Digital Audio, Image and Video

Lecture 2

Dr. Heba Hamdy

Multimedia Elements

- In computers, audio, image and video are stored as files just like other text files.
  - Text
  - For images, these files can have an extension like
    - BMP, JPG, GIF, TIF, PNG, PPM, …
  - For audios, the file extensions include
    - WAV, MP3, …
  - The videos files usually have extensions:
    - AVI, MOV, …
  - Animation & Interaction

Multimedia Elements: Text

- For effective visual communication in a presentation, it is best to use large fonts **(i.e., 18 to 36 points),** and **no more than 6 to 8 lines per screen.**
- Sans-serif fonts work better than serif fonts.

Multimedia Elements: **Graphics (2-D and 3-D):**

Vector graphics

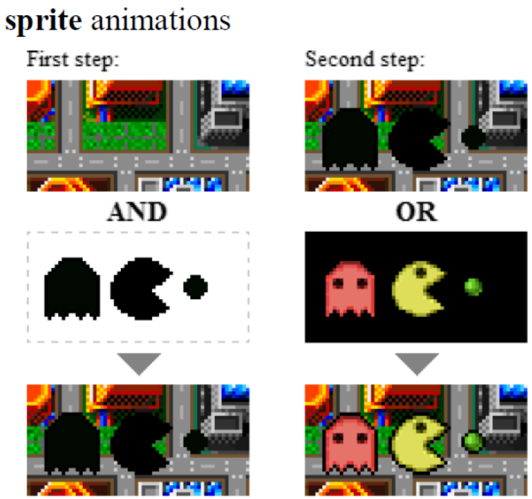
- Vector graphics is the use of geometrical primitives such as **points**, **lines**, **curves**, and **shapes** or **polygons** all of which are based on mathematical expressions to represent images in computer graphics.
- Company logo and brand graphics should be created as a vector.

Multimedia Elements :**Raster images**

- **Bitmap(or raster)** images are stored as a series of tiny dots called pixels. Each pixel is actually a very small square that is assigned a color.
- JPEGs, GIFs and PNGs are common raster image types.

Multimedia Elements : **Animation & Interaction**

- Animation is the rapid
- is play of a sequence of images of 2-D or 3-D art work or model positions in order to create an illusion of movement



Graphics Styles:

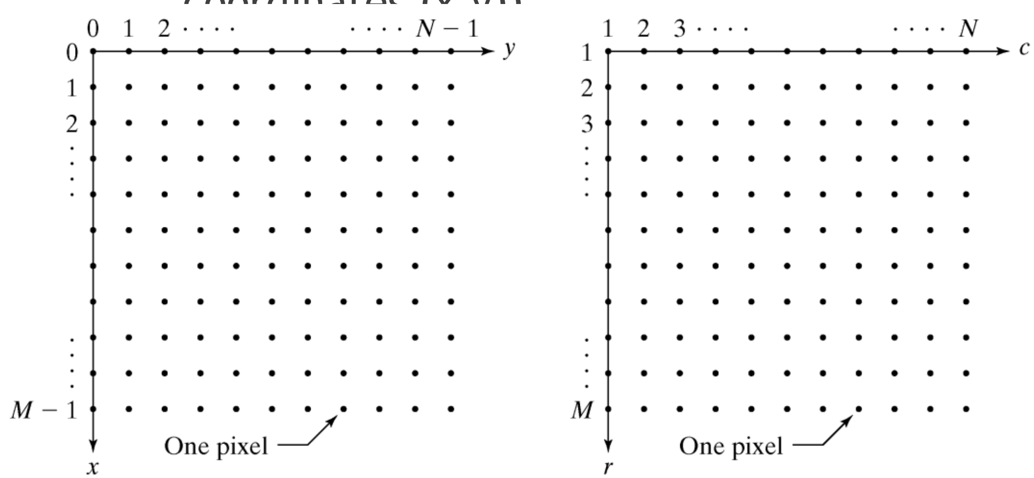
Color principles and guidelines:

- Some color schemes and art styles are best combined with a certain theme or style.
- A general hint is to not use too many colors, as this can be distracting.

Digital Image Representation

2D function f(x,y)

- x,y: (spatial) coordinates
- Amplitude f : intensity (at a certain point with coordinates (x,y))



Digital Image Representation

- A $M \times N$ digital image is often represented as a matrix as

$$\begin{array}{cl} \mathbf{A} = f(x,y) = \begin{bmatrix} f(0,0) = a_{0,0} & f(0,1) = a_{0,1} & \cdots & f(0,N-1) = a_{0,N-1} \\ f(1,0) = a_{1,0} & f(1,1) = a_{1,1} & \cdots & f(1,N-1) = a_{1,N-1} \\ \cdots & \cdots & \cdots & \cdots \\ f(M-1,0) = a_{M-1,0} & f(M-1,1) = a_{M-1,1} & \cdots & f(M-1,N-1) = a_{M-1,N-1} \end{bmatrix} & \begin{array}{l} 1^{\text{st}} \text{ raw} \\ 2^{\text{nd}} \text{ raw} \\ \\ M^{\text{th}} \text{ raw} \end{array} \\ \begin{array}{cccc} 1^{\text{st}} \text{ column} & 2^{\text{nd}} \text{ column} & \qquad & N^{\text{th}} \text{ column} \end{array} & \end{array}$$

