AN INTRODUCTION TO HUMAN SPATIAL HEARING

Richard O. Duda
CIPIC Interface Laboratory
UC Davis

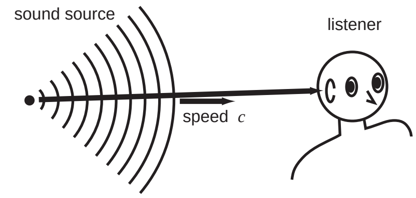
http://phosphor.cipic.ucdavis.edu

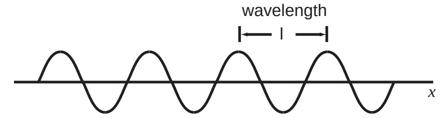
October 12, 2000

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

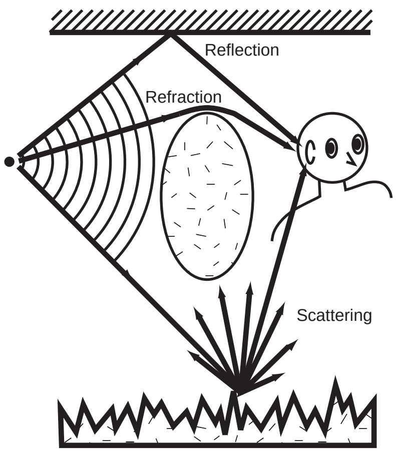
SOUND PROPAGATION





c = f l

MULTIPATH PROPAGATION



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:

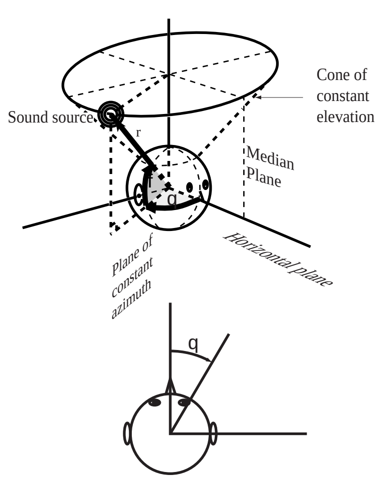
- Bandwidth (20 Hz to 20 kHz)
- Amplitude (1-dB resolution)
- Monaural phase (2-ms resolution)
- Latency (10-ms resolution)
- Spectral fine structure (critical bands, Q = 8)

AXIOM III

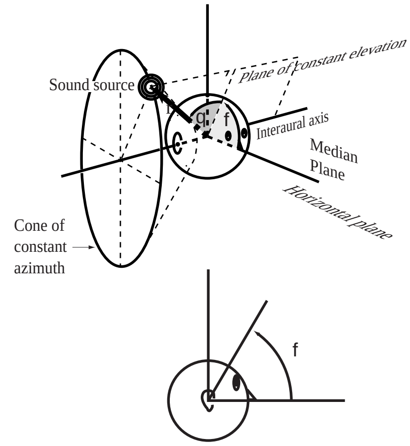
Although it is not necessary to reproduce all of the cues exactly, conflicting cues degrade perception.

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

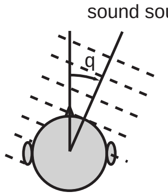
VERTICAL-POLAR COORDINATES



INTERAURAL-POLAR COORDINATES

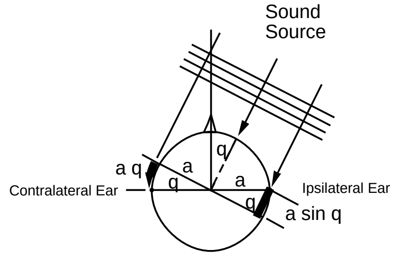


AZIMUTH CUES



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

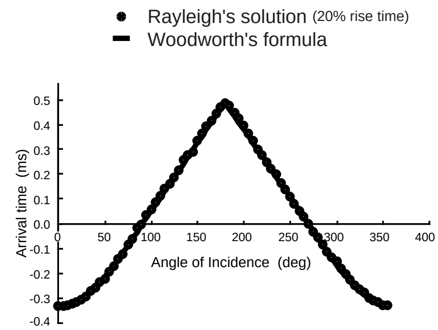
WOODWORTH'S FORMULA



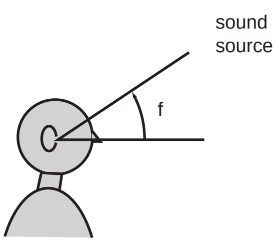
$$DT_{con} = \frac{a\,q}{c} \qquad DT_{ips} = \frac{-\,a \sin q}{c}$$

