3D SOUND

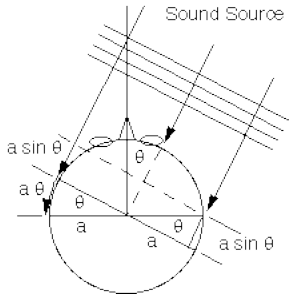
Cues
Head-Related Transfer Functions
Room Models
Speaker Arrays

3D Sound

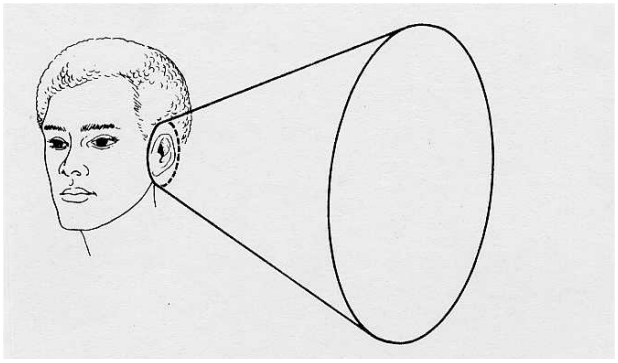
- Various cues to sense direction and distance
  - Inter-aural time delay (direction in horiz plane)
  - Spectral cues (elevation)
  - Reverberation : Direct Sound ratio (distance)
  - Spectral cues (amplitude, distance)
- Computer systems can simulate some of these cues

Head-Related Transfer Functions

- Richard Duda, "3-D Audio for HCI" at
  - http://interface.cipic.ucdavis.edu/CIL_tutorial/3D_home.htm
- Duplex Theory – Lord Rayleigh
  - Interaural Time Difference (ITD)
  - Interaural Level Difference (ILD)
    - ILD is frequency dependent



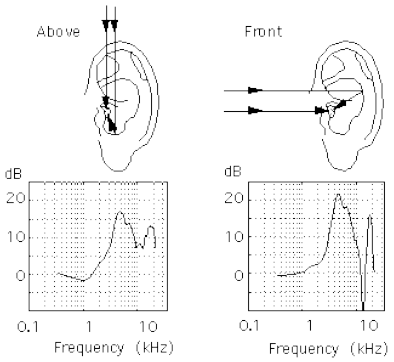
Cone of Confusion



Source: http://www.owlnet.rice.edu/~psyc351/Images/ConeOfConfusion.jpg

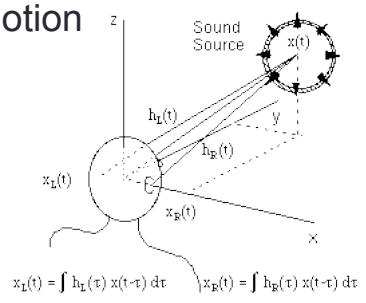
Elevation Cues

- Outer Ear or Pinna
- Pinna reflections create interference that is frequency-dependent



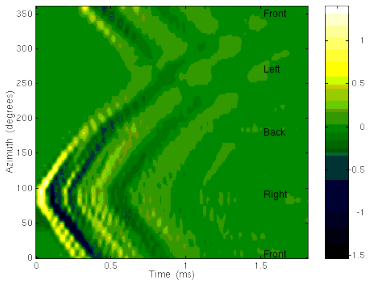
HRTF: Head-Related Transfer Function

- Measure effect of head on sound to each ear as a function of source direction: HRTF
- Apply HRTF to "position" a sound in space
- Interpolate HRTF for sound motion

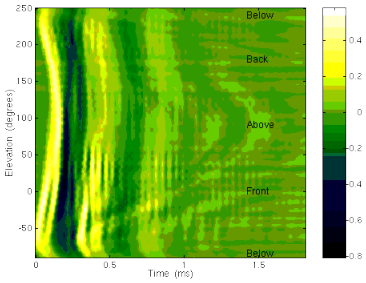


HRIR – Head Related Impulse Response

- Horizontal Plane



- Median Plane



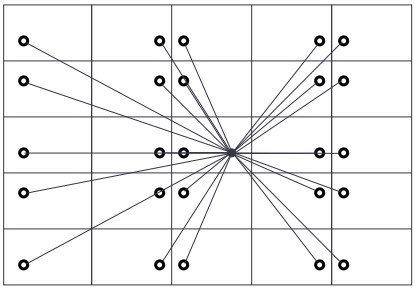
HRTF, Headphones, and Head Tracking

- Present different signal to each ear → headphones
- Sounds must change when you move your head → Use head tracking to update HRTFs