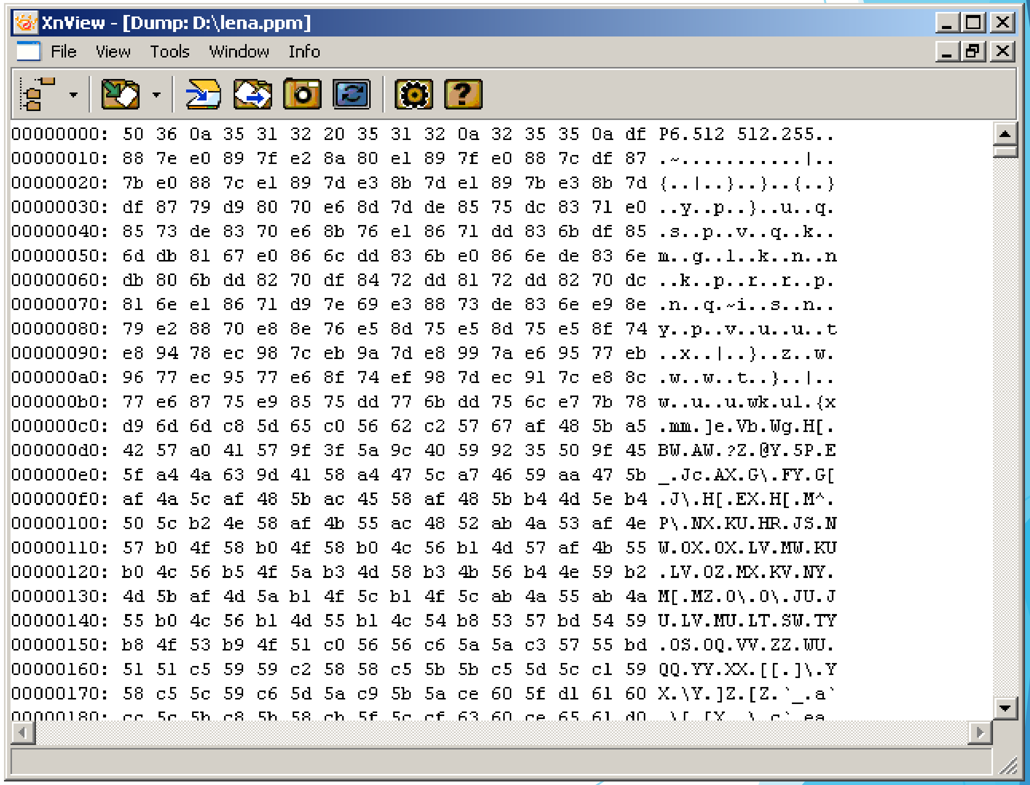
- Each entry is called a picture element “pel” or “pixel”
- In 3D case, each entry is called a volume element “voxel”

Image content values

An image Contains 2 basic parts

1- Image header

2-Data a set of (integer) numbers.



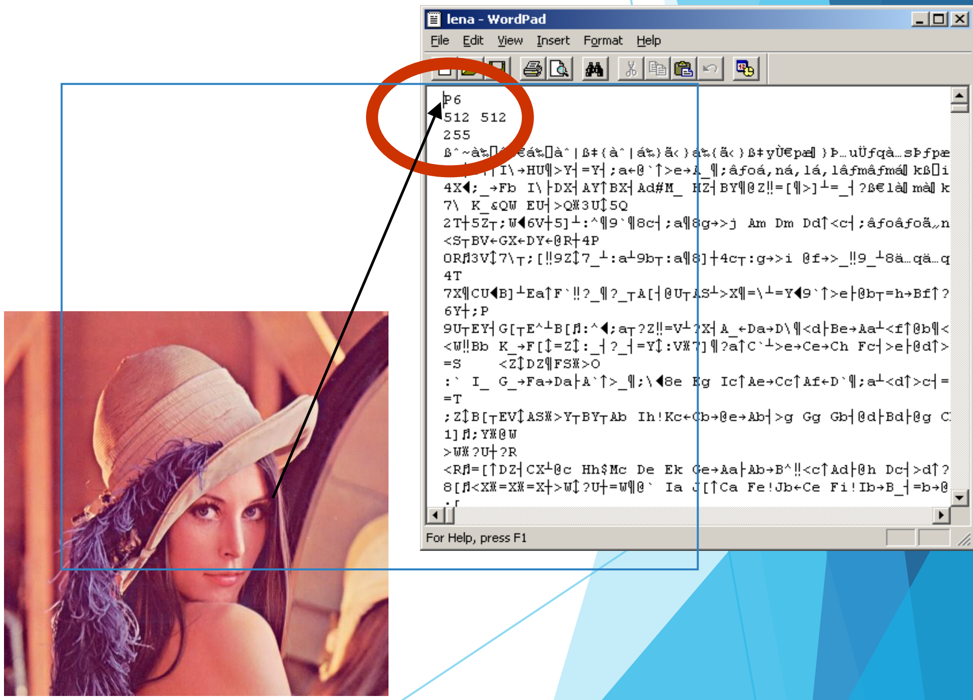
An Digital Image Example

Let's open an image file in its "raw" format:

