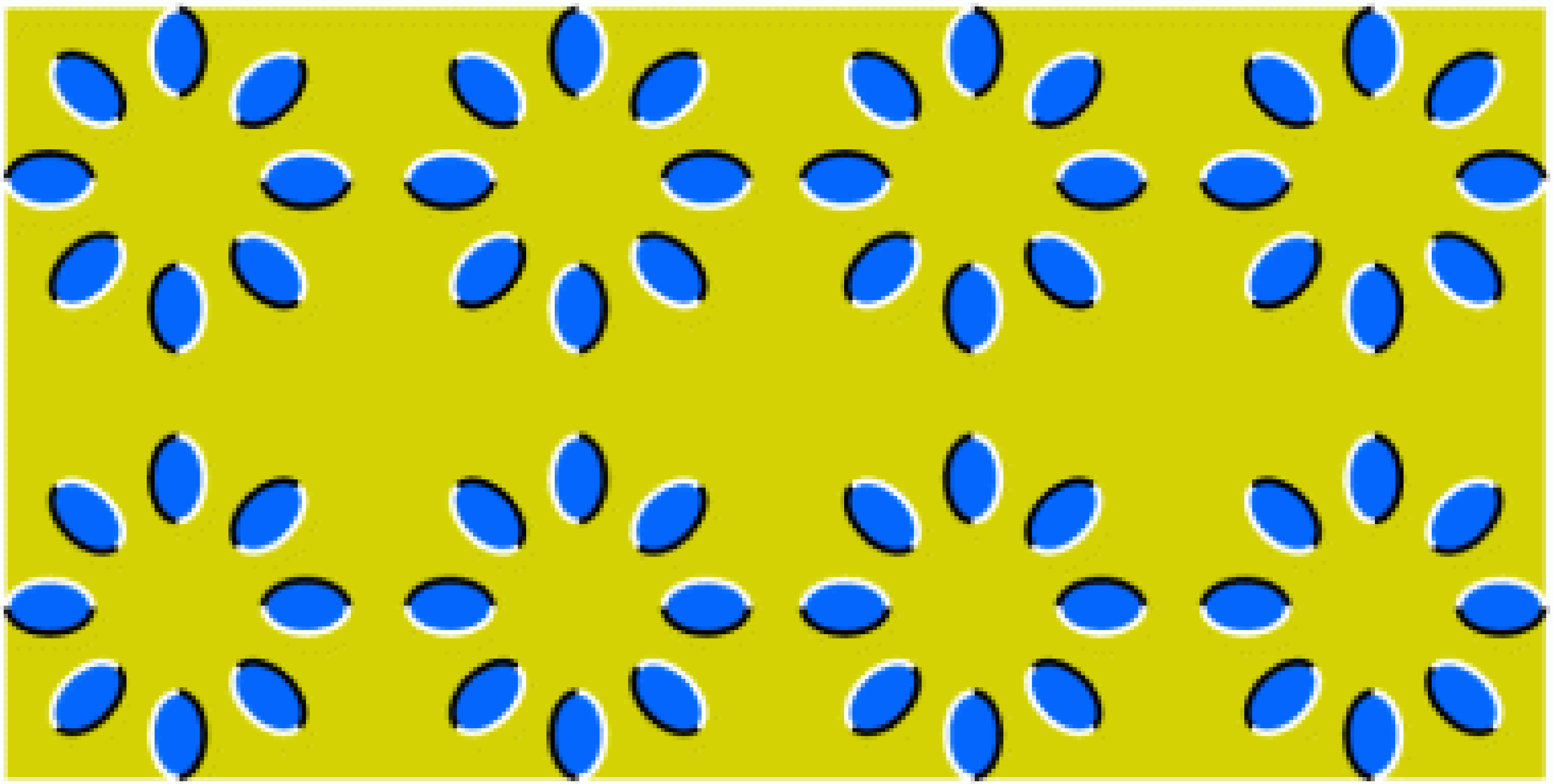
## There are two different images:

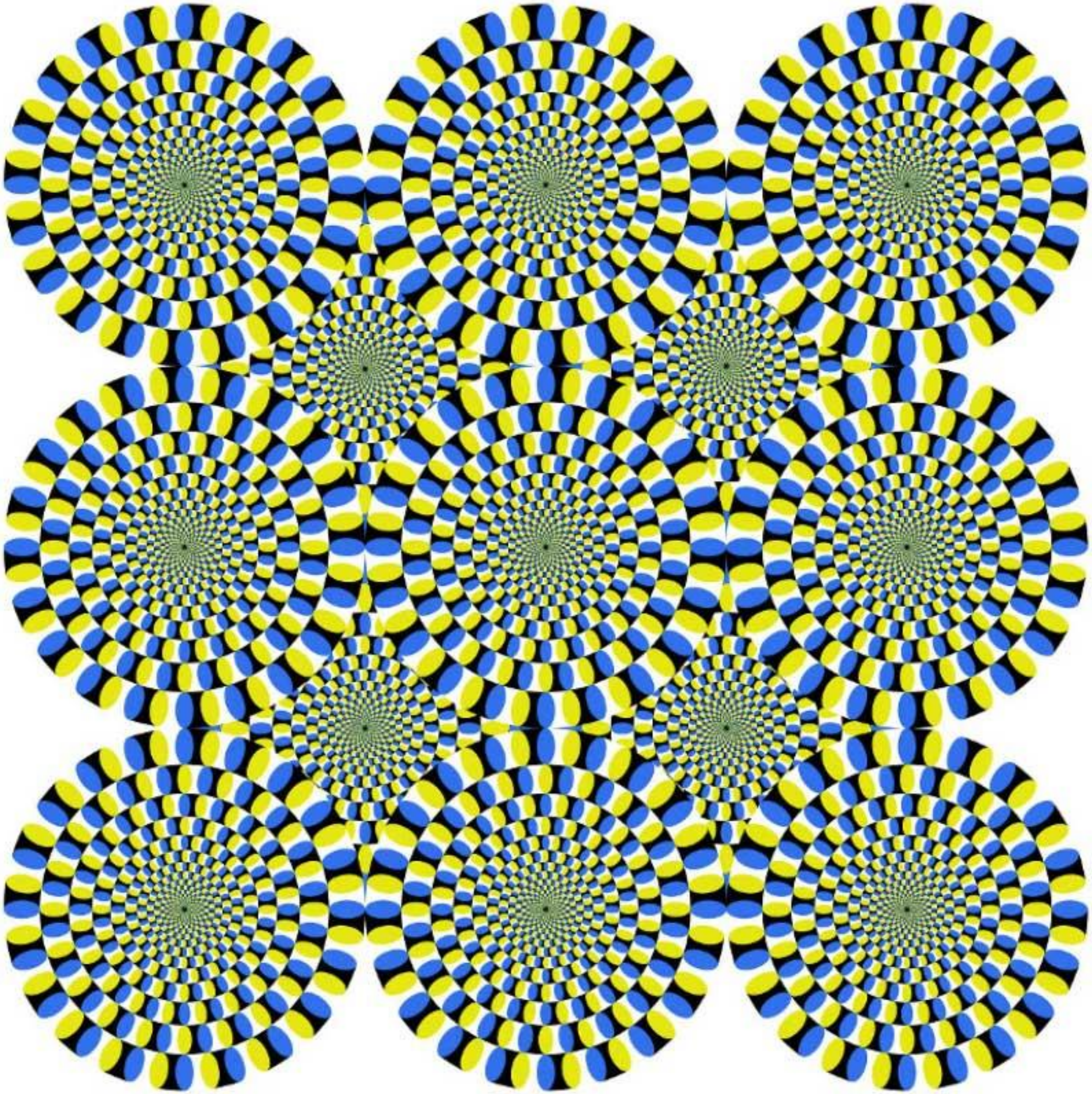
A frog & an horse

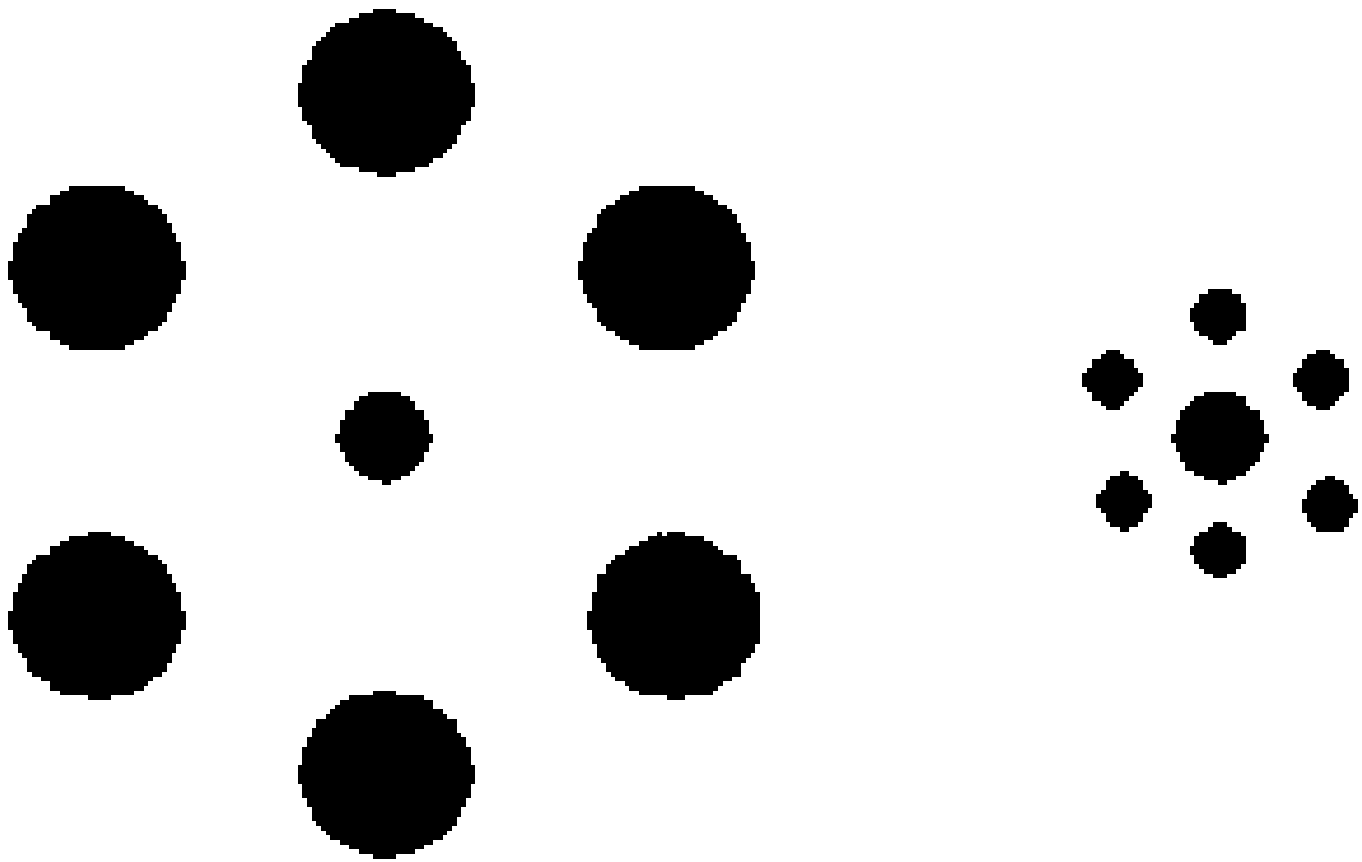
## Whitout a doubt, right?

