Session:  Software
Topic:  MP3 Audio Format

Daniel Chang
Sound
• Recall that sound is a continuous (somewhat bumpy) wave
• CD Audio format digitizes sound to make an approximation
• Even though loses some information, loss is not detectable
• But takes a lot of bits to get this quality
• CD Audio is typically extracted to files ending in ".wav"

Compression
• Other audio formats simply try to get the close to the same sound by using fewer bits
• CD Audio format can be "compressed" by throwing out some information that hopefully does not affect sound quality
• Generally, the more you throw out, the worse it will sound
• "Compression algorithms" are steps or processes used to perform the compression
MP3 Audio Format

- MPEG-1 Audio Layer 3 files, usually ending in ".mp3"
- MPEG-1 is a standards suite, including video and audio compression standards (used in Video CD)
- MPEG-1 suite includes three layers of audio encoding
- MP3 format tries to obtain compression by 1) using a lower bit rate and 2) using tricks

Perceptual Noise Shaping (Psychoacoustics)

- MP3 format uses characteristics of how the human ear operates to eliminate information that does not affect how the ear "hears" the music
- Some sounds cannot be heard by the human ear (out of range, unusual tones)
- Some sounds cannot be heard (well) by most human ears
- If multiple sounds are playing simultaneously, the louder ones can be heard while the softer cannot
### Units

<table>
<thead>
<tr>
<th>Term</th>
<th>Symbol</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>kilo</td>
<td>k</td>
<td>1000</td>
</tr>
<tr>
<td>Kilo</td>
<td>K</td>
<td>$2^{10} = 1024$</td>
</tr>
<tr>
<td>Mega</td>
<td>M</td>
<td>$2^{20} = 1024 \times 1024 = 1,048,576$</td>
</tr>
<tr>
<td>bit</td>
<td>b</td>
<td>(0 or 1)</td>
</tr>
<tr>
<td>Byte</td>
<td>B</td>
<td>8 bits (256 values)</td>
</tr>
<tr>
<td>Hertz (sampling rate)</td>
<td>Hz</td>
<td>samples/sec, [anything]/sec</td>
</tr>
<tr>
<td>Bit Rate</td>
<td>bps (kbps)</td>
<td>bits/sec, kilobits/sec</td>
</tr>
</tbody>
</table>
MP3 Sampling Rate
- Sampling rate is how often a sound sample is taken
- CD Audio used a sampling rate of 44.1 kHz (samples/sec)
- Exactly 16 bits are used for each sample, resulting in a "bit rate" of 1378 Kbps (1411 kbps)
- MP3 allows for 32, 44.1, and 48 kHz sampling rates (typically use 44.1 since same as CD Audio)

MP3 Bit Rate
- However, for MP3 specific "bit rates" are used, which changes the amount of bits that may be used for each sample
- Any single sample may use 16 bits or even more, but some must use less so that overall the total bits per second are not exceeded.
- Sometimes just use a constant number of bits per sample
- Possible MP3 bit rates are 32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256 and 320 kbps.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Bit Rate = Bit Depth × Sampling Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>units</td>
<td></td>
</tr>
<tr>
<td>Bit Rate</td>
<td>Bit Depth</td>
</tr>
<tr>
<td>CD Audio</td>
<td>1,411,200 bps (a result)</td>
</tr>
<tr>
<td>MP3</td>
<td>[varies] (specified)</td>
</tr>
</tbody>
</table>

- Lower bit rates result in smaller files, but lower audio quality
- Lower bit rates can also result in "compression artifacts", which are sounds that are not even present in the original but occur due to failed "encoding"
Variable Bit Rate (VBR)
- Achieve better compression by specifying a maximum bit rate, but allowing the rate to vary during music
- Bit rate can fall lower when less information in sound, thus saving space
- Detectable by an annoying variation in display

General Observations
- 128 Kbps is the de facto minimum acceptable quality, resulting in a compression ratio of about 11:1
- 192 Kbps is starting to become more popular as broadband speed and drive storage increase

<table>
<thead>
<tr>
<th>Quality</th>
<th>Excellent</th>
<th>Very Good</th>
<th>Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit Rate (Kbps)</td>
<td>224 - 320</td>
<td>192 - 224</td>
<td>128 - 192</td>
</tr>
</tbody>
</table>

Comparison
- 4 kbps - minimum necessary for recognizable speech
- 8 kbps - telephone quality
- 16 kbps - minimum videophone quality
- 96 kbps - FM quality
Encoders

- Ripping from CD Audio involves compressing CD Audio format into MP3 format
- Sampling rate may not matter if encoder is poor
- Encoders vary in
  1. How fast they work to be able to sample data at a given bit rate (if they do not keep up, they lose information)
  2. How correctly they even perform the MP3 compression algorithms (software has bugs, in case you did not know)

<table>
<thead>
<tr>
<th>Encoder</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAME</td>
<td>Created by Mike Cheng in 1998</td>
</tr>
<tr>
<td></td>
<td>Fully LGPL and steps all over even new stuff</td>
</tr>
<tr>
<td>Fraunhofer Gesellschaft</td>
<td>Your results will vary</td>
</tr>
<tr>
<td>Xing</td>
<td>Used to come with video player</td>
</tr>
<tr>
<td></td>
<td>No longer widely used</td>
</tr>
</tbody>
</table>

MP3 Contributions

- Space savings - allows storage as well as faster transfer
- Portability - MP3 players, fit more onto single data CD
- Editing - once digital, can manipulate with a computer

MP3 Limitations

- Maximum bit rate is 320 kbps
- "Lossy" compression, where some information is discarded in order to reduce size
- Compare to "Lossless" compression formats
  1. FLAC - Free Lossless Audio Codec
  2. Monkey's Audio (.ape)
  3. WavPack (.wav) - compression of CD Audio format
  4. Many more
Metadata

- Information about Information
- The actual information is what you want, metadata simply describes features of the information, but is not the actual thing you are interested in

<table>
<thead>
<tr>
<th>&quot;Information&quot;</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movie</td>
<td>Show time, cast, rating</td>
</tr>
<tr>
<td>Text in a book</td>
<td>Author, Number of pages, Genre</td>
</tr>
<tr>
<td>Online purchase</td>
<td>Time of purchase, purchase history, similar purchases</td>
</tr>
</tbody>
</table>

- Metadata is more prevalent due to computing power
- Previously could barely keep the information itself, now can process facts about the information
- Metadata has become more important than data

ID3 Tag

- MP3 files can store tags, which can contain metadata about the actual music
- Some example items are "Title", "Artist", "Album", "Track Number" and "Genre"
- Since the information is digital, a computer can work with it - sorting, selecting based on parameters
- There are versions "ID3v1" and "ID3v2", and a new "APEv2" which was originally developed for the "MPC" file format