- Often combined with head-mounted display for VR

Room Models

- Ray tracing
- Audio has long wavelengths
- Early reflection models can convey room geometry



Doppler Shift

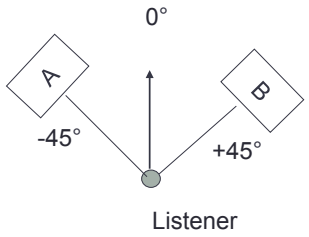
- Distance adds delay
- Change in delay creates doppler shift
- Can also simulate doppler shift using frequency modulation, etc.
- What about reflections?

Panning

- Linear panning:

$$A_{amp} = (45 - \theta)/90$$

$$B_{amp} = 1 - A_{amp}, \quad -45^\circ < \theta < +45^\circ$$



At 0°,
sum of power = $A^2 + B^2$
= ¼ + ¼ = ½, but
at 45° or -45°,
sum of power = 0 + 1 = 1

- "Hole in the Middle" effect

Constant Power Panning

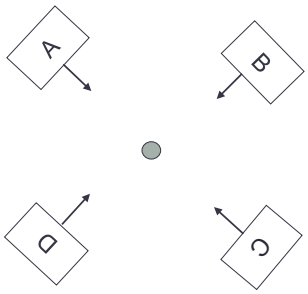
$$A_{amp} = \frac{\sqrt{2}}{2} \times [\cos(\theta) + \sin(\theta)]$$

$$B_{amp} = \frac{\sqrt{2}}{2} \times [\cos(\theta) - \sin(\theta)], \quad -45^\circ < \theta < +45^\circ$$

$A^2 + B^2 = 1$ for all angles

Multi-speaker playback

- Generally use speaker placement for direction
- Use panning between speakers for direct sound, with Doppler shift
- As distance increases,
  - Intensity decreases
  - Reverb decreases more slowly
- One reverberator feeds all channels because reverb is diffuse, non-directional



Loudspeaker Orchestras

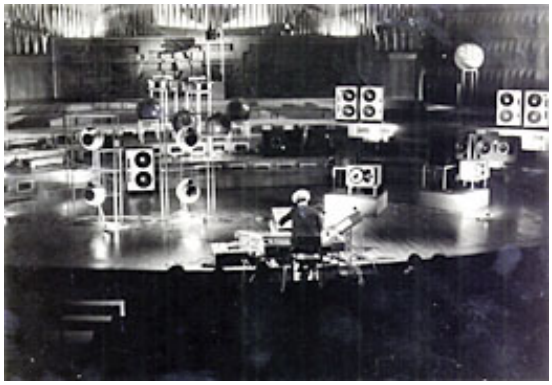
Francois Bayle (seen from behind) seated at the Acousmonium control panel in the Salle Olivier Messiaen, Maison de Radio France, Paris, in 1980.
Photo courtesy GRM, from EMF website (http://emfinstitute.emf.org)

Wavefield Synthesis

- Large array of speakers to reproduce the wave front of an imaginary 3D or 2D scene
- Technical University of Berlin: 832 channels, 10cm spacing + woofers at 40cm

http://www.four-audio.de/en/references/wave-field-synthesis.html

Loudspeaker Orchestras



Francois Bayle (seen from behind) seated at the Acousmonium control panel in the Salle Olivier Messiaen, Maison de Radio France, Paris, in 1980.
Photo courtesy GRM, from EMF website (http://emfinstitute.emf.org)

Summary

- Multiple cues for sound location, distance
- HRTF's model direction cues, but require headphones
- Multiple loudspeakers provide multiple point sources, but usually in a plane and panning is a crude approximation of ideal perceptual cues
- Wavefield Synthesis solves many problems, but very expensive even for 2D
- Loudspeaker orchestra "accepts" speakers as legitimate sound sources rather than surrogates