Week 5: Digital Audio

DIGITAL ASSET DEVELOPMENT

Contents

- Digital sound
- Digital audio parameters
- Working with sound

Sound Waves

- Sound is the vibration of air (or some other medium) in response to pressure
- Sound waves are longitudinal - vibration is back and forth along direction of travel
- As opposed to transverse waves, which vibrate at right angles to wave direction
  - Waves in the sea are transverse
- Sound is generated by a vibrating string or surface (vocal cord, loudspeaker)

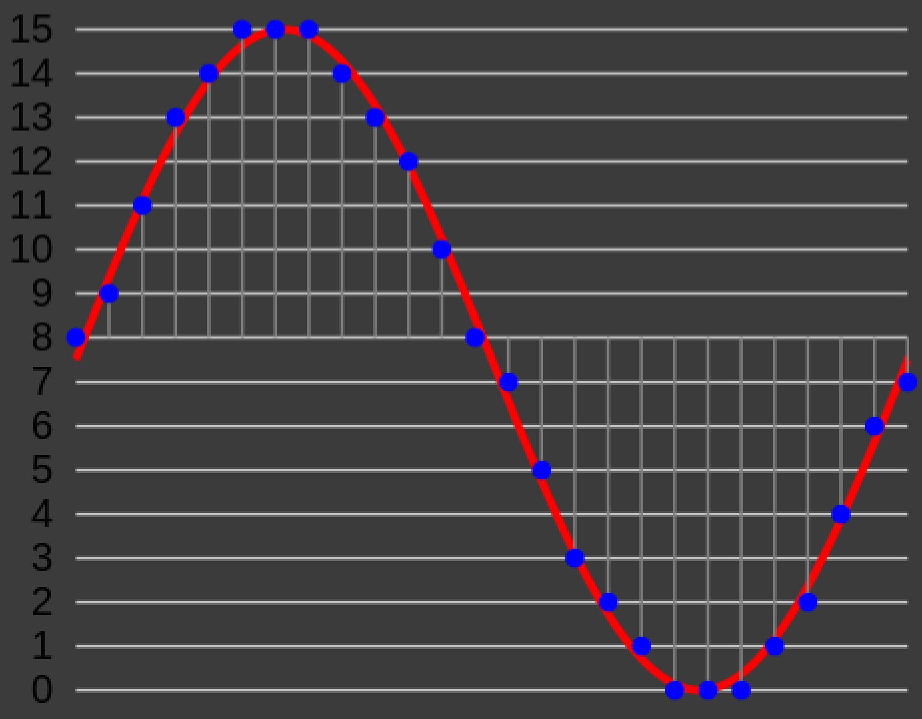
Pitch and Intensity

- The pitch of a sound corresponds to sound wave frequency
  - Number of vibrations per second (Hertz)
  - Natural sounds are a combination of tones at various pitches
  - Musical notes have harmonics – frequencies that are multiples of the basic note
- The loudness of a sound is its intensity
  - Corresponds to wave's amplitude (height)

Digital Audio Data

- Digital audio is a representation of sound waves in digital form
- Data consists of samples taken at regular time intervals
  - Most common technique is known as pulse code modulation (PCM)
- Data quality will depend on:
  - How often we acquire data samples
  - Precision of these samples

PCM Sampling (4 bit)



"Pcm". Licensed under CC BY-SA 3.0 via Commons -
https://commons.wikimedia.org/wiki/File:Pcm.svg#/media/File:Pcm.svg

Digital Audio Parameters

- For raw audio data we have three major parameters:
  - Sample rate: number of samples taken per second
  - Bit-depth: number of bits needed to represent each sample
  - Number of audio channels – mono, stereo or multichannel sound
- As usual, higher numbers give better quality but generate larger data files

Sample Rate

- There is underlying theory that defines suitable sample rates for given sounds
- Assume we are digitising sound where *n* is the maximum frequency present
  - A sample rate of *2n* is sufficient to represent that sound in full (called the Nyquist rate)
  - A lower rate will lose some of the top end frequencies
  - This may cause some degradation in sound quality

Bit Depth

- As in other areas of digital data, certain bit depths are favoured
- For a given audio channel, typical values are 8, 16 or 24 bits per sample
- 8 bit audio is associated with retro sound effects
- CD data uses 16 bit audio
- Specialist audio production tools often work with 24 bits

Multichannel Audio

- Stereo sound requires left and right audio channels to be present
- Allows separation of instrumentation in a musical recording
- Also enables sound to be panned between speakers
  - Useful for some in-game audio
- More channels allows possibilities for immersive “surround sound”

Human Hearing

- Sound processing tasks are generally aimed at enhancing audio quality
- Must bear in mind the target system
  - In this case, the human auditory system
- Human hearing has evolved to perceive certain types of sounds very effectively
  - Most importantly, speech
  - We can also handle a wide dynamic range
  - Our hearing is directional (stereo effect)

Human Audio Perception

- Human hearing is usually defined as covering the range 20 - 20000 Hz
  - The top end decreases with age
  - Hence “mosquito” devices for driving away teenagers!
- Dominant frequencies for human speech are 100 - 3000 Hz
- We can also “filter out” quite high levels of background noise when required

Typical Parameter Settings

Sample Rate	Resolution	Channels	Output
192 kHz	24-bit	6.1	DVD-A / SACD
44.1 kHz	16-bit	2	CD quality
44.1 kHz	8-bit	2	decent PC audio
11.025 kHz	8-bit	1	basic effects

- Uncompressed CD-quality sound uses up around 10 Megabytes per minute
- The lowest quality setting above uses ~0.6 Megabytes per minute

Audio Compression

- In practice, most audio distributed today is significantly compressed
  - Saves greatly on file space and bandwidth
- Compression levels are as important as sample rate / bit depth for audio quality
- As with image data there are a variety of compression formats available
  - Most use variants of standard methods
  - Details vary depending on the type of sound

Frequency-based Compression

- Audio compression algorithms usually take the following basic approach:
  - Break down the sound waves into their component frequencies
  - Remove or suppress those frequencies which humans will not hear (much!)
  - Store quantised frequency data
- This has many similarities to the JPEG transform coding method for images

Working With Audio

- We use sound in games and other media for (among other purposes):
  - Background music
  - As a response to user actions
  - Sound effects
  - Speech (dialogue, narration,...)
- In each case:
  - Audio must be of appropriate quality
  - Sound should be correctly synchronised

Audio Editing Tasks

- Common audio tasks in animation and games production include:
  - Trimming or splicing sound clips
  - Converting clips into appropriate format and quality level
  - Adjusting pitch, tempo or other sound characteristics
  - Applying effects
  - Mixing together disparate audio tracks
  - Panning sounds for “3D” effects

Audio Editing Tools

- Audacity is freeware and has a good range of editing features
  - This week and next in the lab
- Adobe Audition is more fully featured, but less intuitive (and costs more)
- At the pro level there are a range of specialist tools for:
  - Music composition and production
  - Sound design for film and games