Steganography

COMP620

Steganography

- **Steganography** is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message
- A form of security through obscurity

Review of Graphic File Formats

- **Vector** – Pictures are drawn as a series of lines. The hardware draws a line from point A to point B.
- **Raster** – Pictures are represented as a matrix of picture elements or pixels.

Picture Size and Resolution

- Screen resolution is measured in horizontal and vertical pixels per inch.
- A good monitor can display about 90 pixels per inch.
- A good printer can print about 600 to 1200 pixels per inch.
- True size of a picture depends on the number of pixels and the device.

Pixel

- Each pixel has a color.
- Monochrome pictures have only two colors.
- The color is usually represented as three numbers that are the intensity of the three primary colors, Red, Green and Blue (RGB)
- The pixel represents both the color and the intensity or darkness of the pixel.

Bits / Pixel

- The number of bits it takes to represent a pixel depends on the number of possible colors.
- Black and white requires only 1 bits per pixel.
- Full color pictures require 24 bits per pixel
- Simple graphics can use 4 or 8 bits per pixel

Graphics Formats

- There are many formats for graphical data
  - BMP
  - JPEG
  - GIF
  - TIFF
  - PNG
  - etc.

Bit Mapped File

- BMP is a very simple format for graphics.
- Used frequently by Microsoft system.
- Each pixel is stored as a number. The number of bits per pixel is determined by the number of different colors.

BMP format

BMP File Header	Stores general information about the BMP file.
Bitmap Information (DIB header)	Stores detailed information about the bitmap image.
Color Palette	Stores the definition of the colors being used for indexed color bitmaps.
Bitmap Data	Stores the actual image, pixel by pixel.

BMP Efficiency

- Each pixel in a BMP file is represented by a number.
- There is no compression.
- It does not matter how "complex" the image, the file size is determined by the number of pixels or the size of the picture.

JPEG files

- The JPEG format was created by the Joint Photographic Experts Group
- JPEG is a commonly used method of compression for photographic images
- JPEG uses lossy compression. Some image quality is lost.
- You can control the level of compression and therefore the image quality.
- JPEG uses a Discrete cosine transform for compression.

JPEG Example

- Q = 100
- 81.3 KB
- 219,726 bytes if BMP
- 37% of BMP

JPEG Example

- Q = 50
- 14.7 KB
- 18% of Q100
- 6.7% of BMP

JPEG Example

- Q = 25
- 9.32 KB
- 11.5% of Q100
- 4.2% of BMP

JPEG Example

- Q = 10
- 4.67 KB
- 5.7% of Q100
- 2.1% of BMP

JPEG Example

- Q = 1
- 1.48 KB
- 1.8% of Q100
- 0.67% of BMP

JPEG Artifacts

Appropriate Use

- BMP files are not distorted in any way, but they are large.
- JPEG works well for photographs
- GIF works well for diagrams or charts.

Hiding Data in Images

- Data can easily be hidden in the least significant bits of the pixels of a complex graphics file
- One bit of text can replace the lowest bit of one (or all) of the colors
- Changing only the least significant bit is very hard to detect visually
- Statistically the text actually changes the bit only 50% of the time

Example Images 1 blue bit / pixel

Original image

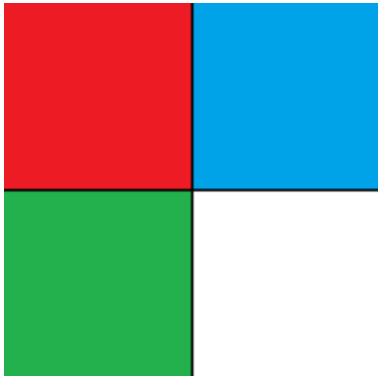
with hidden text

Example Images 8 bit / pixel

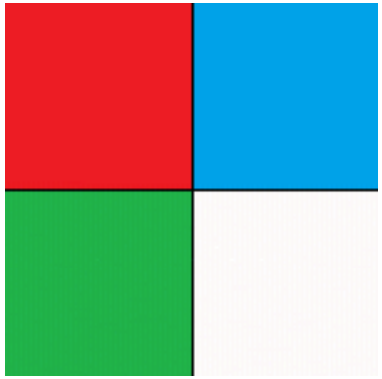
Original image

with hidden text

Example Images 8 bit / pixel



Original image



with hidden text

Legitimate Use

- Some brands of color printers add tiny yellow dots giving the printer serial number plus the time and date.
- Digital watermarking can be used to identify ownership
  - May be visible or invisible
  - Robust watermarking can survive certain transformations to the file

Steganography Assignment Notes

- You may want to read the text from a file instead of from the keyboard
- Do not worry about marking the end of data in case the text and BMP are different sizes
- The Java source code is available on the class website

Chaffing and Winnowing

- Chaffing and winnowing is a cryptographic technique to achieve confidentiality without using encryption when sending data over an insecure channel.
- The technique was conceived by Ron Rivest
- Different from encryption or steganography
- Appears to be useful in places that do not allow encryption

Hiding in Plain Sight

- The idea of chaffing and winnowing is to embed the data in a large stream of irrelevant information
- Only the receiver should be able to determine what data is part of their message
- A “Message Authentication Code” can be created by encrypting a hash of the data

Example chaffing and winnowing

- Alice divides her message into bits and sends one bit per packet to Bob
- Each packet contains the data bit, a serial number and an authentication code
- The authentication code is a symmetric key encryption of the data and serial number
- In addition to the data, Alice sends many similar packets that have no meaning
- Bob has the key to authenticate the real packets to reproduce the message