Audio

- Transforming signals from the analog realm into the digital realm.
- Storing and processing of audio numbers
- Transforming numbers back into sounds

Capturing Audio

Three Questions:

1: How many channels (independent audio streams)?
   - 1: Monaural
   - 2: Stereo
2: How “good” is each sample?
   - 1 Byte per Sample (256 levels, OK for voice)
   - 2 Bytes per Sample (65536 levels, OK for music)
3: How many samples per second are needed?

3: How many Samples per Second?

- 3A: What is the Highest Frequency that Humans can hear?
- 3B: How does Frequency map onto Samples per Second?
2-Second Sound Samples
- 1000 Hz
- 2000 Hz
- 4000 Hz
- 8000 Hz
- 12000 Hz
- 16000 Hz
- 18000 Hz
- 19000 Hz
- 20000 Hz
- 21000 Hz
- 22000 Hz

3A: What is the Highest Frequency that Humans can hear?
- 22KHz (22,000 cycles/second) as infants
- 12KHz in middle age
- <6KHz in old age, as hairs in ear canal stiffen with age and become less capable of vibrating at the higher frequencies.

3B: How does Frequency map onto Samples per Second?
- Answer: Nyquist Sampling Theorem
- Harry Nyquist (1889-1976), AT&T, Bell Telephone Laboratories
- You must sample at least twice the rate of the highest frequency you wish to capture.
- For example: to capture 1000 Hz with enough information to play it back correctly, you need to sample at least 2000 times per second.

3: How many Samples per Second?
- For Compact Discs:
  To capture 22 KHz we will need to sample at:
  44,000 samples/second/channel
  Actually, we use:
  44,100 samples/second/channel because 44,100 is $2^2 \times 3^2 \times 5^2 \times 7^2$ which is divisible by 2, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 18, 20, ...
Lossy Compression

- Perform Fourier Analysis (Spectrum Analysis) to convert samples into a sum of instantaneous frequencies,
- Discard high frequencies that some people can’t hear anyway,
- Discard audible frequencies that are being “shouted out” by other louder frequencies,
- Drop sampling rate on quiet passages.

Compression

- **.WAV** – Uncompressed (raw data as ripped directly from a CD)
- **.MP3** – Lossy Compression, roughly 10:1
- **.WMA** – Supported by Windows, Roughly 20:1 compression for same quality as .MP3
- **.AAC** – Supported by Apple

Thinking About Compression

Analyzing Audio in Spreadsheet Land

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Channels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Bytes per Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>44100</td>
<td>Samples per Second per Channel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8+8</td>
<td>Samples per Second</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8+8</td>
<td>Bytes per Second per Channel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>Seconds per Minute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8+8</td>
<td>Bytes per Minute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>Length of song in Minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>18*1</td>
<td>Size of song in Bytes (compressed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>Compression Factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>+0.3784</td>
<td>Size of song in Bytes (compressed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>700000000</td>
<td>Size of CD in Bytes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>+0.3784</td>
<td>Minutes of Music per CD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analyzing Audio in Spreadsheet Land

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Channels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Bytes per Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>44100</td>
<td>Samples per Second per Channel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8+8</td>
<td>Samples per Second</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8+8</td>
<td>Bytes per Second per Channel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>Seconds per Minute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8+8</td>
<td>Bytes per Minute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>Length of song in Minutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>18*1</td>
<td>Size of song in Bytes (compressed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>Compression Factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>+0.3784</td>
<td>Size of song in Bytes (compressed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>700000000</td>
<td>Size of CD in Bytes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>+0.3784</td>
<td>Minutes of Music per CD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Video

• Has all the issues associated with audio, and...
• ...much more!
• Video is a series of frames (images) rapidly shown on screen to simulate movement.
• Often has a linked audio track (maybe more than one).

Video Issues

• How Wide is each Frame (pixels)?
• How High is each Frame (pixels)?
• How many Bytes per Pixel (colors)?
• How many Frames per Second?
• Is the video Compressed in some way?

Frame Size (Bytes)

• Number of Pixels per Frame = Width × Height
• Number of Bytes per Frame = Pixels × Bytes per Pixel
  – 24-Bit RGB Color: 3 Bytes per Pixel
  – 256 Color: 1 Byte per Palette Index
  – 16 Color: ½ Byte per Palette Index
  – 2 Color: ¼ Byte (1 Bit) per Palette Index

Bytes and Frames

• Bytes per Second = Bytes per Frame × Frames per Second
• Frames per Second:
  – Traditional Movies: 24 frames per second
  – Old Analog TV: 30 frames per second
  – Modern HDTV: up to 120 frames per second
  – Minimum: 10-12 frames per second (any slower and you will see each individual frame)

An Example

• A 3½ minute Video which is 360 × 240 pixels (very small), 3 bytes per pixel, playing at 10 frames per second (very slow) would have:
  – Bytes per Frame = 360 × 240 × 3 = 259200
  – Bytes per Second = 259200 × 10 = 2592000
  – Length of Video = 3.5 × 60 = 210 seconds
  – Length of Video = 2592000 × 210 = 544,320,000 Bytes (uncompressed)
Another Example

- A 60 minute Video which is $1024 \times 768$ pixels (full screen), 3 bytes per pixel, playing at 30 frames per second (full speed) would have:
  - Bytes per Frame = $1024 \times 768 \times 3 = 2,359,296$
  - Bytes per Second = $2,359,296 \times 30 = 70,778,880$
  - Length of Video = $60 \times 60 = 3600$ seconds
  - Length of Video = $70,778,880 \times 3600 = 254,803,968,000$ Bytes (uncompressed)

Compression Helps!

- **.AVI** – Audio Video Interleave (Windows)
  - A frame of Video is followed by a chunk of Audio
  - Several compression methods
  - Every frame stored, played at constant rate
- **.MOV** – Quicktime (Apple)
  - Frames can play at variable rates
  - Multiple Audio tracks (secondary audio programs)
- **.MP4** – Variation of **MOV** for small devices

Compression Helps!

- **.MPG** – Motion Picture Experts Group
  - Sometimes **MPEG**
  - Start with key-frame (full image),
  - For next frames only the differences between current frame and previous frame are stored (much smaller amount of information).
  - Takes a LONG time to compress, but...
  - Very fast Playback (video over HDTV)

Cut to the Videos...