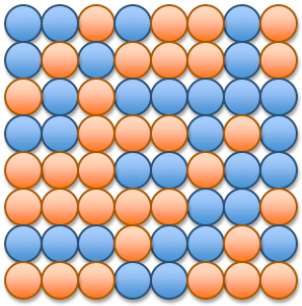
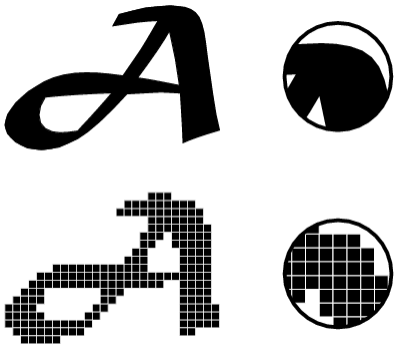


UNIT 7C
Data Representation: Images and Sound

Encoding Images:
Vector vs. Raster / Bitmap

- There are two major ways to store images:
  - Vector graphics: a series of lines or curves. Expensive to compute but smoothly rescales.
  - Raster or Bitmap graphics: an array of pixels. Cheap to compute, but scales poorly.



Common Standards

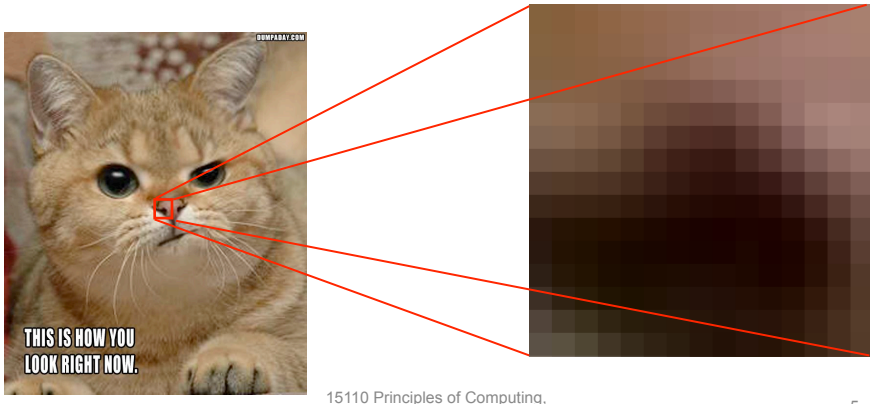
- Vector: SVG, EPS, AI, CDR.
  - Special-purpose: commonly used for high-quality illustrations, graphics, etc.
- Raster: JPEG (compression), GIF (transparency), PNG (web portability), TIFF (printing, huge), BMP (huge)
  - Commonly used for photos and pretty much everything

Image file formats

- File type
  - Usually a "magic number" in the first 4 bytes tells us if it is a JPEG, GIF, PNG, etc. file
- File format (or subtype)
  - Pixel format (color or B&W, how many bytes?)
  - Data compression algorithm
  - Etc.
- Size of image
  - Rows
  - Cols
- Image data

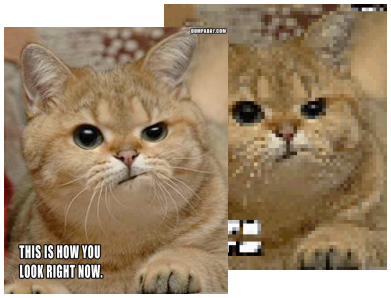
Pixels

- A raster/bitmap image is stored in a computer as a sequence of *pixels*, picture elements.



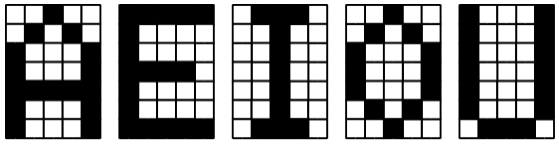
Resolution

- The resolution of an image is the number of pixels used to represent the image (e.g. 1024 X 768).
- Each pixel represents the average color in that region.
- The more pixels per area, the higher the resolution, and the more accurate the image will appear.



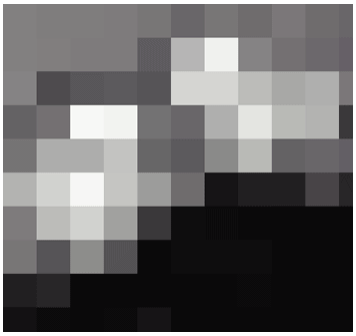
Storing Bitmap Images

- In bitmapped images, each pixel is represented in computer memory in binary, just like other data types.
- If pixels of an image are black or white only, then we only need 1 bit per pixel to store the image, e.g. 00100 might be first row of "A".



