Sound 1

Sound in Multimedia and HCI

Sound in Multimedia and HCI

Overview

- Sound in multimedia and user interfaces
- The physical characteristics of sound
- The psychological characteristics of sound
- Fidelity and interpretation
- Sound file formats

Nature of Sound Interfaces

Like graphics, sound can be output from the computer

- to a loudspeaker
- to headphones

Sound also can be an input

- usually speech input
  - text input
  - speaker recognition
- even Auditory Scene Analysis (cf Visual Scene Analysis)

HCI is not necessarily the end…

- why keep computers sessile with their input mediated by users?
- Instead, move towards machines that "understand" (or at least interact appropriately with) their environment.

Sound as Input (I)

Infrequently used… but clearly possible and useful

Speech input issues:

- language, dialect
- single or multiple user?
- isolated words or continuous speech?
- vocabulary size

Command interfaces are generally simpler

- handling applications or commands inside an application
- generally small (known) vocabulary
- form is usually <verb> <noun>
  - verb might be run, stop, hide, minimise, maximise (for applications)
- e.g. run firefox, hide word, select excel etc.

Sound as Input (II)

Inside an application (once selected)

- verb might be print, save, open, new,
- note that vocabulary may be updated for each new environment
  - e.g. open letter3; goto A67
- in general, the range of commands is defined by the nature of the application so although there may be a larger vocabulary altogether, the system knows which verbs (and nouns) to expect
- some applications have quite a large range of possible commands

Free text interfaces for WP input are altogether harder

- much larger possible vocabulary
- very difficult to use context to limit vocabulary effectively.

Other issues, like the joker who shouts "delete all files!" from the doorway...

Sound as Output

Very common usage, earliest interaction being a simple beep

- usually meaning "something is wrong"
- could be anything
- uninformative!

… to much more sophisticated systems

- but what form should they take?
- Spoken output?
  - "Your printer is out of paper"
  - telephone synthesised speech/response systems
    - Telephone banking systems, ticket buying systems, …
    - Generally multi-user, and keep the vocabulary size low
- sound (non-speech) output?
  - Musical sounds, everyday sounds?

Sound Associations

- Many sounds have associations, these may be obvious and usable
  - breaking glass
  - scream
  - door slamming
- or may be personal (and different for each individual)
  - dog barking
  - keys rattling
- and they may be culturally determined as well
  - a champagne cork being pulled
  - cicadas chirping

Sound in game worlds (I)

Game worlds are virtual worlds

- or a form of interactive simulation
- or a form of virtual reality
- and sound is part of normal reality

So, appropriate sound:

- adds to the "reality" of the game world
- can make the interactive simulation feel more interactive
  - as the user performs some action the sound alters

Sound in game worlds (II)

Sound effects: Bangs and crashes

- identifying what has happened
- who has shot (in shoot-em-up games)
  - and whether they hit or missed
  - and what have they shot
- try pinball...

Sounds may be musical or everyday

- sounds must be of short duration, otherwise they slow down the pace of the game

Games were amongst the first applications to go beyond the simple beep

Sound in game worlds (III)

Sounds identify or provide information about all the entities and events. Identity may be associated with a particular pitch or timbre. The timing of events may be associated with a particular sound being emitted

With stereo sounds, location information can also be provided

- many advanced sound cards have on-board digital signal processing hardware to perform convolutions

Sound is used in conjunction with graphical output

- to provide the game player with the information needed to play
- the better this information, the more effective the simulation

