Basic definitions

What is the difference between *Image processing* and *Computer Vision?*

There are many views – answer depends on who you ask

- Ans 1: None
- Ans 2: IP deals with 2D and CV with 3D-2D relationship
- Ans 3: IP's focus is on manipulating an image while CV's focus is on understanding given images
  - interpret, recognise, decide etc
- Ans 4: CV =**Artificial Intelligence married to IP**!

Syllabus

- Spatial domain proc. - Point and neighbourhood processing, histogram processing
- Image transforms – Fourier, Wavelets
- Colour image processing
- Morphological processing
- Image representation – Pyramids, Scale-space, Multi-resolution
- Image segmentation
- Feature detection and representation – corners, edges, blobs, SIFT, SURF

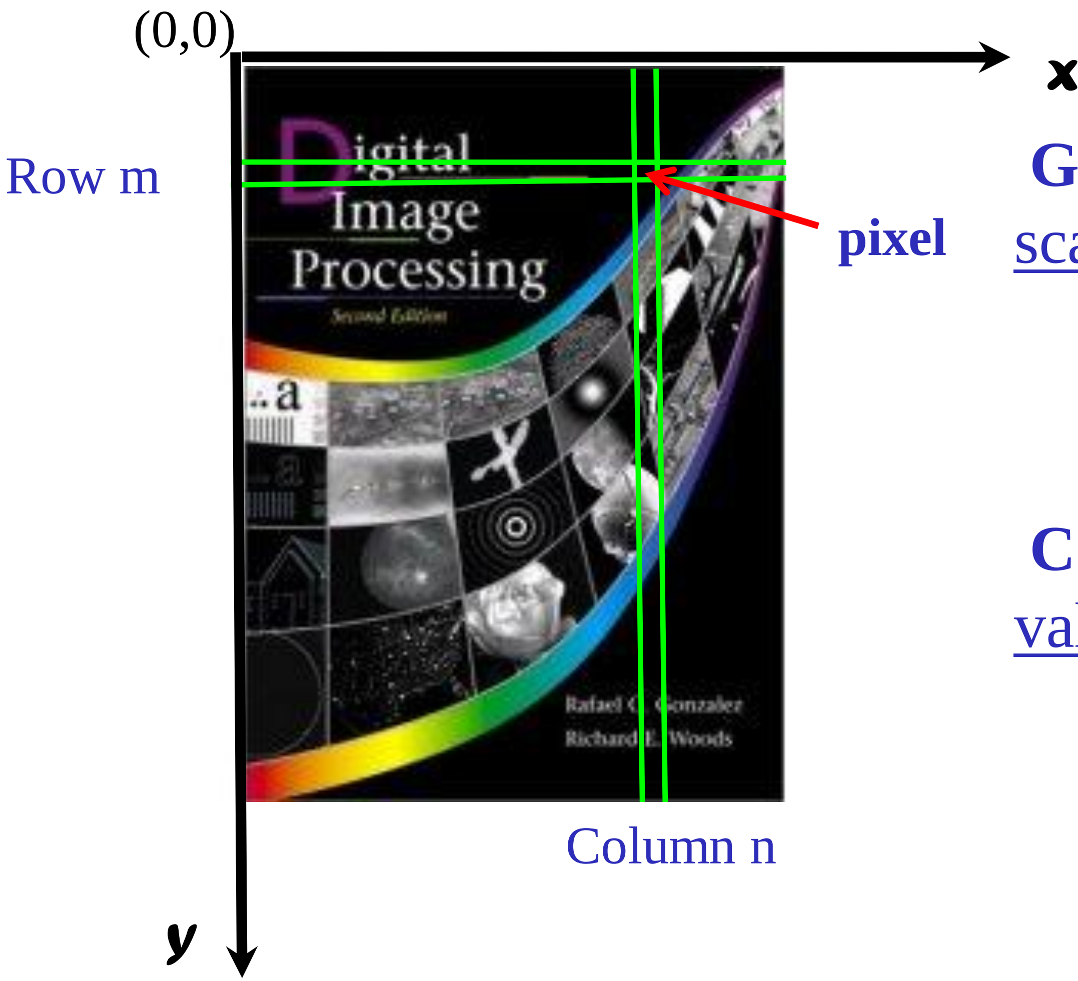
Common types of images

- Binary

- Grey scale

- Colour

Digital Images – basic element



Greyscale image: Image is a <u>scalar valued</u> function

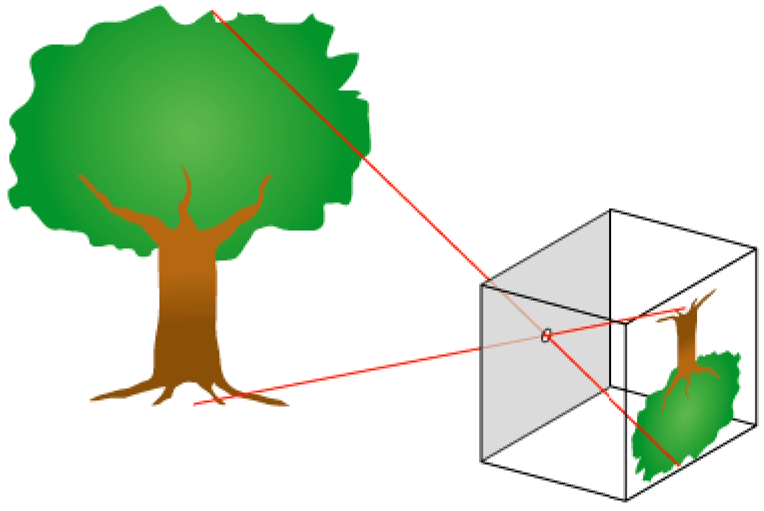
- Pixel value ∈ [0, 255] for 8-bit image

Colour image: Image is a <u>vector valued</u> function

- vector has 3 elements to represent the colour
- each element is usually allocated 8 bits