$$ITD = a\,\frac{q + \sin q}{c}$$

ARRIVAL TIME

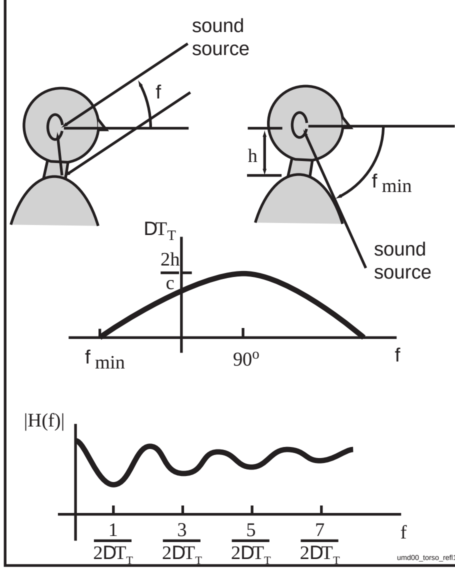


ELEVATION CUES

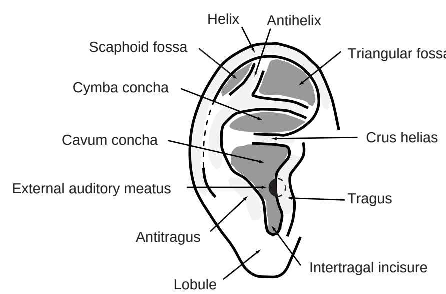


- Pinna reflections and resonances
- Torso and shoulder reflections

TORSO REFLECTION

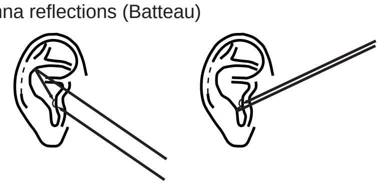


THE PINNA

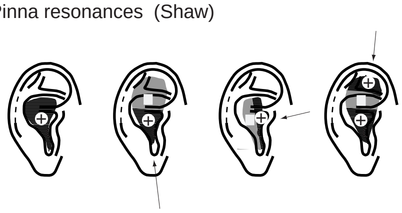


PINNA PHENOMENA

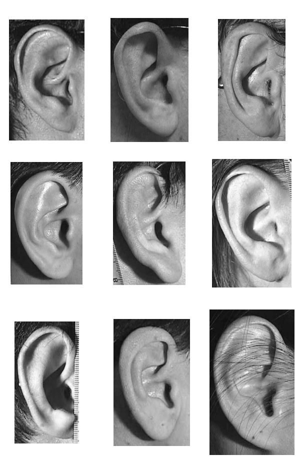
Pinna reflections (Batteau)



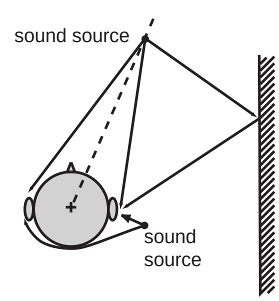
Pinna resonances (Shaw)



PINNAE

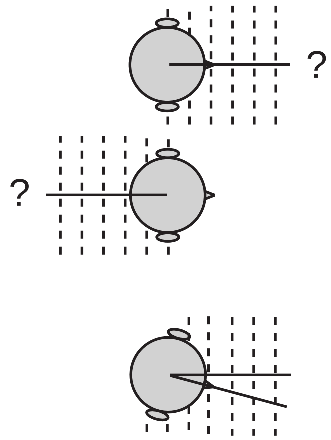


RANGE CUES

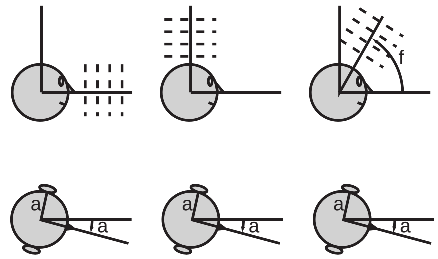


- Loudness (for familiar sources)
- Excess ILD (for close sources)
- Direct/reverberant (for distant sources)

