Data Representation II

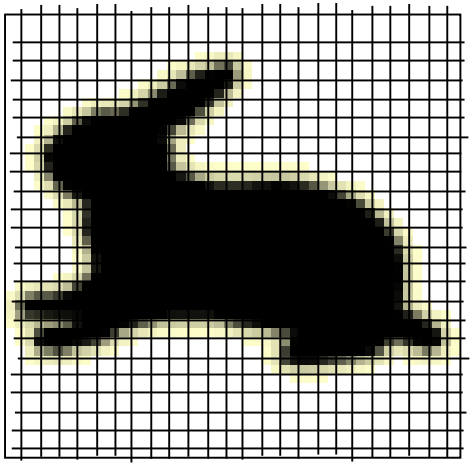
Sabah Sayed

Topics

1. Black and White Images.
2. Gray Images.
3. Colored Images.
4. Bitmap Images.
5. Vector Images.
6. Data Compression.
7. Run Length Compression.
8. Dictionary Compression.
9. Image Compression

How can we represent Images?

- We can break it into pixels
- The simplest form is black/white
- We can represent a pixel by one bit



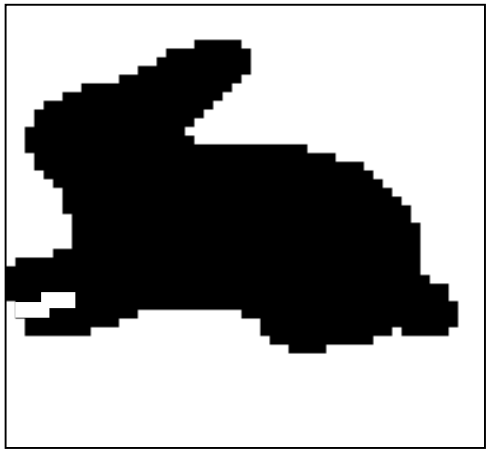


Image representation

- An image can be divided into many tiny squares, called pixels.
- Each pixel has a particular color.
- The quality of the picture depends on 2 factors:
  - the **density** of pixels.
  - The **number of bits** representing colors.
- The resolution of an image is the pixels density.
- The higher the resolution the more information the image contains.
  - 600 x 800 is better than 400 x 600

Image representation

- Black and white images
- Gray images
- Colored images
- Bitmap images
- Vector images

(a) (b) (c)

Figure 1 Different version of the same image; black and white (a), grey-level (b), and colour (c).

