RECONFIGURABLE MULTITASK AUDIO DYNAMICS PROCESSING SCHEME

Jun Yang, Amit S. Chhetri, Carlo Murgia, and Philip Hilmes

Amazon Lab126, Sunnyvale, CA 94089, USA

ICASSP 2019
May 12 - 17, 2019

Overview

- Voice over Internet Protocol (VoIP) and Automatic Speech Recognition (ASR) enabled smart speakers are changing the way we live.
- Audio dynamics processing is a crucial component of smart speakers' VoIP and ASR.
- Existing single-band and multiband dynamics processing (MBDP) schemes fail to maximize bass and loudness, can produce distortions and nonlinear echo, which result in poor ASR and full-duplex voice communication performance.
- A novel reconfigurable multitask MBDP scheme is proposed by using a perfect reconstruction filterbank, a flexible multiband compressor, and a scalable multiband limiter.
- The proposed MBDP is integrated with an acoustic echo cancellation (AEC) system in smart speakers.
- The proposed MBDP maximizes bass and loudness, enhances listening experience.
- The proposed MBDP minimizes speakers' total harmonic distortion (THD), prevents audio from clipping, drives speakers in linear range, reduces nonlinear echo.
- The proposed MBDP can achieve the optimal VoIP and ASR performance.

The Proposed MBDP Algorithm

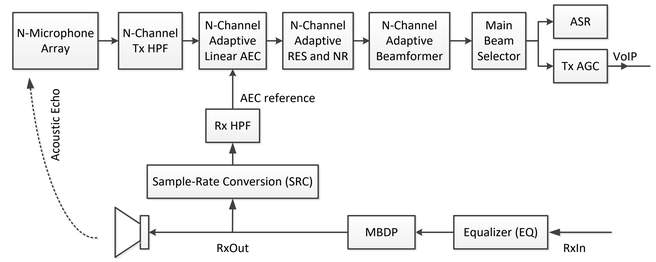

Fig. 1 The Proposed MBDP Scheme Integrated with AEC System

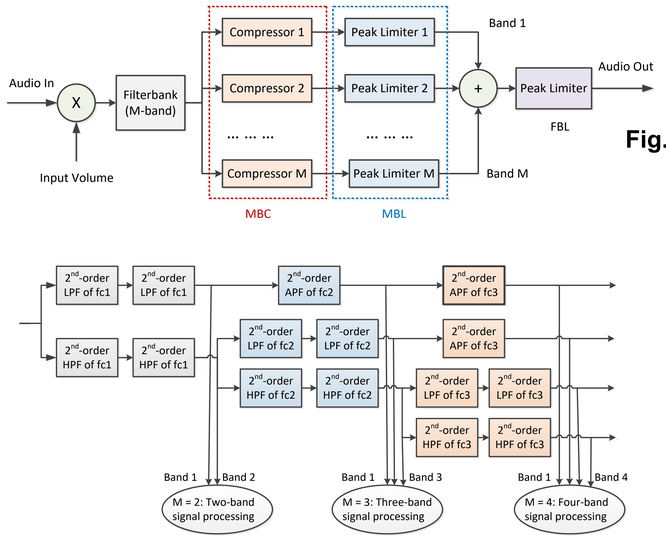

Fig. 2 The Proposed MBDP Scheme

Fig. 3 The Reconfigurable Filterbank Alg.

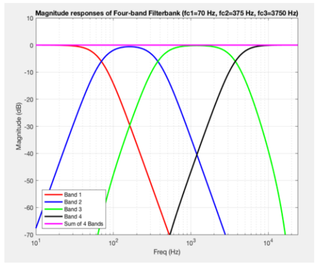

Fig. 4 Magnitude Responses of the Proposed 4-band Filterbank Alg.

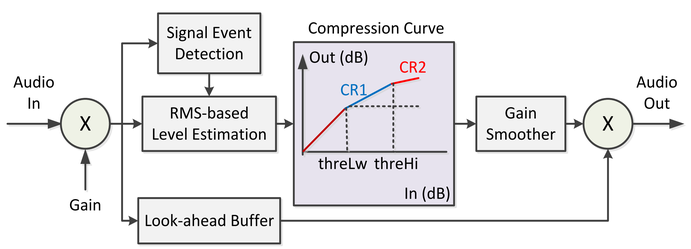

Fig. 5 The Proposed Flexible Compressor Alg.

Evaluation: Audio Bass and Loudness Enhancement

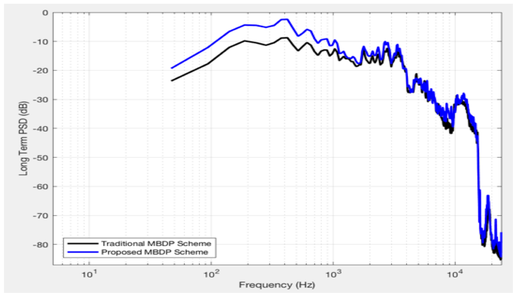

Fig. 9 Long-term Power Spectral Density Plots of Traditional (black) and the Proposed (blue) MBDP Schemes

Evaluation: Audio Distortion Reduction

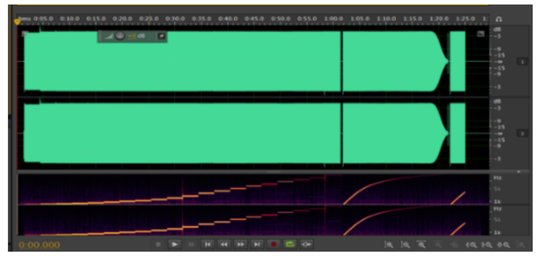

Fig. 6 Test Input Signals

Fig. 7 Output Signal of a Traditional MBDP Scheme

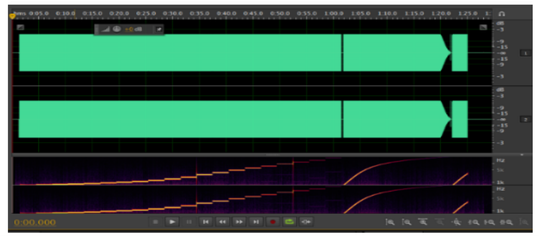

Fig. 8 Output Signal of the Proposed MBDP Scheme

Evaluation: Wake-up Word Recognition Improvement

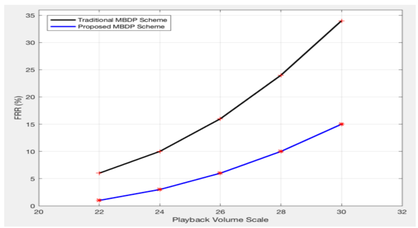

Fig. 10 False-Rejection-Rate (FRR) (%) versus Playback Volume.

Features of the Proposed MBDP Algorithm

- Maximize bass and loudness
- Prevent audio from clipping, drive speakers in linear range
- Minimize speakers' THD and audio playback distortion
- Provide good listening experience
- Reduce microphone signal saturation
- Reduce nonlinear echo, improve AEC barge-in performance
- Low latency, low computational complexity
- Easy and straightforward integration with any related systems and platforms