

Master Thesis

voice conversion for speech synthesis and speech recognition applications

Motivation

Speech synthesis is the artificial production of human speech. With the state-of-the-art methods using neural networks, the new systems reach the level of the natural human voice. With voice conversion, using one-shot learning for example, the synthesis system could adapt to different human voices. Voice conversion has many applications, such as personalizing the synthesis, adding emotions or prosody to it, and generating synthetic data for speech recognition systems. A very promising technique of this field is the Generative Adversarial Networks (GANs) which revolutionized high-fidelity image generation.

Goal:

- Development of a voice conversion technique for speech synthesis system with few target voice data.

Task:

- Literature research on state of the art technology.
- Train and test the most promising architectures.

Qualifications:

- Good knowledge of machine learning and neural networks.
- Good grades in machine learning related subjects.

Interested?

Send your application to:

M.SC. Juan Hussain

ISL Room 239

+49 721 608 44733

juan.hussain@kit.edu