

Variational Speech Coding With Mean-Scale Hyperprior Model

Master Thesis

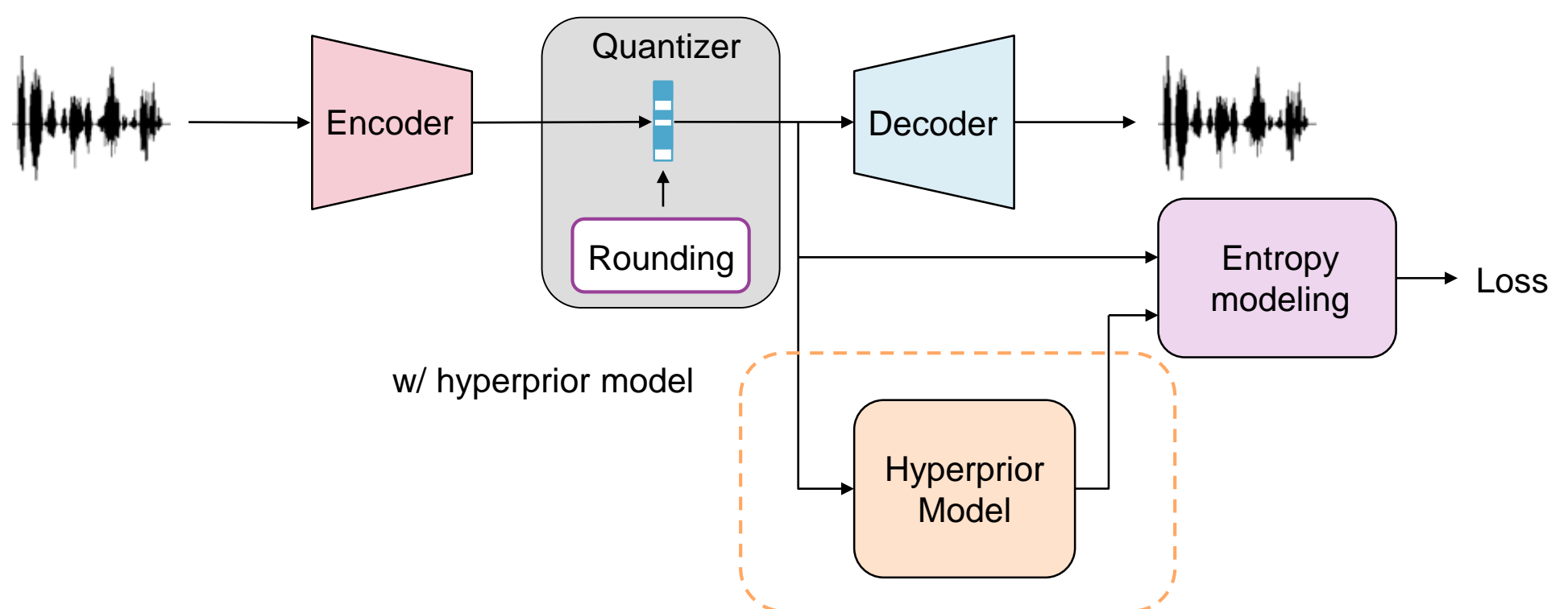
Speech Coding

Data Compression

Deep Learning

Entropy Modeling

Speech coding is essential for efficient speech transmission. The core is to transmit speech signals with fewer bits while preserving the intelligibility and quality as much as possible. Existing traditional and neural-based coding methods are characterized by an autoencoder embedded with a quantization module. Inspired by the work of neural image compression, where a hyperprior model is used to remove the marginal, e.g., structure, for images to achieve more efficient coding, the dependency within the latent features of speech obtained from the encoder has drawn more attention. In this work, we aim to investigate the impact of the hyperprior model in a speech coding application and optimize the whole pipeline.



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