Evolution of Audio and Speech Synthesis

Karan Dhingra
@kdhingra307

Support Staff

- Jupyter Notebook

- Tensorflow

- Librosa

- Matplotlib

BTS - Signal Synthesis

- Synthesis, what does it mean?

- Is synthesis of the sound wave is different than that of sign wave??

- What's the need for synthesizing??

- MOS, what is it??

HOW To - Signal Synthesis

* Input the signal and normalize it.

* Align it in a sequential manner if it is not.

* Clean the signal i.e. remove silence and extra noise.

* Feed the signal into a sequential neural learning architecture (RNNs).

* Raw Output from RNNs (As regression) vs Discretisation into levels(As Classification)

raw_signal.ipynb

BTS - Statistical Techniques for Speech

- Do statistical techniques uses a neural network to synthesize speech??

- Why can't we just synthesize audio just like we synthesized sine wave??

- MFCCs or Mel Spectrogram - What are these??

griff_lim.ipynb

BTS - Neural Techniques For Speech

- How are these different from statistical techniques??

- How much computation differences do both shows??

- How much quality gain is generally observed??

wavernn.ipynb

Current Work

- Wavenet [wavenet.ipynb]

- Tacotron 1 [tacotron.ipynb]

- Tacotron 2 [tacotron_wavenet.ipynb]

- WaveRNN [wavernn.ipynb]

My Work

* Fast Architecture, that synthesizes audio using the neural vocoder.

* Stream-Based Approach to synthesize speech.

* Sparsed Speech Generator.

Thank you