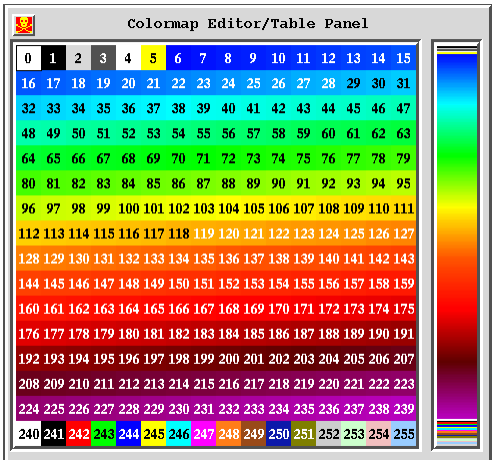
Grayscale Images

- Grayscale images contain pixels that are various shades of gray, from black (maximum gray) to white (minimum gray).
- If there are 256 levels of gray for pixels, we can represent each pixel using 8 bits.
  11111111 = white
  : (shades of gray)
  00000000 = black
- E.g., (in hex) 7F 7D … 00 00 00 00



256-color images (8-bit color)

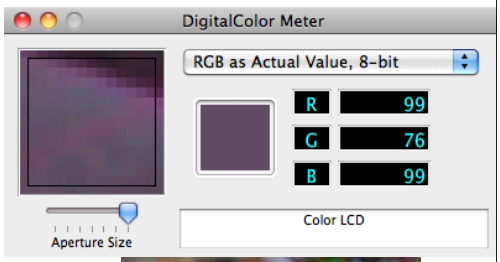
- Each pixel is represented with a 8-bit value that is an index into a *palette* of 256 colors.



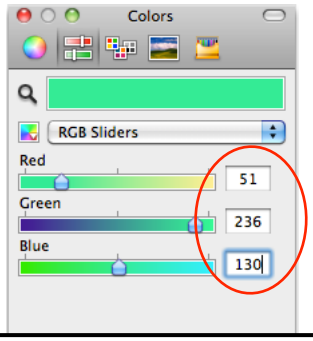
RGB-color images (24-bit color)

- Colors are represented as mixtures of red (R), green (G), and blue (B).
- Each pixel is represented using three 8-bit values, one for each color component.
- This representation allows for 2^{24} = 16,777,216 different colors.
- This representation is also called "true color".

(image from Wikipedia)



RGB example



	RED	GREEN	BLUE
dec:	51	236	130
bin:	00110011	11101100	10000010
hex:	3 3	E C	8 2

Comparing Representations

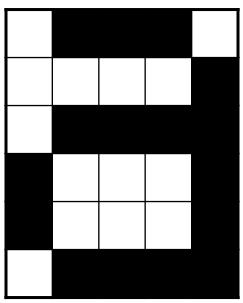
- If an image has a resolution of 640 X 480 (307,200 pixels), how many bytes does each representation require?
  - B&W 38,400 bytes (307200/8)
  - 8-bit grayscale 307,200 bytes
  - 256-color (8-bit color) 307,200 bytes
  - 24-bit color 921,600 byte(307200*24/8)
- A single RGB image of size 1600 X 1200 requires over 5.76 million bytes!

Compressing Raster Data

- Run-length encoding (lossless, limited)
- Color maps (GIF) (usually lossy)
- JPEG (lossy - a suite of techniques exploiting human visual perception)

RLE compression

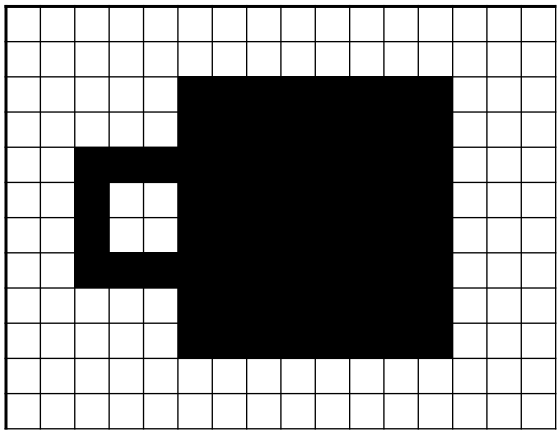
- Run-Length Encoding is a lossless compression technique used in early image files.
- For a black-and-white image, Instead of storing the 8-bit value for every pixel, we store how many of each color occurs in a row (run).
- This saves a lot **when there are large runs of the same color**.
- If there are more than 2 colors, then we store color, run, color, run, ...



white, black, white, black, etc.

1, 3, 1
4, 1
1, 4
0, 1, 3, 1
0, 1, 3, 1
1, 4

RLE Comparison



RLE(B&W)	Bitmap
1 byte	16 bytes
1 byte	16 bytes
3 bytes	16 bytes
3 bytes	16 bytes
3 bytes	16 bytes
5 bytes	16 bytes
5 bytes	16 bytes
3 bytes	16 bytes
3 bytes	16 bytes
3 bytes	16 bytes
1 byte	16 bytes
3 bytes	16 bytes
32 bytes	192 bytes

GIF: Graphic Interchange Format

- Developed by CompuServe in the late 1980s to represent 8-bit images efficiently.
- Each pixel is an 8-bit value, mapping to a table of 256 24-bit RGB colors.
- A *codebook* stores recurring sequences.
- Useful for representing images with fewer colors or large areas of color like company logos.

JPEG (JPG): Joint Photographic Experts Group

- A lossy compression technique used generally for photographic images.
  - Perceptual Coding: based on what we can/cannot see.
- Supports varying levels of compression.