HEAD-MOTION CUES AND FRONT/BACK CONFUSION



HEAD-MOTION CUES AND ELEVATION MAGNITUDE

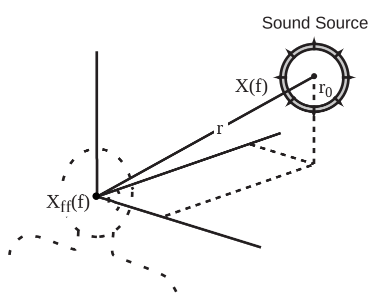


ITD = $\frac{2a}{c}$ a   ITD = 0   ITD = $\frac{2a}{c}$ a cos f

OTHER CUES

- **Visual cues**
  - **Synchronized motion**
  - **Absence**
- **Knowledge of source**
- **Knowedge of environment**

FREE-FIELD RADIATION FROM A SPHERICAL SOURCE

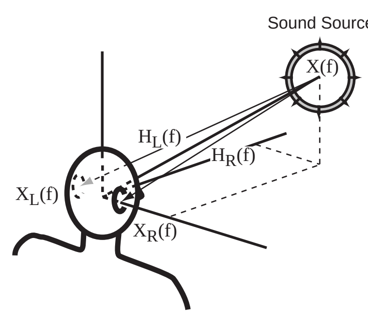


$X(f)$ = Fourier transform of source pressure
$X_{ff}(f)$ = Free-field pressure at head center

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

$H_{ff}(f) = \underbrace{\frac{r_0}{r}}_{\text{Inverse range}}\ \underbrace{e^{-jkr}}_{\text{Propagation delay}}$, $k = \frac{2 p f}{c}$

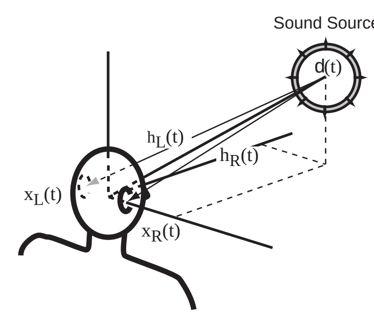
THE HEAD-RELATED TRANSFER FUNCTION



$X(f)$ = Fourier transform of source pressure
$X_L(f)$ = Fourier transform of left ear pressure
$X_R(f)$ = Fourier transform of right ear pressure
$X_{ff}(f)$ = Free-field pressure at the origin

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

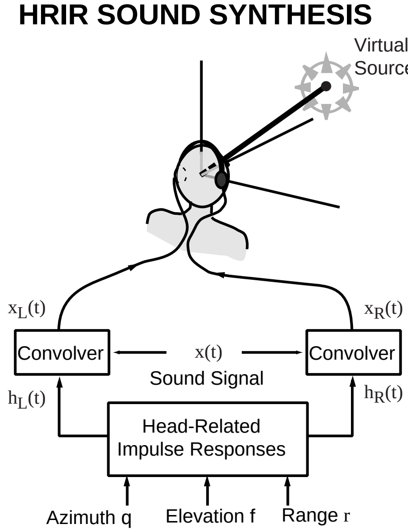
THE HEAD-RELATED IMPULSE RESPONSE



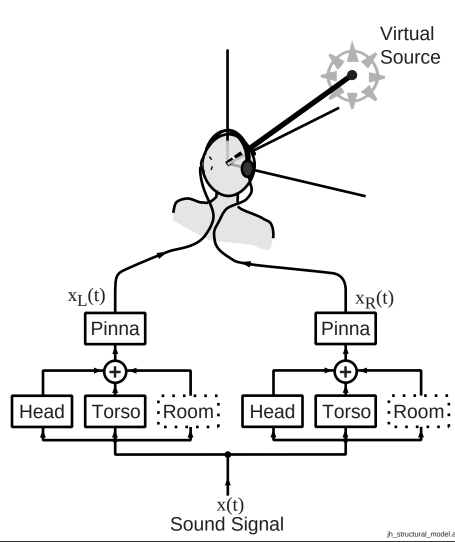
$x_L(t)$ = Left ear pressure
$x_R(t)$ = Right ear pressure
$x_{ff}(t)$ = Free-field pressure at the origin

