Comparing the influence of spectro-temporal integration in computational speech segregation

The goal of computational speech segregation systems is to automatically segregate a target speaker from interfering maskers. Typically, these systems include a feature extraction stage in the front-end and a classification stage in the back-end. A spectrotemporal integration strategy can be applied in either the frontend, using the so-called delta features, or in the back-end, using a second classifier that exploits the posterior probability of speech from the first classifier across a spectro-temporal window. This study systematically analyzes the influence of such stages on segregation performance, the error distributions and intelligibility predictions. Results indicated that it could be problematic to exploit context in the back-end, even though such a spectro-temporal integration stage improves the segregation performance. Also, the results emphasized the potential need of a single metric that comprehensively predicts computational segregation performance and correlates well with intelligibility. The outcome of this study could help to identify the most effective spectro-temporal integration strategy for computational segregation systems.

General information
Publication status: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Contributors: Bentsen, T., May, T., Kressner, A. A., Dau, T.
Number of pages: 5
Publication date: 2016

Host publication information
Title of host publication: Proceedings of Interspeech 2016
Publisher: International Speech Communication Association
Keywords: Computational speech segregation, Binary masks, Supervised learning, Spectro-temporal integration
DOIs: 10.21437/Interspeech.2016-1025
Source: PublicationPreSubmission
Source ID: 124288371
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2016 › Research › peer-review