Environment-Aware Reconfigurable Noise Suppression

Jun Yang, Joshua Bingham

Facebook AR/VR

05/08/2020

Outline

1) Background and Motivation
2) Products of Noise Suppression (NS) Technology
3) Traditional NS Algorithm
4) Our Proposed NS Solution
5) Evaluations
6) Features of Our Proposed NS Algorithm
7) Acknowledgments

1. Why Do We Need NS?

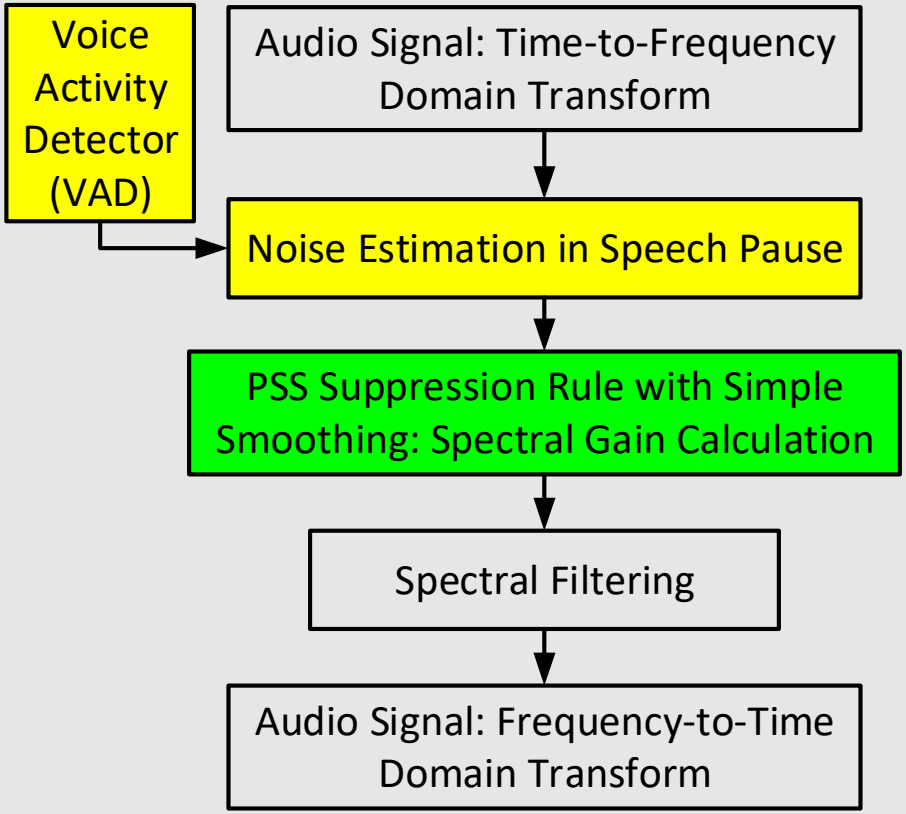
- Audio devices have been playing an important role in our daily life.
- Our environment is full of various types of noises.
- These noises can significantly degrade
  - voice quality
  - acoustic echo cancellation (AEC) performance
  - barge-in performance
  - automatic speech recognition (ASR) performance
  - voice over internet protocol (VoIP) performance

2. What Products Need NS?

Any audio device with microphone(s) and/or speaker(s)

- Home devices
- Mobile devices
- Wearable devices
- etc.