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

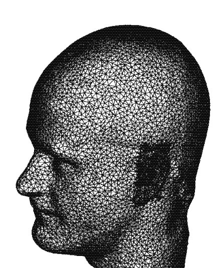
HRIR SOUND SYNTHESIS



A STRUCTURAL MODEL



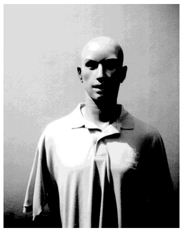
COMPUTING HRTFs BY BOUNDARY ELEMENT METHODS



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

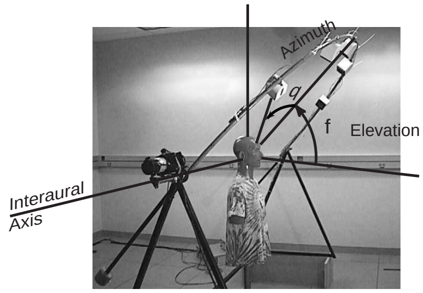
* See Kahana et al.

THE KEMAR ACOUSTIC MANIKIN



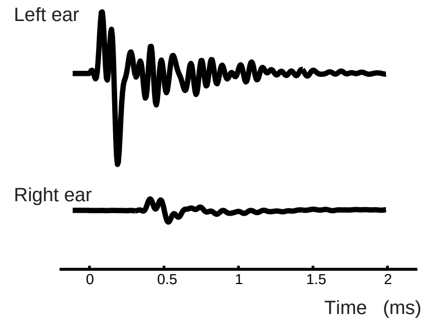
ubc_kemar.ai

ACOUSTIC HRTF MEASUREMENT



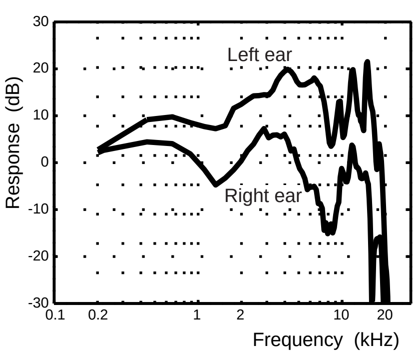
umd00_hoop.ai

KEMAR HRIR



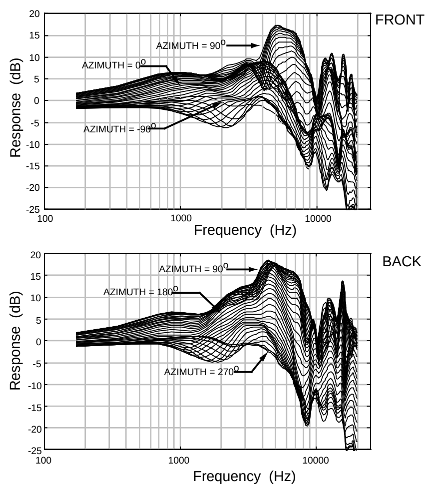
jh_kemar_hrir_m45.ai

KEMAR HRTF



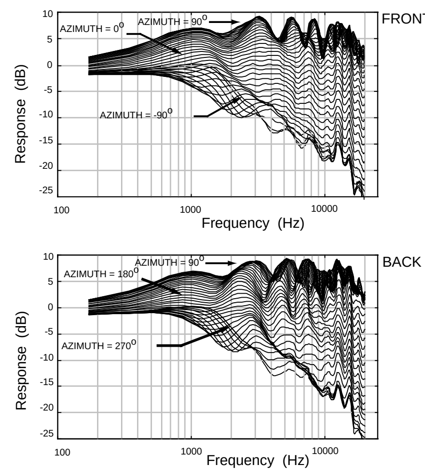
jh_kemar_hrtf_m45.ai

RIGHT-EAR HRTF FOR KEMAR (Horizontal Plane)



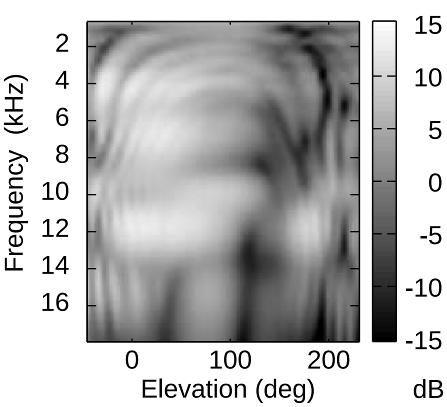
ubc_ke_freq.ai

HRTF FOR KEMAR, NO PINNA (Horizontal Plane)



