RUHR-UNIVERSITÄT BOCHUM

McMaster University

RUB

Department of Electrical Engineering and Information Technology
Institute of Communication Acoustics

A Lightweight Music Processing Algorithm for Cochlear Implants Based on an Auditory Adaptation Mechanism

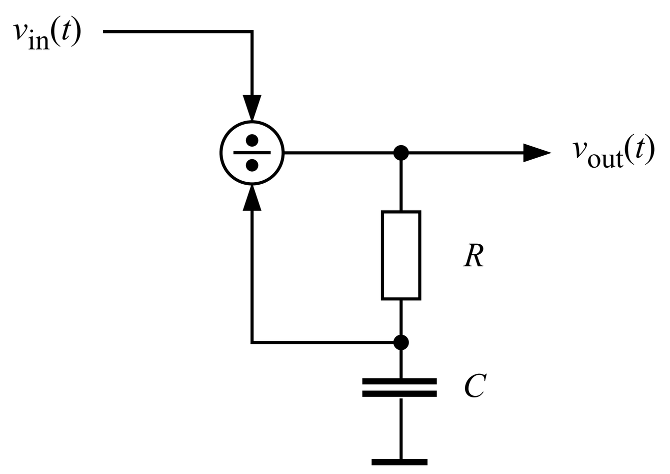
Anil Nagathil[1,2], Benjamin Lentz[1], Giselle Mojica[1], Ian C. Bruce[2], Christiane Völter[3], and Rainer Martin[1]

Introduction

- Listening to music is challenging for most cochlear implant (CI) listeners.
- Music processing can reduce polyphony (spectral complexity) and improve music appreciation.
- Most existing music processing algorithms cannot be easily implemented in current CIs (high computational complexity and latency).
- We propose a **computationally efficient music processing algorithm** that is based on a simple auditory adaptation model. The algorithm attenuates slowly varying signal components such as sustained accompaniments and retains important sounds like vocals and drums.

Auditory Adaptation Model (AAM)

Analog model [1, 2]

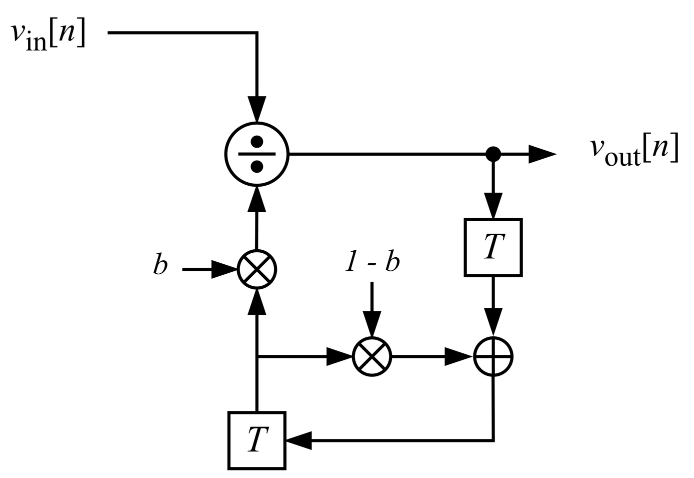


$$\frac{dv_C(t)}{dt} = \frac{1}{\tau}\left(v_{\text{out}}(t) - v_C(t)\right)$$

with $\tau = RC$.