I[m,n] of size MxN has M rows and N columns and hence MN pixels

Image Acquisition

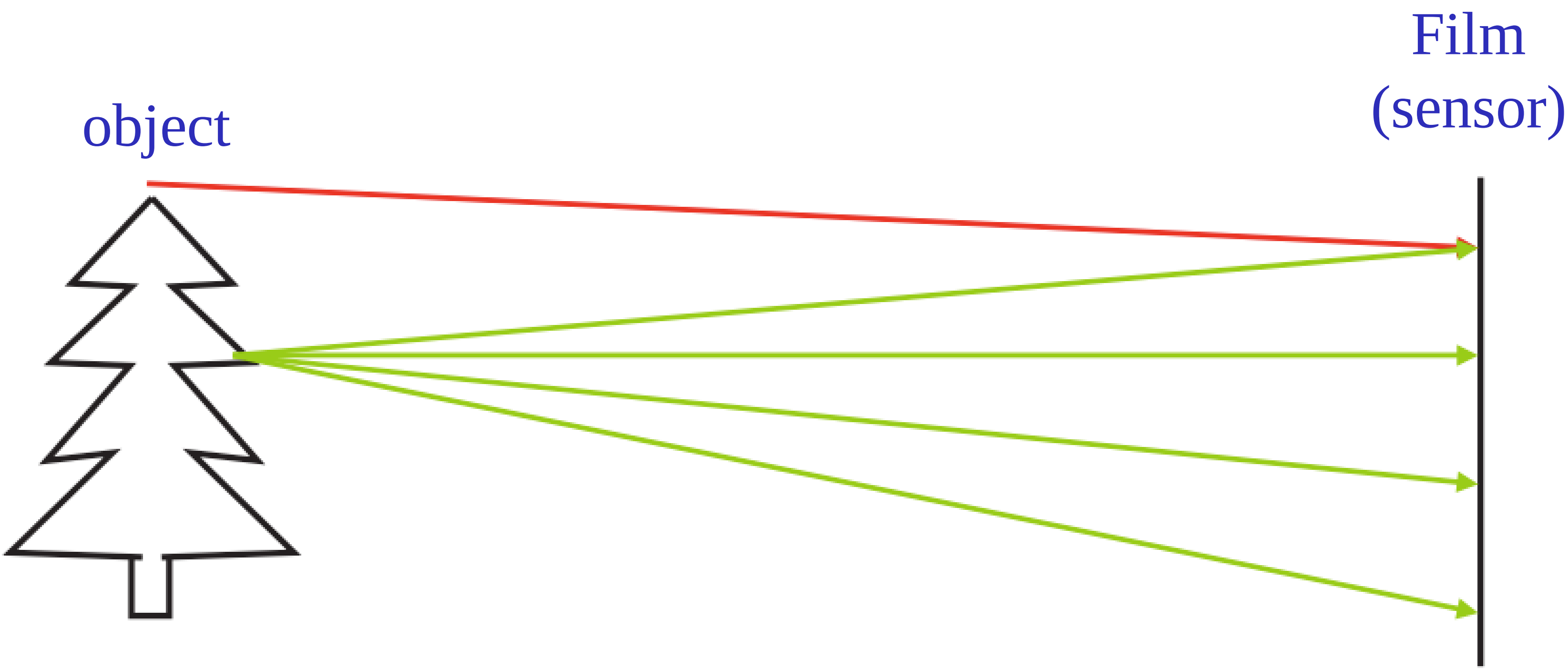


a) Image properties depend on:

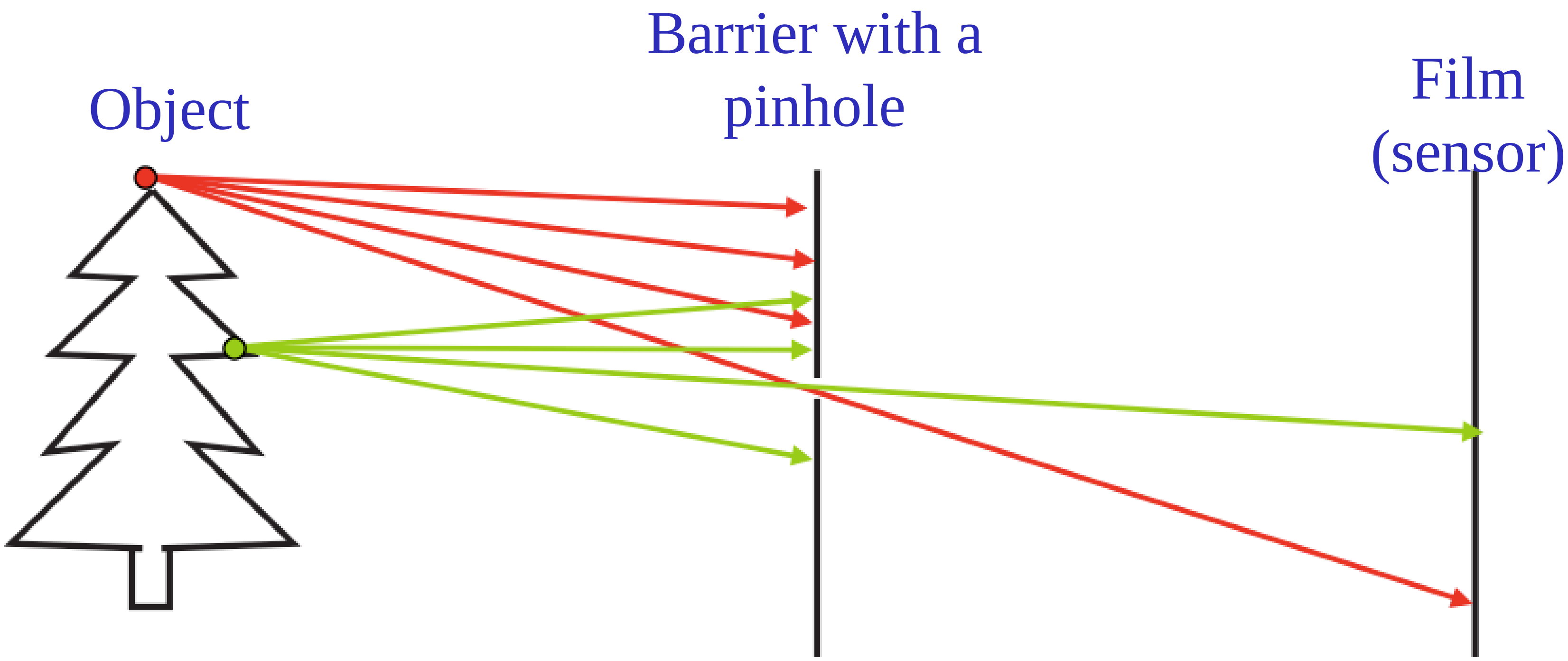
- Camera distance, viewpoint, motion
- Camera intrinsic parameters (e.g., lens aberration)
- Number of cameras
- Ambient illumination

b) Visual properties of the 3D world captured

Working of a camera



Difficult to interpret the image when the film is <u>directly</u> in front of an object

