EECS225d Audio Signal Processing in Humans and Machines **Spring 1995**

Lecture 13: February 15, 1995

Lecturer: John Lazzaro *Scribe: Kwong-Wen Wei*

13.1 It's Hard To Have A New Idea

Analogy is a way of speeding the process. Hearing is the example for this lecture. But let's start with the analogy of birds and planes.

13.1.1 Learning to Fly

Biological flight is the "existence proof" of the ability of flying. Human flight started with analogy of birds. Abstract analogies helped human design parts of the flying machine, such as wings and landing gear of the planes. However, specific analogies didn't help, such as the flapping of the birds' wings and the feathers on their wings.

13.1.2 Why?

Birds and Boeing plane solve different problems. For example, muscles and motors are really different. Also, birds don't carry passengers.

13.2 Auditory Airplane Analogies

The ear does spectral decomposition. Hearing is roughly log(f) and log(A). Also, the ear has limited frequency resolution(critical bands). The hearing range is 300-3300Hz, 48dB for intelligibility. These analogies are the jumping-off point for most speech signal analysis - these are the "wings and landing gear" of speech DSP. A few groups are exploring more detailed analogies.

13.2.1 Guidelines for Detailed Analogy

There is ample room for more detailed analogy. We should study physiology and psychoacoustics non-superficially. We should also form theories about the underlying computations. And finally implement theories at the highest level of abstraction possible, but no higher. The last point is a tricky one-too high lands you in airplane abstraction territory, too low then computational issues dominate.

13.3 Points of General Consensus

- Preserve cycle-by-cycle waveform information
- preserve time-frequency balance

13.4 Points of General Controversy

How important are:

- Modeling accuracy for filter shapes
- Modeling phase reponse
- Modeling non-linearity of filter shape
- Modeling gain control

13.4.1 One example - Gammatone filter

Gammatone filter is a good replacement for general purpose auditory nerve model. It captures a lot of the characteristics of the auditory nerve. In the equation gt[t] is the output as a function of time and u[t] is the step function. Given the step response, we are able to find out the filter specifications. For different frequencies, we normalize the phase differences.

$gt[t] = t^{n-1} exp[-2 * pi * t] cos[2 * pi * f0 * t + phi] u[t]$

$gt[t] = (t + tc)^{n-1} exp[-2 * pi * b * (t + tc)] cos[2 * pi * f0 * t] u[t]$

13.5 Viewgraphs

(1) The model for measuring and encoding interaural time differences:

In 1948, Jeffress proposed a model to explain the encoding of interaural time differences in neural circuits, as applied to the barn owl. In the model, fibers from the left and right nucleus magnocellularis(NM) converge on the nucleus laminaris(NL); The NM fibers encode timing information in the temporal patterning of their action potentials. Each fiber has a uniform time delay. The NL neurons (circles labeled a through e) act as coincidence detectors, firing maximally when signals from two sources arrive simultaneously. The pattern of innervation of the NM fibers in the NL create left-right asymmetries in transmission delays; when the binaural disparities in acoustic signals exactly compensate for the asymmetry of a particular neuron, this neuron fires maximally. The position in the array thus encodes the interaural time difference.

(2) Innervation of nucleus laminaris:

Axons from the ipsilateral and contralateral NM enter, respectively, the dorsal and ventral surfaces of the NL. Each axon runs along the respective surface until it comes to the iso-frequency band that it is destined to innervate. Within this band, the axon sends two or three collaterals into the NL, each of which divides into several branches. These fibers, which might be as long as 1mm, run along a relatively straight course toward the opposite surface. These fibers appear to establish contacts with the NL cell bodies either by en passant endings or by small branchlets.