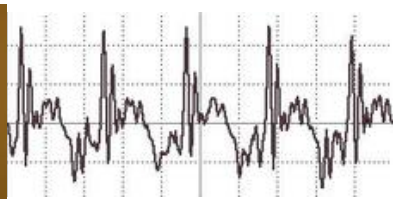
Higher quality
Compression 2.6:1
(images from Wikipedia)

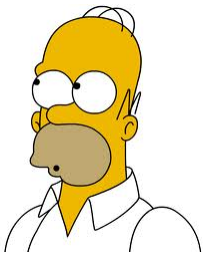
Medium quality
Compression 23:1

Lowest quality
Compression 144:1

Sound

- When an instrument is played or a voice speaks, changes occur in air pressure which our ears interpret as sound.

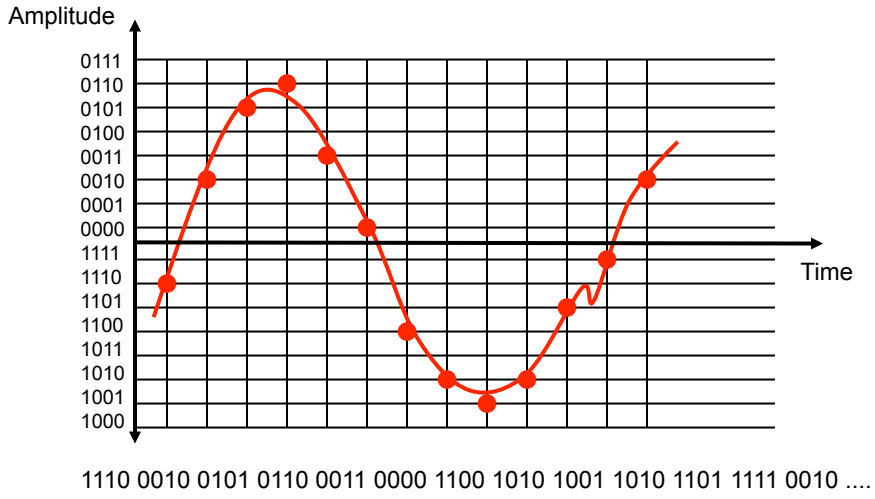




Sampling

- A sound is represented digitally by *sampling* an electronic version of the audio signal.
- Accuracy determined by
  - Sampling rate
  - Sample size
- Sampling rate: The amplitude (pressure) of the signal is measured (sampled) at equally-spaced time intervals.
- Sample size: The amplitude axis is divided into equally-spaced intervals depending on how many bits we want to store per sample.

Sampling



Sampling

- In order to reproduce the audio waveform as accurately as possible, we need to increase the sampling rate (samples per second) and the number of amplitude levels (bits per sample).
- Note in the previous picture how some of the samples had to be moved up or down to match an amplitude level and some finer changes in the sound signal could be missed if the sampling rate is too low.

Sampling

- Digital recordings are typically sampled at 44,100 samples per second (Hertz – Hz).
  - This is due to the "sampling theorem" that states that *the sampling rate must be at least twice the highest frequency in the sound*, and humans can hear up to approx. 20,000 Hz.
- For accurate amplitude readings, sound is often sampled at 16-bits per sample (so there are 65,536 amplitude levels that can be measured).
  - *Quantization* (rounding to integer sample values) *introduces noise.* Adding one bit cuts the noise in half.

Compressing Sound Files

- *codecs* (compression/decompression) implement various compression/ decompression techniques
- Lossless: WMA Lossless, ALAC, MPEG-4 ALS, …
- Lossy: MPEG, like JPEG, a family of perceptually-based techniques

MP3

- MP3 (MPEG3) is a lossy compression technique.
- This format takes advantage of some facts about human hearing.
  - Our hearing is better in mid range frequencies than on the low and high ends.
  - If a loud and soft sound play at about the same time or about the same frequencies, we can't hear the soft sound: this is called *masking*
  - *Masking can hide noise introduced by compression.*

MP3 Compression

- Like JPEG, MP3 has various levels of compression:

Bit Rate	Compression Ratio	Quality
256Kbps	5:1	Supreme (near best)
192Kbps	7:1	Excellent (better)
128Kbps	11:1	(good)
96Kbps	19:1	(fair)
64Kbps	22:1	FM quality (poor)

- MP3 also has Variable Bit Rate (VBR) since compression ability can vary at different segments of the digital recording.

MP4

- MPEG = Moving Picture Experts Group
- MP4 (MPEG4) is a compression technique developed for video.
- Since most of the time there are only small changes from one frame of a video to the next frame, large savings are possible.
- DVD movies and smart phones use the MP4 encoding.

YouTube, Vimeo, etc.

- Video hosting sites such as YouTube and Vimeo support a variety of formats, including MP4, AVI (Microsoft), QuickTime (Apple), and Flash (Adobe).
- You can download videos from these sites in your preferred format using tools such as KeepVid (visit www.keepvid.com).
- Uploading and then downloading a video may reduce the quality due to lossy compression.