Non-Speech Sound

Phil Gray

Non-Speech Sound

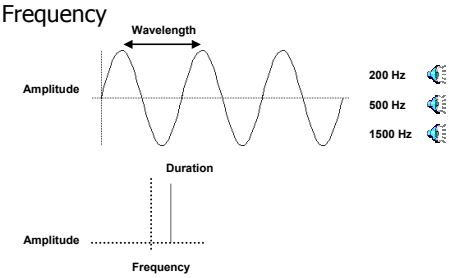
- What is sound?
- Sampling sounds
- Synthesizing sounds
- MIDI
- Sound in User Interfaces

What is Sound?

- Sounds are pressure variations that propagate through an elastic medium (e.g. air)
- Sounds are structured at their source, from which they inherit their characteristic properties:
  - amplitude (dB)
  - frequency (Hz)
  - duration
  - timbre
  - directivity

What is Sound?

- Frequency



How Do We Hear?

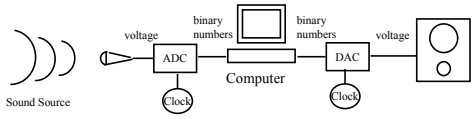
- Ear has three parts
  - inner ear
  - middle ear
  - outer ear
- Outer ear
  - Pinnae and ear canal
- Middle ear
  - ear drum connects to inner ear
  - amplification
- Inner ear
  - Cochlea - vibrations stimulate auditory nerve

How Do We Hear?

- Humans can hear 20Hz to 20kHz frequency range
- Can hear differences of around 1.5Hz
- Hearing loss occurs with age
  - age 50 max 14kHz, at age 70 max 10kHz
  - important for interface design

Digital Sound Recording and Playback

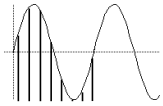
- Convert *analogue* signal into *digital* signal for storage and manipulation ("sampling")
- Convert digital signal to analogue for playback
- Sampler may be PC or dedicated h/w sampler



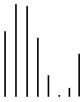
Sampling - Sample Rate

- Signal measured at a set of distinct times - sample rate

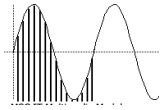
• 8 samples per cycle



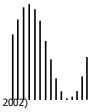
• Resulting sample data



• 16 samples per cycle



• Resulting sample data



Sampling - Aliasing

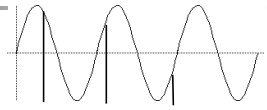
- Higher the sampling rate, the higher the match between the original signal and that reconstructed from sample data
- Sample rate must be greater than 2 x frequency
  - aliasing problems
  - distortion

Aliasing

- Good sampling rate is 44,100 samples/second
  - maximum frequency is therefore 22,050Hz
  - humans can only hear up to ~20,000Hz
  - can therefore filter sound source to avoid aliasing
- New DVD Audio specification
  - 192kHz sample rate / 24bit sample size (max)
  - max frequency 96,000Hz

Aliasing

- 10Hz sound, 5 samples/Sec.



- Resulting sample data

Sampling - Sample Size

- Larger the bit size the better the amplitude range (*dynamic range*)
  - usually 8 bit (only 256 possible values)
  - or 16 bit (65,535 possible vales)
- Quantization problems
  - with 8 bit numbers cannot represent many amplitude levels
  - can result in *noise*
- Lower quality can affect usability

Sample Formats

- Many, many formats for sampled sound
- CD quality 44.1kHz, 16bit (per channel)
  - 44,100 x 2 x 60 = 5.2 MB/minute
  - .wav for PC, AIFF for Mac
- Medium quality 22kHz, 8 bit
- Low quality 11kHz, 8 bit, mono
  - 11,000 x 60 = 660KB/minute

44.1kHz, 16bit
11kHz, 8bit
5.5kHz, 8bit

Sample Formats

- New DVD Audio specification
  - 34MB/min per channel (max)
- Telephony quality 8kHz, 8bit, mono
  - mu-law : samples encoded logarithmically in 8 bits -> 12 bit linear range
- See the audio FAQ for more formats

Audio Compression

- Want high quality sounds but need to reduce size
- MP3 is now common
  - MPEG-1 (Motion Pictures Expert Group) Audio Layer 3
- Uses a range of perceptual coding techniques to reduce file size but keep quality
  - lossy compression

MP3 Compression

- Main components
  - minimum audition threshold
  - masking effect
  - joint stereo
  - Huffman coding

MP3 Compression

- Can reduce file sizes by a factor of 12
- Bitrate - number of bits for one second of audio

Bitrate	Quality	MB/min
1411	CD	10.584
192	good CD quality MP3	1.440
112	near CD quality MP3	0.840
64	FM quality MP3	0.480
32	AM quality MP3	0.240
16	Shortwave quality MP3	0.120

Sound Synthesis

- Many synthesisers available
  - mostly controlled by MIDI (see below)
  - maybe separate h/w, a sound card or s/w synthesiser built into PC
- Many different types of synthesis techniques
  - additive synthesis
  - FM synthesis
  - wavetable
  - physical modelling

MIDI - Musical Instrument Digital Interface

- Allows real-time control of electronic musical instruments
  - synthesisers, samplers, etc.
- Specifies a h/w interconnection scheme + protocol for data communications + grammar for encoding musical performance data

MIDI

- MIDI data is like a 'piano roll' - gives note on, off, instrument info
- Much less information is contained than for samples
  - MIDI files much smaller than sample files
- MIDI does not encode timbre so synthesis is left up to synthesiser
  - things may sound different when played back

Advantages of MIDI

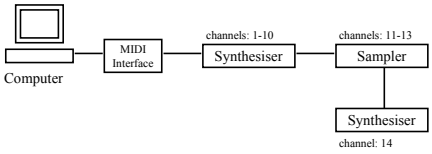
- Allows play back of sounds on many different types of synthesiser
  - almost!
  - General MIDI
- Separates i/p device from sound generator
  - one keyboard can play many synths
- Many different types of i/p device can be used
  - piano keyboard, computer, MIDI guitar, etc.
- MIDI can control a wide variety of devices
  - synthesiser, sampler, audio effects, lighting

MIDI Setup

- PC (maybe piano style keyboard) used to control the MIDI system
- S/w running on PC will generate MIDI messages
  - might be a sequencer or other s/w
  - these sent to synthesiser (or sampler) to play sound

MIDI Setup

- A standard MIDI setup might look like:



Sound in Interface Design

- More natural information representation
  - 'sonification' of high dimensional data
  - spatialized audio for more immersive displays
- Essential for eyes-free applications
  - aids for blind people
  - compact mobile devices
  - telephone displays

Why Use Sound in User Interfaces?

- Why use non-speech sound ?
  - interdependence of vision and hearing
  - natural means of presentation
  - reduce the load on the visual sense
  - increase the bandwidth of communication
  - omni-directional
  - attention grabbing
  - your eyes can only do one thing at once

Resources

- The computer music tutorial, Roads, C. (1996), MIT Press
  - Lots and lots of info
- Roads and Strawn. (1985) Foundations of computer music. MIT Press
  - old but good
- There are many WWW sites on audio, a good one is
  - mitpress.mit.edu/e-journals/Computer-Music-Journal/CMJ.html

Resources

- See book chapter on auditory interfaces on Web site
- Lots of work done at Glasgow
  - www.dcs.gla.ac.uk/~stephen/publications.shtml
- MIDI information
  - www.midi.org
- Audio in interfaces
  - www.dcs.gla.ac.uk/~stephen/otherlinks.shtml