LBSC 690 Session #11

Multimedia

Jimmy Lin
The iSchool
University of Maryland

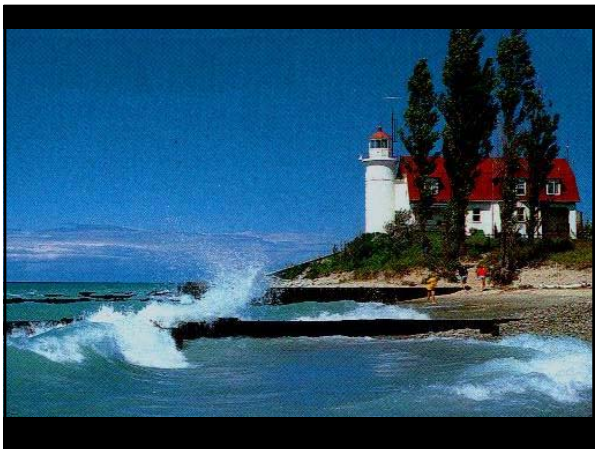
Wednesday, November 12, 2008

Take-Away Messages

- Human senses are gullible
  - Images, video, and audio are all about "trickery"
- Compression: storing a lot of information in a little space
  - So that it fits on your hard drive
  - So that you can send it quickly across the network

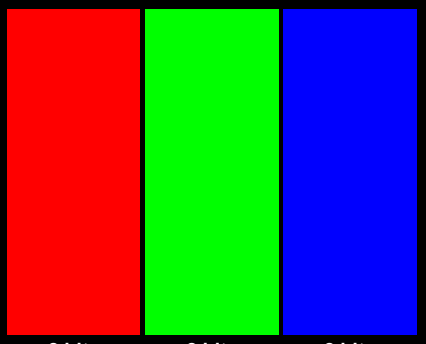
How do you make a picture?

Georges Seurat - A Sunday Afternoon on the Island of La Grande Jatte



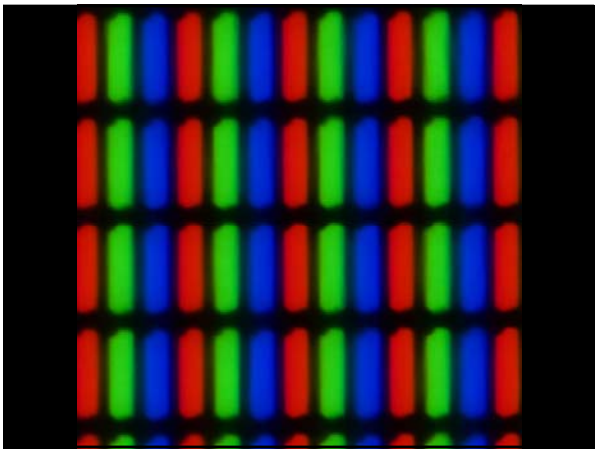
What's a pixel?
What's "resolution"?

How do you get color?



#99FF66
#9999FF

How do LCDs work?



How do digital cameras work?

2,048 x 1,536 = 3,145,728 ≈ 3 MP
2,560 x 1,920 = 4,915,200 ≈ 5 MP
3,264 x 2,448 = 7,990,272 ≈ 8 MP
3,648 x 2,736 = 9,980,928 ≈ 10 MP

Is a picture really worth 1000 words?

(consider an image with 1024 x 768 resolution)

Compression

- Goal: represent the same information using fewer bits
- Two basic types of data compression:
  - Lossless: can reconstruct exactly
  - Lossy: can't reconstruct, but looks the same
- Two basic strategies:
  - Reduce redundancy
  - Throw away stuff that doesn't matter

Run-Length Encoding

- Opportunity:
  - Large regions of a single color are common
- Approach:
  - Record # of consecutive pixels for each color
- An example with text:

Sheep go baaaaaaaaa and cows go moooooooooo
→ Sheep go ba<10> and cows go mo<10>

Using Dictionaries

- Opportunity:
  - Data often has shared substructure, e.g., patterns
- Approach:
  - Create a dictionary of commonly seen patterns
  - Replace patterns with shorthand code
- An example with text:

The rain in Spain falls mainly in the plain
→ The r* ^ Sp* falls m*ly ^ the pl* (*=ain,^=in)

Palette Selection

- Opportunity:
  - No picture uses all 16 million colors
- Approach:
  - Select a palette of 256 colors
  - Indicate which palette entry to use for each pixel
  - Look up each color in the palette
- What happens if there are more than 256 colors?