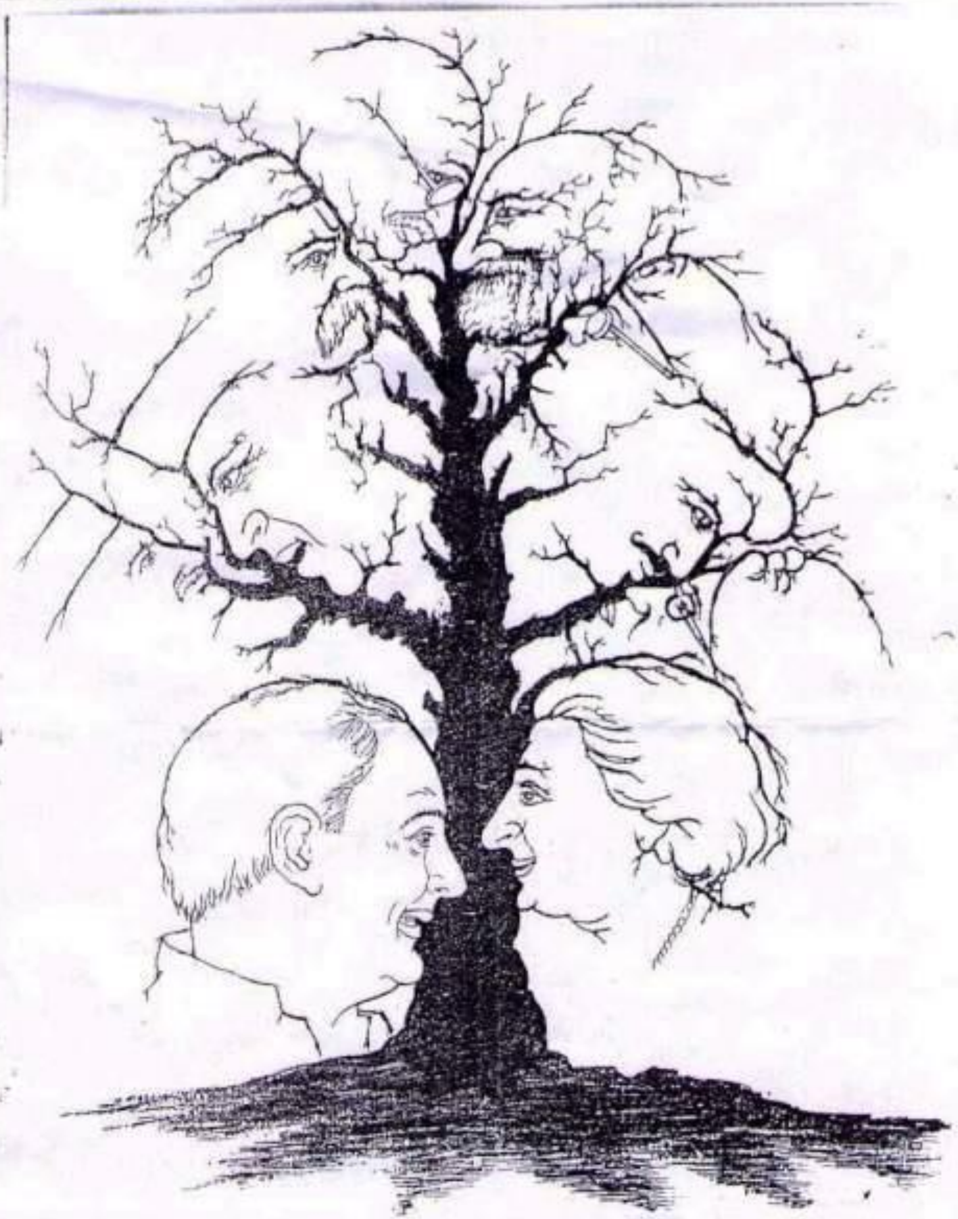
**Now, turn your head sideways...**







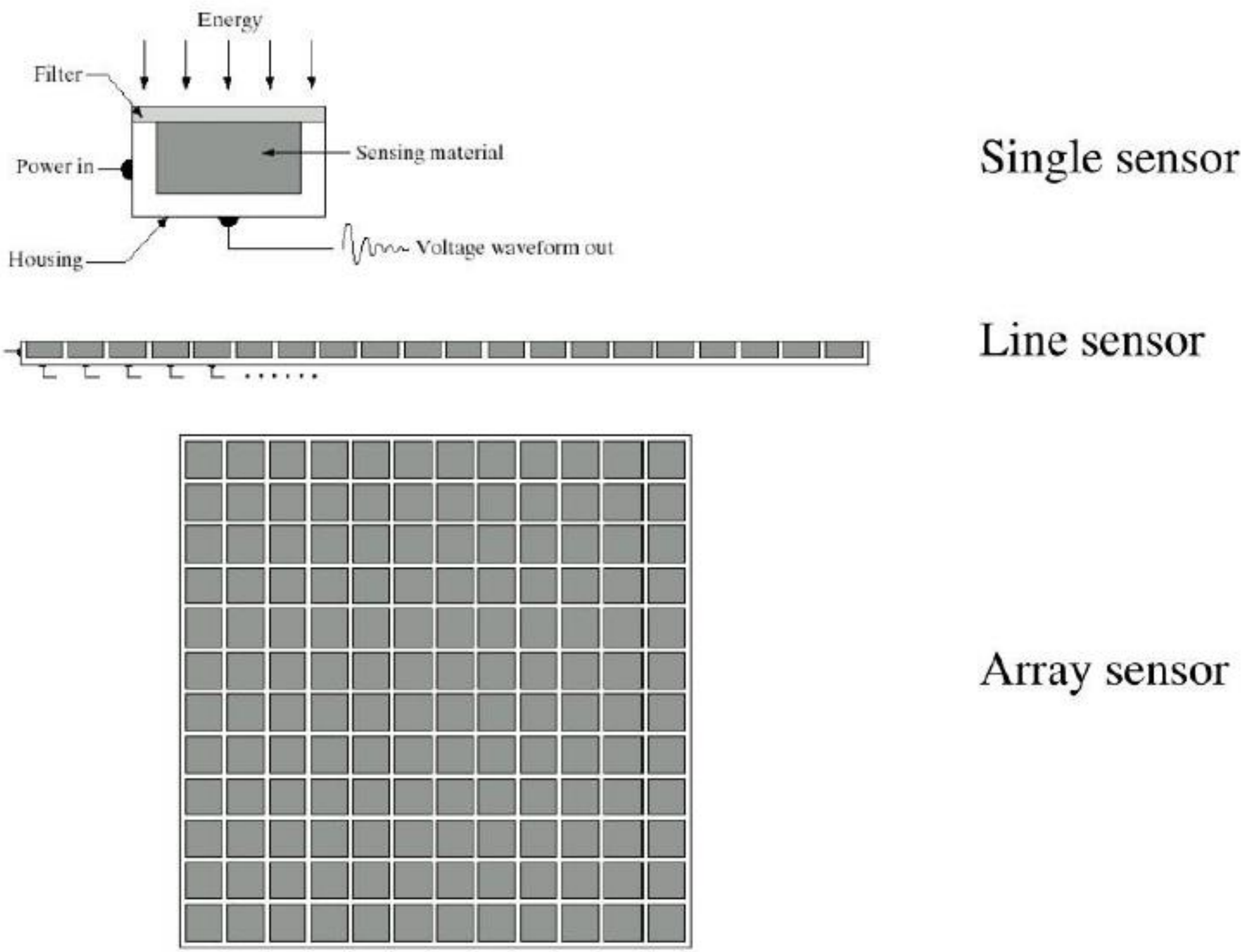




NATIONAL LEADERS TREE

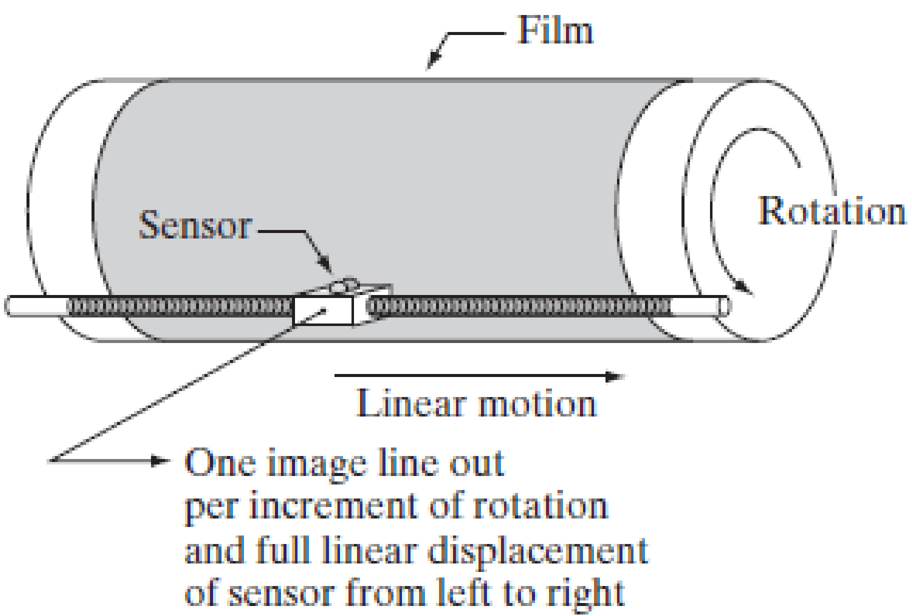
# Image sensors

- Incoming energy lands on a sensor material responsive to that type of energy and this generates a voltage
- Collections of sensor are arranged to compute images



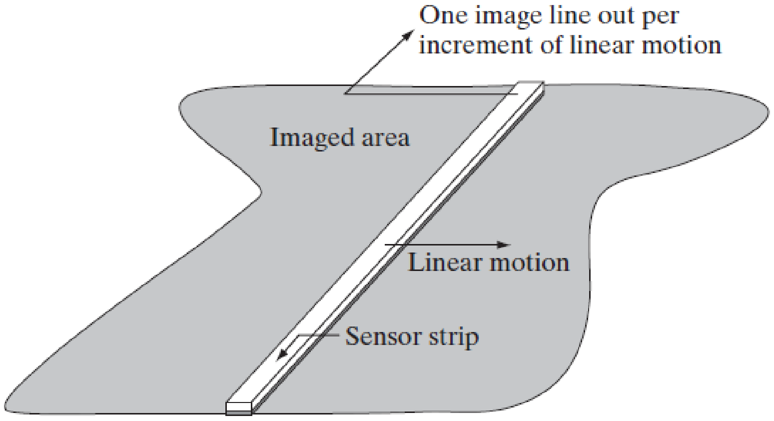
# Image Acquisition Using a Single Sensor

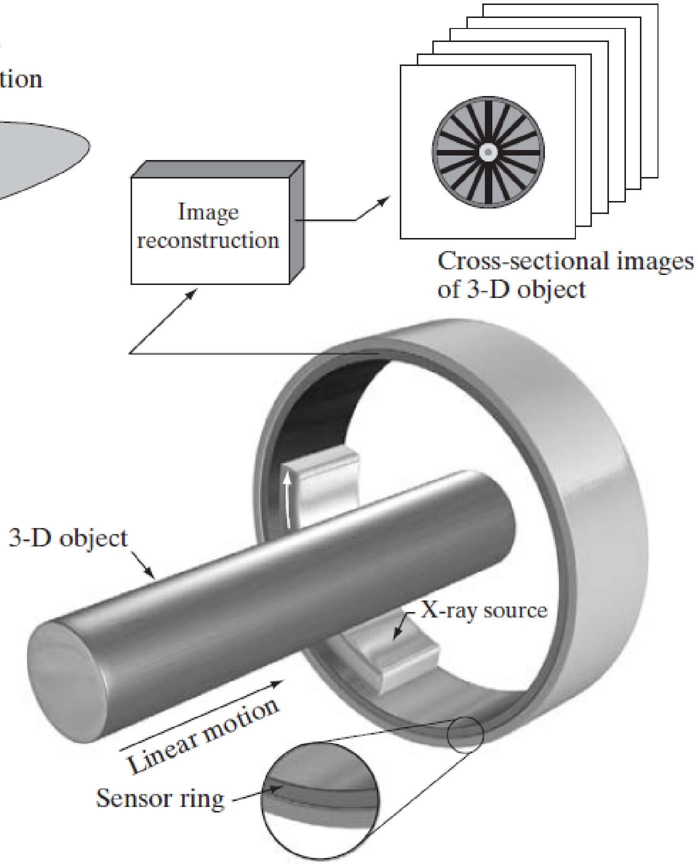
- In order to generate a 2-D image using a single sensor, there has to be relative displacements in both the x- and y-directions between the sensor and the area to be imaged
- An arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension



# IMAGE ACQUISITION USING SENSOR STRIPS

- Image acquisition using a linear sensor strip and
- Image acquisition using a circular sensor strip



# Simple Image Model

- $0 < f(x, y) < inf$
- $f(x, y)$ characterized by two components

The amount of source light incident on the scene being viewed $i(x, y)$

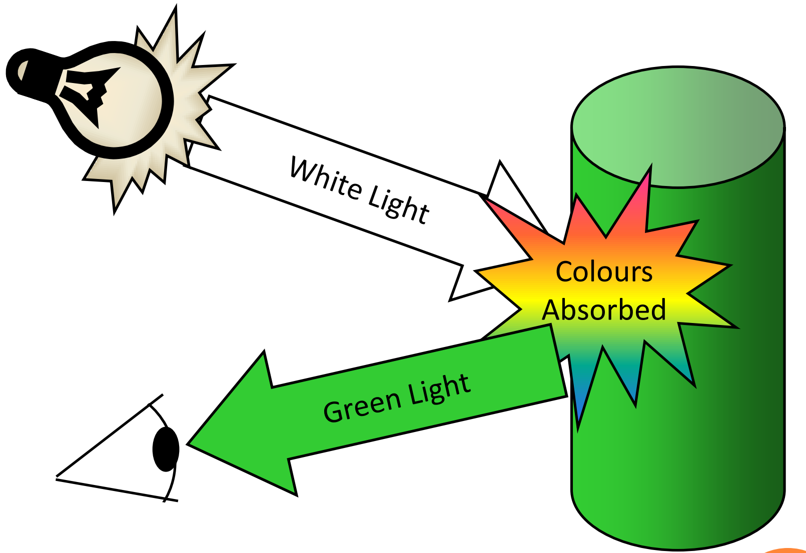
The amount of light reflected by the objects in the scene $r(x, y)$

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

Where, $0 < i(x, y) < inf$ and $0 < r(x, y) < 1$

# Reflected Light

- The colours that we perceive are determined by the nature of the light reflected from an object
- For example, if white light is shone onto a green object most wavelengths are absorbed, while green light is reflected from the object



# Simple Image Model

- The intensity of monochrome image at any coordinates $(x_0, y_0)$ is the gray level $\ell$ at that point
- That is

$$\ell = f(x_0, y_0)$$

$$\text{and, } L_{min} \leq \ell \leq L_{max}$$

$$L_{min} = i_{min}\, r_{min}$$

$$L_{max} = i_{max}\, r_{max}$$

- The interval $[L_{min}, L_{max}]$ is called gray scale
- Common practice, the interval is [0 L-1], where $\ell = 0$ is considered black and $\ell = L-1$ is considered white on the gray scale
- All intermediate values are shades of gray varying from black to white

