
Unsupervised Speech Decomposition Via Triple Information Bottleneck

This repository provides a PyTorch implementation of SpeechSplit, which enables more detailed speaking style conversion by disentangling speech into content, timbre, rhythm and pitch.

This is a short video that explains the main concepts of our work. If you find this work useful and use it in your research, please consider citing our paper.



```
1 @article{qian2020unsupervised,  
2   title={Unsupervised speech decomposition via triple information  
3     bottleneck},  
4   author={Qian, Kaizhi and Zhang, Yang and Chang, Shiyu and Cox, David  
5     and Hasegawa-Johnson, Mark},  
6   journal={arXiv preprint arXiv:2004.11284},  
7   year={2020}  
8 }
```

Audio Demo

The audio demo for SpeechSplit can be found [here](#)

Dependencies

- Python 3.6
- Numpy
- Scipy

-
- PyTorch \geq v1.2.0
 - librosa
 - pysptk
 - soundfile
 - matplotlib
 - wavenet_vocoder `pip install wavenet_vocoder==0.1.1` for more information, please refer to https://github.com/r9y9/wavenet_vocoder

To Run Demo

Download pre-trained models to [assets](#)

Download the same WaveNet vocoder model as in AutoVC to [assets](#)

The fast and high-quality hifi-gan v1 (<https://github.com/jik876/hifi-gan>) pre-trained model is now available here.

Run `demo.ipynb`

Please refer to AutoVC if you have any problems with the vocoder part, because they share the same vocoder scripts.

To Train

Download training data to [assets](#). The provided training data is very small for code verification purpose only. Please use the scripts to prepare your own data for training.

1. Extract spectrogram and f0: `python make_spect_f0.py`
2. Generate training metadata: `python make_metadata.py`
3. Run the training scripts: `python main.py`

Please refer to Appendix B.4 for training guidance.

Final Words

This project is part of an ongoing research. We hope this repo is useful for your research. If you need any help or have any suggestions on improving the framework, please raise an issue and we will do our best to get back to you as soon as possible.