(3) Floorplan of the silicon model of the time-coding pathway of the owl

Sounds for the left ear and right ear enter the respective silicon cochleas at the lower left and lower right of the figure. Silicon inner hair cells tap each silicon cochlea at 62 equally spaced locations; each silicon inner hair cell connects directly to a silicon spiral-ganglion cell. The square box marked with a pulse represents both the silicon inner hair cell and the silicon spiral-ganglion cell. Each silicon spiral-ganglion cell generates action potentials; these signal travel down silicon axons, which propagate from left to right for silicon spiral-ganglion cell from the left cochlea, and from right to left for silicon spiral-ganglion cells from the right cochlea. The rows of small rectangular boxes, marked with the symbol "dt," represent the silicon axons. One hundred seventy silicon NL cells, represented by small circles, lie between each pair of antiparallel silicon axons. Each silicon NL cell connects directly to both axons, and reponds maximally when action potentials present in both axons reach that particular neuron at the same time. In this way, interaural time differences map into a neural place code. Each vertical wire that spans the array combines the outputs of all silicon NL neurons with a specific interaural time difference. These 170 vertical wires form a temporally smoothed map of interaural time difference, which responds to a wide range of input sound frequencies. The nonlinear inhibition circuit near the bottom of the figure increases the selectivity of this map. The time- multiplexing scanner transforms this map into a signal suitable for display on an oscilloscope.

(4) Tuning properties of the auditory nerve:

We can characterize the tuning properties of the auditory-neve-fiber circuit model, using pure tones as input. In response to a pure tone of sufficient intensity and appropriate frequency, the auditory fiber produces spikes at a constant mean rate. The mean spike rate of a silicon fiber, in response to a constant tone, does not decrease over time, unlike that of a physiological fiber.

(5) Timing properties of the auditory nerve:

The temporal firing patterns of the auditory-nerve fibers encode information. This viewgraph shows period histograms of a chip fiber, in response to 40- to 90-db pure tones at the fiber's characteristic frequency; these histograms show the probability of a spike output during within a particular time interval during a single cycle of the input sinusoid. The fiber preserves the shape of the input sinusoid throughout this intensity range. The fiber, however, does not preserve absolute phase at higher intensities.

(5) The main ascending auditory pathways of the brain stem:

Many minor pathways are not shown on the viewgraph. The abbreviations used are: IC–inferior colliculus; MGB–medial geniculate body; NLL–nucleus of the lateral lemniscus. The branching and joining of arrows does not mean that the fibers branch or join.

(6) Frequency response area of single fiber in cochlear nerve:

For low-amplitude tones, the fiber responds to a narrow range of frequencies; for higher-intensity tones, the fiber responds to a wider range of frequencies. The saturating nonlinearities of the basilar-membrane circuit and of the inner-hair-cell circuit cause the bandwidth of the fiber to increase with sound intensity.

(7) Frequency reponse of a auditory nerve fiber:

This viewgraph shows a frequency-domain characterization of auditory-nerve fiber. The data was taken by presenting sinusoidal signals at different frequencies and amplitudes as input, and noting the average firing rate from a spiking output. Note that the fiber bandwidth asymmetrically increases with increasing amplitude; the high-frequency filter cutoff remains fairly stationary, but the low-frequency fiber cutoff shifts dramatically downward. The wide bandwidth of auditory nerve fibers at moderate intensities allows cycle-by-cycle timing information to be accurately preserved in the timing of individual spikes.

(8) Mean spike rate vs. amplitude:

This view graph shows the mean spike rate of an auditory fiber as a function of pure tone amplitude, at frequencies below, at, and above the best frequency of the fiber. In response to its characteristic frequency, the fiber encodes about 65 dB of tone amplitude before saturation.

(9) "Critical delay" cell:

This viewgraph shows the response of a cell in inferior colliculus as a function of the time delay between signals arriving at the two ears. A response maximum occurs at the so-called critical delay of 160 micro seconds largely irrespective of the frequency of the tone (or its level). Similar responses are found in the superior olive.

(10) Sphere model:

This viewgrah shows the differences between the distances of the ears from a source of sound far enough to produce a nearly plane wave front. The time delay between two ears depends on angle. The construction shows the geometry of the formula: delta d = r (theta + sin(theta)).

(11) Interaural time differences:

This viewgraph shows the interaural time differences as a function of the direction of the source of clicks. - are the measured values for five subjects and — are the values computed from sphere. These two sets of values agree fairly well.

References

[1] J. Lazzaro, "Biologically-based Auditory Signal Processing in Analog VLSI".

[2] J. Lazzaro and C. Mead, "Circuit Models of Sensory Transduction in The cochlea".