Unveiling the auditory system with a neural network

Abstract

Although there has been much research on human auditory characteristics, it is difficult to directly address the question of what kinds of input and training lead to the acquisition of these characteristics. In this work, we tackled the clinical and academic aspects of the question by using artificial neural networks (ANNs), and obtained new findings in each case. (1) It is known that people with hearing loss who wear cochlear implants (CIs) have difficulty with pitch perception, but we confirmed that the cochlear implant signal contains a certain amount of pitch information, suggesting that the difficulty in pitch perception is mainly due to physiological factors. (2) By measuring the response of a single unit in an artificial neural network trained to recognize natural sounds, we found out the ANN units (neurons) with the binaural processing characteristics were equivalent to those found in the auditory system of animals. We believe that cochlear implant users may be able to achieve normal pitch perception under a clean environment after an appropriate rehabilitation. We also hope to further develop AI technology and CI devices that behave in a human-like manner by advancing auditory information processing technology that is consistent with the auditory nervous system.

Understanding auditory mechanisms with artificial neural networks (ANNs)

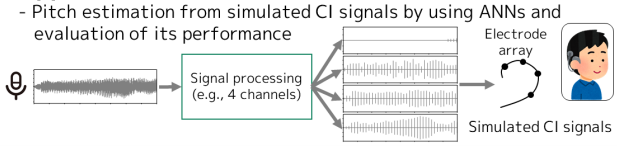
Topic (1): Pitch information transmitted by the cochlear implant (CI)※

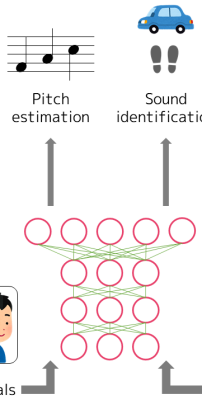
※ An artificial organ to partially restore hearing loss caused by damage to the inner ear

Background:
- Cochlear implantation significantly restores speech perception
- Pitch perception is difficult and varies considerably between individuals

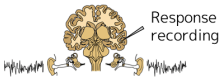
Question:
- Does the signal transmitted by the CI contain the information necessary for pitch estimation?

Approach:
- Pitch estimation from simulated CI signals by using ANNs and evaluation of its performance





Topic (2): Characteristics of neuronal response to binaural sound



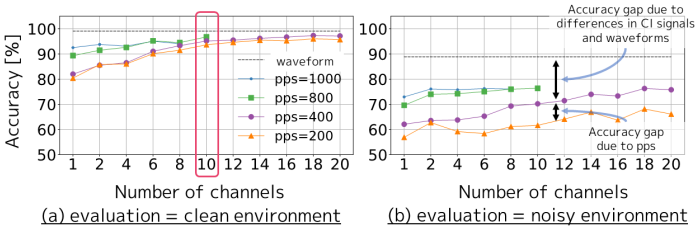
Background:
- The time differences between when sound reaches the left and right ear (interaural time difference (ITD)) is important for sound source localization and identification
- Neuronal response characteristics have been investigated

Question:
- How are response characteristics to ITD acquired?

Approach:
- Training ANNs to identify sound sources from natural binaural sounds and measuring single cell responses to stimuli with ITD

(1) A certain amount of pitch information is contained in the cochlear signal

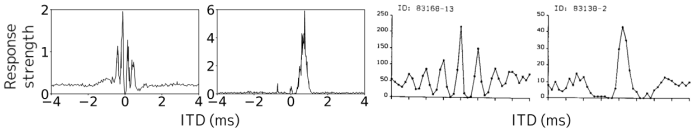
- When the target sound is presented alone (clean environment (a)), if CI signals have a sufficient number of channels (almost 10), the accuracy is comparable to that of the waveform
  - Containing pitch information in CI signals comparable to that in the waveform under a clean environment
- In the presence of background noise (noisy environment (b)), the accuracy of CI signals is worse than that of the waveform, and improves as pulse per second (pps) increases
  - Pitch perception becomes difficult under noisy environments
  - Finer temporal resolution plays an important role



- The difficulty in pitch perception is more likely due to physiological factors than to the signal transmitted by the CI device

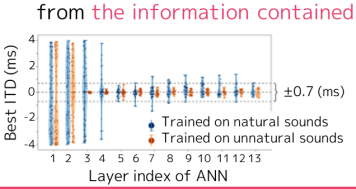
(2) Emergence of ITD response characteristics in natural sound identification

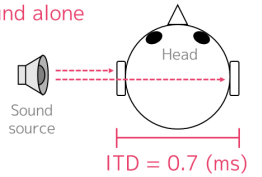
- Response strength of ANN neurons varies with stimulus ITD
  - ITD response characteristics qualitatively similar to those of animals are also evident in ANNs



(Yin et al., 1986, J. Neurophysiol.)

- ANN shows high response strength in the range of ITDs that humans naturally experience
  - No use of human body shape information for ANN training
  - The range becomes narrower when trained on unnatural sounds
  - ANN captures the natural environmental structure for humans from the information contained in sound alone





References

[1] T. Ashihara, S. Furukawa, M. Kashino, “F0 estimation from simulated cochlear-implant signals by using a DNN model,” *Spring Meeting of Acoustic Society of Japan*, 2022.

[2] T. Koumura, H. Terashima, S. Furukawa, “Emergence of ITD selectivity in a deep neural network trained for binaural natural sound detection,” in *Proc. 42nd Association for Research in Otolaryngology (ARO) MidWinter Meeting*, 2019.

[3] TC. Yin, JC. Chan, DR. Irvine, “Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise,” *Journal of Neurophysiology*, pp. 280–300, 1986.

Contact

Takanori Ashihara / Human Informatics Laboratories
Takuya Koumura / Sensory Representation Research Group, Human and Information Science Laboratory
Email: cs-openhouse-ml@hco.ntt.co.jp