COMP 546

Lecture 21

Head and Ear

Tues. April 4, 2017
The human ear.
Overview of Today

• outer ear
  • head related impulse response
  • head related transfer function

• inner ear
  • cochlea

• critical bands, and masking
Head related impulse response (HRIR)

Suppose sound source is coming from direction \((\phi, \theta)\) and emits \(I_{src}(t; \phi, \theta)\).

The wave is planar when it arrives at the end. If the source is an impulse then sound measured at the ear drum of ear \(i\) is:

\[
I(t) = \mathbb{D}_i(t; \phi, \theta) \ast \delta(r - vt)
\]

For a general source, the sound is

\[
I(t) = \mathbb{D}_i(t; \phi, \theta) \ast I_{src}(t; \phi, \theta)
\]

(Ignoring time delay here)
In following slides, I will show HRIR measurements $\mathbf{h}_i (t; \phi, \theta)$ from a KEMAR mannequin.

Note coordinate system below is different from one used for cone of confusion last lecture.

azimuth $\theta$

elevation $\phi$

Note: only need to represent elevations from -90 to 90.
Azimuth $\theta$  (Elevation $\phi = 0$)

Suppose sound is measured at right ear drum.
HRIR

Source direction (azimuth)
Arrival time differences are not as significant when azimuth = 0 and elevation is varied.

HRIR

Source direction (elevation)
ASIDE: If head is symmetric about the medial plane (left/right), then HRIR satisfies:

$$\mathcal{H}_{left}(t; \phi, \theta) = \mathcal{H}_{right}(t; \phi, -\theta)$$

Note: only need to represent elevations from -90 to 90.
HRTF: Head Related Transfer Function

\[ I_r(t; \theta, \phi) = h_r(t; \theta, \phi) \ast I_{src}(t; \theta, \phi) \]

For each incoming direction, what is the Fourier transform with respect to t?

\[ \hat{I}_r(\omega; \theta, \phi) = \hat{h}_r(\omega; \theta, \phi) \cdot \hat{I}_{src}(\omega; \theta, \phi) \]
HRTF $\hat{I}(\omega; \theta, \phi)$

(elevation $\phi = 0$)

Roughly sinusoidal for each frequency, with max at 90 degrees (right ear)
HRTF \( \hat{I}(\omega; \theta, \phi) \)

(azimuth \( \theta = 0 \))

Curves shifted for visualization

Valley is “pinnal notch” (it distinguishes elevations)

(medial plane)
Middle Ear

“Ear drum”

Ossicles (bones)
Ossicles act as a lever, transferring vibrations from ear drum to fluid in cochlea.
Vestibular apparatus (also inner ear)

cochlea
Cochlea (unrolled)
Basilar Membrane (BM)

• If sound is a pure tone, then all positions on basilar membrane vibrate at that frequency

• For each position on the basilar membrane, there is a pure tone that gives the largest vibration amplitude at that position (called the center frequency or CF)


http://auditoryneuroscience.com/ear/bm_motion_2
Recall vibrating string  \[ \omega = \frac{c}{L} \]

Both \(L\) and \(c\) vary on fibres on basilar membrane.
Basilar membrane mechanical impulse response function (at one point on the BM) is commonly modelled as a “gammatone” filter:

\[ g(t) = t^{n-1} e^{-\beta t} \cos(\omega t + \phi) \quad t > 0 \]

bandwidth center frequency
Bandwidth depends on frequency

\[ \Delta \omega \sim 100 \text{ Hz} \quad \Delta \omega = \frac{1}{3} \text{ octave} \]

(later today) The above model is derived from psychophysical experiments on "critical band" masking, but similar models can be made for BM vibration, hair cell or ganglion cell responses.
Place code (“tonotopic”)

Nerve cells (hair + ganglion) are distributed along the basilar membrane and have similar response functions as BM.
Analogy

**Eye**
- Lens and camera
- Photoreceptors (light $\rightarrow$ chemical)
- Ganglion cells (spikes) & optic nerve

**Ear**
- Outer ear
- Cochlea hair cells (mechanical $\rightarrow$ chemical)
- Ganglion cells (spikes) & cochlear/auditory nerve
Ganglion cells produce action potentials that carry encoded sound signals to the brain.

Hair cells are on basilar membrane
Sequence of events in cochlea

- **Basilar membrane** vibrates at some location

- **Hair cells** at that location release neurotransmitter that signal BM position as function of time (hair cells can signal detailed timing up to 2 kHz; beyond 2 kHz only the envelope is signalled)

- **Ganglion cells** respond with spikes
Timing of spikes ("phase locking")

Ganglion cells: spikes occur at peaks of basilar membrane deviations

- Louder sound (within band)
  - greater mechanical deviation (amplitude)
  - greater probability of spike at any peak (intensity coding)
- 3,000 hair cells } in each (left, right)
- 30,000 ganglion cells
- each hair cell feeds many ganglion cells
- each ganglion cell receives input from just one hair cell

Why?
Ganglion cells cannot spike faster than 500 Hz. Phase locking allows each ganglion cell to synchronize spikes with peaks of BM.

Many ganglion cells per hair cell allows all peaks to be coded by at least one ganglion cell.
“Volley” code

Different ganglion cells

spikes

time
Overview of Today

• outer ear
  • head related impulse response
  • head related transfer function

• inner ear
  • cochlea

• critical bands, and masking
Absolute threshold of hearing $I_0$ is frequency dependent.
Masking

How does the presence of one frequency limit our ability to hear other frequencies?

Two similar frequencies "mask" each other more than two different frequencies.

[Masking is the basis of MP3 audio coding]
Masking experiment

Which interval contains the test tone?
Which interval contains the test tone?

As we increase the intensity of the masking tone, it becomes more difficult to detect the test tone.
Choose a test frequency $\omega_0$ with SPL $I_0$.
For each masking frequency $\omega_M$
Measure a masking threshold $I_M(\omega_M)$
Define critical bandwidth for $\omega_0$ by $\Delta \omega$.

One of great discoveries of audition research is that there is close correspondence between basilar membrane response, hair cells and ganglion cell responses, and psychophysics of masking.
"Critical Bands" (Fletcher 1940)

For each $w_0$, there is a neighborhood (bandwidth) $\Delta w$ for which masking has a low threshold.

$\Delta w \approx \begin{cases} 
100 \text{ Hz, } w_0 < 1000 \text{ Hz}, \\
\frac{1}{3} \text{ octave, } w_0 > 1000 \text{ Hz}.
\end{cases}$

\[
\log_2 \frac{w_0 + \frac{\Delta w}{2}}{w_0 - \frac{\Delta w}{2}} = \frac{1}{3}
\]
Cochlear implants are used for profoundly deaf people (e.g. hair cells destroyed by disease but auditory nerve intact)

Microphone + speech/sound processor

Electrode array (inserted into cochlea)