- The portable pixmap format (PPM),
- The portable graymap format (PGM),
- The portable bitmap format (PBM)

are image file formats

- **P6:** (this is a ppm image)
- **Resolution:** 512x512
- **Depth**: 255 (8bits per pixel in each channel)



Digital Media Capturing

To get a digital image, an audio or a video clip.

- we use some media capturing device such as:
  - Digital camera or a scanner,
  - Digital audio recorder,
  - Digital camcorder (an electronic device that combines a video camera and a video recorder into one unit).
- Capturing Device have to perform the following tasks:

- Sampling:

- convert a continuous media into discrete formats.

-Digitization:

- convert continuous samples into finite number of digital numbers.

-Compression:

- some devices can perform compression process.

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.

Image Acquisition

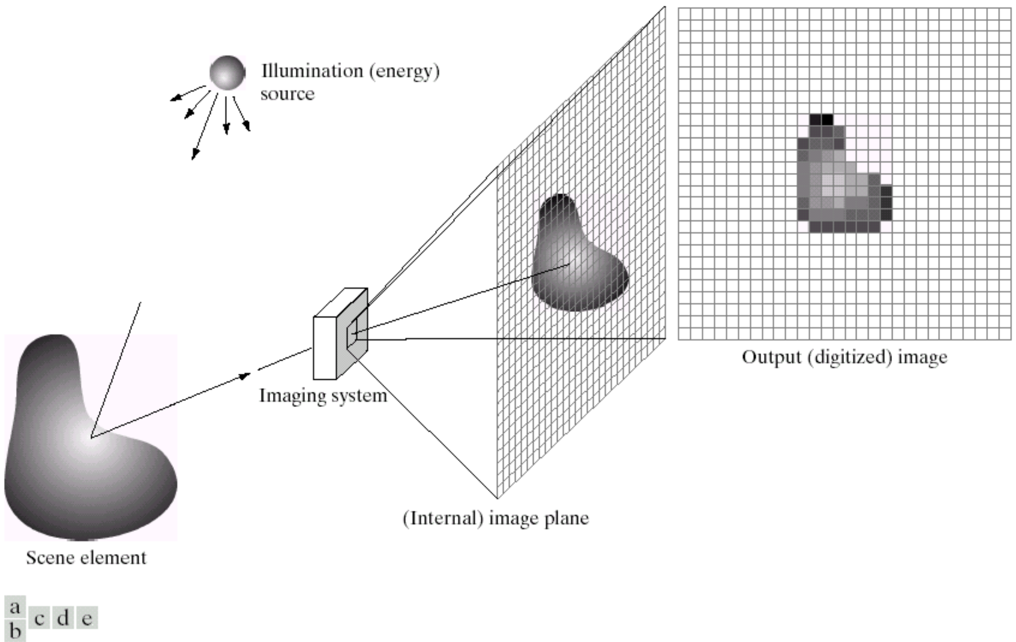


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

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 $0 < r(x,y) < 1$ ıt source

- Determined by the nature of the objects in the scene bounded from total absorption to total reflectance.

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.

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

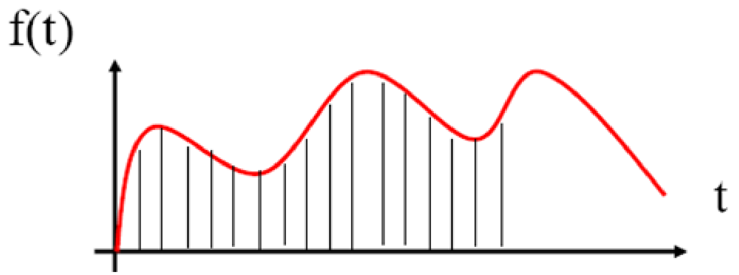
signal has discrete domain and range = digital

Analog & Digital

Sampling.	Quantization
• Digitalizing the coordinate values	• Digitalizing the amplitude values

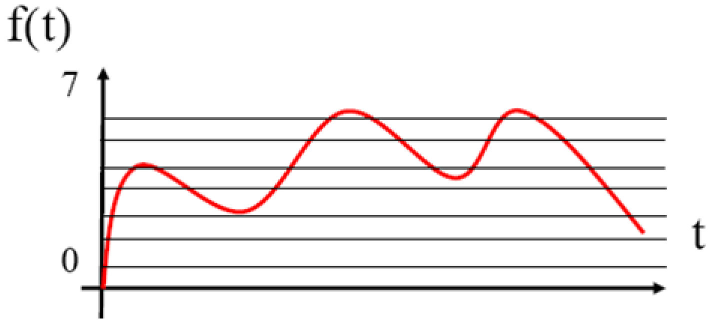
Sampling

- 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.

