

SpeechPy - Speech Processing and Recognition

Amirsina Torfi^{1,2}

¹ Virginia Tech ² Center for Identification Technology Research, West Virginia University

DOI: [10.21105/joss.00734](https://doi.org/10.21105/joss.00734)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 14 May 2018

Published: 17 May 2018

Licence

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

SpeechPy is an open source Python package that contains speech preprocessing techniques, speech features, and important post-processing operations. It provides most frequent used speech features including MFCCs and filterbank energies alongside with the log-energy of filter-banks. The aim of the package is to provide researchers with a simple tool for speech feature extraction and processing purposes in applications such as Automatic Speech Recognition and Speaker Verification.

Considering the recent advent of deep learning in ASR and SR and the importance of the accurate speech feature extraction, here are the motivations behind **SpeechPy** package: + Developing a free open source package which covers important preprocessing techniques, speech features, and post-processing operations required for ASR and SR applications. + A simple package with a minimum degree of complexity should be available for beginners. + A well-tested and continuously integrated package for future developments should be developed.

SpeechPy has been developed to satisfy the aforementioned needs. It contains the most important preprocessing and post-processing operations and a selection of frequently used speech features. The package is free and released as an open source software. Continuous integration using for instant error check and validity of changes has been deployed for **SpeechPy**. Moreover, prior to the latest official release of **SpeechPy**, the package has successfully been utilized for research purposes

Acknowledgements

This work is based upon a work supported by the Center for Identification Technology Research and the National Science Foundation under Grant #1650474.