$$v_{\text{out}}(t) = \frac{v_{\text{in}}(t)}{v_C(t)}$$

Equivalent discrete-time model

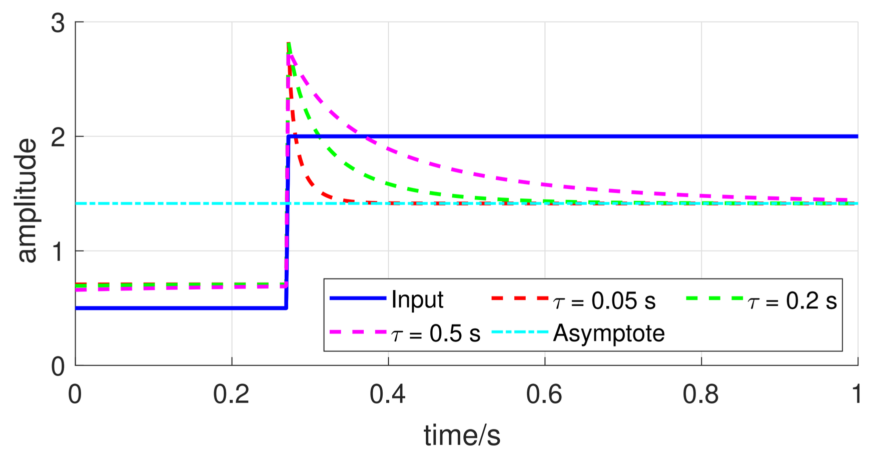


$$v_C[n] = v_{\text{out}}[n-1] + (1-b)\cdot v_C[n-1]$$

with $b = \Delta t/\tau$, Δt: sampling period

$$v_{\text{out}}[n] = \frac{1}{b}\cdot\frac{v_{\text{in}}[n]}{v_C[n-1]}$$

Effect of adaptation model

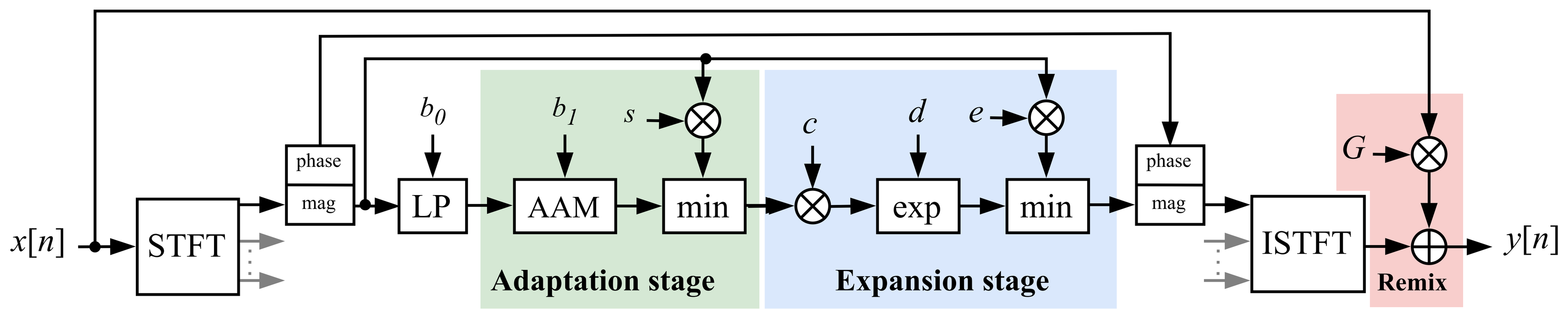


- Enhancement of onsets
- Square-root compression of steady-state components

References

[1] D. Püschel, *Prinzipien der zeitlichen Analyse beim Hören*, Ph.D. thesis, Georg-August-Universität zu Göttingen, 1988.
[2] T. Dau et al., "A quantitative model of the "effective" signal processing in the auditory system. I. Model structure," *J. Acoust. Soc. Am.*, vol. 99, no. 6, pp. 3615–3622, June 1996.
[3] F. Chen et al., "Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure," *Biomedical Signal Processing and Control*, vol. 8, no. 3, pp. 311–314, May 2013.
[4] N. Ono et al., "Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram," in *Proc. EUSIPCO*, Aug. 2008.
[5] W. Buyens et al., "A harmonic/percussive sound separation based music pre-processing scheme for cochlear implant users," in *Proc. EUSIPCO*, Sept. 2013.

Proposed Music Processing Algorithm



- LP: low-pass filter for temporal smoothing of magnitude spectrum.
- AAM enhances onsets, *min* operators + parameters s, e limit overshoots.
- Expansion stage controls output dynamics (d: exponent).
- Remix stage adds attenuated input signal to processed signal (G: gain).