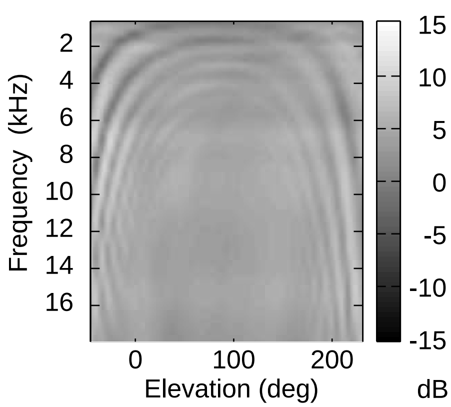
ubc_ke_np_freq.ai

HRTF ELEVATION DEPENDENCE



umd00_full_HRTF.ai

HRTF WITHOUT PINNA

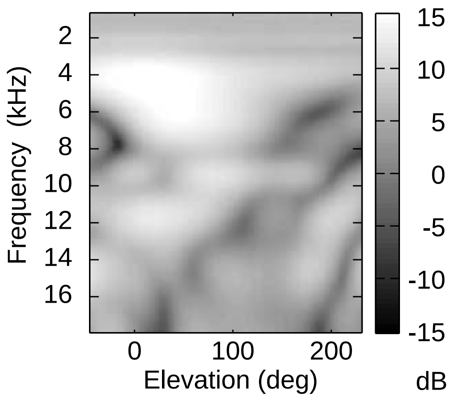


umd00_HRTF_nopinna.ai

A PINNA ON A PLANE

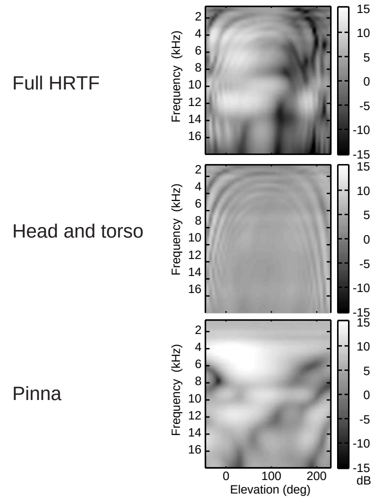
umd00_pinplane.ai

HRTF FOR ISOLATED PINNA



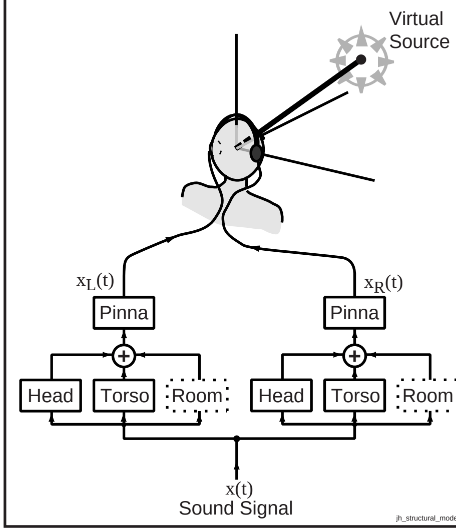
umd00_HRTF_pinna.ai

CONTRIBUTIONS TO THE HRTF



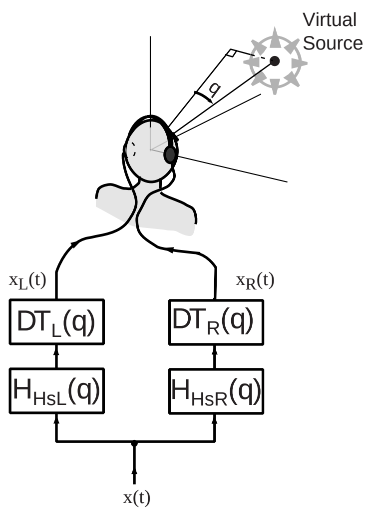
umd00_HRTF_contributions.ai

A STRUCTURAL MODEL



jh_structural_model.ai

THE SPHERICAL-HEAD MODEL



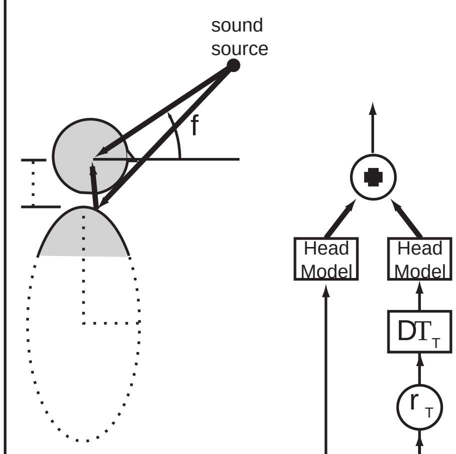
ubc_sphere_model.ai

ASSESSING THE SPHERICAL HEAD MODEL

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

jh_sphere_assess.ai

ELLIPSOIDAL-TORSO MODEL



r_T = torso reflection coefficient

DT_T = torso reflection delay

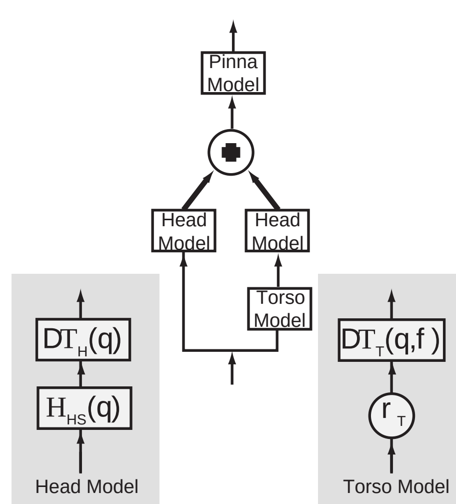
jh_torso_reflections.ai

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

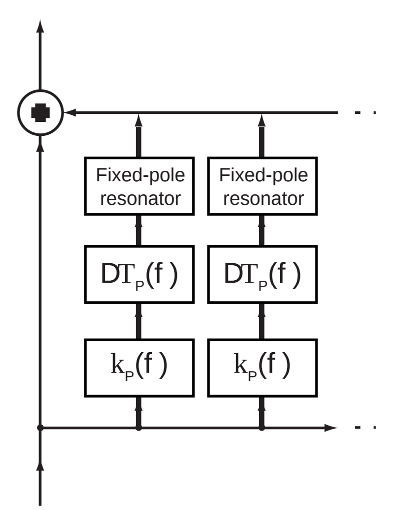
jh_ellipsoid_assess.ai

STRUCTURAL HRTF MODEL



jh_structural_model_2.ai

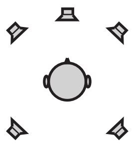
SIMPLIFIED PINNA MODEL



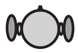
jh_structural_model_3.ai

SPATIAL SOUND SYSTEMS

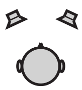
Multichannel



Two-channel:
