1479: CCI-MOBILE: AUTO-LSP BASED SPEECH ENHANCEMENT WITH COCHLEAR IMPLANT LISTENERS USING CONVOLUTIONAL NEURAL NETWORK CONSTRAINT MAPPING

Nursadul Mamun, John H. L. Hansen

Center for Robust Speech Systems (CRSS): Cochlear Implant Processing (CI) Lab
Erik Jonsson School of Engineering & Computer Science,
University of Texas at Dallas, Richardson, Texas, U.S.A.

(Nursadul.Mamun, John.Hansen)@utdallas.edu

Cochlear Implant Laboratory

1. INTRODUCTION

- **Observation:** Cochlear Implant (CI) recipients enjoy near-to-normal speech intelligibility (SI) in quiet conditions [1].
- **Challenge:** CI recipients generally experience reduced speech intelligibility in environmentally rich spaces due to limited frequency resolution [2].
- **Iterative Wiener Filter:** traditional iterative Wiener filtering is not always stable [3,4]
- **Auto-LSP based SE:** apply speech production and auditory based time-frequency spectral constrains to achieve improved speech quality [4,5].
- **CNN-based Speech Enhancement (SE):** considers local temporal-spectral structures of speech signals and effectively dis-integrates the clean LSP parameters and noise from noisy input.
- **Proposed networks:** deploy a CNN to predict the enhanced line-spectral pair (LSP) parameters from their noisy version.

2. METHODS

- **Method#1:** Calculates LPC and LSP parameters from each input signal frame over time
- **Method#2:** Apply Inter- and Intra-frame constraints to LPC parameters.

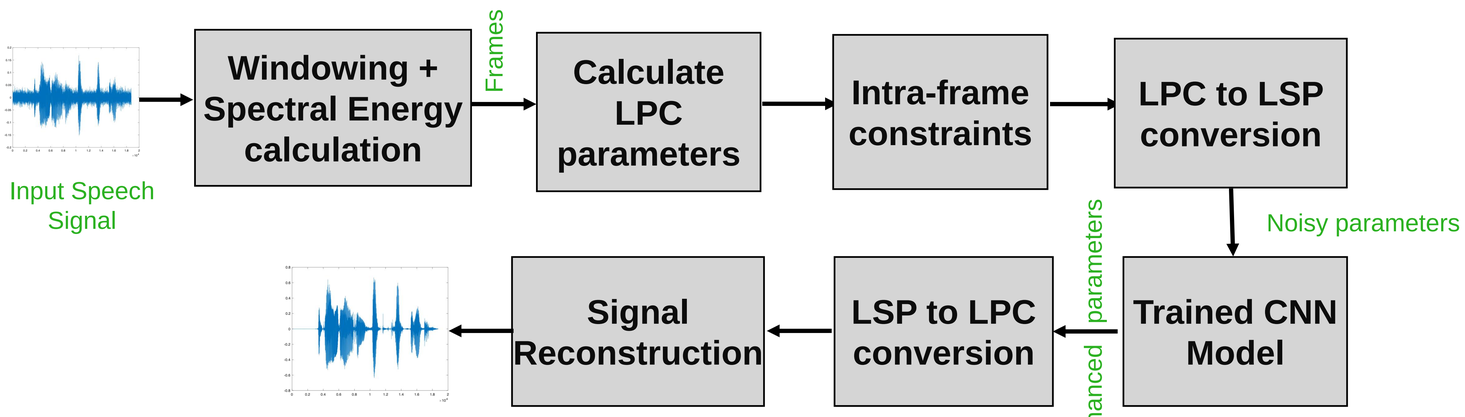


Fig.1. Basic block diagram of Auto-LSP based SE algorithm.

