

ANALYSIS OF SENSORY TRANSFORMATIONS IN RESPONSE TO COMPLEX SENSORY STIMULI IN THE AUDITORY PATHWAY

Alexander Elman, Nan Kong*, Edward Bartlett, Kevin J. Otto
Weldon School of Biomedical Engineering
Purdue University
West Lafayette, IN
nkong@purdue.edu

Abstract

We develop a computational modeling framework to study information processing in the auditory pathway. Complex sounds consisting of a series of frequency modulated sweeps were used as test stimuli for our system. Frequency modulation is an important factor in sound recognition and speech perception. Human behavioral data demonstrated that perception of the direction modulation (UP or DOWN) could be easily controlled in the stimuli by manipulating a single parameter. In order to analyze the representations to these stimuli in the peripheral auditory system, a two-stage process was employed. First, the stimuli were decomposed by a well-accepted model of the auditory nerve to produce simulated auditory nerve outputs. Next, a neural network was developed to classify whether a given stimulus was going UP or DOWN. These data were compared to human behavioral data and neural activity in the rat auditory cortex. Outputs from the neural network model demonstrated that neurons could classify frequency change properly for simple stimuli but introduce significant bias errors in their classifications for more difficult stimuli. These results suggest ways by which pattern recognition techniques can be used to disambiguate neural responses that share some common characteristics.

Keywords: auditory nerve, neural networks, medial geniculate, neural coding transformation

1. Introduction

Biological sensory systems represent the physical world with a complex network of neurons that signal to one another via rapid electrical pulses called action potentials, or spikes. Although the neural spike codes for simple stimuli in peripheral sensory systems are understood relatively well, the neural coding transformations that take place between the auditory nerve and the auditory cortex, where conscious perception of sound originates, are poorly understood. We in this research utilize a computational modeling framework to study the way in which information is processed in the auditory pathway, including the auditory nerve, the auditory cortex, and its main sensory input, the auditory thalamus.

In this research, we integrate psychophysical and electrophysiological techniques to collect data that will inspire computational models (see Fig 1). Our research goal is to develop a neural network computational model that successfully maps auditory stimuli to neural and behavioral responses as well as revealing the neural coding transformations that occur between different

auditory nuclei. Our first purpose is to develop approaches with which pattern recognition techniques can be used to disambiguate neural responses that share some common characteristics (average spike rate, for example). Furthermore, the ability to understand how the analysis techniques are successful will provide insight into the underlying neural mechanisms.

Despite employing very different decision-making processes, a shared goal of animal intelligence and artificial intelligence is to make optimal decisions efficiently in complex and stressful environments. Many machine sensors are advanced pattern recognition devices that will generate consistent decisions for repetitions of the same pattern. Although this is often useful, it does not provide the flexibility to adapt decisions to changing situational context and unpredictable events. Generally speaking, artificially engineered algorithms are very successful in certain well-crafted problems when the system and parameters are relatively known, but do poorly at complex decision tasks under uncertainty such as speech recognition in natural environments. Unlike machines, humans and animals are able to make rapid, reasonably accurate decisions and adapt their behaviors to maximize benefits in uncertain environments.

Our vision of this line of research is to develop a theoretical framework to understand decision-making algorithms used by neural systems. Ideally, one would like to be able to predict behavior to a wide stimulus set in a wide range of behavioral contexts based on the neural activities of a relatively small number of neurons. For reasonably complex stimuli and environments, this becomes quite difficult because the perception of a stimulus may be shaped by factors such as the stimulus history and the consequences of potential behaviors in response to the stimulus.

2. Experimental Design

Stimulus design

Our basic experimental task is to discriminate upward changes in frequency versus downward changes in frequency. This flexible paradigm allows for the use of a wide variety of sound stimuli, including FM sweeps, tone sequences, and harmonic complexes. It also allows for excellent control over the stimulus complexity. The basic stimulus used was the “miniFM” stimulus (see Figures 2A&2B), which is a series of FM sweeps with semi-random starting times that produce a continually shifting set of temporally overlapping FM sweeps (see Figure 2B). Complex FM stimuli were designed to be

