A NOISE-ROBUST SIGNAL PROCESSING STRATEGY FOR COCHLEAR IMPLANTS USING NEURAL NETWORKS

Nengheng Zheng[1,2], Yupeng Shi[3], Yuyong Kang[1], Qinling Meng[4]

[1]The Guangdong Key Laboratory of Intelligent Information Processing, College of Electronics and Information Engineering, Shenzhen University, China; [2]Peng Cheng Laboratory, Shenzhen, China ; [3]Tencent Media Lab, Shenzhen, China; [4]Acoustic Lab., School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China

E-mail: nhzheng@szu.edu.cn

Abstract

Background:
Most clinical cochlear implants (CIs) extract and transmit coarse speech envelopes to stimulate the auditory neurons that help recipients restore partial hearing ability. The incomplete representation of the rich fine structures in speech has significantly degraded the CI recipients' ability in high-level perception, including their speech understanding in noise.

What we do:
This paper proposed a neural network-based CI strategy, namely NNACE, which has the following features:

- compatible with Nucleus ACE-based CI system and can serve as the modulator to generate the electric stimuli
- more noise-robust
- might bear a certain degree of the temporal fine structures of speech

Results:
Subjective and objective evaluations with vocoder simulated speech show that NNACE outperforms the other methods and further actual CI experiments are warranted.