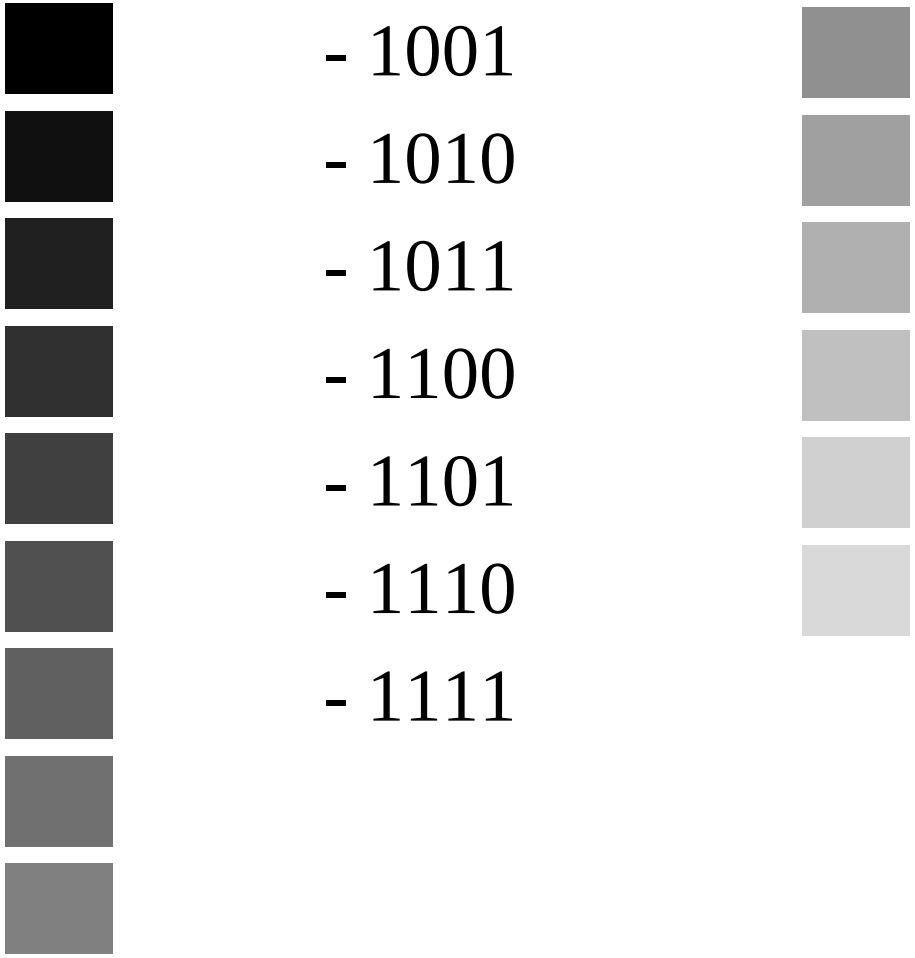
Gray Bitmap Images

- Each individual pixel in a graphic stored as a binary number
  - Pixel: A small area with associated coordinate location
  - Example: 2-bit per pixel
    - $4=2^2$ choices
    - 00 (off, off)=white
    - 01 (off, on)=light grey
    - 10 (on, off)=dark grey
    - 11 (on, on)=black

Example: 4 bit per pixel

- Each point below is represented by a 4-bit code corresponding to 1 of 16 shades of gray

- 16 =2^4 choices
  - 0000
  - 0001
  - 0010
  - 0011
  - 0100
  - 0101
  - 0110
  - 0111
  - 1000
  - 1001
  - 1010
  - 1011
  - 1100
  - 1101
  - 1110
  - 1111



Color Representation

- Colors are represented with a sequence of bits.
- 256 colors – how many bits?
  - Hint for calculating
    - To figure out how many bits are needed to represent a range of values, figure out the smallest power of 2 that is equal to or bigger than the size of the range.
    - That is, find *x* for $2^x => 256$
- 24-bit color – how many possible colors?
  - 16 million possible colors
- 24-bit color is often referred to as the *true colour.*
  - Any real-life shade, detected by the naked eye, will be among the 16 million possible colours.

RGB

- uses additive color mixing, because it describes what kind of light needs to be emitted to produce a given color.
- RGB stores individual values for red, green and blue.
  - (r,g,b)=>(0,0,0) black
  - (r,g,b)=>((1,1,1) white
  - ranges 0 to 1

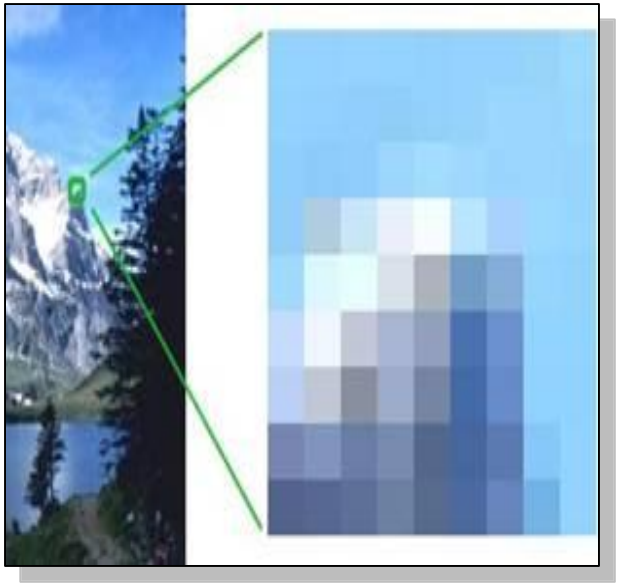
Image Formats

- The two main image categories are:
  1. **Bitmaps:** A bitmap is a method for storing images using pixels. It is called a bitmap because it is a 'map' of where the 'bits' of information are stored. Example of image types with bitmap formats are BMP, JPEG, GIF.
  2. **Vector (geometric) Images:** they are images that are completely described using mathematical definitions.

❖ You can convert a vector image into a bitmap, but you cannot convert a bitmap into a vector image.

