

Master Thesis title / MAL izenburua

Comparison of neural vocoders to be used with Tacotron2

Proposer(s) / Proposatzailea(k): names / izenak

Eva Navas

Contact / Kontaktua: email

eva.navas@ehu.eus

Description / Deskribapena

Tacotron 2 is one of the current state of the art text to spectrogram generation systems based on deep neural networks. It takes a text at its input and produces a spectrogram that can be converted into speech by applying a signal processing algorithm or by using a neural vocoder. Several neural vocoders have been proposed and most produce good results for the language they have been trained for (usually English). Some of these vocoders claim to be language independent, although there is usually a degradation in performance when used for languages other than English. The goal of this work is to test several of these vocoders for Spanish and compare their capability to generate high quality synthetic speech. Vocoders already implemented for English in Coqui-TTS (<https://github.com/coqui-ai/TTS/tree/main/TTS/vocoder>) will be the first to be tried, although more models can also be considered.

Goals / Helburuak

Adapting and training neural vocoders for Spanish

Integrating the vocoders with Tacotron 2

Evaluating the quality of the resulting synthetic speech

Requirements / Betebeharrak

Be able to read and understand Python programs

Framework / Esparrua

Framework here if needed

Tasks and plan / Atazak eta plana

Tasks and plan here if needed