3. Block Diagram of Traditional NS Algorithm



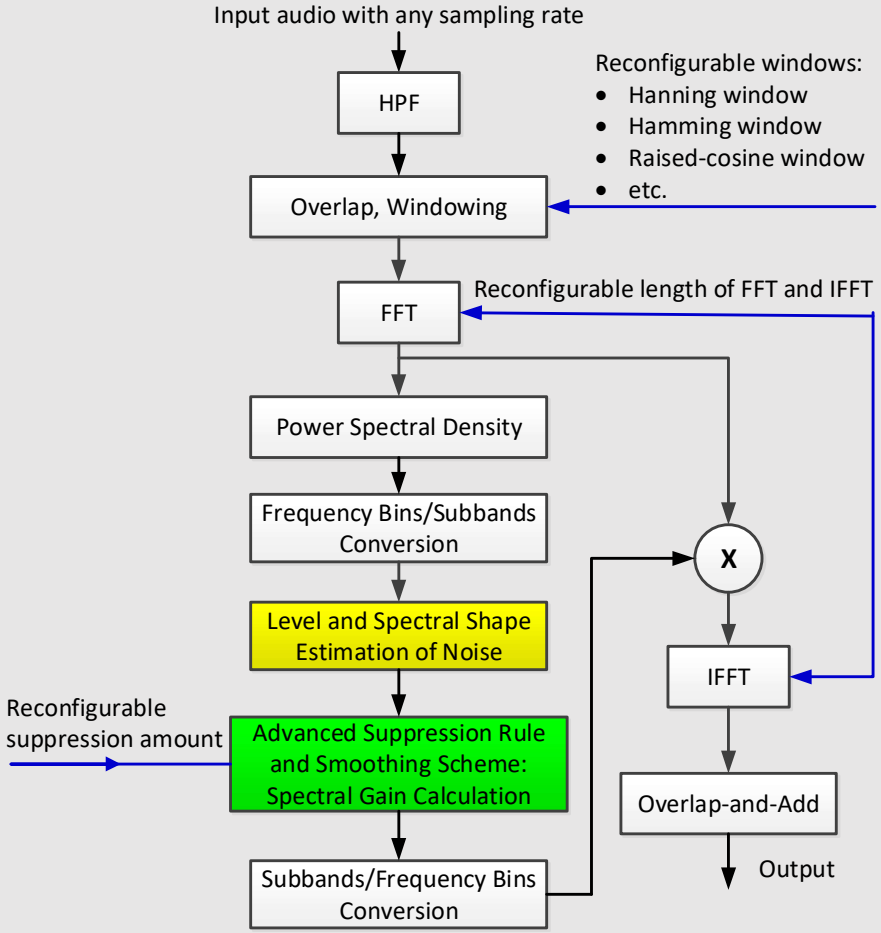
3.1 Traditional NS Algorithm: Suppression Rule

$$G(k,m) = [1 - (\frac{\lambda(k,m)}{|Y(k,m)|^2})^{\frac{\alpha}{2}}]^{\beta}$$ (or other similar formula)

where $\lambda(k, m)$ is an estimation of the noise power spectrum, α and β are two fixed parameters .

Method	(α, β)
Power Spectral Subtraction	(2, 1/2)
Magnitude Spectral Subtraction	(1, 1)
Short-time Wiener Filtering	(2, 1)