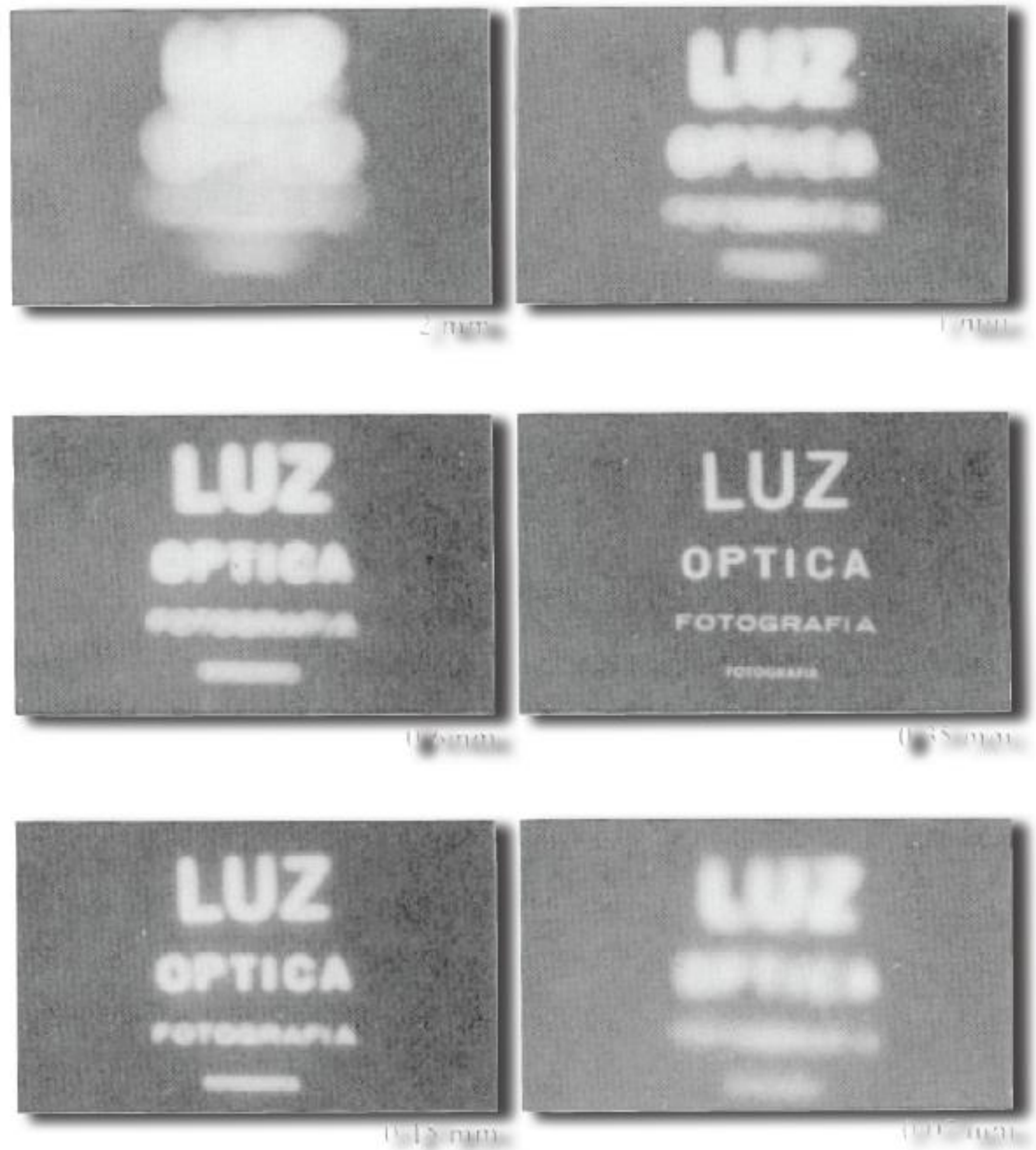


- The barrier blocks off most of the rays
- This reduces blurring
- Aperture = Opening of the pinhole