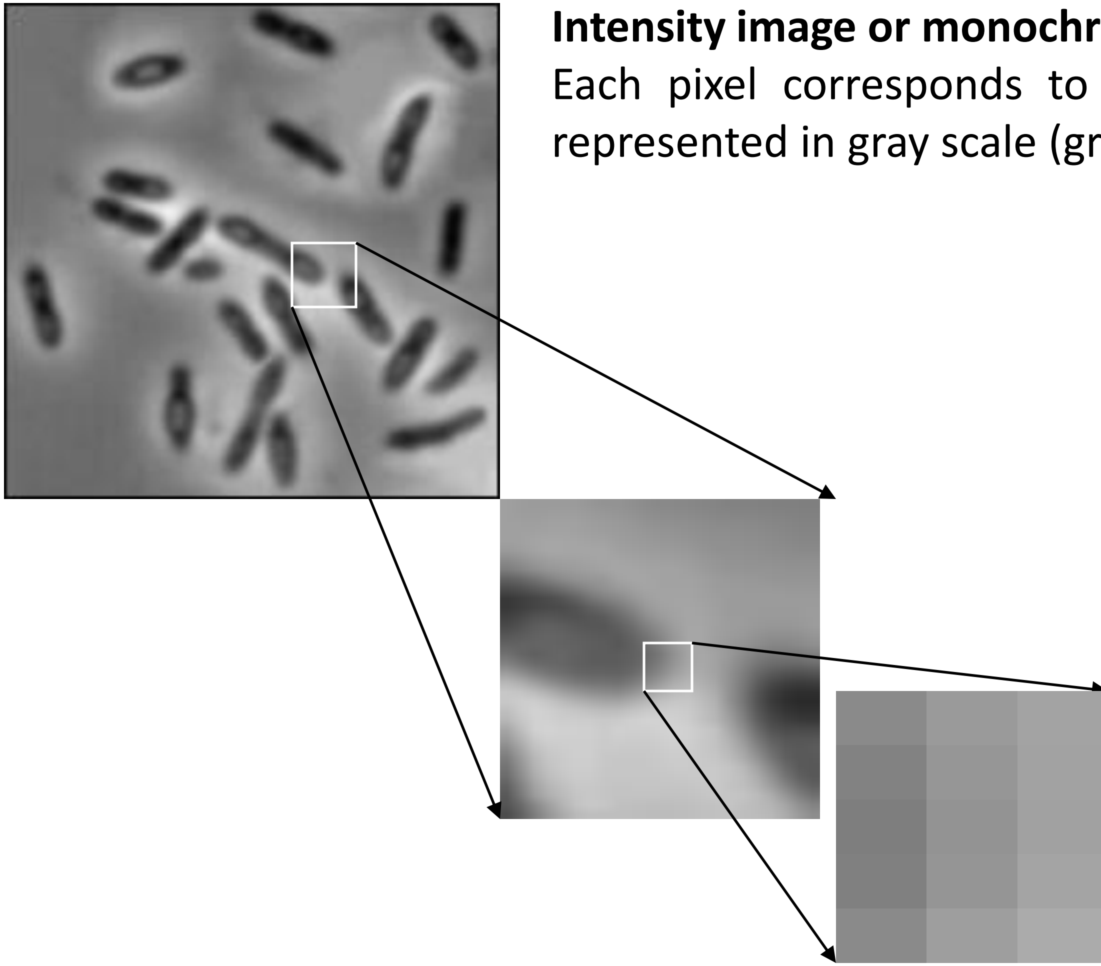
# Image Types

- ❑ Intensity (or grayscale) images
- ❑ RGB (or true color) images
- ❑ Binary Image
- ❑ Indexed images

# Intensity Image

**Intensity image or monochrome image**

Each pixel corresponds to light intensity normally represented in gray scale (gray level)

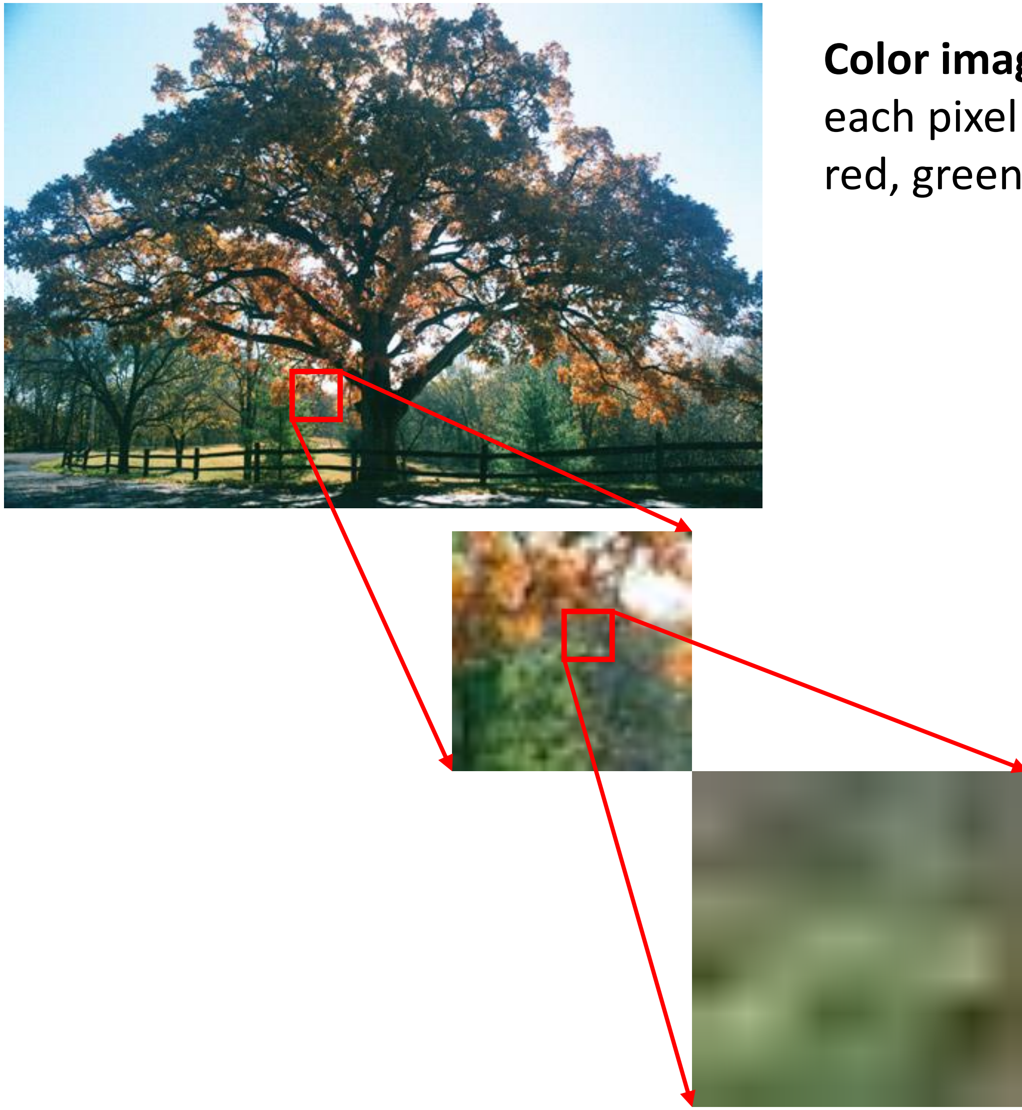


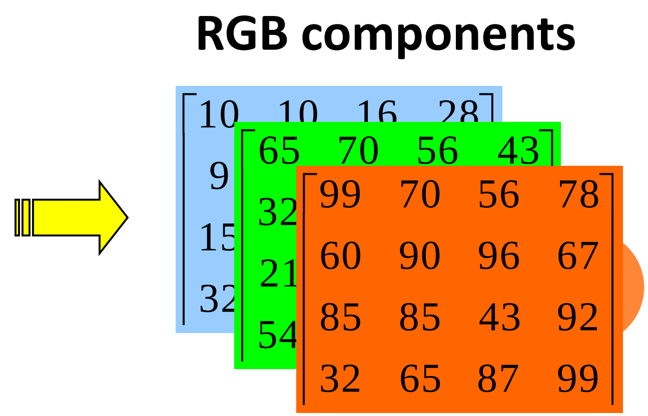
**Gray scale values**

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

# RGB Image

**Color image or RGB image**:
each pixel contains a vector representing red, green and blue components





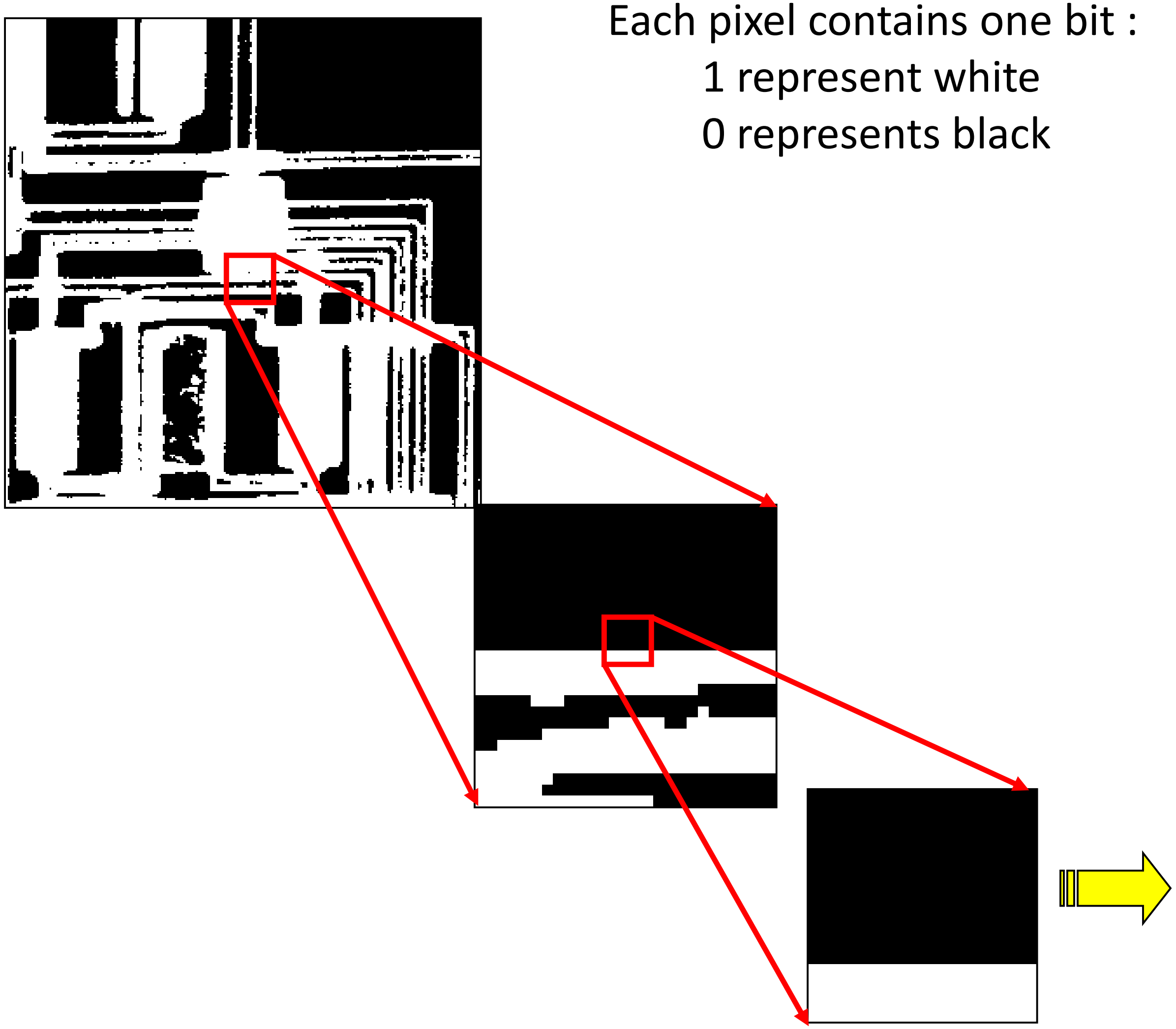
# Binary Image

**Binary image or black and white image**

Each pixel contains one bit :

1 represent white

0 represents black

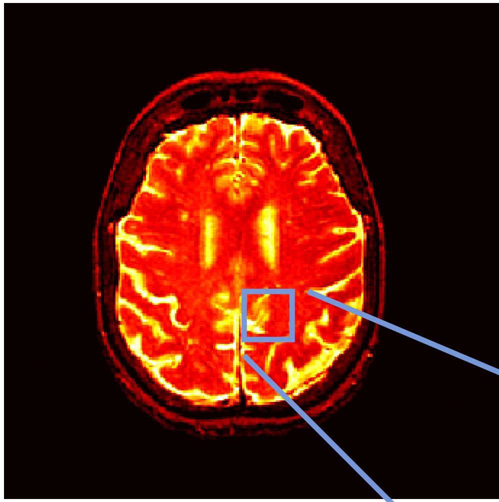


**Binary data**

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

# Index Image

**Index image**
Each pixel contains index number
pointing to a color in a color table



$$\begin{bmatrix} 1 & 4 & 9 \\ 6 & 4 & 7 \\ 6 & 5 & 2 \end{bmatrix}$$

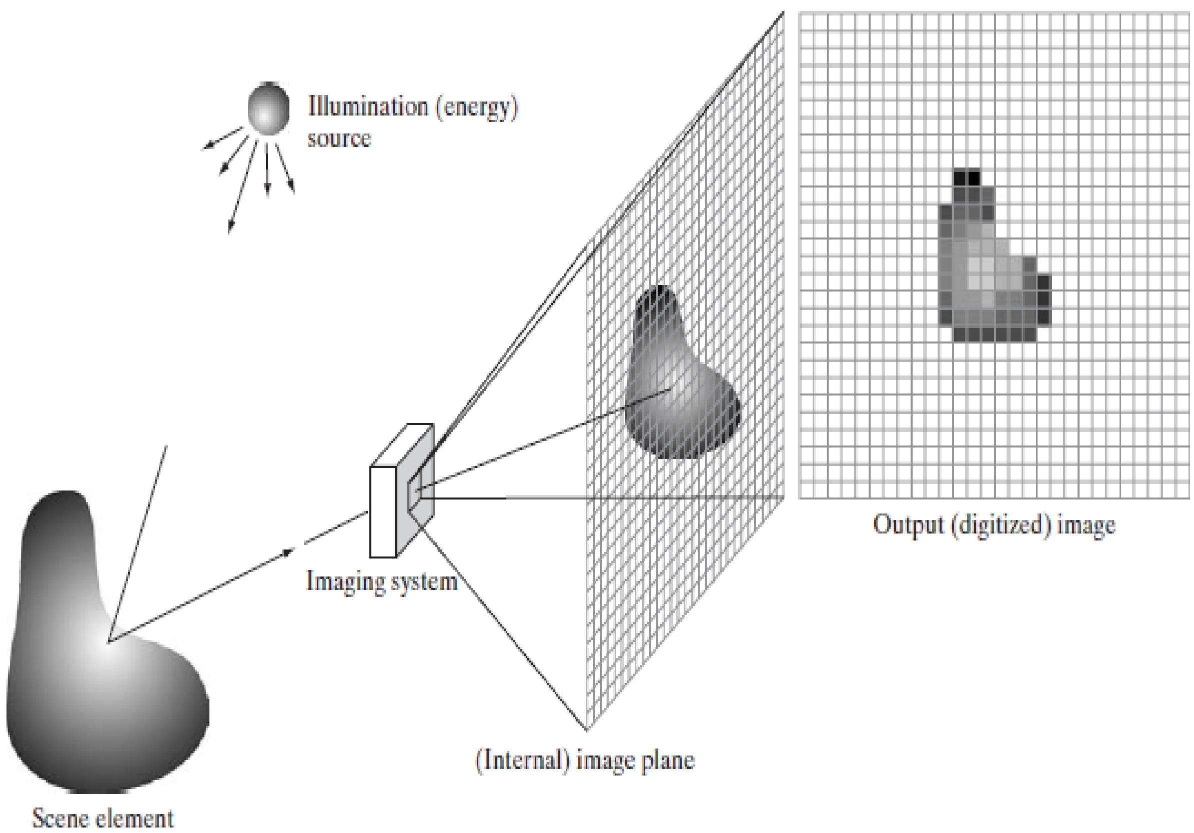
**Index value**

**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 |
| … | … | … | … |

# IMAGE ACQUISITION

- Images are typically generated by illuminating a scene and absorbing the energy reflected by the objects in that scene



