UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering
Instructor: Mark Hasegawa-Johnson

ECE 537 Speech Processing

Problem Set 8

Fall 2009

Issued: Fri Oct. 30, 2009 **Due:** NOT DUE

Reading for problem set 8: Licklider, 1951, *A Duplex Theory of Pitch Perception*; Wickesberg slides

---

Problem 8.1

Voltage differences across the length of a neuron cause current flow. In a copper wire, current flow takes the form of electrons diffusing gradually in the direction opposite the current. What is the chemical carrier of current in a neuron, and in what direction does it flow?

Problem 8.2

Describe briefly the function of each of the following organs, organelles, or compounds:

(a) Neurotransmitter

(b) Dendrites

(c) Axon

(d) Stellate cells

(e) Bushy cells

(f) Octopus cells

(g) Cochlear nucleus

(h) Olivary nucleus

(i) Cerebellum

(j) Temporal lobe

Problem 8.3

A plane wave $P_+(\Omega)$ arrives at the head from direction $[\theta, \phi]$, where $\theta \in [-\pi, \pi]$ is azimuth relative to straight ahead ($|\theta| < \pi/2$ is the front of the head, $\theta > 0$ is the right side of the head), and $\phi \in [-\pi/2, \pi/2]$ is elevation ($\phi > 0$ is the top of the head). The ratio between sound pressures at the left and right eardrums is called the head-related transfer function (HRTF),

$$H(\Omega, \theta, \phi) = \frac{P_L(\Omega)}{P_R(\Omega)}$$

(a) Assume that the head is a sphere of radius R, there are no other obstacles, the concha has zero length (eardrums are on the surface of the head), and diffusion around the head does not change the amplitude of the sound (i.e., you should model only ITD, not IID). What is $H(\Omega, \theta, \phi)$?

(b) Now assume the following model of IID: diffusion around the head reduces SPL at the left ear by $L(\Omega, \theta, \phi)$dB relative to SPL at the right ear. Modeling both IID and ITD, what is the HRTF?

(c) Suppose sound bounces off the pinna, producing an echo with an amplitude of α times that of the direct sound. The left pinna echo arrives at the left ear $T_L = T_P(1 + \sin\theta)$ seconds after the direct sound, where $T_P = D_P/c$ and D_P is the lateral width of the pinna. The right pinna echo arrives at the right ear $T_R = T_P(1 - \sin\theta)$ seconds after the direct sound. Adding this effect to the effects modeled in parts (a) and (b), what is the HRTF?

(d) Suppose now that when sound arrives from behind the head ($|\theta| > \pi/2$), the pinna echoes are all that you hear; the direct sound is wiped out entirely (a coarse model of diffusion). What is the HRTF?

Problem 8.4

The following sound is played to a subject, where $w(t)$ is a rectangular window of length 100ms:

$$p(t) = w(t)\left(\cos(2\pi 800t) + \cos(2\pi 1000t)\right)$$

(a) What pitch frequency does the subject hear,

  (1) In Hertz?
  (2) In octaves relative to A4?
  (3) In tones relative to A4?
  (4) In semitones relative to A4?
  (5) In mel?

(b) Vertical velocity of the basilar membrane is $\dot{z}(t, F_c) = h(t, F_c) * p(t)$, where $h(t, F_c)$ is the cochlear filter centered at F_c Hertz, and $\dot{z}(t, F_c)$ is the velocity of the basilar membrane at position $X(F_c)$ at time t. Assume for simplicity that the cochlear filters are ideal bandpass filters with a bandwidth of one ERB. Assume, further, that $w(t)$ is sufficiently smooth (sufficiently little out-of-band splatter) so that $h(t) * w(t)\cos(\Omega t) = w(t)\cos(\Omega t)$ if Ω is inside the passband of the filter, and zero otherwise. Find $\dot{z}(t, F_c)$.

(c) Licklider's duplex model of pitch perception can be understood as a model in which stellate cells accumulate a correlogram,

$$\frac{dC(t,T,F_c)}{dt} = -\alpha C(t,T,F_c) + y(t,T,F_c)$$

where $y(t,T,F_c)$ is an instantaneous coincidence counter; if we ignore the transduction nonlinearity, we can write

$$y(t,T,F_c) = \dot{z}(t,F_c)\dot{z}(t-T,F_c)$$

Assume the basilar membrane velocities are as computed in part (b). What is the resulting correlogram, $C(t,T,F_c)$? Assume that any part of $y(t,T,F_c)$ with a sinusoidal dependence on t is zeroed out by the integrator, and does not contribute to $C(t,T,F_c)$.

(d) Suppose that an octopus cell integrates across frequency bands,

$$r_{oct}(t,T) = \int C(t,T,F_c)dF_c$$

Demonstrate that $r_{oct}(t,T)$ has its biggest peak at $T = T_0$, the pitch period. (Note: this model is probably backward: the coincidence counters probably act on the output of the octopus cells, rather than vice versa. However, this is the model that Licklider proposed.)