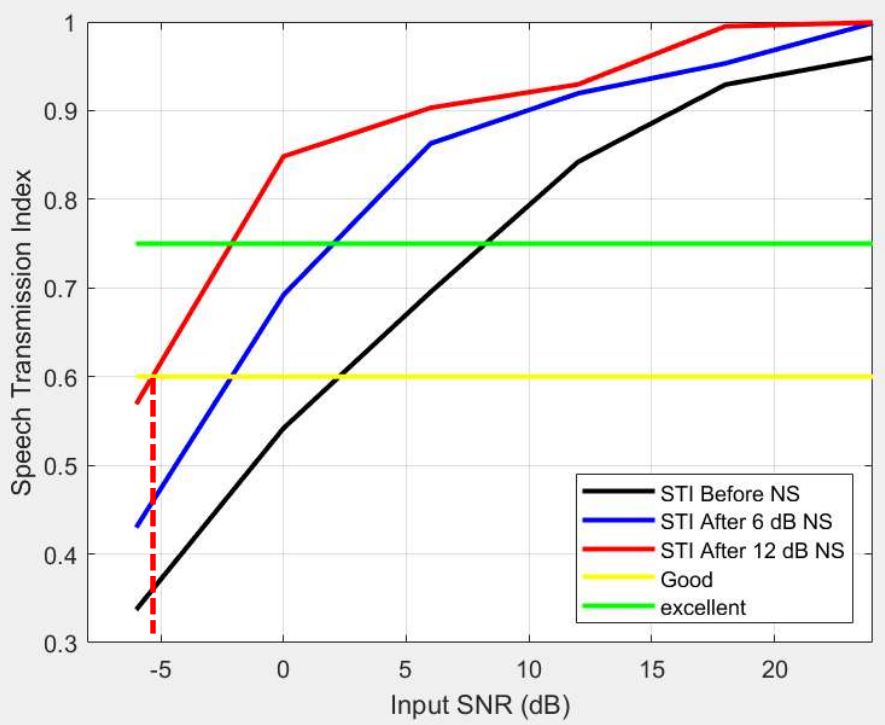
4. Our Proposed NS Solution

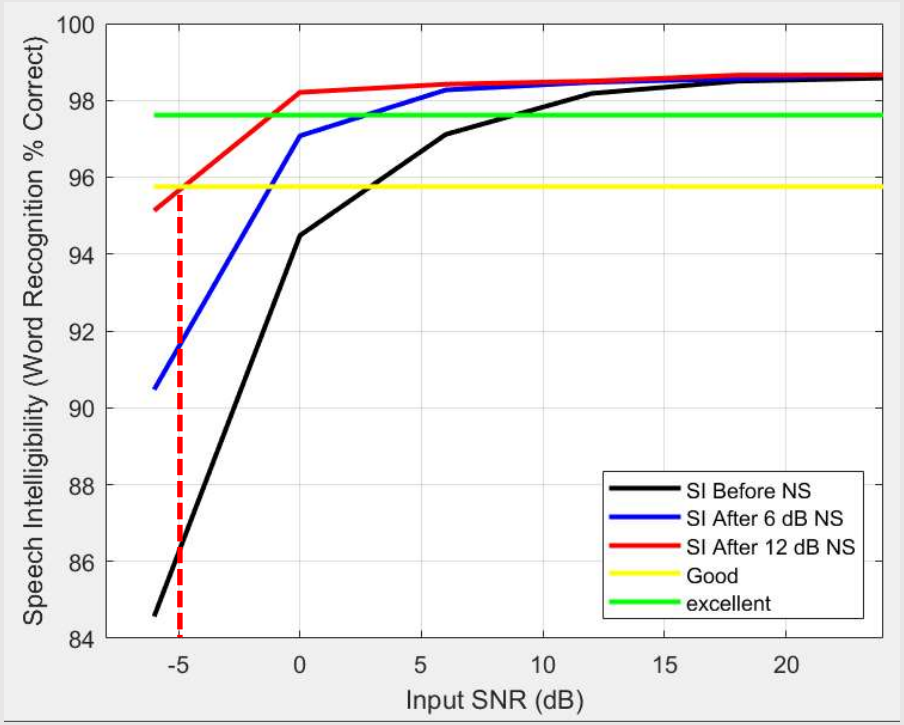


5. Evaluations

1) Speech Transmission Index (STI) and Speech Intelligibility (SI)
2) Signal-to-noise ratio (SNR) Improvement and Listening Tests
   - Test Conditions
     a) Two Input SNRs: 6 dB and 12 dB
     b) Two speech-levels at mouth reference point: 89 and 95 dBC
     c) Two distances between device-under-test and head-and-torso-simulator (HATS): 1 meter and 4 meters
     d) 9 types of noise: air condition noise, Café noise, fan noise, living-room noise, office noise, pink noise, Pub noise, rain noise, rock music noise
   - Total test cases: 2 * 2 * 2 * 9 = **72 cases**

