OVERVIEW

- Physics of sound
- Acoustic cues for sound localization
  - Azimuth
  - Elevation
  - Range
- Head-related transfer functions (HRTFs)
- Approaches to synthesizing spatial sound
- Opportunities and challenges

 đaopotat: 2c

\[ c = \frac{f}{l} \]

MULTIPATH PROPAGATION

- Reflection
- Refraction
- Scattering

AXIOM I

The sound pressure at the two ear drums is a sufficient stimulus. Producing the same sound pressure will produce the same auditory perception.

Caveats:
- Bone conduction
- Adaptation
- Conflicting visual cues
- Conflicting expectations

AXIOM II

Exact reproduction of the sound pressure is not necessary for producing the same auditory perception. The limitations of neural responses allow different (and simpler) stimuli to produce the same response.

Examples:
- Bandwidth (20 Hz to 20 kHz)
- Amplitude (1-dB resolution)
- Monaural phase (2-ms resolution)
- Latency (10-ms resolution)
- Spectral fine structure (critical bands, Q = 8)

AXIOM III

Although it is not necessary to reproduce all of the cues exactly, conflicting cues degrade perception.

Key engineering challenge -- find the most cost-effective approximation.

VERTICAL-POLAR COORDINATES

INTERAURAL-POLAR COORDINATES
AZIMUTH CUES

• ITD (Interaural Time Difference)
• ILD (Interaural Level Difference)

WOODWORTH'S FORMULA

\[
T_{\text{con}} = \frac{a}{c} \\
T_{\text{ips}} = -\frac{a \sin c}{c} \\
\text{ITD} = a \frac{\sin c}{c}
\]

ARRIVAL TIME

- Rayleigh's solution (20% rise time)
- Woodworth's formula

ELEVATION CUES

• Pinna reflections and resonances
• Torso and shoulder reflections

TORSO REFLECTION

\[
T_T = \frac{2h}{v} \\
[H(f)]
\]

THE PINNA

PINNA PHENOMENA

- Pinna reflections (Batteau)
- Pinna resonances (Shaw)

PINNAE

RANGE CUES

• Loudness (for familiar sources)
• Excess ILD (for close sources)
• Direct/reverberant (for distant sources)
HEAD-MOTION CUES AND FRONT/BACK CONFUSION

HEAD-MOTION CUES AND ELEVATION MAGNITUDE

ITD = $\frac{2\alpha}{c}$  
ITD = 0  
ITD = $\frac{2\alpha}{c} \cos f$

OTHER CUES

- Visual cues
- Synchronized motion
- Absence

- Knowledge of source
- Knowledge of environment

FREE-FIELD RADIATION FROM A SPHERICAL SOURCE

$X(f) = \text{Fourier transform of source pressure}$

$X_{ff}(f) = \text{Free-field pressure at head center}$

$X_{ff} = H_{ff} X$

$H_{ff}(f) = \frac{t_0}{r} e^{-jkr}$, $k = \frac{2\pi f}{c}$

Inverse range  Propagation delay

THE HEAD-RELATED TRANSFER FUNCTION

$X(f) = \text{Fourier transform of source pressure}$

$X_L(f) = \text{Fourier transform of left ear pressure}$

$X_R(f) = \text{Fourier transform of right ear pressure}$

$X_{ff}(f) = \text{Free-field pressure at the origin}$

$X_L(f) = H_L(f) X_{ff}(f)$  
$X_R(f) = H_R(f) X_{ff}(f)$

THE HEAD-RELATED IMPULSE RESPONSE

$X_L(t) = \text{Left ear pressure}$

$X_R(t) = \text{Right ear pressure}$

$X_{ff}(t) = \text{Free-field pressure at the origin}$

$x_L(t) = \int_{-\infty}^{\infty} h_L(\tau) x_{ff}(t-\tau) \, d\tau$  
$x_R(t) = \int_{-\infty}^{\infty} h_R(\tau) x_{ff}(t-\tau) \, d\tau$

HRIR SOUND SYNTHESIS

A STRUCTURAL MODEL

COMPUTING HRTFs BY BOUNDARY ELEMENT METHODS

* Digitize with a 3-D scanner
* Solve wave equation numerically

* See Kahana et al.
**HRTF FOR ISOLATED PINNA**

![HRTF FOR ISOLATED PINNA](image)

**CONTRIBUTIONS TO THE HRTF**

![CONTRIBUTIONS TO THE HRTF](image)

**A STRUCTURAL MODEL**

![A STRUCTURAL MODEL](image)

**THE SPHERICAL-HEAD MODEL**

![THE SPHERICAL-HEAD MODEL](image)

**ASSESSING THE SPHERICAL HEAD MODEL**

- Only one parameter -- easily customized
- Well focused
- Good left/right position
- No up/down control -- image elevated
- With a head tracker:
  - Moderately externalized
  - Little front/back confusion
- Without a head tracker:
  - Internalized
  - Usually seems to be in back

**ELLIPSOIDAL-TORSO MODEL**

![ELLIPSOIDAL-TORSO MODEL](image)

**ASSESSING THE ELLIPSOIDAL TORSO MODEL**

- Five parameters; still easily customized
- Provides an elevation cue
  - Significant below 3 kHz
  - Ineffective in median plane
- Only one component of a full model

**STRUCTURAL HRTF MODEL**

![STRUCTURAL HRTF MODEL](image)

**SIMPLIFIED PINNA MODEL**

![SIMPLIFIED PINNA MODEL](image)
SPATIAL SOUND SYSTEMS

Multichannel

Two-channel: headphones

Two-channel: crosstalk-cancelled loud speakers

MULTICHLANEL SYSTEMS

Pros
• Works with a large audience
• No customization needed
• Conceptually simple

Cons
• Speakers must be distant
• Many channels needed for full 3-D
• Space consuming, expensive

TWO-CHANNEL: HEADPHONES

Pro
• Can reproduce full 3-D with only 2 channels
• Private and non-interfering
• Conceptually simple

Cons
• Uncomfortable for extended use
• Unencumbered listening

TWO-CHANNEL: CROSSTALK-CANCELLED LOUD SPEAKERS

Pros
• Can reproduce full 3-D with only 2 channels
• Unencumbered listening

Cons
• Small "sweet spot"
• Cannot be used with a large audience
• Requires customization for full 3-D
• Difficult to get near or rear locations

APPROACHES TO CUSTOMIZATION

• Measure exact HRTF for each person
• Acoustic
• Computational
• Nearest-neighbor
• Trial and error
• Anthropometry
• Scale a standard HRTF
• Global
• Pinna/head/torso components
• Use an adaptive model
• Match to anthropometry
• Match to exact HRTF

CHALLENGES AND OPPORTUNITIES

• Frequency range
• Elevation perception
• Range perception
• Transducers
• Headphone compensation
• Loudspeaker "sweet spot"
• Latency in dynamic systems
• Room acoustics