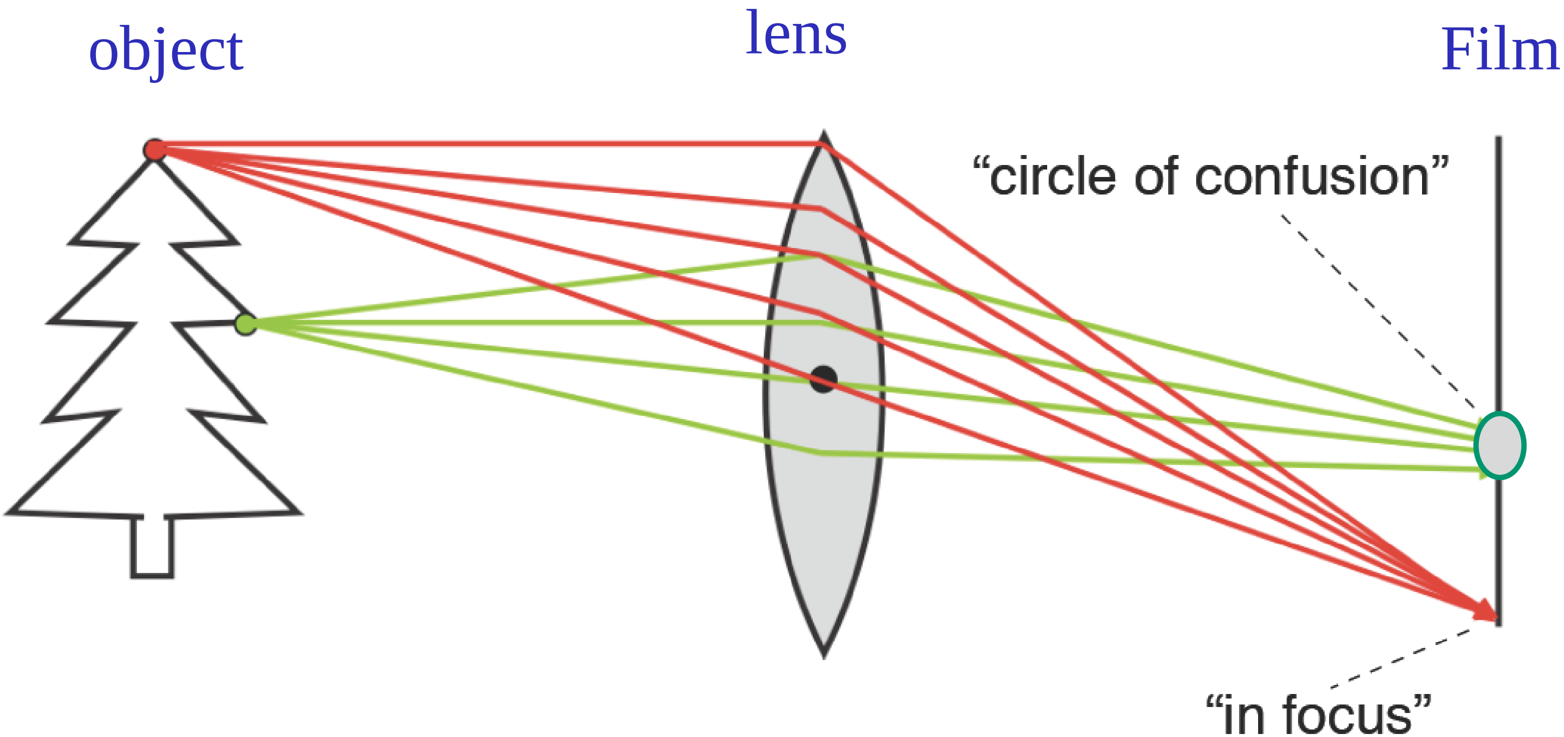
pinhole too big:
bright and blurred

pinhole right size:
dark and crisp

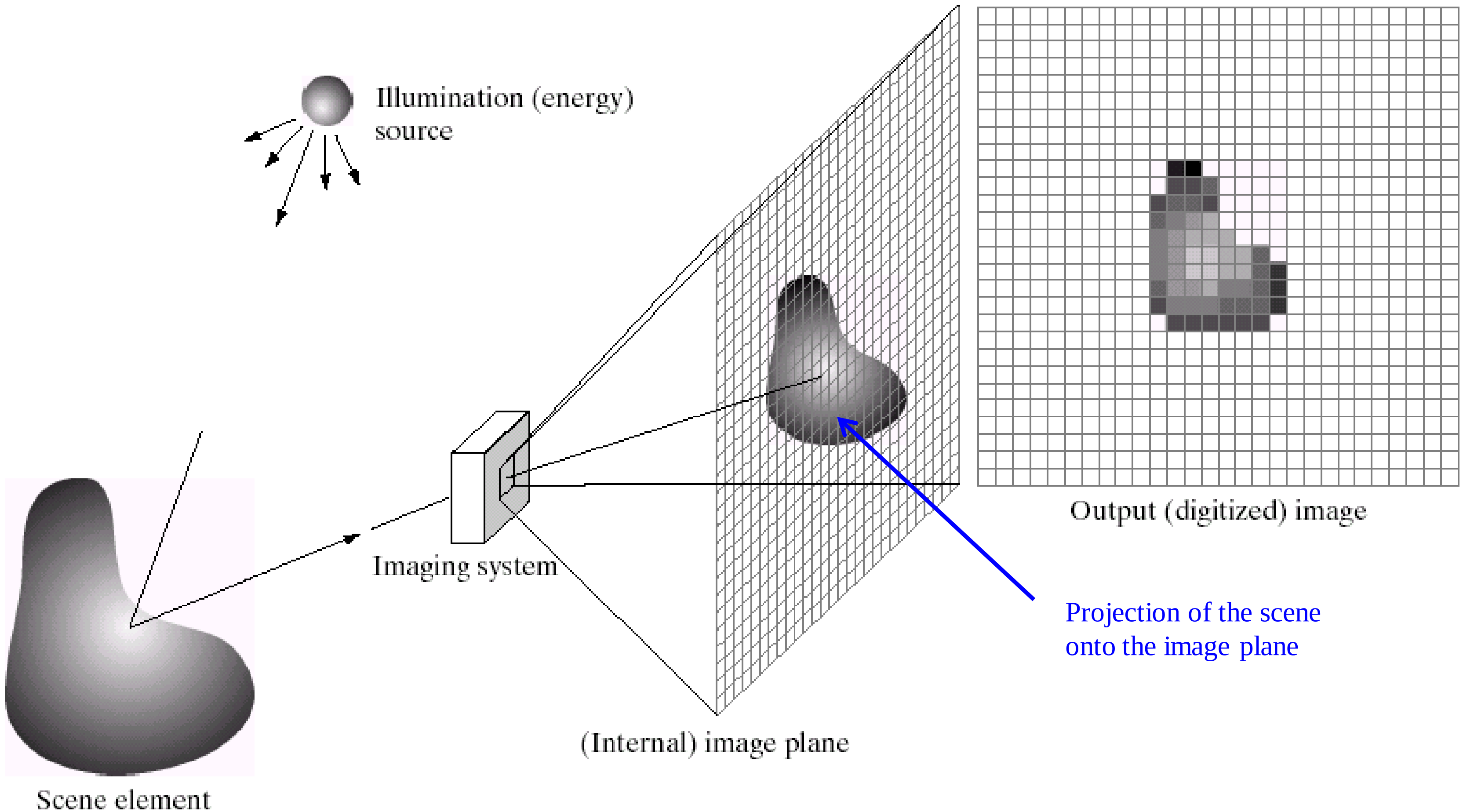
pinhole too small:
dark and diffraction blur



Adding lens ..