5.1 STI and SI Before & After Our NS Processing

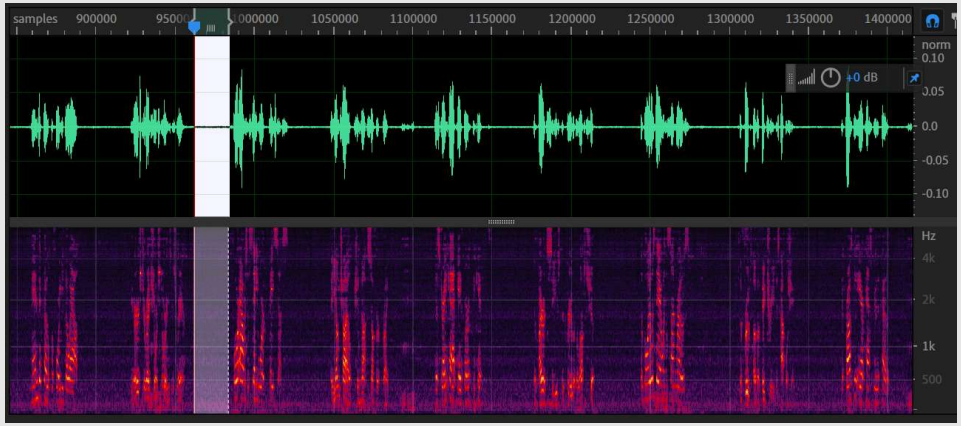




The Modified Rhyme Tests according to standard IEC 60268-16

5.2 Waveform & Spectrogram Before & After Our NS Processing

Before NS Processing



After Our NS Processing

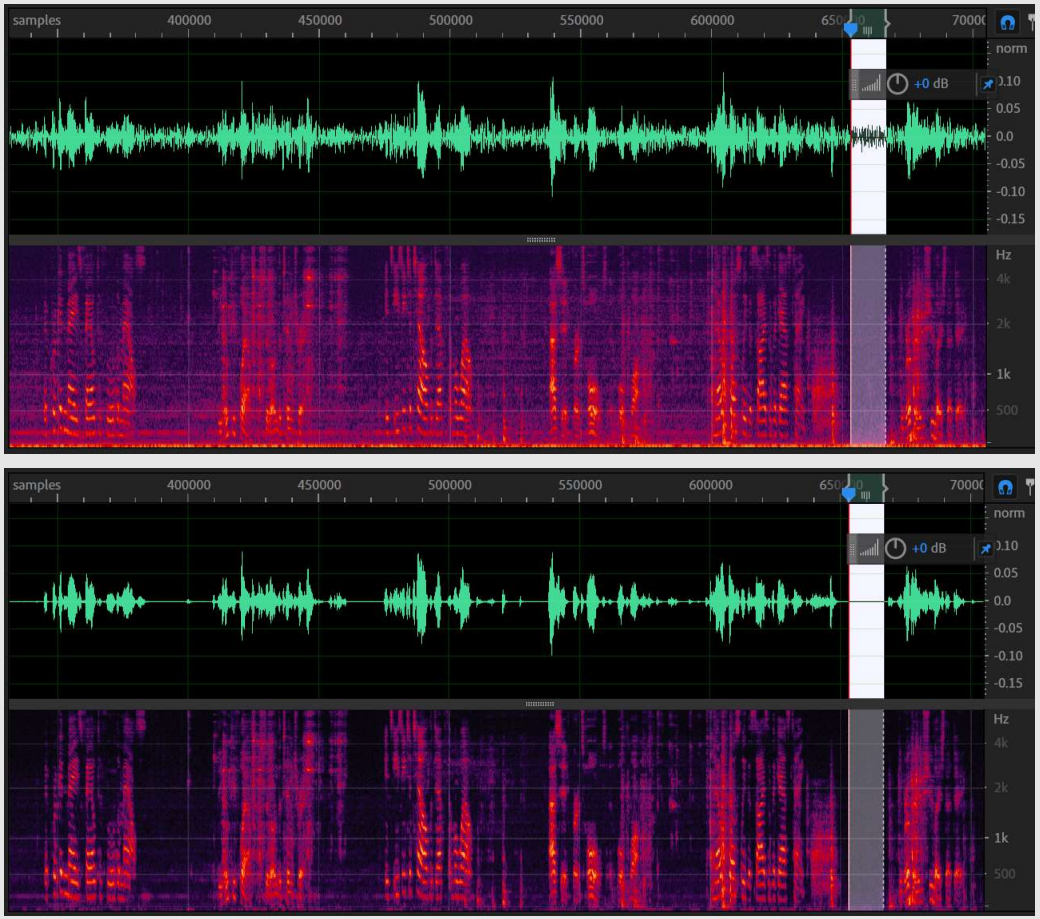
Test Condition:

6 dB input SNR, 89 dBC, 4 meters, AC noise

Result:

Noise is suppressed by ~24 dB after our NS.

5.3 Waveform & Spectrogram Before & After Our NS Processing



Before NS Processing

Speech in living-room noise:

(1). Input SNR = 6 dB

(2). 89 dBC at MRP

(3). 4 meters between DUT and head

(4). The noise is suppressed by ~37 dB by our NS processing.

After our NS Processing

6. Features of Our Proposed NS Algorithm

1) Natural sounding voice quality
2) Stable and comfortable residual noise, no “musical noise”
3) User adjustable noise suppression amount
4) Robust to input level
5) Support any sampling rate
6) Support multi-channel system with independent settings

7. Acknowledgments

Sincere thanks to:

- Jens Nilsson (Facebook Audio Tech @ Portal) for providing noisy speech database.
- Gongqiang Yu (Facebook Audio Tech @ AR) for providing Speech Transmission Index (STI) and Speech Intelligibility (SI) evaluation tool, the clean speech data and noise data.

Thank You!