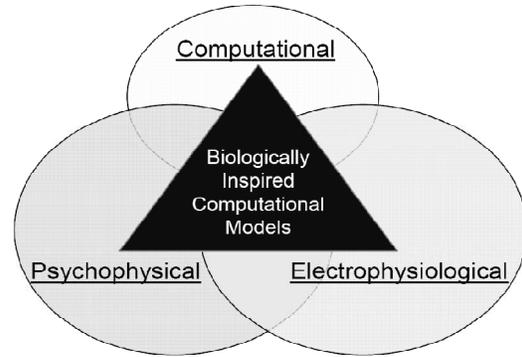


Figure 1: Schematic description of our research approach

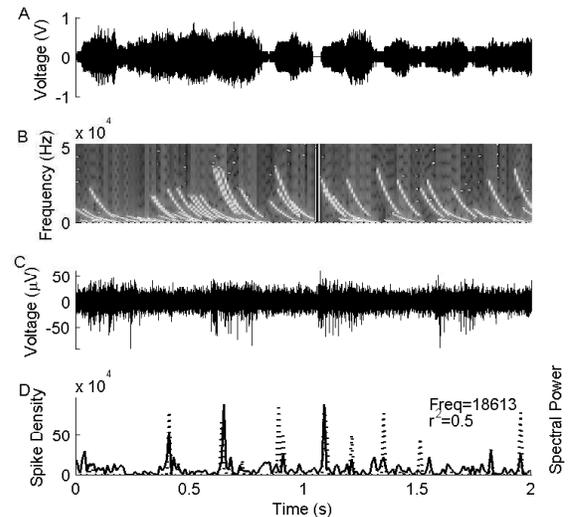


Figure 2: MiniFM stimulus and electrophysiology

analogous in terms of stimulating the peripheral receptor organs to moving dot patterns which have been used successfully to study neural correlates of visual perception [1]. FM sweeps were also chosen because proper perception of FM stimuli is critical for speech perception in realistic conditions [2].

Electrophysiological Data

Electrophysiological recordings were performed in rat primary auditory cortex in response to the miniFM stimulus. Extracellular electrophysiological recordings from a single microelectrode site are shown as Figure 2C. The evoked recordings were made 12 days post-surgery in response to the stimulus shown in Figure 2A. From the response to 20 repetitions of the stimulus, spike density functions were estimated from the action potential firing patterns (solid black lines, Figure 2D). A correlation coefficient was calculated for the spike density function vs. the spectral power of each band of the stimulus frequency. The spectral power for 18,613 Hz is shown as the dotted line. The correlation coefficient between the spike density function and the spectral power function is 0.5, suggesting that acoustic selectivity accounted for some but not all of the response variability at this recording site.

Psychophysical Data

Test paradigm. The testing paradigm utilized four different sets of parameters in order to systematically vary the uncertainty. Once subjects have mastered discriminating Up100 from Down100 stimuli (see Figure 2), stimulus uncertainty was introduced by adding a controlled amount of FM sweep components whose sweep rates are significantly different than the primary sweep rate. As Figure 3 demonstrates, human listeners have difficulty discriminating upward from downward sweeps when stimulus uncertainty is present in the form of interfering stimuli. It is easy to control task difficulty by controlling the proportion of FM sweeps that share a common rate of frequency change. As the proportion of FM sweeps with the same sweep rate decreases, the task becomes more difficult.

To investigate the task difficulty of miniFM and learning capabilities in humans, two subjects were tested for their ability to discriminate upward from downward miniFM stimuli. Subjects were not given verbal instructions, and the response panel consisted simply of an UP button and a DOWN button. Subjects were informed whether they responded correctly or not on each trial. These conditions were a partial attempt to simulate the limited information available to rats during training. During the first training block (Figure 3, Training Block 1) the subjects did poorly on the miniFM discrimination, with performance only somewhat

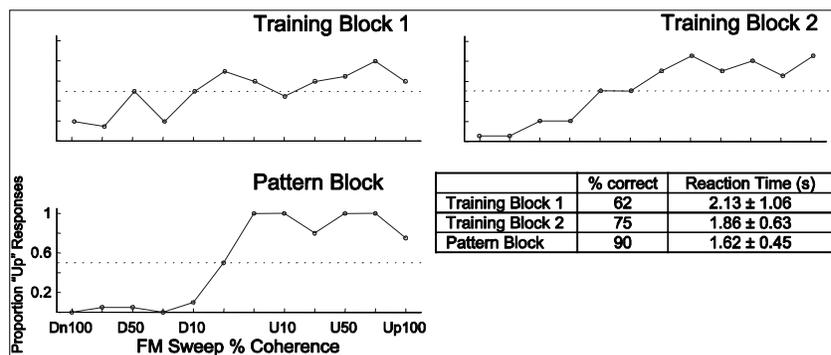


Figure 3 and Table 1: (Training Block 1) Psychophysical data from two naïve human subjects tested with Up vs. Down sweeps, 10 trials/stimulus. (Training Block 2) Two days later, subjects were retested and significantly improved their responses and shortened their reaction times ($p < 0.01$, χ^2 and ranksum tests). (Pattern Block) Subjects were also tested with 10 repeated stimulus blocks that progressed from Up100 to Down100 in 12 steps. Subjects significantly improved their responses and shortened their reaction times ($p < 0.01$, χ^2 and ranksum tests).