headphones

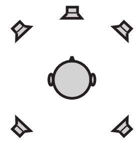


Two-channel:
crosstalk-canceled
loud speakers



umd00_systems.ai

MULTICHANNEL SYSTEMS



Pros

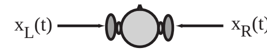
- Works with a large audience
- No customization needed
- Conceptually simple

Cons

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

umd00_systems2.ai

TWO-CHANNEL: HEADPHONES



Pros

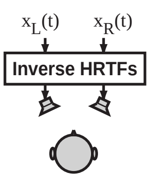
- Can reproduce full 3-D with only 2 channels
- Private and non-interfering
- Conceptually simple

Cons

- Uncomfortable for extended use
- Clumsy for a large audience
- Requires customization for full 3-D
- Difficult to achieve frontal externalization

umd00_systems3.ai

TWO-CHANNEL: CROSSTALK-CANCELED 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

umd00_systems4.ai

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

umd00_customization.ai

CHALLENGES AND OPPORTUNITIES

- Frequency range (combining partial HRTFs)
- Elevation perception
  - Front/back confusion
  - Low elevations
- Range perception
  - Headphones: externalization
    - Median plane
    - Frontal
  - Speakers: back locations
- Transducers
  - Headphone compensation
  - Loudspeaker "sweet spot"
- Latency in dynamic systems
- Room acoustics

umd00_problems.ai