Bitmap Images

- Bitmap images (also called raster images) are made with pixels (picture element), which look like rectangles.
- All the pixels, when combined for visual images, are called continuous tone images (contones).
- Bitmap images are resolution dependent, and this must be taken into consideration when producing images of different size and quality.
- Bitmap images can be:
  - Monochrome (1 bit)
  - Gray scale (8 bits)
  - RGB (24 bits)
  - RGB Lookup (8 bits)



Direct Coding (RGB)

- Storage space for each pixel to code the color
- Use 3 bytes per pixel (1 for R, 1 for G and 1 for B) [Industry standard]
- 256 different intensity level for each color

Bit 1 R	Bit 2 G	Bit 3 B	Color Name
0	0	0	Black
0	0	1	Blue
0	1	0	Green
1	0	0	Red

Lookup table (RGB Lookup)

- Pixel values do not code colors directly
- Refer to a table of color values
- A table with 256 colors with RGB values

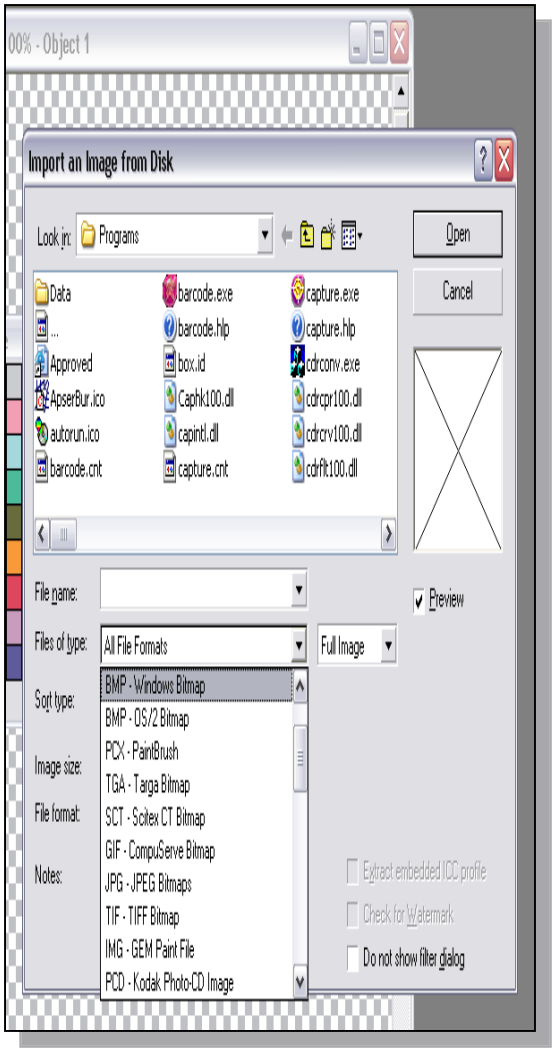
	r	g	b
0	**11111111**	**11111111**	**11111111**
1	**00100000**	**1000000**	**00100000**
2			
3			
…			
255			

Bitmap Images

- Resolution is the density of the pixel grid. It is the number of pixels in an image and is referred to as dpi or dots per inch.
- Resolution is based on the number of pixels in an image, which is determined, by its width and height of the pixels.
- Example: Image size = width in pixels * height in pixels

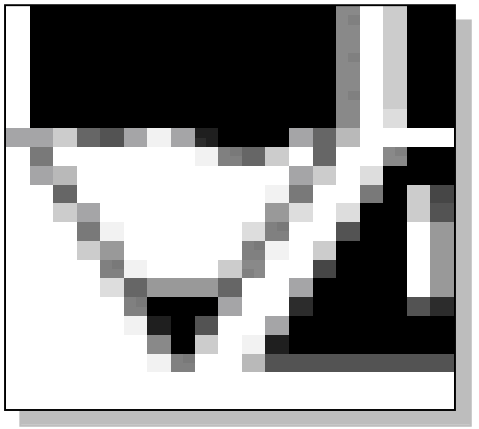
Advantages of Bitmap Images

1. Bitmap images are easily converted to different formats.
2. Bitmap images are easier to import into different software applications.
3. Bitmap images produce a variety of continuous tone images.
4. Bitmap images are better suited for most high quality renderings and web page graphics.



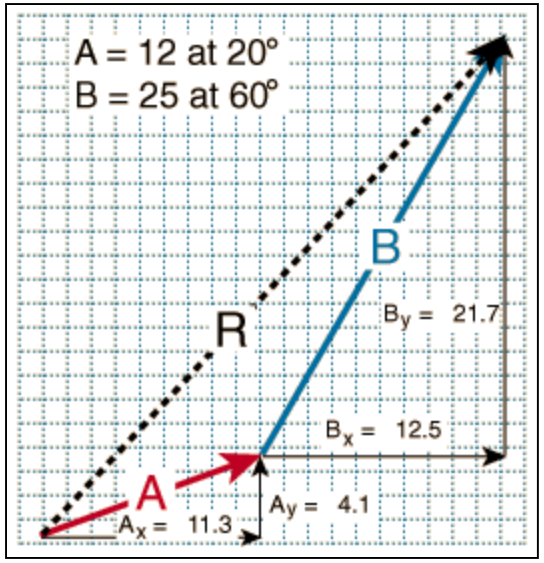
Disadvantages of Bitmap Images

1. Bitmap images produce larger files sizes.
2. Bitmap images have restrictions in regards to alterations and modifications such as scale (resizing). Because of this, bitmap files must usually be printed at the resolution in which they were originally stored



