CMSC 498M: Chapter 11a
Sound in Games

Source:

- Lecture notes from Prof. Ramani Duraiswami at UMCP.

Overview:

- Nature of sound and human hearing.
- Audio design in games.
- Audio engines and the aural pipeline.

Overview

- Nature of sound and human hearing.
- Audio design in games.
- Audio engines and the aural pipeline.

The Nature of Sound

What is Sound?:

- Vibratory energy caused by movement of physical objects.
- As mechanical wave energy, it requires a medium such as air or water in which to propagate.

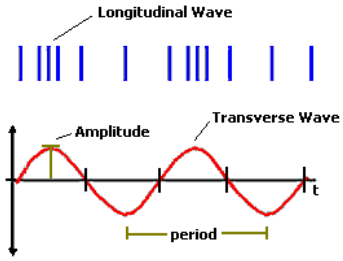
Properties:

Frequency: **Rate** of vibration.

- Experience as **pitch** (high or low).
- We hear 20-20,000 Hz (cycles/sec

Amplitude: **Intensity** of vibration.

- Experience as **loudness**
- Measured in **decibels** (dB) (too loud too long ⇒ hearing loss)



The Nature of Sound

Frequency:

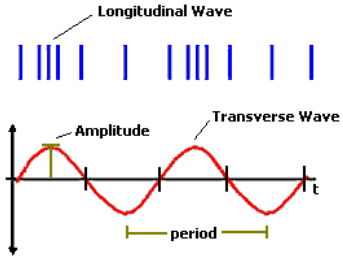
- Number of sound pulses that travel past a fixed point in a second.
- Measured in **Hertz** (Hz), that is, **cycles per second**.
- **Humans**: 20-20,000 Hz; **Bats**: 100,000 Hz.

Speed of Sound:

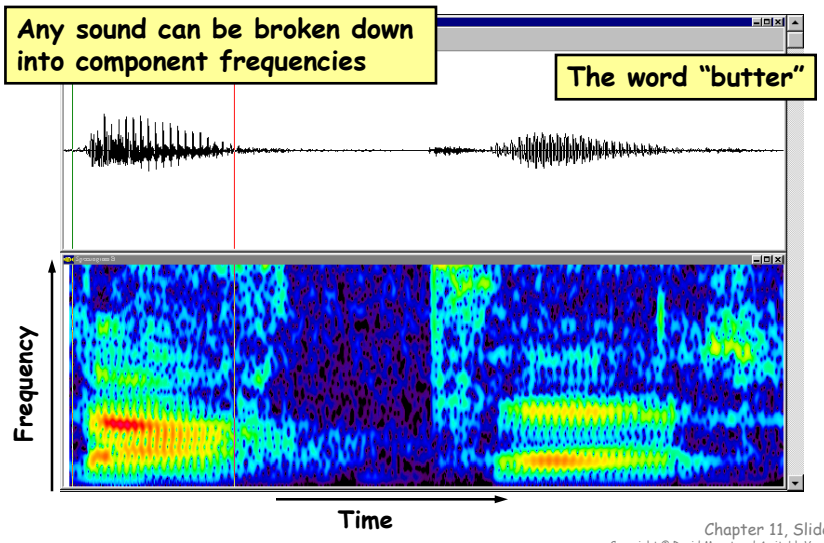
- How fast sound pulse travels.
- Sound travels at **same speed** in a given medium, irrespective of pitch.
- **344 meters/sec** in air.

Pitch:

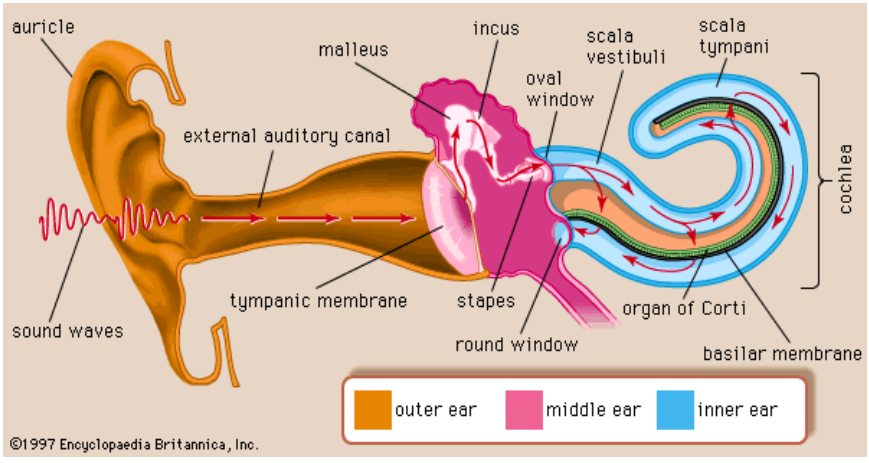
- As with color in vision, this is **perception**.
- A **high frequency** sound is heard as a **high pitch**.