Methods

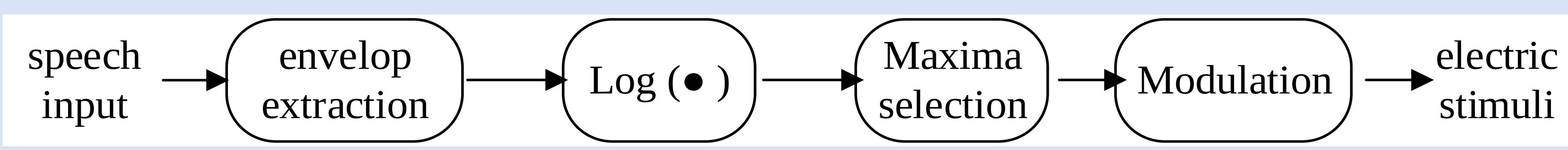

Fig.1 Block diagram for the conventional ACE

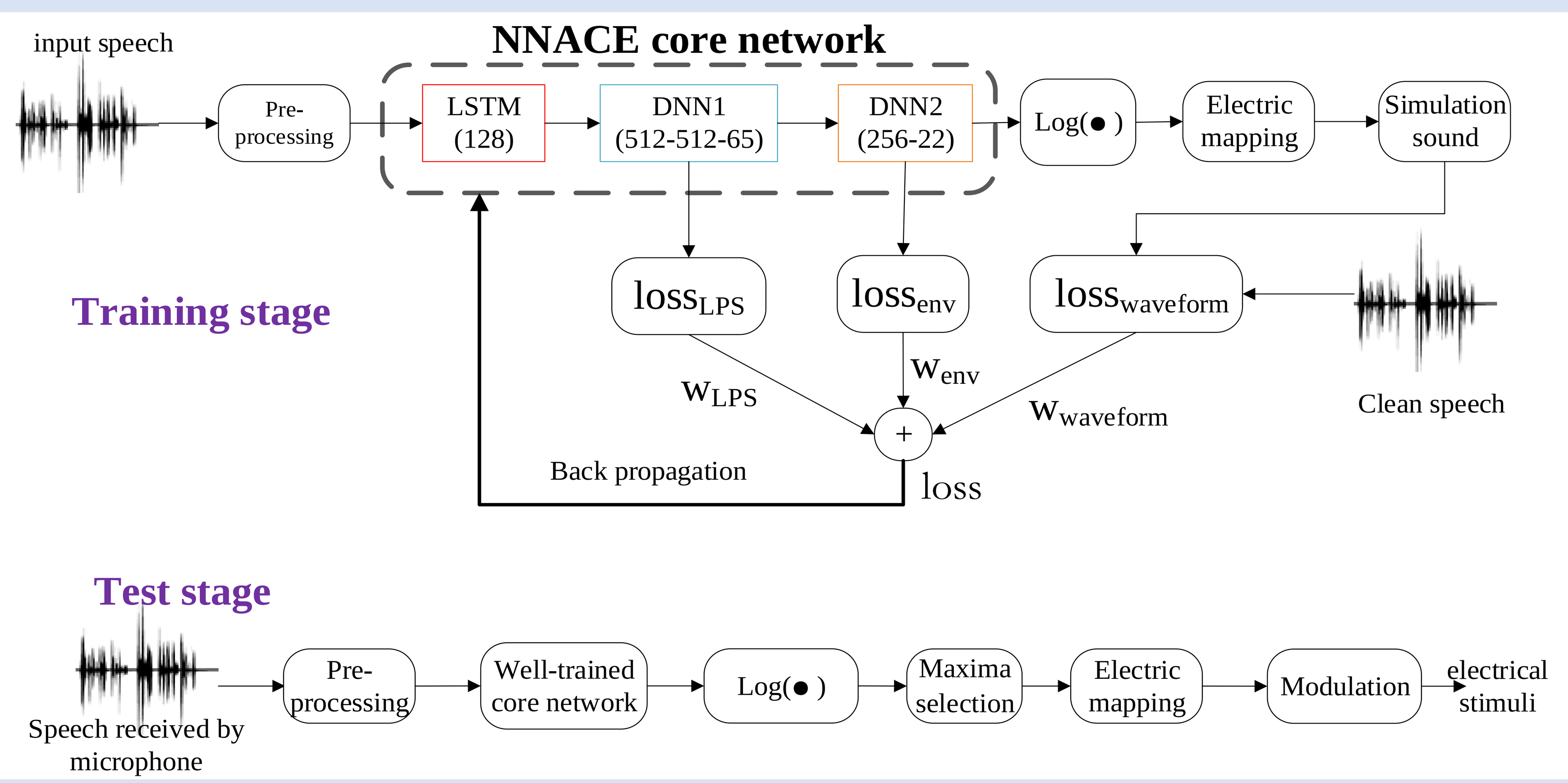

Fig.2 Systematic block diagram of the proposed NNACE

Note: the loss function are calculated as follow equations

$$loss_{LSP} = \sum_t \sum_f \left\| X(t,f) - \hat{X}(t,f) + log(|w(t,f)| + 10^{-8}) \right\|_1 \quad (1)$$

$$loss_{ENV} = \sum_t \sum_b \left\| E(t,b) - \hat{E}(t,f)) \right\|_1 \quad (2)$$

$$loss_{VOC} = \sum_n \left\| x(t,n) - \hat{x}(t,n)) \right\|_1 \quad (3)$$

$$loss = w_{LPS} \cdot l_{LPS} + w_{ENV} \cdot l_{ENV} + w_{VOC} \cdot l_{VOC} \quad (4)$$

Where the weights $w_{LPS}, w_{ENV}\ w_{VOC}$ are set to be 0.7, 0.3, and 1

Experiments

Pre-process: 128 points STST (16kHz sampling, frames size 8ms, shift 1ms)
Neural network setting: neural units as shown in Fig. 2 , Adam optimizer, learning rate 0.0001,
Baseline: ACE, wiener filtering as front-end to ACE (**Wiener-ACE**), DNN as front-end to ACE (**DNN-ACE**).

database
- speech
  - THCHS-30 for NN training and objective test
  - MHINT-M for subjective test
- noise: speech shape noise (SSN), Babble, from NoiseX-92

SNR
- Train data: 0dB, 5dB, 10dB, and ∞
- Objective test data: -5dB, 0dB, 5dB, 10dB, and ∞
- Subjective test data: -5dB to 15dB, in steps of 2dB

Note: for training, each clean speech mix with noise at a randomly select SNR;
for objective and subjective test, each clean speech mix with noise at all SNRs.

Evaluation
- **objective**
  - weighted spectral slope (WSS), the lower the better
  - normalized covariance metric (NCM), the higher the better
- **Subjective**
  - preference test
  - speech reception threshold (SRT), the lower the better

