A computational model for discrimination of natural sounds

Rajiv Narayan[1,2], **Kamal Sen**[1,2,3]

1.Dept. of Biomedical Engineering, 2.Hearing Research Center,

3.Center for Biodynamics, Boston University, Boston MA

Complex natural sounds like human speech and birdsongs are characterized by temporal variations over multiple timescales. These variations could provide cues for distinguishing between different sounds even at the single neuron level. Here we assess the contribution of temporal patterns in neuronal firing to the discriminability of songs.

Our model is based on the spectral temporal receptive field (STRF), a quantitative description of the stimulus-response properties of auditory neurons. We use experimentally derived STRFs from field L (the avian analog of auditory cortex) to generate an estimate of the firing rate of a neuron in response to an ensemble of conspecific birdsongs [4,5]. This firing rate is used to drive a stochastic spike train generator to generate trials of model spike trains [1].

Conventional methods for characterizing neural discrimination such as total spike counts and mean firing rates typically neglect spike train dynamics. We investigate the use of a recently proposed spike distance metric that accounts for the temporal structure in the spike trains [3]. A significant advantage of the spike distance metric is that by varying a single free parameter, a time-constant tau, one can quantify discriminability over different time-scales of the neural response. At small time scales the distance measure acts like a "coincidence detector", with small differences in spike timing contributing to the distance, whereas at long time scales the distance metric acts like a "rate difference counter", where long term average firing rates contribute to the distance. This measure allows us to quantify the discriminability of spike trains for a range of different temporal resolutions and also examine the dynamics of discriminability.

We used the distance metric and a supervised classification scheme to quantify the neural discriminability of conspecific songs [2]. First, 10 trials of spike trains were obtained from the model for each of 20 different songs. We then picked a "template" spike train for each of the songs. Every spike train was assigned to the song with the closest template based on the spike-distance. The templates were chosen randomly and error bars were obtained by permuting the templates. The percentage of songs correctly classified (% correct) was used as a measure of the discriminability of songs.

We examined how discriminability changes as a function of the time scale of neuronal response by varying the time-constant in the distance metric. As discussed before, most previous methods have considered only the contribution of the long-term average firing rate in discrimination, which would correspond to choosing a very long time constant in the distance computation. Song discriminability improves significantly when one takes into account the temporal variations in the firing rate on much faster time scales reaching an optimal value between 10-20 ms (Figure 1). At time scales shorter than 10 ms, however, discriminability degrades again. Thus, the best discriminability is obtained in a range of time-scales that is intermediate between the "coincidence-detector" and the "rate difference counter" regimes.

To characterize how discriminability evolves over time as the neural response accumulates we computed % correct as a function of time over increasingly longer windows of the spike train (Figure 2). As seen in the figure, discrimination starts out at chance level, improves steadily and then saturates at long durations. In addition to providing information about the discriminability, this curve also provides information about the speed of discrimination.

Finally, we assessed how changes in parameters of the model STRF affect the reliability and time-course of discrimination. Figure 3 shows the discriminability curves for an experimentally derived STRF, which had a significant inhibitory component, and a synthetic STRF in which we removed the inhibitory component, normalizing the average firing rate of both models to be the same. As can be seen, the STRF with the inhibitory region produced faster and more accurate discrimination. This result suggests that the relationship between excitation and inhibition may play a critical role in shaping discriminability of natural sounds. We are currently investigating the effects due to more graded changes in inhibition as well as the effect of changing other STRF parameters. The computational model should help us identify candidate parameters that may be important for the discrimination of behaviorally relevant natural sounds.

References

1. M.J. Berry, M. Meister Refractoriness and neural precision. J. Neurosci. 18 (1998) 2200-2211.
2. C.K. Machens, H. Schutze, A. Franz, O. Kolesnikova, M.B. Stemmler, B. Ronacher, A.V. Herz Single auditory neurons rapidly discriminate conspecific communication signals, Nature Neuroscience 6 (2003) 341-342.
3. M.C.W. van Rossum, Novel Spike Distance, Neural Computation 13 (2001) 751-763.
4. K. Sen, F.E. Theunissen, A.J. Doupe Feature Analysis of Natural Sounds in the Songbird Auditory Forebrain, J Neurophysiol. 86 (2001) 1445-1458.
5. F.E. Theunissen, K. Sen, A.J. Doupe Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci. 20 (2000) 2315-2331.