Generalize Digital Image

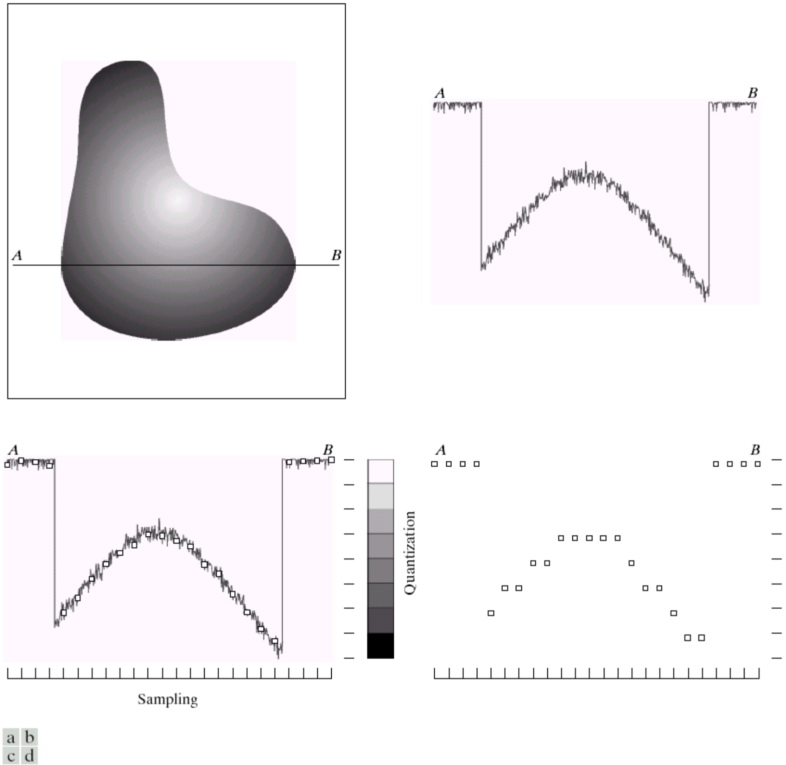
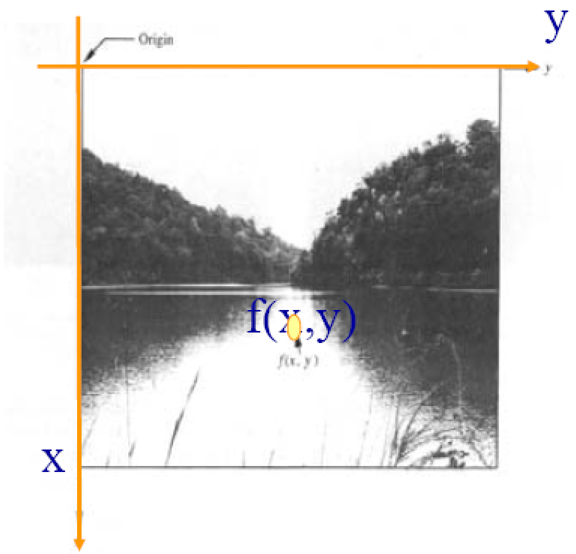


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Digital Image Representation

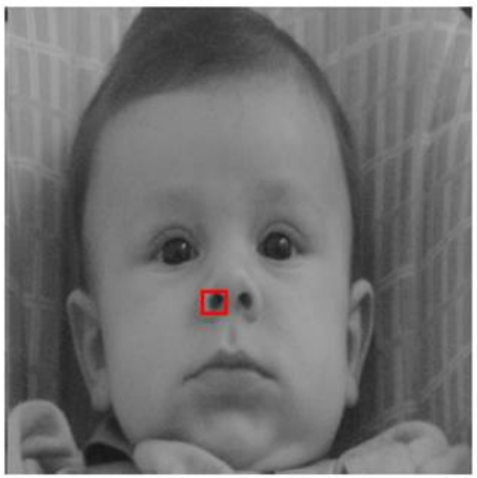


- A digital image is an image *f(x,y)* that has been digitized both in spatial coordinates and brightness.
- the value of f at any point *(x,y)* is proportional to the brightness (or gray level) of the image at that point.