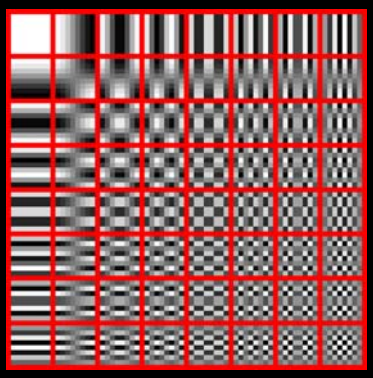
This is GIF!

The iSchool
University of Maryland

Discrete Cosine Transform

- Opportunity:
  - Images can be approximated by a series of patterns
  - Complex patterns require more information than simple patterns
- Approach:
  - Break an image into little blocks (8 x 8)
  - Represent each block in terms of "basis images"

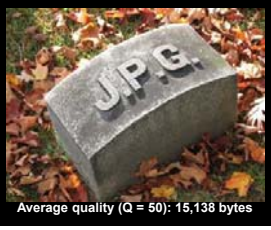
The iSchool
University of Maryland



This is JPEG!

Full quality (Q = 100): 83,261 bytes

Medium quality (Q = 25): 9,553 bytes



Average quality (Q = 50): 15,138 bytes

Low quality (Q = 10): 4,787 btes

When should you use jpegs?
When should you use gifs?

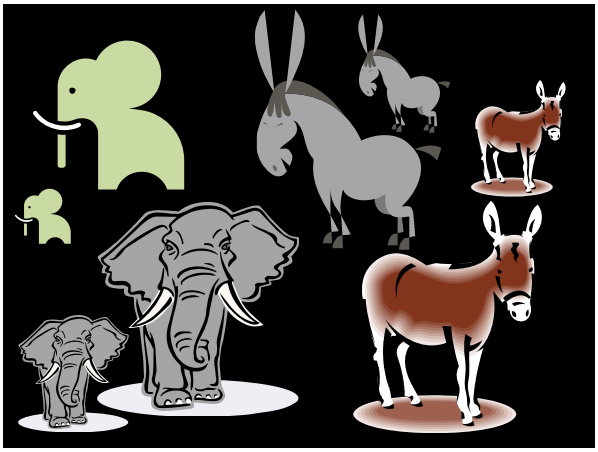
Demo!

Raster vs. Vector Graphics

- Raster images = bitmaps
  - Actually describe the contents of the image
- Vector images = composed of mathematical curves
  - Describe *how* to draw the image

The iSchool
University of Maryland

What happens when you scale vector images?
What happens when you scale raster images?



How do you make video?

Basic Video Coding

- Display a sequence of images…
  - Fast enough to trick your eyes
  - (At least 30 frames per second)
- NTSC Video
  - 60 “interlaced” half-frames/sec, 720x486
- HDTV
  - 30 “progressive” full-frames/sec, 1280x720

The iSchool
University of Maryland

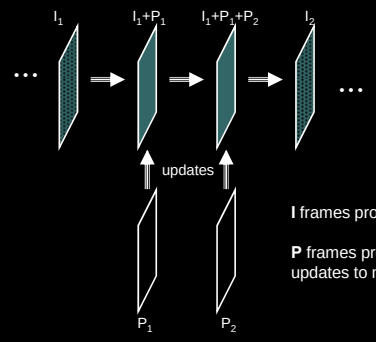
Video Example

- Typical low-quality video:
  - 640 x 480 pixel image
  - 3 bytes per pixel (red, green, blue)
  - 30 frames per second
- Storage requirements:
  - 26.4 MB/second!
  - A CD-ROM would hold 25 seconds
  - 30 minutes would require 46.3 GB
- Some form of compression required!

