Sound Decoding from Auditory Nerve Activity

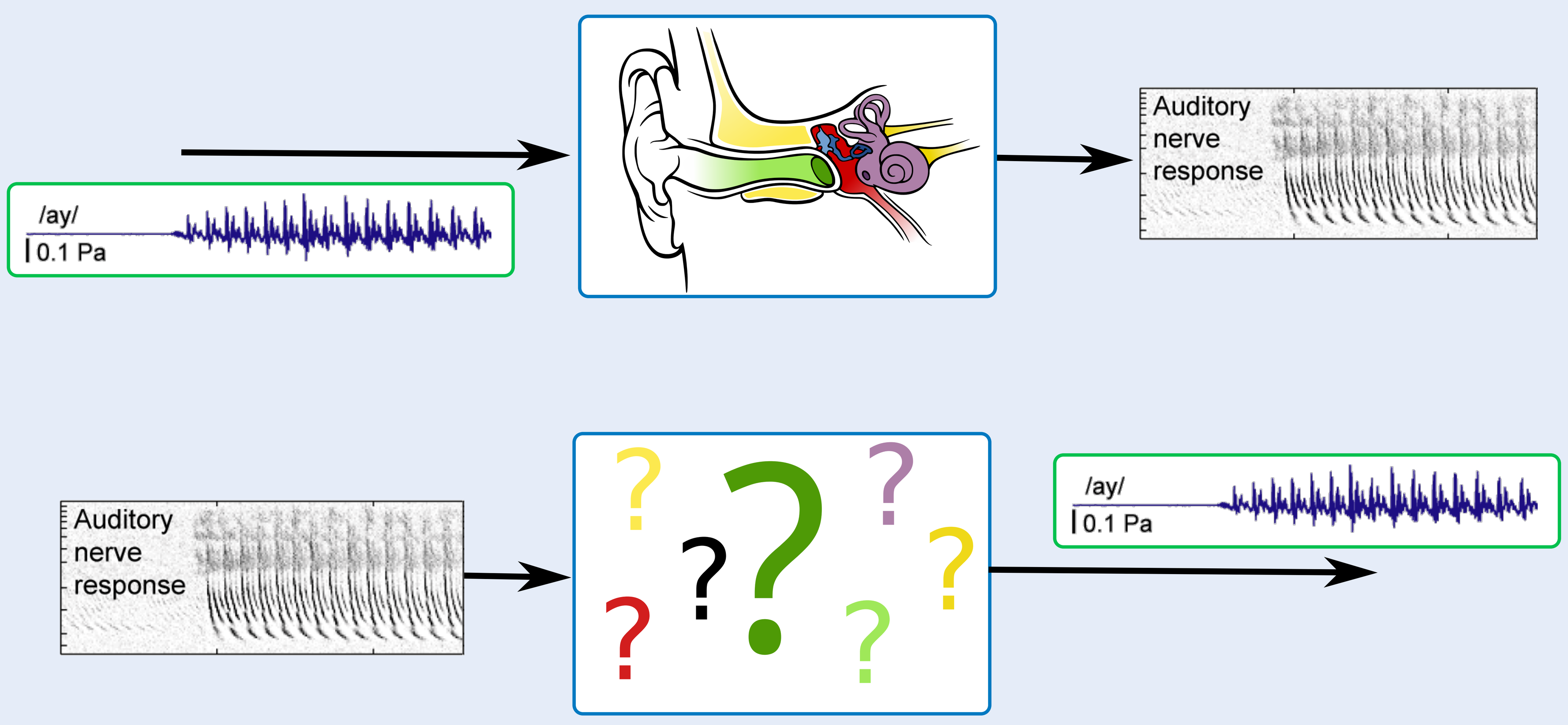
Marek Rudnicki, Marcelo K. Zuffo and Werner Hemmert

Bio-Inspired Information Processing, Technische Universität München
Laboratory of Integrated Systems, University of São Paulo