Results & Discussion

Network complexity (model parameters)
DNN-ACE: 0.57Mb; NNACE: 0.49Mb

- The complexity of NNACE and DNN-ACE are comparable

Objective evaluation (with vocoder speech) results

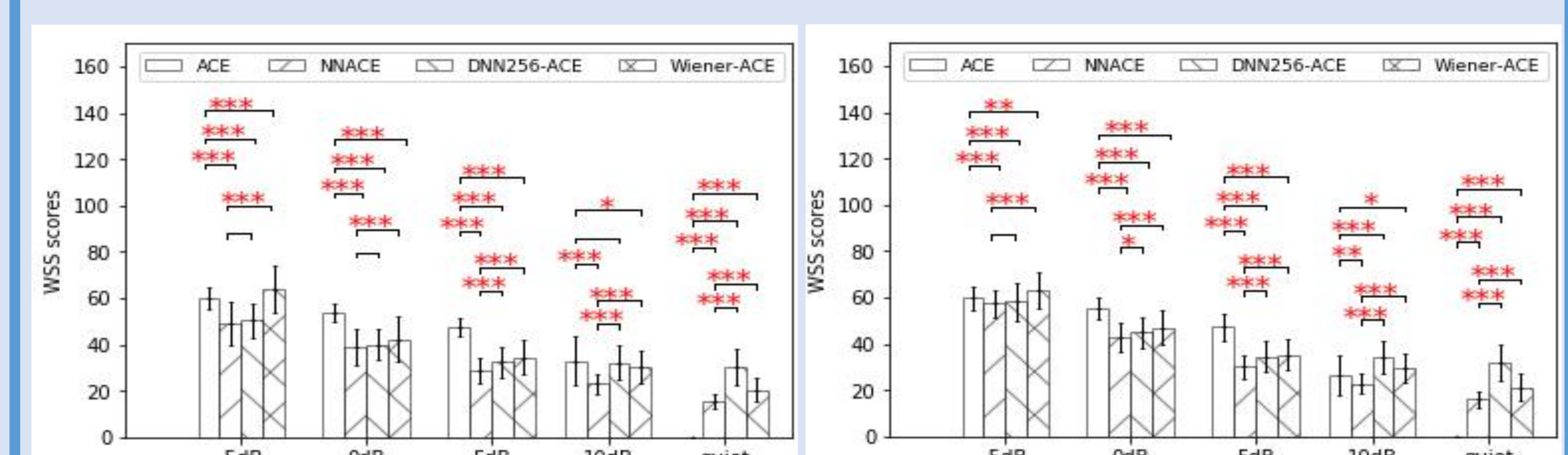

Fig.3 WSS scores and significance between systems, at SSN (left) and Babble (right) noises

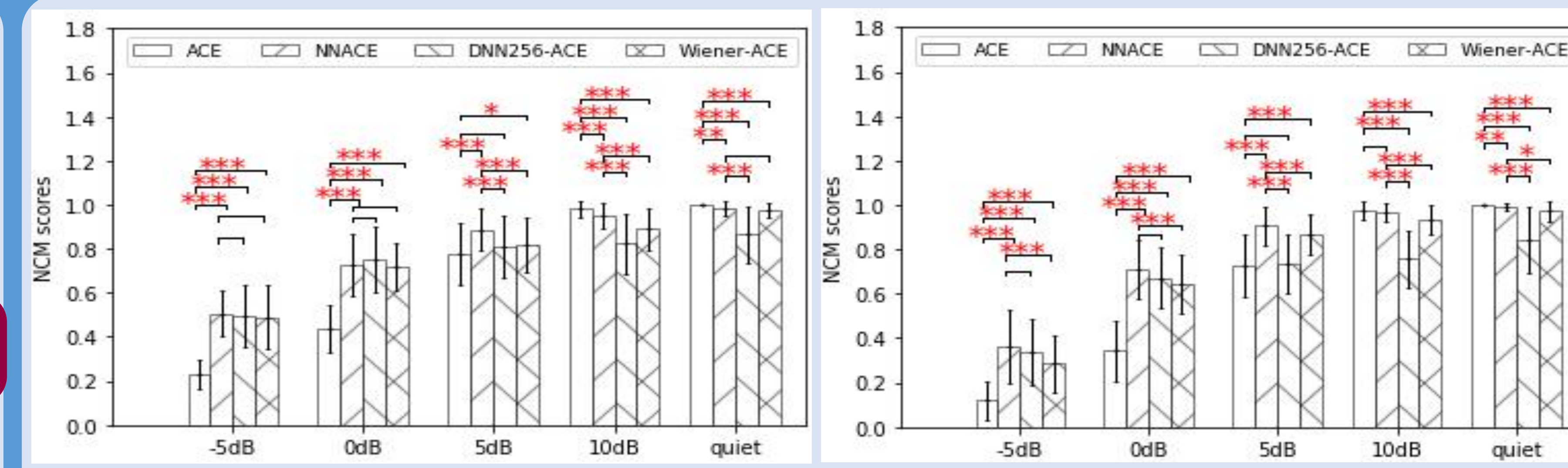

Fig.4 NCM scores and significance between systems, at SSN (left) and Babble (right) noises