better than chance (Table 1, top row). A second training block (Figure 3, Training Block 2) significantly improved their performance and decreased their reaction time, indicating that they were able to improve their discrimination abilities with training (Table 1, middle row). A third set of stimuli was a pattern that was repeated 10 times (Figure 3, Pattern Block). Both subjects significantly improved their performance and decreased their reaction times compared to Training Block 2 (Table 1, bottom row). Both subjects were aware that stimuli repeated in a predictable pattern, but neither could explicitly report the pattern. These data suggest that behavior can be substantially improved when stimuli are predictable, that is, when stimulus uncertainty is low.

3. Modeling

3.1. Auditory Nerve Model

Synaptic data was obtained using a feline auditory periphery model developed by Heinz and Bruce *et al.* [3] to characterize the response of mammalian auditory-nerve (AN) fibers to high-level stimuli. The underpinning of the model's accuracy is in its faithful representation of the component 1-component 2 (C1/C2) transition and peak-splitting phenomena [4]. The model consists of eight separate processing blocks that accept an input stimulus in sound pressure level and discharges spike trains from a model of the IHC-AN synapse. The first block is the middle-ear (ME) filter which accepts an instantaneous pressure waveform of the miniFM stimulus. The ME filter is a fifth-order digital filter that discretizes the input using the bilinear transformation at a sampling rate of 500 kHz. The output of the ME filter is fed into a parallel path tenth order C2 filter, a signal-path tenth order C1 chirping filter, and a feed-forward control path. The C1 and C2 filters include transduction functions that model the behavior of the basilar membrane such as throttling frequency selectivity. The output of the C1 and C2 transduction functions are summed and passed through a seventh order low pass filter of the inner hair cell (IHC) block. The output of the IHC block is passed to the input of the IHC-AN synapse block. The IHC-AN synapse model consists of a nonlinear time-varying three store diffusion model. The terminating block serves as a discharge generator that outputs spike times by a renewal process driven by the synapse output [3].

3.2. Computational Model Development

A computational model was developed to accurately discriminate correlation between the synaptic output data of the AN model and the sweep direction of the miniFM input stimulus. The choice to use a neural network (NN) to characterize the input/output relationship between stimulation and synaptic response follows from its ability to quickly discern patterns from nonlinear temporally organized data [5]. While a NN model does not yield any insight into how biological coding is performed, it is a useful tool in interpreting the neural coding transformations between auditory nuclei.

3.2.1. Description of the neural network

The NN described in this experiment consists of a triple-layer perceptron network trained with a standard supervised feed-forward error back-propagation algorithm. Backpropagation is an appropriate NN architecture choice for this particular model because a well-trained backpropagation network faithfully generalizes outputs to foreign inputs. The input vector consists of selected characteristic frequency (CF) slices of firing rates (spikes/s) from the synaptic output of the AN model in response to the miniFM sweeps. The training vector consists

of binary values corresponding to whether or not the input sweep is up (1) or down (0). During the training session, the network undergoes dynamic weight and bias adjustments until the mean squared error is minimized to a chosen goal. Training terminates when the training goal is realized. To prevent noise from degrading the quality of the NN model, the synaptic outputs for each characteristic frequency and sweep count are summed to create ten 0.5 s bins. Besides eliminating noise, binning the spike rates decreases the complexity of the system and improves training speed and performance.

The Levenberg-Marquardt training algorithm was employed since it approximates the Hessian matrix instead of computing it directly, saving computation time at the expense of memory [5]. The hidden layer consisted of twenty neurons each with a linear transfer function. The output layer utilized a tan-sigmoid transfer function allowing output values to converge to a binary value.