Abstract

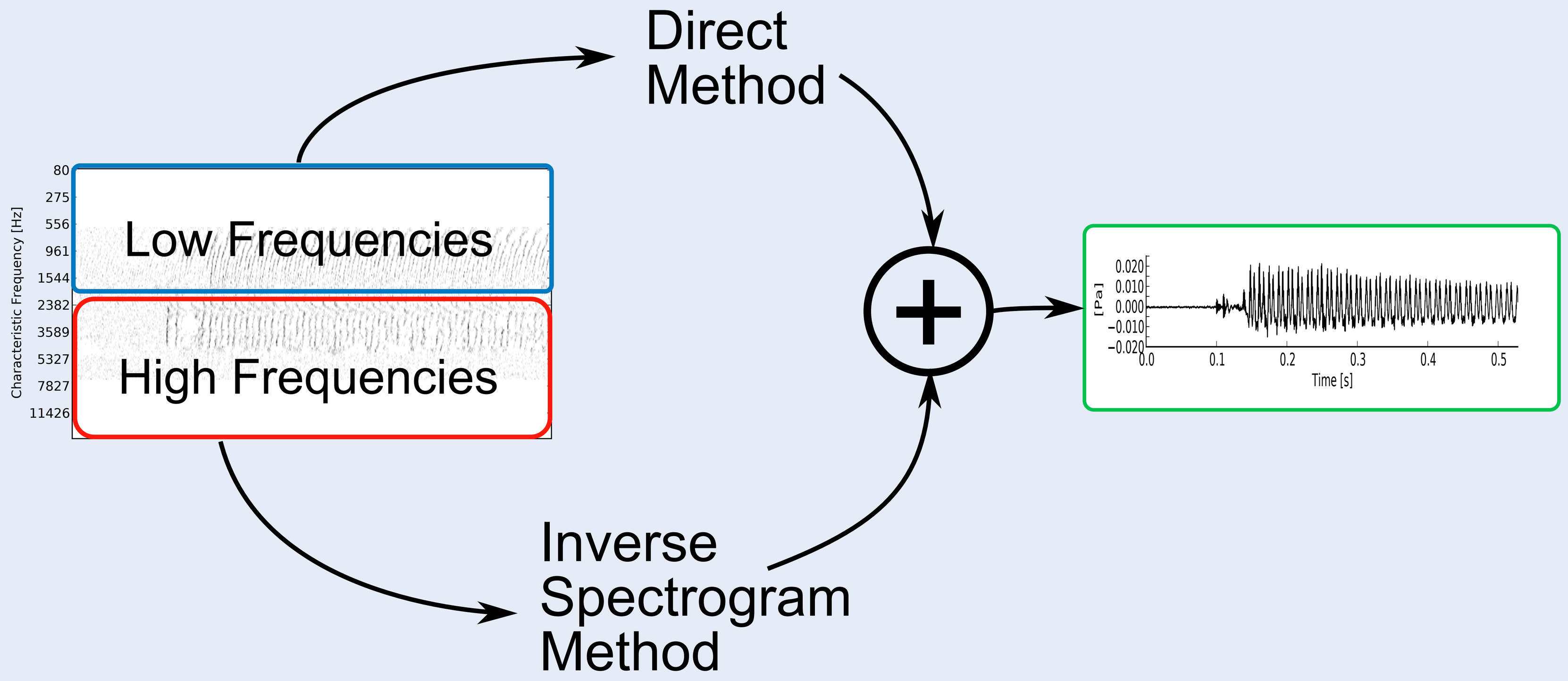
In the inner ear sounds are converted to discrete action potentials and sent to the central nervous system. This transformation is non-linear and results in massive information loss. However, we still can hear and analyze sounds with high fidelity. This is because the crucial features of sounds are still present in the auditory nerve signals.

Here we present a method which decodes sounds from a large population of simulated auditory nerve fibers (ANFs). We also use the procedure to reconstruct sounds from a model of an impaired cochlea. This way we are able to mimic how hearing impaired subjects perceive sounds.