The lens focuses light onto the film



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

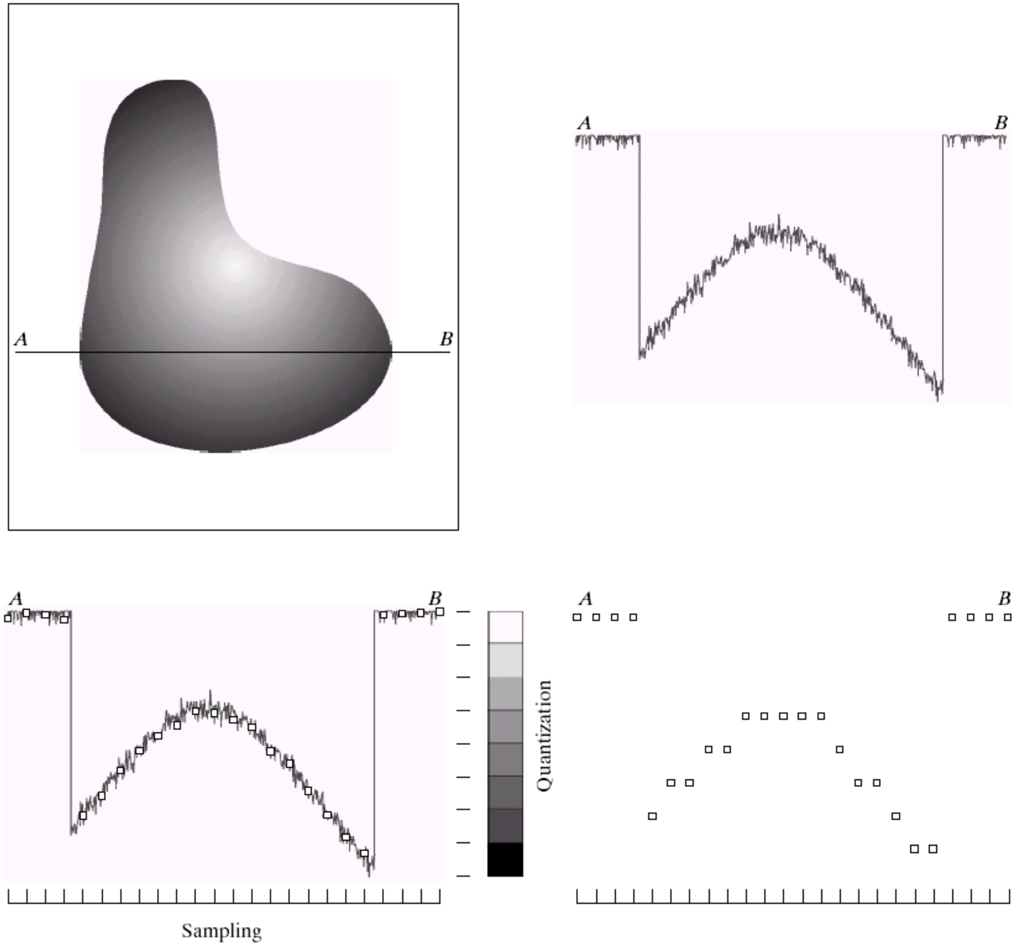
$$0 < i(x,y) < \infty; 0 < r(x,y) < 1$$

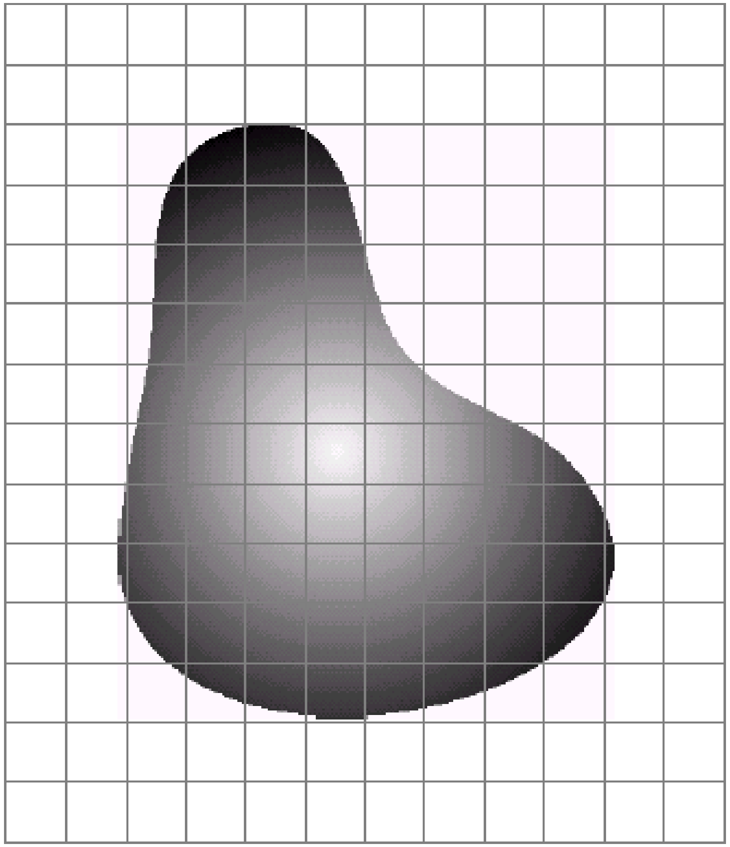
i : incident illumination

r : illumination reflected by the objects

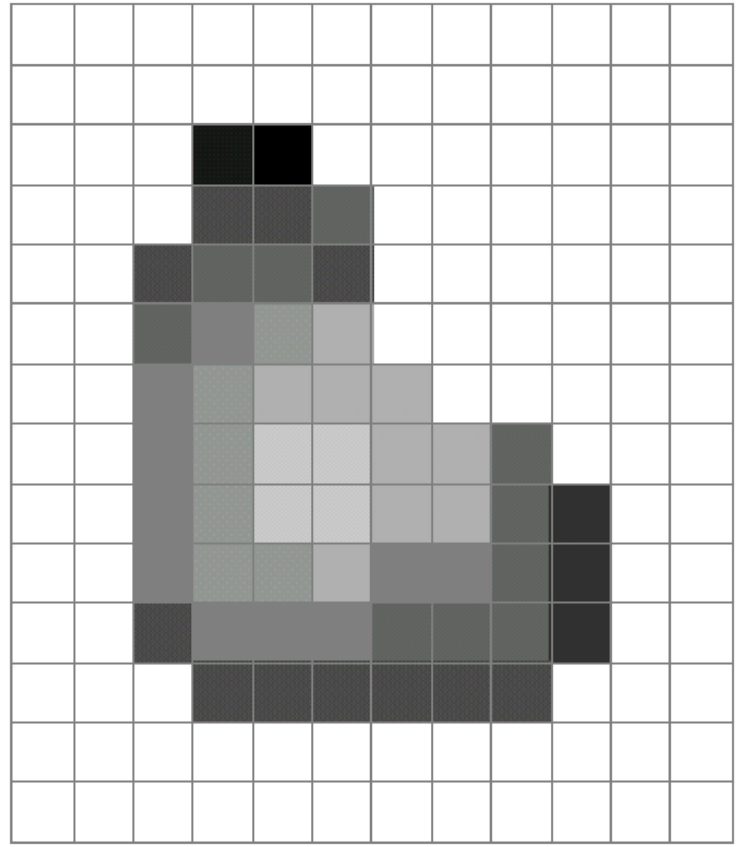
Sampling and Quantization

- *f*(*x*,*y*) is a continuous image w.r.to
  - the spatial variables *x*, *y*
  - amplitude *f*
- Digital conversion needs
  - ***Sampling***: discretising the spatial variables
  - ***Quantisation***: discretising the (intensity) amplitude





Continuous image projected on image grid



Result of sampling and quantization

The quality of a digital image primarily depends on

1. The <u>number of samples</u> used for the spatial variables
2. The <u>number of discrete gray levels</u> used in quantization
3. Image content

What does HD format focus on? 1, 2 or both?

Digital Image Repn. - alternatives

- As a vector I(x)
  - Using row or column ordering of a matrix
  - Used for computational efficiency
- Positional indexing
  - Each pixel has a single index address

Image resolution

$$f(x,y) \xrightarrow{\text{Sample and quantise}} I[m,n]$$

- **Spatial resoultion**: (x,y) → [m,n]
  - Discretising the (spatial) domain of the function by choosing a sampling rate
  - Size of the image MxN = mega/giga pixels
- **Pixel or radiometric**: f → I
  - Discretising the range of the function by selecting the quantisation levels
  - Bit-depth of the image *b* bits/pixel

Insufficient image resolution -effects

$$f(x,y) \rightarrow I[m,n]$$

- Too few quantisation levels for *I* (bits)
  - *False contouring*
- Undersampling the space (x,y)
  - *Ringing* around edges due to *aliasing*
  - *Blockiness/checkerboard effect* with physical size of the pixel becoming larger

What is the ideal resolution?