# Image Sampling and Quantization

- An image is a continuous signal (distribution of continuous) light energy over a (continuous) plane
- A raw images are continuous in x and y coordinates and also in terms of amplitude
- To convert into digital form we have to sample the f(x, y) in both coordinates as well as in amplitude
- Digitizing the coordinate value is called **sampling** and digitizing amplitude is called **quantization**

# Digitization

- Conversion of the continuous (in space and value) signal into a digital (signal) image
- <u>We must determine</u>:
  - Spatial resolution (how many samples to take)
  - Signal or Gray-level resolution (dynamic range of values)
  - Tessellation Pattern (how to distribute the spatial samples on the image plane)

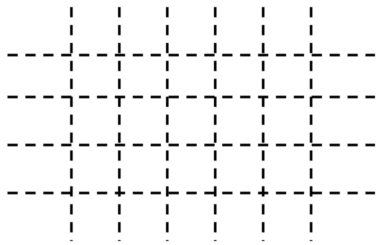
# DIGITIZATION

- **Spatial Resolution (Sampling)**
  - What should be the smallest perceivable image detail?
  - How much information of the original signal is kept?
  - Can we go back from the sampled signal to the continuous signal?
- **Gray-level resolution (Quantization)**
  - What should be the smallest discernible change in the gray level or color value?
  - Is there an optimal quantizer?

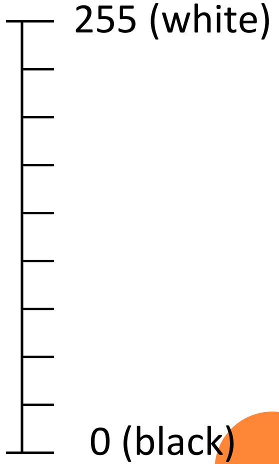
# Obtain Digital Image by Sampling + Quantization

- Computer handles “discrete” data

- Sampling
  - Sample the value of the image at the nodes of a regular grid on the image plane.
  - A pixel (picture element) at (i, j) is the image intensity value at grid point indexed by the integer coordinate (i, j)



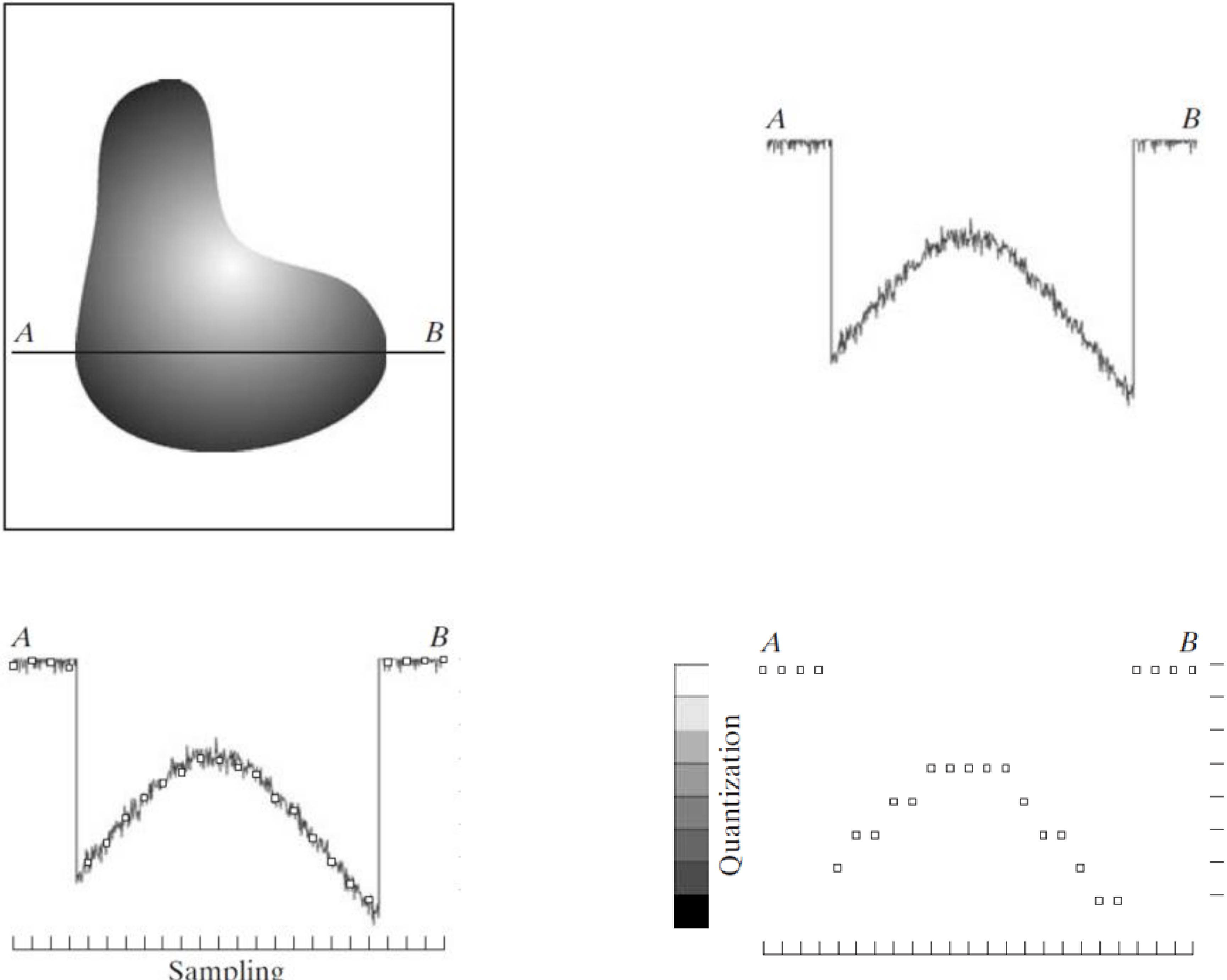
- Quantization
  - Is a process of transforming a real valued sampled image to one taking only a finite number of distinct values.
  - Each sampled value in a 256-level grayscale image is represented by 8 bits.



# Image Sampling and Quantization

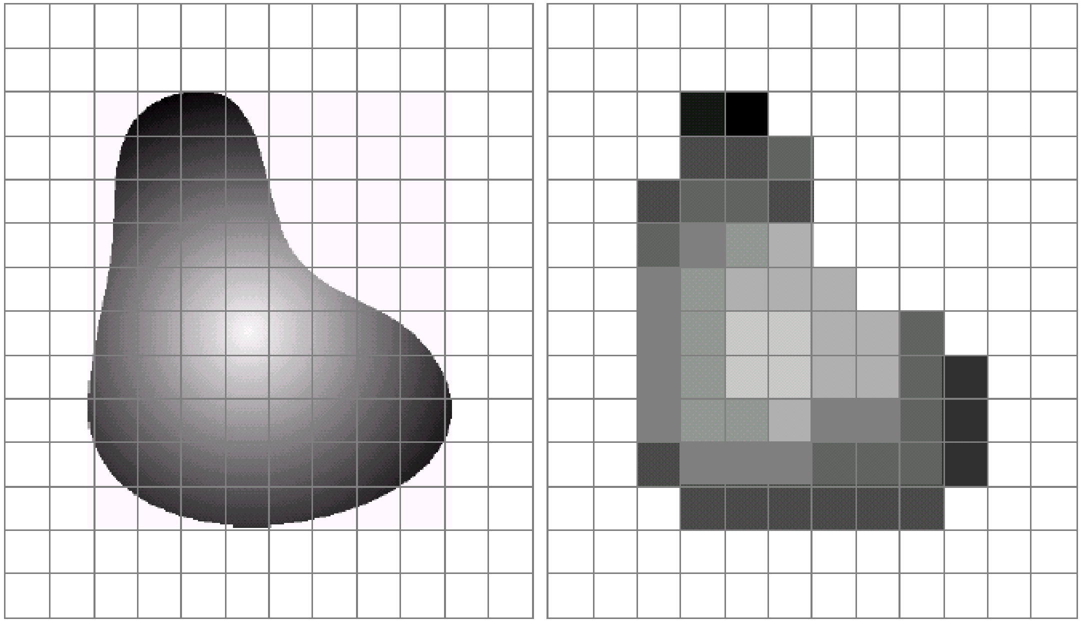
- A digital sensor can only measure a limited number of **samples** at a **discrete** set of energy levels

- *Quantization* is the process of converting a continuous **analogue** signal into a digital representation of this signal


A
B
A
B
A
B
Sampling
A
B
Quantization

# IMAGE SAMPLING AND QUANTISATION (CONT…)

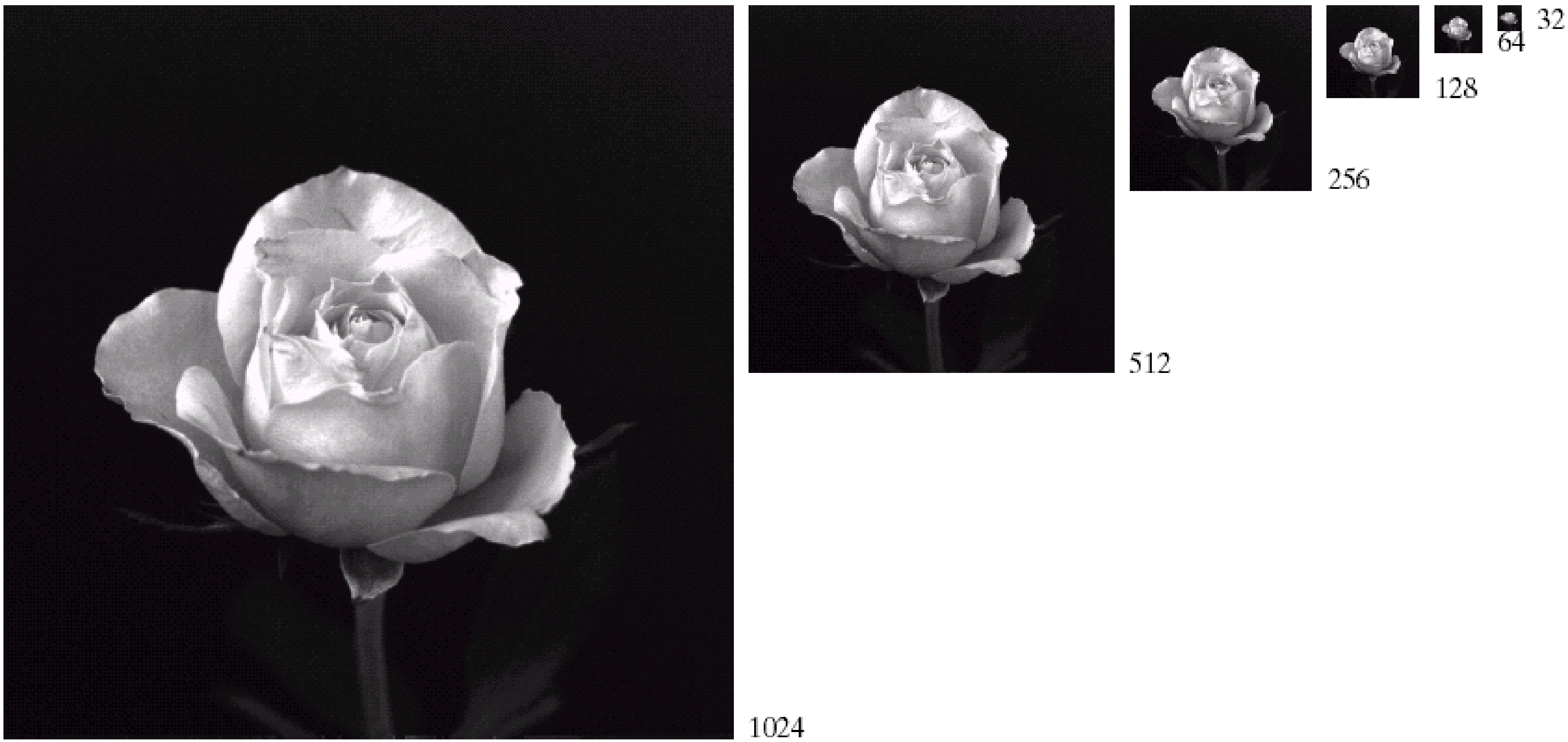
- Remember that a digital image is always only an **approximation** of a real world scene