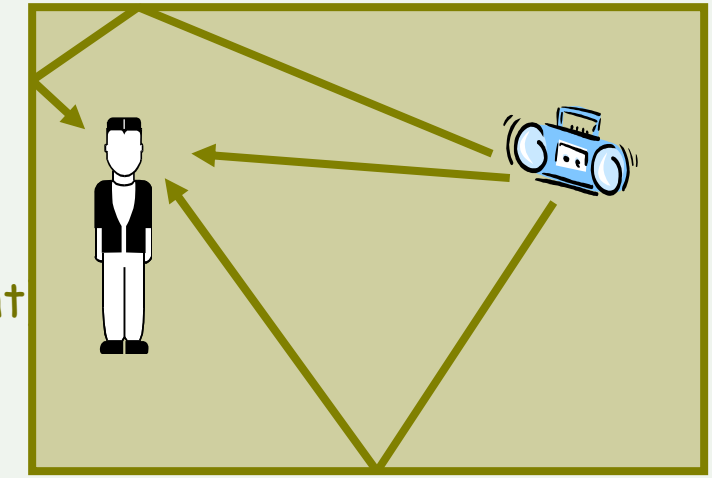
Sound output strengthens the illusion of reality.

Distance and Direction of a Sound Source

- We can often also tell (roughly) the direction and distance of a sound source
- This comes partly from the loudness of the sound, and partly from other characteristics of the sound - e.g. number and form of reflections

Physical correlates are:

- Loudness
- Reflections
- Spectral shape (frequency content



Room Acoustics

Why does a particular sound seem different in different environments?

- Because of the reflections that the sound makes off the elements of the room
- walls, floor, ceiling, furniture, etc.

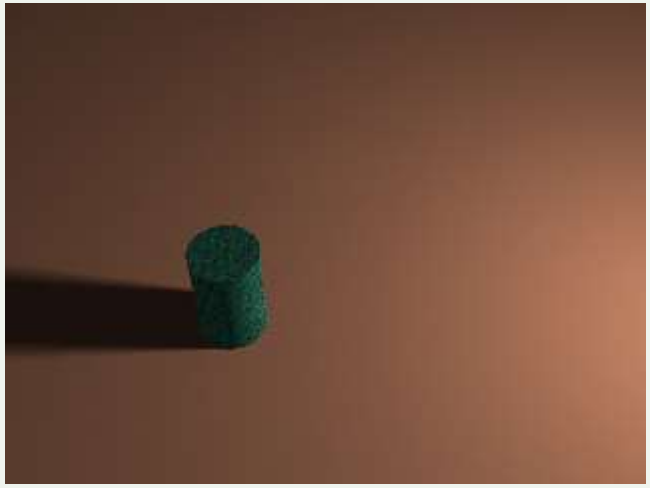
These can be modelled as well, so that the illusion that a sound comes from a particular environment can be created.

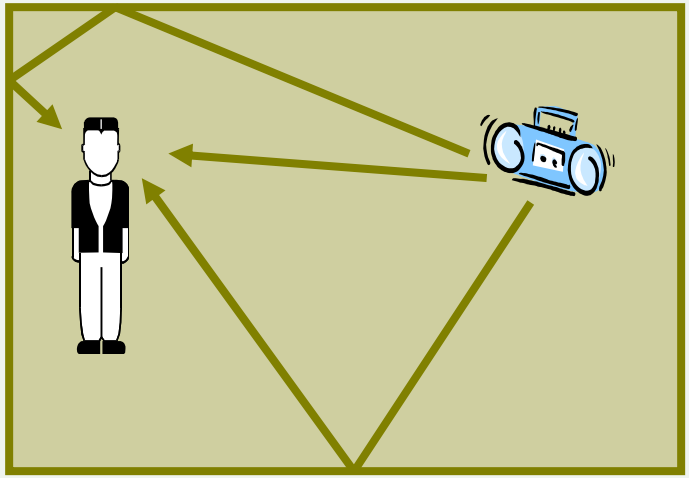
How is this achieved?

Signal processing and soundscapes

Realistic images use accurate models of light reflections

Realistic (located) sounds use accurate models of sound reflections. The processing used is based on linear signal processing theory





Data Auralisation or Sonification

Generating sound output from datasets

- (excluding reading text aloud)

why?

- The auditory system is good at picking out patterns
- possibly even better than eyes

Example: see

- http://www.artificialvision.com/javoice.htm

Next we look at the physical characteristics of sound…