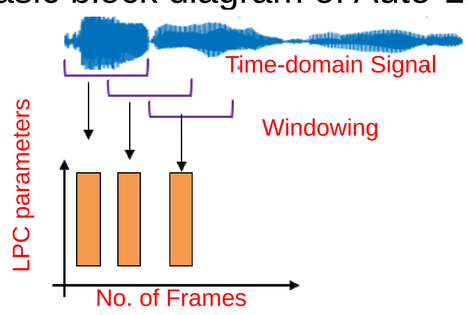


Fig.2. (a) Intra-frame constraints[4,5]

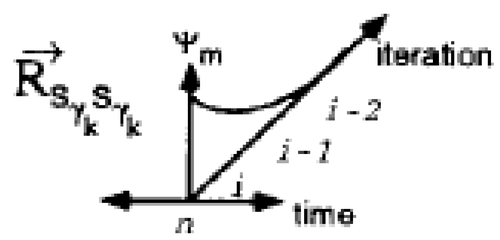


Fig.2. (b) Inter-frame constraints [4,5]

3. METHODS cont.

- **Method#3:** Deploy a CNN network to predict enhanced LSP parameters
- **Method#4:** A constrained Weiner filter is used to predict the clean speech signal from the enhanced time-frequency speech constrained LPC parameters.

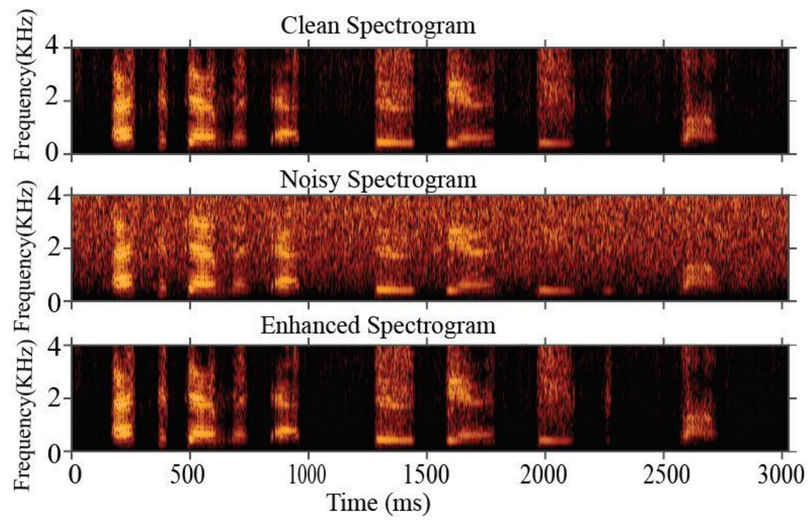
CNN Parameters:

- Training/Testing Sets: 5000/500 utterances
- Convolutional Layer: 5.
- Epochs: 100. Kernels: 17
- Activation Function: 'ReLu'

Test Corpora:

- Database: TIMIT sentence corpus
- SNR: 10 and 5 dB.
- Noise: Speech shape noise (SSN).

4. RESULTS



Intelligibility/Quality scores:

	IS	STOI	SOPM	PESQ
Noisy	4.19	0.722	0.717	1.432
Enhanced	2.42	0.757	0.965	1.941

Effect of Wiener filter parameter : Beta

	IS	STOI	SOPM	PESQ
Beta = 0.5	3.34	0.74	0.91	1.63
Beta = 1.5	2.42	0.757	0.965	1.941

Enhanced signal provides improved intelligibility with Beta = 1.5

5. CONCLUSIONS

- Proposed algorithms shown to successfully enhanced signals under noisy conditions for CI users.
- Advancements show processing time requirements are less for system convergence to optimum enhanced signal
- Wiener filter coefficient values have a great impact on output enhanced signal quality
- Best Intra-frame constraints are found for the combination: 0.3 and 0.7.

REFERENCES

1) F.G. Zeng; S. Rebscher (**2008**). *IEEE reviesw in Biomedical Engineering 1, **115**.*
2) L. M. Friesen, R. V. Shannon, (**2001**) *J. Assoc. Res. in Otolaryngology* 110(2):1150.
3) J. S. Lim, A.V. Oppenheim, (**1978**) *IEEE Trans. ASSP*, 26(3):**197-210.**
4) J.H.L. Hansen; M. A. Clements (**1987**). *IEEE ICASSP-1987, **189-192**.*
5) J.H.L. Hansen; M. A. Clements (1991) *IEEE Trans. Signal Proc.,* 39(4), **795-805.**

2021 Conference on Implantable Auditory Prostheses
Lake Tahoe, CA (Virtual), July 12 – 16, 2021

This work was supported by the grant R01 DC016839 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.