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



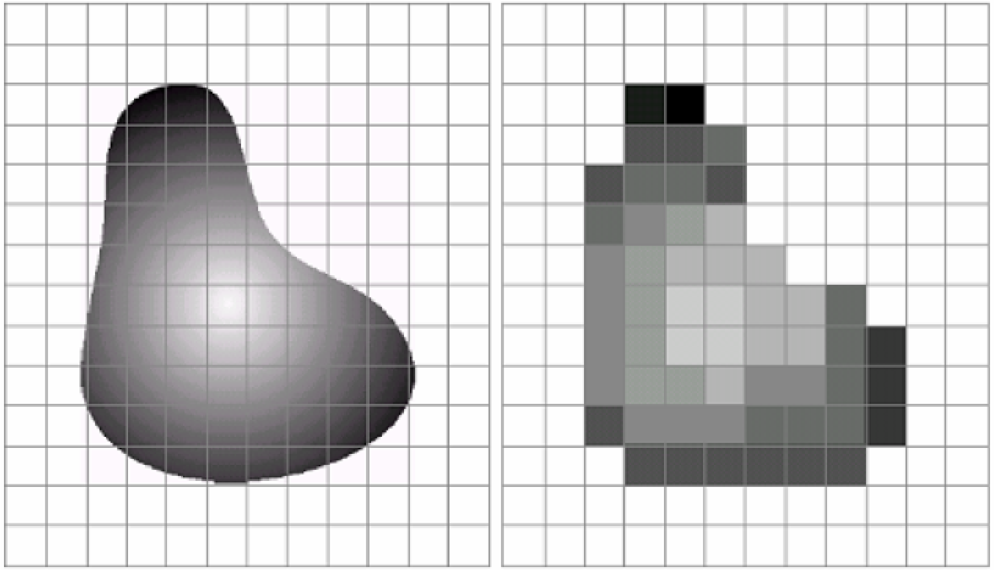
Pixel values in highlighted region

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

CAMERA → DIGITIZER → A set of number in 2D grid

Samples the analog data and digitizes it.

Examples of Digital Images



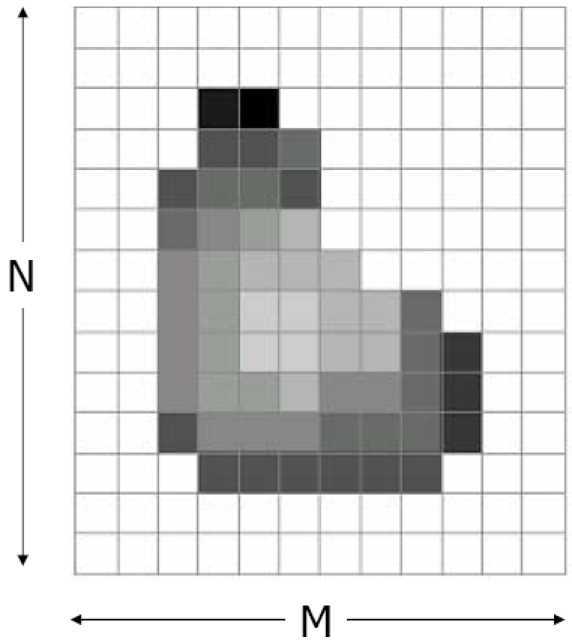
Continuous image projected onto a sensor array

Result of image sampling and quantization

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 $L_{\min} \leq l \leq L_{\max}$
- 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.

Spatial Resolution

- Smallest discernible detail about the scene in an image
- Determined by the sampling rate (number of samples taken per unit distance, ex: dpi)
- Equivalently determined by the physical size of a pixel