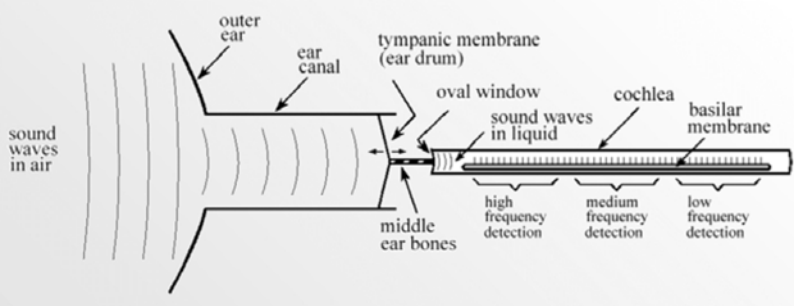
The Nature of Sound



Human Hearing



Human Hearing



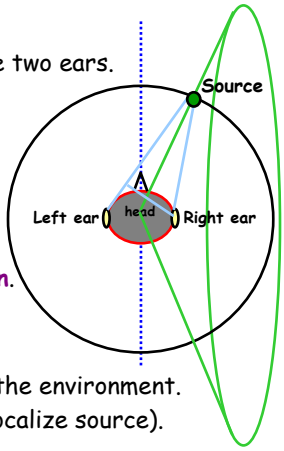
- **Outer ear** collects sound waves from the environment.
- **Ear drum** → **middle ear bones** → **fluid-filled cochlea**.
- 12,000 nerve cells form the **cochlear nerve**.
- Each nerve cell only responds to **narrow range** of audio frequencies.
- The ear is a **frequency spectrum analyzer**!

How do we perceive sound location?

Initial idea:
- Measure attributes of received sound at the two ears.

Compare sound received at two ears:
- Interaural **Level** Differences (ILD)
- Interaural **Time** Differences (ITD)

Surfaces of constant Time Delay:
- $| x - x_L | - | x - x_R | = c \, \delta \, t$.
- Delays same for points on **cone-of-confusion**.
- Level differences are **small**.



Other mechanisms necessary:
- **Scattering** of sound off our **bodies** and off the environment.
- **Purposive motion** (e.g. moving left-right to localize source).

Sound and Human Spaces

Sound Wavelengths:

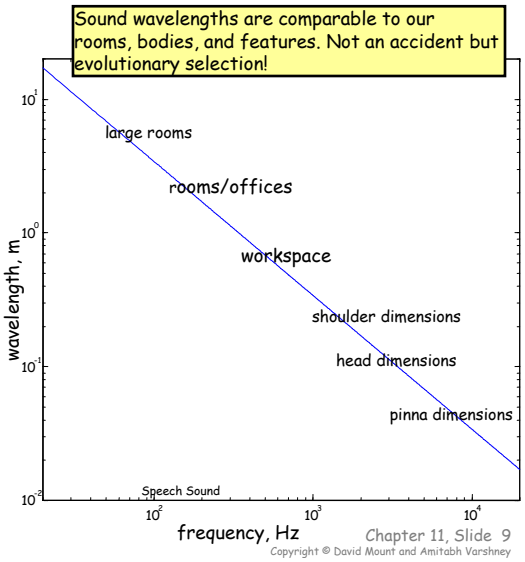
- Comparable to **human dimensions** and of the spaces we live in.
- $f\lambda = c$.

Cases:

$\lambda >> a$: wave is **unaffected** by object.

$\lambda \sim a$: behavior of scattered wave is complex and **diffraction effects are important**.

$\lambda << a$: wave behaves like a **ray**.



Modeling Sound Scattering

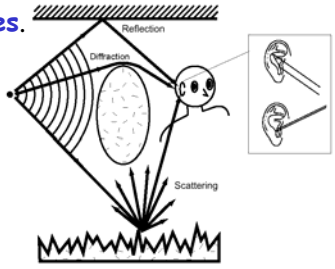
Interactions change received sound waves.

Scattering of body and ears:

- Bodies ~ 50 cm
- Heads ~ 25 cm
- Ears ~ 4 cm
- Not much multiple scattering

Scattering off surroundings:

- Rooms ~ 2m – 10m
- More multiple scattering
- Larger sizes → lower frequencies



Overview

- Nature of sound and human hearing.
- Audio design in games.
- Audio engines and the aural pipeline.