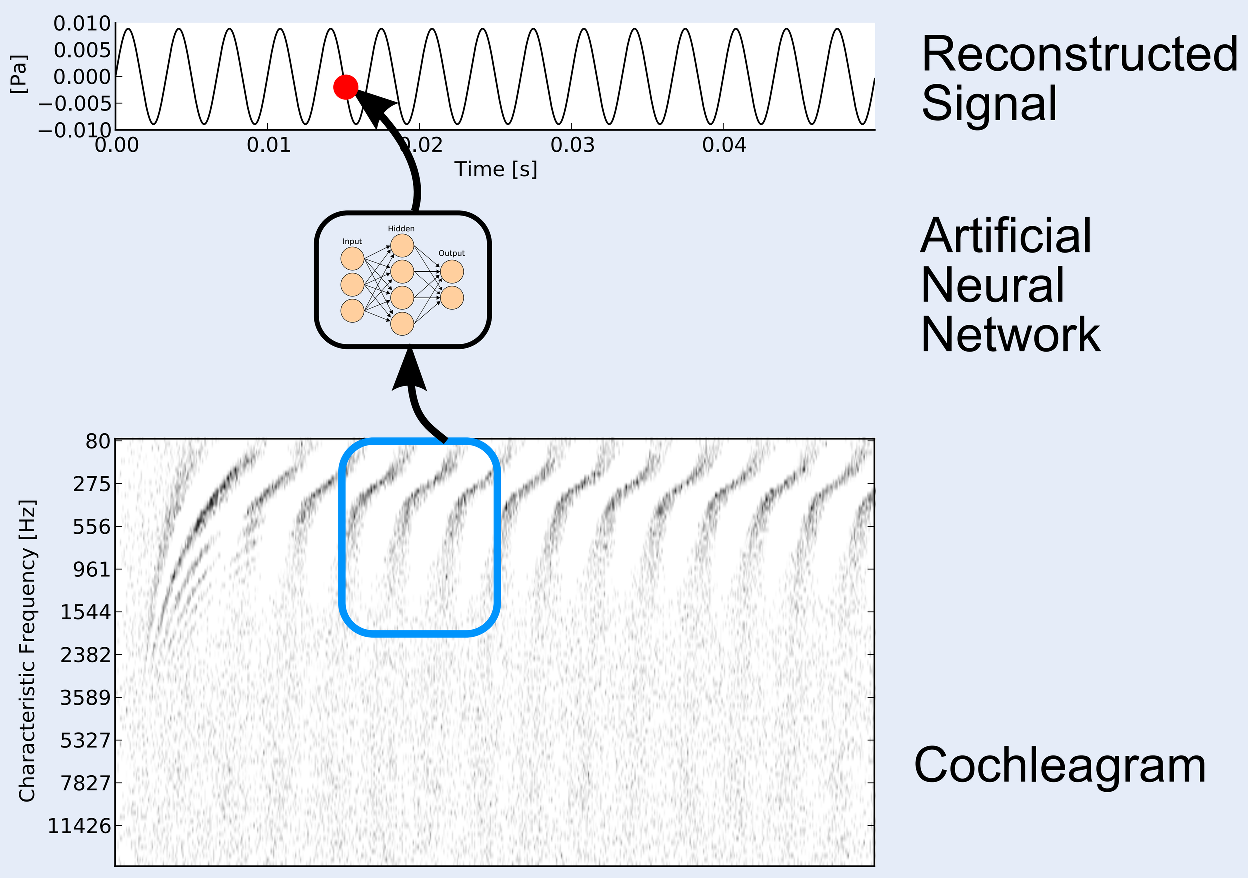


Methods

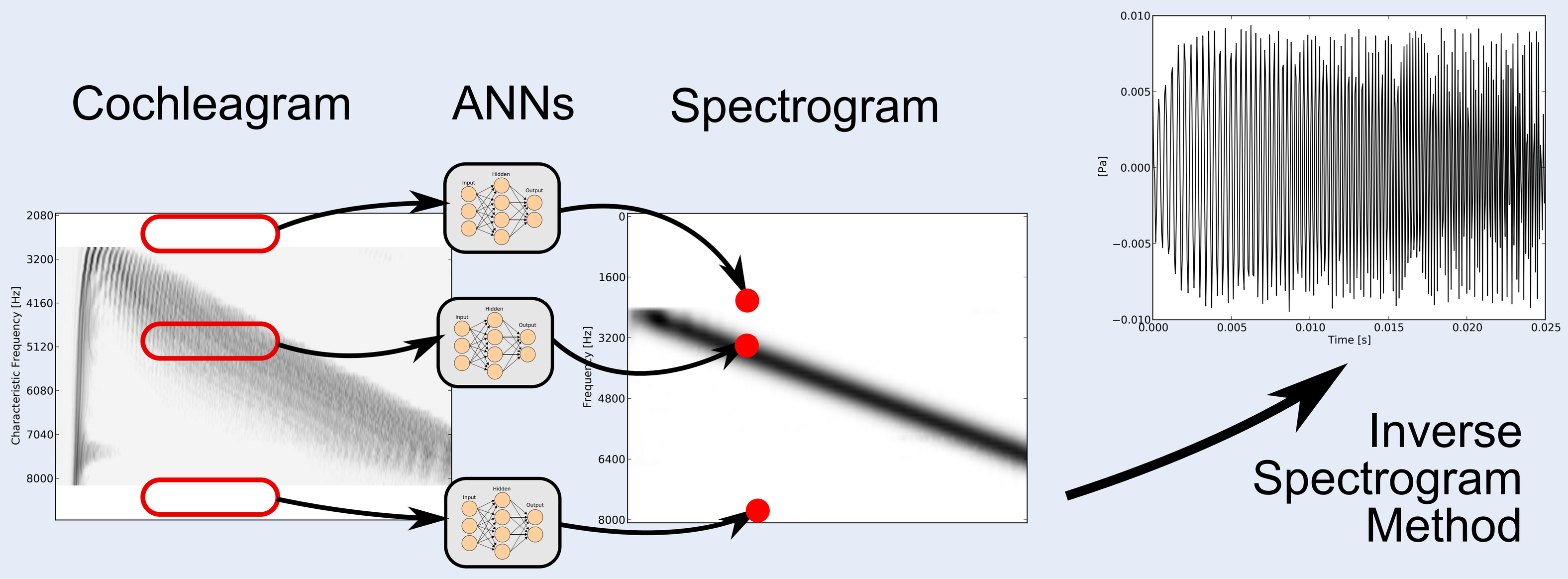
Our approach leverages responses of a large population of ANFs (close to the number present in the human ear) and non-linear reconstruction using an artificial neural network (ANN). In the reconstruction we employed dual approach: 1) direct ANN reconstruction for frequencies below 2kHz and 2) inverse spectrogram reconstruction for frequencies above 2kHz. To generate ANF responses we used the biophysical auditory periphery model from Zilany et al. (2009). The model was adopted to replicate the human hearing range and thresholds.



