

POSTER PRESENTATION

Open Access

Spiking neural network model of cortical auditory source segregation

Lakshmi Krishnan^{1*}, Michael Campos², Shihab Shamma^{1,3}

From The Twenty Third Annual Computational Neuroscience Meeting: CNS*2014
Québec City, Canada. 26-31 July 2014

Humans have the remarkable ability to tune into a particular voice even in loud noisy environments. The neural underpinnings of this amazing perceptual phenomenon are not yet fully understood. Recent EcoG [1] and MEG studies [2] have established that the neural representation of the attended speaker's speech is much stronger than the unattended (distractor) speech when human subjects are asked to pay attention to a target speaker in a mixture of speech. How the brain sieves through the mixture waveform to enhance the target speaker's speech and attenuate the background acoustic scene is still being investigated. In this work, we propose a spiking neural network architecture based on the theory of temporal coherence [3] to achieve auditory source segregation.

Our model does not require training on the background noise or prior exposure to the target speech. Along with using bottom-up spectro-temporal features and pitch features, the model can also accommodate top-down attentional mechanisms to generate segregated phase locked neural representations to target speaker's speech envelope. The model comprises of a feature extraction stage followed by clustering stage. The feature extraction stage mimics the auditory pathway starting from a cochlear representation followed by a multi-resolution analysis of the cochlear output using a bank of band-pass filters (cortical stage), to provide a rich timbre representation. Dominant pitch tracks are extracted from the sound mixture and processed through the same set of band-pass filters as the timbre channels. The output of the feature extraction stage comprising of the pitch and timbre channels are transduced into a spike-based representation using leaky integrate and fire neurons with

time constants tuned to the bandwidth of the multi-resolution band-pass filters. The clustering stage comprises of a bank of coincidence detector neurons. Using the pitch signals as anchors the coincidence detector neurons can segregate the two sources from the mixture timbre representation. Thus, the output of the coincidence detector neurons comprises only of responses phase locked to the envelope of a single source.

This model does not require any weight learning, is unsupervised and can segregate sources online. Previous studies on correlation based sound segregation employed network of neurons with intrinsic oscillator dynamics [4]. In this work, clustering of features belonging to a single source is driven only by the temporal coherence of spectro-temporal features of the given source. This spike-based representation provides an easy mechanism to group coherent features, which otherwise would require computationally expensive numerical routines for online, adaptive principal components analysis. Future work is aimed at reconstructing the speech waveform from the segregated spike trains.

Acknowledgements

The spike-based architecture for source segregation was developed when the first author was interning with Qualcomm Research, San Diego.

Authors' details

¹Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20783, USA. ²Qualcomm Research, San Diego, CA 92121, USA. ³Department Etude Cognitive, Ecole Normale Supérieure, Paris 75005, France.

Published: 21 July 2014

References

1. Mesgarani N, Chang E: **Selective cortical representation of attended speaker in multi-talker speech perception.** *Nature* 2012, **485**:233-236.
2. Ding N, Simon JZ: **Emergence of neural encoding of auditory objects while listening to competing speakers.** *Proceedings of the National Academy of Sciences* 2012, **109**:11854-11859.

* Correspondence: lakshmik@umd.edu

¹Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20783, USA

Full list of author information is available at the end of the article

3. Shamma S, Elhilali M, Micheyl C: **Temporal coherence and attention in auditory scene analysis.** *Trends in Neurosciences* 2011, **34**:114-123.
4. Wang D, Brown G: **Separation of speech from interfering sounds based on oscillatory correlation.** *IEEE Transactions on Neural Networks* 1999, **10**:684-697.

doi:10.1186/1471-2202-15-S1-P50

Cite this article as: Krishnan *et al.*: Spiking neural network model of cortical auditory source segregation. *BMC Neuroscience* 2014 **15**(Suppl 1): P50.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