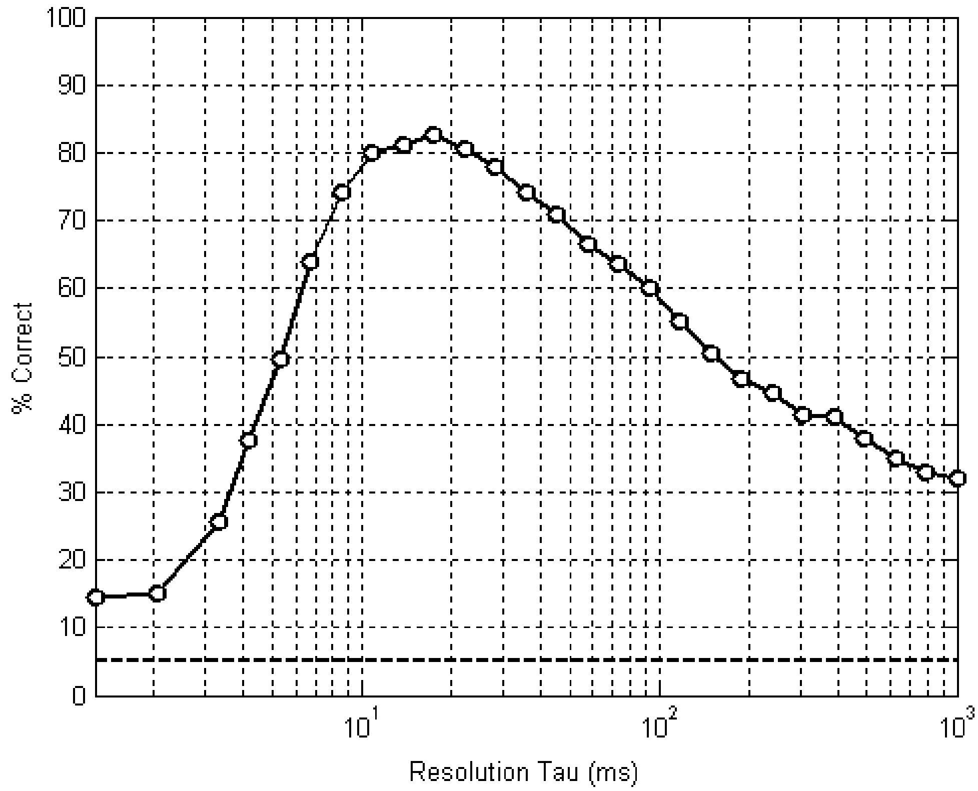

Fig 1. Dependence of discrimination on time-scale of neural response.