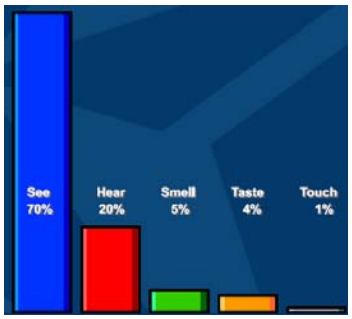
Sound in Games

Role of Sound in Games:

- "[Sound] serves the **story**, creates a **mood**, … and can be the key to bringing the **visuals to life**." [LoBrutto,1994]
- "Game audio is **judged** against all audio played on that system. We must not just meet those standards but **exceed them**." [Marty O'Donnell 2002, about Halo]
- **20% of perception is acoustic**. [Dobbler et. al. 2002]

Importance of Sound:

- Screen space is **limited**.
- Can hear what is going on around you and what is approaching you from **off screen**.



Audio Elements in Games

Sound Effects:

- Indicates **what is going** on in the game.
- May alert player to a **significant event**.
- Broadly defined, may include **background sounds** and **musical score**.

Narration:

- **Reinforces action**.
- Explains **non-obvious events** and story.

Sound Effects

Background (BG or ambient) sound effects:

- Do not explicitly synchronize with the picture. **Indicate setting**.
- Examples: Wind, rain, traffic, crowd noises (called "walla").

Hard sound effects:

- **Common sounds** that appear **on screen**.
- Examples: Door slams, weapons firing, and cars driving by.

Foley sound effects:

- Sounds that **synchronize precisely** with action on screen.
- Require the **expertise of a Foley artist** to record properly.
- Examples: Footsteps, punching, rustling of cloth.

Design sound effects and Background score:

- Sounds that **do not occur in nature** or impossible to record.
- Commonly used in **science fiction and fantasy** genres.
- Use of **music** to create an emotional mood.

Foley Artists

Ever wondered?

- Why footsteps sound louder in movies?
- Why they echo so much?
- Why the clanging of swords is so ringing and metallic?

Answer:

- **Foley artists**.
- Named after Jack Foley, a famous 1930's sound engineer.

Web sites:

- http://www.marblehead.net/foley/index.html
- http://www2.digidesign.com/digizine/-dz_main.cfm?edition_id=142&navid=1053

Foley Artists

Effect	How It's Made
Galloping horses	Banging empty coconut shells together
Kissing	Kissing back of hand
Punching someone	Thumping watermelons
High heels	High heels on wooden platform
Bone-breaking blow	Breaking celery or bamboo or twisting a head of lettuce
Footsteps in snow	Squeezing a box of corn starch
Star Trek sliding doors	Pulling a piece of paper from an envelope
Star Wars sliding doors	Flare gun plus sneakers squeak
Bird flapping its wings	Flapping a pair of gloves
Grass or leaves crunching	Balling up audio tape

Spoken Dialog

Two approaches:

- Text to speech (analogous to procedural animation).
- Sound capture (analogous to motion capture).

Text to speech:

- Sound is **artificially synthesized**.
- May sound **artificial** or **computer-like**.
- **Significant progress** over the last decade in generating natural sounding dialog from text.
- It is still **hard to convey emotion** and patterns of intonation realistically.

Examples:

- http://www.microsoft.com/speech/default.mspx
- http://www.research.att.com/~ttsweb/tts/demo.php
- http://www.neospeech.com/

Spoken Dialog

Sound Capture:

- Hire **actors** to read a script.
- Approach that is mostly followed in **expensive games** and **animated movies**.
- **Good vocals** can make up for **weak animation**.
- Pixar and Disney movies feature **top talent** for voices.

Overview

- Nature of sound and human hearing.
- Audio design in games.
- Audio engines and the aural pipeline.

Aural Rendering Pipeline: Goals

3D localization:

- Head-related impulse response (**HRIR**).
- Head-related transfer function (**HRTF**). Fourier transform of HRIR.
- Captures all the physical cues to **source localization**.
- Can synthesize **accurate binaural signals** from a **monaural source**.

Room Simulation:

- Room-related impulse response.

Speed and efficiency:

- Balance number of sources against real time constraints.

Output:

- Stereo, Surround Sound (e.g. Dolby 5.1, 7.1).