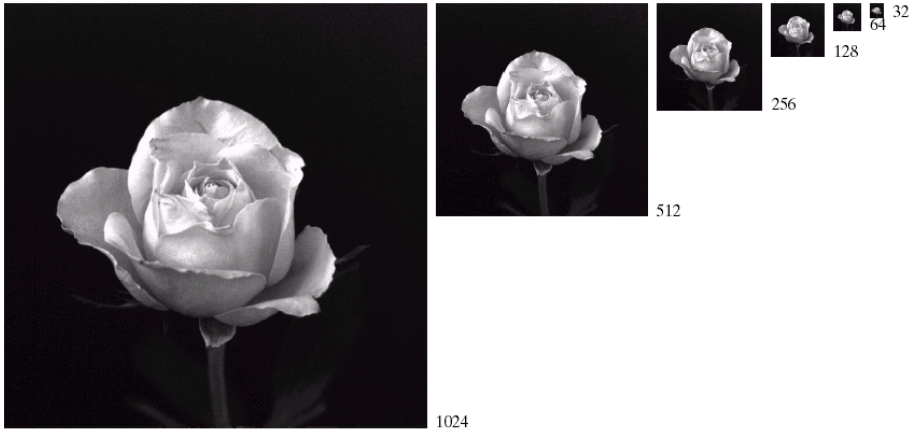


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Spatial Resolution

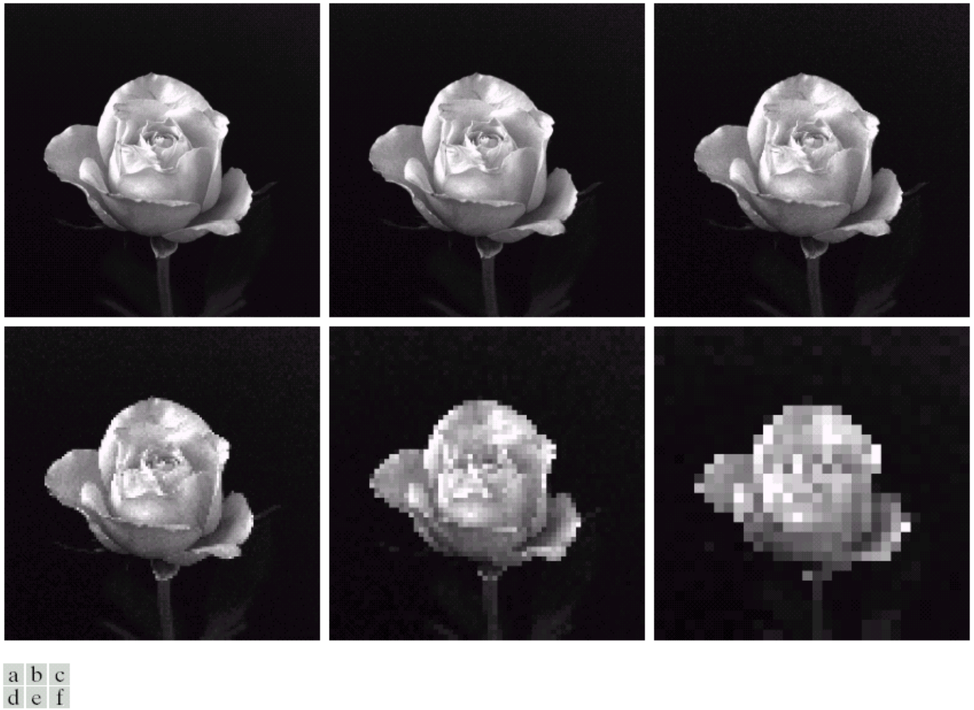
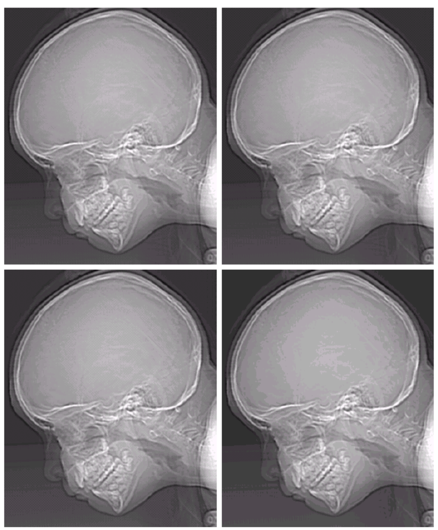


FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

Checkboard Effect

Gray-Level Resolution

- Refers to the smallest discernible change in gray level
- Specified by the number of gray levels used to represent an image

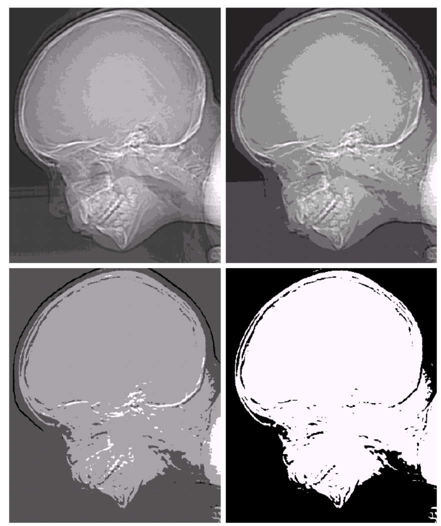


a b
c d

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

e f
g h

FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Gray-Level Resolution

- Both the spatial and gray level resolutions determine the (storage) size of an image (bytes)
  - Ex: scene area: 1km2 , spatial resolution 20 lines/km, gray level resolution: 64

$$\#\ \text{of pixels}: 20 \cdot 20 = 400$$

$$400\,(\text{pixels}) \cdot 6\,(\text{bits/pixel}) = 2400\,(\text{bits}) = 300\ \text{bytes}$$

$$= 0.3\ \text{kb}$$

(without compression, overheads...)

Types of Digital Images

- Grayscale image
  - Usually we use 256 levels for each pixel. Thus we need 8bits to represent each pixel (2^8 == 256)
  - Some images use more bits per pixel, for example MRI (Magnetic resonance imaging) images could use 16bits per pixel.

A 8bit grayscale Image.

