

Online-capable single-channel voice filter improves speech perception in speech-on-speech masking conditions

Jan Rennies¹, Ragini Sinha¹, Christian Rollwage¹, Ann-Christin Scherer¹, Simon Doclo^{1,2}

¹ Fraunhofer-Institut für Digitale Medientechnologie (IDMT) Oldenburg

² Carl von Ossietzky Universität Oldenburg, Signalverarbeitung, Cluster of Excellence „Hearing4all“

Abstract

Focusing on a specific talker in the presence of competing speech is a very challenging task. Previous studies have shown that both speech intelligibility and perceived listening effort can be considerably improved by algorithms that employ oracle knowledge to perform the segregation task for the listener (e.g., ideal binary masks). However, such algorithms are not suitable for hearing devices. This study therefore evaluated if an online-capable, single-channel voice filter algorithm could improve speech perception when a target talker was masked by two simultaneous competing talkers. The algorithm requires auxiliary information, i.e., “voice print” obtained from a couple of seconds of unmasked target speech using a speaker embedder, which is used as input to the speech extractor along with the mixture. Unprocessed and processed target speech of a male talker was evaluated in listening experiments with normal-hearing listeners for both male and female maskers. The results show that improvements in speech recognition thresholds are difficult to achieve, because listeners were able to perform well down to very low SNRs already in the unprocessed conditions. However, the effort related to understanding the target was considerably reduced as indicated by paired-comparisons and categorical listening effort scaling, indicating the potential of the algorithm for hearing device applications.