- NNACE significantly outperforms ACE and Wiener-ACE in most noisy conditions
- The two NN-based systems have comparable performances at low SNRs
- NNACE significantly outperforms DNN-ACE at high SNRs

Subjective evaluation (with vocoder speech) results

Table1. Listeners' preference on ACE- and NNACE-outputs in noise-free conditions

ACE	Comparable	NNACE
28.75%	41.25%	30%

- The subjects have the similar preference on ACE and NNACE in quiet.

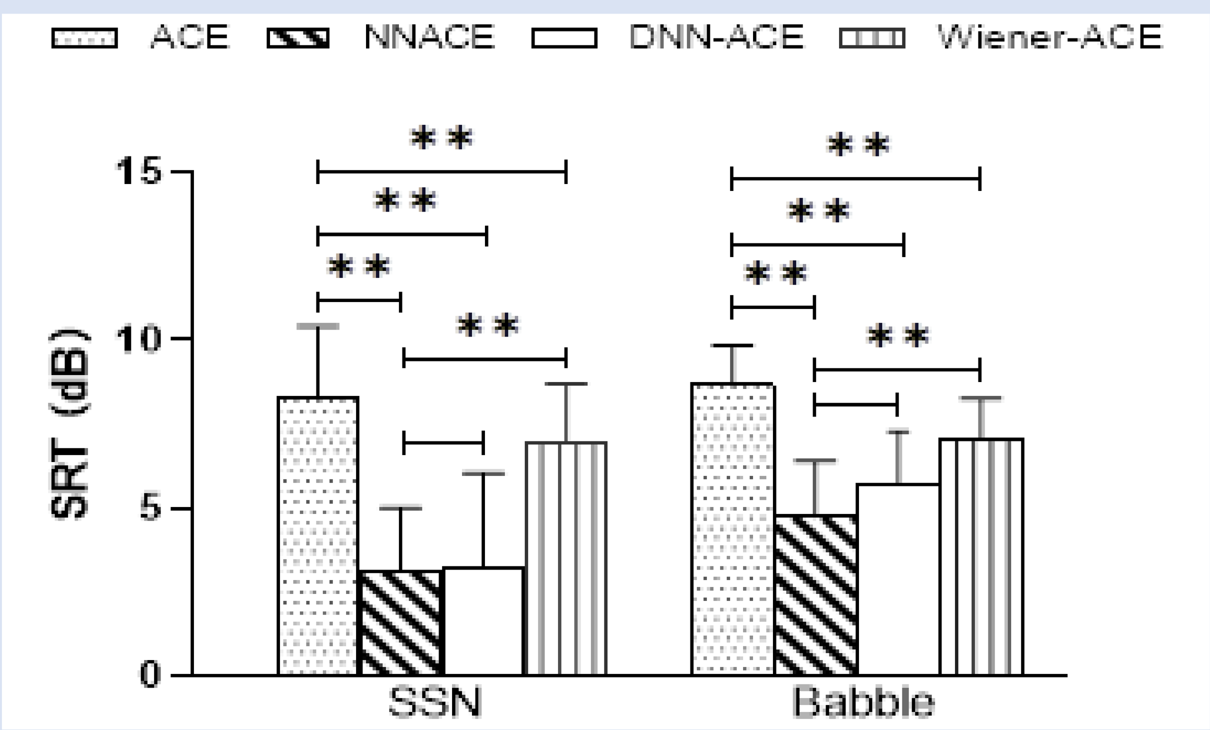

Fig.5 Mean SRTs (averaged over all participants) and significance between systems

- All SE systems significantly improve speech intelligibility in noisy envelopment
- NNACE outperforms Wiener-ACE
- Two NN-based systems obtained the similar performance in each noise type

Conclusions & Acknowledgement

- A noise-robust neural network-based strategy NNACE is proposed for CI signal processing.
- Both objective and subjective evaluations with vocoded speech imply that NNACE has potential to improve CI speech understanding in noisy conditions.

This work is jointly supported by Guangdong Key Area R&D Project (No. 2018B030338001), National Natural Science Foundation of China (61771320, 11704129).