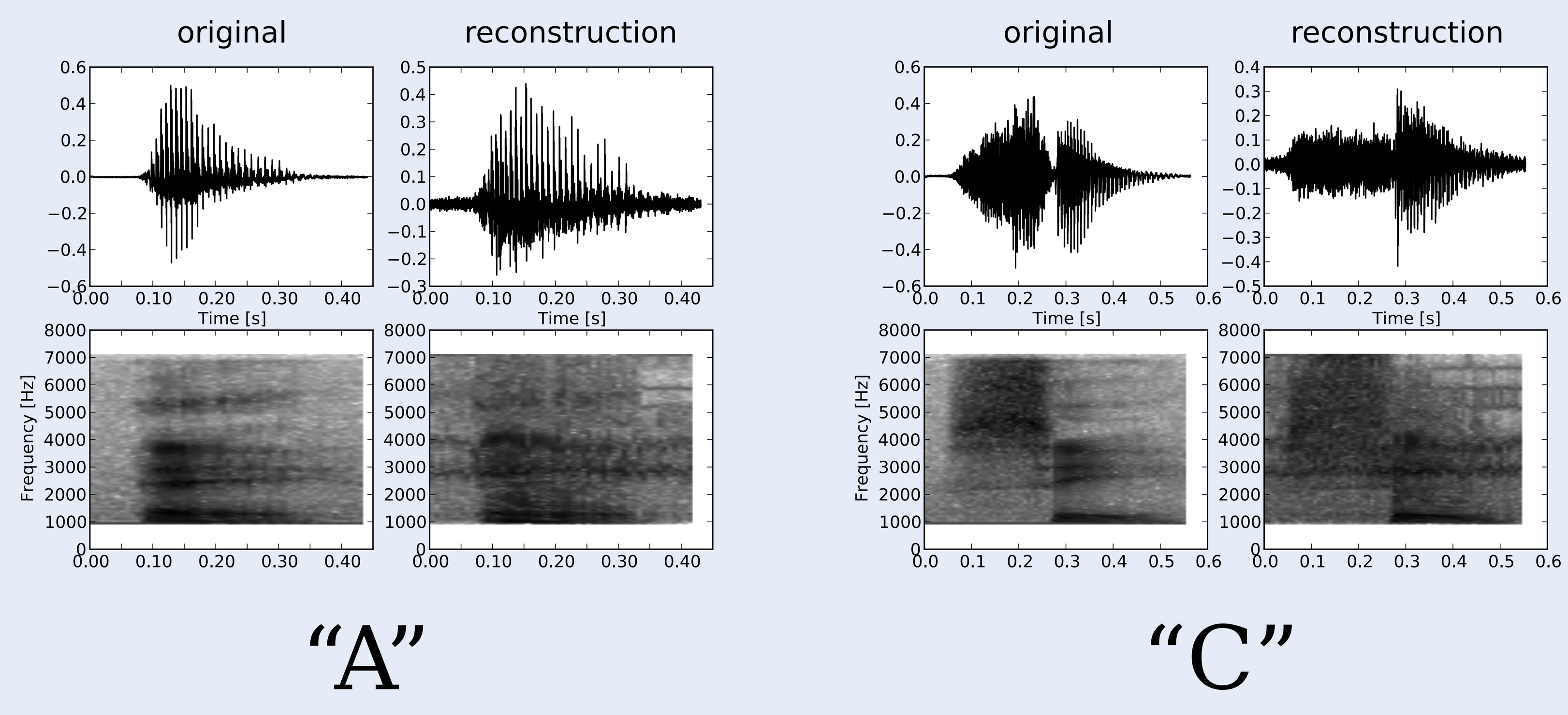
The particular ANN used was a multi-layer perceptron (MLP) with a single hidden layer. The input to the MLP was a 10 ms sliding window from multiple spike trains across 10 different characteristic frequencies. The output was a single value of the reconstructed signal. In this way we trained and tested the MLP with sounds below 2 kHz.



The direct MLP approach did not work for frequencies above 2 kHz, because of the lacking phase information (phase locking) in spike trains. We therefore developed a two-stage algorithm to reconstruct high frequency signals. First, spike trains were converted to a spectrogram by MLPs. Second, the spectrogram was transformed to an acoustic signal using an iterative method (Decorsiere et al. 2011). In order to convert spike trains to a spectrogram we trained 51 MLPs. The input to each MLP was a sliding window of 5 ms from multiple spike trains and the output one of 51 frequency channels of a spectrogram.



Demo and Results



Acknowledgment

This work was supported by the German Federal Ministry of Education and Research within the Munich Bernstein Center of Computational Neuroscience (reference number: 01GQ1004B) and TUM Graduate School.

Conclusions

- decoding of sounds from AN activity is possible
- speech quality is understandable and well perceived

marek.rudnicki@tum.de