lecture 21

ear

Monday April 8, 2013

Outer Ear
- pinna + auditory canal
- transforms the incoming sound wave by attenuating some frequencies but not others
  (Shoulders and head also transform incoming sounds.)

Head related impulse response (HRIR)
Suppose source in direction $(\theta, \phi)$ emits
$I_{\text{src}}(t; \theta, \phi)$. Assume source is distant so wave is planar when it arrives.
If the source is an impulse:
$I_{\text{src}}(t) = \delta(t)$.
The right ear measures
$I_{r}(t) = \delta(t) * h_{r}(t; \theta, \phi)
= h_{r}(t; \theta, \phi)

For a general source from $(\theta, \phi)$
$I_{\text{src}}(t; \theta, \phi)
= \sum_{t'} I_{\text{src}}(t'; \theta, \phi) \delta(t-t')$
So, the right ear measures
$I_{r}(t) = h_{r}(t; \theta, \phi) * I_{\text{src}}(t; \theta, \phi)$

HRIR measurements from a KEMAR mannequin
(see green data slides below)

azimuth
elevation
Suppose sound is measured at the right ear drum.

Elevation - arrival time differences are not as significant.

If head is symmetric about the median plane (left/right), then HRIR satisfies:

$h_e(t; \theta, \phi) = hr(t; \theta, 180^\circ - \phi)$

For each incoming direction $(\theta, \phi)$, take Fourier transform with respect to $t$:

$\hat{h}_e(\omega; \theta, \phi) = \mathcal{F}_\omega [h_e(t; \theta, \phi)]$

$\hat{h}(\omega, \theta, \phi) \mid$ for Azimuth angle $\theta$ (elevation fixed, $\phi = 0$).

Head related transfer function (HRTF):

$I_r(t; \theta, \phi) = h_r(t; \theta, \phi) * \text{Isrc}(t; \theta, \phi)$

$\text{HRIR}$
Challenging Situation: Cocktail party

Sound is coming from all directions ($\theta, \phi$)

"Cocktail Party Effect"
Ability to pay attention ("focus on") one sound source when many are present.
People with hearing in only one ear are much worse at this. (Source localization seems to be involved.)

Next lecture, we will discuss source localization i.e. how left & right ear's sounds can be combined to estimate $\theta, \phi$. 

Now, we look at how sounds are measured and coded in the ear. For simplicity, we are assuming there is a single source direction.
**Inner Ear**
(semicircular canals + cochlea)

- Basilar Membrane
- Recall vestibular system (head motion sensor)

If we could unroll the cochlea

If sound is a pure tone, 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).

recall vibrating string \( \omega = \frac{C}{L} \)

- Short (small L)
- Long (large L)
- Tense (large C)
- Loose (small C)
Basilar Membrane mechanical impulse response (is. response to a single 'click' sound) commonly modelled as

\[ g(t) = N e^{-at^2} \cos(\omega t + \phi), \quad t > 0 \]

Called a gamma-tone filter.

Similar to a Gabor but skewed.

Eye

- lens + cornea
- retina
- photoreceptors
- hair cells
- ganglion cells
- retinal nerve

Analog

- outer ear
- cochlea
- mechanical (IKs, neurotransmitters)
- ganglion cells (spikes)
- auditory nerve

Ear

- basilar membrane vibrations at same location
- hair cell at that location release neurotransmitter that signal BM position as function of time
- ganglion cell neurotransmitter levels with spikes
- ganglion cells: spikes occur at peaks of basilar membrane deviations
- Louder sound (within band) => greater mechanical deviations (amplitude) => greater probability of spike at any peak i.e. intensity coding

Sequence of events in inner ear:

- 3,000 hair cells
- 30,000 ganglion cells
- each hair cell feeds many ganglion cells
- each ganglion cell receives input from just one hair cell

Timing ("Phase Locking")

BM vibration \( \rightarrow \) time

Spike train

- ganglion cells: spikes occur at peaks of basilar membrane deviations
- Louder sound (within band) => greater mechanical deviations (amplitude) => greater probability of spike at any peak i.e. intensity coding
Ganglion cells cannot spike faster than 500 Hz. Phase locking allows each ganglion cell to synchronize spikes with peaks of BM.

\[ \text{e.g. } 2000 \text{ Hz (2 kHz)} \]

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

**Example**
- Glottal pulse
- Transfer function of vocal tract
- Head related impulse response
- Basilar membrane (gamma tone filter)
- Spikes

**Example:** Histograms of responses of cat auditory nerve fibres to human voice (vowel) at 12.5 ms (41 = 80 Hz).

**VOICE**

Ganglion cell's characteristic frequency (CF) determined by pure tone stimuli.

**Cochlear Implant** - used for profoundly deaf people.

- Microphone
- Speech/sound processor
- Transmitter
- Receiver + electrode array (inserted into the cochlea)