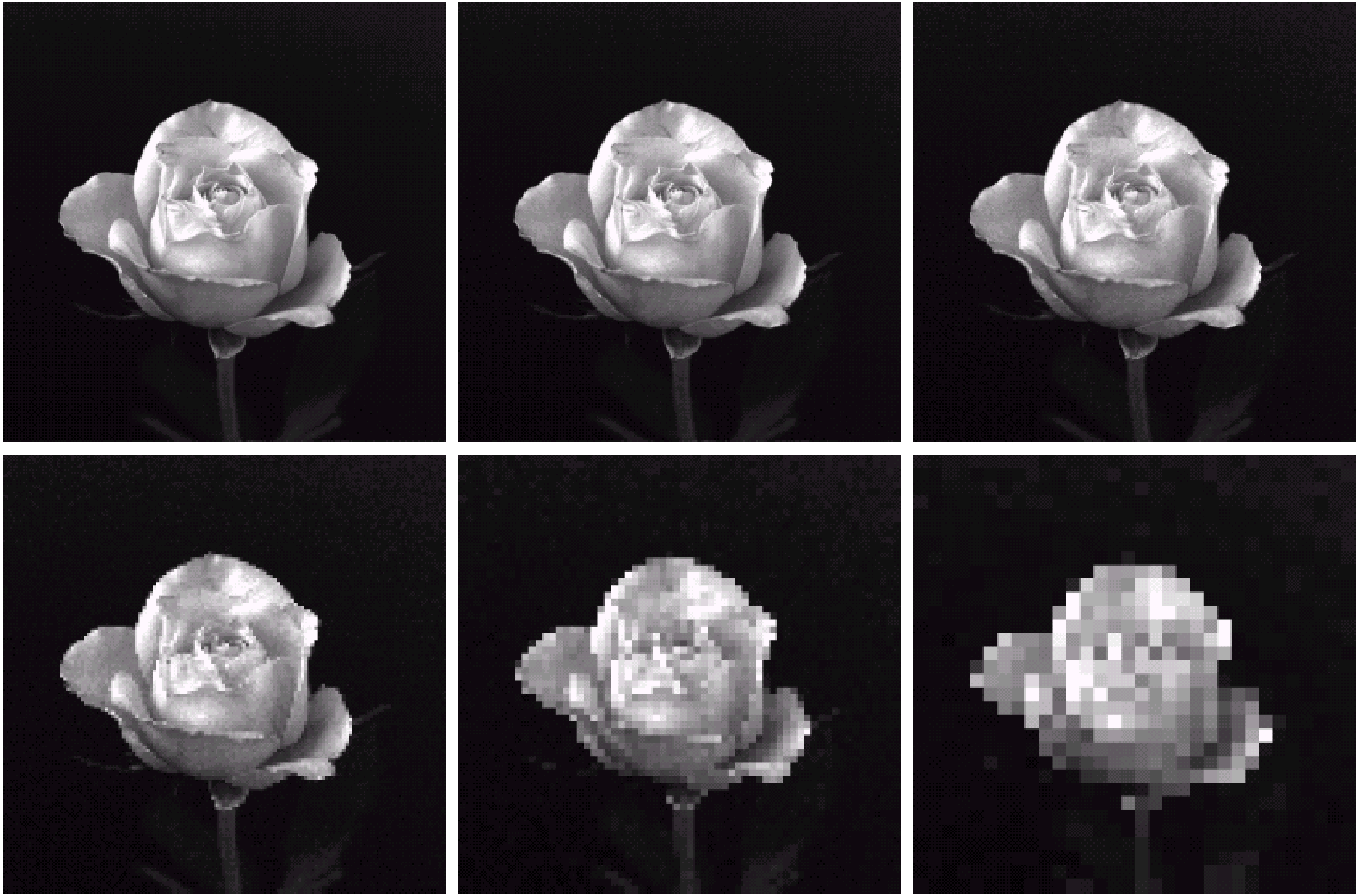
- Depends on scene being imaged

In general:

Globally, slow variations are visually important

Locally, fast variations are visually important

1024 X 1024 512 X 512 256 X 256



128 X 128 64 X 64 32 X 32

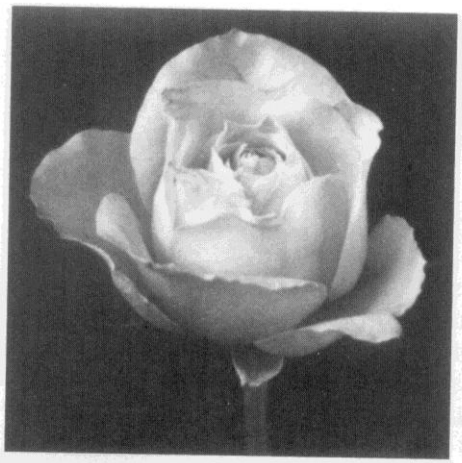
Spatial and Gray-level resolution

- Resolution
  - Spatial: the smallest discernible detail in an image *(sampling)*
  - Gray-level: the smallest discernible change in gray

Effect of reducing spatial resolution (fix number of gray-levels- 256)

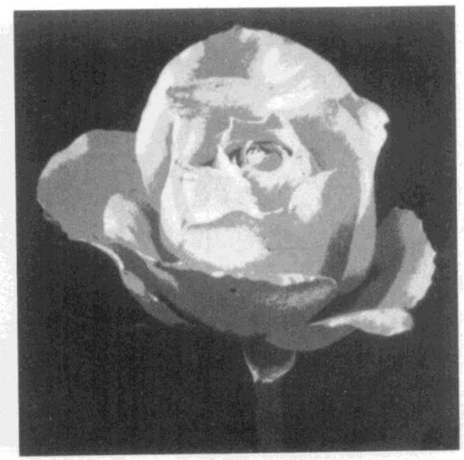
Bit/radiometric resolution – effect of pixel depth

8 bits



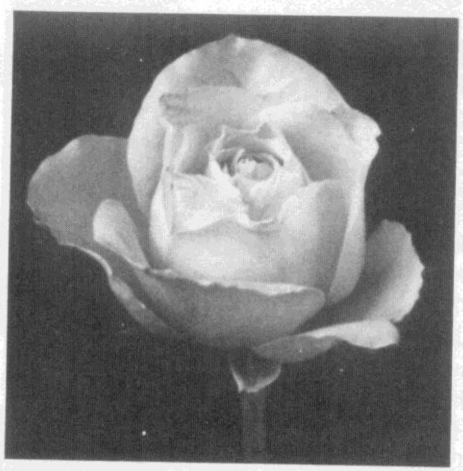
(a)

3 bits



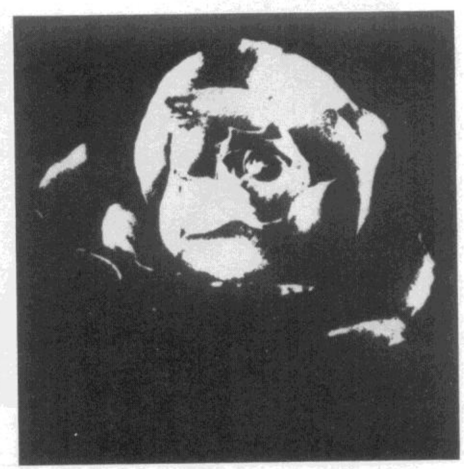
(f)

6 bits



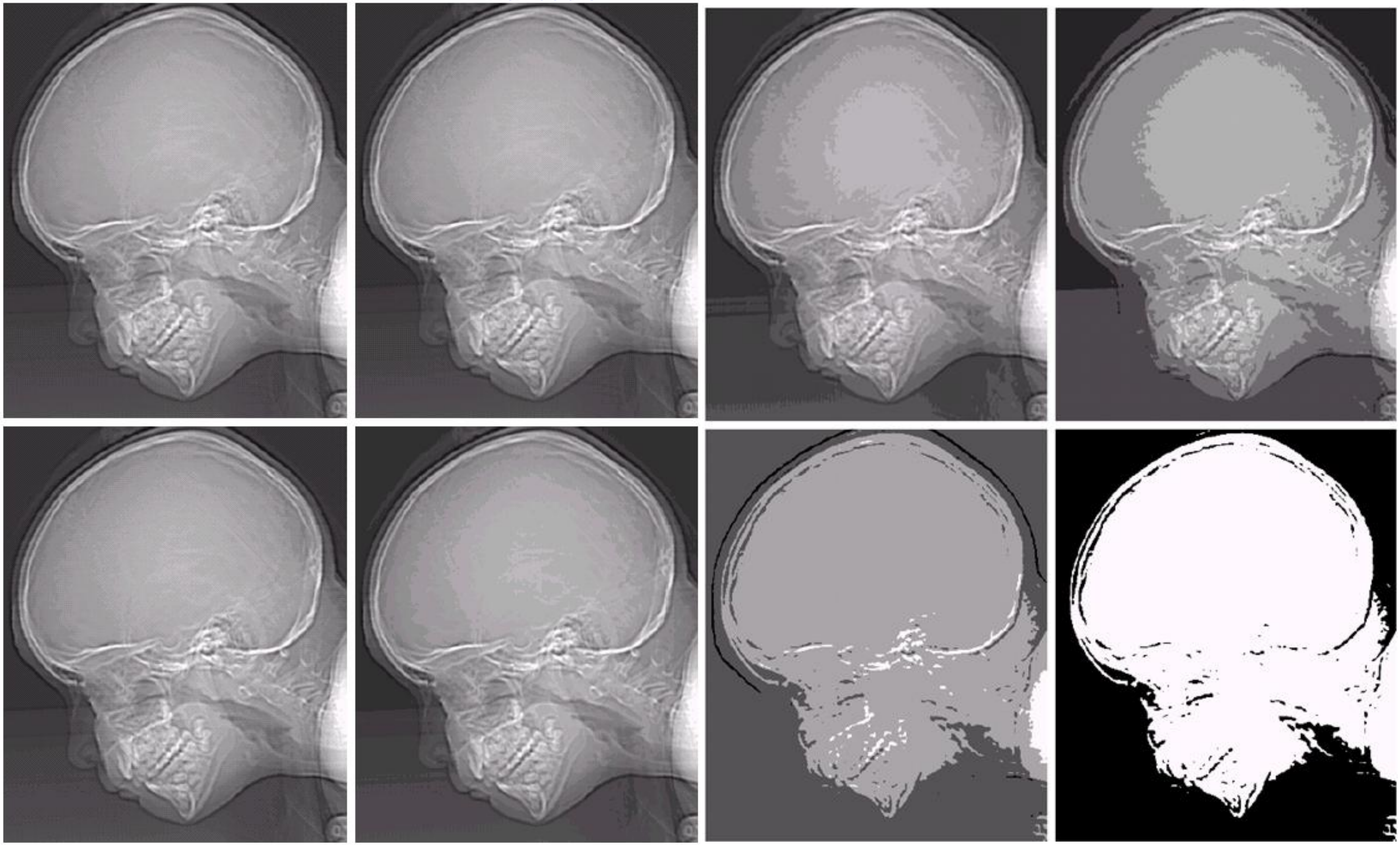
(c)