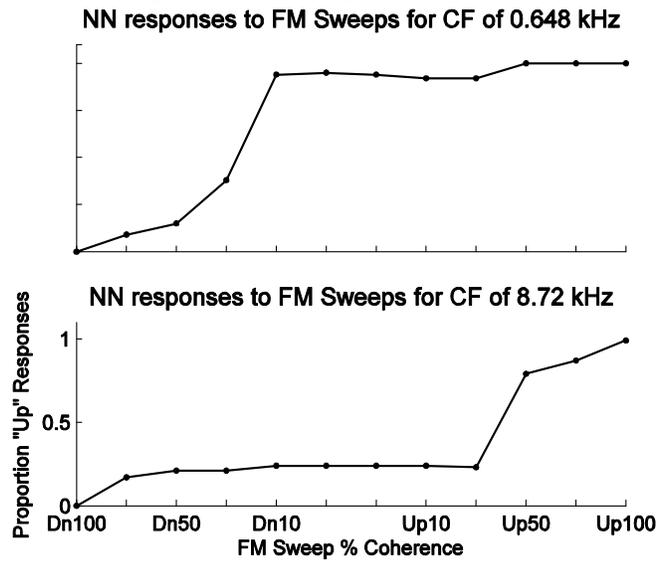


Figure 4: NN reported % of “Up” responses as a function of miniFM sweep direction and coherence. Lower CF have biases toward up responses.

3.3.2. Implementation of the neural network model

The computational work was performed in MATLAB (The Mathworks, Natick, MA) running on a Windows-based (Microsoft, Redmond, WA) PC. Twelve miniFM stimuli were fed into the AN model and an m-file was created for each stimulus. From the m-file, a matrix consisting of 57 spike-train vectors separated on an equally spaced characteristic frequency (CF) continuum from 125 Hz to 16 kHz contained the synaptic firing rate data. From the synaptic output matrix, a representative set of seven different CF fibers was extracted from the matrix and the values were summed into ten 0.5 s bins. The Neural Network Toolbox™ within MATLAB was used to automate the creation of the NN. Four training sessions were performed and then trained using the NN over 7 CF. Sweep pairs of firing rate vectors served as the input vectors and the binary values identifying sweep direction was used as the training vector. Training was carried out over 100 trials for each CF. The stochastic nature of the NN output is the result of how the toolbox uses pseudo-random values to initialize the network weights and biases before training. Training session 1 consisted of the 100up/100down pair. Each subsequent training session added the next highest sweep pair with training session 4, the final session, including all pairs except for the 10up/10down and 5up/5down pair. Each training session was tested against all twelve miniFM stimuli.

4. Results

The results are reported in terms of percentage of responses made to up sweeps of the FM stimuli. The accuracy of the NN varied over the continuum of CF values. The NN was faithful in

response to sweep coherence between 25% and 100%. The data show that the output representing the low frequency CF of 0.648 kHz was biased towards responding with up responses to difficult downward sweeps (Down5, Down 10) but was able to discriminate the direction of more coherent stimuli. Interestingly, each individual fiber had individual as represented by the comparison of two CF in figure 4. The CF of 8.72 kHz is biased towards responding down with difficult upward sweeps. The change in bias occurs between 3.6 and 8.7 kHz.

5. Conclusions and Future Work

Our results demonstrate the feasibility of investigating computational models that predict behavioral and neural data. Future work will extend these results to include rat behavioral data, more complete electrophysiological responses, and test between different computational methods. Our goals are twofold. First, we can use these models to gain insight into the neural mechanisms responsible for representing the sensory and behavioral variables. Second, we can apply these insights to the generation of advanced decision making and machine learning algorithms that can correctly identify patterns under substantial uncertainty.

References

1. Kajikawa, Y., et al., 2005, "A comparison of Neuron Response Properties in Areas A1 and CM of the Marmoset Monkey Auditory Cortex: Tones and Broadband Noise," *Journal of Neurophysiology*, 93(1), 22-34.
2. Kajikawa, Y., et al., 2008, "Coding of FM Sweep Trains and Twitter Calls in Area CM of Marmoset Auditory Cortex," *Hearing Research*, 238(1-2), 107-125.
3. Zhang, X., Heinz, M.G., Bruce, I.C., and Carney, L.H., 2001, "A Phenomenological Model for the Responses of Auditory-nerve Fibers: I. Nonlinear Tuning with Compression and Surpression," *Journal of the Acoustical Society of America*, 109(2), 648-670.
4. Zilany, M.S.A. and Bruce, I.C., 2006, "Modeling Auditory-nerve Responses for High Sound Pressure Levels in the Normal and Impaired Auditory Periphery," *Journal of the Acoustical Society of America*, 120(3), 1446-66.
5. Bishop, C. M., 1995, *Neural Networks for Pattern Recognition*, Oxford University Press, New York.