BIT Plane

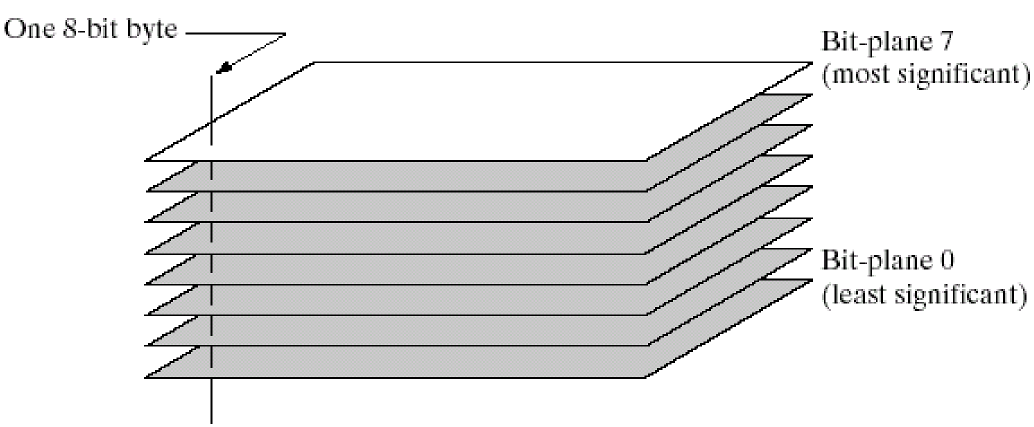


FIGURE 3.12 Bit-plane representation of an 8-bit image.

8-bit Image

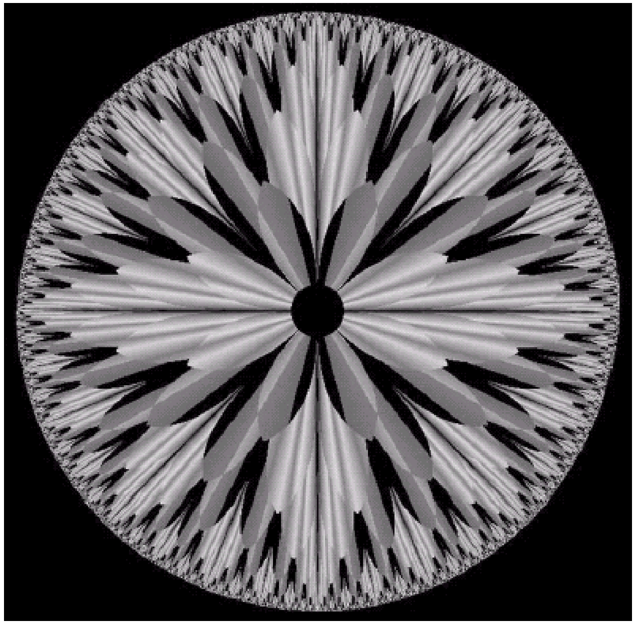


FIGURE 3.13 An 8-bit fractal image. (A fractal is an image generated from mathematical expressions). (Courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA.)

8 bit plane slicing

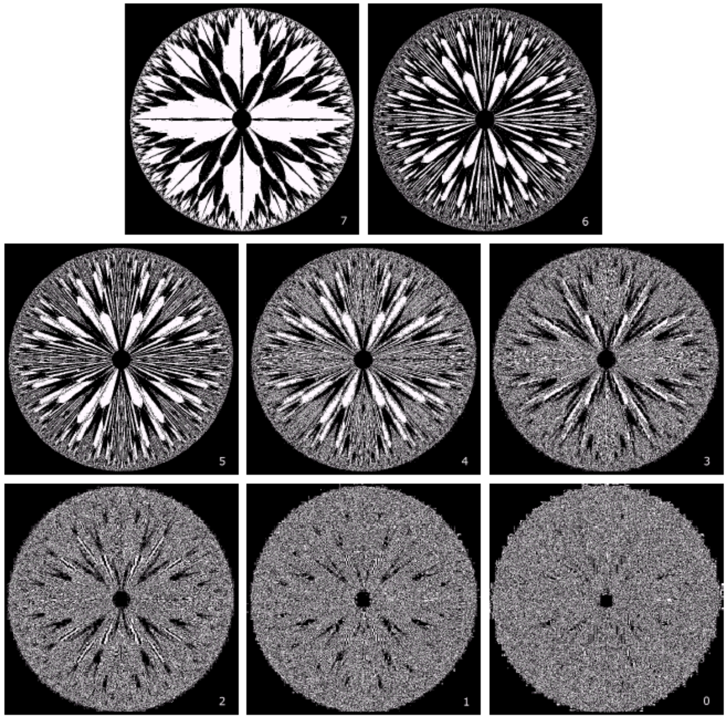


FIGURE 3.14 The eight bit planes of the image in Fig. 3.13. The number at the bottom, right of each image identifies the bit plane.