# Spatial Resolution

- Spatial resolution is a measure of the smallest discernible detail in an image
- Quantitatively, spatial resolution can be stated in a number of ways, with line pairs per unit distance, and dots (pixels) per unit distance being among the most common measures

# Spatial Resolution



# Spatial Resolution (cont…)

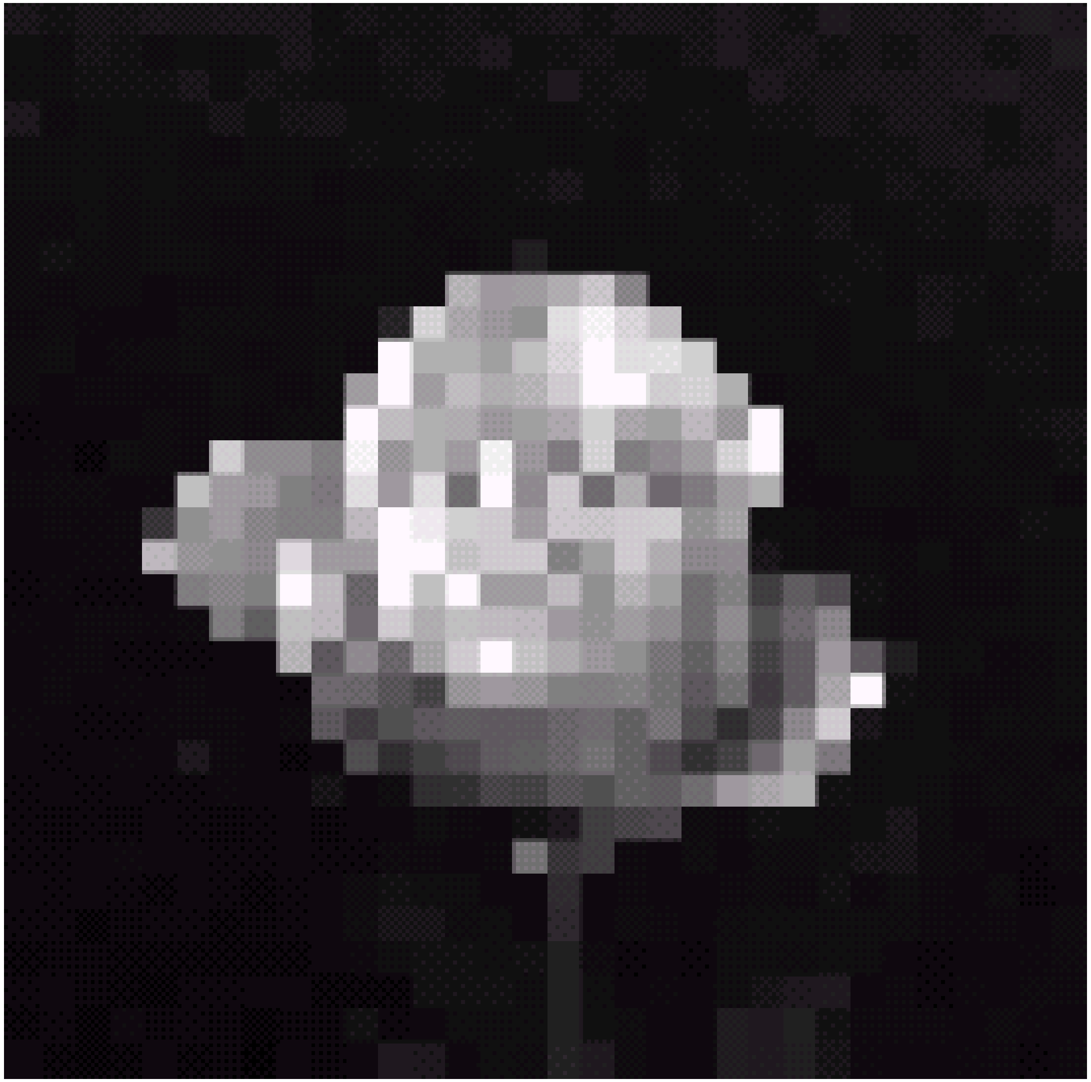
# Spatial Resolution (cont…)

# Spatial Resolution (cont…)

# Spatial Resolution (cont…)

# Spatial Resolution (cont...)

# Spatial Resolution (cont…)



# Intensity Level Resolution

- *Intensity level resolution* refers to the number of intensity levels used to represent the image
  - The more intensity levels used, the finer the level of detail discernible in an image
  - Intensity level resolution is usually given in terms of the number of bits used to store each intensity level

| Number of Bits | Number of Intensity Levels | Examples |
|---|---|---|
| 1 | 2 | 0, 1 |
| 2 | 4 | 00, 01, 10, 11 |
| 4 | 16 | 0000, 0101, 1111 |
| 8 | 256 | 00110011, 01010101 |
| 16 | 65,536 | 1010101010101010 |

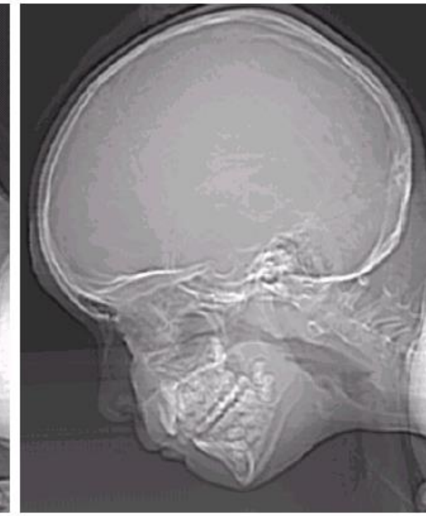
# Intensity Level Resolution (cont…)

256 grey levels (8 bits per pixel)

128 grey levels (7 bpp)

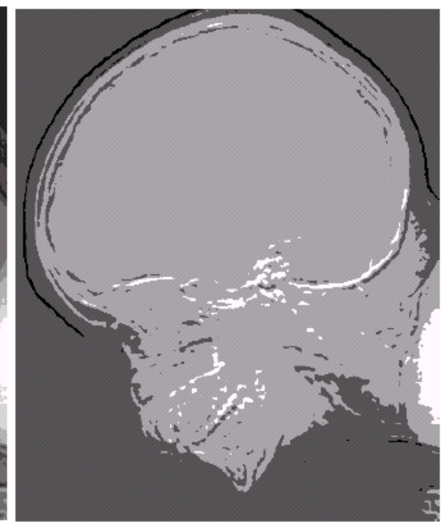
64 grey levels (6 bpp)

32 grey levels (5 bpp)



16 grey levels (4 bpp)

8 grey levels (3 bpp)



4 grey levels (2 bpp)

2 grey levels (1 bpp)

# Intensity Level Resolution (cont…)

# Intensity Level Resolution (cont…)

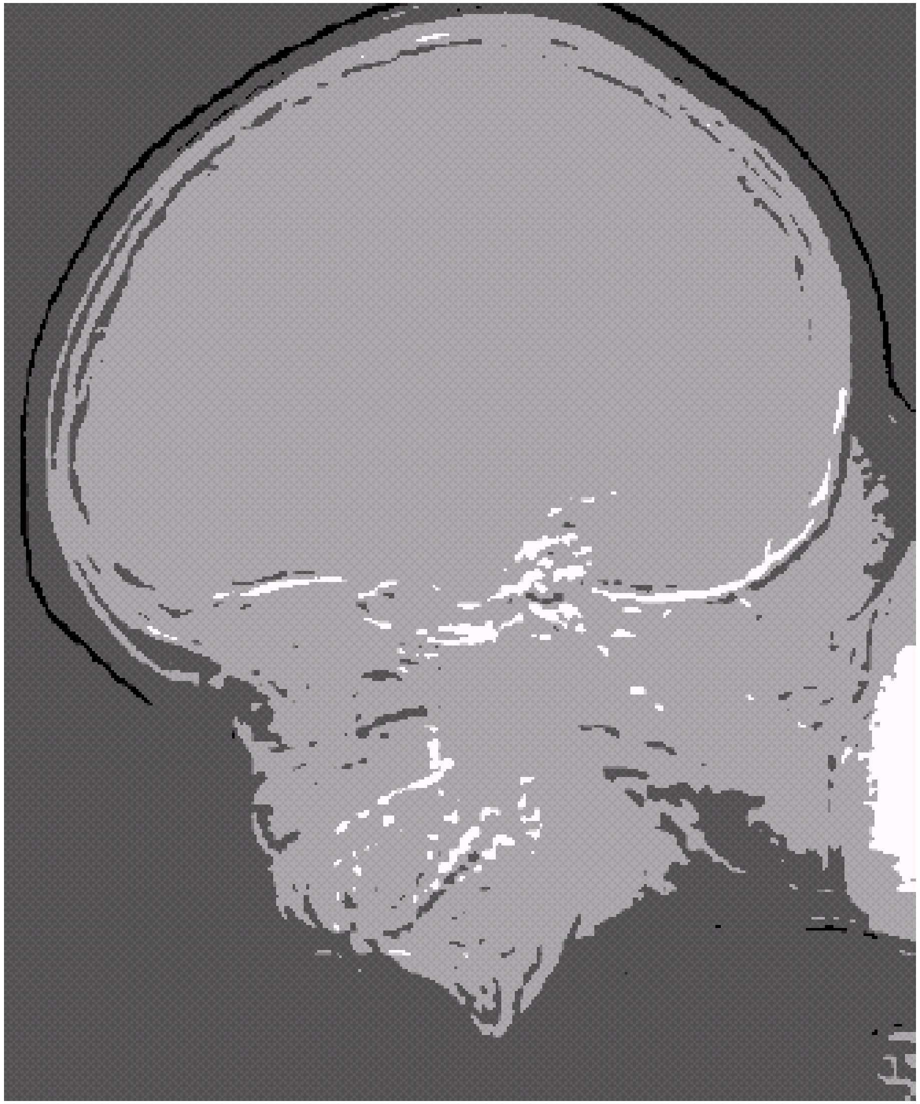
# INTENSITY LEVEL RESOLUTION (CONT…)

# Intensity Level Resolution (cont…)

# Intensity Level Resolution (cont...)

# Intensity Level Resolution (cont…)

# INTENSITY LEVEL RESOLUTION (CONT…)



# Intensity Level Resolution (cont…)

# Resolution: How Much Is Enough?

- The big question with resolution is always *how much is enough*?
  - This all depends on what is in the image and what you would like to do with it
  - Key questions include
    - Does the image look aesthetically pleasing?
    - Can you see what you need to see within the image?

# Resolution: How Much Is Enough? (cont…)


- The picture on the right is fine for counting the number of cars, but not for reading the number plate

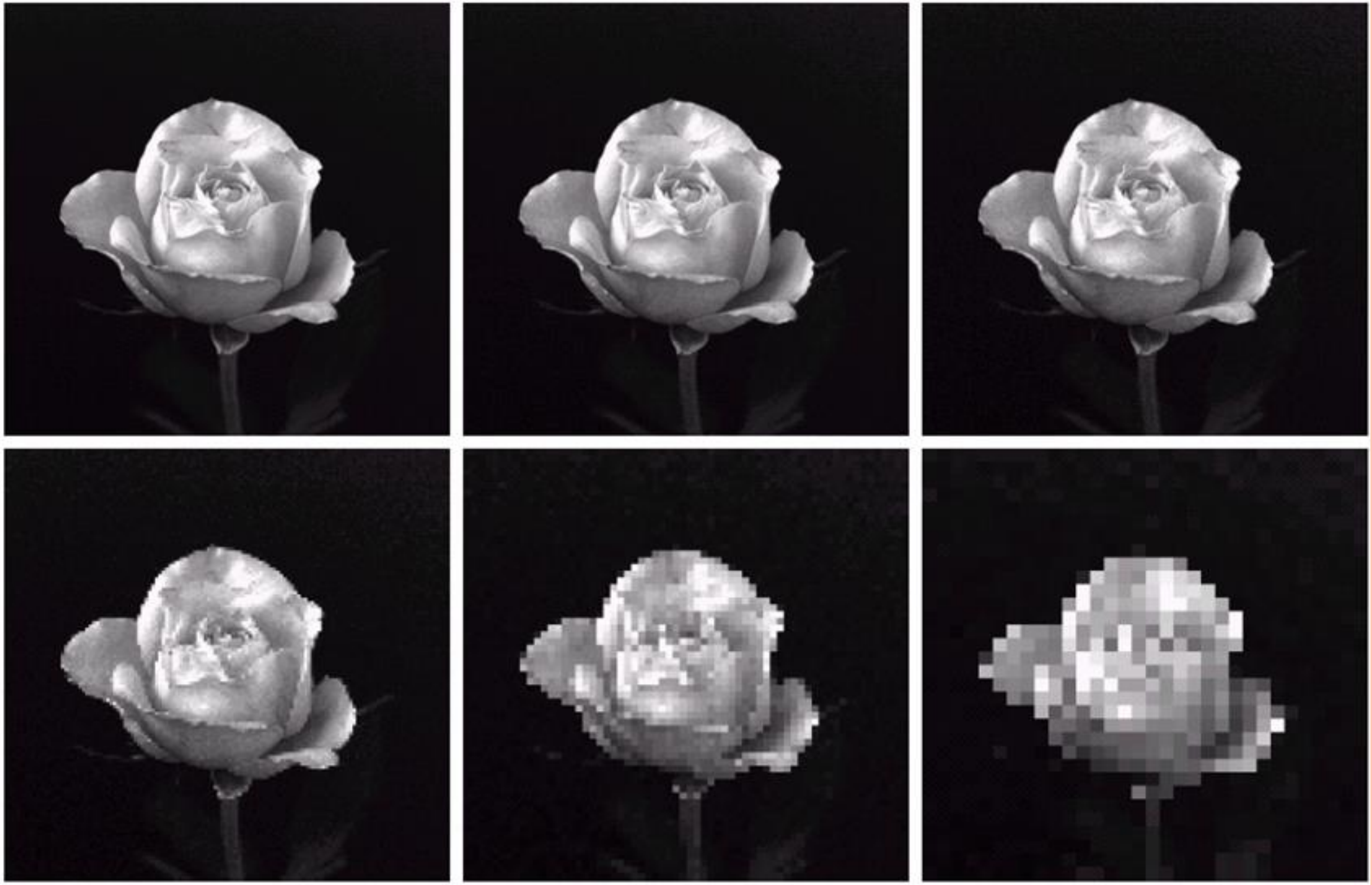
# INTENSITY LEVEL RESOLUTION (CONT…)