Aural Pipline

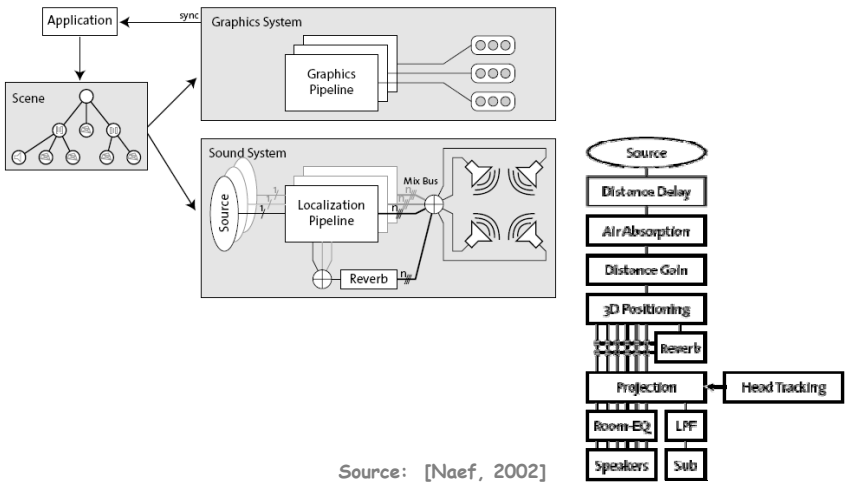
Basic Elements:

- **Buffers**:
  - Primary
  - Secondary
- **Sound sources**:
- **Listener**:
  - Position
  - Orientation
  - Velocity
- **Enhanced elements**:
  - Directionality
  - Doppler Shift
  - Effects

Functionality:

- **Playback**:
  - Play, pause, stop, rewind, …
- **Notifications**:
  - For synchronization.
- **Volume/Gain**:
  - Adjustable:
    - Smooth fades.
    - Panning / Positioning.
  - Relative vs. absolute:
    - Relative more practical
- **Frequency**:
  - Resampling.
  - Pitch shifting.

Aural Pipeline



Source: [Naef, 2002]

Aural Pipeline: Distance

Inverse Square Law:

- When distance is **doubled**, sound intensity decreases to **one fourth**.
- Measured in **decibels** (dB).

Problem:

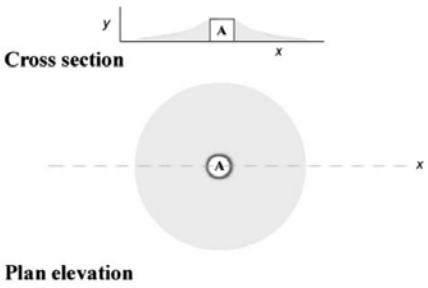
- **Limited dynamic range**.
- Solution is analogous to near and far clipping planes.

Minimum distance:

- Sound **does not** get **louder** when source is closer than this.

Maximum distance:

- Volume **does not decrease** when source is farther.



Audio Engines

Common Audio Engines:

- Microsoft DirectX
  - XNA and XACT
- OpenAL
- fMod
- Miles
- Java 3D

OpenAL

OpenAL:

- A **cross-platform 3D audio API**.
  - Styled after OpenGL.
  - ALUT similar to GLUT.
  - Under active development.
- Appropriate for use with **gaming applications** and many other types of audio applications.
- **History**:
  - Founded by Terry Sikes and Bernd Kreimeier at Loki Software.
  - Ported Windows games to Linux.
  - Subsequently taken over by Creative Technology with support from Apple.
- **http://www.openal.org/**

Microsoft DirectSound

DirectSound:

- **Part of DirectX**.
- Platform for **Windows and Xbox**.

Features:

- **HRTF** (head-related transfer function) integration.
- Audio effects:
  - Filters.
  - One primary buffer.
  - Secondary buffers limited by sound card.

XACT:

- Microsoft "Cross-platform" Audio Creation Tool (with API).
- Supports **Windows and Xbox**.

Other Audio Engines

Commercial Engines:

Miles:

- RAD Game Tools
- Various Render Engines

fmod:

- Free for noncommercial use.

Java3D sound API:

- Intelligent **class design**.
- Limited to **stereo output**.
- Not widely used.

	Direct Sound	OpenAL	fmod	Miles	Java3D
Multichannel support	X	X	X	X	
HRTF filtering	X	X	X	X	
MP3 support			X	X	
Audio effects (filters)	X	X	X	X	X
Platform independent		X	X	X	X
Available for free	X	X			X

Summary

Summary:

- Nature of sound and human hearing.
- Audio design in games.
- Audio engines and the aural pipeline.