Neuroscience 359 (2017) 130–141

MULTIDIMENSIONAL RECEPTIVE FIELD PROCESSING BY CAT PRIMARY AUDITORY CORTICAL NEURONS

CRAIG A. ATENCIO [a]* AND TATYANA O. SHARPEE [b,c]

[a] *Coleman Memorial Laboratory, UCSF Center for Integrative Neuroscience, Kavli Institute for Fundamental Neuroscience, Department of Otolaryngology-HNS, University of California, San Francisco, USA*

[b] *Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA*

[c] *Center for Theoretical Biological Physics and Department of Physics, University of California, San Diego, La Jolla, CA, USA*

Abstract—The receptive fields of many auditory cortical neurons are multidimensional and are best represented by more than one stimulus feature. The number of these dimensions, their characteristics, and how they differ with stimulus context have been relatively unexplored. Standard methods that are often used to characterize multidimensional stimulus selectivity, such as spike-triggered covariance (STC) or maximally informative dimensions (MIDs), are either limited to Gaussian stimuli or are only able to recover a small number of stimulus features due to data limitations. An information theoretic extension of STC, the maximum noise entropy (MNE) model, can be used with non-Gaussian stimulus distributions to find an arbitrary number of stimulus dimensions. When we applied the MNE model to auditory cortical neurons, we often found more than two stimulus features that influenced neuronal firing. Excitatory and suppressive features coded different acoustic contexts: excitatory features encoded higher temporal and spectral modulations, while suppressive features had lower modulation frequency preferences. We found that the excitatory and suppressive features themselves were sensitive to stimulus context when we employed two stimuli that differed only in their short-term correlation structure: while the linear features were similar, the secondary features were strongly affected by stimulus statistics. These results show that multidimensional receptive field processing is influenced by feature type and stimulus context.

Key words: spectrotemporal, strf, stimulus statistics, modulation, primary auditory cortex, cat.

*Correspondence to: C. A. Atencio, 675 Nelson Rising Lane, Box 0444, Room 535, San Francisco, CA 94158, USA.
E-mail address: craig@phy.ucsf.edu (C. A. Atencio).

Abbreviations: BMF, best modulation frequency; DMR, dynamic moving ripple; MAD, median/median absolute deviation; MIDs, maximally informative dimensions; MNE, maximum noise entropy; MNE, maximum noise entropy; NRF, network receptive field; RTF, ripple transfer function; SMF, spectral modulation frequency; SMTF, spectral modulation transfer function; STA, spike-triggered average; STC, spike-triggered covariance; STRFs, spectrotemporal receptive fields; TMF, function of temporal; TMTF, temporal modulation transfer function.

http://dx.doi.org/10.1016/j.neuroscience.2017.07.003

INTRODUCTION

The spectrotemporal receptive fields (STRFs) of auditory cortical neurons have been extensively studied in recent years (Blake and Merzenich, 2002; Elhilali et al., 2004; Gourevitch et al., 2009). The standard approach has been to estimate a single spectrotemporal feature in conjunction with a static nonlinearity (Chichilnisky, 2001). The feature may be seen as a stimulus that drives the neuron to respond, or it may be interpreted as a descriptor of neuronal stimulus processing (Sharpee et al., 2004). The single feature STRF description has provided important insights, though in recent years it has become apparent that many auditory cortical neurons simultaneously encode information about more than one stimulus feature in their spiking activity, and thus the single feature may not reveal the true richness of auditory cortical processing (Atencio et al., 2008, 2009, 2012; Harper et al., 2016; Kozlov and Gentner, 2016).

Multiple stimulus features can be estimated through dimensionality reduction techniques such as spike-triggered covariance (STC) and maximally informative dimensions (MIDs). STC accounts for pairwise stimulus interactions and can be applied in conjunction with Gaussian stimuli (Paninski, 2003; Samengo and Gollisch, 2013). The STC approach decomposes a spike-triggered stimulus covariance matrix into a set of eigenvectors (or stimulus features), where the contribution of each eigenvector is determined by the corresponding eigenvalue. Each feature can be classified as excitatory (increases neural responsiveness) or suppressive (decreases neural responsiveness) by examining the corresponding eigenvalue (Touryan et al., 2002; Rust et al., 2005; Chen et al., 2007). A more stimulus-robust approach is MID analysis (Kouh and Sharpee, 2009). MID analysis can be used with any stimulus type, making it useful when analyzing responses to naturalistic stimulation. However, though the MID procedure can account for all stimulus correlations, in practice it is limited to identifying a limited number of stimulus features due to the limits of data collection in standard physiological experiments (Rowekamp and Sharpee, 2011).

The tradeoff between stimulus type and filter number can be bridged by constraining the neural model. By assuming a functional form for the nonlinear input/output function, and by accounting for a restricted subset of stimulus statistics, an increased number of filters may be estimated. A recently developed approach, the maximum noise entropy (MNE) model, is able to account for the first- and second-order stimulus correlations that drive a neuron to fire. The MNE model produces an unbiased estimate of stimulus processing because it does not make assumptions regarding the input statistics, and parameter optimization is accomplished using maximum likelihood (Fitzgerald et al., 2011a,b). For non-specified higher ordered correlations, the model remains as unbiased as possible. Thus, the MNE model can recover multiple features that account for the linear and pairwise correlations in the stimulus while remaining maximally uncommitted toward other stimulus correlations.