Experimental Setup

Signal processing parameters:

- Sampling frequency: $f_s = \frac{1}{\Delta t} = 22.05$ kHz
- DFT length: $N = 512$, Hann window, 75% overlap

Parameter optimization:

- Parameter ranges: $\tau_i/\text{sec.} \in [0.005, 0.2]$, $s \in [0.9, 3]$, $d \in [1, 4]$, $e \in [0.5, 3]$
- Optimization criterion: ModA measure [3] (quantifies the modulation strength by calculating the area under the modulation spectrum)

Data set:

- DSD100 multi-track data set (vocals, bass, drums, accompaniment)

Baseline algorithm:

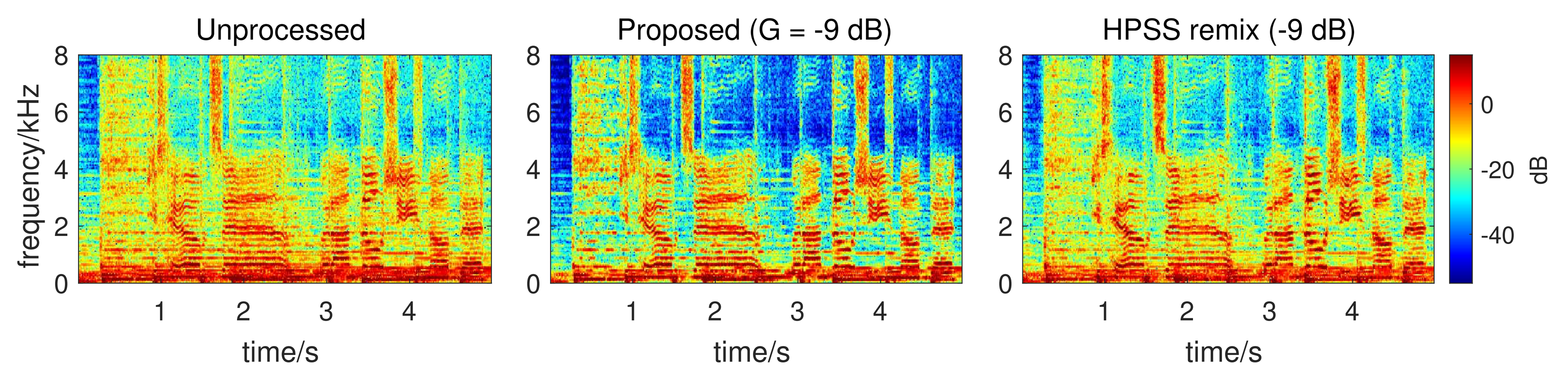
- Iterative Harmonic-Percussive Sound Separation (HPSS) [4] & remix [5]
- Estimates sustained ("harmonic") & transient ("percussive") sound components and attenuates sustained components.
- $f_s = 22.05$ kHz, DFT length: $N = 4096$, Hann window, 75% overlap

Listening experiment:

- 10 postlingually deafened CI listeners (5 female, 5 male)
- Mean age: 65.1 ± 14.53 years (range: 40 to 86 years)
- 16 music excerpts (not used during optimization)
- 2AFC tests: Proposed algorithm ($G = -9$ dB)
  - vs. unprocessed input signal
  - vs. HPSS remix (harmonic attenuated by -9 dB w.r.t. percussive)
- Signed-rank significance tests with Bonferroni-Holm correction.