1 bit



Effect of varying the number of gray level in a digital image

256 levels *128 levels* *16 levels* *8 levels*



64 levels *32 levels* *4 levels* *2 levels*

Digital Image Storage

- **Image**: M x N x bits/pixel
  - Ex. For grey scale, medium resolution images

    M = 256 = N and 8 bits /pixel = 1 byte

    Storage requirement (raw image) ~ 65 Kbytes

  - Ex. For color images = MN x 3 colour planes

    M = N = 256 and bits /pixel = 3 bytes

    (1 byte each for R,G and B)

    Storage requirement ~ 196 Kbytes

- **Video**: MN bytes/frame x frame rate x time
  - Ex. For a 10 minute video

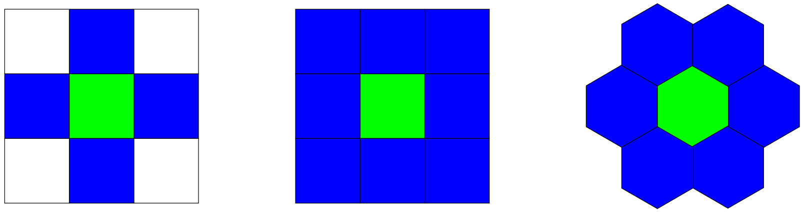
    65Kb/frame x 20 frames/sec x 10 x 60 sec = 780 Mb

Pixel Neighbourhoods

A *neighbourhood* is a region adjacent to a given pixel
- set of pixels that share an edge or corner

- 4-neighbours (square)
  - 2 horiz. plus 2 vertical
- 8-neighbours (square)
  - 4 diag. plus 4 neighbours
- 6-neighbours
  - Used in hexagonal tessellations

8-neighbours



4-neighbours
D-neighbours

6-neighbours

Neighbours

Given a pixel **p** at [m,n], its neighbours are:

- 4-Neighbours of **p**: $N_4(p)$ = {2 vert. and 2 horiz. pixels)}
  - Pixels at [m-1,n], [m+1,n], [m,n-1], [m,n+1]
- D-Neighbours of **p:** $N_D(p)$ = {4 diagonal pixels}
  - Pixels at [m+1,n+1], [m+1,n-1], [m-1,n+1], [m-1,n-1]
- 8-Neighbours of **p**: $N_8(\mathbf{p}) = N_D(\mathbf{p}) \cup N_4(\mathbf{p})$

Distance Measures

Given pixels p and q at (x_p,y_p) and (x_q,y_q)

- Euclidean distance L_2 *norm ('as the crow flies')*

  $[(x_p-x_q)^2 + (y_p-y_q)^2]^{1/2}$

- City block distance L_1 *norm*

  $|(x_p-x_q)| + |(y_p-y_q)|$

- Chebychev or chessboard distance L_∞ *norm*

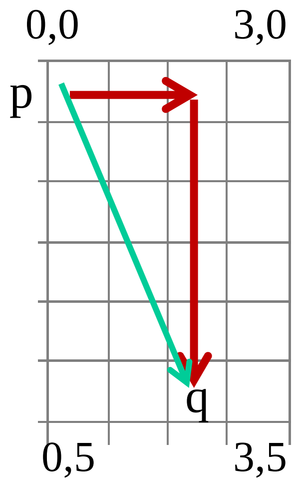
  $\max\{|(x_p-x_q)|, |(y_p-y_q)|\}$

Example

Euclidean distance:

City block distance:

Chess board distance:



What is the locus of equidistant points in each of these cases?

Connectivity

Connectivity is a measure of the **adjacency** between pixels of similar values

Pixels **p** is 4 or 8-connected to **q** if they have <u>similar</u> values AND $\mathbf{q} \in N_{4 \text{ or } 8}(\mathbf{p})$

Mixed connectivity

$\mathbf{q} \in N_4(\mathbf{p})$ <u>OR</u>

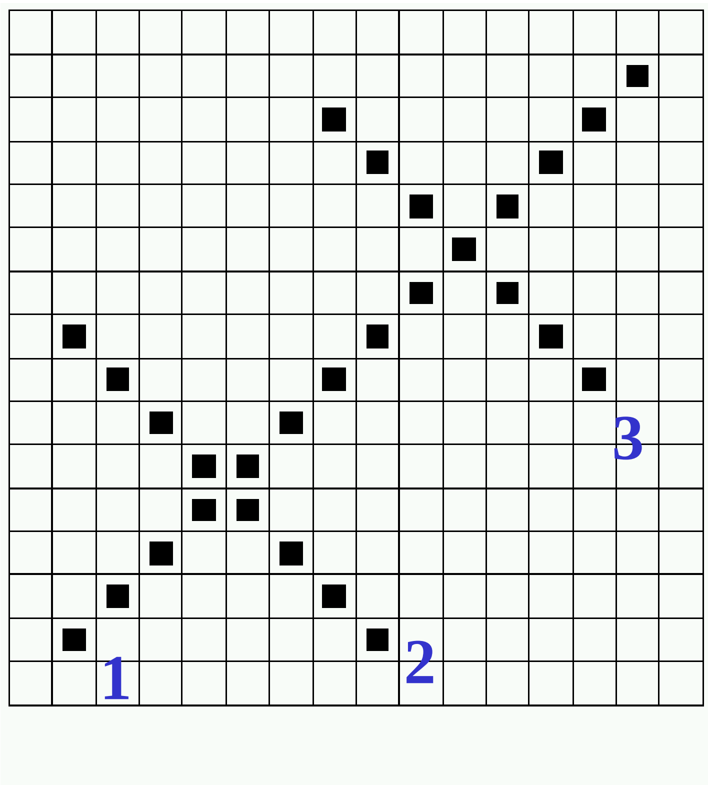
$\mathbf{q} \in N_D(\mathbf{p})$ AND $N_4(\mathbf{p}) \cap N_4(\mathbf{q}) = \Phi$

