

Neural Speech Coding With Latent-Domain Predictive Coding

Master Thesis

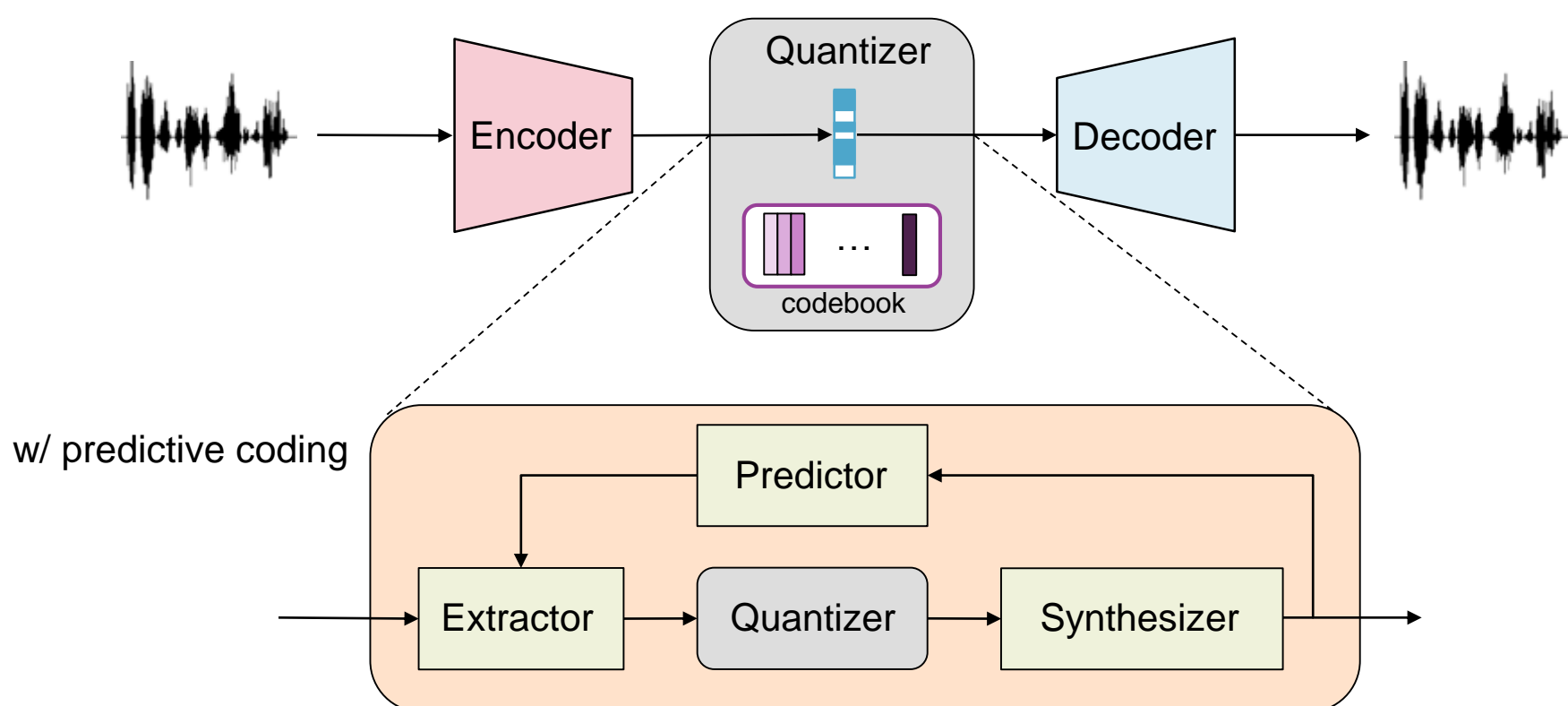
Speech Coding

Data Compression

Deep Learning

Audible Results

Speech coding aims to reconstruct the original speech with the least quality degradation. It is usually realized via an encoder to compress the given speech data, a quantizer to quantize the latent feature, and a decoder to reconstruct the speech, as shown in the figure (upper part). Due to the strong temporal correlation in the speech signal, the latent features obtained from the encoder often contain temporal redundancy, leading to a higher bitrate as the model needs to allocate more bits to represent the features. In video coding, redundancy removal techniques e.g., predictive coding, have been widely employed. Thus, we aim to investigate predictive coding in the latent domain to achieve more efficient speech coding.



What is the thesis about?

- Introduce predictive coding into the existing speech coding pipeline
- Evaluate the performance and analyze the behavior with predictive coding

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

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

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.