Low Detail

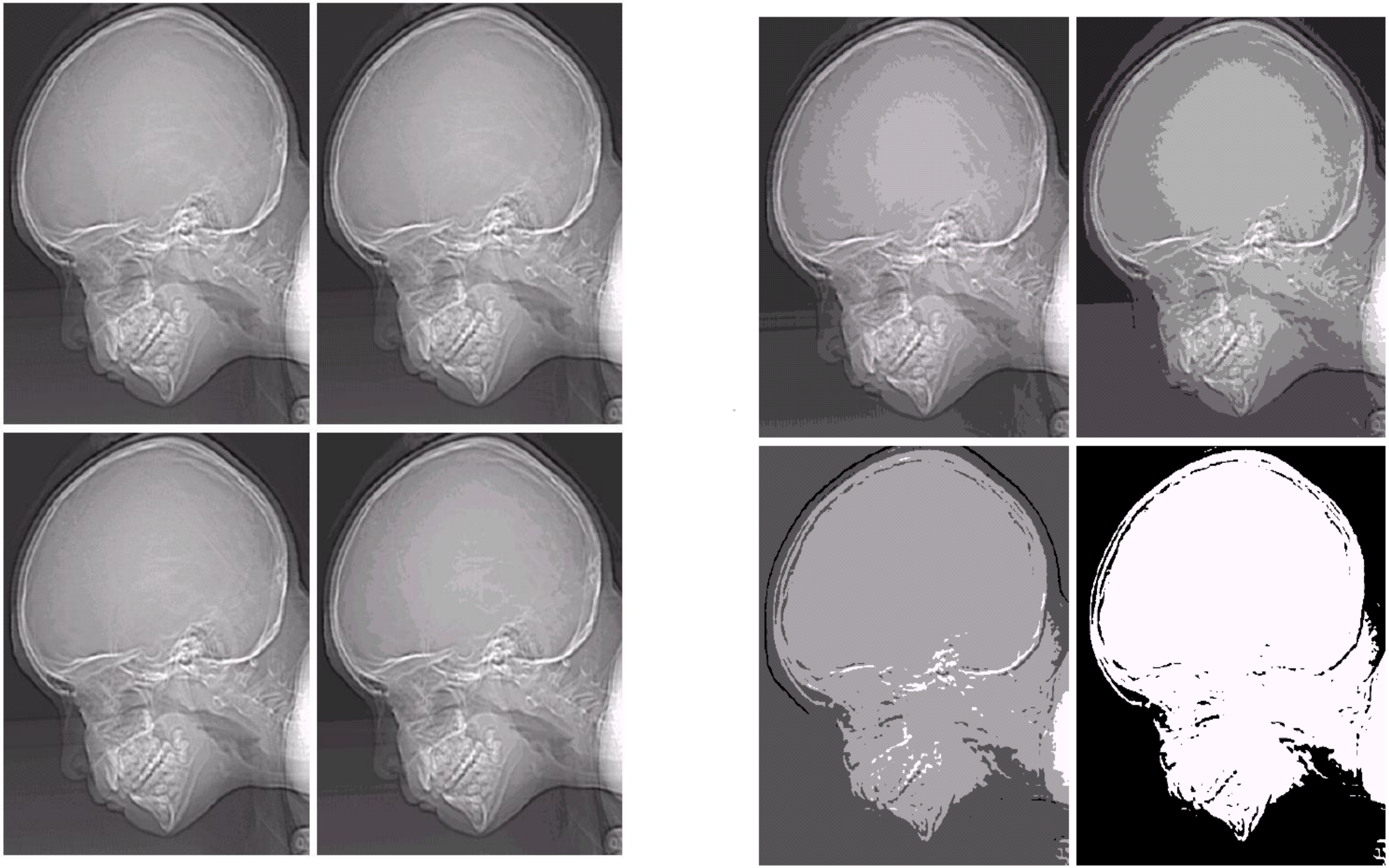
Medium Detail

High Detail

# Reduction of Space Resolution:
# Chessboard Effect



# Reduction of Grey Level Resolution:
# False Contouring



# Storing of Image

- Assume, Image size (M,N), Quantization level L, where $L=2^k$
- Assume that the discrete levels are equally spaced and that they are integers in the interval [0 L-1].
- The number, b of bits required to store a digitized image is

$$b = M * N * k.$$

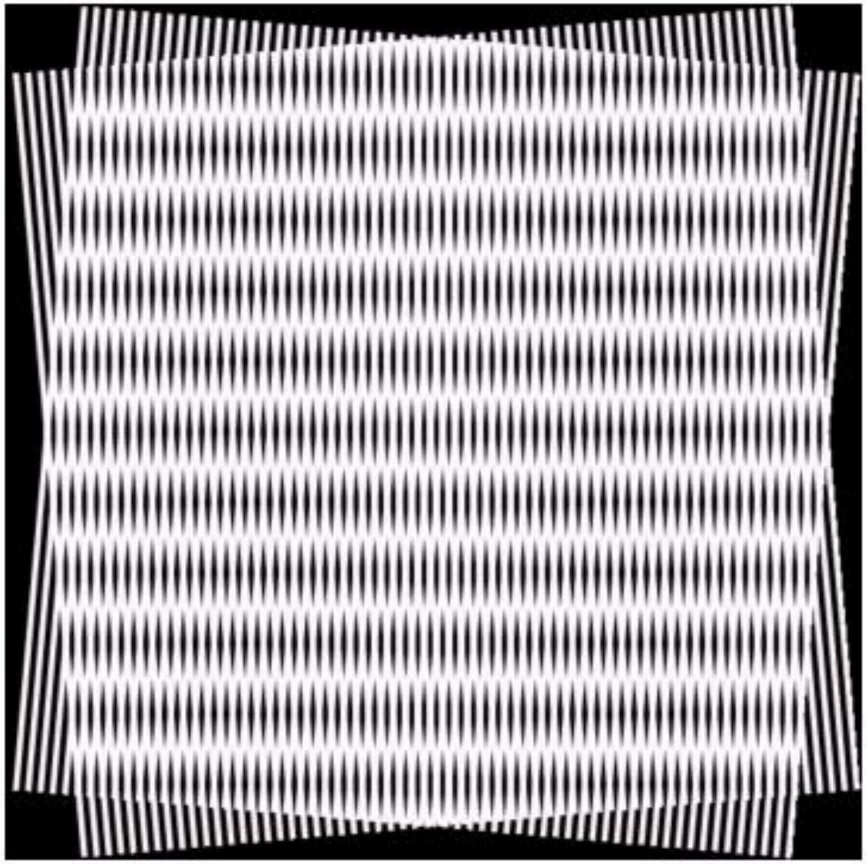
- If M=N, this equation become,

$$b = N^2 k$$

# High contrast and low contrast image

- Discrete levels are equally spaced in the range [0 L - 1]
- Sometimes the range of values spanned by the gray scale is called the dynamic range of an image
- Image whose gray level span a significant portion of the gray scale as having a high dynamic range is called high contrast image
- Conversely, an image with low dynamic range tends to have a dull, washed out gray look called low contrast image

# Moire Pattern Effect : Special Case of Sampling



Moire patterns occur when frequencies of  two superimposed periodic patterns are close to each other.

# Image Interpolation

- Interpolation is a basic tool used extensively in tasks such as zooming, shrinking, rotating, and geometric corrections
- *Interpolation* is the process of using known data to estimate values at unknown locations
- Zooming may be viewed as a over sampling and shrinking may be viewed as a under sampling
- Zooming requires two steps: the creation of new pixels locations, and the assignment of gray level to those location

# Image Interpolation

- ***Nearest neighbor interpolation:*** assigns to each new location the intensity of its nearest neighbor in the original image

- This method has the tendency to produce undesirable artifacts, such as severe distortion of straight edges

# Image Interpolation

- ***Bilinear interpolation:*** In this method, we use the four nearest neighbors to estimate the intensity at a given location

- Let $(x, y)$ denote the coordinates of the location to which we want to assign an intensity value and let $v(x, y)$ denote that intensity value

- For bilinear interpolation, the assigned value is obtained using the equation

$$v(x, y) = ax + by + cxy + d$$

- where the four coefficients are determined from the four equations in four unknowns that can be written using the four nearest neighbors of point $(x, y)$

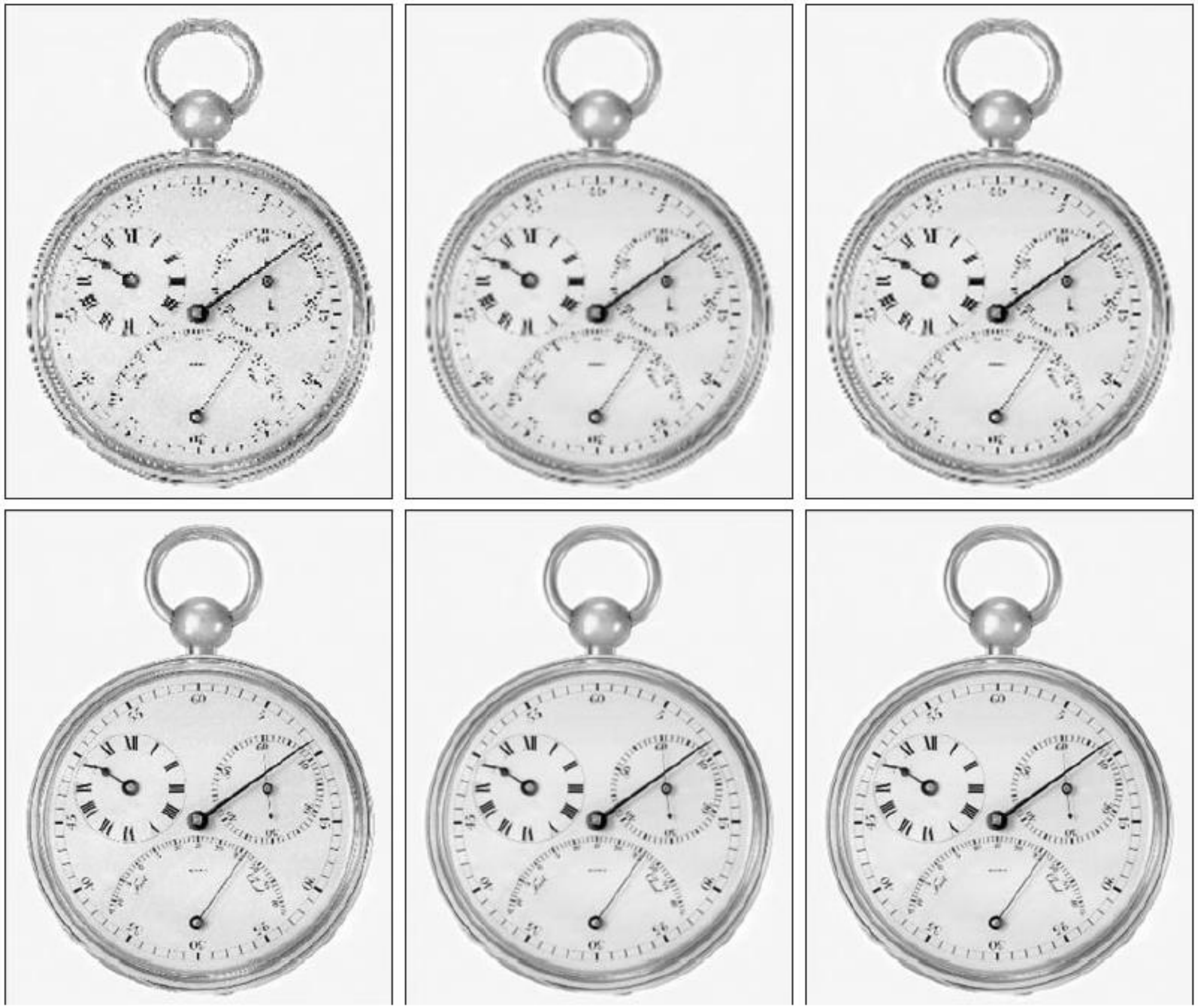
# Image Interpolation

- ***Bicubic interpolation:*** This method involves the sixteen nearest neighbors of a point

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

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

- where the sixteen coefficients are determined from the sixteen equations in sixteen unknowns that can be written using the sixteen nearest neighbors of point *(x, y)*
- Bicubic interpolation is the standard used in commercial image editing programs, such as Adobe Photoshop and Corel Photo paint