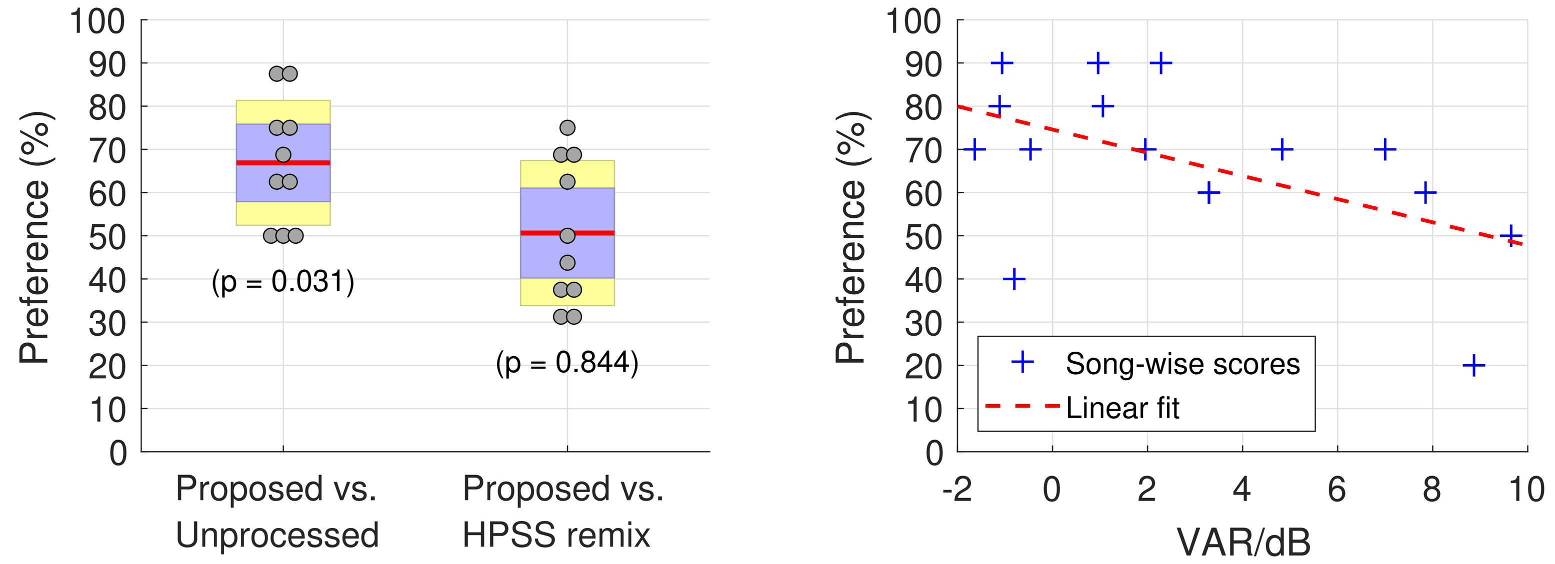
Results

Example spectrograms:



- Proposed algorithm increases spectro-temporal contrast and yields distinct vocals and drums.

Listening experiment:



- Significant preference over unprocessed music.
- Most effective at lower vocal-to-accompaniment ratios (VAR).
- No significant difference in preference compared to HPSS remix.

Computational complexity and latency:

(Signal length: 10 s; PC: Intel Core Ultra 7 165U, 64 GB RAM)

	Proposed	HPSS remix
Processing time	0.234 s ± 0.009 s	0.582 s ± 0.024 s
Latency (N/f_s)	23.2 ms	185.8 ms

Conclusions

Proposed music processing algorithm

- **improves music appreciation** and is on par with HPSS remix,
- has **lower computational complexity** and latency than HPSS,
- could be implemented as a **real-time application** in a CI.

Acknowledgments: This work has received funding from the European Union's Framework Programme for Research and Innovation Horizon Europe (2021-2027) under the Marie Sklodowska-Curie grant agreement no. 101151006 (AN), and the German Research Foundation (DFG), research grant 544277892 (BL, GM).

[1]Institute of Communication Acoustics, Ruhr-Universität Bochum, Bochum, Germany
[2]Department of Electrical & Computer Engineering, McMaster University, Hamilton, ON, Canada
[3]Cochlear Implant Center, Department of Otorhinolaryngology, Head and Neck Surgery, St. Elisabeth Hospital, Ruhr-Universität Bochum, Germany

21st Annual NeuroMusic Conference 2025, November 21st - 23rd, Hamilton, ON, Canada