Vector Images

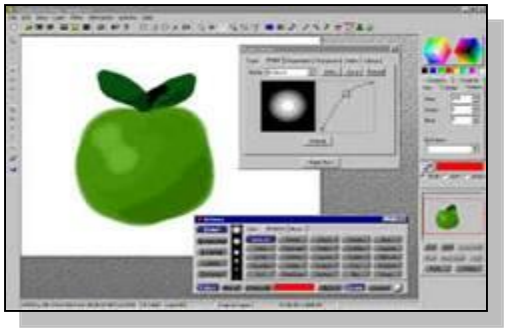
- Vector images (also called outline images) are images made with lines, text, curves and shapes.
- Vector graphics are resolution independent.



Advantages of Vector Images

- They can be output to the highest quality at any scale. i.e. an image can be enlarged or reduced without affecting the quality of the image.

- Vector graphic images normally have much smaller file sizes than raster-based bitmaps.

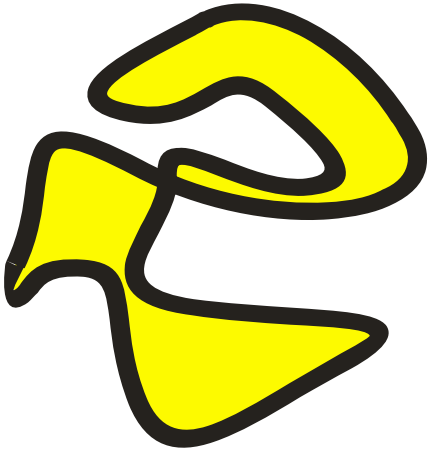


- Changing or transforming the characteristics of a vector object does not affect on the object.

- There is no background unless it is placed behind the image as a layer

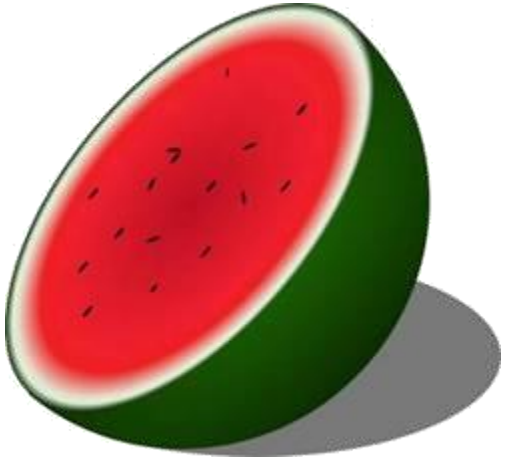
Advantages of Vector Images

- Vector images are not limited to rectangular shapes like bitmaps.
- Vector images have the appearance of artistic form such as cartoons.
- Vector images can be easily **converted to bitmap** images.
- Lines and curves are easily defined and will always be smooth and retain their continuity.

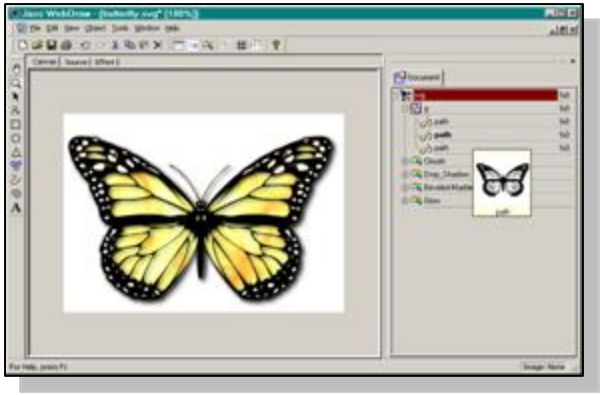


Disadvantages of Vector Images

- The main disadvantage is they are not photorealistic.



- Vector images are usually filled with solid or gradient colors but lack in depth and appearance in the values and colors of a true continuous tone image.



Data Compression

- Lossy versus lossless

1. Run-length encoding (lossless)
2. Dictionary encoding (Includes adaptive dictionary encoding such as LZW encoding.) (lossless)

Run Length Encoding

- The process of replacing sequences of identical data elements with a **code** indicating the element that is repeated and the number of times it occurs in the sequence.