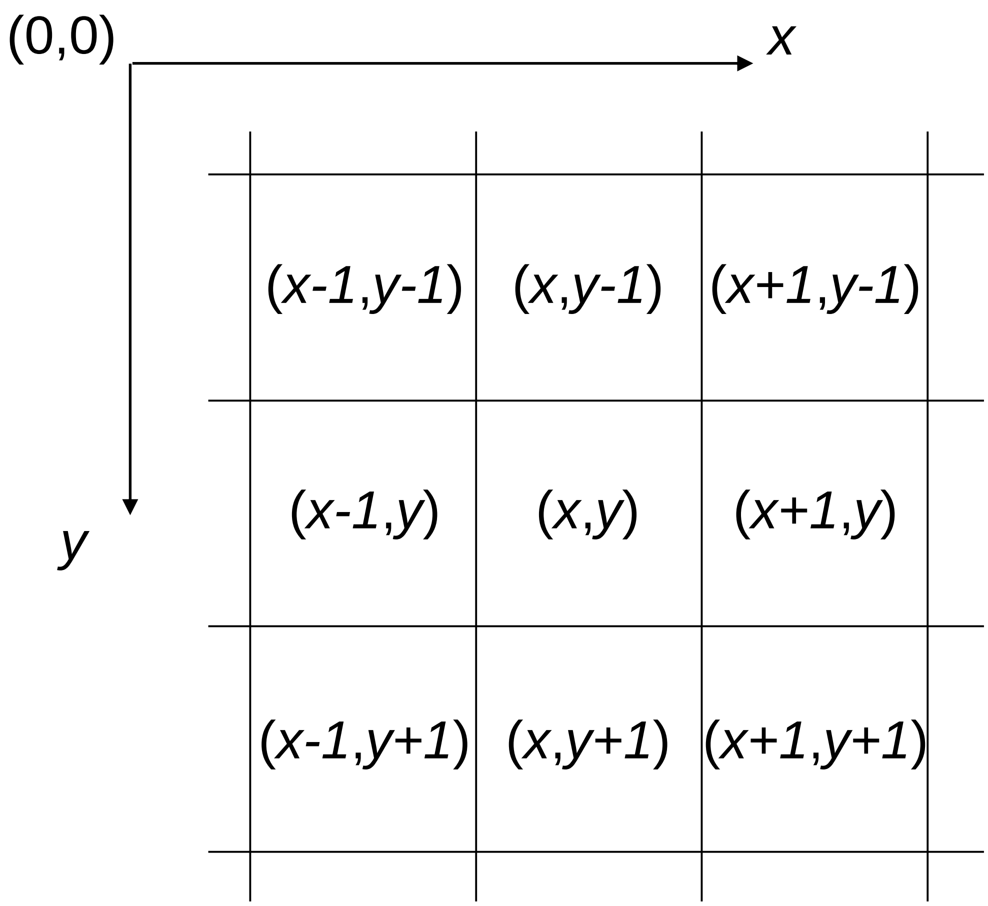
a b c
d e f

FIGURE (a) Image reduced to 72 dpi and zoomed back to its original size (pixels) using nearest neighbor interpolation. This figure is the same as Fig. (d)

(b) Image shrunk and zoomed using bilinear interpolation, (c) Same as (b) but using bicubic interpolation

(d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. (d) is the same as Fig. (c)]. Compare Figs. (e) and (f), especially the latter, with the original image in Fig. (a).

# Basic Relationships between Pixels




Conventional indexing method

# *Neighbors of a Pixel*

Neighborhood relation is used to tell adjacent pixels. It is useful for analyzing regions.

| | | |
|---|---|---|
| | $(x,y-1)$ | |
| $(x-1,y)$ | $p$ | $(x+1,y)$ |
| | $(x,y+1)$ | |

**4-neighbors of *p*:**

$$N_4(p) = \left\{ \begin{array}{c} (x-1, y) \\ (x+1, y) \\ (x, y-1) \\ (x, y+1) \end{array} \right\}$$

4-neighborhood relation considers only vertical and horizontal neighbors.

Note: $q \in N_4(p)$ implies $p \in N_4(q)$

# *Neighbors of a Pixel (cont.)*

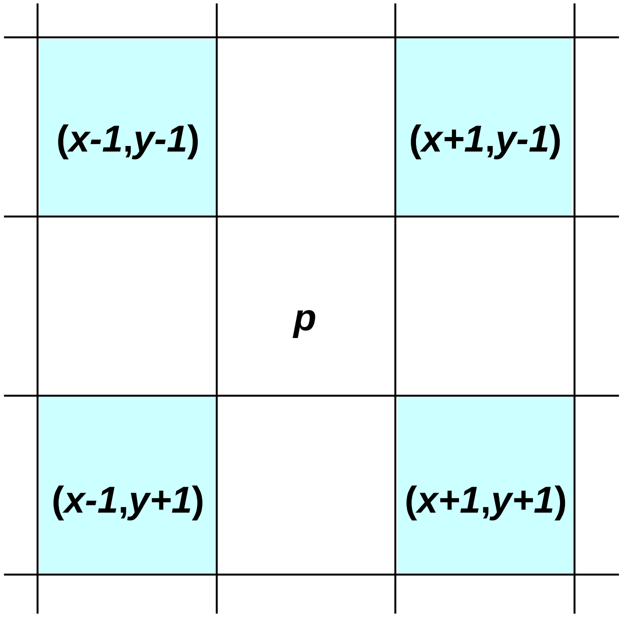
| | | |
|---|---|---|
| **(*x-1,y-1*)** | **(*x,y-1*)** | **(*x+1,y-1*)** |
| **(*x-1,y*)** | ***p*** | **(*x+1,y*)** |
| **(*x-1,y+1*)** | **(*x,y+1*)** | **(*x+1,y+1*)** |

**8-neighbors of *p*:**

$$N_8(p) = \left\{ \begin{array}{c} (x-1,\ y-1) \\ (x,\ y-1) \\ (x+1,\ y-1) \\ (x-1,\ y) \\ (x+1,\ y) \\ (x-1,\ y+1) \\ (x,\ y+1) \\ (x+1,\ y+1) \end{array} \right\}$$

8-neighborhood relation considers all neighbor pixels.

# *Neighbors of a Pixel (cont.)*




## Diagonal neighbors of *p*:

$$N_D(p) = \left\{ \begin{array}{c} (x-1,\ y-1) \\ (x+1,\ y-1) \\ (x-1,\ y+1) \\ (x+1,\ y+1) \end{array} \right\}$$

Diagonal -neighborhood relation considers only diagonal neighbor pixels.

# Connectivity

- Connectivity between pixels is an important concept used in establishing boundaries of objects and components of regions in an image
- To establish whether two pixels are connected, it must be determined if they are adjacent in some sense and if their gray levels satisfy a specified criterion of similarity

# Connectivity

- Connectivity is adapted from neighborhood relation
- Two pixels are connected if they are in the same class (i.e. the same color or the same range of intensity) and they are neighbors of one another

For $p$ and $q$ from the same class

- 4-connectivity: $p$ and $q$ are 4-connected if $q \in N_4(p)$
- 8-connectivity: $p$ and $q$ are 8-connected if $q \in N_8(p)$
- mixed-connectivity (m-connectivity):

$p$ and $q$ are m-connected if $q \in N_4(p)$ or
$q \in N_D(p)$ and $N_4(p) \cap N_4(q) = \varnothing$

# Connectivity

0 1 1
0 1 0 Arrangement of pixels
0 0 1

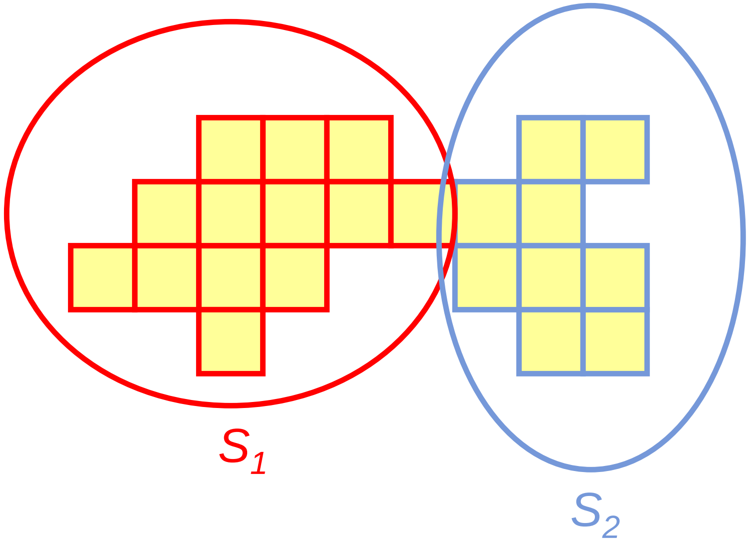
0 1 1
0 1 0 4-connectivity
0 0 1

0 1 1
0 1 0 8-connectivity
0 0 1

0 1 1
0 1 0 m-connectivity
0 0 1

# Adjacency

- A pixel $p$ is *adjacent* to pixel $q$ if they are connected
- Two image subsets $S_1$ and $S_2$ are adjacent if some pixel in $S_1$ is adjacent to some pixel in $S_2$




We can define type of adjacency: 4-adjacency, 8-adjacency or m-adjacency depending on type of connectivity

# Path

- A ***path*** from pixel $p$ at $(x,y)$ to pixel $q$ at $(s,t)$ is a sequence of distinct pixels:

  $$(x_0,y_0),\ (x_1,y_1),\ (x_2,y_2),\ldots,\ (x_n,y_n)$$

  such that

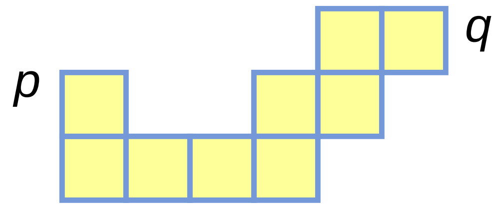
  $$(x_0,y_0) = (x,\ y),\ (x_n,y_n) = (s,t)$$

  and

  $$(x_i,y_i) \text{ is adjacent to } (x_{i-1},y_{i-1}), \qquad i = 1,\ldots,n$$

- In this case, n is the length of the path

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

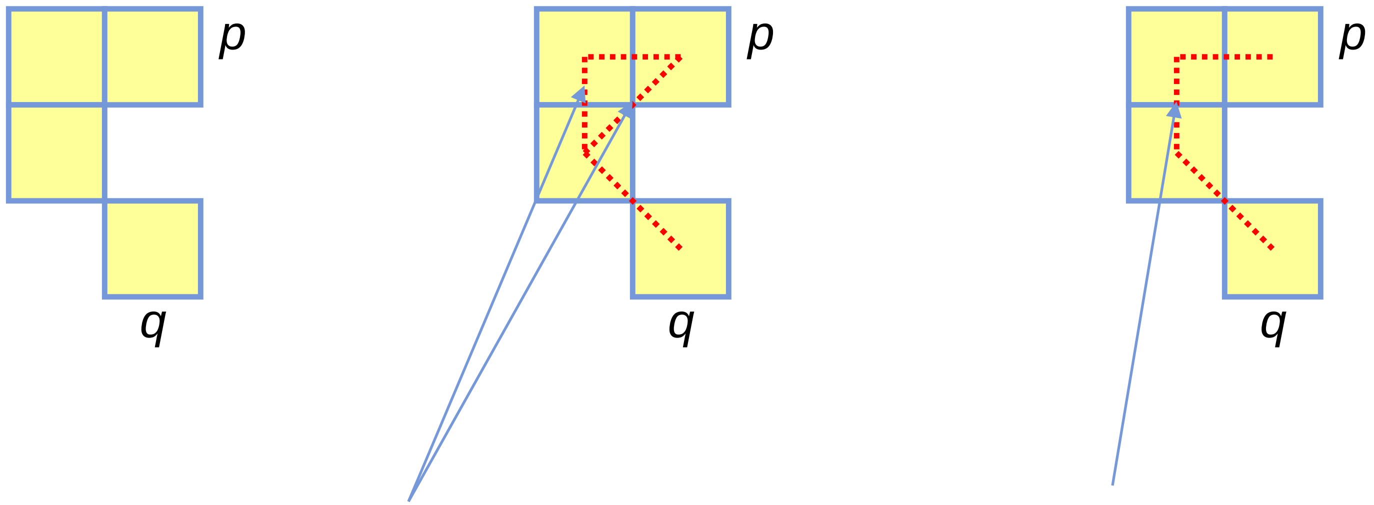



- We can define type of path: 4-path, 8-path or m-path depending on type of adjacency

# Path (cont.)

8-path

m-path




8-path from *p* to *q*
results in some ambiguity

m-path from *p* to *q*
solves this ambiguity

# Distance

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

- $D(p, q) \geq 0$ $(D(p,q) = 0$ if and only if $p = q)$
- $D(p, q) = D(q,p)$
- $D(p, z) \leq D(p,q) + D(q,z)$

Example: Euclidean distance

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

$D_e$ distance measure, the pixels having a distance less than or equal to some value $r$ from $(x, y)$ are the points contained in a disk of radius r centered at $(x, y)$

# Distance (cont.)

***$D_4$-distance*** (*city-block distance*) is defined as

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

| | | | | |
|---|---|---|---|---|
| | | 2 | | |
| | 2 | 1 | 2 | |
| 2 | 1 | 0 | 1 | 2 |
| | 2 | 1 | 2 | |
| | | 2 | | |

Pixels with $D_4(p) = 1$ is 4-neighbors of $p$.

# Distance (cont.)

***$D_8$-distance*** (*chessboard distance*) is defined as

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

| | | | | |
|---|---|---|---|---|
| 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 |

Pixels with $D_8(p) = 1$ is 8-neighbors of $p$.