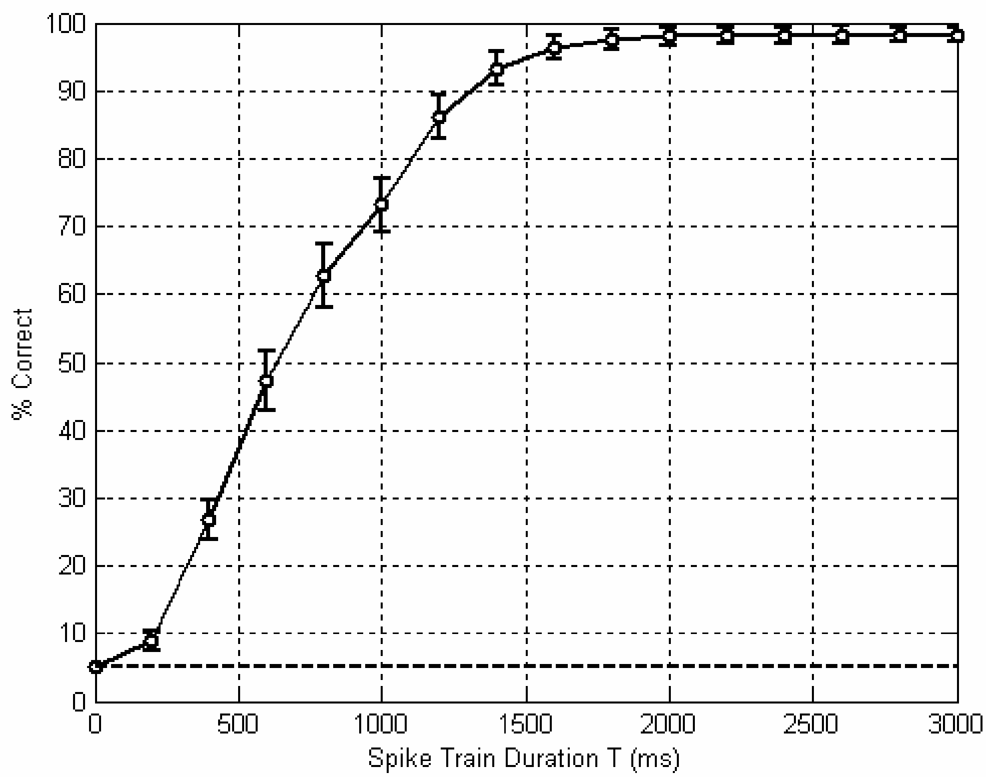

Fig 2. Dynamics of Song Discrimination.

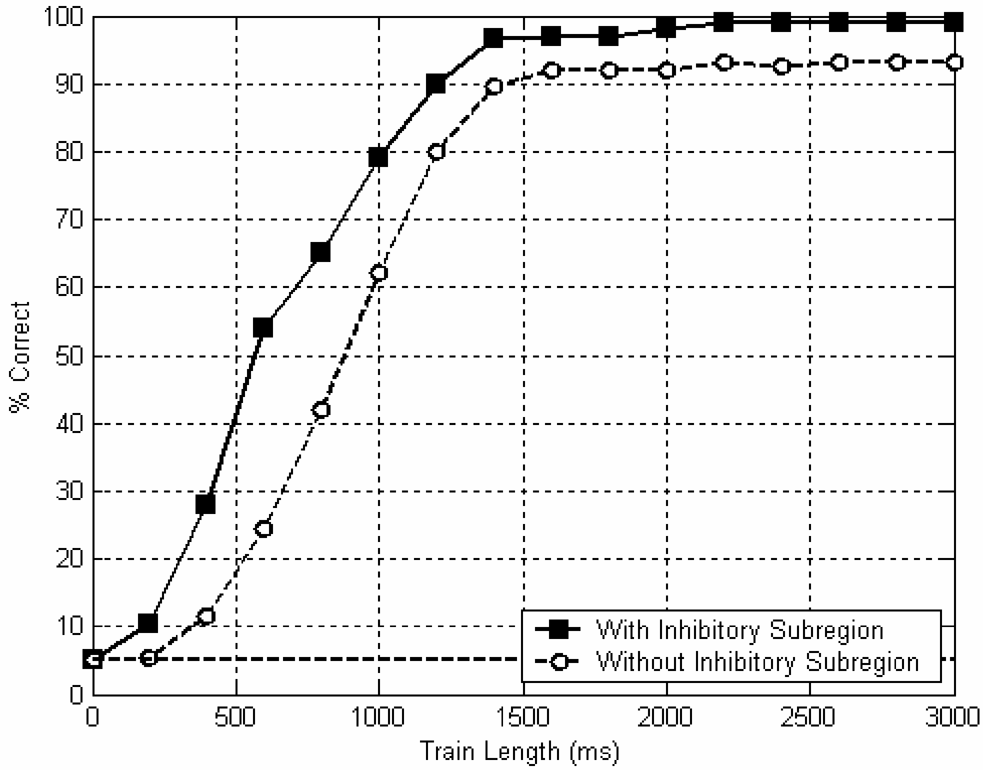

Fig 3. Dependence of discrimination on excitatory-inhibitory subregions of STRF.