

Speech Enhancement in the Latent Space Based on Discrete Representations

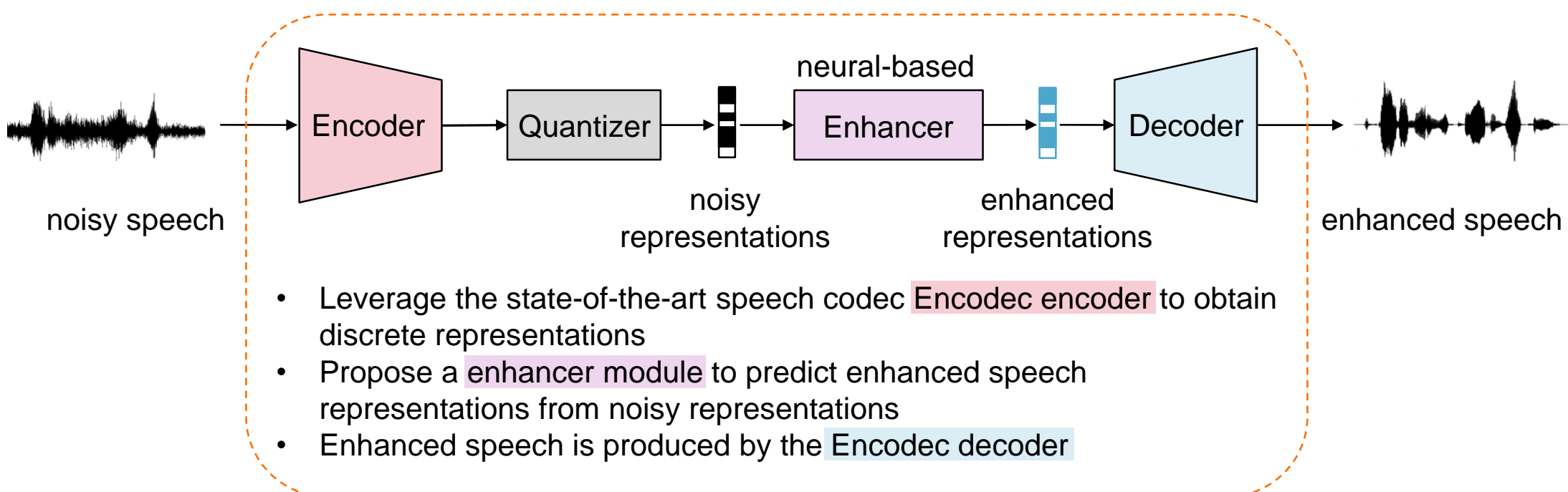
Master Thesis

Speech Enhancement

Latent Space Enhancement

Discrete representations

Speech enhancement aims to generate speech with high quality and intelligibility from distorted speech. Recent neural-based speech enhancement models have demonstrated superior ability in suppressing background noise, reverberation, and other acoustic interferences while preserving the clarity of the speech signal. However, the generated speech suffers from residual noise and artifacts for low SNR scenarios. Speech representations extracted by large self-supervised learning (SSL) models or speech codecs have facilitated many down-stream tasks. In this work, the state-of-the-art speech codec Encodec is to be investigated for improving the quality of enhanced speech.



What is the thesis about?

- Introduce hyperprior model into the existing speech coding pipeline
- Investigate the temporal dependency in the latent domain and improve coding efficiency
- Evaluate the performance and analyze the behavior

Your skills

- Good programming skills, ideally Python
- First knowledge or practical experience in the field of machine learning is valuable
- Enthusiasm to solve problems
- Read and understand scientific texts in English

How to get in touch?

Just send me an e-mail (re.shi@tu-Braunschweig.de) with your field of study, your grades, and why you are interested in this topic.

Our offer to you

- Insights into our current research
- Personal supervision and frequent discussion
- Invitations to the final presentation of other students to get an insight into the different research topics
- A workplace at our institute and the possibility to get in contact with other students
- Access to our own GPU cluster
- We aim to publish the results at a peer-reviewed conference