- Example: 11111111111000001111111111111000000000000000000000011111111111111
  - 70 bits
  - 11 (1) + 5 (0) + 14 (1) + 23 (0) + 17 (1)

Dictionary Encoding

- In particular, an entire word can be encoded as a single reference to this dictionary rather than as a sequence of individual characters encoded using a system such as UTF-8.

- Dictionary encoding can be used by word processors to compress text documents because the dictionaries already contained in these processors for the purpose of spell checking make excellent compression dictionaries.

Dictionary Encoding Example

- A typical dictionary in a word processor contains approximately 25,000 entries, which means an individual entry can be identified by an integer in the range of 0 to 24,999.
- This means that a particular entry in the dictionary can be identified by a pattern of only 15 bits. In contrast, if the word being referenced could be 4 or 8 bytes long

LZW Encoding

- **Adaptive dictionary encoding:** the dictionary is allowed to change during the encoding process.
- We have a dictionary

x	y	space	xxy	yyx
1	2	3	4	5

xxy yyx xxy xxy yyx

1123221343435

Compressing Images

- Compression is **lossy**!
- GIF: Good for cartoons
  - a dictionary encoding system
- JPEG: Good for photographs
  - most digital cameras use JPEG as their default compression technique.

Graphics Interchange Format GIF

- GIF approaches the compression problem by reducing the number of colors that can be assigned to a pixel to only 256.

- The red-green-blue combination for each of these colors is encoded using three bytes (same technique as lookup for BMP)

- These 256 encodings are stored in a table (a dictionary) called the palette..

Graphics Interchange Format GIF

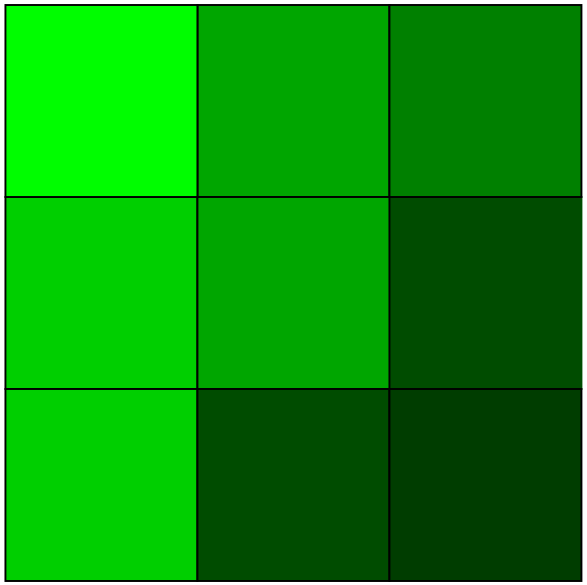
- GIF is a **lossy** compression system when applied to arbitrary images because the colors in the palette may not be identical to the colors in the original image.

- GIF can obtain additional compression by extending this simple dictionary system to an adaptive dictionary system using LZW techniques.

Joint Photographic Experts Group JPEG

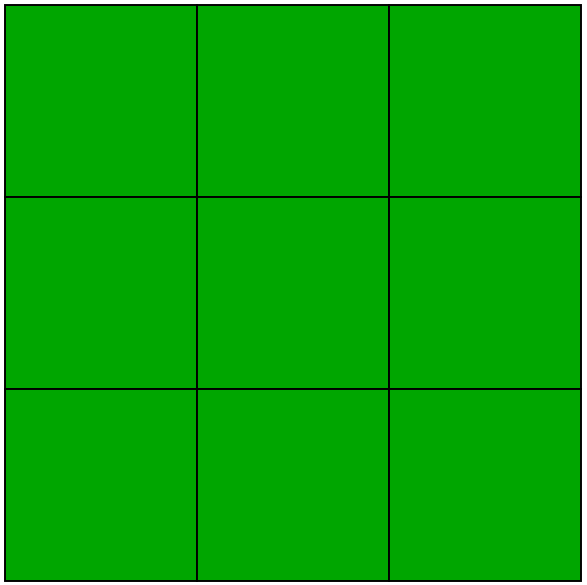
- JPEG's baseline standard (also known as JPEG's lossy sequential mode) has become the standard of choice in many applications.
- It takes advantage of a human eye's limitations.
- Eye is more sensitive to changes in brightness than to changes in color.
- Encoded in terms of **luminance** and **chrominance**.
  - ➢Average the chrominance values over two-by-two pixel squares. This reduces the size of the chrominance information by a factor of four while preserving all the original brightness information.
  - ➢ Then, divide the image into eight-by-eight pixel blocks and Compress each block as a unit.

JPEG



9:1

“averaging”



The result is a significant degree of compression without a noticeable loss of image quality.