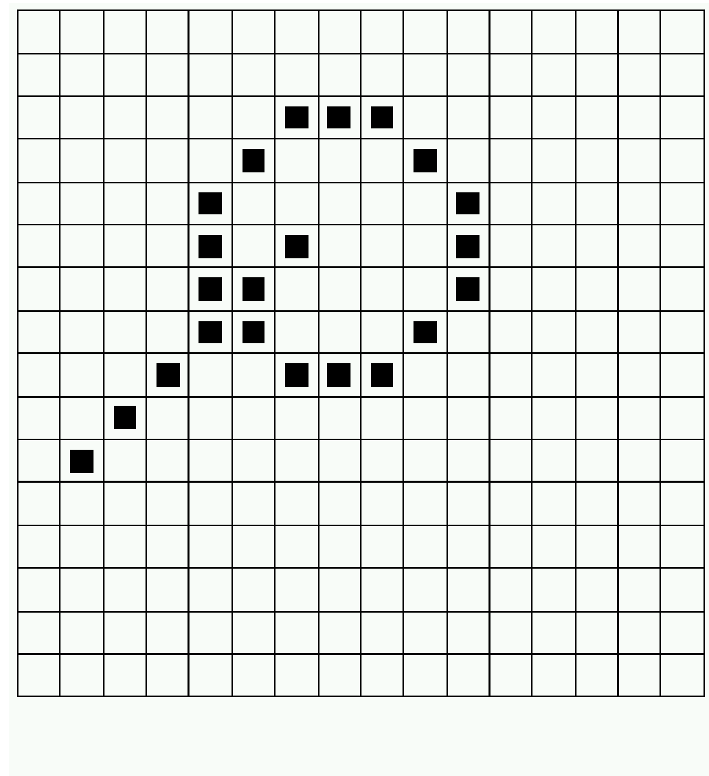
no loops can be present as is possible with N_8 .

Regions

- A region is a contiguous set
  - neighbouring pixels should “touch”each other
- Square grid poses problems to defining regions
- Hexagonal grid does not pose problems

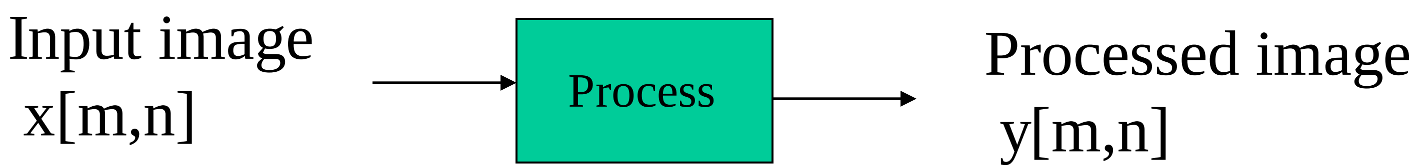
Examples



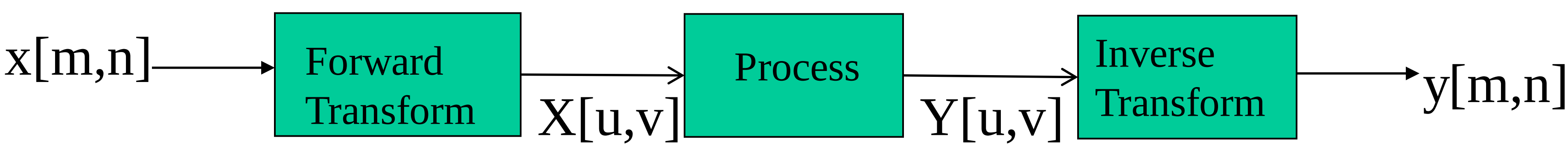


Processing an Image

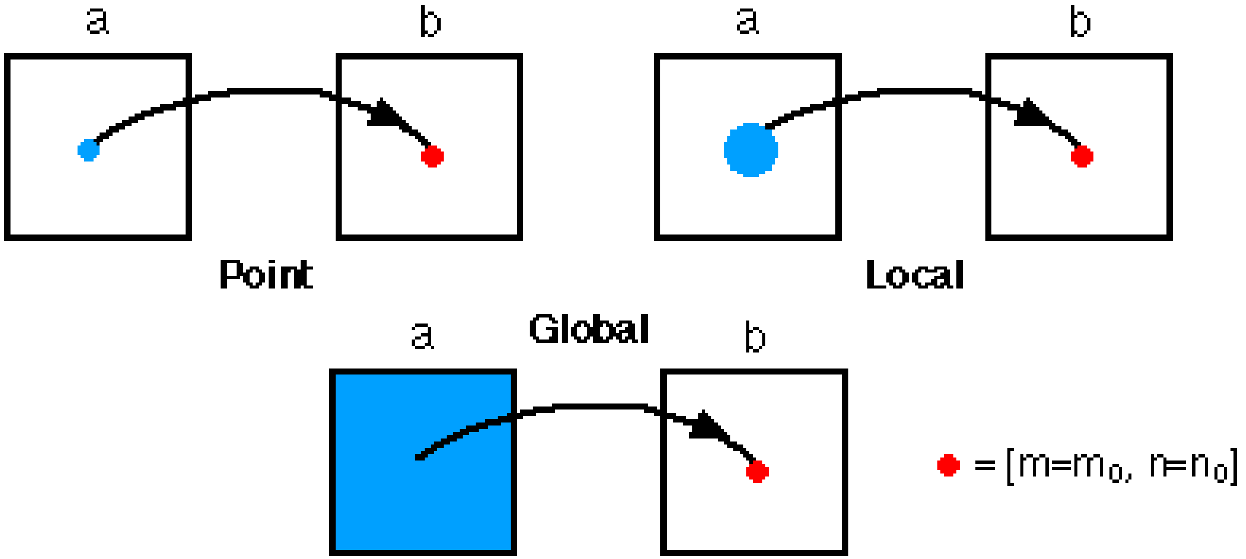
Spatial domain processing



Transform (frequency) domain processing



Spatial domain processing



$b[m_0,n_0] = f(a[m_0,n_0])$

$b[m_0,n_0] = f(a[N(m_0,n_0)])$

$b[m_0,n_0] = f(a[m,n])$

a: input image

b: processed image

Techniques for spatial domain processing

- Local or Point Operations
  - Each pixel's value is changed according to its original value
- Neighbourhood operations
  - Each pixel's value is changed according to the values of its neighbouring pixels
  - Linear or nonlinear
- Global operations
  - Each pixel's value is changed according to some global property of the entire input image

- Challenges: pixel addressing load, computational load, memory requirements

Type of manipulations on $I[m,n]$

On the **dependent variable (I)**

- Arithmetic (last class)
  - {+,-, *, / }
- Logical (binary IP)
  - {AND, OR, NOT}
- Set (binary morphological proc.)
- Statistical
  - {Mean, variance, median, mode etc}
- Convolution, correlation
  - linear vs circular convolution
  - Window/mask based computation

On the **independent variables (m,n)**

- Geometric
  - Affine transformation
- Up/down sampling
  - zoom, pan