Data Representation

Interpreting bits to give them meaning

Part 4: Media - Sound, Video, Compression

Notes for CSC 100 - The Beauty and Joy of Computing
The University of North Carolina at Greensboro

Reminders

Big thing for this week:

Project Proposal Presentations: This Friday

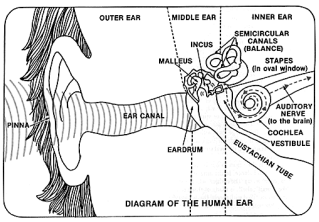
Homework 3

- Should have completed online fractal tutorial
- Hopefully have started playing around with drawing in BYOB
- HW 3 due: Wednesday, November 6

Sound

What is sound?

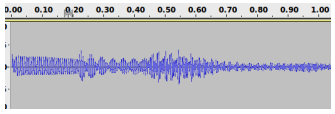
Sound is just rapid fluctuations in air pressure, detected by the (somewhat delicate!) organs in our ears



Sound

Sound waveforms

We can plot changes in pressure over time:



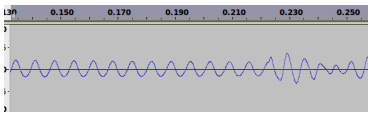
Main components:

- *Intensity* (how much pressure changes): We perceive this as "loudness" and in graph would be reflected in larger fluctuations
- *Frequency*: How rapid are the fluctuations?

Sound

Sound waveforms - Zooming in!

"Pure" tone is a sine wave (real world sounds are generally not pure!)



One *cycle* here is approximately 0.150 seconds to 0.157 seconds:

- *Period* is 0.007 seconds
- *Frequency* is 1/0.007 = 142.857... Hz (for "Hertz")
  - For reference, "middle C" is around 261.626 Hz
  - An octave doubles/halves frequency, so this note is a probably something like a "D below middle C" (which is 146.8 Hz)

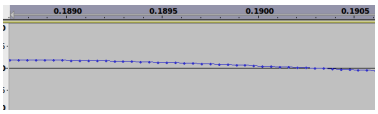
Question: How do we make this digital?

Sound

Sound waveforms - Zooming in even more!

Answer: We *sample* the waveform many times per second.

This is zoomed in enough where you can see actual samples:



Quality of sound reproduction depends on sample rate (samples per second):

- In this example, 22 samples between 0.1890 and 0.1900
  - *So 22/(0.190-0.189) = 22,000 samples per second*
- CD sound: 44,100 samples/second
- Typical DVD sound: 48,000 samples/second

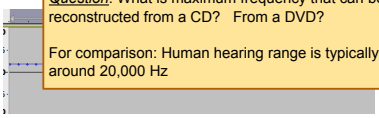
Nyquist Theorem: Perfect reconstruction of signals with frequency <= F if you sample at (2/F) samples/second

Sound

Sound waveforms - Zooming in even more!

Answer: We ***sample*** the waveform many times per second.

This is zoomed in



Question: What is maximum frequency that can be reconstructed from a CD? From a DVD?

For comparison: Human hearing range is typically 20 Hz to around 20,000 Hz

Quality of sound reproduction depends on sample rate (samples per second):

- In this example, 22 samples between 0.1890 and 0.1900
  - *So 22/(0.190-0.189) = 22,000 samples per second*
- CD sound: 44,100 samples/second
- Typical DVD sound: 48,000 samples/second

Nyquist Theorem: Perfect reconstruction of signals with frequency <= F if you sample at (2/F) samples/second

Video

Basics

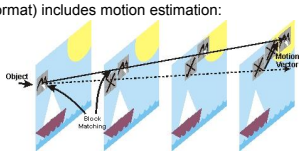
Can be viewed as a series of still images

- 24 frames per second (fps) in movies
- 30 fps in US television

Motion-JPEG (M-JPEG) is exactly this: JPEG image for each frame

- Benefit: Very simple format to work with and edit
- Drawback: Doesn't take advantage of temporal similarities between frames

MPEG (DVD format) includes motion estimation:



Video

A few more details...

Frames are no longer independent!

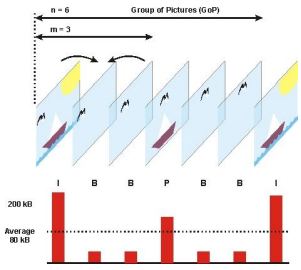
MPEG has three frame types:

- I-frames (intra-coded - independent)
- P-frames (predicted)
- B-frames (bi-predictive)

Must buffer B-frames until the next P-frame

Can only "enter" a video stream at an I-frame (or you see very blocky artifacts).

Video editors need to be very careful about this (splicing at non-I frames can be tricky!)



Video and Sound

A movie typically has multiple "streams" multiplexed together:

- Video stream
- Audio stream (maybe multiple for multi-language)
- Subtitles

Rendering software must synchronize streams - otherwise sound and video may be off (probably everyone has seen this happen!)

Compression

Taking advantage of redundancies and other structure to give smaller file sizes.

Two main types:

- Lossless: Allows perfect reconstruction of original data
  - Zip, RAR, FLAC, ... (JPEG has a lossless mode too!)
- Lossy: Reconstruction is an approximation of original
  - Most media formats: JPEG, MPEG, MP3, ...
  - Can usually trade off quality for compression

Note that digital sampling/capture is already a lossy process

(Remember taking advantage of human color vision?)

Compression

Examples, and what you can expect

Text: "Pride and Prejudice"

Original (uncompressed)	685 kB
Zip	250 kB
GZip	250 kB
RAR	217 kB
7Zip	204 kB
BZip	176 kB

Audio: "London Calling" (3:19 long)

CD audio (uncompressed)	35.2 MB
Zip (lossless, general)	33.9 MB
FLAC (lossless, audio)	25.4 MB
MP3 (lossy, 128 kbps)	3.2 MB
Ogg (lossy, quality 3)	3.1 MB

Notes:

- Zip is not designed for audio
- Both MP3 and Ogg sound good at this rate
- MP3 plays on almost all players
- MP3 encoding (using LAME) took 11.2 sec
- Ogg encoding took 6.1 sec

Compression

Examples, and what you can expect - cont'd

Picture: 3648 x 2736 (9.98 MPixel)

Raw	29.9 MB
Zip (lossless)	17.0 MB
BZip (lossless)	10.9 MB
PNG (lossless)	9.8 MB
JPEG (lossy - Q=95)	2.1 MB
JPEG (lossy - Q=85)	1.1 MB

Video: "Wizard of Oz" (1:41:42)
480x720 @30fps

Raw	190 GB
HQ DVD	3.6 GB

Notes:

- DVD compression is over 50:1
- DiVX / MP4 can give 200:1 or more

Summary

There's a *lot* more we could talk about

- Logarithmic scale of human perception (intensities, frequencies, etc.)
- Image formats: bitmapped vs vector formats
- Compression techniques
- Other imagery formats (multispectral images)
- ...

Explore this if it interests you! Following your curiosity is a great way to learn...