# Arithmetic operation

- The principal use of image addition is for image averaging to reduce noise
- Image subtraction is a basic tool in medical imaging, where it is used to remove static background information
- One of the principal uses of image multiplication is to correct gray-level shading resulting from non-uniformities in illumination or in the sensor used to acquire the image

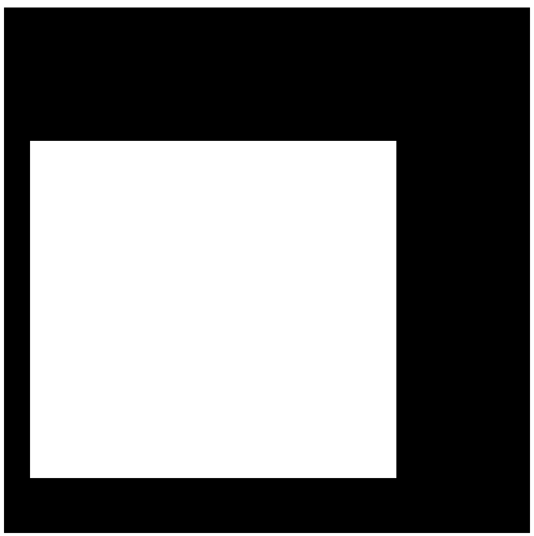
# Logic operation

- Logic operation apply only in binary image whereas arithmetic operation apply to multivalued pixels

- Logic operations are basic tools in binary image processing, where they are used for task such as masking, feature detection, and shape analysis

# NOT

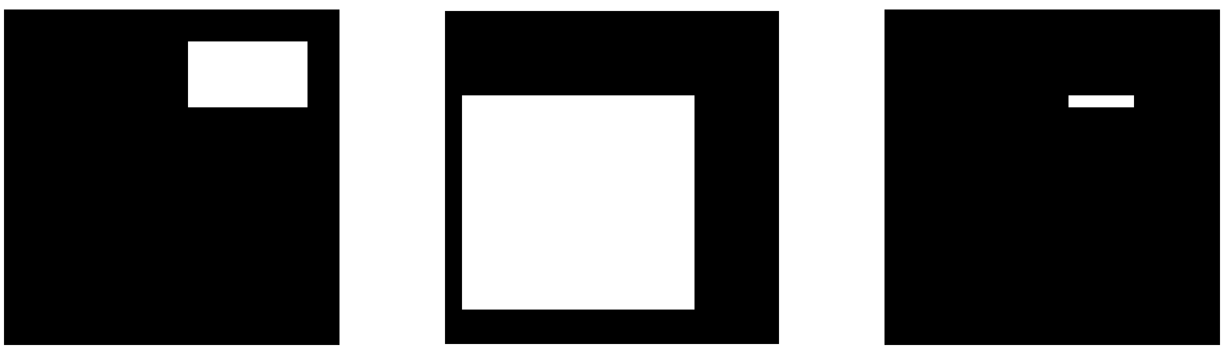
B



NOT(B)

# AND

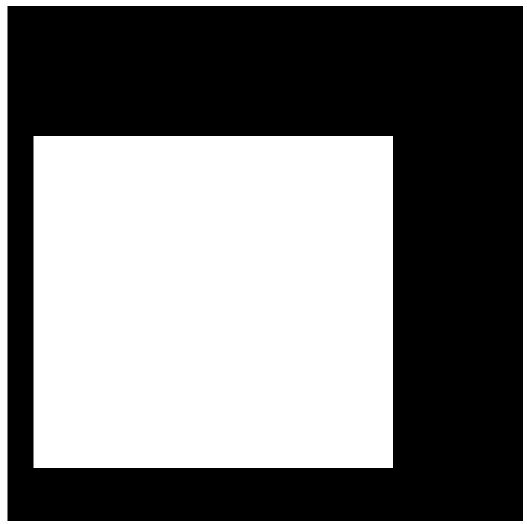
A B A AND B



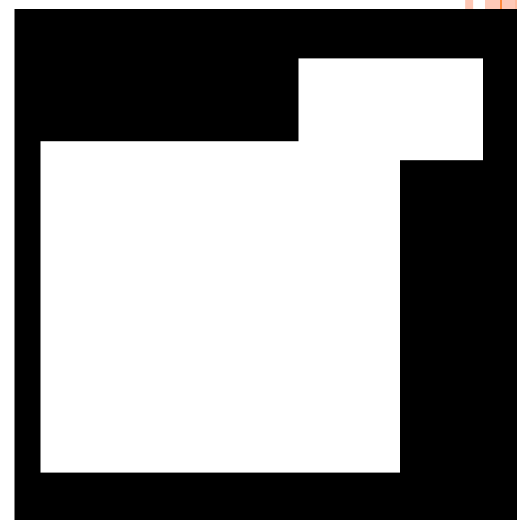
# OR

A

B



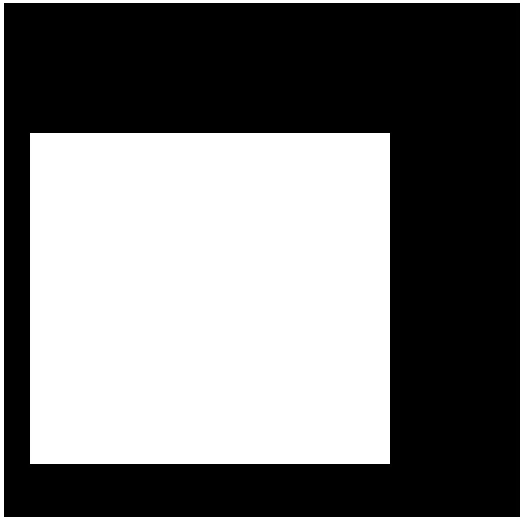
A OR B



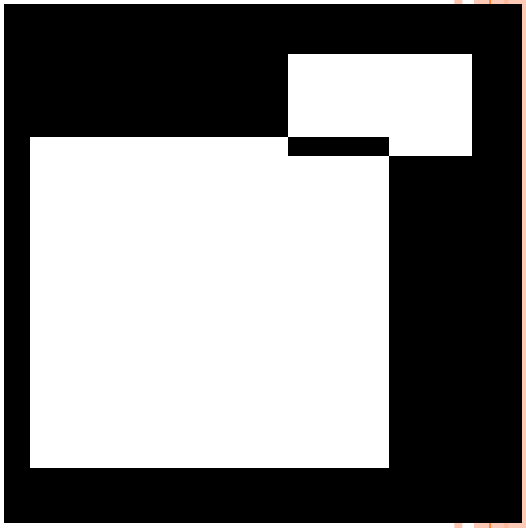
# XOR

A

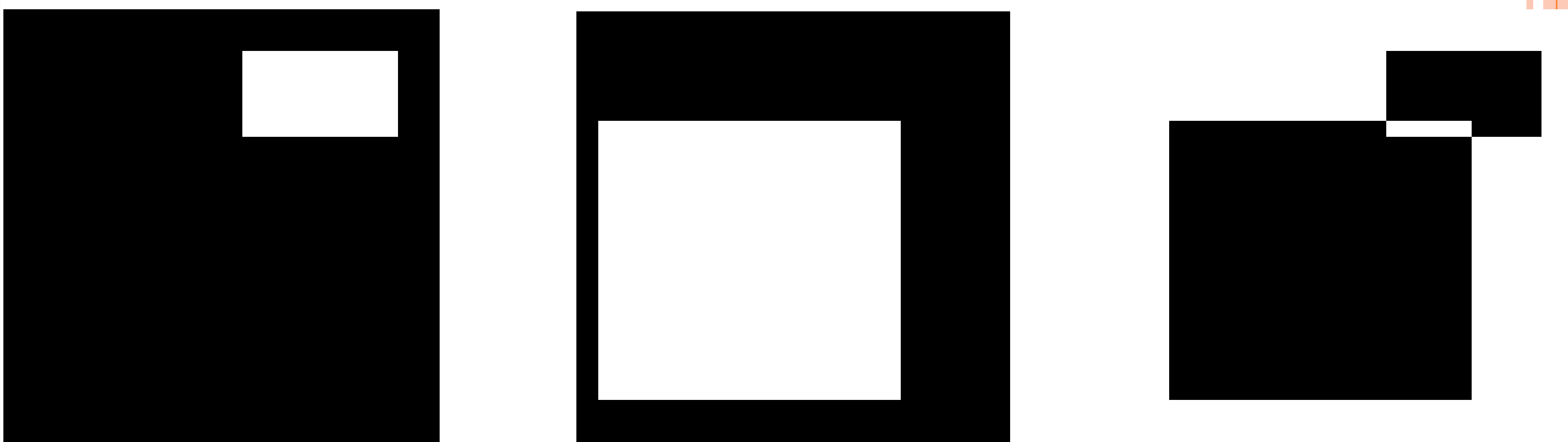
B



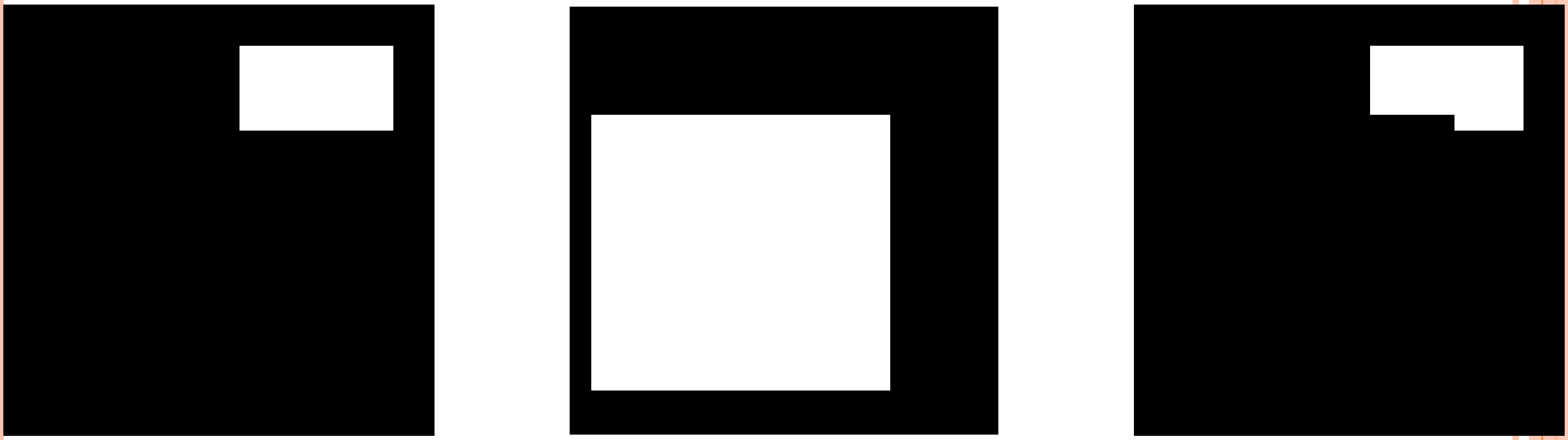
AxorB

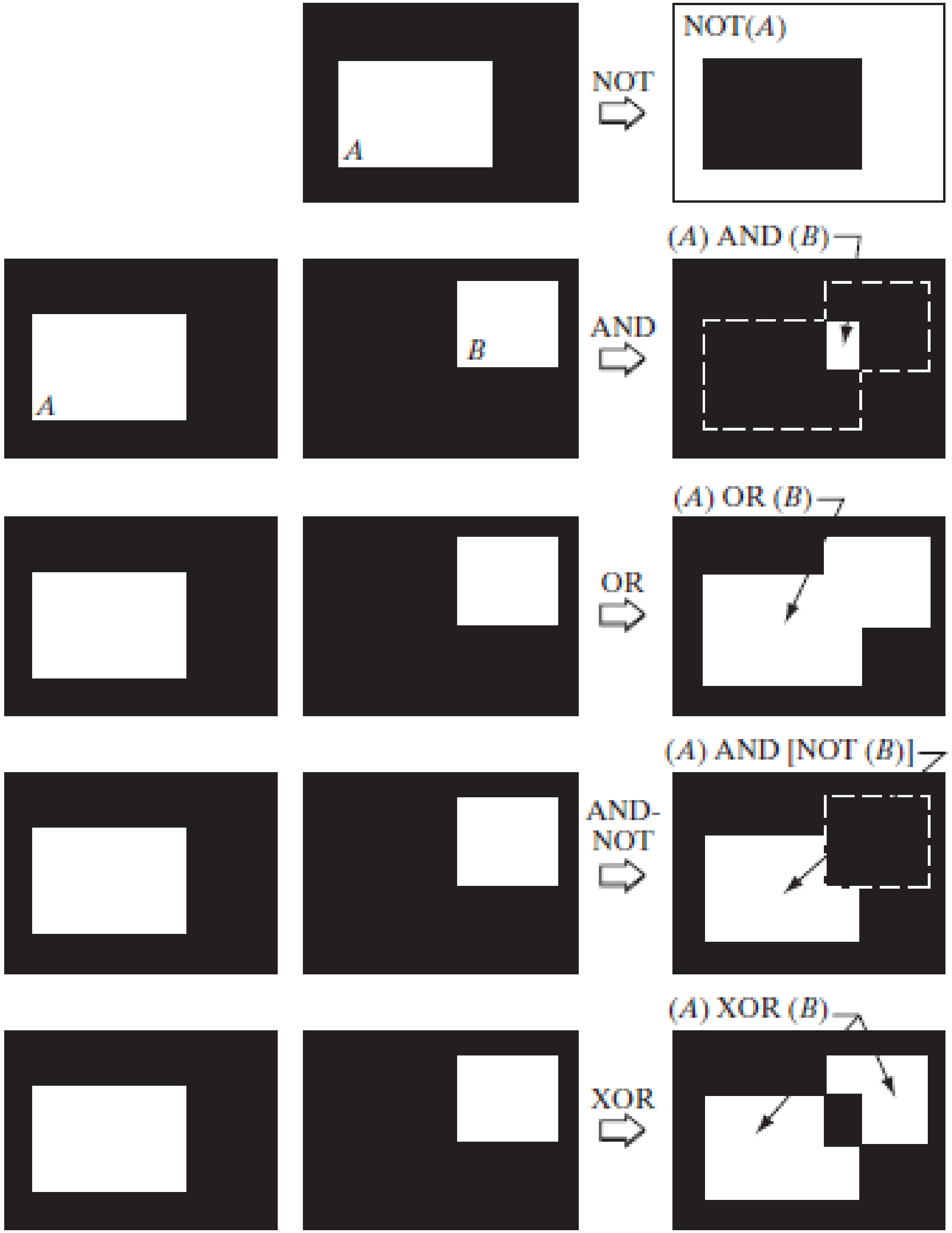


# XNOR



# A AND (NOT B)





A
NOT
NOT(A)
A
B
AND
(A) AND (B)
OR
(A) OR (B)
AND-
NOT
(A) AND [NOT (B)]
XOR
(A) XOR (B)

# Thank you