We applied the MNE approach to study the neural coding of auditory cortical neurons. We show that responses of auditory cortical neurons are affected by multiple excitatory or suppressive features and identify a specific relationship between the relevant features. We found that excitatory features encode finer temporal and spectral details compared to the broader contexts encoded by suppressive filters. We further found that the number of identifiable features depended on stimulus context. Thus, our results show that stimulus processing depends on feature type and stimulus context.

EXPERIMENTAL PROCEDURES

Surgical procedures

All experimental procedures were approved by the University of California, San Francisco Committee for Animal Research under protocol AN086113-01B. The experimental procedures used in this study have been previously described (Atencio and Schreiner, 2010b,a). Briefly, young female adult cats ($N = 4$) were given an initial dose of ketamine (22 mg/kg) and acepromazine (0.11 mg/kg), and then anesthetized with pentobarbital sodium (Nembutal, 15–30 mg/kg) during the surgical procedure. The animal's temperature was maintained with a thermostatic heating pad. Bupivacaine was applied to incisions and pressure points. Surgery consisted of a tracheotomy, reflection of the soft tissues of the scalp, craniotomy over AI, and durotomy. After surgery, to maintain an areflexive state, the animal received a continuous infusion of ketamine/diazepam (2–5 mg/kg/h ketamine, 0.2–0.5 mg/kg/h diazepam in lactated Ringer solution).

Recording

With the animal inside a sound-shielded anechoic chamber (IAC, Bronx, NY), stimuli were delivered via a closed speaker system to the ear contralateral to the exposed cortex (electrostatic diaphragms, model SRX MK3, from Stax, Japan). The system frequency transfer function was nearly flat ($\pm$6 dB) for frequencies $\leq$14 kHz, and attenuated 10 dB/octave for frequencies above 14 kHz.

Extracellular recordings were made using linear multichannel silicon recording probes, which were provided by Neuronexus (Michigan). We used probes with channel impedances between 2 and 3 MΩ, since these impedances allowed us to resolve single units. Probes were carefully inserted using a microdrive (David Kopf Instruments, Tujunga, CA, USA) into the center of the ectosylvian gyrus, allowing for recording away from the anterior and posterior ectosylvian sulci. The cortical depth position of each recorded neuron was estimated from microdrive readings, which have previously been shown to allow for accurate laminar estimates (Atencio and Schreiner, 2010b).

Neural traces were bandpass filtered between 0.6 and 6 kHz and recorded to disc with a Neuralynx Cheetah A/D system at sampling rates between 18 kHz and 27 kHz. The traces were sorted off-line with a Bayesian spike sorting algorithm (Lewicki, 1994, 1998; Atencio and Schreiner, 2013). The average over the entire recording trace was estimated, and only events in the traces that exceeded the average by 5 RMS noise levels were used in the spike sorting procedure. All recording locations were in AI, as verified through initial multi-unit mapping and determined by the layout of the tonotopic gradient and bandwidth modules on the crest of the ectosylvian gyrus (Imaizumi and Schreiner, 2007).

Stimulation

All neurons were also probed with a broadband (0.5–40 kHz) dynamic moving ripple (DMR) stimulus (Escabi and Schreiner, 2002; Atencio et al., 2008). The maximum spectral modulation frequency (SMF) of the DMR was 4 cyc/oct, and the maximum temporal modulation frequency was 40 cyc/s (Escabi and Schreiner, 2002). The maximum modulation depth of the spectrotemporal envelope was 40 dB. Mean intensity was set at 30–50 dB above the average pure tone threshold. A subset of neurons was probed with a ripple noise (RN) stimulus. The RN is the sum of 16 independently created DMRs, and therefore it has the same carrier structure and modulation depth as the DMR. The RN and DMR differ in short-term, but not long-term, correlations: the DMR has short-term correlations but no long-term correlations, while the RN has neither short-term nor long-term correlations (Escabi and Schreiner, 2002). The duration of each stimulus was either 10 min or 15 min.

Receptive field estimation and analysis

The stimulus envelope was sampled at 5-ms resolution in time and six carriers per octave in frequency. For each neuron, we first chose a set of 25 frequencies and 20 time bins that encompassed the stimulus bandwidth and history to which the neuron responded (500 stimulus dimensions per feature). We then applied the binary noise MNE analysis described in (Fitzgerald et al., 2011a,b). The MNE analysis code is available on Github: http://github.com/MarvinT/pyMNE. Briefly, for a given stimulus **s**, the probability of a spiking response was modeled through a logistic nonlinearity having the form:

$$P(spike|s) = \frac{1}{1 + \exp(-(a + s \cdot h + s^T \cdot J \cdot s))}$$

The parameters a, **h**, and **J** were estimated so that the model matched the experimentally observed mean firing rate, the spike-triggered average (STA) statistics, and STC statistics. Maximum likelihood estimation was used to determine a, **h**, and **J** while remaining as unbiased as