Binary Image

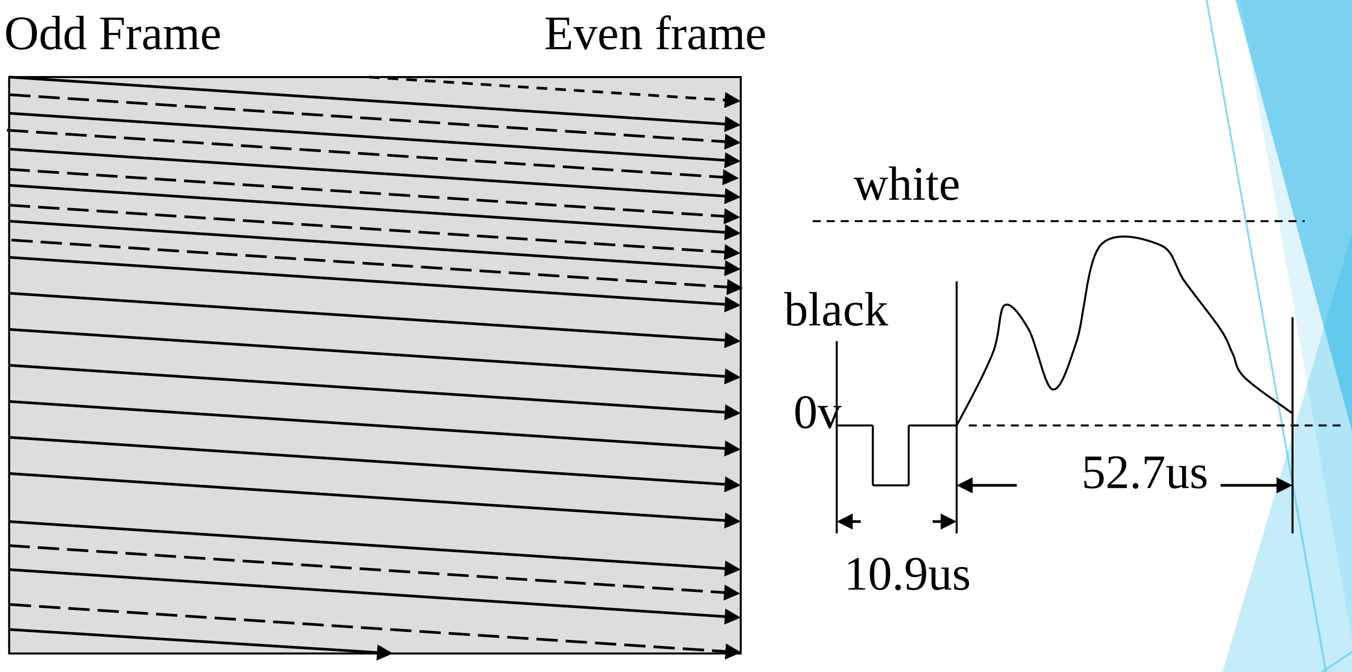
- Binary Image

A binary image has only two values (0 or 1).

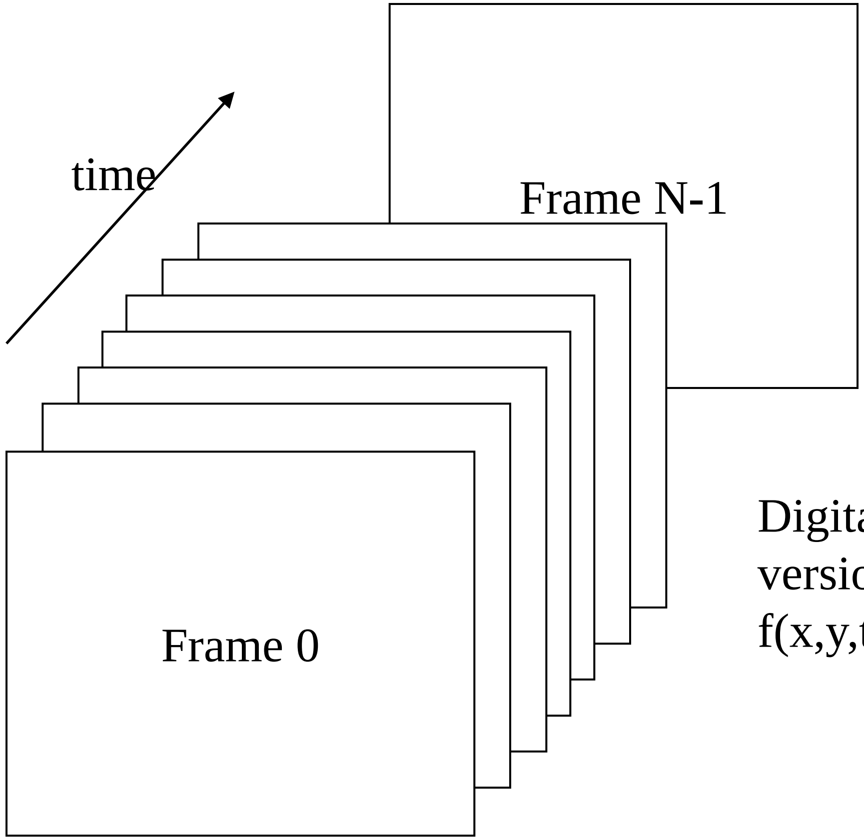
Binary image is quite important in image analysis and object detection applications.

Video

Analog video



Digital Video



Digital video is digitized version of a 3D function f(x,y,t)