Video Compression

- Opportunity:
  - One frame looks very much like the next
- Approach:
  - Record only the pixels that change

Frame Reconstruction



I frames provide complete image

P frames provide series of updates to most recent I frame

What is sound?
How does hearing work?
How does a speaker work?
How does a microphone work?

Basic Audio Coding

- Sample at twice the highest frequency
  - 8 bits or 16 bits per sample

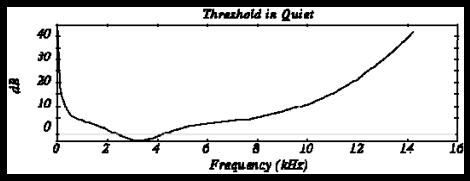
- Speech (0-4 kHz) requires 8 KB/s
  - Standard telephone channel (8-bit samples)
- Music (0-22 kHz) requires 172 KB/s
  - Standard for CD-quality audio (16 bit samples)

How do MP3s work?

- Opportunity:
  - The human ear cannot hear all frequencies at once, all the time
- Approach:
  - Don't represent things that the human ear cannot hear

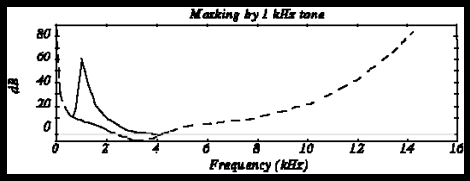
Human Hearing Response

Experiment: Put a person in a quiet room. Raise level of 1kHz tone until just barely audible. Vary the frequency and plot the results.



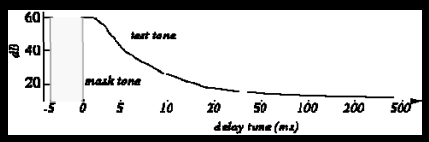
Frequency Masking

Experiment: Play 1kHz tone (masking tone) at fixed level (60 db). Play test tone at a different level and raise level until just distinguishable. Vary the frequency of the test tone and plot the threshold when it becomes audible.



Temporal Masking

If we hear a loud sound, then it stops, it takes a while until we can hear a soft tone at about the same frequency.



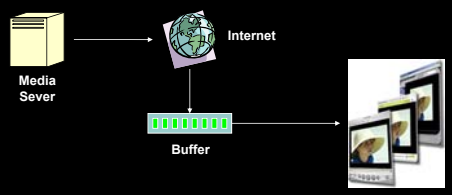
MP3s: Psychoacoustic compression

- Eliminate sounds below threshold of hearing
- Eliminate sounds that are frequency masked
- Eliminate sounds that are temporally masked
- Eliminate stereo information for low frequencies

How do you deliver continuous data over packet-switched networks?

Streaming Audio and Video

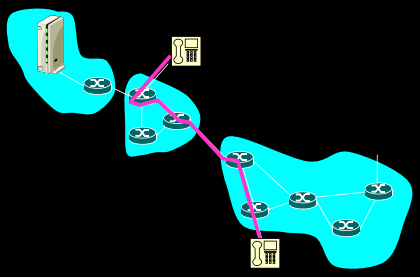
- Simultaneously:
  - Receive downloaded content in buffer
  - Play current content of buffer
- Analogy: filling and draining a basin concurrently



to buffer or not to buffer…

Internet radio
YouTube
Skype
Instant Messenger

Example: Internet Telephony



IP Phones: Network Issues

- **Network loss:** packets lost due to network congestion
- **Delay loss:** packets arrives too late for playout at receiver
- **Loss tolerance:** depending on voice encoding packet loss rates between 1% and 10% can be tolerated

IP Phones: Playout Delay

- Receiver attempts to playout each chunk exactly *q* ms after chunk was generated
  - Chunk has time stamp *t*: play out chunk at *t+q*
  - Chunk arrives after *t+q*: data arrives too late for playout, data "lost"
- Tradeoff for *q*:
  - Large *q*: less packet loss
  - Small *q*: better interactive experience

Take-Away Messages

- Human senses are gullible
  - Images, video, and audio are all about "trickery"
- Compression: storing a lot of information in a little space
  - So that it fits on your hard drive
  - So that you can send it quickly across the network